



**H2020-MSCA-ITN-2018-813545**

**HELICAL**

**Health Data Linkage for Clinical Benefit**

**Deliverable 4.3**

**Governance framework for research in rare diseases**

*This deliverable reflects only the authors' views, and the European Commission Research Executive Agency is not responsible for any use that may be made of the information it contains*



This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 813545



<b><i>Introduction</i></b> .....	<b>4</b>
<b><i>Framework Implementation Approach</i></b> .....	<b>4</b>
Regulatory Mechanisms – Policies and DPIA .....	5
Data Management Plan .....	5
Material Transfer Agreements .....	5
Ethical Approvals and Checks.....	5
Outreach and Engagement for Oversight – Patient Communities.....	6
Operational Management and Conclusions .....	6
<b><i>Annex A – Data Protection Impact Assessment Template</i></b> .....	<b>7</b>
IG Assessment Checklist – [Project Title] .....	7
Appendix B – Broad Privacy Risk Assessment: .....	21
<b><i>Annex B – Early-Stage Researcher Data Protection Impact Assessments Snapshots</i></b> .....	<b>23</b>
IG Assessment Checklist ESR1– Semantic Combining for Exploration of Environmental and Disease data: ANCA vasculitis in Ireland case study .....	24
IG Assessment Checklist ESR2 – T cell repertoires in giant cell arteritis.....	52
IG Assessment Checklist: ESR4, Harnessing the power of integrated data to investigate environmental exposures on ANCA vasculitis risk.....	68
IG Assessment Checklist ESR5 – ANCA-associated vasculitis & environmental risk factors: a case-control study.....	90
IG Assessment Checklist ESR6 – Atmospheric monitoring and time series analysis of climate and pollution impact on vasculitis onset.....	107
IG Assessment Checklist ESR7 – Identification of functionally relevant genetic variants associated with giant-cell arteritis (GCA).....	121
IG Assessment Checklist ESR8 – Linking public and GCA datasets to identify novel pathogenic pathways .....	140
IG Assessment Checklist ESR9 – Systems biology and bioinformatics approaches to provide a holistic understanding of GCA biology.....	159
IG Assessment Checklist ESR10 – [Functional characterisation of inflammation and vascular remodeling pathways in GCA, IDIBAPS, Barcelona].....	176



<b>IG Assessment Checklist ESR11 – Exosomes as biomarkers in ANCA-associated vasculitis .....</b>	<b>193</b>
<b>IG Assessment Checklist ESR12 – Computer assisted morphometry of pathological changes in renal biopsies from patients with AAV .....</b>	<b>211</b>
<b>IG Assessment Checklist ESR13 – Profiling the autoantibody repertoire in the context of systemic vasculitis flare .....</b>	<b>226</b>
<b><i>Annex C – Material Transfer Agreements .....</i></b>	<b><i>244</i></b>
<b><i>Annex D – Data Transfer Agreements.....</i></b>	<b><i>248</i></b>
<b><i>Annex E – List of Legal Bases and Special Category Personal Data Justifications.....</i></b>	<b><i>249</i></b>
<b><i>Annex F – Module 2 Curriculum and Transparency Workshop Materials .....</i></b>	<b><i>254</i></b>
<b>HELICAL Mid Term Module – Preparatory Work .....</b>	<b>254</b>



## Introduction

This deliverable describes the design, development, and deployment of the governance framework for HELICAL. The governance framework for rare disease research across the consortium partners is designed to facilitate the adoption of consistent as well as legally and regulatorily compliant practices for all of the research studies across all work packages. The existing instruments (policies, codes, rules, assessments and template agreements) that are being defined for governing the conduct of research across the consortium have been designed and are regularly updated to ensure that the autonomy and decision-making of each data/sample source is respected, including adherence to any local governance arrangements the source may be obliged to follow. Coupled with the Data Management Plan as provided in Deliverable 4.1 and the Information Governance Policies in Deliverable 4.2, the data governance framework has been informed by the activities described in these deliverables and represents their implementation.

The General Data Protection Regulation (GDPR) establishes the paradigm of “data protection by design and default” where protection of data needs to be considered and built in from the outset of developing any data intensive activity, and this was established as a first requirement for the development of the HELICAL. The second requirement was to run an impact assessment for any processing of data to assess whether there were particular risks to the rights and freedoms of individuals, and to the controllers and processors of data required to achieve a particular purpose. This requirement is embodied in the Data Protection Impact Assessment (DPIA) and has established itself as an essential tool for any individual or organisation to discharge their responsibility for protecting data and the people about whom it is being recorded.

This deliverable describes the approach taken to develop on the understanding of the research space at play across HELICAL and to describe the steps taken to ensure that the requirements for privacy are met, whilst achieving the educational goals of the ITN, preparing the Early Stage Researchers (ESRs), for a career steeped in robust and effective sensitive data handling. In this Deliverable we present the steps taken to tackle challenges that HELICAL would be addressing in the regulatory space, summarise the educational and engagement outreach and provide the policy items that have been developed as a result.

## Framework Implementation Approach

During the initial meetings which took place in December 2019 and the subsequent communication which followed with the individual sites, a series of information governance issues were identified which went on to lead to the creation of a programme aiming to give rise to policies tackling these. Building on these, the Information Governance Policies under Deliverable 4.2 were created, which encompassed the delivery of Module 2 to the ESRs the data flows of their project as well as the overall risk mitigation strategy. Please refer to Deliverable 8.5 for a full treatise on the approach and particulars taken for policy development.



## Regulatory Mechanisms – Policies and DPIA

As explained in further detail in Deliverable 4.2, the varied nature of the data, materials and regulatory requirements used in HELICAL mandated a trusted and authoritative approach for navigating these. Following the GDPR data protection by design and default approach, the European Institute for Innovation through Health Data (i~HD) adapted a DPIA Template that had been developed by its experts and used in the context of secondary use health data sets. This tailor made DPIA was deployed for the overall project (**Annex A**) as well as for each of the projects conducted by the ESRs (**Annex B**), thereby helping to comprehend and establish data flows, overarching legislation and compliance with the GDPR principles. Those individual DPIAs informed and guided how the information governance policies would then be implemented into practice. This has also been balanced by (i) the requirement that each ESR update the DPIA and the information governance team periodically, (ii) the existing ethics approvals obtained by the individual ESRs for their projects, and (iii) through overarching biobank registry ethics approvals and governance. The end result is ensuring that appropriate policies and agreements are set up based upon an understanding of the data flows and lawful purposes of the data processing.

## Data Management Plan

As part of the submitted HELICAL Periodic Technical Report Part B, the Data Management Plan (DMP) that was first prepared and submitted for HELICAL as part of as Deliverable 4.1 was reviewed with a view to be revised if necessary, in order to ensure that it provided an accurate outlook and solid framework for the work to continue. Following careful review and evaluation, it was decided that the DMP was rigorous, whilst the data flows have moved forward as expected, the key standards made in the DMP remain reasonable. A copy of the DMP is available in D4.1.

## Material Transfer Agreements

Similarly, to enable the sharing of data between the various sites and transfer of material across partners of the consortium, Material Transfer Agreements (MTAs) and Data Sharing Agreements (DSAs) were developed and / or reviewed by i-HD in collaboration with Trinity College, to ensure that the agreements enabled the parties to obtain the necessary data and material for the project in a legal and regulatory compliant manner. Copies of these (both draft and executed versions) are enclosed in **Annex C** and **Annex D** respectively (available upon request and with agreement of the relevant parties).

## Ethical Approvals and Checks

In July 2020, the Ethics Check Report was communicated to HELICAL requesting further information and documentation. As part of that work, HELICAL has, in relation to the aspect of the protection aspect for the clinical data and biosample list used in the project, been liaising with the ESRs to obtain the relevant information where this below is applicable to their projects:

1. Any information leaflets participants were provided and, where participant consent was sought, the blank consent form templates used.



2. Where geolocation data was used, a brief explanation which details and explains:
  - a. What data is used;
  - b. What this data is being used for;
  - c. The sources of the data; and
  - d. Confirmation that the data is used for research and not for stigmatization.

The responses obtained have been placed in a table which encompasses, in addition to the above, the legal basis under the GDPR through which ESR uses data in their project. The documentation obtained and table are enclosed in **Annex E**.

### Outreach and Engagement for Oversight – Patient Communities

In addition, a significant portion of the DPIA structure relates to GDPR's transparency requirements where ESRs would need to be able to meaningfully articulate their work to the patient community and wider public. We have been closely liaising with the patient advocate communities through Vasculitis Ireland Awareness (VIA), part of the RITA European Reference Network, who were able to join the sessions through Module 2 in 2020 (see D4.2) to provide training to the ESRs on effective public engagement and who have been a key part in developing Transparency Materials and a relevant workshop (**Annex F**), aimed at both to educate and ensure that the patient voice is represented in HELICAL in a foundational way.

### Operational Management and Conclusions

Through the establishment of the ESR DPIAs, overview of the data management plan, involvement of patient representative partners and ongoing ethical oversight, the governance framework proceeds on a basis whereby these particulars require continual update and evaluation.

WP4 and i~HD have, in close collaboration with TCD, overseen the ongoing reviews of the DPIAs and required that the ESRs periodically update the DPIAs and engage with i~HD for their review. The ethical oversight aspects have fed in to ensuring that data protection compliance requirements are met. These aspects are further balanced by strong communication with the Patient Association Organisations where the ESRs have the opportunities to address their membership and ensure that the nature of their research is clearly articulated and their privacy notices and efforts for transparency remain effective and well received.

The DPIAs, DMP, agreements, Ethics review responses, Transparency Materials and Information Governance policies resulting from this work are attached as appendices. A review by the Information Governance Board in March 2022 is proposed for review and assessment where appropriate.



## Annex A – Data Protection Impact Assessment Template

### IG Assessment Checklist – [Project Title]

#### Introduction to IG Assessment process

Under the General Data Protection Regulation (GDPR), a Data Protection Impact Assessment (DPIA) is only required where proposed data processing is “likely to result in a high risk to the rights and freedoms of natural persons” (Article 35(1)). However, Article 35(3) explicitly requires one where there is ‘large-scale’ processing of ‘special category’ (e.g. healthcare) data then a DPIA is required.

One other possibility is that the data being processed is already anonymised (see Recital 26) so falls outside GDPR altogether so that no DPIA is actually required.

However, good project management and information governance suggests that there should be a general approach to risk assessment for any project or business enterprise – if only to determine whether a DPIA might be required.

Ideally, one should work from a simple initial Checklist (this document) which identifies possible areas of information risk and compliance requirements to a ‘discussion note’ which explores any issues in more depth and may help identify the necessary mitigation methods and mechanisms to offset most if not all risks. Only if risks are unmitigated or remain ‘high’ would you move to a formal DPIA report.

#### The IG Assessment approach

There should be an overview of the proposed project or business change to explain what processing is envisaged as well as the purpose and intended outcome. The ‘purpose’ is important to establish the legal basis for the processing as well as ensuring that any possible mitigations or counter-measures do not undermine the main rationale for the processing.

The next step is to establish what compliance requirements may apply: GDPR, contractual or other regulatory restrictions, consent requirements, or obligations to preserve the data for legal or other reasons (including the benefit of posterity perhaps).

Once the precise range of obligations has been established, then appropriate checks can be made and recorded within the document.

The most obvious of these being GDPR compliance. There must be a ‘High Risk’ assessment (Appendix A) to determine whether the supervisory authority needs to be informed – generally, it is expected that it will not be necessary; if so, then a formal DPIA report will be needed.

Appendix B has a broader Privacy Impact Assessment that may throw up some broader issues.

Initial conclusions as to next steps or particular countermeasures to be considered should be detailed below.







Project Background/Overview

[Explain business background, including any existing processes and procedures; outline the project including stages, deliverables, and timelines]

Comparison of process steps (simplified): [optional]

This allows identification of what processing is new or changed through the project:

Step	Current	Proposed
Project initiation, including any ISAC approval, up to Task Order from client		No change

Initial Conclusions

concerning further counter-measures or business viability [possibly tentative]

1. ...
2. ...



Compliance Checks required:

Tick	Requirement	Notes [replace guide text with response]
<input type="checkbox"/>	Does the project involve processing 'personal data' of any sort?	Note: not just patient data; may need clear assessment of any anonymization to establish outside GDPR
<input type="checkbox"/>	Does the project involve processing 'confidential data' of any sort?	Note: may be 'commercial in confidence', medical confidentiality, or organisational confidentiality (internally sensitive); may need to check contractual limitations
<b>Data Availability requirements</b>		
<input type="checkbox"/>	Does data need to be held for GCP compliance?	
<input type="checkbox"/>	Does data need to be held to meet 'Open Data' requirements?	
<input type="checkbox"/>	Does data need to be held to meet ICMJE requirements or commitments?	



GDPR Compliance Checklist – where ‘personal data’ is processed:

Tick	Requirement	Notes [replace guide text with response]
<b>Article 5: Principles compliance checks</b>		
<input type="checkbox"/>	a) Is processing lawful, fair, and transparent?	
<input type="checkbox"/>	b) Is the purpose (or purposes) of the processing clearly defined	['purpose limitation' so should cover any subsequent or later processing]
<input type="checkbox"/>	c) adequate, relevant and limited to what is necessary	['data minimisation']
<input type="checkbox"/>	d) accurate and, where necessary, kept up to date	
<input type="checkbox"/>	e) kept and permits identification of data subjects for no longer than is necessary	['storage limitation']
<input type="checkbox"/>	f) processed securely	
<input type="checkbox"/>	2) can you demonstrate this compliance?	['accountability']
<b>Articles 13 &amp; 14 compliance</b>		[See detailed Transparency Checklist below]
<input type="checkbox"/>	Did the data come from publicly accessible sources?	[if so then transparency requirements may be reduced, but need to ensure data is accurate & up-to-date]
<input type="checkbox"/>	Are data subjects informed before processing starts for any new purpose if incompatible with original purpose where the controller wants to use data for a different purpose to the purpose for which they currently hold data	
<input type="checkbox"/>	Does the Privacy Notice and/or PIL cover this processing?	
<input type="checkbox"/>	What patient choices are available? Are these explained?	[see also Data Subject Rights below]
<b>Articles 6 and 9: legal bases</b>		
<input type="checkbox"/>	What are legal bases under Article 6	



Tick	Requirement	Notes [replace guide text with response]
<input type="checkbox"/>	What are legal bases under Article 9 (if 'special category' data)	
<input type="checkbox"/>	Are Article 6 legitimate interests explained where relevant?	[Complete an LIA form]
<input type="checkbox"/>	Are details of statutory obligations for Article 6 explained where relevant.	[Quote statutes or regulation]
<input type="checkbox"/>	Is this proposed processing compatible with the declared purposes?	[Check against any privacy notices and public information]
<b>Article 89(1) research exemption</b>		
<input type="checkbox"/>	If for research, do we meet Art 89(1) data minimisation	
<b>Articles 15-23: Data Subject Rights</b>		[See detailed table below]
<input type="checkbox"/>	Do we support data subject rights?	[If data is pseudo-/anonymised, then it would be difficult/impossible to do so]
<input type="checkbox"/>	There is no use of automated decision making (e.g. profiling)	[Otherwise need at least a 'discussion note']
<b>Articles 24-43: Controller-Processor</b>		
<input type="checkbox"/>	A28 & 29: What measures are there to ensure processors comply?	[Is there a formal Data Processing Agreement]
<input type="checkbox"/>	A30: Is there an entry for this processing/data held in the register?	
<input type="checkbox"/>	A32-34: Do we ensure appropriate security, including protection against unauthorised or unlawful processing and against accidental loss, destruction or damage, using appropriate technical or organisational measures?	[separate security checklist?]
<input type="checkbox"/>	A37-39: Is there a DPO and have they been or will they be consulted?	[part of sign-off of the DPIA]



Tick	Requirement	Notes [replace guide text with response]
<b>Articles 44-50: International transfers</b>		
	What form of data will be transferred to a third country or international organisation	[describe nature of data and whether identified, identifiable, de-identified or anonymous]
<input type="checkbox"/>	Are there safeguards for international transfers?	[e.g. US Privacy Shield, anonymisation, GDPR equivalence, approved contractual clauses, or BCR]
<b>Article 90: Obligations of secrecy</b>		
<input type="checkbox"/>	Do we meet medical confidentiality requirements?	[Note any national case law and statutory requirements that may affect this]

Data Subject Rights:

Note if supported and what process/procedure applies; if not, then describe the legal justification for not supporting this right.

<input type="checkbox"/>	To be informed: about processing, about choices, about rights, about controller	
<input type="checkbox"/>	the right of access to see or receive a printed copy	
<input type="checkbox"/>	the right to rectification – to correct any material errors in the personal data	
<input type="checkbox"/>	the right to erasure – where appropriate, to ask that all personal data is erased	
<input type="checkbox"/>	the right to restrict processing – to ask that some or all processing ceases [see opt-out]	
<input type="checkbox"/>	the right to data portability – this only applies to data provided directly by individual	
<input type="checkbox"/>	the right to object to and not to be subject to automated decision-making, including profiling	



<input type="checkbox"/>	Right to object to a Data Processing Authority (typically the relevant supervisory authority of each Member State)	
<input type="checkbox"/>	Where consent is the legal basis, the right to withdraw consent	

#### Detailed Transparency Checklist<sup>1</sup>

Does privacy information provided to data subjects include:

<input type="checkbox"/>	The name and contact details of our organisation	
<input type="checkbox"/>	The name and contact details of our representative (if applicable)	
<input type="checkbox"/>	The contact details of our data protection officer (if applicable)	
<input type="checkbox"/>	The purposes of the processing	
<input type="checkbox"/>	The lawful bases for the processing	[Art6 for 'personal data' & Art9 for 'special category']
<input type="checkbox"/>	The legitimate interests for the processing (if applicable)	
<input type="checkbox"/>	The categories of personal data obtained (if the personal data is not obtained from the individual it relates to)	[for Art14]
<input type="checkbox"/>	The recipients or categories of recipients of the personal data	
<input type="checkbox"/>	The details of transfers of the personal data to any third countries or international organisations (if applicable)	
<input type="checkbox"/>	The retention periods for the personal data.	

<sup>1</sup> Taken from UK Information Commissioner's Office template



<input type="checkbox"/>	The rights available to individuals in respect of the processing	
<input type="checkbox"/>	The right to withdraw consent (if applicable)	
<input type="checkbox"/>	The right to lodge a complaint with a supervisory authority	
<input type="checkbox"/>	The source of the personal data (if the personal data is not obtained from the individual it relates to)	[For Art14]
<input type="checkbox"/>	The details of whether individuals are under a statutory or contractual obligation to provide the personal data  (if applicable, and if the personal data is collected from the individual it relates to)	
<input type="checkbox"/>	The details of the existence of automated decision-making, including profiling (if applicable)	
<input type="checkbox"/>	We provide individuals with privacy information at the time we collect their personal data from them – or where we obtain personal data from a source other than the individual it relates to, we provide them with privacy information	
<input type="checkbox"/>	within a reasonable of period of obtaining the personal data and no later than one month	
<input type="checkbox"/>	if we plan to communicate with the individual, at the latest, when the first communication takes place	
<input type="checkbox"/>	if we plan to disclose the data to someone else, at the latest, when the data is disclosed	
<input type="checkbox"/>	We provide the information in a way that is:  <input type="checkbox"/> concise;	[Describe how we check is Plain English, etc.]



	<input type="checkbox"/> transparent; <input type="checkbox"/> intelligible; <input type="checkbox"/> easily accessible; and <input type="checkbox"/> uses clear and plain language.	
<input type="checkbox"/>	<p>When drafting the information, we:</p> <input type="checkbox"/> undertake an information audit to find out what personal data we hold and what we do with it. <input type="checkbox"/> put ourselves in the position of the people we're collecting information about. <input type="checkbox"/> carry out user testing to evaluate how effective our privacy information is	<p>[Note: best practice advice]</p>
<input type="checkbox"/>	<p>When providing our privacy information to individuals, we use a combination of appropriate techniques, such as:</p> <input type="checkbox"/> a layered approach; <input type="checkbox"/> dashboards; <input type="checkbox"/> just-in-time notices; <input type="checkbox"/> icons; and <input type="checkbox"/> mobile and smart device functionalities.	<p>[Note: best practice advice]</p>





### Security & Access Control Checklist

Controls need to be appropriate to level of risk: identified special category data needs more protection against potential misuse than non-personal data.

	Data Security classification (above Official)	<input type="checkbox"/> - Official-Sensitive <input type="checkbox"/> - Secret <input type="checkbox"/> - Top Secret <input type="checkbox"/> - Public Domain
<input type="checkbox"/>	Personal Data involved [GDPR]	
<input type="checkbox"/>	Special Category of personal data involved [GDPR]	
<input type="checkbox"/>	Electronic Communications (inc. cookies) [PECR]	
<input type="checkbox"/>	Credit Card data	
<input type="checkbox"/>	Legal enforcement [LED2018]	
<input type="checkbox"/>	Financial data	
<input type="checkbox"/>	Intellectual Property (detail owner)	
<input type="checkbox"/>	Commercial in confidence (detail owner)	
	Data Location (storage or processing)  (include any back-up site(s))	<input type="checkbox"/> - UK <input type="checkbox"/> - EU/EEA <input type="checkbox"/> - EU White-list <input type="checkbox"/> - USA <input type="checkbox"/> - Other:
<input type="checkbox"/>	Is data held in secure data centre?	[detail centre and what certification supports assertion]
<input type="checkbox"/>	Is this new supplier, location, or system?	[If so, need specific IS check; also need formal contract]
<input type="checkbox"/>	Is all user access subject to 2-factor authentication?	<input type="checkbox"/> - no control



		<input type="checkbox"/> - single factor (e.g. just password) <input type="checkbox"/> - 2-factor (e.g. password & fob) <input type="checkbox"/> - biometric [note: GDPR reqs] <input type="checkbox"/> - Other control:
<input type="checkbox"/>	Are there established JML procedures?	[Joiners, Movers, Leavers]
<input type="checkbox"/>	Are there checks that passwords are robust and secure enough?	[]
<input type="checkbox"/>	Are all administrator & user accounts routinely monitored?	[Particularly for redundant or little used accounts]
<input type="checkbox"/>	Are systems protected against malware and other attacks?	[provide details of protection software and procedures]

[Need some aspect of CIA/impact-likelihood assessment]

#### Information Asset Register Checklist

<input type="checkbox"/>	Are there new IAs being created?	[provide details]
<input type="checkbox"/>	Are old IAs being retired?	[provide details]
<input type="checkbox"/>	Have IAOs & IACs been consulted?	
<input type="checkbox"/>	Has IAR been updated/amended?	[at least create project task to do so]
<input type="checkbox"/>	Data Retention classification & period	
<input type="checkbox"/>	Data retention procedure/functionality in place	



## Appendix A – Supervisory Authority ‘High Risk’ Check

If the DPIA shows ‘high risk’ processing which cannot be mitigated, then the DPIA should be sent to the relevant authority for review before any processing starts. Note that their review may take several weeks to process. A ‘High Risk’ assessment represents a ‘risk to the rights and freedoms of individuals’ – so may extend beyond GDPR consideration, including Human Rights.

GDPR Article 35(3) provides three examples:

- a) a systematic and extensive evaluation of personal aspects relating to natural persons which is based on automated processing, including profiling, and on which decisions are based that produce legal effects concerning the natural person or similarly significantly affect the natural person;
- b) processing on a large scale of special categories of data referred to in Article 9(1), or of personal data relating to criminal convictions and offences referred to in Article 1013; or
- c) a systematic monitoring of a publicly accessible area on a large scale

ICO cites:

1. Systematic and extensive profiling with significant effects
2. Large scale use of sensitive data [viz. ‘special category’ in GDPR terms]
3. Public monitoring

These being the same as (a)-(c) above. They further identify:

1. **New technologies:** processing involving the use of new technologies, or the novel application of existing technologies (including AI).
2. **Denial of service:** Decisions about an individual’s access to a product, service, opportunity or benefit which is based to any extent on automated decision-making (including profiling) or involves the processing of special category data.
3. **Large-scale profiling:** any profiling of individuals on a large scale.
4. **Biometrics:** any processing of biometric data.
5. **Genetic data:** any processing of genetic data, other than that processed by an individual GP or health professional for the provision of health care direct to the data subject.
6. **Data matching:** combining, comparing or matching personal data obtained from multiple sources.
7. **Invisible processing:** processing of personal data that has not been obtained direct from the data subject in circumstances where the controller considers that compliance with Article 14 would prove impossible or involve disproportionate effort.
8. **Tracking:** processing which involves tracking an individual’s geolocation or behaviour, including but not limited to the online environment.
9. **Targeting of children or other vulnerable individuals:** The use of the personal data of children or other vulnerable individuals for marketing purposes, profiling or other automated decision-making, or if you intend to offer online services directly to children.



10. **Risk of physical harm:** Where the processing is of such a nature that a personal data breach could jeopardise the [physical] health or safety of individuals.

**'High Risk' assessment using ICO criteria:**

Criterion:	Assessment	Comments
New technologies		
Denial of service		
Large-scale profiling		
Biometrics		
Genetic data		
Data matching		
Invisible processing		
Tracking		
Targeting of children or other vulnerable individuals		
Risk of physical harm		

[The assessment can be one of N/A (not applicable), Low, Medium, or High. The comments should explain how the assessment is justified.]



Appendix B – Broad Privacy Risk Assessment:

#	Risk Description/detail	Discussion
1.	Data accuracy and timeliness	[Is data accurately recorded & kept up-to-date?]
2.	Differential treatment of patients/data subjects	[Might certain categories of people be adversely affected, e.g. children, vulnerable adults]
3.	Data Accuracy and identification	[Is the identification of individual reliable? Is there a danger of mis-attribution or incorrect linkage of data?]
4.	Holding / sharing / use of excessive data within [Company] systems	[Might too much data be held or for long? Is there a clear justification for data retention? Not 'just in case']
5.	Data held too long within [Company] systems	[Is there a clear data retention period specified and are there processes to ensure its deletion when no longer needed? Are copies tracked and deleted as well?]
6.	Excessive range of access in terms of users to personal data (consider new users/change of access privileges)	[Do more users have access than strictly necessary? Are user roles clear distinguished and reflected in the access privileges? Is there a clear process for granting and revoking access privileges?]
7.	Potential for misuse of data, unauthorised access to systems	[What are the likely threats to the data? What countermeasures are or might be applied? Is it possible for access to be granted inappropriately?]
8.	New sharing of data with other organisations, including new or change of suppliers	[Is data being shared from new data providers or with new data users? Are there new suppliers or data processors? What controls will apply?]
9.	Variable and inconsistent adoption / implementation	[How well will this system work end-to-end? How robust is it against partial adoption or system failure?]
10.	Legal compliance, particularly DP transparency requirements and support for data subject rights	[How well does this system meet legal requirements – or appear to meet legal requirements? Does it meet the 'No surprises' rule? What would happen if an individual requests data erasure or ceasing processing, etc.]



#	Risk Description/detail	Discussion
11.	Medical confidentiality	[Are there any addition sensitivities over confidentiality? Might specific approval (e.g. REC) be required to support this processing?]



## Annex B – Early-Stage Researcher Data Protection Impact Assessments Snapshots

Note that the DPIAs are living documents and subject to periodic review and update.



## IG Assessment Checklist ESR1– Semantic Combining for Exploration of Environmental and Disease data: ANCA vasculitis in Ireland case study

### Introduction to IG Assessment process

Under the General Data Protection Regulation (GDPR), a Data Protection Impact Assessment (DPIA) is only required where proposed data processing is “likely to result in a high risk to the rights and freedoms of natural persons” (Article 35(1)). However, Article 35(3) explicitly requires one where there is ‘large-scale’ processing of ‘special category’ (e.g. healthcare) data then a DPIA is required.

One other possibility is that the data being processed is already anonymised (see Recital 26) so falls outside GDPR altogether so that no DPIA is actually required.

However, good project management and information governance suggests that there should be a general approach to risk assessment for any project or business enterprise – if only to determine whether a DPIA might be required.

Ideally, one should work from a simple initial Checklist (this document) which identifies possible areas of information risk and compliance requirements to a ‘discussion note’ which explores any issues in more depth and may help identify the necessary mitigation methods and mechanisms to offset most if not all risks. Only if risks are unmitigated or remain ‘high’ would you move to a formal DPIA report.

### The IG Assessment approach

There should be an overview of the proposed project or business change to explain what processing is envisaged as well as the purpose and intended outcome. The ‘purpose’ is important to establish the legal basis for the processing as well as ensuring that any possible mitigations or counter-measures do not undermine the main rationale for the processing.

The next step is to establish what compliance requirements may apply: GDPR, contractual or other regulatory restrictions, consent requirements, or obligations to preserve the data for legal or other reasons (including the benefit of posterity perhaps).

Once the precise range of obligations has been established, then appropriate checks can be made and recorded within the document.

The most obvious of these being GDPR compliance. There must be a ‘High Risk’ assessment (Appendix A) to determine whether the supervisory authority needs to be informed – generally, it is expected that it will not be necessary; if so, then a formal DPIA report will be needed.

Appendix B has a broader Privacy Impact Assessment that may throw up some broader issues.

Initial conclusions as to next steps or particular countermeasures to be considered should be detailed below.





## Project Background/Overview

[Explain business background, including any existing processes and procedures; outline the project including stages, deliverables, and timelines]

The aetiology and treatment for the vast majority of rare diseases remains unknown. Researchers studying rare diseases have several challenges to better understand health outcomes. One of these challenges is the requirement of additional data types outside the health sector, like environmental data. KG approaches are being used in this domain since they facilitate the data integration process for diverse data. However, Health Data Researchers would require expertise in Computer Science to integrate, access, navigate and export these linked data to be used in environmental research.

Therefore, we formulated the following research question to be addressed during my PhD:

- 'To what extent can a graph-based methodology that integrates environmental data with longitudinal and geospatial diverse clinical data, support Health Data Researchers (HDR) to identify appropriate environmental variables to validate their hypothesis validation for rare disease research?'

The solution proposed is to develop a framework to support researchers that require a flexible methodology to integrate environmental with longitudinal and geospatial diverse clinical data. An initial framework has been developed, called SERDIF (Semantic Environmental and Rare Disease data Integration Framework), which is illustrated in Figure 1.

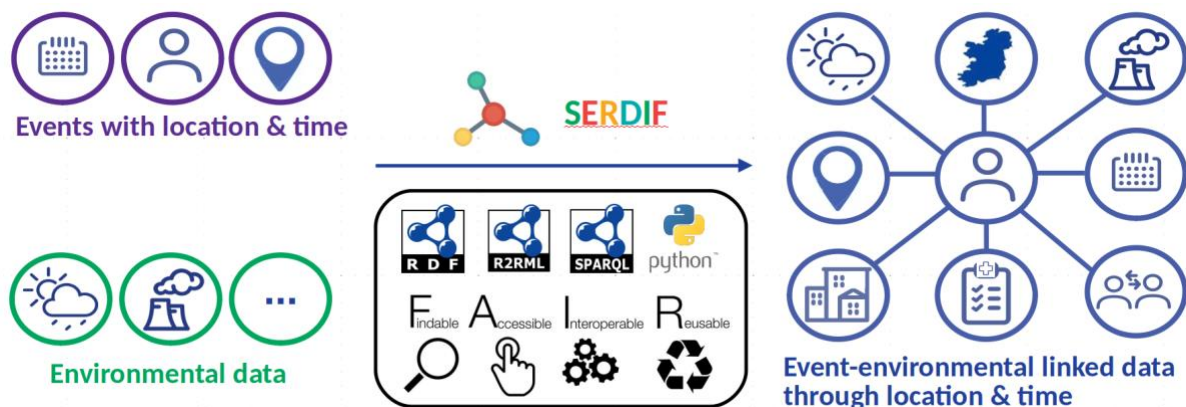


Figure 1: SERDIF graphical abstract.

The foreseen contributions of my PhD at the current state are in the (i) SERDIF framework: a methodology<sup>2</sup>, the associated knowledge graph<sup>3</sup> structure and a dashboard to provide meaningful access to this linked data<sup>4</sup>. This novel contribution (ii) uses Semantic Web technologies to bridge rare

<sup>2</sup> <https://github.com/navarral/ijckg2021-serdif-paper>

<sup>3</sup> <https://serdif-example.adaptcentre.ie>

<sup>4</sup> <https://serdif-example-dash.herokuapp.com/>



disease research, environmental science and data protection disciplines addressing a gap in the state-of-the-art. Furthermore, (iii) the description of the framework in-use application for the three projects mentioned above, which use Semantic Web technologies to link patient and scientific data. (iv) SERDIF is envisaged to be used in a series of health data linkage projects at a European and International level.

Three rare diseases will be used as case studies for this research: Anti-neutrophil cytoplasm antibody (ANCA)-associated vasculitis (AAV) in Ireland, Kawasaki disease in Japan and AAV vasculitis in Europe. These particular case studies were chosen due to the opportunity to interact with the HDRs afforded through the AVERT<sup>5</sup>, Marie Curie ITN Helical<sup>6</sup>, and EJD FAIRVASC<sup>7</sup> projects. The data used in each case study is further described in the following section.

#### AAV in Ireland Data

##### **Clinical:**

- the Rare Kidney Disease Registry and Biobank<sup>8</sup> from AVERT;
- patient's medical records Ethics and Data Management documents references: AVERT's Data Management plan<sup>9</sup>, AVERT's participant consent form<sup>10</sup> and information sheet<sup>11</sup>.

##### **Environmental:**

- Met Éireann historical weather data<sup>12</sup> and EPA Ireland air quality data<sup>13</sup> (land-based stations);
- ERA5 climate data<sup>14</sup> and EAC4 pollution data<sup>15</sup> (reanalysis data).

##### **Geometry:**

- Republic of Ireland electoral division boundaries<sup>16</sup>

---

<sup>5</sup> <https://www.tcd.ie/medicine/thkc/avert/>

<sup>6</sup> <http://helical-itn.eu/>

<sup>7</sup> <https://fairvasc.eu/>

<sup>8</sup> <https://www.tcd.ie/medicine/thkc/research/rare.php>

<sup>9</sup> [https://www.tcd.ie/medicine/thkc/assets/pdf/Data%20Management%20Plan\\_v1.1.pdf](https://www.tcd.ie/medicine/thkc/assets/pdf/Data%20Management%20Plan_v1.1.pdf)

<sup>10</sup>

[https://www.tcd.ie/medicine/thkc/assets/pdf/AVERT\\_Participant%20Consent%20Form%20RKD\\_v2\\_May%202018.pdf](https://www.tcd.ie/medicine/thkc/assets/pdf/AVERT_Participant%20Consent%20Form%20RKD_v2_May%202018.pdf)

<sup>11</sup> [https://www.tcd.ie/medicine/thkc/assets/pdf/AVERT-patient-info-sheet\\_RKD\\_v2.1\\_May-2018.pdf](https://www.tcd.ie/medicine/thkc/assets/pdf/AVERT-patient-info-sheet_RKD_v2.1_May-2018.pdf)

<sup>12</sup> <https://www.met.ie//climate/available-data/historical-data>

<sup>13</sup> <https://www.epa.ie/air/quality/data/>

<sup>14</sup> <https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels>

<sup>15</sup> <https://ads.atmosphere.copernicus.eu/cdsapp#!/dataset/cams-global-reanalysis-eac4>

<sup>16</sup> <http://census.cso.ie/censusasp/saps/boundaries/ED%20Disclaimer.htm>



## Kawasaki Disease Data

### Clinical

- Kawasaki epidemiological national survey in Japan

### Environmental

- JMA weather data<sup>17</sup>, NIES air quality data<sup>18</sup> (pollution) and AD-Net lidar data (aerosol).
- ERA5 climate data and EAC4 pollution data (reanalysis data).

### Geometry

- Japan prefecture boundaries

## AAV in Europe Data

### Clinical

- FAIRVASC AAV disease registry

### Environmental

- European Air Quality data from EEA<sup>19</sup>
- ERA5 climate data and EAC4 pollution data (reanalysis data).

### Geometry

- Europe countries, regions and counties boundaries

## The SERDIF metadata

In this project, dataset descriptor, provenance, lineage and data protection metadata is generated when querying event-environmental linked data. The metadata is described using the Resource Description Framework (RDF) following W3C standards and recommendations:

1. dataset descriptions (DCAT, <https://www.w3.org/TR/vocab-dcat-2/>),
2. statistical data (RDF Data Cube, <https://www.w3.org/TR/vocab-data-cube/>),

---

<sup>17</sup> <https://www.jma-net.go.jp/kousou/information/data/index-e.html>

<sup>18</sup> <http://www.nies.go.jp/igreen/index.html>

<sup>19</sup> <https://discomap.eea.europa.eu/map/fme/AirQualityExport.htm>  
<https://discomap.eea.europa.eu/map/fme/AirQualityExportAirbase.htm>



3. provenance data (PROV-O, <https://www.w3.org/TR/prov-o/>),
4. data protection (DPV, <https://w3c.github.io/dpv/dpv/>).

Metadata is key in this research project since environmental data associated with individual health events is considered pseudonymised data. Identification risks exist for the patients in terms of singling out an individual, data linking with other sources or inferencing certain data from the former.

Environmental data cannot be generalized when using regional or small area approaches. In addition, anonymization methods cannot be applied effectively without losing the value of the data for rare disease research. For example, permuting the environmental observations would affect the temporality of the data or introducing noise would affect the magnitude of the values hiding the signal researchers are looking for. However, example data and real metadata could be shared as Open Data following the FAIR guiding principles as in the DOI below:

<https://doi.org/10.5281/zenodo.5544257>

#### Usability metrics

In this project, we aim to address the challenge of integrating multiple heterogeneous data sources using Knowledge Graphs (KG). KGs have a steep learning curve which can present an obstacle for non-technical researchers who want to access and explore the data to meet their needs. That is why we designed the SERDIF dashboard as an artefact of the framework, a visual tool designed for use by Health Data researchers. The SERDIF dashboard allows to safely combine, access and export environmental data associated with clinical rare disease data; whilst hiding the complexities.

The SERDIF dashboard is evaluated and refined in a usability evaluation together with the framework requirements. The usability evaluation starts with an experiment that involves questioning the usability of this dashboard. The participants will be asked to complete a series of tasks starting from sending a query to downloading the data of interest. Visualizing the query results from a table and a plot is also assessed since there is an interest in testing basic comprehension of the results prior to downloading the data. The time spent per task will be recorded during the completion of each task with a stopwatch. While performing the experiment the participants are asked to think aloud, the statements and feedback are recorded with an automatic transcription feature of the video conferencing platform. This recording will be used to correct the statements that the note-taker will write down during the experiment.

The experiment will contribute to evaluating through different case studies my main target research contribution of my PhD. This contribution is to design a framework to support Health Data researchers identify the appropriate environmental variables to enable their hypothesis exploration.



Comparison of process steps (simplified): [optional]

This allows identification of what processing is new or changed through the project:

Case study	Step	Current	Proposed
AAV in Ireland	Data collection	Signed a data sharing agreement to use RKD registry data -- 20/02/2020	
		Agreed to terms and conditions, acknowledgment of the source and data download disclaimer per environmental and geometry dataset -- 03/10/2019	
	Semantic Uplift	Clinical and Geometry data already available as Resource Description Framework (RDF) files within a triple store in the School of Computer Science and Statistics (SCSS) network within the Trinity College of Dublin (TCD).	Design the mapping to convert tabular environmental data files to RDF.  Imported the RDF generated files in a triplestore separated from the clinical data hosted in the SCSS network within TCD.
	Data Querying and Filtering		Defined a spatio-temporal query as a SPARQL template, which allows the user to input their parameters of interest.
	Data Visualization		SERDIF dashboard
	Data export/downlift		SERDIF dashboard



		User experiment (see Background section) got Submitted and accepted proposal to the Research Ethics Committee of the School of Computer Science and Statistics in TCD -- 16/12/2020	
	Usability evaluation		Evaluation ongoing
			Tool Delivery to be started

### Initial Conclusions

concerning further counter-measures or business viability [possibly tentative]

1. Robustness and accuracy of data: validated environmental data from trusted sources (see references on Figure 1) and patient's medical records validated by the AVERT Information and Governance Board (see AVERT data management plan)
2. Re-identifiable patient data: Although the data under consideration is de-identified, due to its nature, in practice the data cannot be assumed to be anonymised. A linkage table exists that maps the study ID to the identifiable medical record but I won't have access to the table.

### Compliance Checks required:

Tick	Requirement	Notes [replace guide text with response]
X	Does the project involve processing 'personal data' of any sort?	<p>Yes. For the <b>clinical data</b>, the <b>location</b> is needed: individual patient's electoral district, county or/and hospital location.</p> <p>For the <b>usability evaluation</b> participants will be contacted using their email addresses that may identify them, but their identities would not be retained during the evaluations and will be listed with their own participant ID. There will be no record linking the email address to the study ID.</p>



		Demographics of the participants as unique categories: the researchers are international professors, researchers and PhD students with fluent English, who are analysing AAV clinical data in their research.
X	Does the project involve processing 'confidential data' of any sort?	Yes. <b>Dates:</b> individual patient dates for diagnosis, flare events and disease activity (e.g. yyyy-mm-dd).  Patient's location and related dates could lead to re-identification when combined with environmental data. Therefore, demonstrations, reports and publications about my project will not display actual individual level patient data but metadata with example data.
<b>Data Availability requirements</b>		
0	Does data need to be held for GCP compliance?	No. My intention is not to run clinical trials but to provide accurate associated environmental data for individual health events.
X	Does data need to be held to meet 'Open Data' requirements?	Yes. Any data that I generate is going to be accessible to the scientific community as example data, metadata, code and workflow. Results published in scientific, computer science and biomedical conferences and journals will remain archived to meet research governance requirements.
X	Does data need to be held to meet ICMJE requirements or commitments?	Yes. Required for most leading biomedical journals.

DPR Compliance Checklist – where 'personal data' is processed:

Tick	Requirement	Notes [replace guide text with response]
------	-------------	--



**Article 5: Principles compliance checks**

X	a) Is processing lawful, fair, and transparent?	<p>Yes. A data sharing agreement (DSA) was signed for the use of clinical data (AVERT) in the research. RKD data (patient’s medical records) have been ethically approved for the purpose of conducting scientific research.</p> <p>Two outstanding data sharing agreements will be signed in the course of my research for the KD and FAIRVASC projects regarding the use of clinical data.</p> <p>The data obtained from the dashboard usability experiments is gathered with the ethics approval from SCSS in TCD.</p> <p>The compliance for linked data generated as a result of the SERDIF framework is assessed in this DPIA.</p>
X	b) Is the purpose (or purposes) of the processing clearly defined	<p>Yes. The purpose is to conduct scientific research developing a framework to support researchers that require a flexible methodology to integrate environmental data with longitudinal and geospatial diverse clinical data. Towards the goal of predicting flares for rare disease patients with statistical models. The above are also included in the ethics documentation for the usability study conducted within this research.</p>
X	c) adequate, relevant and limited to what is necessary	<p>Yes. Data minimization has been applied: patient’s location and dates (as defined above) are necessary to link clinical and environmental data.</p>
X	d) accurate and, where necessary, kept up to date	<p>The data sources listed above ensure the validation of the data. Versioning of the data</p>





		will be included as part of the metadata for each data source.
X	e) kept and permits identification of data subjects for no longer than is necessary	<p>As represented in Figure 1, the data remains in the original database within each research project (refer to Project Background section). Only the required fields of the clinical data will be used to associate environmental data to it 8 (as described above).</p> <p>The data gathered during the usability experiment follows the ethics approval from the SCSS in TCD</p> <p>Standard scientific research retention periods will apply and receive support under GDPR Article 89.</p>
X	f) processed securely	<p>Encrypted personal laptop. The research data will be held in an approved “safe haven” in TCD. Advanced data protection methods to protect electronic information, coding the data (name, location data or other identifiable information is not used).</p> <p>Patient’s personal details and medical record numbers will be recorded on the research log which allows, with consent, to link personal information with information from other relevant studies you may have participated in, in particular the RKD registry.</p> <p>This log will be kept securely in a separate place to the coded research data. Access to these identifiable data will be strictly controlled in accordance with the data management plan (see AVERT data management plan).</p>



		Governing data sharing agreements will be in place to access and use the clinical data (Refer section above).
X	2) can you demonstrate this compliance?	Yes. Data sources ensure the robustness of the data: the data controller demonstrates this compliance in the data management plan (see AVERT data management plan) and data sharing agreements.
<b>Articles 13 &amp; 14 compliance</b>		[See detailed Transparency Checklist below]
X	Did the data come from publicly accessible sources?	Environmental and geometry data are from a public source but clinical registry data requires a data sharing agreement with the corresponding data controller entity.  Data obtained in the usability studies comes from each participant and their performance in the experiment.
X	Are data subjects informed before processing starts for any new purpose if incompatible with the original purpose where the controller wants to use data for a different purpose to the purpose for which they currently hold data?	Yes. Data processing and use for other purposes will require explicit patient approval. The approval given by data subjects is contained within the patient's permission in the consent form. The same applies for the participants in the usability study.
X	Does the Privacy Notice and/or PIL cover this processing?	Refer to patient's information sheet and participant's information sheet sent to participate in the usability study  The data use is transparent with regards of the RKD details <sup>20</sup> .
X	What patient choices are available? Are these explained?	Refer to the patient's consent form and participant's signed consent to participate in the usability study.

<sup>20</sup> <https://www.tcd.ie/medicine/thkc/research/rare.php>



		The data use is transparent with regards of the RKD details <sup>21</sup> .
<b>Articles 6 and 9: legal bases</b>		
X	What are legal bases under Article 6	6.1(e) [Monitor consent legal basis secondary legislation requirement] <sup>22</sup>
X	What are legal bases under Article 9 (if 'special category' data)	Yes. 9.2(j) public interest research
0	Are Article 6 legitimate interests explained where relevant?	N/A
X	Are details of statutory obligations for Article 6 explained where relevant.	Yes. Under 6.1(e) processing is necessary for the performance of a task carried out in the public interest.
X	Is this proposed processing compatible with the declared purposes?	[Check against any privacy notices and public information]  Yes. The purpose stated above and in the data sharing agreement for verification I will have to check with the two remaining DSA from the other projects.
<b>Article 89(1) research exemption</b>		
X	If for research, do we meet Art 89(1) data minimisation	Yes. As above to be confirmed. I am using the minimum personal and confidential data to achieve my research goal
<b>Articles 15-23: Data Subject Rights</b>		[See detailed table below]
X	Do we support data subject rights?	[If data is pseudo-/anonymised, then it would be difficult/impossible to do so]  The work in the project will be conducted under the RKD, KD, FAIRVASC privacy notices and/or Public Interest Litigation (PIL)

<sup>21</sup> <https://www.tcd.ie/medicine/thkc/research/rare.php>

<sup>22</sup> Irish regulation currently updated



		document. Data used in part of this project will in any event be anonymized.
0	There is no use of automated decision making (e.g. profiling)	[Otherwise need at least a 'discussion note']  No. Classification of patients cannot have a negative impact on the individuals.
<b>Articles 24-43: Controller-Processor</b>		
X	A28 & 29: What measures are there to ensure processors comply?	EU Model Contract Clauses and a Data Processing Amendment which means all data must remain in countries which meet the EU's "adequacy" standard for privacy protection.  Data sharing agreement formally signed and two other data sharing agreements being developed.
X	A30: Is there an entry for this processing/data held in the register?	Check with the institutes  Part of the purpose or a new purpose?
X	A32-34: Do we ensure appropriate security, including protection against unauthorised or unlawful processing and against accidental loss, destruction or damage, using appropriate technical or organisational measures?	Trinity College Dublin is committed to adopting a security model in line with the ISO27001/ISO27002 international best practice standards.
X	A37-39: Is there a DPO and have they been or will they be consulted?	Reviewed by the DPO in TCD
<b>Articles 44-50: International transfers</b>		
0	What form of data will be transferred to a third country or international organisation	[describe nature of data and whether identified, identifiable, de-identified or anonymous]  KD data from Japan has already been transferred to the EU and I won't be transferring this data back to a non-EU or third country. All signing parties of the data



		sharing agreements and collaborators are within the EU.
0	Are there safeguards for international transfers?	[e.g. US Privacy Shield, anonymisation, GDPR equivalence, approved contractual clauses, or BCR]  The data that I am currently using are not being transferred to any laboratories or places out-side the consortium. If any international transfers are to occur, the DPIA will be revisited.
<b>Article 90: Obligations of secrecy</b>		
X	Do we meet medical confidentiality requirements?	[Note any national case law and statutory requirements that may affect this]  Yes. Working in line with RKD, KD and FAIRVASC registries governance (see projects Data Management Plans).

Data Subject Rights:

Note if supported and what process/procedure applies; if not, then describe the legal justification for not supporting this right.

This section refers to the AVERT, KD and FAIRVASC Participant Information sheet and Consent form (see references from Section Project Background)

X	To be informed: about processing, about choices, about rights, about controller	<b>For RKD data:</b> please contact Professor Mark Little on 01-896-2145, <a href="mailto:mlittle@tcd.ie">mlittle@tcd.ie</a> or the Study Coordinator on 085-150-7587, <a href="mailto:rkd nurse@tcd.ie">rkd nurse@tcd.ie</a> .  For KD and FAIRVASC, the contact data for this section will be updated in the following year.
X	the right of access to see or receive a printed copy	Participants may ask to see a copy of the information we hold (except where it is



		de-identified) and for a 'portable' copy of any data provided.
X	the right to rectification – to correct any material errors in the personal data	<p>If the participant identifies inaccuracies in the data held by the AVERT study team, they can notify the research nurse about this.</p> <p>KD and FAIRVASC right to rectification will be provided in the following year.</p>
X	the right to erasure – where appropriate, to ask that all personal data is erased	<p>In accordance with GDPR legislation, if a participant requests for their data to be erased, this will be managed by Professor Mark Little. It will not be possible to erase data that have already been used in a scientific manuscript or collaboration.</p> <p>For KD and FAIRVASC, the data for this section will be updated in the following year.</p>
X	the right to restrict processing – to ask that some or all processing ceases [see opt-out]	<p>Participants may object to any further processing of the information we hold (except where it is de-identified).</p> <p>For KD and FAIRVASC, the data for this section will be updated in the following year.</p>
0	the right to data portability – this only applies to data provided directly by individual	Only identifiable by the data controllers from the AVERT project
0	the right to object to and not to be subject to automated decision-making, including profiling	No automatic decision-making
X	Right to object to a Data Processing Authority (typically the relevant supervisory authority of each Member State)	<p>Depends on legal bases that we use</p> <p>Data protection officer will receive the request.</p>



X	Where consent is the legal basis, the right to withdraw consent	If you withdraw from the study, the information that we have obtained up to the point of you coming out of the study will continue to be used for the purpose of the study. If your data have already been used at the time you withdraw, it may be impossible to withdraw the results once they have been compiled with the results of others participating in the study, or if they have contributed to a published paper.
---	---	--

Detailed Transparency Checklist<sup>23</sup>

Does privacy information provided to data subjects include:

This section will be updated with KD and FAIRVASC projects relevant data within the following year

X	The name and contact details of our organisation	AVERT, HELICAL, RKD Registry and HSE
X	The name and contact details of our representative (if applicable)	Prof. Mark Little, <a href="mailto:mlittle@tcd.ie">mlittle@tcd.ie</a>
X	The contact details of our data protection officer (if applicable)	Data Protection Officer, <a href="mailto:dataprotection@tcd.ie">dataprotection@tcd.ie</a>
X	The purposes of the processing	To conduct scientific research
X	The lawful bases for the processing	[Art6 for 'personal data' & Art9 for 'special category']  The legal basis of processing reflected in the above section.
0	The legitimate interests for the processing (if applicable)	Not applicable

<sup>23</sup> Taken from UK Information Commissioner's Office template



0	The categories of personal data obtained (if the personal data is not obtained from the individual it relates to)	[for Art14] Personal data is only obtained from the individual it relates to
0	The recipients or categories of recipients of the personal data	No data sending
X	The details of transfers of the personal data to any third countries or international organisations (if applicable)	The General Data Protection Regulation (GDPR) (Regulation (EU)2016/679) as enacted in May 2018 addresses the export of personal data outside the EU
X	The retention periods for the personal data.	Directive 2006/24/EC of 15 March 2006 on the retention of data generated or processed .
X	The rights available to individuals in respect of the processing	The General Data Protection Regulation (GDPR) (Regulation (EU)2016/679) as enacted in May 2018, which strengthens and unifies data protection for all individuals within the European Union (EU).
X	The right to withdraw consent (if applicable)	If the patient decides to withdraw from the study at any stage, the research nurse/research team member will document this decision clearly in the patient's medical notes and CRF and ECRF, detailing the reason if known.
X	The right to lodge a complaint with a supervisory authority	Concerns or complaints about any aspect of the way you have been approached or treated during this study, you should contact Professor Mark Little on 01-896-2145 or the St James's Hospital CRF governance unit 01 410 3906 or <a href="mailto:dataprotection@stjames.ie">dataprotection@stjames.ie</a> .
X	The source of the personal data (if the personal data is not obtained from the individual it relates to)	[For Art14] Personal data is obtained directly from patients and participants in





		the SERDIF dashboard usability experiment.
X	The details of whether individuals are under a statutory or contractual obligation to provide the personal data  (if applicable, and if the personal data is collected from the individual it relates to)	Patients sign a consent form (see AVERT's Participant Consent form)
0	The details of the existence of automated decision-making, including profiling (if applicable)	No automated decision-making or profiling
X	We provide individuals with privacy information at the time we collect their personal data from them – or where obtain personal data from a source other than the individual it relates to, we provide them with privacy information	Participant Info sheet and Consent form are provided to the individuals. (see AVERT's Participant Information sheet)
X	within a reasonable of period of obtaining the personal data and no later than one month	Participant Info sheet and Consent form are provided to the individuals.
X	if we plan to communicate with the individual, at the latest, when the first communication takes place	Participant Info sheet and Consent form are provided to the individuals.
X	if we plan to disclose the data to someone else, at the latest, when the data is disclosed	Participant Info sheet and Consent form are provided to the individuals.
X	We provide the information in a way that is:  <input type="checkbox"/> concise;  <input type="checkbox"/> transparent;  <input type="checkbox"/> intelligible;  <input type="checkbox"/> easily accessible; and  <input type="checkbox"/> uses clear and plain language.	[Describe how we check is Plain English, etc.]  Refer to Participant Information sheet
X	When drafting the information, we:	[Note: best practice advice]



	<input type="checkbox"/> undertake an information audit to find out what personal data we hold and what we do with it.  X put ourselves in the position of the people we're collecting information about.  <input type="checkbox"/> carry out user testing to evaluate how effective our privacy information is	We had interactions as presentations and informal talks with the participants to try to comprehend their position.  This is with reference to the process of RKD regarding transparency and for the other two projects this will be revisited.
X	When providing our privacy information to individuals, we use a combination of appropriate techniques, such as:  <input type="checkbox"/> a layered approach;  <input type="checkbox"/> dashboards;  <input type="checkbox"/> just-in-time notices;  <input type="checkbox"/> icons; and  X mobile and smart device functionalities.	[Note: best practice advice]  This is handled by the registers and biobanks themselves.

### Security & Access Control Checklist

Controls need to be appropriate to the level of risk: identified special category data needs more protection against potential misuse than non-personal data.

This section will be updated with KD and FAIRVASC projects relevant data within the following year

	Data Security classification (above Official)	X - Official-Sensitive  <input type="checkbox"/> - Secret  <input type="checkbox"/> - Top Secret  <input type="checkbox"/> - Public Domain
X	Personal Data involved [GDPR]	Patient's location: Electoral Division.
X	Special Category of personal data involved [GDPR]	De-identified patient's medical registries



0	Electronic Communications (inc. cookies) [PECR]	
0	Credit Card data	
0	Legal enforcement [LED2018]	
0	Financial data	
X	Intellectual Property (detail owner)	
X	Commercial in confidence (detail owner)	No personal data will be used for commercial purposes, although knowledge derived from the research using the personal data may be brought forward to such use as appropriate.
X	Data Location (storage or processing) (include any back-up site(s))	<input type="checkbox"/> - UK <input checked="" type="checkbox"/> - EU/EEA <input type="checkbox"/> - EU White-list <input type="checkbox"/> - USA <input type="checkbox"/> - Other:
X	Is data held in a secure data centre?	[detail centre and what certification supports assertion]  ADAPT server: located on the TCD Virtual Machine and Docker cluster.
0	Is this a new supplier, location, or system?	[If so, need specific IS check; also need formal contract]  No, it is the same Trinity College Dublin in the ADAPT servers.
X	Is all user access subject to 2-factor authentication?	<input type="checkbox"/> - no control <input checked="" type="checkbox"/> - single factor (e.g. just password) <input type="checkbox"/> - 2-factor (e.g. password & fob)



		<input type="checkbox"/> - biometric [note: GDPR reqs] <input type="checkbox"/> - Other control:
0	Are there established JML procedures?	[Joiners, Movers, Leavers] Check policy from TCD
X	Are there checks that passwords are robust and secure enough?	Information Governance Board (IGB) using password protection, and aligning to the United States National Institute of Standards and Technology (NIST) digital authentication guidelines, NIST SP 800-63B-3
X	Are all administrator & user accounts routinely monitored?	<p>[Particularly for redundant or little used accounts]</p> <p>Staff access to the database and content system is restricted and monitored.</p>
X	Are systems protected against malware and other attacks?	<p>[provide details of protection software and procedures]</p> <p>Two firewalls: between our subnet and the host School of Computer Science and Statistics network and TCD firewall. For Apache web servers, we use a tool called Nikto (<a href="https://cirt.net/nikto2">https://cirt.net/nikto2</a>) to scan every month all the websites hosted in our cluster for known vulnerabilities. For all web servers, we expose them through our reverse proxy.</p>

[Need some aspect of CIA/impact-likelihood assessment]

#### Information Asset Register Checklist

0	Are there new IAs being created?	<p>[provide details]</p> <p>Check TCD</p>
0	Are old IAs being retired?	[provide details]
0	Have IAOs & IACs been consulted?	



0	Has IAR been updated/amended?	[at least create project task to do so]
X	Data Retention classification & period	Directive 2006/24/EC of 15 March 2006 on the retention of data generated or processed in connection with the provision of publicly available electronic communication services or of public communications networks.
X	Data retention procedure/functionality in place	Until the end of the HELICAL project, taking into consideration data sharing agreements, Irish and European jurisdiction.



## Appendix A – Supervisory Authority ‘High Risk’ Check

If the DPIA shows ‘high risk’ processing which cannot be mitigated, then the DPIA should be sent to the relevant authority for review before any processing starts. Note that their review may take several weeks to process. A ‘High Risk’ assessment represents a ‘risk to the rights and freedoms of individuals’ – so may extend beyond GDPR consideration, including Human Rights.

GDPR Article 35(3) provides three examples:

- a) a systematic and extensive evaluation of personal aspects relating to natural persons which is based on automated processing, including profiling, and on which decisions are based that produce legal effects concerning the natural person or similarly significantly affect the natural person;
- b) processing on a large scale of special categories of data referred to in Article 9(1), or of personal data relating to criminal convictions and offences referred to in Article 1013; or
- c) a systematic monitoring of a publicly accessible area on a large scale

ICO cites:

1. Systematic and extensive profiling with significant effects
2. Large scale use of sensitive data [viz. ‘special category’ in GDPR terms]
3. Public monitoring

These being the same as (a)-(c) above. They further identify:

1. **New technologies:** processing involving the use of new technologies, or the novel application of existing technologies (including AI).
2. **Denial of service:** Decisions about an individual’s access to a product, service, opportunity or benefit which is based to any extent on automated decision-making (including profiling) or involves the processing of special category data.
3. **Large-scale profiling:** any profiling of individuals on a large scale.
4. **Biometrics:** any processing of biometric data.
5. **Genetic data:** any processing of genetic data, other than that processed by an individual GP or health professional for the provision of health care direct to the data subject.
6. **Data matching:** combining, comparing or matching personal data obtained from multiple sources.



7. **Invisible processing:** processing of personal data that has not been obtained directly from the data subject in circumstances where the controller considers that compliance with Article 14 would prove impossible or involve disproportionate effort.
8. **Tracking:** processing which involves tracking an individual’s geolocation or behaviour, including but not limited to the online environment.
9. **Targeting of children or other vulnerable individuals:** The use of the personal data of children or other vulnerable individuals for marketing purposes, profiling or other automated decision-making, or if you intend to offer online services directly to children.
10. **Risk of physical harm:** Where the processing is of such a nature that a personal data breach could jeopardise the [physical] health or safety of individuals.

**‘High Risk’ assessment using ICO criteria:**

How data is used in the fields below

Criterion:	Assessment	Comments
New technologies	Low	Environmental data is associated with personal health data (location and time of an event) using semantic queries.  <b>Mitigation:</b> the personal health data is only consulted.
Denial of service	N/A	No decisions about an individual’s access to a product, service, opportunity or benefit
Large-scale profiling	N/A	No profiling, only associating environmental data to health events.
Biometrics	N/A	I do not use biometric data in this project
Genetic data	N/A	At this stage of the process I do not handle genetic data. It is possible that towards the end of the project I might include patient’s biomarkers data in the study.
Data matching	Medium	The location and time from personal health events is used to associate an environmental record to each event.



		<p><b>Mitigation:</b> The processing is computed on encrypted laptops that access and consult the health data. Event-environmental linked data won't be published as open data, only example data, a generic metadata record together with the workflows and code.</p>
Invisible processing	N/A	All personal data has been obtained directly from the patients and participants with their consent.
Tracking	Medium	<p>Personal health data used includes the location of the event (i.e. electoral district or hospital), which is used to infer the associated region (i.e. county) in the semantic query.</p> <p><b>Mitigation:</b> only the country will be published in the metadata record together with the period, several years, of the study.</p>
Targeting of children or other vulnerable individuals	N/A	Not using personal data of children or other vulnerable individuals for marketing purposes, profiling or other automated decision-making.
Risk of physical harm	N/A	I work with de-identified patient's medical records within a patient registry (RKD and FAIRVASC) with an ID per patient, and a national epidemiologic survey data for KD in Japan. A linkage table exists that maps the study ID to the identifiable medical record but I won't have access to the table. However, there is no impact on [physical] health or safety of individuals during my research.

[The assessment can be one of N/A (not applicable), Low, Medium, or High. The comments should explain how the assessment is justified.]





Appendix B – Broad Privacy Risk Assessment:

#	Risk Description/detail	Discussion
1.	Data accuracy and timeliness	<p>[Is data accurately recorded &amp; kept up-to-date?]</p> <p>There is no data gathering per se, only linkage of diverse data sources. The linkage is provided with provenance metadata; therefore, if an accuracy is spotted we can address the issues without jeopardizing the existing data structure. The only data collected is the usability metrics for the dashboard evaluation.</p>
2.	Differential treatment of patients/data subjects	<p>[Might certain categories of people be adversely affected, e.g. children, vulnerable adults]</p> <p>The clinical data is combined with specific county aggregated environmental data. This association could have an impact on certain neighbourhoods if the result of the study is that the environment variables are related to patients' flares, which are more likely to be attributed to geographic areas.</p>
3.	Data Accuracy and identification	<p>[Is the identification of individuals reliable? Is there a danger of mis-attribution or incorrect linkage of data?]</p> <p>Doctors, nurses and researchers check the attribution of the clinical data, and the environmental data is from trusted sources.</p>
4.	Holding / sharing / use of excessive data within i~HD systems	<p>[Might too much data be held or for long? Is there a clear justification for data retention? Not 'just in case']</p> <p>Patient's information and national survey epidemiological data is meant to be kept as records in the RKD, KD, and FAIRVASC databases beyond the duration of my PhD - in a perpetual manner. Environmental and geometry data is publicly available. Participant's performance metrics recorded during the usability experiment are anonymized.</p>
5.	Data held too long within i~HD systems	<p>[Is there a clear data retention period specified and are there processes to ensure its deletion when no longer needed? Are copies tracked and deleted as well?]</p>



		<p>Clinical data retention period is specified in each data management plan and consent is signed by the patients of the cohorts (see section above for the source of the documents).</p> <p>Consent is sought from patients for both perpetual data preservation and for the sharing with relevant research groups, as approved by the information governance board. AVERT intends to retain these data since it is unclear at this point which data sources will be the most relevant, or which further data sources may become available in future (potentially making seemingly insignificant variables more important) (See AVERT data management plan)</p> <p>If the AVERT group decides to disband we will deposit the datasets and algorithm / algorithm provenance in an accredited archive, respecting data protection requirements. Copies of the data are tracked by the data controller.</p>
6.	Excessive range of access in terms of users to personal data (consider new users/change of access privileges)	<p>[Do more users have access than strictly necessary? Are user roles clear distinguished and reflected in the access privileges? Is there a clear process for granting and revoking access privileges?]</p> <p>Access is only granted manually through a request to the data controllers of the AVERT, KD and FAIRVASC projects.</p>
7.	Potential for misuse of data, unauthorised access to systems	<p>[What are the likely threats to the data? What countermeasures are or might be applied? Is it possible for access to be granted inappropriately?]</p> <p>The data controllers of the AVERT, KD and FAIRVASC projects are responsible for the security of the hosted clinical data. The framework can only integrate the clinical, environmental and geometry data if the user opens an ssh tunnel with the approved credentials.</p>
8.	New sharing of data with other organisations, including new or change of suppliers	<p>[Is data being shared from new data providers or with new data users? Are there new suppliers or data processors? What controls will apply?]</p> <p>New environmental data sources could be incorporated during the PhD, which will be only for validated data sources.</p>



9.	Variable and inconsistent adoption / implementation	<p>[How well will this system work end-to-end? How robust is it against partial adoption or system failure?]</p> <p>The dashboard gives error messages if the user selects an option that is not available. If a critical error appears, the user has only to refresh the web app from the reload button on the browser. The data structures are built with Semantic Web technologies, a robust information architecture by nature.</p>
10.	Legal compliance, particularly DP transparency requirements and support for data subject rights	<p>[How well does this system meet legal requirements – or appear to meet legal requirements? Does it meet the ‘No surprises’ rule? What would happen if an individual requests data erasure or ceasing processing, etc.]</p> <p>The RDF triplestore used in the SERDIF framework allows the removal of environmental data with specific commands. If a patient asks for removal of their data, this will be handled by the data controllers of the clinical data.</p>
11.	Medical confidentiality	<p>[Are there any additional sensitivities over confidentiality? Might specific approval (e.g. REC) be required to support this processing?]</p> <p>Signing a data sharing agreement form is needed to access and process the de-identified medical records. Only the data owner and controllers have access to the identifiable medical records, the sensitive and confidential data.</p>



## IG Assessment Checklist ESR2 – T cell repertoires in giant cell arteritis

### Introduction to IG Assessment process

Under the General Data Protection Regulation (GDPR), a Data Protection Impact Assessment (DPIA) is only required where proposed data processing is “likely to result in a high risk to the rights and freedoms of natural persons” (Article 35(1)). However, Article 35(3) explicitly requires one where there is ‘large-scale’ processing of ‘special category’ (e.g. healthcare) data then a DPIA is required.

One other possibility is that the data being processed is already anonymised (see Recital 26) so falls outside GDPR altogether so that no DPIA is actually required.

However, good project management and information governance suggests that there should be a general approach to risk assessment for any project or business enterprise – if only to determine whether a DPIA might be required.

Ideally, one should work from a simple initial Checklist (this document) which identifies possible areas of information risk and compliance requirements to a ‘discussion note’ which explores any issues in more depth and may help identify the necessary mitigation methods and mechanisms to offset most if not all risks. Only if risks are unmitigated or remain ‘high’ would you move to a formal DPIA report.

### The IG Assessment approach

There should be an overview of the proposed project or business change to explain what processing is envisaged as well as the purpose and intended outcome. The ‘purpose’ is important to establish the legal basis for the processing as well as ensuring that any possible mitigations or counter-measures do not undermine the main rationale for the processing.

The next step is to establish what compliance requirements may apply: GDPR, contractual or other regulatory restrictions, consent requirements, or obligations to preserve the data for legal or other reasons (including the benefit of posterity perhaps).

Once the precise range of obligations has been established, then appropriate checks can be made and recorded within the document.

The most obvious of these being GDPR compliance. There must be a ‘High Risk’ assessment (Appendix A) to determine whether the supervisory authority needs to be informed – generally, it is expected that it will not be necessary; if so, then a formal DPIA report will be needed.

Appendix B has a broader Privacy Impact Assessment that may throw up some broader issues.

Initial conclusions as to next steps or particular countermeasures to be considered should be detailed below.



### Project Background/Overview

[Explain business background, including any existing processes and procedures; outline the project including stages, deliverables, and timelines]

The project aims at training a machine learning algorithm to predict the specificity of T cell receptors. This could then be used to identify possible autoantigens in vasculitis patients. For the first stage of the project, public databases with T cell receptor sequences linked to epitope sequences are used to train the algorithm. No patient specific data is needed. In the second stage of the project (starting around mid-2021) T cell receptor sequencing data as well as bulk RNA sequencing data from patient's blood and biopsies will be used to identify expanded T cell clones and likely autoantigens. The algorithm can then be used to predict binding between them and predict the most likely autoantigens involved.

### Comparison of process steps (simplified): [optional]

This allows identification of what processing is new or changed through the project:

Step	Current	Proposed
Training machine learning algorithm	First training completed	No change
Data sharing agreement with Leeds	In place	
Analysis of RNAseq and TCR sequencing data	Not started yet	



### Initial Conclusions

concerning further counter-measures or business viability [possibly tentative]

3. Since all the potentially personal data is going to be handled by the collaborators in Leeds, who will also have patient contact, inform them etc., I should contact them about information on ethics approvals, DPIA, patient leaflets,...

Compliance Checks required:

Tick	Requirement	Notes [replace guide text with response]
x	Does the project involve processing 'personal data' of any sort?	Patients not identifiable through data, but sequences linked to anonymous patients
<input type="checkbox"/>	Does the project involve processing 'confidential data' of any sort?	RNA and TCR sequences are not considered confidential data
<b>Data Availability requirements</b>		
<input type="checkbox"/>	Does data need to be held for GCP compliance?	Not by me
x	Does data need to be held to meet 'Open Data' requirements?	yes
x	Does data need to be held to meet ICMJE requirements or commitments?	yes



GDPR Compliance Checklist – where ‘personal data’ is processed:

Tick	Requirement	Notes [replace guide text with response]
<b>Article 5: Principles compliance checks</b>		
x	g) Is processing lawful, fair, and transparent?	Yes, ethics approval obtained by experimentalists
x	h) Is the purpose (or purposes) of the processing clearly defined	Yes, detecting T cells reacting against autoantigen
x	i) adequate, relevant and limited to what is necessary	Patient sample data will be limited and relevant
x	j) accurate and, where necessary, kept up to date	Accurate and not changing
x	k) kept and permits identification of data subjects for no longer than is necessary	Shouldn't reveal identification at all
x	l) processed securely	Yes, secure data sharing and access restricted, data safe haven at IBM
<input type="checkbox"/>	3) can you demonstrate this compliance?	How?
<b>Articles 13 &amp; 14 compliance</b>		[See detailed Transparency Checklist below]
x	Did the data come from publicly accessible sources?	Yes, data cannot be out of date
<input type="checkbox"/>	Are data subjects informed before processing starts for any new purpose if incompatible with original purpose where the controller wants to use data for a different purpose to the purpose for which they currently hold data	Data subjects can only be contacted by medical team in Leeds
<input type="checkbox"/>	Does the Privacy Notice and/or PIL cover this processing?	Need to ask Leeds collaborators
<input type="checkbox"/>	What patient choices are available? Are these explained?	[see also Data Subject Rights below]
<b>Articles 6 and 9: legal bases</b>		



Tick	Requirement	Notes [replace guide text with response]
x	What are legal bases under Article 6	<i>(e) processing is necessary for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller;</i>
x	What are legal bases under Article 9 (if 'special category' data)	Article 9.2 (j) scientific research in the public interest
<input type="checkbox"/>	Are Article 6 legitimate interests explained where relevant?	[Complete an LIA form]
<input type="checkbox"/>	Are details of statutory obligations for Article 6 explained where relevant.	[Quote statutes or regulation]
x	Is this proposed processing compatible with the declared purposes?	yes
<b>Article 89(1) research exemption</b>		
x	If for research, do we meet Art 89(1) data minimisation	yes
<b>Articles 15-23: Data Subject Rights</b>		[See detailed table below]
<input type="checkbox"/>	Do we support data subject rights?	Data anonymized
<input type="checkbox"/>	There is no use of automated decision making (e.g. profiling)	No profiling
<b>Articles 24-43: Controller-Processor</b>		
x	A28 & 29: What measures are there to ensure processors comply?	Data Sharing agreement with Leeds
x	A30: Is there an entry for this processing/data held in the register?	No





Tick	Requirement	Notes [replace guide text with response]
x	A32-34: Do we ensure appropriate security, including protection against unauthorised or unlawful processing and against accidental loss, destruction or damage, using appropriate technical or organisational measures?	IBM high security measures, backups, Box (safe cloud), data safe haven
x	A37-39: Is there a DPO and have they been or will they be consulted?	No
<b>Articles 44-50: International transfers</b>		
	What form of data will be transferred to a third country or international organisation	Data of UK patients is stored in UK and accessed from Switzerland  Anonymous TCR and RNA sequencing data
<input type="checkbox"/>	Are there safeguards for international transfers?	Data sharing agreement
<b>Article 90: Obligations of secrecy</b>		
<input type="checkbox"/>	Do we meet medical confidentiality requirements?	[Note any national case law and statutory requirements that may affect this]

#### Data Subject Rights:

Note if supported and what process/procedure applies; if not, then describe the legal justification for not supporting this right.

- ➔ These points will have to be addressed as soon as the patient samples are being collected by the medical team in Leeds; I need to ask them about the information and choices of the patients.

<input type="checkbox"/>	To be informed: about processing, about choices, about rights, about controller	
<input type="checkbox"/>	the right of access to see or receive a printed copy	



<input type="checkbox"/>	the right to rectification – to correct any material errors in the personal data	
<input type="checkbox"/>	the right to erasure – where appropriate, to ask that all personal data is erased	
<input type="checkbox"/>	the right to restrict processing – to ask that some or all processing ceases [see opt-out]	
<input type="checkbox"/>	the right to data portability – this only applies to data provided directly by individual	
<input type="checkbox"/>	the right to object to and not to be subject to automated decision-making, including profiling	
<input type="checkbox"/>	Right to object to a Data Processing Authority (typically the relevant supervisory authority of each Member State)	
<input type="checkbox"/>	Where consent is the legal basis, the right to withdraw consent	

#### Detailed Transparency Checklist<sup>24</sup>

Does privacy information provided to data subjects include:

<input type="checkbox"/>	The name and contact details of our organisation	
<input type="checkbox"/>	The name and contact details of our representative (if applicable)	
<input type="checkbox"/>	The contact details of our data protection officer (if applicable)	
<input type="checkbox"/>	The purposes of the processing	
<input type="checkbox"/>	The lawful bases for the processing	[Art6 for 'personal data' & Art9 for 'special category']

<sup>24</sup> Taken from UK Information Commissioner's Office template



<input type="checkbox"/>	The legitimate interests for the processing (if applicable)	
<input type="checkbox"/>	The categories of personal data obtained (if the personal data is not obtained from the individual it relates to)	[for Art14]
<input type="checkbox"/>	The recipients or categories of recipients of the personal data	
<input type="checkbox"/>	The details of transfers of the personal data to any third countries or international organisations (if applicable)	
<input type="checkbox"/>	The retention periods for the personal data.	
<input type="checkbox"/>	The rights available to individuals in respect of the processing	
<input type="checkbox"/>	The right to withdraw consent (if applicable)	
<input type="checkbox"/>	The right to lodge a complaint with a supervisory authority	
<input type="checkbox"/>	The source of the personal data (if the personal data is not obtained from the individual it relates to)	[For Art14]
<input type="checkbox"/>	The details of whether individuals are under a statutory or contractual obligation to provide the personal data  (if applicable, and if the personal data is collected from the individual it relates to)	
<input type="checkbox"/>	The details of the existence of automated decision-making, including profiling (if applicable)	
<input type="checkbox"/>	We provide individuals with privacy information at the time we collect their personal data from them – or where we obtain personal data from a	



	source other than the individual it relates to, we provide them with privacy information	
<input type="checkbox"/>	within a reasonable of period of obtaining the personal data and no later than one month	
<input type="checkbox"/>	if we plan to communicate with the individual, at the latest, when the first communication takes place	
<input type="checkbox"/>	if we plan to disclose the data to someone else, at the latest, when the data is disclosed	
<input type="checkbox"/>	We provide the information in a way that is: <ul style="list-style-type: none"> <li><input type="checkbox"/> concise;</li> <li><input type="checkbox"/> transparent;</li> <li><input type="checkbox"/> intelligible;</li> <li><input type="checkbox"/> easily accessible; and</li> <li><input type="checkbox"/> uses clear and plain language.</li> </ul>	[Describe how we check is Plain English, etc.]
<input type="checkbox"/>	When drafting the information, we: <ul style="list-style-type: none"> <li><input type="checkbox"/> undertake an information audit to find out what personal data we hold and what we do with it.</li> <li><input type="checkbox"/> put ourselves in the position of the people we're collecting information about.</li> <li><input type="checkbox"/> carry out user testing to evaluate how effective our privacy information is</li> </ul>	[Note: best practice advice]
<input type="checkbox"/>	When providing our privacy information to individuals, we use a combination of appropriate techniques, such as: <ul style="list-style-type: none"> <li><input type="checkbox"/> a layered approach;</li> <li><input type="checkbox"/> dashboards;</li> </ul>	[Note: best practice advice]



<ul style="list-style-type: none"><li><input type="checkbox"/> just-in-time notices;</li><li><input type="checkbox"/> icons; and</li><li><input type="checkbox"/> mobile and smart device functionalities.</li></ul>	
--	--



### Security & Access Control Checklist

Controls need to be appropriate to level of risk: identified special category data needs more protection against potential misuse than non-personal data.

	Data Security classification (above Official)	<input checked="" type="checkbox"/> - Official-Sensitive <input type="checkbox"/> - Secret <input type="checkbox"/> - Top Secret <input type="checkbox"/> - Public Domain
x	Personal Data involved [GDPR]	yes
<input type="checkbox"/>	Special Category of personal data involved [GDPR]	
<input type="checkbox"/>	Electronic Communications (inc. cookies) [PECR]	
<input type="checkbox"/>	Credit Card data	
<input type="checkbox"/>	Legal enforcement [LED2018]	
<input type="checkbox"/>	Financial data	
x	Intellectual Property (detail owner)	Probably intellectual property of Leeds team
<input type="checkbox"/>	Commercial in confidence (detail owner)	
	Data Location (storage or processing) (include any back-up site(s))	<input checked="" type="checkbox"/> - UK <input checked="" type="checkbox"/> - EU/EEA <input type="checkbox"/> - EU White-list <input type="checkbox"/> - USA <input type="checkbox"/> - Other:
x	Is data held in secure data centre?	IBM data center
<input type="checkbox"/>	Is this new supplier, location, or system?	[If so, need specific IS check; also need formal contract]
<input type="checkbox"/>	Is all user access subject to 2-factor authentication?	<input type="checkbox"/> - no control



		x - single factor (e.g. just password) <input type="checkbox"/> - 2-factor (e.g. password & fob) <input type="checkbox"/> - biometric [note: GDPR reqs] <input type="checkbox"/> - Other control:
<input type="checkbox"/>	Are there established JML procedures?	no
x	Are there checks that passwords are robust and secure enough?	High requirements and changing every 3 months
x	Are all administrator & user accounts routinely monitored?	Every 3 months
x	Are systems protected against malware and other attacks?	High security environment at IBM

[Need some aspect of CIA/impact-likelihood assessment]

#### Information Asset Register Checklist

<input type="checkbox"/>	Are there new IAs being created?	[provide details]
<input type="checkbox"/>	Are old IAs being retired?	[provide details]
<input type="checkbox"/>	Have IAOs & IACs been consulted?	
<input type="checkbox"/>	Has IAR been updated/amended?	[at least create project task to do so]
<input type="checkbox"/>	Data Retention classification & period	
<input type="checkbox"/>	Data retention procedure/functionality in place	



## Appendix A – Supervisory Authority ‘High Risk’ Check

If the DPIA shows ‘high risk’ processing which cannot be mitigated, then the DPIA should be sent to the relevant authority for review before any processing starts. Note that their review may take several weeks to process. A ‘High Risk’ assessment represents a ‘risk to the rights and freedoms of individuals’ – so may extend beyond GDPR consideration, including Human Rights.

GDPR Article 35(3) provides three examples:

- d) a systematic and extensive evaluation of personal aspects relating to natural persons which is based on automated processing, including profiling, and on which decisions are based that produce legal effects concerning the natural person or similarly significantly affect the natural person;
- e) processing on a large scale of special categories of data referred to in Article 9(1), or of personal data relating to criminal convictions and offences referred to in Article 1013; or
- f) a systematic monitoring of a publicly accessible area on a large scale

ICO cites:

- 4. Systematic and extensive profiling with significant effects
- 5. Large scale use of sensitive data [viz. ‘special category’ in GDPR terms]
- 6. Public monitoring

These being the same as (a)-(c) above. They further identify:

- 11. **New technologies:** processing involving the use of new technologies, or the novel application of existing technologies (including AI).
- 12. **Denial of service:** Decisions about an individual’s access to a product, service, opportunity or benefit which is based to any extent on automated decision-making (including profiling) or involves the processing of special category data.
- 13. **Large-scale profiling:** any profiling of individuals on a large scale.
- 14. **Biometrics:** any processing of biometric data.
- 15. **Genetic data:** any processing of genetic data, other than that processed by an individual GP or health professional for the provision of health care direct to the data subject.
- 16. **Data matching:** combining, comparing or matching personal data obtained from multiple sources.
- 17. **Invisible processing:** processing of personal data that has not been obtained direct from the data subject in circumstances where the controller considers that compliance with Article 14 would prove impossible or involve disproportionate effort.
- 18. **Tracking:** processing which involves tracking an individual’s geolocation or behaviour, including but not limited to the online environment.
- 19. **Targeting of children or other vulnerable individuals:** The use of the personal data of children or other vulnerable individuals for marketing purposes, profiling or other automated decision-making, or if you intend to offer online services directly to children.





20. **Risk of physical harm:** Where the processing is of such a nature that a personal data breach could jeopardise the [physical] health or safety of individuals.

**'High Risk' assessment using ICO criteria:**

Criterion:	Assessment	Comments
New technologies	N/A	
Denial of service	N/A	
Large-scale profiling	N/A	
Biometrics	N/A	
Genetic data	Medium	The genetic data generated is from T cell receptors. Every individual has millions of different T cell receptor sequences in their body, only a tiny fraction of this is seen in the sample. It is therefore almost impossible to reconstruct patient identity from T cell receptor sequences, in contrast to other genetic material.
Data matching	N/A	
Invisible processing	N/A	
Tracking	N/A	



Criterion:	Assessment	Comments
Targeting of children or other vulnerable individuals	N/A	
Risk of physical harm	N/A	

[The assessment can be one of N/A (not applicable), Low, Medium, or High. The comments should explain how the assessment is justified.]

Appendix B – Broad Privacy Risk Assessment:

#	Risk Description/detail	Discussion
12.	Data accuracy and timeliness	Cannot be out of date
13.	Differential treatment of patients/data subjects	No differentiation will be made
14.	Data Accuracy and identification	My research will not make individuals identifiable; correct data linkage will be secured.
15.	Holding / sharing / use of excessive data within [Company] systems	Data will and must be held until publication.
16.	Data held too long within [Company] systems	See above
17.	Excessive range of access in terms of users to personal data (consider new users/change of access privileges)	Only people directly working with the data will have access.
18.	Potential for misuse of data, unauthorised access to systems	Only explicitly invited people can view the folders with the data.



#	Risk Description/detail	Discussion
19.	New sharing of data with other organisations, including new or change of suppliers	Not planned
20.	Variable and inconsistent adoption / implementation	Not much processing, so very robust
21.	Legal compliance, particularly DP transparency requirements and support for data subject rights	[How well does this system meet legal requirements – or appear to meet legal requirements? Does it meet the ‘No surprises’ rule? What would happen if an individual requests data erasure or ceasing processing, etc.]
22.	Medical confidentiality	no



## IG Assessment Checklist: ESR4, Harnessing the power of integrated data to investigate environmental exposures on ANCA vasculitis risk

### Introduction to IG Assessment process

Under the General Data Protection Regulation (GDPR), a Data Protection Impact Assessment (DPIA) is only required where proposed data processing is “likely to result in a high risk to the rights and freedoms of natural persons” (Article 35(1)). However, Article 35(3) explicitly requires one where there is ‘large-scale’ processing of ‘special category’ (e.g. healthcare) data then a DPIA is required.

One other possibility is that the data being processed is already anonymised (see Recital 26) so falls outside GDPR altogether so that no DPIA is actually required.

However, good project management and information governance suggests that there should be a general approach to risk assessment for any project or business enterprise – if only to determine whether a DPIA might be required.

Ideally, one should work from a simple initial Checklist (this document) which identifies possible areas of information risk and compliance requirements to a ‘discussion note’ which explores any issues in more depth and may help identify the necessary mitigation methods and mechanisms to offset most if not all risks. Only if risks are unmitigated or remain ‘high’ would you move to a formal DPIA report.

### The IG Assessment approach

There should be an overview of the proposed project or business change to explain what processing is envisaged as well as the purpose and intended outcome. The ‘purpose’ is important to establish the legal basis for the processing as well as ensuring that any possible mitigations or counter-measures do not undermine the main rationale for the processing.

The next step is to establish what compliance requirements may apply: GDPR, contractual or other regulatory restrictions, consent requirements, or obligations to preserve the data for legal or other reasons (including the benefit of posterity perhaps).

Once the precise range of obligations has been established, then appropriate checks can be made and recorded within the document.

The most obvious of these being GDPR compliance. There must be a ‘High Risk’ assessment (Appendix A) to determine whether the supervisory authority needs to be informed – generally, it is expected that it will not be necessary; if so, then a formal DPIA report will be needed.

Appendix B has a broader Privacy Impact Assessment that may throw up some broader issues.

Initial conclusions as to next steps or particular countermeasures to be considered should be detailed below.



### Project Background/Overview

[Explain business background, including any existing processes and procedures; outline the project including stages, deliverables, and timelines]

#### Overview:

There is limited evidence about the role of the environment in mediating or driving the risk for Anti-Neutrophilic Cytoplasmic Antibody-Associated Vasculitis (AAV). This area of research has received little attention, in part, because the causes of AAV are complex and involves a combination of genetic, epigenetic and environmental factors. Recent data from other similar inflammatory rheumatic diseases (like Rheumatoid Arthritis) have shown that short -term exposure to environmental air pollutants are associated with markers of inflammation and disease activity as well as overall disease risk. In the cases AAV, current data are inconclusive and show no consistent association between air pollution exposures and the risk for disease onset.

#### Objectives and deliverables:

1. To investigate the association between environmental exposures (e.g.: air pollution, ambient temperature) and the risk for ANCA-Associated Vasculitis onset.
2. Identify spatial circumstances associated with disease risk

**Project stage:** Currently, access to UK Biobank data has been approved and preliminary analyses are underway. Additional data linkage of environmental data is underway with support from Professor Duncan Lee and Dr Breda Cullen at the university of Glasgow who are experts in this area.

#### Comparison of process steps (simplified): [optional]

This allows identification of what processing is new or changed through the project:

Step	Current	Proposed
Project initiation (data application, GDPR training, curation of research and analysis plan)	Complete	No change
Selection of study participants (cases and c from the UK Biobank)	Complete	No change
Preliminary analysis	Underway	No change



Step	Current	Proposed
Data linkage of environmental exposures to the UK Biobank	Third quarter of 2020	No change
Detailed analysis and reporting of results.	Ongoing	No change

#### Initial Conclusions

concerning further counter-measures or business viability [possibly tentative]

4. None so far
5. None so far

#### Compliance Checks required:

Tick	Requirement	Notes [replace guide text with response]
<input checked="" type="checkbox"/>	Does the project involve processing 'personal data' of any sort?	Yes, self-reported and census-derived demographic data will be used as factors such as income and education are associated with many health outcomes and could mask the association between environmental exposures and AAV risk, if not taken into account.
<input checked="" type="checkbox"/>	Does the project involve processing 'confidential data' of any sort?	Yes, Individual's medical records and postcodes will be requested in order to link neighbourhood-level environmental data to each participant. This will allow us to assess the short and long-term effects of environmental exposures on the risk for disease onset.
<b>Data Availability requirements</b>		
<input checked="" type="checkbox"/>	Does data need to be held for GCP compliance?	Yes



Tick	Requirement	Notes [replace guide text with response]
<input checked="" type="checkbox"/>	Does data need to be held to meet 'Open Data' requirements?	Yes
<input checked="" type="checkbox"/>	Does data need to be held to meet ICMJE requirements or commitments?	Yes



GDPR Compliance Checklist – where ‘personal data’ is processed:

Tick	Requirement	Notes [replace guide text with response]
<b>Article 5: Principles compliance checks</b>		
<input checked="" type="checkbox"/>	m) Is processing lawful, fair, and transparent?	The lawful basis for processing personal data are covered under the new category of the GDPR law, namely <u>legitimate interests</u> and <u>explicit consent</u> (an updated version of <i>consent</i> under the previous law). As and when required, the UKB were specific about which basis was being used for a particular activity; for example, when UK Biobank currently links secondary health care data (such as hospital events and death and cancer information) through NHS Digital, it uses legitimate interests as the appropriate lawful basis.
<input checked="" type="checkbox"/>	n) Is the purpose (or purposes) of the processing clearly defined	The two purposes for processing UKB data are defined in our study objectives, this is namely, to understand the role of the environment in explaining the risk of AAV onset. Secondly, to investigate the spatial circumstances associated with the disease risk.
<input checked="" type="checkbox"/>	o) adequate, relevant and limited to what is necessary	The use of data will be limited to the goal and specific outcomes of the project.
<input checked="" type="checkbox"/>	p) accurate and, where necessary, kept up to date	Both personal data and environmental data will be kept up to date in accordance to industry best practice. Participants who withdraw from the UKB will be removed from our records and the analyses will be updated accordingly.





Tick	Requirement	Notes [replace guide text with response]
<input checked="" type="checkbox"/>	q) kept and permits identification of data subjects for no longer than is necessary	The data will be kept in accordance to the <a href="#">material transfer agreement (MTA)</a> between the University of Glasgow and the UKB. The <a href="#">UofG record and information management services (RIMS)</a> will hold the data as long as the project is deemed possible.
<input checked="" type="checkbox"/>	r) processed securely	Processing of the UKB data is done on the University safe haven platform. For access, a one-step authentication is required when working at the university and a two-step authentication if working remotely via the university secure VPN connection (Cisco AnyConnect Secure Mobility Client)
<input checked="" type="checkbox"/>	4) can you demonstrate this compliance?	The university keeps a record of each of their members user activity as well as provide secure access to the data on its platform. The UKB has right to audit this process with the UofG as stated in the MTA and therefore each data user is recommended to be compliant.
<b>Articles 13 &amp; 14 compliance</b>		[See detailed Transparency Checklist below]
<input type="checkbox"/>	Did the data came from publicly accessible sources?	No, data access is limited to bona fide researchers who are registered to a specific UKB project and with the UofG.
<input checked="" type="checkbox"/>	Are data subjects informed before processing starts for any new purpose if incompatible with original purpose where the controller wants to use data for a different purpose to the purpose for which they currently hold data	Yes, though not applicable to this project as the study objectives are in line with the overall legal obligation and purpose for processing the data; this including that the processing of the data is for legitimate interest.



Tick	Requirement	Notes [replace guide text with response]
<input checked="" type="checkbox"/>	Does the Privacy Notice and/or PIL cover this processing?	Yes – this was communicated to each participant at the beginning of the study. See <a href="#">consent form</a> and <a href="#">information leaflet</a> for more details
<input checked="" type="checkbox"/>	What patient choices are available? Are these explained?	The following choices were communicated to each participant at the time of entering in the UKB. They were made aware they had a choice not to take part in the study, <i>to restrict processing of their data, to be forgotten, erasure and withdraw from the study.</i>
<b>Articles 6 and 9: legal bases</b>		
<input type="checkbox"/>	What are legal bases under Article 6	Explicit consent and Legitimate interest
<input type="checkbox"/>	What are legal bases under Article 9 (if 'special category' data)	Explicit consent, legitimate interest and public interest (Archiving, research and statistics)



<input checked="" type="checkbox"/>	<p>Are Article 6 legitimate interests explained where relevant?</p>	<p>Legitimate interests were explained using a series of questions that covers the 3-step test for “legitimate interest”.</p> <p><i>Purpose test: what are UK Biobank’s legitimate interests?</i></p> <ul style="list-style-type: none"> <li>• <i>What is UK Biobank trying to achieve?</i> Our objective is to set up and manage a major international research resource for health-related research that is in the public interest.</li> <li>• <i>Who benefits from UK Biobank’s processing?</i> Patients and the wider public benefit from the advances made in the prevention, diagnosis and treatment of disease.</li> <li>• <i>How significant/important are these benefits?</i> UK Biobank is now one of the largest and most used health research resources in the world. Over 6,000 institutions are registered with us and over 1,000 health-related research applications have been approved.</li> </ul> <p><i>Necessity test: is the processing necessary for the legitimate interests?</i></p> <ul style="list-style-type: none"> <li>• <i>Is processing personal data a reasonable way to achieve the objective?</i> Without the personal data provided voluntarily by you and the other participants, UK Biobank would not exist.</li> <li>• <i>Is there another less obtrusive way to meet our purposes?</i> Your data are stored in a way that makes it extremely difficult even for UK Biobank to re-identify you. Only a very few individuals within UK Biobank are allowed to do so (and they are strictly monitored) in order that further</li> </ul>
-------------------------------------	---	--



Tick	Requirement	Notes [replace guide text with response]
		<p>information about you can be added. Data provided to researchers have personal identifiers removed so that an individual participant cannot be identified. There are no circumstances in which your data can be processed in a manner that could have an adverse impact on you.</p> <p><i>Balancing test: UK Biobank had to weigh up the participant's interests.</i></p> <ul style="list-style-type: none"> <li>• <i>Would participants expect UK Biobank to use their data this way? Yes; this is what we set out in the information materials provided to participants and in the consent form each of them signed.</i></li> <li>• <i>How likely would a participant be to object? In UKB view, this was very unlikely. During the past 10 years since participants joined UK Biobank during 2006-10, fewer than 800 of the 500,000 participants have withdrawn from the study and asked that we delete all of their information.</i></li> </ul>



Tick	Requirement	Notes [replace guide text with response]
<input type="checkbox"/>	Are details of statutory obligations for Article 6 explained where relevant.	<p>No further action was need here – the Recital 41 of the gdpr law exempts the legal obligation for specific processing activity. This is as long as the overall purpose for processing personal data is compliant with the legal obligation which has sufficiently clear basis in either common law or statute</p> <p>Recital 41 – <i>“Where this Regulation refers to a legal basis or explicit statutory obligation, this does not necessarily require a legislative act as long as the application of the law is foreseeable to those individuals subject to it”</i></p>
<input checked="" type="checkbox"/>	Is this proposed processing compatible with the declared purposes?	Yes, the overall purpose for collecting and processing the data is to support a diverse range of research intended to improve the prevention, diagnosis and treatment of illness, and the promotion of health throughout society. Our purpose aligns with this declared purpose.
<b>Article 89(1) research exemption</b>		
<input checked="" type="checkbox"/>	If for research, do we meet Art 89(1) data minimisation	Yes, only data related to the project objectives will be requested and held.
<b>Articles 15-23: Data Subject Rights</b>		[See detailed table below]



Tick	Requirement	Notes [replace guide text with response]
<input checked="" type="checkbox"/>	Do we support data subject rights?	<p>Yes – subjects rights were communicated to each participant at the beginning of the study via the information leaflet. This included the <i>right to restrict processing, to be forgotten, erasure and withdrawal</i></p> <p>Furthermore, all of personal data are anonymised, and key identifiers have been removed. Only few members of the UKB have access to identifiable information and these individuals are monitored.</p>
<input checked="" type="checkbox"/>	There is no use of automated decision making (e.g. profiling)	N/A
<b>Articles 24-43: Controller-Processor</b>		
<input checked="" type="checkbox"/>	A28 & 29: What measures are there to ensure processors comply?	Measures of compliance are underline in the <a href="#">MTA</a> , setting out a series of obligations incumbent on individual researchers, such as using the UK Biobank data for the approved purpose, paying the access fees, keeping the data secure, returning their findings to UK Biobank and not trying to re-identify any participants.
<input type="checkbox"/>	A30: Is there an entry for this processing/data held in the register?	Yes, copies of the MTA are held by both the data processor and controller (individual researchers, the university and the UK Biobank)



Tick	Requirement	Notes [replace guide text with response]
<input checked="" type="checkbox"/>	A32-34: Do we ensure appropriate security, including protection against unauthorised or unlawful processing and against accidental loss, destruction or damage, using appropriate technical or organisational measures?	The data is stored on the UofG servers. The <a href="#">university also provides guidance</a> on data handling – recommending that users must take care when handling personal data, ensuring that selected passwords are unique and access to the password are restricted to authorised personnel only, among others.
<input type="checkbox"/>	A37-39: Is there a DPO and have they been or will they be consulted?	Yes, if and when necessary. See <a href="#">UofG DPO contacts</a>
<b>Articles 44-50: International transfers</b>		
	What form of data will be transferred to a third country or international organisation	N/A
<input checked="" type="checkbox"/>	Are there safeguards for international transfers?	N/A
<b>Article 90: Obligations of secrecy</b>		
<input checked="" type="checkbox"/>	Do we meet medical confidentiality requirements?	Yes, under the common law – “ <i>where the individual to whom the information relates has consented and where disclosure is necessary to safeguard the individual, or others, or is in the public interest</i> ”

#### Data Subject Rights:

Note if supported and what process/procedure applies; if not, then describe the legal justification for not supporting this right.

<input checked="" type="checkbox"/>	To be informed: about processing, about choices, about rights, about controller	Information provided in the <a href="#">UKB information leaflet</a>
<input checked="" type="checkbox"/>	the right of access to see or receive a printed copy	Information provided in the <a href="#">UKB information leaflet</a>



<input checked="" type="checkbox"/>	the right to rectification – to correct any material errors in the personal data	Information provided in the <a href="#">UKB information leaflet</a>
<input checked="" type="checkbox"/>	the right to erasure – where appropriate, to ask that all personal data is erased	Information provided in the <a href="#">UKB information leaflet</a>
<input checked="" type="checkbox"/>	the right to restrict processing – to ask that some or all processing ceases [see opt-out]	Information provided in the <a href="#">UKB information leaflet</a>
<input checked="" type="checkbox"/>	the right to data portability – this only applies to data provided directly by individual	Information provided in the <a href="#">UKB information leaflet</a>
<input checked="" type="checkbox"/>	the right to object to and not to be subject to automated decision-making, including profiling	Information provided in the <a href="#">UKB information leaflet</a>
<input checked="" type="checkbox"/>	Right to object to a Data Processing Authority (typically the relevant supervisory authority of each Member State)	Information provided in the <a href="#">UKB information leaflet</a>
<input checked="" type="checkbox"/>	Where consent is the legal basis, the right to withdraw consent	Information provided in the <a href="#">UKB consent form</a>





Detailed Transparency Checklist<sup>25</sup>

Does privacy information provided to data subjects include:

<input checked="" type="checkbox"/>	The name and contact details of our organisation	Information provided in the <a href="#">UKB information leaflet</a>
<input checked="" type="checkbox"/>	The name and contact details of our representative (if applicable)	Information provided in the <a href="#">UKB information leaflet</a>
<input checked="" type="checkbox"/>	The contact details of our data protection officer (if applicable)	Information provided in the <a href="#">UKB information leaflet</a>
<input checked="" type="checkbox"/>	The purposes of the processing	Information provided in the <a href="#">UKB information leaflet</a>
<input type="checkbox"/>	The lawful bases for the processing	Information provided in the <a href="#">UKB information leaflet</a>
<input checked="" type="checkbox"/>	The legitimate interests for the processing (if applicable)	Information provided in the <a href="#">UKB information leaflet</a>
<input checked="" type="checkbox"/>	The categories of personal data obtained (if the personal data is not obtained from the individual it relates to)	Information provided in the <a href="#">UKB information leaflet</a> [for Art14]
<input checked="" type="checkbox"/>	The recipients or categories of recipients of the personal data	N/A
<input checked="" type="checkbox"/>	The details of transfers of the personal data to any third countries or international organisations (if applicable)	N/A
<input checked="" type="checkbox"/>	The retention periods for the personal data.	Unsure

<sup>25</sup> Taken from UK Information Commissioner's Office template



<input checked="" type="checkbox"/>	The rights available to individuals in respect of the processing	Information provided in the <a href="#">UKB information leaflet</a>
<input checked="" type="checkbox"/>	The right to withdraw consent (if applicable)	Information provided in the <a href="#">UKB information leaflet</a>
<input checked="" type="checkbox"/>	The right to lodge a complaint with a supervisory authority	Information provided in the <a href="#">UKB information leaflet</a>
<input type="checkbox"/>	The source of the personal data (if the personal data is not obtained from the individual it relates to)	N/A
<input type="checkbox"/>	The details of whether individuals are under a statutory or contractual obligation to provide the personal data (if applicable, and if the personal data is collected from the individual it relates to)	N/A
<input type="checkbox"/>	The details of the existence of automated decision-making, including profiling (if applicable)	N/A
<input checked="" type="checkbox"/>	We provide individuals with privacy information at the time we collect their personal data from them – or where we obtain personal data from a source other than the individual it relates to, we provide them with privacy information	Information provided in the <a href="#">UKB information leaflet</a>
<input checked="" type="checkbox"/>	within a reasonable of period of obtaining the personal data and no later than one month	Information provided in the <a href="#">UKB information leaflet</a>
<input checked="" type="checkbox"/>	if we plan to communicate with the individual, at the latest, when the first communication takes place	One month after entering the study



<input type="checkbox"/>	if we plan to disclose the data to someone else, at the latest, when the data is disclosed	N/A
<input checked="" type="checkbox"/>	<p>We provide the information in a way that is:</p> <ul style="list-style-type: none"> <li><input checked="" type="checkbox"/> concise;</li> <li><input checked="" type="checkbox"/> transparent;</li> <li><input checked="" type="checkbox"/> intelligible;</li> <li><input checked="" type="checkbox"/> easily accessible; and</li> <li><input checked="" type="checkbox"/> uses clear and plain language.</li> </ul>	Information provided in the <a href="#">UKB information leaflet</a>
<input checked="" type="checkbox"/>	<p>When drafting the information, we:</p> <ul style="list-style-type: none"> <li><input checked="" type="checkbox"/> undertake an information audit to find out what personal data we hold and what we do with it.</li> <li><input checked="" type="checkbox"/> put ourselves in the position of the people we're collecting information about.</li> <li><input type="checkbox"/> carry out user testing to evaluate how effective our privacy information is</li> </ul>	[Note: best practice advice]
<input checked="" type="checkbox"/>	<p>When providing our privacy information to individuals, we use a combination of appropriate techniques, such as:</p> <ul style="list-style-type: none"> <li><input type="checkbox"/> a layered approach;</li> <li><input type="checkbox"/> dashboards;</li> <li><input checked="" type="checkbox"/> just-in-time notices;</li> <li><input type="checkbox"/> icons; and</li> <li><input type="checkbox"/> mobile and smart device functionalities.</li> </ul>	[Note: best practice advice]



### Security & Access Control Checklist

Controls need to be appropriate to level of risk: identified special category data needs more protection against potential misuse than non-personal data.

	Data Security classification (above Official)	<input checked="" type="checkbox"/> - Official-Sensitive <input type="checkbox"/> - Secret <input type="checkbox"/> - Top Secret <input type="checkbox"/> - Public Domain
<input checked="" type="checkbox"/>	Personal Data involved [GDPR]	Yes
<input checked="" type="checkbox"/>	Special Category of personal data involved [GDPR]	Yes
<input type="checkbox"/>	Electronic Communications (inc. cookies) [PECR]	N/A
<input type="checkbox"/>	Credit Card data	N/A
<input type="checkbox"/>	Legal enforcement [LED2018]	N/A
<input type="checkbox"/>	Financial data	N/A
<input type="checkbox"/>	Intellectual Property (detail owner)	N/A
<input type="checkbox"/>	Commercial in confidence (detail owner)	N/A
	Data Location (storage or processing) (include any back-up site(s))	<input checked="" type="checkbox"/> - UK <input type="checkbox"/> - EU/EEA <input type="checkbox"/> - EU White-list <input type="checkbox"/> - USA <input type="checkbox"/> - Other:
<input checked="" type="checkbox"/>	Is data held in secure data centre?	Yes, within the University of Glasgow Servers  [detail centre and what certification supports assertion]
<input type="checkbox"/>	Is this new supplier, location, or system?	[If so, need specific IS check; also need formal contract]



<input checked="" type="checkbox"/>	Is all user access subject to 2-factor authentication?	<input type="checkbox"/> - no control <input checked="" type="checkbox"/> - single factor (e.g. just password) <input type="checkbox"/> - 2-factor (e.g. password & fob) <input type="checkbox"/> - biometric [note: GDPR reqs] <input type="checkbox"/> - Other control:
<input type="checkbox"/>	Are there established JML procedures?	[Joiners, Movers, Leavers]
<input checked="" type="checkbox"/>	Are there checks that passwords are robust and secure enough?	[]
<input checked="" type="checkbox"/>	Are all administrator & user accounts routinely monitored?	Carried out by the University of Glasgow  [Particularly for redundant or little used accounts]
<input checked="" type="checkbox"/>	Are systems protected against malware and other attacks?	Carried out by the University of Glasgow  [provide details of protection software and procedures]

[Need some aspect of CIA/impact-likelihood assessment]

#### Information Asset Register Checklist

<input type="checkbox"/>	Are there new IAs being created?	N/A
<input type="checkbox"/>	Are old IAs being retired?	N/A
<input type="checkbox"/>	Have IAOs & IACs been consulted?	N/A
<input type="checkbox"/>	Has IAR been updated/amended?	N/A
<input type="checkbox"/>	Data Retention classification & period	Unsure
<input type="checkbox"/>	Data retention procedure/functionality in place	Unsure



#### Appendix A – Supervisory Authority ‘High Risk’ Check

If the DPIA shows ‘high risk’ processing which cannot be mitigated, then the DPIA should be sent to the relevant authority for review before any processing starts. Note that their review may take several weeks to process. A ‘High Risk’ assessment represents a ‘risk to the rights and freedoms of individuals’ – so may extend beyond GDPR consideration, including Human Rights.

GDPR Article 35(3) provides three examples:

- g) a systematic and extensive evaluation of personal aspects relating to natural persons which is based on automated processing, including profiling, and on which decisions are based that produce legal effects concerning the natural person or similarly significantly affect the natural person;
- h) processing on a large scale of special categories of data referred to in Article 9(1), or of personal data relating to criminal convictions and offences referred to in Article 1013; or
- i) a systematic monitoring of a publicly accessible area on a large scale

ICO cites:

- 7. Systematic and extensive profiling with significant effects
- 8. Large scale use of sensitive data [viz. ‘special category’ in GDPR terms]
- 9. Public monitoring

These being the same as (a)-(c) above. They further identify:

- 21. **New technologies:** processing involving the use of new technologies, or the novel application of existing technologies (including AI).
- 22. **Denial of service:** Decisions about an individual’s access to a product, service, opportunity or benefit which is based to any extent on automated decision-making (including profiling) or involves the processing of special category data.
- 23. **Large-scale profiling:** any profiling of individuals on a large scale.
- 24. **Biometrics:** any processing of biometric data.
- 25. **Genetic data:** any processing of genetic data, other than that processed by an individual GP or health professional for the provision of health care direct to the data subject.
- 26. **Data matching:** combining, comparing or matching personal data obtained from multiple sources.
- 27. **Invisible processing:** processing of personal data that has not been obtained direct from the data subject in circumstances where the controller considers that compliance with Article 14 would prove impossible or involve disproportionate effort.
- 28. **Tracking:** processing which involves tracking an individual’s geolocation or behaviour, including but not limited to the online environment.
- 29. **Targeting of children or other vulnerable individuals:** The use of the personal data of children or other vulnerable individuals for marketing purposes, profiling or other automated decision-making, or if you intend to offer online services directly to children.
- 30. **Risk of physical harm:** Where the processing is of such a nature that a personal data breach could jeopardise the [physical] health or safety of individuals.



**‘High Risk’ assessment using ICO criteria:**

<b>Criterion:</b>	<b>Assessment</b>	<b>Comments</b>
New technologies	N/A	N/A
Denial of service	N/A	N/A
Large-scale profiling	N/A	N/A
Biometrics	N/A	N/A
Genetic data	N/A	N/A
Data matching	Low	When UK Biobank releases data to researchers, these data are released with project-specific randomised ID codes for each participant (i.e. they have been “de-identified”: please see UK Biobank's note on data <a href="#">de-identification protocol</a> )
Invisible processing	N/A	N/A
Tracking	N/A	N/A



<b>Criterion:</b>	<b>Assessment</b>	<b>Comments</b>
Targeting of children or other vulnerable individuals	N/A	N/A
Risk of physical harm	N/A	N/A

[The assessment can be one of N/A (not applicable), Low, Medium, or High. The comments should explain how the assessment is justified.]

Appendix B – Broad Privacy Risk Assessment:

<b>#</b>	<b>Risk Description/detail</b>	<b>Discussion</b>
23.	Data accuracy and timeliness	[Is data accurately recorded & kept up-to-date?]
24.	Differential treatment of patients/data subjects	[Might certain categories of people be adversely affected, e.g. children, vulnerable adults]
25.	Data Accuracy and identification	[Is the identification of individual reliable? Is there a danger of mis-attribution or incorrect linkage of data?]
26.	Holding / sharing / use of excessive data within [Company] systems	[Might too much data be held or for long? Is there a clear justification for data retention? Not 'just in case']
27.	Data held too long within [Company] systems	[Is there a clear data retention period specified and are there processes to ensure its deletion when no longer needed? Are copies tracked and deleted as well?]
28.	Excessive range of access in terms of users to personal data (consider new users/change of access privileges)	[Do more users have access than strictly necessary? Are user roles clear distinguished and reflected in the access privileges? Is there a clear process for granting and revoking access privileges?]





#	Risk Description/detail	Discussion
29.	Potential for misuse of data, unauthorised access to systems	[What are the likely threats to the data? What countermeasures are or might be applied? Is it possible for access to be granted inappropriately?]
30.	New sharing of data with other organisations, including new or change of suppliers	[Is data being shared from new data providers or with new data users? Are there new suppliers or data processors? What controls will apply?]
31.	Variable and inconsistent adoption / implementation	[How well will this system work end-to-end? How robust is it against partial adoption or system failure?]
32.	Legal compliance, particularly DP transparency requirements and support for data subject rights	[How well does this system meet legal requirements – or appear to meet legal requirements? Does it meet the ‘No surprises’ rule? What would happen if an individual requests data erasure or ceasing processing, etc.]
33.	Medical confidentiality	[Are there any addition sensitivities over confidentiality? Might specific approval (e.g. REC) be required to support this processing?]



## IG Assessment Checklist ESR5 – ANCA-associated vasculitis & environmental risk factors: a case-control study

### Introduction to IG Assessment process

Under the General Data Protection Regulation (GDPR), a Data Protection Impact Assessment (DPIA) is only required where proposed data processing is “likely to result in a high risk to the rights and freedoms of natural persons” (Article 35(1)). However, Article 35(3) explicitly requires one where there is ‘large-scale’ processing of ‘special category’ (e.g. healthcare) data then a DPIA is required.

One other possibility is that the data being processed is already anonymised (see Recital 26) so falls outside GDPR altogether so that no DPIA is actually required.

However, good project management and information governance suggests that there should be a general approach to risk assessment for any project or business enterprise – if only to determine whether a DPIA might be required.

Ideally, one should work from a simple initial Checklist (this document) which identifies possible areas of information risk and compliance requirements to a ‘discussion note’ which explores any issues in more depth and may help identify the necessary mitigation methods and mechanisms to offset most if not all risks. Only if risks are unmitigated or remain ‘high’ would you move to a formal DPIA report.

### The IG Assessment approach

There should be an overview of the proposed project or business change to explain what processing is envisaged as well as the purpose and intended outcome. The ‘purpose’ is important to establish the legal basis for the processing as well as ensuring that any possible mitigations or counter-measures do not undermine the main rationale for the processing.

The next step is to establish what compliance requirements may apply: GDPR, contractual or other regulatory restrictions, consent requirements, or obligations to preserve the data for legal or other reasons (including the benefit of posterity perhaps).

Once the precise range of obligations has been established, then appropriate checks can be made and recorded within the document.

The most obvious of these being GDPR compliance. There must be a ‘High Risk’ assessment (Appendix A) to determine whether the supervisory authority needs to be informed – generally, it is expected that it will not be necessary; if so, then a formal DPIA report will be needed.

Appendix B has a broader Privacy Impact Assessment that may throw up some broader issues.

Initial conclusions as to next steps or particular countermeasures to be considered should be detailed below.



## Project Background/Overview

### Swiss Case-control study

AAV cases recruited among main Swiss hospitals and one rheumatologic centre.

Controls from the Swiss Household Panel will be matched to AAV cases by age, sex and area of residence.

Recruited AAV cases will complete the FORs questionnaire (only questions focusing on occupational history, tobacco smoke exposure, demographic, clinical and socioeconomic status).

Project will be submitted to the St Gallen Kantonsspital legal department. After its approval it will be submitted to BASEC, the online Swissethic portal.

### Comparison of process steps (simplified): [optional]

This allows identification of what processing is new or changed through the project:

Step	Current	Proposed
Project initiation	Project submission to St Gallen legal department	No change
Project initiation	Project submission to Swissethics (national ethic committee)	
Cases recruitment	Cases recruitment from main Swiss hospitals and a rheumatologic centre	
Controls recruitment	Controls recruitment from SHP	
Data collection	Questionnaires sending to cases.	
Data analysis	Data logging to secure software and statistical analysis	
Publication of results	Writing of results in a scientific and medical journal	

### Initial Conclusions

concerning further counter-measures or business viability [possibly tentative]

6. ...
7. ...



Compliance Checks required:

Tick	Requirement	Notes [replace guide text with response]
✓	Does the project involve processing 'personal data' of any sort?	Yes, demographic data, socio-economic status, nationality and clinical data will be retrieved from cases.
✓	Does the project involve processing 'confidential data' of any sort?	Yes, socio-economic status and nationality may be sensitive data. Furthermore, clinical data is confidential.
<b>Data Availability requirements</b>		
✓	Does data need to be held for GCP compliance?	This project may have an impact on the safety and well-being of human subjects. Furthermore, it is intended to be submitted to regulatory authorities. For those reasons data need to be held for GCP compliance.
☐	Does data need to be held to meet 'Open Data' requirements?	No, data does not need to be held to meet "Open Data" requirements. Data collection is intended to answer specific questions and no to be re-used.
✓	Does data need to be held to meet ICMJE requirements or commitments?	Yes, data need to be held to meet ICMJE requirements or commitments as the project will be submitted to scientific and medical journals.



GDPR Compliance Checklist – where ‘personal data’ is processed:

Tick	Requirement	Notes [replace guide text with response]
<b>Article 5: Principles compliance checks</b>		
✓	s) Is processing lawful, fair, and transparent?	The processing of personal data will be easily accessible and easy to understand. Clear and plain language will be used. Data subject will receive information on the identity of controllers and purposes of the processing of personal data.
✓	t) Is the purpose (or purposes) of the processing clearly defined	Personal data will be collected only for specified, explicit and legitimate purpose and will not be further processed for other purposes.
✓	u) adequate, relevant and limited to what is necessary	Data required will be limited to what is necessary. For instance, geographical location of subjects will be provided only by their postal code in order to comply with the data minimization principle.
✓	v) accurate and, where necessary, kept up to date	Cases will be asked to provide updates or corrections of data after completing the questionnaire in order to comply with the accuracy principle.
✓	w) kept and permits identification of data subjects for no longer than is necessary	Data will be kept until scientific and medical journal publish the results. However, according to Swiss ordinance, “the investigator must retain all documents required for the identification and follow-up of participants, and all other original data, for at least ten years after the completion or discontinuation of the clinical trial.”
✓	x) processed securely	RedCap software will be used to process data. It is a highly secure Health Insurance Portability and Accountability Act (HIPAA).



Tick	Requirement	Notes [replace guide text with response]
✓	5) can you demonstrate this compliance?	Data is stored and hosted at St Gallen Kantonsspital. No project data is ever transmitted at any time by REDCap from that institution to another institution or organization.
<b>Articles 13 &amp; 14 compliance</b>		[See detailed Transparency Checklist below]
✓	Did the data come from publicly accessible sources?	No
✓	Are data subjects informed before processing starts for any new purpose if incompatible with original purpose where the controller wants to use data for a different purpose to the purpose for which they currently hold data	Yes, data subjects will be informed of any new purpose if incompatible with original purpose.
✓	Does the Privacy Notice and/or PIL cover this processing?	Yes, the Privacy Notice will cover this processing.
✓	What patient choices are available? Are these explained?	Patients are fully free to participate to the study, to complete the questionnaire and to withdraw to the study. Those choices will be explained.
<b>Articles 6 and 9: legal bases</b>		
<input type="checkbox"/>	What are legal bases under Article 6 Lawfulness of processing	Data subject has given consent to the processing of his or her personal data for one or more specific purposes;
<input type="checkbox"/>	What are legal bases under Article 9 (if 'special category' data)	Data subject has also given consent to the processing of his or her 'special category' data as nationality and data concerning health will be processed.
✓	Are Article 6 legitimate interests explained where relevant?	An LIA Legitimate interest Analysis form will be completed.



Tick	Requirement	Notes [replace guide text with response]
✓	Are details of statutory obligations for Article 6 explained where relevant.	Yes, details of statutory obligations for Article 6 will be explained.
✓	Is this proposed processing compatible with the declared purposes?	Yes, the proposed processing is compatible with the declared purposes.
<b>Article 89(1) research exemption</b>		
<input type="checkbox"/>	If for research, do we meet Art 89(1) data minimisation	Not applicable. Genetic data will not be used in this project.
<b>Articles 15-23: Data Subject Rights</b>		[See detailed table below]
<input type="checkbox"/>	Do we support data subject rights?	We do support data Subject right to be informed, right of access, right to rectification, right to erasure, right to restrict processing, right to data portability and the right to object. As data will be pseudonymised, it would be possible to do so.
✓	There is no use of automated decision making (e.g. profiling)	There is no use of automated decision making.
<b>Articles 24-43: Controller-Processor</b>		
✓	A28 & 29: What measures are there to ensure processors comply?	Data will be processed by one centre so there is no formal Data Processing Agreement. Moreover, Data provided by other centres will be depersonalised at source so personal identifiers will be removed and replaced with a secret ID. Saint Gallen centre will not retain any means of identifying subjects recruited by other hospitals.
<input type="checkbox"/>	A30: Is there an entry for this processing/data held in the register?	



Tick	Requirement	Notes [replace guide text with response]
✓	A32-34: Do we ensure appropriate security, including protection against unauthorised or unlawful processing and against accidental loss, destruction or damage, using appropriate technical or organisational measures?	RedCap software is an appropriate technical measure that ensures appropriate security. Data will be kept in closed and secured areas.
<input type="checkbox"/>	A37-39: Is there a DPO (data protection officer) and have they been or will they be consulted?	CTU at Saint Gallen and Swissethics will be consulted.
<b>Articles 44-50: International transfers</b>		
	What form of data will be transferred to a third country or international organisation	Study will be restricted to Switzerland and no data will be transferred to a third country or international organisation.
<input type="checkbox"/>	Are there safeguards for international transfers?	Not applicable
<b>Article 90: Obligations of secrecy</b>		
<input type="checkbox"/>	Do we meet medical confidentiality requirements?	Only investigators have access to medical records and data will be securely processed.

#### Data Subject Rights:

Note if supported and what process/procedure applies; if not, then describe the legal justification for not supporting this right.

✓	To be informed: about processing, about choices, about rights, about controller	Informed consent will notify data subject about processing, about choices, about rights and about controller.
✓	the right of access to see or receive a printed copy	Yes
✓	the right to rectification – to correct any material errors in the personal data	Yes





✓	the right to erasure – where appropriate, to ask that all personal data is erased	Yes
✓	the right to restrict processing – to ask that some or all processing ceases [see opt-out]	Yes
✓	the right to data portability – this only applies to data provided directly by individual	Yes
✓	the right to object to and not to be subject to automated decision-making, including profiling	There will not be automated decision-making
✓	Right to object to a Data Processing Authority (typically the relevant supervisory authority of each Member State)	Yes
✓	Where consent is the legal basis, the right to withdraw consent	Yes



Detailed Transparency Checklist<sup>26</sup>

Does privacy information provided to data subjects include:

✓	The name and contact details of our organisation	Yes
✓	The name and contact details of our representative (if applicable)	Yes
✓	The contact details of our data protection officer (if applicable)	Not applicable
✓	The purposes of the processing	Yes
✓	The lawful bases for the processing	Yes
✓	The legitimate interests for the processing (if applicable)	Yes
✓	The categories of personal data obtained (if the personal data is not obtained from the individual it relates to)	Yes
✓	The recipients or categories of recipients of the personal data	Yes
☐	The details of transfers of the personal data to any third countries or international organisations (if applicable)	Not applicable
✓	The retention periods for the personal data.	Yes
✓	The rights available to individuals in respect of the processing	Yes
✓	The right to withdraw consent (if applicable)	Yes
✓	The right to lodge a complaint with a supervisory authority	Yes

<sup>26</sup> Taken from UK Information Commissioner's Office template



✓	The source of the personal data (if the personal data is not obtained from the individual it relates to)	Yes
✓	The details of whether individuals are under a statutory or contractual obligation to provide the personal data  (if applicable, and if the personal data is collected from the individual it relates to)	Individual will receive the information that they are under a contractual obligation to provide the personal data.
☐	The details of the existence of automated decision-making, including profiling (if applicable)	Not applicable
✓	We provide individuals with privacy information at the time we collect their personal data from them – or where we obtain personal data from a source other than the individual it relates to, we provide them with privacy information	Yes
✓	within a reasonable of period of obtaining the personal data and no later than one month	Yes
✓	if we plan to communicate with the individual, at the latest, when the first communication takes place	Yes
✓	if we plan to disclose the data to someone else, at the latest, when the data is disclosed	Yes
✓	We provide the information in a way that is:  ✓ concise;  ✓ transparent;  ✓ intelligible;  ✓ easily accessible; and  ✓ uses clear and plain language.	Informed consent will be read by other persons to ask for its understandability
✓	When drafting the information, we:	Yes



	<ul style="list-style-type: none"> <li>✓ undertake an information audit to find out what personal data we hold and what we do with it.</li> <li>✓ put ourselves in the position of the people we're collecting information about.</li> <li>✓ carry out user testing to evaluate how effective our privacy information is</li> </ul>	
<input type="checkbox"/>	<p>When providing our privacy information to individuals, we use a combination of appropriate techniques, such as:</p> <ul style="list-style-type: none"> <li><input type="checkbox"/> a layered approach;</li> <li><input type="checkbox"/> dashboards;</li> <li>✓ just-in-time notices;</li> <li><input type="checkbox"/> icons; and</li> <li><input type="checkbox"/> mobile and smart device functionalities.</li> </ul>	<p>[Note: best practice advice]</p>



### Security & Access Control Checklist

Controls need to be appropriate to level of risk: identified special category data needs more protection against potential misuse than non-personal data.

	Data Security classification (above Official)	<input checked="" type="checkbox"/> - Official-Sensitive <input type="checkbox"/> - Secret <input type="checkbox"/> - Top Secret <input type="checkbox"/> - Public Domain
<input checked="" type="checkbox"/>	Personal Data involved [GDPR]	Demographic data (sex, age, postcode)
<input checked="" type="checkbox"/>	Special Category of personal data involved [GDPR]	Health data, race, ethnicity
<input type="checkbox"/>	Electronic Communications (inc. cookies) [PECR]	No
<input type="checkbox"/>	Credit Card data	No
<input type="checkbox"/>	Legal enforcement [LED2018]	No
<input type="checkbox"/>	Financial data	No
<input type="checkbox"/>	Intellectual Property (detail owner)	No
<input type="checkbox"/>	Commercial in confidence (detail owner)	No
	Data Location (storage or processing)  (include any back-up site(s))	<input type="checkbox"/> - UK <input checked="" type="checkbox"/> - EU/EEA <input type="checkbox"/> - EU White-list <input type="checkbox"/> - USA <input type="checkbox"/> - Other:
<input checked="" type="checkbox"/>	Is data held in secure data centre?	Yes, data is held in St gallen rheumatologic department which is a secure data centre
<input type="checkbox"/>	Is this new supplier, location, or system?	Saint Gallen rheumatologic department is not a new location. However RedCap software and will be a new system.



<input type="checkbox"/>	Is all user access subject to 2-factor authentication?	<input type="checkbox"/> - no control <input type="checkbox"/> - single factor (e.g. just password) <input checked="" type="checkbox"/> - 2-factor (e.g. password & fob) <input type="checkbox"/> - biometric [note: GDPR reqs] <input type="checkbox"/> - Other control:
<input type="checkbox"/>	Are there established JML procedures?	There are not JML procedures yet
<input type="checkbox"/>	Are there checks that passwords are robust and secure enough?	[]
<input type="checkbox"/>	Are all administrator & user accounts routinely monitored?	[Particularly for redundant or little used accounts]
<input type="checkbox"/>	Are systems protected against malware and other attacks?	[provide details of protection software and procedures]

[Need some aspect of CIA/impact-likelihood assessment]

#### Information Asset Register Checklist

<input type="checkbox"/>	Are there new IAs being created?	[provide details]
<input type="checkbox"/>	Are old IAs being retired?	[provide details]
<input type="checkbox"/>	Have IAOs & IACs been consulted?	
<input type="checkbox"/>	Has IAR been updated/amended?	[at least create project task to do so]
<input type="checkbox"/>	Data Retention classification & period	Clinical data must be retained during 10 years in Switzerland
<input type="checkbox"/>	Data retention procedure/functionality in place	



## Appendix A – Supervisory Authority ‘High Risk’ Check

If the DPIA shows ‘high risk’ processing which cannot be mitigated, then the DPIA should be sent to the relevant authority for review before any processing starts. Note that their review may take several weeks to process. A ‘High Risk’ assessment represents a ‘risk to the rights and freedoms of individuals’ – so may extend beyond GDPR consideration, including Human Rights.

GDPR Article 35(3) provides three examples:

- j) a systematic and extensive evaluation of personal aspects relating to natural persons which is based on automated processing, including profiling, and on which decisions are based that produce legal effects concerning the natural person or similarly significantly affect the natural person;
- k) processing on a large scale of special categories of data referred to in Article 9(1), or of personal data relating to criminal convictions and offences referred to in Article 1013; or
- l) a systematic monitoring of a publicly accessible area on a large scale

ICO cites:

- 10. Systematic and extensive profiling with significant effects
- 11. Large scale use of sensitive data [viz. ‘special category’ in GDPR terms]
- 12. Public monitoring

These being the same as (a)-(c) above. They further identify:

- 31. **New technologies:** processing involving the use of new technologies, or the novel application of existing technologies (including AI).
- 32. **Denial of service:** Decisions about an individual’s access to a product, service, opportunity or benefit which is based to any extent on automated decision-making (including profiling) or involves the processing of special category data.
- 33. **Large-scale profiling:** any profiling of individuals on a large scale.
- 34. **Biometrics:** any processing of biometric data.
- 35. **Genetic data:** any processing of genetic data, other than that processed by an individual GP or health professional for the provision of health care direct to the data subject.
- 36. **Data matching:** combining, comparing or matching personal data obtained from multiple sources.
- 37. **Invisible processing:** processing of personal data that has not been obtained direct from the data subject in circumstances where the controller considers that compliance with Article 14 would prove impossible or involve disproportionate effort.
- 38. **Tracking:** processing which involves tracking an individual’s geolocation or behaviour, including but not limited to the online environment.
- 39. **Targeting of children or other vulnerable individuals:** The use of the personal data of children or other vulnerable individuals for marketing purposes, profiling or other automated decision-making, or if you intend to offer online services directly to children.



40. **Risk of physical harm:** Where the processing is of such a nature that a personal data breach could jeopardise the [physical] health or safety of individuals.

**'High Risk' assessment using ICO criteria:**

Criterion:	Assessment	Comments
New technologies	N/A	
Denial of service	Low	
Large-scale profiling	N/A	
Biometrics	N/A	
Genetic data	N/A	
Data matching	N/A	
Invisible processing	Low	
Tracking	Medium	Postcode location
Targeting of children or other vulnerable individuals	N/A	





Criterion:	Assessment	Comments
Risk of physical harm	Low	

[The assessment can be one of N/A (not applicable), Low, Medium, or High. The comments should explain how the assessment is justified.]

Appendix B – Broad Privacy Risk Assessment:

#	Risk Description/detail	Discussion
34.	Data accuracy and timeliness	Yes, data is accurately recorded & kept up-to-date
35.	Differential treatment of patients/data subjects	There is no certain categories of people be adversely affected.
36.	Data Accuracy and identification	Identification of individual will be reliable. There will be no danger of misattribution or incorrect linkage of data
37.	Holding / sharing / use of excessive data within [Company] systems	Data will be retained to comply with Swiss ordinance.
38.	Data held too long within [Company] systems	Data retention period is 10 years according to Swiss ordinance. There are no processes to ensure its deletion when no longer needed, yet. Copies will be tracked and deleted as well.
39.	Excessive range of access in terms of users to personal data (consider new users/change of access privileges)	Number of users with access to data will be restricted to the strictly necessary. User roles will be clearly distinguished and reflected in the access privileges. There will be a clear process for granting and revoking access privileges.
40.	Potential for misuse of data, unauthorised access to systems	Threats to the data are scarce.
41.	New sharing of data with other organisations, including new or change of suppliers	Data will not be shared from new data providers or with new data users. There are no new suppliers or data processors.



#	Risk Description/detail	Discussion
42.	Variable and inconsistent adoption / implementation	[How well will this system work end-to-end? How robust is it against partial adoption or system failure?]
43.	Legal compliance, particularly DP transparency requirements and support for data subject rights	[How well does this system meet legal requirements – or appear to meet legal requirements? Does it meet the ‘No surprises’ rule? What would happen if an individual requests data erasure or ceasing processing, etc.]
44.	Medical confidentiality	Nationality and geographical location may be sensitive data. Specific approval may be required to support the processing.



## IG Assessment Checklist ESR6 – Atmospheric monitoring and time series analysis of climate and pollution impact on vasculitis onset

### Introduction to IG Assessment process

Under the General Data Protection Regulation (GDPR), a Data Protection Impact Assessment (DPIA) is only required where proposed data processing is “likely to result in a high risk to the rights and freedoms of natural persons” (Article 35(1)). However, Article 35(3) explicitly requires one where there is ‘large-scale’ processing of ‘special category’ (e.g. healthcare) data then a DPIA is required.

One other possibility is that the data being processed is already anonymised (see Recital 26) so falls outside GDPR altogether so that no DPIA is actually required.

However, good project management and information governance suggests that there should be a general approach to risk assessment for any project or business enterprise – if only to determine whether a DPIA might be required.

Ideally, one should work from a simple initial Checklist (this document) which identifies possible areas of information risk and compliance requirements to a ‘discussion note’ which explores any issues in more depth and may help identify the necessary mitigation methods and mechanisms to offset most if not all risks. Only if risks are unmitigated or remain ‘high’ would you move to a formal DPIA report.

### The IG Assessment approach

There should be an overview of the proposed project or business change to explain what processing is envisaged as well as the purpose and intended outcome. The ‘purpose’ is important to establish the legal basis for the processing as well as ensuring that any possible mitigations or counter-measures do not undermine the main rationale for the processing.

The next step is to establish what compliance requirements may apply: GDPR, contractual or other regulatory restrictions, consent requirements, or obligations to preserve the data for legal or other reasons (including the benefit of posterity perhaps).

Once the precise range of obligations has been established, then appropriate checks can be made and recorded within the document.

The most obvious of these being GDPR compliance. There must be a ‘High Risk’ assessment (Appendix A) to determine whether the supervisory authority needs to be informed – generally, it is expected that it will not be necessary; if so, then a formal DPIA report will be needed.

Appendix B has a broader Privacy Impact Assessment that may throw up some broader issues.

Initial conclusions as to next steps or particular countermeasures to be considered should be detailed below.



Project Background/Overview

[Explain business background, including any existing processes and procedures; outline the project including stages, deliverables, and timelines]

ANCA vasculitis is a rare disease that has been difficult to study and explain until now. Particularly, we have limited knowledge about its causes and factors that are associated with it. What we know so far, is that a person’s environment could explain much of the risk for acquiring the disease. To date, however, this knowledge is still very limited. This study therefore aims to address this gap in knowledge by studying the environmental causes of ANCA vasculitis. In achieving this, clinical data from patient registries will be linked with environmental datasets from weather and air quality monitoring registries, and self generated aerobiome measurements. With this linkage, we will be able to study the link between the environment and the risk of acquiring the disease. We will also look at the geographical and seasonal pattern of the disease onset, with the hope of identifying people at particular risk, as well as conditions that may explain the risk in disease and onset.

**Rationale:** Environmental exposures are likely to play a role in the onset of systemic vasculitis, however the precise factors have yet to be delineated.

**Objectives:** To identify and quantify relationships between multiple environmental entities and the spatial circumstances of systemic vasculitis disease onset.

**Design** – Multiple study designs will be used. These include - case crossover design, longitudinal and cross-sectional time series study designs

**Analysis Method** – Regression and time-series analyses will be conducted to determine important environmental predictors of systemic vasculitis onset.

**Expected Results:** The identification of several candidate environmental factors implicated in the onset of AAV

Comparison of process steps (simplified): [optional]

This allows identification of what processing is new or changed through the project:

Step	Current	Proposed
Project initiation, including any ISAC approval, up to Task Order from client		No change



Initial Conclusions

concerning further counter-measures or business viability [possibly tentative]

1. ...
2. ...

Compliance Checks required:

Tick	Requirement	Notes [replace guide text with response]
x	Does the project involve processing 'personal data' of any sort?	Yes, raw pre-processed data will be patient-specific data with information about comorbidities, postcode of residence, age, gender, and several variables that could be used to identify the patient.
x	Does the project involve processing 'confidential data' of any sort?	Medical confidentiality (patient's health records). Personal data such as location.
<b>Data Availability requirements</b>		
x	Does data need to be held for GCP compliance?	Yes, good clinical practice compliance is necessary for any kind of research using clinical data.
x	Does data need to be held to meet 'Open Data' requirements?	The results of working on the data will be publicly disseminated in scientific and medical journals.
x	Does data need to be held to meet ICMJE requirements or commitments?	Yes, the aim is to publish in biomedical journals.



GDPR Compliance Checklist – where ‘personal data’ is processed:

Tick	Requirement	Notes [replace guide text with response]
<b>Article 5: Principles compliance checks</b>		
x	a) Is processing lawful, fair, and transparent?	Yes.
x	b) Is the purpose (or purposes) of the processing clearly defined	Yes, purpose is defined in each proposal to each registry. Mainly: Research into the pathogenesis, diagnosis and treatment of vasculitis
x	c) adequate, relevant and limited to what is necessary	Yes, only required variables are queried and used, leaving unnecessary personal data outside.
x	d) accurate and, where necessary, kept up to date	This falls outside of my control, but the registries queried should ensure so.
?	e) kept and permits identification of data subjects for no longer than is necessary	Registry asked for unlimited access in time since many scientific purposes might arise from it. Data sharing with HELICAL entities will be limited by the duration of the process.
?	f) processed securely	Yes.
?	2) can you demonstrate this compliance?	Compliance checks are run by the local DPO.
<b>Articles 13 &amp; 14 compliance</b>		[See detailed Transparency Checklist below]
x	Did the data come from publicly accessible sources?	All the non-clinical data comes from either public databases or self-generated experiments. Clinical data is accessible through registries by going over the required process in each case.
<input type="checkbox"/>	Are data subjects informed before processing starts for any new purpose if incompatible with original purpose where the controller wants to use data for a different purpose to the purpose for which they currently hold data	From the RKD DPIA, which applies here too: This will be outlined in the information leaflet and consent form. It will not be possible to inform participants of processing that is carried out on individual results (this is also addressed in the consent form) as it is not known when such processing will occur and researchers will only receive coded or completed de-identified samples and data only and may analyse pooled data. Information on publications using the Registry and Biobank will be disseminated to participants via Tara open access



		publications repository <a href="http://www.tara.tcd.ie/">http://www.tara.tcd.ie/</a> .
<input type="checkbox"/>	Does the Privacy Notice and/or PIL cover this processing?	
<input type="checkbox"/>	What patient choices are available? Are these explained?	[see also Data Subject Rights below]
<b>Articles 6 and 9: legal bases</b>		
<input type="checkbox"/>	What are legal bases under Article 6	<i>Article 6(1)(e)</i> - Public Interest and Article
<input type="checkbox"/>	What are legal bases under Article 9 (if 'special category' data)	<i>Article 9(2)(j)</i> Scientific Research.
<input type="checkbox"/>	Are Article 6 legitimate interests explained where relevant?	Yes, this is done in the UKIVAS data request.
<input type="checkbox"/>	Are details of statutory obligations for Article 6 explained where relevant.	
<input type="checkbox"/>	Is this proposed processing compatible with the declared purposes?	Yes.
<b>Article 89(1) research exemption</b>		
x	If for research, do we meet Art 89(1) data minimisation	Yes (5c)
<b>Articles 15-23: Data Subject Rights</b>		
?	Do we support data subject rights?	[See detailed table below]
?	Do we support data subject rights?	Data is aggregated, so individual removal of patient data would be almost impossible and would deem the research non-viable.
?	There is no use of automated decision making (e.g. profiling)	There will be no decisions taken on the individual patient level, but automated clustering could be done (which, in some way, could be considered 'profiling').  However, as the data in this study is coded and only intended for research purposes, outcomes will not be used for monitoring individuals or making automated decisions that will affect individuals - this is not considered to constitute automated decision making or profiling.
<b>Articles 24-43: Controller-Processor</b>		
?	A28 & 29: What measures are there to ensure processors comply?	Yes.
x	A30: Is there an entry for this processing/data held in the register?	Yes.
x	A32-34: Do we ensure appropriate security, including protection against unauthorised or unlawful processing and	Yes.



	against accidental loss, destruction or damage, using appropriate technical or organisational measures?	
x	A37-39: Is there a DPO and have they been or will they be consulted?	There is an institutional DPO, and they are consulted.
<b>Articles 44-50: International transfers</b>		
	What form of data will be transferred to a third country or international organisation	All of the data is international, as the research is not focused on a single country. Both clinical and environmental data spans several countries.
x	Are there safeguards for international transfers?	Yes.
<b>Article 90: Obligations of secrecy</b>		
x	Do we meet medical confidentiality requirements?	Yes, all data access is managed by medical institutions which checked the appropriate requirements.

#### Data Subject Rights:

Note if supported and what process/procedure applies; if not, then describe the legal justification for not supporting this right.

<input type="checkbox"/>	To be informed: about processing, about choices, about rights, about controller	
<input type="checkbox"/>	the right of access to see or receive a printed copy	
<input type="checkbox"/>	the right to rectification – to correct any material errors in the personal data	
<input type="checkbox"/>	the right to erasure – where appropriate, to ask that all personal data is erased	As I mentioned before, withdrawing data retroactively would make the research non-viable.
<input type="checkbox"/>	the right to restrict processing – to ask that some or all processing ceases [see opt-out]	
<input type="checkbox"/>	the right to data portability – this only applies to data provided directly by individual	
<input type="checkbox"/>	the right to object to and not to be subject to automated decision-making, including profiling	
<input type="checkbox"/>	Right to object to a Data Processing Authority (typically the relevant supervisory authority of each Member State)	
<input type="checkbox"/>	Where consent is the legal basis, the right to withdraw consent	SOP for removing data from the database and returning to the individual (if desired). PIL includes a statement that data already used in research cannot be removed.





### Detailed Transparency Checklist<sup>27</sup>

Does privacy information provided to data subjects include:

x	The name and contact details of our organisation	<b>ISGlobal (Institut de Salut Global de Barcelona)</b> Rosselló, 132, 7è 08036 Barcelona Phone: +34 93 227 1806 info@isglobal.org
x	The name and contact details of our representative (if applicable)	
x	The contact details of our data protection officer (if applicable)	Joana Porcel joana.porcel@isglobal.org
x	The purposes of the processing	
x	The lawful bases for the processing	[Art6 for 'personal data' & Art9 for 'special category']
x	The legitimate interests for the processing (if applicable)	
x	The categories of personal data obtained (if the personal data is not obtained from the individual it relates to)	[for Art14]
x	The recipients or categories of recipients of the personal data	
?	The details of transfers of the personal data to any third countries or international organisations (if applicable)	
?	The retention periods for the personal data.	
?	The rights available to individuals in respect of the processing	
?	The right to withdraw consent (if applicable)	
<input type="checkbox"/>	The right to lodge a complaint with a supervisory authority	
<input type="checkbox"/>	The source of the personal data (if the personal data is not obtained from the individual it relates to)	[For Art14]
<input type="checkbox"/>	The details of whether individuals are under a statutory or contractual obligation to provide the personal data (if applicable, and if the personal data is collected from the individual it relates to)	

<sup>27</sup> Taken from UK Information Commissioner's Office template



<input type="checkbox"/>	<p>The details of the existence of automated decision-making, including profiling (if applicable)</p>	
<input type="checkbox"/>	<p>We provide individuals with privacy information at the time we collect their personal data from them – or where we obtain personal data from a source other than the individual it relates to, we provide them with privacy information</p>	
<input type="checkbox"/>	<p>within a reasonable of period of obtaining the personal data and no later than one month</p>	
<input type="checkbox"/>	<p>if we plan to communicate with the individual, at the latest, when the first communication takes place</p>	
<input type="checkbox"/>	<p>if we plan to disclose the data to someone else, at the latest, when the data is disclosed</p>	
<input type="checkbox"/>	<p>We provide the information in a way that is:</p> <ul style="list-style-type: none"> <li><input type="checkbox"/> concise;</li> <li><input type="checkbox"/> transparent;</li> <li><input type="checkbox"/> intelligible;</li> <li><input type="checkbox"/> easily accessible; and</li> <li><input type="checkbox"/> uses clear and plain language.</li> </ul>	
<input type="checkbox"/>	<p>When drafting the information, we:</p> <ul style="list-style-type: none"> <li><input type="checkbox"/> undertake an information audit to find out what personal data we hold and what we do with it.</li> <li><input type="checkbox"/> put ourselves in the position of the people we’re collecting information about.</li> <li><input type="checkbox"/> carry out user testing to evaluate how effective our privacy information is</li> </ul>	<p>[Note: best practice advice]</p>
<input type="checkbox"/>	<p>When providing our privacy information to individuals, we use a combination of appropriate techniques, such as:</p> <ul style="list-style-type: none"> <li><input type="checkbox"/> a layered approach;</li> <li><input type="checkbox"/> dashboards;</li> <li><input type="checkbox"/> just-in-time notices;</li> <li><input type="checkbox"/> icons; and</li> <li><input type="checkbox"/> mobile and smart device functionalities.</li> </ul>	<p>[Note: best practice advice]</p>



## Security & Access Control Checklist

Controls need to be appropriate to level of risk: identified special category data needs more protection against potential misuse than non-personal data.

	Data Security classification (above Official)	<input type="checkbox"/> - Official-Sensitive <input type="checkbox"/> - Secret <input type="checkbox"/> - Top Secret <input type="checkbox"/> - Public Domain
x	Personal Data involved [GDPR]	
x	Special Category of personal data involved [GDPR]	
<input type="checkbox"/>	Electronic Communications (inc. cookies) [PECR]	
<input type="checkbox"/>	Credit Card data	
<input type="checkbox"/>	Legal enforcement [LED2018]	
<input type="checkbox"/>	Financial data	
<input type="checkbox"/>	Intellectual Property (detail owner)	
<input type="checkbox"/>	Commercial in confidence (detail owner)	
	Data Location (storage or processing) (include any back-up site(s))	EU
<input type="checkbox"/>	Is data held in secure data centre?	[detail centre and what certification supports assertion]
<input type="checkbox"/>	Is this new supplier, location, or system?	[If so, need specific IS check; also need formal contract]
x	Is all user access subject to 2-factor authentication?	<input type="checkbox"/> - no control <input type="checkbox"/> - single factor (e.g. just password) <input checked="" type="checkbox"/> - 2-factor (e.g. password & fob) <input type="checkbox"/> - biometric [note: GDPR reqs] <input type="checkbox"/> - Other control:
x	Are there established JML procedures?	<p>Yes. Provided in attached DPO rules of IT Services.</p> <p>Procedure for leavers: If a User ends their relationship with the Institution or changes jobs, they must leave all IT applications, files, information, data and electronic documents they have used in their professional activity, without prejudice. Once the relationship with the Institution has finished, they shall no longer have access to the IT equipment and information incorporated therein, having to return those they have in their possession. They shall continue to be bound to maintain the strictest confidentiality and discretion, not only</p>



		of the information and documents, but also of IT applications, analysis and keys that they have known during or because of their relationship with the Foundation.
x	Are there checks that passwords are robust and secure enough?	Yes. The credentials (username and password) shall be given to the User by HR on a paper document the first time (with a single use password) once their relationship with the Institution has been formalised. When the User first accesses their computer, they must change the password to one of their choosing. The password shall be at least 8 characters long. The characters must be a combination of letters and numbers or special characters. It is obligatory to change the password every 180 days and it can also be voluntarily changed through the ISM.
x	Are all administrator & user accounts routinely monitored?	Yes, by local IT.
x	Are systems protected against malware and other attacks?	Yes.

[Need some aspect of CIA/impact-likelihood assessment]

#### Information Asset Register Checklist

<input type="checkbox"/>	Are there new IAs being created?	[provide details]
<input type="checkbox"/>	Are old IAs being retired?	[provide details]
<input type="checkbox"/>	Have IAOs & IACs been consulted?	
<input type="checkbox"/>	Has IAR been updated/amended?	[at least create project task to do so]
<input type="checkbox"/>	Data Retention classification & period	
<input type="checkbox"/>	Data retention procedure/functionality in place	



## Appendix A – Supervisory Authority ‘High Risk’ Check

If the DPIA shows ‘high risk’ processing which cannot be mitigated, then the DPIA should be sent to the relevant authority for review before any processing starts. Note that their review may take several weeks to process. A ‘High Risk’ assessment represents a ‘risk to the rights and freedoms of individuals’ – so may extend beyond GDPR consideration, including Human Rights.

GDPR Article 35(3) provides three examples:

- a) a systematic and extensive evaluation of personal aspects relating to natural persons which is based on automated processing, including profiling, and on which decisions are based that produce legal effects concerning the natural person or similarly significantly affect the natural person;
- b) processing on a large scale of special categories of data referred to in Article 9(1), or of personal data relating to criminal convictions and offences referred to in Article 1013; or
- c) a systematic monitoring of a publicly accessible area on a large scale

ICO cites:

1. Systematic and extensive profiling with significant effects
2. Large scale use of sensitive data [viz. ‘special category’ in GDPR terms]
3. Public monitoring

These being the same as (a)-(c) above. They further identify:

1. **New technologies:** processing involving the use of new technologies, or the novel application of existing technologies (including AI).
2. **Denial of service:** Decisions about an individual’s access to a product, service, opportunity or benefit which is based to any extent on automated decision-making (including profiling) or involves the processing of special category data.
3. **Large-scale profiling:** any profiling of individuals on a large scale.
4. **Biometrics:** any processing of biometric data.
5. **Genetic data:** any processing of genetic data, other than that processed by an individual GP or health professional for the provision of health care direct to the data subject.
6. **Data matching:** combining, comparing or matching personal data obtained from multiple sources.
7. **Invisible processing:** processing of personal data that has not been obtained direct from the data subject in circumstances where the controller considers that compliance with Article 14 would prove impossible or involve disproportionate effort.
8. **Tracking:** processing which involves tracking an individual’s geolocation or behaviour, including but not limited to the online environment.
9. **Targeting of children or other vulnerable individuals:** The use of the personal data of children or other vulnerable individuals for marketing purposes, profiling or other automated decision-making, or if you intend to offer online services directly to children.



10. **Risk of physical harm:** Where the processing is of such a nature that a personal data breach could jeopardise the [physical] health or safety of individuals.

**'High Risk' assessment using ICO criteria:**

Criterion:	Assessment	Comments
New technologies	Low	There will be processing of data both using new technologies and with novel usage of existing technology.
Denial of service	N/A	
Large-scale profiling	Low	Patient data profiling/clustering might be done to run the analysis and compare outcomes in a most faithful as possible manner.
Biometrics	Low	Biometric data is part of the data the registry uses.
Genetic data	N/A	No genetic data will be used.
Data matching	Low	Data matching is necessary and a core need for the project, as external environmental data needs to be matched to healthcare patient data. The spatial resolution of the research will be done only up to the level required to get relevant results.
Invisible processing	Low	The data subjects already have the information since the moment they are recruited to be part of the registry, so they are aware of any data not coming directly from them, if existing.
Tracking	Low	Geolocation is necessary in order to be able to perform the data matching (since the 'merge' is done based on the spatial location).
Targeting of children or other vulnerable individuals	Low	Kawasaki Disease affects mainly children from 6 months to 5 years old. Their data will not be used for any non-medical purposes, though.



Risk of physical harm	N/A	The scope of the project makes risk of possible harm negligible beyond a reasonable doubt.
-----------------------	-----	--

[The assessment can be one of N/A (not applicable), Low, Medium, or High. The comments should explain how the assessment is justified.]

Appendix B – Broad Privacy Risk Assessment:

#	Risk Description/detail	Discussion
1.	Data accuracy and timeliness	Hopefully, yes. However, this is one of the challenges, actual date of disease onset might be estimated rather than measured, thus providing a certain degree of uncertainty. The actual prodrome
2.	Differential treatment of patients/data subjects	Certain demographics might be found to be more likely to suffer adverse consequences. However, this would lead to better diagnosis and treatment for them.
3.	Data Accuracy and identification	Proper linkage of environmental data to clinical data is one of the difficult points in our research. There is always a degree of aggregation or approximation needed to calculate environmental exposures.
4.	Holding / sharing / use of excessive data within i~HD systems	Duration of the data in the biobanks/ is clearly defined beforehand.
5.	Data held too long within i~HD systems	As 5. Duration of the data in the biobanks/ is clearly defined beforehand.
6.	Excessive range of access in terms of users to personal data (consider new users/change of access privileges)	Users are only given access to the data on a case per case basis, so no excessive range of access should be happening.
7.	Potential for misuse of data, unauthorised access to systems	It is possible in case of an identity theft, but otherwise not possible, and systems are in place in order to avoid so.
8.	New sharing of data with other organisations, including new or change of suppliers	I believe this is not applicable to our case.



9.	Variable and inconsistent adoption / implementation	I believe the concept of adoption/implementation is not completely relevant to our case. The study tries to pinpoint specific environmental candidates of exacerbated autoimmune response by investigating relative changes of <i>population</i> -level incidence by time. Results might vary from very clear, to non-conclusive, but this would not pose a Privacy-risk in any case.
10.	Legal compliance, particularly DP transparency requirements and support for data subject rights	Legal compliance has been checked by the UKIVAS registry.
11.	Medical confidentiality	Medical confidentiality approvals have already been done.





## IG Assessment Checklist ESR7 – Identification of functionally relevant genetic variants associated with giant-cell arteritis (GCA)

### Introduction to IG Assessment process

Under the General Data Protection Regulation (GDPR), a Data Protection Impact Assessment (DPIA) is only required where proposed data processing is “likely to result in a high risk to the rights and freedoms of natural persons” (Article 35(1)). However, Article 35(3) explicitly requires one where there is ‘large-scale’ processing of ‘special category’ (e.g. healthcare) data then a DPIA is required.

One other possibility is that the data being processed is already anonymised (see Recital 26) so falls outside GDPR altogether so that no DPIA is actually required.

However, good project management and information governance suggests that there should be a general approach to risk assessment for any project or business enterprise – if only to determine whether a DPIA might be required.

Ideally, one should work from a simple initial Checklist (this document) which identifies possible areas of information risk and compliance requirements to a ‘discussion note’ which explores any issues in more depth and may help identify the necessary mitigation methods and mechanisms to offset most if not all risks. Only if risks are unmitigated or remain ‘high’ would you move to a formal DPIA report.

### The IG Assessment approach

There should be an overview of the proposed project or business change to explain what processing is envisaged as well as the purpose and intended outcome. The ‘purpose’ is important to establish the legal basis for the processing as well as ensuring that any possible mitigations or counter-measures do not undermine the main rationale for the processing.

The next step is to establish what compliance requirements may apply: GDPR, contractual or other regulatory restrictions, consent requirements, or obligations to preserve the data for legal or other reasons (including the benefit of posterity perhaps).

Once the precise range of obligations has been established, then appropriate checks can be made and recorded within the document.

The most obvious of these being GDPR compliance. There must be a ‘High Risk’ assessment (Appendix A) to determine whether the supervisory authority needs to be informed – generally, it is expected that it will not be necessary; if so, then a formal DPIA report will be needed.

Appendix B has a broader Privacy Impact Assessment that may throw up some broader issues.

Initial conclusions as to next steps or particular countermeasures to be considered should be detailed below.



### Project Background/Overview

Giant cell arteritis (GCA) is a granulomatous vasculitis that affects large and medium-sized blood vessel with a predisposition for the aorta, branches of the ophthalmic artery, and extra-cranial branches of the carotid artery. This pathology occurs mainly in Caucasian people over 50 years of age. Among its most relevant clinical manifestations are visual loss, limb anoxia and stroke [1, 2].

The etiopathogenesis of this pathology is complex and involved, both in its onset and in its progression, an undetermined number of genetic, epigenetic and environmental factors that lead to alterations in the immune regulation mechanisms [3]. So far, it has been clear that the genetic component of the human being plays an important role in the susceptibility to developing this disease [4]. Large-scale genetic studies (ImmunoChip array and Genome-wide association studies - GWAS) and candidate genes, analyzing single nucleotide polymorphisms (SNPs), have contributed to increasing the number of loci associated with this disease, among which are found HLA, PTPN22, IL17A, IL12B, PLG and P4HA2 [5, 6]. However, unlike other vasculitis and immune-mediated diseases, the genetic component of this disease is still largely unknown.

Therefore, the investigation of new strategies, such as DNA methylation and gene expression will allow us to identify and understand the molecular basis of this disease. In addition, considering that the integration of -omics data has proven to be effective in yielding insight into our understanding complex diseases.

The aim of these project is to carry out a Methylome and transcriptome studies, as well as an integrative analysis of these data of CD14+ monocytes and CD4+ lymphocytes two cell groups crucial in the systemic and local inflammatory processes of this disease. We will obtain the DNA and RNA from these two cell types from a large cohort of controls or healthy individuals and patients affected with GCA and we will perform an epigenome- and transcriptome-wide association study. DNA methylation EPIC array and RNA-seq data will be subsequently integrated to identify correlation between methylation and gene expression levels.

In overall, this project will allow us to provide evidence of the genes and pathways that contribute to the pathogenic role of these two cell types in GCA, as well as the molecular response to CG treatment and the potential translation of these findings to clinical practice.

1. Salvarani C, et al. Polymyalgia rheumatica and giant-cell arteritis. *Lancet* (2008);372(9634):234-45.
2. Jennette JC, et al. 2012 revised International Chapel Hill Consensus Conference Nomenclature of Vasculitides. *Arthritis Rheumatology* (2013); 65:1-11
3. Samson M, et al. Recent advances in our understanding of giant cell arteritis pathogenesis. *Autoimmunity Reviews* (2017); 16:833–844.
4. Carmona F.D., Martín J., González-Gay M.A. (2019) Giant Cell Arteritis. In: Martín J., Carmona F. (eds) *Genetics of Rare Autoimmune Diseases. Rare Diseases of the Immune System*. Springer, Cham.
5. Carmona FD, et al. A large-scale genetic analysis reveals a strong contribution of the HLA class II region to giant cell arteritis susceptibility. *The American Journal of Human Genetics* (2015); 96(4):565-80.
6. Carmona FD, et al. A Genome-wide Association Study Identifies Risk Alleles in Plasminogen and P4HA2 Associated with Giant Cell Arteritis. *The American Journal of Human Genetics* (2017); 100(1):64- 4.



Comparison of process steps (simplified):

This allows identification of what processing is new or changed through the project:

Step	Current	Proposed
Project initiation, including any ISAC approval, up to Task Order from client	Our project was evaluated and approved by the Ethics Committee of the Autonomous Community of Andalusia	No change
Project evaluation by a scientific and academic committee	Our project was evaluated and approved by a scientific committee of the PhD program of the University of Granada	No change
Collection of samples and clinical data	With the different academic, clinical and scientific collaborations, we are working on obtaining the samples and clinical information necessary to meet the objectives set out in our project. All the samples that will be obtained will be from all the individuals who agree to participate in our study by signing the voluntary informed consent	No change
Processing of samples and obtaining data from the methylome and transcriptome study	From all the samples, DNA and RNA will be obtained; these will be analyzed using different molecular techniques to obtain all the relevant biological information of our project	No change
Data analysis and interpretation of our data	All the data we obtain will be protected on the servers of the Institute of Parasitology and Biomedicine "López-Neyra" (IPBLN). These data will be analyzed with various bioinformatic tools and will be interpreted to respond to all our hypotheses and thus generate relevant and innovative knowledge	No change



Step	Current	Proposed
Publication and dissemination of results	All the results will be published in high impact scientific journals. In addition, they will be presented at events and scientific meetings. On the other hand, these results are part of the degree thesis of the doctoral program in biomedicine at the University of Granada "Study of the molecular causes of giant cell arteritis through a systemic approach"	No change

#### Initial Conclusions

8. Provide evidence of the genes and pathways that contribute to the pathogenic role of CD13+ monocytes and CD4+ T cell in GCA
9. Development biomarkers to new, more effective and safer therapies to control this disease.

#### Compliance Checks required:

Tick	Requirement	Notes
■	Does the project involve processing 'personal data' of any sort?	Our project requires general demographic description of the patients (ethnicity, age, sex)
■	Does the project involve processing 'confidential data' of any sort?	Our project requires relevant clinical information (comorbidities and treatments)
<b>Data Availability requirements</b>		
■	Does data need to be held for GCP compliance?	In order to ensure that the data and reported results are reliable and accurate and to ensure that the rights, integrity and confidentiality of the individuals who participated in the study are respected and protected



Tick	Requirement	Notes
■	Does data need to be held to meet 'Open Data' requirements?	Some data from our project will be used by other ESRs for the execution and evaluation of the hypotheses of their projects
■	Does data need to be held to meet ICMJE requirements or commitments?	It is necessary to keep the data for the respective publications; however, the data protection guidelines must be followed.



GDPR Compliance Checklist – where ‘personal data’ is processed:

Tick	Requirement	Notes
<b>Article 5: Principles compliance checks</b>		
■	y) Is processing lawful, fair, and transparent?	The procedure will follow were in accordance with the ethical committee and written informed consent will obtain from all individuals where individuals have been informed what their personal data will be used.
■	z) Is the purpose (or purposes) of the processing clearly defined	Yes, all procedures are clearly defined. Personal data collected for one purpose cannot be used for a new and incompatible purpose. However, additional measures can be taken by obtaining the consent of the affected persons or by anonymizing the data.
■	aa) adequate, relevant and limited to what is necessary	Only personal data that is actually needed to achieve our goals will be processed.
■	bb) accurate and, where necessary, kept up to date	All reasonable measures will be taken to ensure that personal data is accurate. Data are collected from the medical records of the participating individuals.
■	cc) kept and permits identification of data subjects for no longer than is necessary	The personal data will be kept in a format that allows the identification of the interested parties during the execution of the project. Personal data may be stored for longer periods subject to the implementation of appropriate safeguards.
■	dd) processed securely	Personal data will be guaranteed to be kept safe, both against external threats (malicious hackers) and internal threats (poorly trained employees).



Tick	Requirement	Notes
■	6) Can you demonstrate this compliance?	All of the above is described in the information sheet and informed consent given to individuals wishing to participate in our study.
<b>Articles 13 &amp; 14 compliance</b>		
■	Did the data come from publicly accessible sources?	Patient data are from the medical records of individuals who agree to participate in our study by signing the informed consent form.
■	Are data subjects informed before processing starts for any new purpose if incompatible with original purpose where the controller wants to use data for a different purpose to the purpose for which they currently hold data	In our informed consent it is contextualized that these data will be used for this study and subsequent studies of the disease.
■	Do the Privacy Notice and/or PIL cover this processing?	In our informed consent the confidentiality of the data is informed and the national laws that support this are written.
■	What patient choices are available? Are these explained?	All procedures of our project are explained in detail in the patient information sheet and in the informed consent form.
<b>Articles 6 and 9: legal bases</b>		
■	What are legal bases under Article 6	- Consent to the processing of his or her personal data for one or more specific purposes
■	What are legal bases under Article 9 (if 'special category' data)	- Processing is necessary for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes in accordance with Article 89(1) based on Union or Member State law which shall be proportionate to the aim pursued, respect the essence of the right to data protection and provide for suitable and specific measures to safeguard the fundamental rights and the interests of the data subject.



Tick	Requirement	Notes
■	Are Article 6 legitimate interests explained where relevant?	
■	Are details of statutory obligations for Article 6 explained where relevant?	Organic Law 3/2018, of 5 December, on the Protection of Personal Data and the guarantee of digital rights (Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales)
■	Is this proposed processing compatible with the declared purposes?	
<b>Article 89(1) research exemption</b>		
■	If for research, do we meet Art 89(1) data minimisation	All the samples that are collected in our project are treated confidentially and are assigned a unique and consecutive alphanumeric code
<b>Articles 15-23: Data Subject Rights</b>		
■	Do we support data subject rights?	Although our samples are pseudo-/ anonymized, our informed consent contextualizes that an individual who agrees to enter the study can be informed of the data we obtain. In addition, the patient can withdraw from the study at any time, without having to give explanations and without affecting their medical care.
■	There is no use of automated decision making (e.g. profiling)	





Tick	Requirement	Notes
<b>Articles 24-43: Controller-Processor</b>		
■	A28 & 29: What measures are there to ensure processors comply?	Processing by a processor shall be governed by a contract or other legal act under Union or Member State law, that is binding on the processor with regard to the controller and that sets out the subject-matter and duration of the processing, the nature and purpose of the processing, the type of personal data and categories of data subjects and the obligations and rights of the controller.
■	A30: Is there an entry for this processing/data held in the register?	In our laboratory, the bioinformatics technicians are in charge of keeping the registration, storage, protection and availability of the data.
■	A32-34: Do we ensure appropriate security, including protection against unauthorised or unlawful processing and against accidental loss, destruction or damage, using appropriate technical or organisational measures?	The confidentiality of the data is essential to have the security of all our data, only and under strict agreements the information is shared with our collaborating centres and the optimal decisions have been made for the adequate treatment and storage of the data.
■	A37-39: Is there a DPO and have they been or will they be consulted?	Our laboratory belongs to the CSIC (Spanish National Research Council), which has the data protection office and is in charge of providing the respective knowledge and training for data protection.
<b>Articles 44-50: International transfers</b>		
■	What form of data will be transferred to a third country or international organisation	The data that is shared in order to cooperate and collaborate with other research centres are genetic, transcriptomic and epigenetic information of all the individuals participating in our study. As previously mentioned, all of our samples are handled under an alphanumeric code and are kept under the principle of confidentiality.



Tick	Requirement	Notes
■	Are there safeguards for international transfers?	For the transfer of information we use anonymization and approved contractual clauses (Data Transfer Agreement (DTA) for Personal Data)
<b>Article 90: Obligations of secrecy</b>		
■	Do we meet medical confidentiality requirements?	All data obtained maintains the confidentiality and security of individuals. In addition, clinical information is managed according to the ethical committees of each centre from which the information comes.

#### Data Subject Rights:

Note if supported and what process/procedure applies; if not, then describe the legal justification for not supporting this right.

■	To be informed: about processing, about choices, about rights, about controller	This is explained in the patient information sheet and the informed consent.
■	the right of access to see or receive a printed copy	This is explained in the patient information sheet and the informed consent.
■	the right to rectification – to correct any material errors in the personal data	This is explained in the patient information sheet and the informed consent.
■	the right to erasure – where appropriate, to ask that all personal data is erased	This is explained in the patient information sheet and the informed consent.
■	the right to restrict processing – to ask that some or all processing ceases [see opt-out]	This is explained in the patient information sheet and the informed consent.
■	the right to data portability – this only applies to data provided directly by individual	This is explained in the patient information sheet and the informed consent.



■	the right to object to and not to be subject to automated decision-making, including profiling	This is explained in the patient information sheet and the informed consent.
■	Right to object to a Data Processing Authority (typically the relevant supervisory authority of each Member State)	This is explained in the patient information sheet and the informed consent.
■	Where consent is the legal basis, the right to withdraw consent	This is explained in the patient information sheet and the informed consent.

#### Detailed Transparency Checklist<sup>28</sup>

Does privacy information provided to data subjects include:

■	The name and contact details of our organisation	
<input type="checkbox"/>	The name and contact details of our representative (if applicable)	
<input type="checkbox"/>	The contact details of our data protection officer (if applicable)	
■	The purposes of the processing	
■	The lawful bases for the processing	[Art6 for 'personal data' & Art9 for 'special category']
■	The legitimate interests for the processing (if applicable)	
■	The categories of personal data obtained (if the personal data is not obtained from the individual it relates to)	[for Art14]
<input type="checkbox"/>	The recipients or categories of recipients of the personal data	

<sup>28</sup> Taken from UK Information Commissioner's Office template



<input type="checkbox"/>	The details of transfers of the personal data to any third countries or international organisations (if applicable)	
<input type="checkbox"/>	The retention periods for the personal data.	
<input checked="" type="checkbox"/>	The rights available to individuals in respect of the processing	
<input checked="" type="checkbox"/>	The right to withdraw consent (if applicable)	
<input checked="" type="checkbox"/>	The right to lodge a complaint with a supervisory authority	
<input type="checkbox"/>	The source of the personal data (if the personal data is not obtained from the individual it relates to)	[For Art14]
<input type="checkbox"/>	The details of whether individuals are under a statutory or contractual obligation to provide the personal data  (if applicable, and if the personal data is collected from the individual it relates to)	
<input type="checkbox"/>	The details of the existence of automated decision-making, including profiling (if applicable)	
<input checked="" type="checkbox"/>	We provide individuals with privacy information at the time we collect their personal data from them – or where we obtain personal data from a source other than the individual it relates to, we provide them with privacy information	
<input type="checkbox"/>	within a reasonable period of obtaining the personal data and no later than one month	
<input checked="" type="checkbox"/>	if we plan to communicate with the individual, at the latest, when the first communication takes place	



<ul style="list-style-type: none"> <li>■</li> </ul>	<p>if we plan to disclose the data to someone else, at the latest, when the data is disclosed</p>	
<ul style="list-style-type: none"> <li>■</li> </ul>	<p>We provide the information in a way that is:</p> <ul style="list-style-type: none"> <li>■ concise;</li> <li>■ transparent;</li> <li>■ intelligible;</li> <li>■ easily accessible; and</li> <li>■ Uses clear and plain language.</li> </ul>	<ul style="list-style-type: none"> <li>- I understand that my participation is voluntary and I am free to participate or not in the study.</li> <li>- I have been informed that all data obtained in this study will be confidential.</li> <li>- I have been informed that the donation / information obtained will only be used for the specific purposes of the study.</li> <li>- I understand that I can withdraw from the study, whenever I want, without having to give explanations, without this having an impact on my medical care.</li> </ul>
<ul style="list-style-type: none"> <li>■</li> </ul>	<p>When drafting the information, we:</p> <ul style="list-style-type: none"> <li>■ undertake an information audit to find out what personal data we hold and what we do with it.</li> <li>■ put ourselves in the position of the people we're collecting information about.</li> <li>■ carry out user testing to evaluate how effective our privacy information is</li> </ul>	<p>[Note: best practice advice]</p>
<ul style="list-style-type: none"> <li>■</li> </ul>	<p>When providing our privacy information to individuals, we use a combination of appropriate techniques, such as:</p> <ul style="list-style-type: none"> <li>■ a layered approach;</li> <li><input type="checkbox"/> dashboards;</li> <li><input type="checkbox"/> just-in-time notices;</li> <li><input type="checkbox"/> icons; and</li> <li><input type="checkbox"/> mobile and smart device functionalities.</li> </ul>	<p>[Note: best practice advice]</p>



### Security & Access Control Checklist

Controls need to be appropriate to level of risk: identified special category data needs more protection against potential misuse than non-personal data.

	Data Security classification (above Official)	<input checked="" type="checkbox"/> - Official-Sensitive <input type="checkbox"/> - Secret <input type="checkbox"/> - Top Secret <input type="checkbox"/> - Public Domain
<input checked="" type="checkbox"/>	Personal Data involved [GDPR]	
<input checked="" type="checkbox"/>	Special Category of personal data involved [GDPR]	
<input checked="" type="checkbox"/>	Electronic Communications (inc. cookies) [PECR]	
<input type="checkbox"/>	Credit Card data	
<input type="checkbox"/>	Legal enforcement [LED2018]	
<input type="checkbox"/>	Financial data	
<input type="checkbox"/>	Intellectual Property (detail owner)	
<input type="checkbox"/>	Commercial in confidence (detail owner)	
	Data Location (storage or processing) (include any back-up site(s))	<input type="checkbox"/> - UK <input checked="" type="checkbox"/> - EU/EEA <input type="checkbox"/> - EU White-list <input type="checkbox"/> - USA <input type="checkbox"/> - Other:
<input checked="" type="checkbox"/>	Is data held in secure data centre?	CSIC - Data protection officer  ( <a href="https://www.pre.sgai.csic.es/en/csic/data-protection">https://www.pre.sgai.csic.es/en/csic/data-protection</a> )



<input type="checkbox"/>	Is this new supplier, location, or system?	[If so, need specific IS check; also need formal contract]
<input checked="" type="checkbox"/>	Is all user access subject to 2-factor authentication?	<input type="checkbox"/> - no control <input type="checkbox"/> - single factor (e.g. just password) <input checked="" type="checkbox"/> - 2-factor (e.g. password & fob) <input type="checkbox"/> - biometric [note: GDPR reqs] <input type="checkbox"/> - Other control:
<input type="checkbox"/>	Are there established JML procedures?	[Joiners, Movers, Leavers]
<input checked="" type="checkbox"/>	Are there checks that passwords are robust and secure enough?	
<input checked="" type="checkbox"/>	Are all administrator & user accounts routinely monitored?	[Particularly for redundant or little used accounts]
<input checked="" type="checkbox"/>	Are systems protected against malware and other attacks?	

[Need some aspect of CIA/impact-likelihood assessment]

#### Information Asset Register Checklist

<input type="checkbox"/>	Are there new IAs being created?	[provide details]
<input type="checkbox"/>	Are old IAs being retired?	[provide details]
<input type="checkbox"/>	Have IAOs & IACs been consulted?	
<input type="checkbox"/>	Has IAR been updated/amended?	[at least create project task to do so]
<input type="checkbox"/>	Data Retention classification & period	
<input type="checkbox"/>	Data retention procedure/functionality in place	



## Appendix A – Supervisory Authority ‘High Risk’ Check

If the DPIA shows ‘high risk’ processing which cannot be mitigated, then the DPIA should be sent to the relevant authority for review before any processing starts. Note that their review may take several weeks to process. A ‘High Risk’ assessment represents a ‘risk to the rights and freedoms of individuals’ – so may extend beyond GDPR consideration, including Human Rights.

GDPR Article 35(3) provides three examples:

- m) a systematic and extensive evaluation of personal aspects relating to natural persons which is based on automated processing, including profiling, and on which decisions are based that produce legal effects concerning the natural person or similarly significantly affect the natural person;
- n) processing on a large scale of special categories of data referred to in Article 9(1), or of personal data relating to criminal convictions and offences referred to in Article 1013; or
- o) a systematic monitoring of a publicly accessible area on a large scale

ICO cites:

- 13. Systematic and extensive profiling with significant effects
- 14. Large scale use of sensitive data [viz. ‘special category’ in GDPR terms]
- 15. Public monitoring

These being the same as (a)-(c) above. They further identify:

- 41. **New technologies:** processing involving the use of new technologies, or the novel application of existing technologies (including AI).
- 42. **Denial of service:** Decisions about an individual’s access to a product, service, opportunity or benefit which is based to any extent on automated decision-making (including profiling) or involves the processing of special category data.
- 43. **Large-scale profiling:** any profiling of individuals on a large scale.
- 44. **Biometrics:** any processing of biometric data.
- 45. **Genetic data:** any processing of genetic data, other than that processed by an individual GP or health professional for the provision of health care direct to the data subject.
- 46. **Data matching:** combining, comparing or matching personal data obtained from multiple sources.
- 47. **Invisible processing:** processing of personal data that has not been obtained direct from the data subject in circumstances where the controller considers that compliance with Article 14 would prove impossible or involve disproportionate effort.
- 48. **Tracking:** processing which involves tracking an individual’s geolocation or behaviour, including but not limited to the online environment.
- 49. **Targeting of children or other vulnerable individuals:** The use of the personal data of children or other vulnerable individuals for marketing purposes, profiling or other automated decision-making, or if you intend to offer online services directly to children.





50. **Risk of physical harm:** Where the processing is of such a nature that a personal data breach could jeopardise the [physical] health or safety of individuals.

**'High Risk' assessment using ICO criteria:**

Criterion:	Assessment	Comments
New technologies		
Denial of service		
Large-scale profiling		
Biometrics		
Genetic data		
Data matching		
Invisible processing		
Tracking		
Targeting of children or other vulnerable individuals		



Criterion:	Assessment	Comments
Risk of physical harm		

[The assessment can be one of N/A (not applicable), Low, Medium, or High. The comments should explain how the assessment is justified.]

Appendix B – Broad Privacy Risk Assessment:

#	Risk Description/detail	Discussion
45.	Data accuracy and timeliness	[Is data accurately recorded & kept up-to-date?]
46.	Differential treatment of patients/data subjects	[Might certain categories of people be adversely affected, e.g. children, vulnerable adults]
47.	Data Accuracy and identification	[Is the identification of individual reliable? Is there a danger of mis-attribution or incorrect linkage of data?]
48.	Holding / sharing / use of excessive data within [Company] systems	[Might too much data be held or for long? Is there a clear justification for data retention? Not 'just in case']
49.	Data held too long within [Company] systems	[Is there a clear data retention period specified and are there processes to ensure its deletion when no longer needed? Are copies tracked and deleted as well?]
50.	Excessive range of access in terms of users to personal data (consider new users/change of access privileges)	[Do more users have access than strictly necessary? Are user roles clear distinguished and reflected in the access privileges? Is there a clear process for granting and revoking access privileges?]
51.	Potential for misuse of data, unauthorised access to systems	[What are the likely threats to the data? What countermeasures are or might be applied? Is it possible for access to be granted inappropriately?]
52.	New sharing of data with other organisations, including new or change of suppliers	[Is data being shared from new data providers or with new data users? Are there new suppliers or data processors? What controls will apply?]



#	Risk Description/detail	Discussion
53.	Variable and inconsistent adoption / implementation	[How well will this system work end-to-end? How robust is it against partial adoption or system failure?]
54.	Legal compliance, particularly DP transparency requirements and support for data subject rights	[How well does this system meet legal requirements – or appear to meet legal requirements? Does it meet the ‘No surprises’ rule? What would happen if an individual requests data erasure or ceasing processing, etc.]
55.	Medical confidentiality	[Are there any addition sensitivities over confidentiality? Might specific approval (e.g. REC) be required to support this processing?]



## IG Assessment Checklist ESR8 – Linking public and GCA datasets to identify novel pathogenic pathways

### Introduction to IG Assessment process

Under the General Data Protection Regulation (GDPR), a Data Protection Impact Assessment (DPIA) is only required where proposed data processing is “likely to result in a high risk to the rights and freedoms of natural persons” (Article 35(1)). However, Article 35(3) explicitly requires one where there is ‘large-scale’ processing of ‘special category’ (e.g. healthcare) data then a DPIA is required.

One other possibility is that the data being processed is already anonymised (see Recital 26) so falls outside GDPR altogether so that no DPIA is actually required.

However, good project management and information governance suggests that there should be a general approach to risk assessment for any project or business enterprise – if only to determine whether a DPIA might be required.

Ideally, one should work from a simple initial Checklist (this document) which identifies possible areas of information risk and compliance requirements to a ‘discussion note’ which explores any issues in more depth and may help identify the necessary mitigation methods and mechanisms to offset most if not all risks. Only if risks are unmitigated or remain ‘high’ would you move to a formal DPIA report.

### The IG Assessment approach

There should be an overview of the proposed project or business change to explain what processing is envisaged as well as the purpose and intended outcome. The ‘purpose’ is important to establish the legal basis for the processing as well as ensuring that any possible mitigations or counter-measures do not undermine the main rationale for the processing.

The next step is to establish what compliance requirements may apply: GDPR, contractual or other regulatory restrictions, consent requirements, or obligations to preserve the data for legal or other reasons (including the benefit of posterity perhaps).

Once the precise range of obligations has been established, then appropriate checks can be made and recorded within the document.

The most obvious of these being GDPR compliance. There must be a ‘High Risk’ assessment (Appendix A) to determine whether the supervisory authority needs to be informed – generally, it is expected that it will not be necessary; if so, then a formal DPIA report will be needed.

Appendix B has a broader Privacy Impact Assessment that may throw up some broader issues.

Initial conclusions as to next steps or particular countermeasures to be considered should be detailed below.



## Project Background/Overview

[Explain business background, including any existing processes and procedures; outline the project including stages, deliverables, and timelines]

Well-phenotyped GCA cohorts with genome-wide genotypic and some transcriptomic and histological data will be combined with publicly accessible molecular data of traits related to immune and vascular function and also matrix turnover.

Genetic scores (combining the effects across multiple known loci) and polygenic scores (combining the effects of many genetic variants with a relaxed threshold of statistical significance) will be generated, both from the analysis of transcripts and from relevant immunological, cardiovascular and tissue remodelling clinical phenotypes (e.g. from UK Biobank, published GWAS and publicly available datasets).

A discovery genome-wide association study of eQTLs and polygenic risk scores for GCA will be conducted, weighted by prior information. On the subset of patients with transcriptomic data, we expect that several transcripts will correlate with GCA susceptibility, subtypes and outcomes, but that only a subset of those will be causal. To separate transcripts that merely reflect GCA pathophysiology from those that are causal in disease, we will apply a Mendelian randomization approach by using the eQTLs as instrumental variables in a GCA case/control genetic study.

Transcript data will also be generated from a subset of GCA patients' FFPE temporal artery biopsies using RNASeq. Specific hypotheses relating transcript levels and eQTL to clinical or histological subtype or outcome will be formulated based on the earlier work and tested in this subset.

Novel meta-dimensional methods for combining genetic and transcriptomic data will be explored. Biological interpretation of genetic and genomic summary data is a major bottleneck in medical genomics research. A range of pathway analysis and drug discovery tools will be reviewed to determine those that will have the greatest chance of identifying pathways that are amenable to therapeutic manipulation. For example, the eXploring Genomic Relations (XGR) suite of bioinformatics tools, which utilises input GWAS and eQTL summary data, will initially be explored. This programme uses prior biological knowledge and relationships and has been used to explore the genomic landscape of the activated immune system and common immunological diseases.



Comparison of process steps (simplified): [optional]

This allows identification of what processing is new or changed through the project:

Step	Current	Proposed
Project initiation, including any ISAC approval, up to Task Order from client		No change

Initial Conclusions

concerning further counter-measures or business viability [possibly tentative]

10. ...

11. ...

Compliance Checks required:

All the answers are given from my perspective as a PhD student; I refer exclusively to the data that I have been given access to and handled.



Tick	Requirement	Notes [replace guide text with response]
<input type="checkbox"/>	Does the project involve processing 'personal data' of any sort?	<p>Note: not just patient data; may need clear assessment of any anonymization to establish outside GDPR</p> <p>All data have been anonymised at the stage of data collection and cleansing. All the files that I use, include 'sample IDs' which were assigned to each patient. I don't have access to the master file which includes patients' personal data.</p>
<input type="checkbox"/>	Does the project involve processing 'confidential data' of any sort?	<p>Note: may be 'commercial in confidence', medical confidentiality, or organisational confidentiality (internally sensitive); may need to check contractual limitations</p> <p>Similarly, as above the project involves processing medical records which are confidential data, but everything was anonymised at the very beginning, before I was granted access to it.</p>
<b>Data Availability requirements</b>		
<input type="checkbox"/>	Does data need to be held for GCP compliance?	<p>All clinical staff (i.e. people that generated the data) are obliged to follow 'the standards of Good Clinical Practice described in the UK Policy Framework for Health and Social Care 2017 and if you are working on a drug trial as described in UK law in the Medicines for Human Use (Clinical Trials) Regulation SI:1031 2004 and subsequent amendments.'</p>
<input type="checkbox"/>	Does data need to be held to meet 'Open Data' requirements?	No
<input type="checkbox"/>	Does data need to be held to meet ICMJE requirements or commitments?	No (not sure)



GDPR Compliance Checklist – where ‘personal data’ is processed:

Tick	Requirement	Notes [replace guide text with response]
<b>Article 5: Principles compliance checks</b>		
<input type="checkbox"/>	ee) Is processing lawful, fair, and transparent?	
<input type="checkbox"/>	ff) Is the purpose (or purposes) of the processing clearly defined	['purpose limitation' so should cover any subsequent or later processing]
<input type="checkbox"/>	gg) adequate, relevant and limited to what is necessary	['data minimisation']
<input type="checkbox"/>	hh) accurate and, where necessary, kept up to date	
<input type="checkbox"/>	ii) kept and permits identification of data subjects for no longer than is necessary	['storage limitation']
<input type="checkbox"/>	jj) processed securely	
<input type="checkbox"/>	7) can you demonstrate this compliance?	['accountability']
	The response to all the question from a) to f) is yes, according to the University's code of practice on data protection (source: <a href="https://dataprotection.leeds.ac.uk/data-protection-code-of-practice/">https://dataprotection.leeds.ac.uk/data-protection-code-of-practice/</a> )	
<b>Articles 13 &amp; 14 compliance</b>		[See detailed Transparency Checklist below]
<input type="checkbox"/>	Did the data came from publicly accessible sources?	[if so then transparency requirements may be reduced, but need to ensure data is accurate & up-to-date]  Yes, part of data comes from publicly accessible sources
<input type="checkbox"/>	Are data subjects informed before processing starts for any new purpose if incompatible with original purpose where the controller wants to use data for a different purpose to the purpose for which they currently hold data	Yes, the subjects are re-contacted and asked for consents.





Tick	Requirement	Notes [replace guide text with response]
<input type="checkbox"/>	Does the Privacy Notice and/or PIL cover this processing?	Yes, the Privacy Notice says that where the data need to be used in a new way, or to engage with external 3rd parties, the controller have to ask for explicit consent.
<input type="checkbox"/>	What patient choices are available? Are these explained?	<p>[see also Data Subject Rights below]</p> <p>Right to be informed</p> <p>Right to rectification</p> <p>Right to be forgotten</p> <p>Right to restriction of processing</p> <p>Right to data portability</p> <p>Right to object to automated decision-making, including individual decision-making and profiling</p>
<b>Articles 6 and 9: legal bases</b>		
<input type="checkbox"/>	What are legal bases under Article 6	<p>Art. 6(1)(a) Consent</p> <p>The individuals have given clear consent for you to process their personal data for a specific purpose</p>
<input type="checkbox"/>	What are legal bases under Article 9 (if 'special category' data)	Art. 9(2)(a) Explicit consent
<input type="checkbox"/>	Are Article 6 legitimate interests explained where relevant?	<p>[Complete an LIA form]</p> <p>Yes, the information about lawfulness, fairness and transparency of data processing is included.</p>
<input type="checkbox"/>	Are details of statutory obligations for Article 6 explained where relevant.	<p>[Quote statutes or regulation]</p> <p>'overall purpose must be to comply with a legal obligation which has a sufficiently clear basis in either common law or statute.'</p>



Tick	Requirement	Notes [replace guide text with response]
<input type="checkbox"/>	Is this proposed processing compatible with the declared purposes?	[Check against any privacy notices and public information]
<b>Article 89(1) research exemption</b>		
<input type="checkbox"/>	If for research, do we meet Art 89(1) data minimisation	Yes, Art 89(1) data minimisation is met by applying pseudonymisation measures which do not permit the identification of data subjects.
<b>Articles 15-23: Data Subject Rights</b>		[See detailed table below]
<input type="checkbox"/>	Do we support data subject rights?	[If data is pseudo-/anonymised, then it would be difficult/impossible to do so]
<input type="checkbox"/>	There is no use of automated decision making (e.g. profiling)	[Otherwise need at least a 'discussion note']  No
<b>Articles 24-43: Controller-Processor</b>		
<input type="checkbox"/>	A28 & 29: What measures are there to ensure processors comply?	[Is there a formal Data Processing Agreement]  Yes, there is a formal Data Processor Agreements (DPA), which needs to be signed
<input type="checkbox"/>	A30: Is there an entry for this processing/data held in the register?	
<input type="checkbox"/>	A32-34: Do we ensure appropriate security, including protection against unauthorised or unlawful processing and against accidental loss, destruction or damage, using appropriate technical or organisational measures?	[separate security checklist?]
<input type="checkbox"/>	A37-39: Is there a DPO and have they been or will they be consulted?	[part of sign-off of the DPIA]
<b>Articles 44-50: International transfers</b>		



Tick	Requirement	Notes [replace guide text with response]
	What form of data will be transferred to a third country or international organisation	[describe nature of data and whether identified, identifiable, de-identified or anonymous]  Gene expression data, along with additional information (clinical metadata) that were collected from data subjects.
<input type="checkbox"/>	Are there safeguards for international transfers?	[e.g. US Privacy Shield, anonymisation, GDPR equivalence, approved contractual clauses, or BCR]  Yes, Data Sharing Agreement contains all the information about 'Security of Processing' of data when doing transfers (ANNEX, part A)
<b>Article 90: Obligations of secrecy</b>		
<input type="checkbox"/>	Do we meet medical confidentiality requirements?	[Note any national case law and statutory requirements that may affect this]  Yes, the Ongoing confidentiality and integrity of data is assured by account access controls and restricted permissions.

#### Data Subject Rights:

Note if supported and what process/procedure applies; if not, then describe the legal justification for not supporting this right.

<input type="checkbox"/>	To be informed: about processing, about choices, about rights, about controller	supported
<input type="checkbox"/>	the right of access to see or receive a printed copy	supported
<input type="checkbox"/>	the right to rectification – to correct any material errors in the personal data	supported
<input type="checkbox"/>	the right to erasure – where appropriate, to ask that all personal data is erased	supported



<input type="checkbox"/>	the right to restrict processing – to ask that some or all processing ceases [see opt-out]	supported
<input type="checkbox"/>	the right to data portability – this only applies to data provided directly by individual	supported
<input type="checkbox"/>	the right to object to and not to be subject to automated decision-making, including profiling	supported
<input type="checkbox"/>	Right to object to a Data Processing Authority (typically the relevant supervisory authority of each Member State)	Not sure
<input type="checkbox"/>	Where consent is the legal basis, the right to withdraw consent	Not sure



Detailed Transparency Checklist<sup>29</sup>

Does privacy information provided to data subjects include:

<input type="checkbox"/>	The name and contact details of our organisation	Yes
<input type="checkbox"/>	The name and contact details of our representative (if applicable)	No
<input type="checkbox"/>	The contact details of our data protection officer (if applicable)	Yes
<input type="checkbox"/>	The purposes of the processing	Yes
<input type="checkbox"/>	The lawful bases for the processing	[Art6 for 'personal data' & Art9 for 'special category'  Yes
<input type="checkbox"/>	The legitimate interests for the processing (if applicable)	Yes
<input type="checkbox"/>	The categories of personal data obtained (if the personal data is not obtained from the individual it relates to)	[for Art14]  Yes
<input type="checkbox"/>	The recipients or categories of recipients of the personal data	Yes
<input type="checkbox"/>	The details of transfers of the personal data to any third countries or international organisations (if applicable)	Yes, includes the information about it, but not in great detail
<input type="checkbox"/>	The retention periods for the personal data.	Yes
<input type="checkbox"/>	The rights available to individuals in respect of the processing	Yes
<input type="checkbox"/>	The right to withdraw consent (if applicable)	Yes
<input type="checkbox"/>	The right to lodge a complaint with a supervisory authority	No

<sup>29</sup> Taken from UK Information Commissioner's Office template



<input type="checkbox"/>	The source of the personal data (if the personal data is not obtained from the individual it relates to)	[For Art14]  Not applicable
<input type="checkbox"/>	The details of whether individuals are under a statutory or contractual obligation to provide the personal data  (if applicable, and if the personal data is collected from the individual it relates to)	Not applicable
<input type="checkbox"/>	The details of the existence of automated decision-making, including profiling (if applicable)	No
<input type="checkbox"/>	We provide individuals with privacy information at the time we collect their personal data from them – or where we obtain personal data from a source other than the individual it relates to, we provide them with privacy information	Yes
<input type="checkbox"/>	within a reasonable of period of obtaining the personal data and no later than one month	Not sure
<input type="checkbox"/>	if we plan to communicate with the individual, at the latest, when the first communication takes place	Not sure
<input type="checkbox"/>	if we plan to disclose the data to someone else, at the latest, when the data is disclosed	Not applicable
<input type="checkbox"/>	We provide the information in a way that is:  <input type="checkbox"/> concise;  <input type="checkbox"/> transparent;  <input type="checkbox"/> intelligible;  <input type="checkbox"/> easily accessible; and  <input type="checkbox"/> uses clear and plain language.	[Describe how we check is Plain English, etc.]
<input type="checkbox"/>	When drafting the information, we:	[Note: best practice advice]



	<input type="checkbox"/> undertake an information audit to find out what personal data we hold and what we do with it. <input type="checkbox"/> put ourselves in the position of the people we're collecting information about. <input type="checkbox"/> carry out user testing to evaluate how effective our privacy information is	
<input type="checkbox"/>	<p>When providing our privacy information to individuals, we use a combination of appropriate techniques, such as:</p> <input type="checkbox"/> a layered approach; <input type="checkbox"/> dashboards; <input type="checkbox"/> just-in-time notices; <input type="checkbox"/> icons; and <input type="checkbox"/> mobile and smart device functionalities.	<p>[Note: best practice advice]</p>



### Security & Access Control Checklist

Controls need to be appropriate to level of risk: identified special category data needs more protection against potential misuse than non-personal data.

	Data Security classification (above Official)	<input type="checkbox"/> - Official-Sensitive <input type="checkbox"/> - Secret <input type="checkbox"/> - Top Secret <input type="checkbox"/> - Public Domain
<input type="checkbox"/>	Personal Data involved [GDPR]	Yes
<input type="checkbox"/>	Special Category of personal data involved [GDPR]	Yes
<input type="checkbox"/>	Electronic Communications (inc. cookies) [PECR]	No
<input type="checkbox"/>	Credit Card data	No
<input type="checkbox"/>	Legal enforcement [LED2018]	No
<input type="checkbox"/>	Financial data	No
<input type="checkbox"/>	Intellectual Property (detail owner)	No
<input type="checkbox"/>	Commercial in confidence (detail owner)	No
	Data Location (storage or processing)  (include any back-up site(s))	<input type="checkbox"/> - UK <input type="checkbox"/> - EU/EEA <input type="checkbox"/> - EU White-list <input type="checkbox"/> - USA <input type="checkbox"/> - Other:
<input type="checkbox"/>	Is data held in secure data centre?	[detail centre and what certification supports assertion]  Yes, on encrypted university computers (Sophos 'SafeGuard' software used by the university)





<input type="checkbox"/>	Is this new supplier, location, or system?	[If so, need specific IS check; also need formal contract]  <b>No</b>
<input type="checkbox"/>	Is all user access subject to 2-factor authentication?	<input type="checkbox"/> - no control  <input type="checkbox"/> - single factor (e.g. just password)  <input type="checkbox"/> - 2-factor (e.g. password & fob)  <input type="checkbox"/> - biometric [note: GDPR reqs]  <input type="checkbox"/> - Other control:
<input type="checkbox"/>	Are there established JML procedures?	[Joiners, Movers, Leavers]  <b>No</b>
<input type="checkbox"/>	Are there checks that passwords are robust and secure enough?	[ ]  <b>Yes</b>
<input type="checkbox"/>	Are all administrator & user accounts routinely monitored?	[Particularly for redundant or little used accounts]  <b>Yes</b>
<input type="checkbox"/>	Are systems protected against malware and other attacks?	[provide details of protection software and procedures]  <b>Yes, McAfee VirusScan should be installed on all University PCs and laptops</b>

[Need some aspect of CIA/impact-likelihood assessment]

#### Information Asset Register Checklist

<input type="checkbox"/>	Are there new IAs being created?	[provide details] <b>No</b>
<input type="checkbox"/>	Are old IAs being retired?	[provide details] <b>No</b>
<input type="checkbox"/>	Have IAOs & IACs been consulted?	<b>Not sure</b>
<input type="checkbox"/>	Has IAR been updated/amended?	[at least create project task to do so]



		Not sure
<input type="checkbox"/>	Data Retention classification & period	Not sure
<input type="checkbox"/>	Data retention procedure/functionality in place	Not sure



## Appendix A – Supervisory Authority ‘High Risk’ Check

If the DPIA shows ‘high risk’ processing which cannot be mitigated, then the DPIA should be sent to the relevant authority for review before any processing starts. Note that their review may take several weeks to process. A ‘High Risk’ assessment represents a ‘risk to the rights and freedoms of individuals’ – so may extend beyond GDPR consideration, including Human Rights.

GDPR Article 35(3) provides three examples:

- p) a systematic and extensive evaluation of personal aspects relating to natural persons which is based on automated processing, including profiling, and on which decisions are based that produce legal effects concerning the natural person or similarly significantly affect the natural person;
- q) processing on a large scale of special categories of data referred to in Article 9(1), or of personal data relating to criminal convictions and offences referred to in Article 1013; or
- r) a systematic monitoring of a publicly accessible area on a large scale

ICO cites:

- 16. Systematic and extensive profiling with significant effects
- 17. Large scale use of sensitive data [viz. ‘special category’ in GDPR terms]
- 18. Public monitoring

These being the same as (a)-(c) above. They further identify:

- 51. **New technologies:** processing involving the use of new technologies, or the novel application of existing technologies (including AI).
- 52. **Denial of service:** Decisions about an individual’s access to a product, service, opportunity or benefit which is based to any extent on automated decision-making (including profiling) or involves the processing of special category data.
- 53. **Large-scale profiling:** any profiling of individuals on a large scale.
- 54. **Biometrics:** any processing of biometric data.
- 55. **Genetic data:** any processing of genetic data, other than that processed by an individual GP or health professional for the provision of health care direct to the data subject.
- 56. **Data matching:** combining, comparing or matching personal data obtained from multiple sources.
- 57. **Invisible processing:** processing of personal data that has not been obtained direct from the data subject in circumstances where the controller considers that compliance with Article 14 would prove impossible or involve disproportionate effort.
- 58. **Tracking:** processing which involves tracking an individual’s geolocation or behaviour, including but not limited to the online environment.
- 59. **Targeting of children or other vulnerable individuals:** The use of the personal data of children or other vulnerable individuals for marketing purposes, profiling or other automated decision-making, or if you intend to offer online services directly to children.



60. **Risk of physical harm:** Where the processing is of such a nature that a personal data breach could jeopardise the [physical] health or safety of individuals.

**'High Risk' assessment using ICO criteria:**

Criterion:	Assessment	Comments
New technologies	Low	
Denial of service	N/A	
Large-scale profiling	High	
Biometrics	N/A	
Genetic data	High	
Data matching	Low	
Invisible processing	N/A	
Tracking	N/A	
Targeting of children or other vulnerable individuals	N/A	



Criterion:	Assessment	Comments
Risk of physical harm	N/A	

[The assessment can be one of N/A (not applicable), Low, Medium, or High. The comments should explain how the assessment is justified.]

Appendix B – Broad Privacy Risk Assessment:

#	Risk Description/detail	Discussion
56.	Data accuracy and timeliness	[Is data accurately recorded & kept up-to-date?] <b>Yes</b>
57.	Differential treatment of patients/data subjects	[Might certain categories of people be adversely affected, e.g. children, vulnerable adults] <b>No</b>
58.	Data Accuracy and identification	[Is the identification of individual reliable? <b>Yes</b> Is there a danger of mis-attribution or incorrect linkage of data?] <b>No</b>
59.	Holding / sharing / use of excessive data within [Company] systems	[Might too much data be held or for long? <b>No</b> Is there a clear justification for data retention? Not 'just in case'] <b>Yes, it will be used in the future</b>
60.	Data held too long within [Company] systems	[Is there a clear data retention period specified and are there processes to ensure its deletion when no longer needed? <b>No</b> Are copies tracked and deleted as well?] <b>Yes</b>
61.	Excessive range of access in terms of users to personal data (consider new users/change of access privileges)	[Do more users have access than strictly necessary? <b>No</b> Are user roles clear distinguished and reflected in the access privileges? <b>Yes</b> Is there a clear process for granting and revoking access privileges?] <b>Yes</b>
62.	Potential for misuse of data, unauthorised access to systems	[What are the likely threats to the data? <b>None</b> What countermeasures are or might be applied? Is it possible for access to be granted inappropriately?] <b>No</b>



#	Risk Description/detail	Discussion
63.	New sharing of data with other organisations, including new or change of suppliers	[Is data being shared from new data providers or with new data users? <b>Yes</b> Are there new suppliers or data processors? <b>Yes</b> What controls will apply?]
64.	Variable and inconsistent adoption / implementation	[How well will this system work end-to-end? How robust is it against partial adoption or system failure?]
65.	Legal compliance, particularly DP transparency requirements and support for data subject rights	[How well does this system meet legal requirements – or appear to meet legal requirements? Does it meet the ‘No surprises’ rule? What would happen if an individual requests data erasure or ceasing processing, etc.]
66.	Medical confidentiality	[Are there any addition sensitivities over confidentiality? Might specific approval (e.g. REC) be required to support this processing?] <b>No</b>



## IG Assessment Checklist ESR9 – Systems biology and bioinformatics approaches to provide a holistic understanding of GCA biology

### Introduction to IG Assessment process

Under the General Data Protection Regulation (GDPR), a Data Protection Impact Assessment (DPIA) is only required where proposed data processing is “likely to result in a high risk to the rights and freedoms of natural persons” (Article 35(1)). However, Article 35(3) explicitly requires one where there is ‘large-scale’ processing of ‘special category’ (e.g. healthcare) data then a DPIA is required.

One other possibility is that the data being processed is already anonymised (see Recital 26) so falls outside GDPR altogether so that no DPIA is actually required.

However, good project management and information governance suggests that there should be a general approach to risk assessment for any project or business enterprise – if only to determine whether a DPIA might be required.

Ideally, one should work from a simple initial Checklist (this document) which identifies possible areas of information risk and compliance requirements to a ‘discussion note’ which explores any issues in more depth and may help identify the necessary mitigation methods and mechanisms to offset most if not all risks. Only if risks are unmitigated or remain ‘high’ would you move to a formal DPIA report.

### The IG Assessment approach

There should be an overview of the proposed project or business change to explain what processing is envisaged as well as the purpose and intended outcome. The ‘purpose’ is important to establish the legal basis for the processing as well as ensuring that any possible mitigations or counter-measures do not undermine the main rationale for the processing.

The next step is to establish what compliance requirements may apply: GDPR, contractual or other regulatory restrictions, consent requirements, or obligations to preserve the data for legal or other reasons (including the benefit of posterity perhaps).

Once the precise range of obligations has been established, then appropriate checks can be made and recorded within the document.

The most obvious of these being GDPR compliance. There must be a ‘High Risk’ assessment (Appendix A) to determine whether the supervisory authority needs to be informed – generally, it is expected that it will not be necessary; if so, then a formal DPIA report will be needed.

Appendix B has a broader Privacy Impact Assessment that may throw up some broader issues.

Initial conclusions as to next steps or particular countermeasures to be considered should be detailed below.



Project Background/Overview

- ◆ The project regards the extension of the Anaxomics analytical pipelines for the automatic mining of large biomedical databases. The new pipelines here adopted include the automatic translation of phenotype/clinical information into the molecular description of each patient, and the application of features selection and classification algorithm for the identification of pathway significantly related to GCA biology and its comorbidities.

Comparison of process steps (simplified): [optional]

This allows identification of what processing is new or changed through the project:

Step	Current	Proposed
Project initiation, including any ISAC approval, up to Task Order from client (ethical approval example)		No change

Initial Conclusions

concerning further counter-measures or business viability [possibly tentative]

- 12. ...
- 13. ...





Compliance Checks required:

Tick	Requirement	Notes [replace guide text with response]
<input type="checkbox"/>	Does the project involve processing 'personal data' of any sort?	Yes. Demographical data together with the clinical description of each patient are taken into account in everything related to the NHANES database.  Gene expression data obtained from partners (ESR10) are used too.
<input type="checkbox"/>	Does the project involve processing 'confidential data' of any sort?	Yes, just when gene expression data are used.
<b>Data Availability requirements</b>		
<input type="checkbox"/>	Does data need to be held for GCP (good clinical practice) compliance?	Yes, data production procedure respects the GCP
<input type="checkbox"/>	Does data need to be held to meet 'Open Data' requirements?	All data about NHANES are already public. On contrary, the gene expression data used in the context of the secondment will be public in the moment of the publication.
<input type="checkbox"/>	Does data need to be held to meet ICMJE (International Committee of Medical Journal Editor) requirements or commitments?	No



GDPR Compliance Checklist – where ‘personal data’ is processed:

Tick	Requirement	Notes [replace guide text with response]
<b>Article 5: Principles compliance checks</b>		
<input type="checkbox"/>	kk) Is processing lawful, fair, and transparent?	Yes, we are working in the public interest toward progress in health care related to GCA.
<input type="checkbox"/>	ll) Is the purpose (or purposes) of the processing clearly defined	Yes, the purpose is to perform epidemiological and molecular analysis with the data
<input type="checkbox"/>	mm) adequate, relevant and limited to what is necessary	Yes, we describe patients just with sex, age, ethnicity, clinical profile or gene expression. No other kind is taken into account.
<input type="checkbox"/>	nn) accurate and, where necessary, kept up to date	The data are public, and we keep track of each modification we do with the original database.
<input type="checkbox"/>	oo) kept and permits identification of data subjects for no longer than is necessary	<p>NHANES data are anonymized and public, so we have no identification of the patients.</p> <p>For what concern non-public gene expression data coming for partners, the time of data availability and usage is limited to what the contract with partners providing such states.</p> <p>The data permits are responsibility of the data producer and not ours.</p>



Tick	Requirement	Notes [replace guide text with response]
<input type="checkbox"/>	pp) processed securely	Anaxomics works under standard operation procedures and is ISO 9001 and ISO 27001 certified
<input type="checkbox"/>	8) can you demonstrate this compliance?	Yes, compliance is demonstrated by ISO 27001 certification
<b>Articles 13 &amp; 14 compliance</b>		[See detailed Transparency Checklist below]
<input type="checkbox"/>	Did the data come from publicly accessible sources?	Only the data from the NHANES database comes from a public resource. On contrary the data used in the context of collaboration with partners are private, and are shared in anonymized form.
<input type="checkbox"/>	Are data subjects informed before processing starts for any new purpose if incompatible with original purpose where the controller wants to use data for a different purpose to the purpose for which they currently hold data	Yes
<input type="checkbox"/>	Does the Privacy Notice and/or PIL (patients information leaflet) cover this processing?	For Nhanes, the processing is covered in the PIL.  For gene expression data, refer to ESR10



Tick	Requirement	Notes [replace guide text with response]
<input type="checkbox"/>	What patient choices are available? Are these explained?	Not for Nhanes  For gene expression data, refer to ESR10  [data subject rights:  Right to be informed  Right to rectification  Right to be forgotten  Right to restriction of processing  Right to data portability  Right to object to automated decision-making, including individual decision-making and profiling]
<b>Articles 6 and 9: legal bases</b>		
<input type="checkbox"/>	<b>What are legal bases under Article 6?</b> (Article 6 EU GDPR "Lawfulness of processing" => Recital: 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 155 => administrative fine: Art. 83 (5) lit a 1. Processing shall be lawful only if and to the extent that at least one of the following applies: => Article: 9 (a) the data subject has given consent to the processing of his or her personal data for one or more specific purposes; => Article: 7, 8, 9 => Recital: 32, 42, 43, 171 => Dossier: Consent, Permission (b) processing is necessary for the performance of a contract to which the data subject is party or in order to take steps at the request of the data subject prior to entering into a contract; => Article: 20 => Dossier: Permission (c) processing is necessary for compliance with a legal obligation to which the controller is subject; => Dossier: Permission (d) processing is necessary in order to protect the vital interests of the data subject or of another natural person; => Dossier: Permission (e) processing is necessary for the performance of a task carrier)	The project act under the Public interest as an initiative to further investigate the molecular basis of Giant Cell Arteritis.



Tick	Requirement	Notes [replace guide text with response]
<input type="checkbox"/>	What are legal bases under Article 9 (if 'special category' data)	Sex and Ethnicity are contained in the Nhanes database but this information is only used for risk analysis with a beneficial health aim.  Gene expression data are used uniquely to identify genes behaviour related to the disease and not for the identification of the individual. No other special category data are included in the dataset.
<input type="checkbox"/>	Are Article 6 legitimate interests explained where relevant?	The data subject has given consent to the processing of his or her personal data for one or more specific purposes.
<input type="checkbox"/>	Are details of statutory obligations for Article 6 explained where relevant.	
<input type="checkbox"/>	Is this proposed processing compatible with the declared purposes?	
<b>Article 89(1) research exemption</b>		
<input type="checkbox"/>	If for research, do we meet Art 89(1) data minimisation	Yes, data are limited to what is necessary for the purposes for which they are processed.
<b>Articles 15-23: Data Subject Rights</b>		[See detailed table below]
<input type="checkbox"/>	Do we support data subject rights?	The data subject is informed on any of the applications for which the data will be used
<input type="checkbox"/>	There is no use of automated decision making (e.g. profiling)	No profiling is applied
<b>Articles 24-43: Controller-Processor</b>		
<input type="checkbox"/>	A28 & 29: What measures are there to ensure processors comply?	We are not the data generator both for Nhanes and for Gene expression data from partners.



Tick	Requirement	Notes [replace guide text with response]
<input type="checkbox"/>	A30: Is there an entry for this processing/data held in the register?	We are not the data generator both for Nhanes and for Gene expression data from partners.
<input type="checkbox"/>	A32-34: Do we ensure appropriate security, including protection against unauthorised or unlawful processing and against accidental loss, destruction or damage, using appropriate technical or organisational measures?	Anaxomics works under standard operation procedures and is ISO 9001 and ISO 27001 certified
<input type="checkbox"/>	A37-39: Is there a DPO and have they been or will they be consulted?	Yes, no specific consultation has been done.
<b>Articles 44-50: International transfers</b>		
	What form of data will be transferred to a third country or international organisation	No, there will be no international transfer.
<input type="checkbox"/>	Are there safeguards for international transfers?	No, there will be no international transfer.
<b>Article 90: Obligations of secrecy</b>		
<input type="checkbox"/>	Do we meet medical confidentiality requirements?	Yes

#### Data Subject Rights:

Note if supported and what process/procedure applies; if not, then describe the legal justification for not supporting this right.

<input type="checkbox"/>	To be informed: about processing, about choices, about rights, about controller	We are not the data generator
<input type="checkbox"/>	the right of access to see or receive a printed copy	We are not the data generator
<input type="checkbox"/>	the right to rectification – to correct any material errors in the personal data	We are not the data generator



<input type="checkbox"/>	the right to erasure – where appropriate, to ask that all personal data is erased	We are not the data generator
<input type="checkbox"/>	the right to restrict processing – to ask that some or all processing ceases [see opt-out]	We are not the data generator
<input type="checkbox"/>	the right to data portability – this only applies to data provided directly by individual	We are not the data generator
<input type="checkbox"/>	the right to object to and not to be subject to automated decision-making, including profiling	We are not the data generator
<input type="checkbox"/>	Right to object to a Data Processing Authority (typically the relevant supervisory authority of each Member State)	We are not the data generator
<input type="checkbox"/>	Where consent is the legal basis, the right to withdraw consent	We are not the data generator

#### Detailed Transparency Checklist<sup>30</sup>

Does privacy information provided to data subjects include:

<input type="checkbox"/>	The name and contact details of our organisation	We are not the data generator
<input type="checkbox"/>	The name and contact details of our representative (if applicable)	We are not the data generator
<input type="checkbox"/>	The contact details of our data protection officer (if applicable)	We are not the data generator
<input type="checkbox"/>	The purposes of the processing	We are not the data generator
<input type="checkbox"/>	The lawful bases for the processing	We are not the data generator
<input type="checkbox"/>	The legitimate interests for the processing (if applicable)	We are not the data generator

<sup>30</sup> Taken from UK Information Commissioner's Office template



<input type="checkbox"/>	The categories of personal data obtained (if the personal data is not obtained from the individual it relates to)	We are not the data generator
<input type="checkbox"/>	The recipients or categories of recipients of the personal data	We are not the data generator
<input type="checkbox"/>	The details of transfers of the personal data to any third countries or international organisations (if applicable)	We are not the data generator
<input type="checkbox"/>	The retention periods for the personal data.	We are not the data generator
<input type="checkbox"/>	The rights available to individuals in respect of the processing	We are not the data generator
<input type="checkbox"/>	The right to withdraw consent (if applicable)	We are not the data generator
<input type="checkbox"/>	The right to lodge a complaint with a supervisory authority	We are not the data generator
<input type="checkbox"/>	The source of the personal data (if the personal data is not obtained from the individual it relates to)	We are not the data generator
<input type="checkbox"/>	The details of whether individuals are under a statutory or contractual obligation to provide the personal data  (if applicable, and if the personal data is collected from the individual it relates to)	We are not the data generator
<input type="checkbox"/>	The details of the existence of automated decision-making, including profiling (if applicable)	We are not the data generator
<input type="checkbox"/>	We provide individuals with privacy information at the time we collect their personal data from them – or where we obtain personal data from a source other than the individual it relates to, we provide them with privacy information	We are not the data generator





<input type="checkbox"/>	within a reasonable of period of obtaining the personal data and no later than one month	We are not the data generator
<input type="checkbox"/>	if we plan to communicate with the individual, at the latest, when the first communication takes place	We are not the data generator
<input type="checkbox"/>	if we plan to disclose the data to someone else, at the latest, when the data is disclosed	We are not the data generator
<input type="checkbox"/>	We provide the information in a way that is: <ul style="list-style-type: none"> <li><input type="checkbox"/> concise;</li> <li><input type="checkbox"/> transparent;</li> <li><input type="checkbox"/> intelligible;</li> <li><input type="checkbox"/> easily accessible; and</li> <li><input type="checkbox"/> uses clear and plain language.</li> </ul>	[Describe how we check is Plain English, etc.]
<input type="checkbox"/>	When drafting the information, we: <ul style="list-style-type: none"> <li><input type="checkbox"/> undertake an information audit to find out what personal data we hold and what we do with it.</li> <li><input type="checkbox"/> put ourselves in the position of the people we're collecting information about.</li> <li><input type="checkbox"/> carry out user testing to evaluate how effective our privacy information is</li> </ul>	[Note: best practice advice]
<input type="checkbox"/>	When providing our privacy information to individuals, we use a combination of appropriate techniques, such as: <ul style="list-style-type: none"> <li><input type="checkbox"/> a layered approach;</li> <li><input type="checkbox"/> dashboards;</li> <li><input type="checkbox"/> just-in-time notices;</li> </ul>	[Note: best practice advice]



<input type="checkbox"/> icons; and  <input type="checkbox"/> mobile and smart device functionalities.	
--	--

### Security & Access Control Checklist

Controls need to be appropriate to level of risk: identified special category data needs more protection against potential misuse than non-personal data.

	Data Security classification (above Official)	<input type="checkbox"/> - Official-Sensitive  <input type="checkbox"/> - <b>Secret</b> (For gene expression)  <input type="checkbox"/> - Top Secret  <input type="checkbox"/> - <b>Public Domain</b> (For NHANES)
<input type="checkbox"/>	<b>Personal Data involved [GDPR]</b>	
<input type="checkbox"/>	Special Category of personal data involved [GDPR]	
<input type="checkbox"/>	Electronic Communications (inc. cookies) [PECR]	
<input type="checkbox"/>	Credit Card data	
<input type="checkbox"/>	Legal enforcement [LED2018]	
<input type="checkbox"/>	Financial data	
<input type="checkbox"/>	Intellectual Property (detail owner)	
<input type="checkbox"/>	Commercial in confidence (detail owner)	
	Data Location (storage or processing) (include any back-up site(s))	<input type="checkbox"/> - UK  <input type="checkbox"/> - <b>EU/EEA</b>  <input type="checkbox"/> - EU White-list  <input type="checkbox"/> - <b>USA</b>  <input type="checkbox"/> - Other:
<input type="checkbox"/>	Is data held in secure data centre?	yes
<input type="checkbox"/>	Is this new supplier, location, or system?	



<input type="checkbox"/>	Is all user access subject to 2-factor authentication?	<input type="checkbox"/> - no control <input type="checkbox"/> - single factor (e.g. just password) <input type="checkbox"/> - 2-factor (e.g. password & fob) <input type="checkbox"/> - biometric [note: GDPR reqs] <input type="checkbox"/> - Other control:
<input type="checkbox"/>	Are there established JML procedures?	
<input type="checkbox"/>	Are there checks that passwords are robust and secure enough?	yes
<input type="checkbox"/>	Are all administrator & user accounts routinely monitored?	yes
<input type="checkbox"/>	Are systems protected against malware and other attacks?	yes

[Need some aspect of CIA/impact-likelihood assessment]

#### Information Asset Register Checklist

<input type="checkbox"/>	Are there new IAs being created?	
<input type="checkbox"/>	Are old IAs being retired?	
<input type="checkbox"/>	Have IAOs & IACs been consulted?	
<input type="checkbox"/>	Has IAR been updated/amended?	
<input type="checkbox"/>	Data Retention classification & period	yes
<input type="checkbox"/>	Data retention procedure/functionality in place	



## Appendix A – Supervisory Authority ‘High Risk’ Check

If the DPIA shows ‘high risk’ processing which cannot be mitigated, then the DPIA should be sent to the relevant authority for review before any processing starts. Note that their review may take several weeks to process. A ‘High Risk’ assessment represents a ‘risk to the rights and freedoms of individuals’ – so may extend beyond GDPR consideration, including Human Rights.

GDPR Article 35(3) provides three examples:

- s) a systematic and extensive evaluation of personal aspects relating to natural persons which is based on automated processing, including profiling, and on which decisions are based that produce legal effects concerning the natural person or similarly significantly affect the natural person;
- t) processing on a large scale of special categories of data referred to in Article 9(1), or of personal data relating to criminal convictions and offences referred to in Article 1013; or
- u) a systematic monitoring of a publicly accessible area on a large scale

ICO cites:

- 19. Systematic and extensive profiling with significant effects
- 20. Large scale use of sensitive data [viz. ‘special category’ in GDPR terms]
- 21. Public monitoring

These being the same as (a)-(c) above. They further identify:

- 61. **New technologies:** processing involving the use of new technologies, or the novel application of existing technologies (including AI).
- 62. **Denial of service:** Decisions about an individual’s access to a product, service, opportunity or benefit which is based to any extent on automated decision-making (including profiling) or involves the processing of special category data.
- 63. **Large-scale profiling:** any profiling of individuals on a large scale.
- 64. **Biometrics:** any processing of biometric data.
- 65. **Genetic data:** any processing of genetic data, other than that processed by an individual GP or health professional for the provision of health care direct to the data subject.
- 66. **Data matching:** combining, comparing or matching personal data obtained from multiple sources.
- 67. **Invisible processing:** processing of personal data that has not been obtained direct from the data subject in circumstances where the controller considers that compliance with Article 14 would prove impossible or involve disproportionate effort.
- 68. **Tracking:** processing which involves tracking an individual’s geolocation or behaviour, including but not limited to the online environment.
- 69. **Targeting of children or other vulnerable individuals:** The use of the personal data of children or other vulnerable individuals for marketing purposes, profiling or other automated decision-making, or if you intend to offer online services directly to children.



70. **Risk of physical harm:** Where the processing is of such a nature that a personal data breach could jeopardise the [physical] health or safety of individuals.

**'High Risk' assessment using ICO criteria:**

Criterion:	Assessment	Comments
New technologies	LOW	New AI are developed but only on non-linked anonymized data
Denial of service	NA	
Large-scale profiling	NA	No profiling is applied except for metadata extraction (SEX and AGE)
Biometrics	LOW	Biometric data are used in anonymized form
Genetic data	LOW	Genetic data are used in anonymized form
Data matching	NA	No link/matching applied
Invisible processing	NA	No other source other <u>then</u> the ones with informed consent are used.
Tracking	NA	No geographical / temporal data are used
Targeting of children or other vulnerable individuals	NA	None of this are involved in the process



Criterion:	Assessment	Comments
Risk of physical harm	NA	None

[The assessment can be one of N/A (not applicable), Low, Medium, or High. The comments should explain how the assessment is justified.]

Appendix B – Broad Privacy Risk Assessment:

#	Risk Description/detail	Discussion
67.	Data accuracy and timeliness	Data are guaranteed to be accurate.
68.	Differential treatment of patients/data subjects	No, the same data and same policy are applied to each patient.
69.	Data Accuracy and identification	[Is the identification of individual reliable? Is there a danger of mis-attribution or incorrect linkage of data?]  There is no linkage
70.	Holding / sharing / use of excessive data within [Company] systems	[Might too much data be held or for long? Is there a clear justification for data retention? Not 'just in case']  The non-public data are held just for the time needed for the processing.
71.	Data held too long within [Company] systems	[Is there a clear data retention period specified and are there processes to ensure its deletion when no longer needed? Are copies tracked and deleted as well?]  This aspect will be defined in future by the contract with the partners
72.	Excessive range of access in terms of users to personal data (consider new users/change of access privileges)	[Do more users have access than strictly necessary? Are user roles clear distinguished and reflected in the access privileges? Is there a clear process for granting and revoking access privileges?]  The roles of the user are defined by the finality of the research, and remain strict to it.



#	Risk Description/detail	Discussion
73.	Potential for misuse of data, unauthorised access to systems	[What are the likely threats to the data? What countermeasures are or might be applied? Is it possible for access to be granted inappropriately?]  Data are protected by firewall
74.	New sharing of data with other organisations, including new or change of suppliers	[Is data being shared from new data providers or with new data users? Are there new suppliers or data processors? What controls will apply?]  No, the contract with the partner limits its usage to the member of the project, no other third part can access the data
75.	Variable and inconsistent adoption / implementation	
76.	Legal compliance, particularly DP transparency requirements and support for data subject rights	[How well does this system meet legal requirements – or appear to meet legal requirements? Does it meet the ‘No surprises’ rule? What would happen if an individual requests data erasure or ceasing processing, etc.]
77.	Medical confidentiality	[Are there any addition sensitivities over confidentiality? Might specific approval (e.g. REC) be required to support this processing?]



## IG Assessment Checklist ESR10 – [Functional characterisation of inflammation and vascular remodeling pathways in GCA, IDIBAPS, Barcelona]

### Introduction to IG Assessment process

Under the General Data Protection Regulation (GDPR), a Data Protection Impact Assessment (DPIA) is only required where proposed data processing is “likely to result in a high risk to the rights and freedoms of natural persons” (Article 35(1)). However, Article 35(3) explicitly requires one where there is ‘large-scale’ processing of ‘special category’ (e.g. healthcare) data then a DPIA is required.

One other possibility is that the data being processed is already anonymised (see Recital 26) so falls outside GDPR altogether so that no DPIA is actually required.

However, good project management and information governance suggests that there should be a general approach to risk assessment for any project or business enterprise – if only to determine whether a DPIA might be required.

Ideally, one should work from a simple initial Checklist (this document) which identifies possible areas of information risk and compliance requirements to a ‘discussion note’ which explores any issues in more depth and may help identify the necessary mitigation methods and mechanisms to offset most if not all risks. Only if risks are unmitigated or remain ‘high’ would you move to a formal DPIA report.

### The IG Assessment approach

There should be an overview of the proposed project or business change to explain what processing is envisaged as well as the purpose and intended outcome. The ‘purpose’ is important to establish the legal basis for the processing as well as ensuring that any possible mitigations or counter-measures do not undermine the main rationale for the processing.

The next step is to establish what compliance requirements may apply: GDPR, contractual or other regulatory restrictions, consent requirements, or obligations to preserve the data for legal or other reasons (including the benefit of posterity perhaps).

Once the precise range of obligations has been established, then appropriate checks can be made and recorded within the document.

The most obvious of these being GDPR compliance. There must be a ‘High Risk’ assessment (Appendix A) to determine whether the supervisory authority needs to be informed – generally, it is expected that it will not be necessary; if so, then a formal DPIA report will be needed.

Appendix B has a broader Privacy Impact Assessment that may throw up some broader issues.

Initial conclusions as to next steps or particular countermeasures to be considered should be detailed below.





Project Background/Overview

Testing the effect of only available biological therapy for GCA – tocilizumab (Actemra) and finding biomarkers predictors of response. Project start: February 2020, finish: January 2023.

Comparison of process steps (simplified): [optional]

This allows identification of what processing is new or changed through the project:

Step	Current	Proposed
Project initiation, Clinical Research Ethics Committee (February 2020)	This is the current step	No change
Taking samples from patients included in the study (March)	Delayed due to Covid-19	From 25 <sup>th</sup> of June
Sample processing and testing efficacy of new therapy (April – September)	Delayed due to Covid-19	From July – September 2021
Secondment CSIC Granada - genetic analysis (September – December 2020)	September – December 2020	Completed
Data analyses (January – April 2021)	January – April 2021	Until now
Secondment Tissue Gnostics – algorithm training (April – July)	April – July 2021	Delayed
Data processing and implementing methods learned at TissueGnostics	July 2021 – October 2021	Delayed
Collecting samples of patients (September 2021)	September 2021	No change



Step	Current	Proposed
Testing new available therapy	September - December 2021	No change
Data analysis and publishing papers	2022	No change

#### Initial Conclusions

concerning further counter-measures or business viability [possibly tentative]

14. ...

15. ...

#### Compliance Checks required:

Tick	Requirement	Notes [replace guide text with response]
<input type="checkbox"/>	Does the project involve processing 'personal data' of any sort?	Yes – gene expression, genetic sequences, clinical data (age, sex, symptoms)
<input type="checkbox"/>	Does the project involve processing 'confidential data' of any sort?	Yes – disease diagnosis and location of patients as it is a rare disease
<b>Data Availability requirements</b>		
<input type="checkbox"/>	Does data need to be held for GCP compliance?	Yes
<input type="checkbox"/>	Does data need to be held to meet 'Open Data' requirements?	No
<input type="checkbox"/>	Does data need to be held to meet ICMJE requirements or commitments?	Don't know



GDPR Compliance Checklist – where ‘personal data’ is processed:

Tick	Requirement	Notes [replace guide text with response]
<b>Article 5: Principles compliance checks</b>		
<input type="checkbox"/>	qq) Is processing lawful, fair, and transparent?	Yes
<input type="checkbox"/>	rr) Is the purpose (or purposes) of the processing clearly defined	Yes
<input type="checkbox"/>	ss) adequate, relevant and limited to what is necessary	yes
<input type="checkbox"/>	tt) accurate and, where necessary, kept up to date	yes
<input type="checkbox"/>	uu) kept and permits identification of data subjects for no longer than is necessary	not sure
<input type="checkbox"/>	vv) processed securely	yes
<input type="checkbox"/>	9) can you demonstrate this compliance?	No
<b>Articles 13 &amp; 14 compliance</b>		[See detailed Transparency Checklist below]
<input type="checkbox"/>	Did the data come from publicly accessible sources?	no
<input type="checkbox"/>	Are data subjects informed before processing starts for any new purpose if incompatible with original purpose where the controller wants to use data for a different purpose to the purpose for which they currently hold data	Not sure but guess they are. I am controller of one part of the data, contacting controllers from the first steps of the study is needed to clarify this.
<input type="checkbox"/>	Does the Privacy Notice and/or PIL cover this processing?	Not sure – the same explanation as for previous step.



Tick	Requirement	Notes [replace guide text with response]
<input type="checkbox"/>	What patient choices are available? Are these explained?	The following should be included:  Right to rectification  Right to be forgotten  Right to restriction of processing  Right to data portability  Right to object to automated decision-making, including individual decision-making and profiling
<b>Articles 6 and 9: legal bases</b>		
<input type="checkbox"/>	What are legal bases under Article 6	Research of public interest
<input type="checkbox"/>	What are legal bases under Article 9 (if 'special category' data)	Yes – genetic data (gene expression, genetic sequence)
<input type="checkbox"/>	Are Article 6 legitimate interests explained where relevant?	Not sure
<input type="checkbox"/>	Are details of statutory obligations for Article 6 explained where relevant.	Not sure
<input type="checkbox"/>	Is this proposed processing compatible with the declared purposes?	Not sure
<b>Article 89(1) research exemption</b>		
<input type="checkbox"/>	If for research, do we meet Art 89(1) data minimisation	yes
<b>Articles 15-23: Data Subject Rights</b>		[See detailed table below]
<input type="checkbox"/>	Do we support data subject rights?	Data is pseudo-/anonymised
<input type="checkbox"/>	There is no use of automated decision making (e.g. profiling)	No



Tick	Requirement	Notes [replace guide text with response]
<b>Articles 24-43: Controller-Processor</b>		
<input type="checkbox"/>	A28 & 29: What measures are there to ensure processors comply?	[Is there a formal Data Processing Agreement] – not sure
<input type="checkbox"/>	A30: Is there an entry for this processing/data held in the register?	Yes
<input type="checkbox"/>	A32-34: Do we ensure appropriate security, including protection against unauthorised or unlawful processing and against accidental loss, destruction or damage, using appropriate technical or organisational measures?	Not sure
<input type="checkbox"/>	A37-39: Is there a DPO and have they been or will they be consulted?	[part of sign-off of the DPIA] not sure
<b>Articles 44-50: International transfers</b>		
	What form of data will be transferred to a third country or international organisation	No data will be transferred to third country
<input type="checkbox"/>	Are there safeguards for international transfers?	anonymisation
<b>Article 90: Obligations of secrecy</b>		
<input type="checkbox"/>	Do we meet medical confidentiality requirements?	Yes

#### Data Subject Rights:

Note if supported and what process/procedure applies; if not, then describe the legal justification for not supporting this right.

<input type="checkbox"/>	To be informed: about processing, about choices, about rights, about controller	Not sure since I am not controlling this part of data
<input type="checkbox"/>	the right of access to see or receive a printed copy	



<input type="checkbox"/>	the right to rectification – to correct any material errors in the personal data	
<input type="checkbox"/>	the right to erasure – where appropriate, to ask that all personal data is erased	
<input type="checkbox"/>	the right to restrict processing – to ask that some or all processing ceases [see opt-out]	
<input type="checkbox"/>	the right to data portability – this only applies to data provided directly by individual	
<input type="checkbox"/>	the right to object to and not to be subject to automated decision-making, including profiling	
<input type="checkbox"/>	Right to object to a Data Processing Authority (typically the relevant supervisory authority of each Member State)	
<input type="checkbox"/>	Where consent is the legal basis, the right to withdraw consent	



Detailed Transparency Checklist<sup>31</sup>

Does privacy information provided to data subjects include:

<input type="checkbox"/>	The name and contact details of our organisation	Not sure since I am not controlling this part of data
<input type="checkbox"/>	The name and contact details of our representative (if applicable)	
<input type="checkbox"/>	The contact details of our data protection officer (if applicable)	
<input type="checkbox"/>	The purposes of the processing	
<input type="checkbox"/>	The lawful bases for the processing	[Art6 for 'personal data' & Art9 for 'special category']
<input type="checkbox"/>	The legitimate interests for the processing (if applicable)	
<input type="checkbox"/>	The categories of personal data obtained (if the personal data is not obtained from the individual it relates to)	[for Art14]
<input type="checkbox"/>	The recipients or categories of recipients of the personal data	
<input type="checkbox"/>	The details of transfers of the personal data to any third countries or international organisations (if applicable)	
<input type="checkbox"/>	The retention periods for the personal data.	
<input type="checkbox"/>	The rights available to individuals in respect of the processing	
<input type="checkbox"/>	The right to withdraw consent (if applicable)	
<input type="checkbox"/>	The right to lodge a complaint with a supervisory authority	

<sup>31</sup> Taken from UK Information Commissioner's Office template



<input type="checkbox"/>	The source of the personal data (if the personal data is not obtained from the individual it relates to)	[For Art14]
<input type="checkbox"/>	The details of whether individuals are under a statutory or contractual obligation to provide the personal data  (if applicable, and if the personal data is collected from the individual it relates to)	
<input type="checkbox"/>	The details of the existence of automated decision-making, including profiling (if applicable)	
<input type="checkbox"/>	We provide individuals with privacy information at the time we collect their personal data from them – or where we obtain personal data from a source other than the individual it relates to, we provide them with privacy information	
<input type="checkbox"/>	within a reasonable of period of obtaining the personal data and no later than one month	
<input type="checkbox"/>	if we plan to communicate with the individual, at the latest, when the first communication takes place	
<input type="checkbox"/>	if we plan to disclose the data to someone else, at the latest, when the data is disclosed	
<input type="checkbox"/>	We provide the information in a way that is:  <input type="checkbox"/> concise;  <input type="checkbox"/> transparent;  <input type="checkbox"/> intelligible;  <input type="checkbox"/> easily accessible; and  <input type="checkbox"/> uses clear and plain language.	[Describe how we check is Plain English, etc.]
<input type="checkbox"/>	When drafting the information, we:	[Note: best practice advice]





	<input type="checkbox"/> undertake an information audit to find out what personal data we hold and what we do with it.  <input type="checkbox"/> put ourselves in the position of the people we're collecting information about.  <input type="checkbox"/> carry out user testing to evaluate how effective our privacy information is	
<input type="checkbox"/>	<p>When providing our privacy information to individuals, we use a combination of appropriate techniques, such as:</p> <ul style="list-style-type: none"> <li><input type="checkbox"/> a layered approach;</li> <li><input type="checkbox"/> dashboards;</li> <li><input type="checkbox"/> just-in-time notices;</li> <li><input type="checkbox"/> icons; and</li> <li><input type="checkbox"/> mobile and smart device functionalities.</li> </ul>	<p>[Note: best practice advice]</p>



### Security & Access Control Checklist

Controls need to be appropriate to level of risk: identified special category data needs more protection against potential misuse than non-personal data.

	Data Security classification (above Official)	<input type="checkbox"/> - Official-Sensitive <input type="checkbox"/> - Secret <input type="checkbox"/> - Top Secret <input type="checkbox"/> - Public Domain
<input type="checkbox"/>	Personal Data involved [GDPR]	
<input type="checkbox"/>	Special Category of personal data involved [GDPR]	Data from genetic analysis
<input type="checkbox"/>	Electronic Communications (inc. cookies) [PECR]	
<input type="checkbox"/>	Credit Card data	No data
<input type="checkbox"/>	Legal enforcement [LED2018]	
<input type="checkbox"/>	Financial data	No data
<input type="checkbox"/>	Intellectual Property (detail owner)	No data
<input type="checkbox"/>	Commercial in confidence (detail owner)	No data
	Data Location (storage or processing)  (include any back-up site(s))	<input type="checkbox"/> - UK <input checked="" type="checkbox"/> - EU/EEA <input type="checkbox"/> - EU White-list <input type="checkbox"/> - USA <input type="checkbox"/> - Other:
<input type="checkbox"/>	Is data held in secure data centre?	Digital platform
<input type="checkbox"/>	Is this new supplier, location, or system?	[If so, need specific IS check; also need formal contract]
<input type="checkbox"/>	Is all user access subject to 2-factor authentication?	<input type="checkbox"/> - no control <input checked="" type="checkbox"/> - single factor (e.g. just password)



		<input type="checkbox"/> - 2-factor (e.g. password & fob) <input type="checkbox"/> - biometric [note: GDPR reqs] <input type="checkbox"/> - Other control:
<input type="checkbox"/>	Are there established JML procedures?	[Joiners, Movers, Leavers]
<input type="checkbox"/>	Are there checks that passwords are robust and secure enough?	NO
<input type="checkbox"/>	Are all administrator & user accounts routinely monitored?	YES
<input type="checkbox"/>	Are systems protected against malware and other attacks?	YES but don't know which

[Need some aspect of CIA/impact-likelihood assessment]

#### Information Asset Register Checklist

<input type="checkbox"/>	Are there new IAs being created?	[provide details]
<input type="checkbox"/>	Are old IAs being retired?	[provide details]
<input type="checkbox"/>	Have IAOs & IACs been consulted?	
<input type="checkbox"/>	Has IAR been updated/amended?	[at least create project task to do so]
<input type="checkbox"/>	Data Retention classification & period	
<input type="checkbox"/>	Data retention procedure/functionality in place	



## Appendix A – Supervisory Authority ‘High Risk’ Check

If the DPIA shows ‘high risk’ processing which cannot be mitigated, then the DPIA should be sent to the relevant authority for review before any processing starts. Note that their review may take several weeks to process. A ‘High Risk’ assessment represents a ‘risk to the rights and freedoms of individuals’ – so may extend beyond GDPR consideration, including Human Rights.

GDPR Article 35(3) provides three examples:

- v) a systematic and extensive evaluation of personal aspects relating to natural persons which is based on automated processing, including profiling, and on which decisions are based that produce legal effects concerning the natural person or similarly significantly affect the natural person;
- w) processing on a large scale of special categories of data referred to in Article 9(1), or of personal data relating to criminal convictions and offences referred to in Article 1013; or
- x) a systematic monitoring of a publicly accessible area on a large scale

ICO cites:

- 22. Systematic and extensive profiling with significant effects
- 23. Large scale use of sensitive data [viz. ‘special category’ in GDPR terms]
- 24. Public monitoring

These being the same as (a)-(c) above. They further identify:

- 71. **New technologies:** processing involving the use of new technologies, or the novel application of existing technologies (including AI).
- 72. **Denial of service:** Decisions about an individual’s access to a product, service, opportunity or benefit which is based to any extent on automated decision-making (including profiling) or involves the processing of special category data.
- 73. **Large-scale profiling:** any profiling of individuals on a large scale.
- 74. **Biometrics:** any processing of biometric data.
- 75. **Genetic data:** any processing of genetic data, other than that processed by an individual GP or health professional for the provision of health care direct to the data subject.
- 76. **Data matching:** combining, comparing or matching personal data obtained from multiple sources.
- 77. **Invisible processing:** processing of personal data that has not been obtained direct from the data subject in circumstances where the controller considers that compliance with Article 14 would prove impossible or involve disproportionate effort.
- 78. **Tracking:** processing which involves tracking an individual’s geolocation or behaviour, including but not limited to the online environment.
- 79. **Targeting of children or other vulnerable individuals:** The use of the personal data of children or other vulnerable individuals for marketing purposes, profiling or other automated decision-making, or if you intend to offer online services directly to children.



80. **Risk of physical harm:** Where the processing is of such a nature that a personal data breach could jeopardise the [physical] health or safety of individuals.

**'High Risk' assessment using ICO criteria:**

Criterion:	Assessment	Comments
New technologies	N/A	
Denial of service	N/A	
Large-scale profiling	N/A	
Biometrics	N/A	
Genetic data	N/A	
Data matching	N/A	
Invisible processing	N/A	
Tracking	N/A	
Targeting of children or other vulnerable individuals	N/A	



Criterion:	Assessment	Comments
Risk of physical harm	N/A	

[The assessment can be one of N/A (not applicable), Low, Medium, or High. The comments should explain how the assessment is justified.]

Appendix B – Broad Privacy Risk Assessment:

#	Risk Description/detail	Discussion
78.	Data accuracy and timeliness	[Is data accurately recorded & kept up-to-date?]
79.	Differential treatment of patients/data subjects	[Might certain categories of people be adversely affected, e.g. children, vulnerable adults]
80.	Data Accuracy and identification	[Is the identification of individual reliable? Is there a danger of mis-attribution or incorrect linkage of data?]
81.	Holding / sharing / use of excessive data within [Company] systems	[Might too much data be held or for long? Is there a clear justification for data retention? Not 'just in case']
82.	Data held too long within [Company] systems	[Is there a clear data retention period specified and are there processes to ensure its deletion when no longer needed? Are copies tracked and deleted as well?]
83.	Excessive range of access in terms of users to personal data (consider new users/change of access privileges)	[Do more users have access than strictly necessary? Are user roles clear distinguished and reflected in the access privileges? Is there a clear process for granting and revoking access privileges?]
84.	Potential for misuse of data, unauthorised access to systems	[What are the likely threats to the data? What countermeasures are or might be applied? Is it possible for access to be granted inappropriately?]
85.	New sharing of data with other organisations, including new or change of suppliers	[Is data being shared from new data providers or with new data users? Are there new suppliers or data processors? What controls will apply?]
86.	Variable and inconsistent adoption / implementation	[How well will this system work end-to-end? How robust is it against partial adoption or system failure?]



#	Risk Description/detail	Discussion
87.	Legal compliance, particularly DP transparency requirements and support for data subject rights	[How well does this system meet legal requirements – or appear to meet legal requirements? Does it meet the ‘No surprises’ rule? What would happen if an individual requests data erasure or ceasing processing, etc.]
88.	Medical confidentiality	[Are there any addition sensitivities over confidentiality? Might specific approval (e.g. REC) be required to support this processing?]







## IG Assessment Checklist ESR11 – Exosomes as biomarkers in ANCA-associated vasculitis

### Introduction to IG Assessment process

Under the General Data Protection Regulation (GDPR), a Data Protection Impact Assessment (DPIA) is only required where proposed data processing is “likely to result in a high risk to the rights and freedoms of natural persons” (Article 35(1)). However, Article 35(3) explicitly requires one where there is ‘large-scale’ processing of ‘special category’ (e.g. healthcare) data then a DPIA is required.

One other possibility is that the data being processed is already anonymised (see Recital 26) so falls outside GDPR altogether so that no DPIA is actually required.

However, good project management and information governance suggests that there should be a general approach to risk assessment for any project or business enterprise – if only to determine whether a DPIA might be required.

Ideally, one should work from a simple initial Checklist (this document) which identifies possible areas of information risk and compliance requirements to a ‘discussion note’ which explores any issues in more depth and may help identify the necessary mitigation methods and mechanisms to offset most if not all risks. Only if risks are unmitigated or remain ‘high’ would you move to a formal DPIA report.

### The IG Assessment approach

There should be an overview of the proposed project or business change to explain what processing is envisaged as well as the purpose and intended outcome. The ‘purpose’ is important to establish the legal basis for the processing as well as ensuring that any possible mitigations or counter-measures do not undermine the main rationale for the processing.

The next step is to establish what compliance requirements may apply: GDPR, contractual or other regulatory restrictions, consent requirements, or obligations to preserve the data for legal or other reasons (including the benefit of posterity perhaps).

Once the precise range of obligations has been established, then appropriate checks can be made and recorded within the document.

The most obvious of these being GDPR compliance. There must be a ‘High Risk’ assessment (Appendix A) to determine whether the supervisory authority needs to be informed – generally, it is expected that it will not be necessary; if so, then a formal DPIA report will be needed.

Appendix B has a broader Privacy Impact Assessment that may throw up some broader issues.

Initial conclusions as to next steps or particular countermeasures to be considered should be detailed below.



### Project Background/Overview

Due to the high novel interest in exomes role, their importance in cell to cell communication and the problematic diagnostic process of AAV patients it is essential to analyse whether exosomes can be used as biomarkers of ANCA-associated vasculitis in order to determine the quiescent and active stage of this disease.

It is expected that the exosomes of patients with active stage of ANCA-associated vasculitis (AAV) express different protein pattern on the surface due to the pathological changes during occurrence. Profiling of these proteins in healthy individuals and patients with AVV in well characterized different stages will determine the differences in protein expression and can indicate a new potential biomarker for clinical appliance.

### Comparison of process steps (simplified): [optional]

This allows identification of what processing is new or changed through the project:

Step	Current	Proposed
Method optimisation and characterisation for exosome isolation	In process	
Verification whether microarray is suitable for exosome profiling	Secondment in KTH	
Determining the linkage of between exosomes and their cells by surface protein profiling in various cell lines	Secondment in KTH	
Sample recruitment from patients and healthy control from AKH Biobank	Ethical approval required	
Isolation of exosomes from serum or plasma		



Step	Current	Proposed
Protein profiling on exosomes derived from serum/ plasma	Secondment in KTH Data Transfer Agreement?	
Determination whether the exosomes can be grouped accordingly to their cell origin		
Finding the signatures on exosome surface in patients with multiply flares		
Verification whether this signature can be proposed as a new biomarker with Validation cohort from RKD biobank	MTA with RKD required	
New biomarker validation	Secondment in Firalis Data Transfer Agreement	

#### Initial Conclusions

concerning further countermeasures or business viability [possibly tentative]

16. NA

17. ...

#### Compliance Checks required:

Tick	Requirement	Notes [replace guide text with response]
<input checked="" type="checkbox"/>	Does the project involve processing 'personal data' of any sort?	Patient clinical data



Tick	Requirement	Notes [replace guide text with response]
<input checked="" type="checkbox"/>	Does the project involve processing 'confidential data' of any sort?	<p>Note: may be 'commercial in confidence', medical confidentiality, or organisational confidentiality (internally sensitive); may need to check contractual limitations</p> <p>Patient's medical record</p>
<b>Data Availability requirements</b>		
<input checked="" type="checkbox"/>	Does data need to be held for GCP compliance?	<p>My research is not a Clinical Trial, thus I understand it is not required to be held by GCP (?)</p> <p>Yes, patients could be re-identified</p>
<input checked="" type="checkbox"/>	Does data need to be held to meet 'Open Data' requirements?	
<input checked="" type="checkbox"/>	Does data need to be held to meet ICMJE requirements or commitments?	



GDPR Compliance Checklist – where ‘personal data’ is processed:

Tick	Requirement	Notes [replace guide text with response]
<b>Article 5: Principles compliance checks</b>		
<input checked="" type="checkbox"/>	ww) Is processing lawful, fair, and transparent?	The research is not YET approved by the ethics committee.  The processing will be transparent: patients will be informed about the data transfer.
<input checked="" type="checkbox"/>	xx) Is the purpose (or purposes) of the processing clearly defined	The purpose will be clearly defined in the ethics application
<input checked="" type="checkbox"/>	yy) adequate, relevant and limited to what is necessary	The accessed patient clinical data will be limited to the information, which covers the ethics
<input checked="" type="checkbox"/>	zz) accurate and, where necessary, kept up to date	
<input checked="" type="checkbox"/>	aaa) kept and permits identification of data subjects for no longer than is necessary	Patient Clinical Data will be removed after they have been analysed and the project was finalised
<input checked="" type="checkbox"/>	bbb) processed securely	Data will be processed on a computer and stored on portable drive with the restricted access for only a project supervisor and a PhD student.
<input checked="" type="checkbox"/>	10) can you demonstrate this compliance?	
<b>Articles 13 &amp; 14 compliance</b>		[See detailed Transparency Checklist below]
<input checked="" type="checkbox"/>	Did the data come from publicly accessible sources?	Only part of them came from Human Protein Atlas, which is publicly accessible data.
<input checked="" type="checkbox"/>	Are data subjects informed before processing starts for any new purpose if incompatible with original purpose where the controller wants to use data for a different purpose to the purpose for which they currently hold data	



Tick	Requirement	Notes [replace guide text with response]
<input checked="" type="checkbox"/>	Does the Privacy Notice and/or PIL cover this processing?	Ethics consent contains PIL
<input checked="" type="checkbox"/>	What patient choices are available? Are these explained?	Depending on which samples, I am planning to use for my experiment: samples from the biobank (Patients are consented, and I suppose they are fully informed about their right and choices prior to it; or sample from Theresa, in which case patients are not consented).
<b>Articles 6 and 9: legal bases</b>		
<input checked="" type="checkbox"/>	What are legal bases under Article 6	Consents to process personal data and public interest
<input type="checkbox"/>	What are legal bases under Article 9 (if 'special category' data)	Article 9.2J; Performance of scientific research in a public interest
<input checked="" type="checkbox"/>	Are Article 6 legitimate interests explained where relevant?	
<input checked="" type="checkbox"/>	Are details of statutory obligations for Article 6 explained where relevant.	"the data subject has given consent to the processing of his or her personal data for one or more specific purposes;"
<input checked="" type="checkbox"/>	Is this proposed processing compatible with the declared purposes?	[Check against any privacy notices and public information]
<b>Article 89(1) research exemption</b>		



Tick	Requirement	Notes [replace guide text with response]
<input checked="" type="checkbox"/>	If for research, do we meet Art 89(1) data minimisation	<p>Art. 89 GDPR: “Safeguards and derogations relating to processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes”</p> <p>Processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes, shall be subject to appropriate safeguards, in accordance with this Regulation, for the rights and freedoms of the data subject.</p> <p>2Those safeguards shall ensure that technical and organisational measures are in place in particular in order to ensure respect for the principle of data minimisation.</p> <p>3Those measures may include pseudonymisation provided that those purposes can be fulfilled in that manner.</p> <p>4Where those purposes can be fulfilled by further processing which does not permit or no longer permits the identification of data subjects, those purposes shall be fulfilled in that manner.</p>
<b>Articles 15-23: Data Subject Rights</b>		[See detailed table below]
<input checked="" type="checkbox"/>	Do we support data subject rights?	Although our data about patient samples are pseudonymised, our informed consent contextualises that an individual who agrees can withdrawal from the study at anytime without giving an apparent reason.
<input checked="" type="checkbox"/>	There is no use of automated decision making (e.g. profiling)	No use
<b>Articles 24-43: Controller-Processor</b>		
<input type="checkbox"/>	A28 & 29: What measures are there to ensure processors comply?	There is not



Tick	Requirement	Notes [replace guide text with response]
<input type="checkbox"/>	A30: Is there an entry for this processing/data held in the register?	"Records of processing activities" There is not, should there be?
<input checked="" type="checkbox"/>	A32-34: Do we ensure appropriate security, including protection against unauthorised or unlawful processing and against accidental loss, destruction or damage, using appropriate technical or organisational measures?	We do ensure appropriate security.
<input type="checkbox"/>	A37-39: Is there a DPO and have they been or will they be consulted?	[part of sign-off of the DPIA]
<b>Articles 44-50: International transfers</b>		
	What form of data will be transferred to a third country or international organisation	[describe nature of data and whether identified, identifiable, de-identified or anonymous]  I will receive the pseudonymised data of patients from RKD Biobank in Dublin regarding their clinical stage. This data will be received directly from RKD; or through KTH Stockholm Shaghayegh Bayati ESR13, who is having the same study group involved in her project. Ms. Shaghayegh Bayati is going to receive the pseudonymised data from MedUniWien biobank.
<input checked="" type="checkbox"/>	Are there safeguards for international transfers?	[e.g. US Privacy Shield, anonymisation, GDPR equivalence, approved contractual clauses, or BCR]  Pseudonymisation
<b>Article 90: Obligations of secrecy</b>		
<input checked="" type="checkbox"/>	Do we meet medical confidentiality requirements?	Yes, the transferred data will be pseudonymised.





Data Subject Rights:

Note if supported and what process/procedure applies; if not, then describe the legal justification for not supporting this right.

<input checked="" type="checkbox"/>	To be informed: about processing, about choices, about rights, about controller	
<input checked="" type="checkbox"/>	the right of access to see or receive a printed copy	
<input checked="" type="checkbox"/>	the right to rectification – to correct any material errors in the personal data	
<input checked="" type="checkbox"/>	the right to erasure – where appropriate, to ask that all personal data is erased	
<input checked="" type="checkbox"/>	the right to restrict processing – to ask that some or all processing ceases [see opt-out]	
<input checked="" type="checkbox"/>	the right to data portability – this only applies to data provided directly by individual	
<input checked="" type="checkbox"/>	the right to object to and not to be subject to automated decision-making, including profiling	
<input checked="" type="checkbox"/>	Right to object to a Data Processing Authority (typically the relevant supervisory authority of each Member State)	
<input checked="" type="checkbox"/>	Where consent is the legal basis, the right to withdraw consent	



Detailed Transparency Checklist<sup>32</sup>

Does privacy information provided to data subjects include:

<input checked="" type="checkbox"/>	The name and contact details of our organisation	HELICAL ITN
<input checked="" type="checkbox"/>	The name and contact details of our representative (if applicable)	Prof. Mark Little
<input checked="" type="checkbox"/>	The contact details of our data protection officer (if applicable)	
<input checked="" type="checkbox"/>	The purposes of the processing	
<input checked="" type="checkbox"/>	The lawful bases for the processing	[Art6 for 'personal data' & Art9 for 'special category']
<input checked="" type="checkbox"/>	The legitimate interests for the processing (if applicable)	
<input checked="" type="checkbox"/>	The categories of personal data obtained (if the personal data is not obtained from the individual it relates to)	[for Art14]
<input checked="" type="checkbox"/>	The recipients or categories of recipients of the personal data	
<input checked="" type="checkbox"/>	The details of transfers of the personal data to any third countries or international organisations (if applicable)	
<input type="checkbox"/>	The retention periods for the personal data.	
<input checked="" type="checkbox"/>	The rights available to individuals in respect of the processing	
<input checked="" type="checkbox"/>	The right to withdraw consent (if applicable)	
<input checked="" type="checkbox"/>	The right to lodge a complaint with a supervisory authority	

<sup>32</sup> Taken from UK Information Commissioner's Office template



<input checked="" type="checkbox"/>	The source of the personal data (if the personal data is not obtained from the individual it relates to)	[For Art14]
<input checked="" type="checkbox"/>	The details of whether individuals are under a statutory or contractual obligation to provide the personal data  (if applicable, and if the personal data is collected from the individual it relates to)	
<input checked="" type="checkbox"/>	The details of the existence of automated decision-making, including profiling (if applicable)	
<input type="checkbox"/>	We provide individuals with privacy information at the time we collect their personal data from them – or where we obtain personal data from a source other than the individual it relates to, we provide them with privacy information	
<input type="checkbox"/>	within a reasonable of period of obtaining the personal data and no later than one month	
<input type="checkbox"/>	if we plan to communicate with the individual, at the latest, when the first communication takes place	
<input checked="" type="checkbox"/>	if we plan to disclose the data to someone else, at the latest, when the data is disclosed	
<input checked="" type="checkbox"/>	We provide the information in a way that is:  <input type="checkbox"/> concise;  <input type="checkbox"/> transparent;  <input type="checkbox"/> intelligible;  <input type="checkbox"/> easily accessible; and  <input type="checkbox"/> uses clear and plain language.	[Describe how we check is Plain English, etc.]  Plain English, Plain German, brochure user friendly
<input checked="" type="checkbox"/>	When drafting the information, we:	[Note: best practice advice]



	<input type="checkbox"/> undertake an information audit to find out what personal data we hold and what we do with it.  <input type="checkbox"/> put ourselves in the position of the people we're collecting information about.  <input type="checkbox"/> carry out user testing to evaluate how effective our privacy information is	
<input checked="" type="checkbox"/>	<p>When providing our privacy information to individuals, we use a combination of appropriate techniques, such as:</p> <input type="checkbox"/> a layered approach; <input type="checkbox"/> dashboards; <input type="checkbox"/> just-in-time notices; <input type="checkbox"/> icons; and <input type="checkbox"/> mobile and smart device functionalities.	<p>[Note: best practice advice]</p>



### Security & Access Control Checklist

Controls need to be appropriate to level of risk: identified special category data needs more protection against potential misuse than non-personal data.

	Data Security classification (above Official)	<input type="checkbox"/> - Official-Sensitive <input type="checkbox"/> - Secret <input type="checkbox"/> - Top Secret <input type="checkbox"/> - Public Domain
<input checked="" type="checkbox"/>	Personal Data involved [GDPR]	
<input type="checkbox"/>	Special Category of personal data involved [GDPR]	
<input type="checkbox"/>	Electronic Communications (inc. cookies) [PECR]	
<input type="checkbox"/>	Credit Card data	
<input type="checkbox"/>	Legal enforcement [LED2018]	
<input type="checkbox"/>	Financial data	
<input type="checkbox"/>	Intellectual Property (detail owner)	
<input type="checkbox"/>	Commercial in confidence (detail owner)	
	Data Location (storage or processing) (include any back-up site(s))	<input type="checkbox"/> - UK <input checked="" type="checkbox"/> - EU/EEA <input type="checkbox"/> - EU White-list <input type="checkbox"/> - USA <input type="checkbox"/> - Other:
<input type="checkbox"/>	Is data held in secure data centre?	[detail centre and what certification supports assertion]
<input type="checkbox"/>	Is this new supplier, location, or system?	[If so, need specific IS check; also need formal contract]
<input type="checkbox"/>	Is all user access subject to 2-factor authentication?	<input type="checkbox"/> - no control



		<input checked="" type="checkbox"/> - single factor (e.g. just password) <input type="checkbox"/> - 2-factor (e.g. password & fob) <input checked="" type="checkbox"/> - biometric [note: GDPR reqs] <input type="checkbox"/> - Other control:
<input type="checkbox"/>	Are there established JML procedures?	[Joiners, Movers, Leavers]
<input checked="" type="checkbox"/>	Are there checks that passwords are robust and secure enough?	Password is changed periodically and has the required length
<input type="checkbox"/>	Are all administrator & user accounts routinely monitored?	[Particularly for redundant or little used accounts]
<input checked="" type="checkbox"/>	Are systems protected against malware and other attacks?	Avast antivirus  Windows Security Defender

[Need some aspect of CIA/impact-likelihood assessment]

#### Information Asset Register Checklist

<input type="checkbox"/>	Are there new IAs being created?	[provide details]
<input type="checkbox"/>	Are old IAs being retired?	[provide details]
<input type="checkbox"/>	Have IAOs & IACs been consulted?	
<input type="checkbox"/>	Has IAR been updated/amended?	[at least create project task to do so]
<input type="checkbox"/>	Data Retention classification & period	
<input type="checkbox"/>	Data retention procedure/functionality in place	



## Appendix A – Supervisory Authority ‘High Risk’ Check

If the DPIA shows ‘high risk’ processing which cannot be mitigated, then the DPIA should be sent to the relevant authority for review before any processing starts. Note that their review may take several weeks to process. A ‘High Risk’ assessment represents a ‘risk to the rights and freedoms of individuals’ – so may extend beyond GDPR consideration, including Human Rights.

GDPR Article 35(3) provides three examples:

- y) a systematic and extensive evaluation of personal aspects relating to natural persons which is based on automated processing, including profiling, and on which decisions are based that produce legal effects concerning the natural person or similarly significantly affect the natural person;
- z) processing on a large scale of special categories of data referred to in Article 9(1), or of personal data relating to criminal convictions and offences referred to in Article 1013; or
- aa) a systematic monitoring of a publicly accessible area on a large scale

ICO cites:

- 25. Systematic and extensive profiling with significant effects
- 26. Large scale use of sensitive data [viz. ‘special category’ in GDPR terms]
- 27. Public monitoring

These being the same as (a)-(c) above. They further identify:

- 81. **New technologies:** processing involving the use of new technologies, or the novel application of existing technologies (including AI).
- 82. **Denial of service:** Decisions about an individual’s access to a product, service, opportunity or benefit which is based to any extent on automated decision-making (including profiling) or involves the processing of special category data.
- 83. **Large-scale profiling:** any profiling of individuals on a large scale.
- 84. **Biometrics:** any processing of biometric data.
- 85. **Genetic data:** any processing of genetic data, other than that processed by an individual GP or health professional for the provision of health care direct to the data subject.
- 86. **Data matching:** combining, comparing or matching personal data obtained from multiple sources.
- 87. **Invisible processing:** processing of personal data that has not been obtained direct from the data subject in circumstances where the controller considers that compliance with Article 14 would prove impossible or involve disproportionate effort.
- 88. **Tracking:** processing which involves tracking an individual’s geolocation or behaviour, including but not limited to the online environment.
- 89. **Targeting of children or other vulnerable individuals:** The use of the personal data of children or other vulnerable individuals for marketing purposes, profiling or other automated decision-making, or if you intend to offer online services directly to children.



90. **Risk of physical harm:** Where the processing is of such a nature that a personal data breach could jeopardise the [physical] health or safety of individuals.

**'High Risk' assessment using ICO criteria:**

Criterion:	Assessment	Comments
New technologies	Low	Biomarkers have been used in clinical practice for many years.
Denial of service	NA	
Large-scale profiling	Medium	Exosomes derived from the patient serum will be profiled by the chosen antibody set. This will be limited to a study group and number of antibodies.
Biometrics	NA	
Genetic data	NA	
Data matching	Low	Data may be matched with the other data set generated from the same patient. Exosome profiles of an individual will be matched with his autoantibody Repertoire.
Invisible processing	NA	
Tracking	NA	
Targeting of children or other vulnerable individuals	NA	





Criterion:	Assessment	Comments
Risk of physical harm	NA	

[The assessment can be one of N/A (not applicable), Low, Medium, or High. The comments should explain how the assessment is justified.]

Appendix B – Broad Privacy Risk Assessment:

#	Risk Description/detail	Discussion
89.	Data accuracy and timeliness	Data will be accurately recorded and updated.
90.	Differential treatment of patients/data subjects	All patients are treated the same.
91.	Data Accuracy and identification	Data will be double checked each time they are in use.
92.	Holding / sharing / use of excessive data within [Company] systems	The data will be held until the project is finalised.
93.	Data held too long within [Company] systems	The period is not yet specified. Ideally, it will be held until the last date of a project. The copies will be tracked and deleted.
94.	Excessive range of access in terms of users to personal data (consider new users/change of access privileges)	Each person, that want to get access to this data must be confirmed by the supervisor team of a project.
95.	Potential for misuse of data, unauthorised access to systems	Data will be stored on a laptop, which is in a risk of being stolen. This laptop has biometric protection (fingerprint).
96.	New sharing of data with other organisations, including new or change of suppliers	Data may or may not be shared with additional organisations, if so password control will be applied with pre authentication of person with granted access



#	Risk Description/detail	Discussion
97.	Variable and inconsistent adoption / implementation	<p>[How well will this system work end-to-end? How robust is it against partial adoption or system failure?]</p> <p>The system failure is considered, thus the additional USB drive with data under the password in kept with the access available only for the PhD student.</p>
98.	Legal compliance, particularly DP transparency requirements and support for data subject rights	<p>[How well does this system meet legal requirements – or appear to meet legal requirements? Does it meet the ‘No surprises’ rule? What would happen if an individual requests data erasure or ceasing processing, etc.]</p>
99.	Medical confidentiality	<p>Medical records data is confidential and sensitive, thus the data is encoded.</p>



## IG Assessment Checklist ESR12 – Computer assisted morphometry of pathological changes in renal biopsies from patients with AAV

### Introduction to IG Assessment process

Under the General Data Protection Regulation (GDPR), a Data Protection Impact Assessment (DPIA) is only required where proposed data processing is “likely to result in a high risk to the rights and freedoms of natural persons” (Article 35(1)). However, Article 35(3) explicitly requires one where there is ‘large-scale’ processing of ‘special category’ (e.g. healthcare) data then a DPIA is required.

One other possibility is that the data being processed is already anonymised (see Recital 26) so falls outside GDPR altogether so that no DPIA is actually required

However, good project management and information governance suggests that there should be a general approach to risk assessment for any project or business enterprise – if only to determine whether a DPIA might be required.

Ideally, one should work from a simple initial Checklist (this document) which identifies possible areas of information risk and compliance requirements to a ‘discussion note’ which explores any issues in more depth and may help identify the necessary mitigation methods and mechanisms to offset most if not all risks. Only if risks are unmitigated or remain ‘high’ would you move to a formal DPIA report.

### The IG Assessment approach

There should be an overview of the proposed project or business change to explain what processing is envisaged as well as the purpose and intended outcome. The ‘purpose’ is important to establish the legal basis for the processing as well as ensuring that any possible mitigations or counter-measures do not undermine the main rationale for the processing.

The next step is to establish what compliance requirements may apply: GDPR, contractual or other regulatory restrictions, consent requirements, or obligations to preserve the data for legal or other reasons (including the benefit of posterity perhaps).

Once the precise range of obligations has been established, then appropriate checks can be made and recorded within the document.

The most obvious of these being GDPR compliance. There must be a ‘High Risk’ assessment (Appendix A) to determine whether the supervisory authority needs to be informed – generally, it is expected that it will not be necessary; if so, then a formal DPIA report will be needed.

Appendix B has a broader Privacy Impact Assessment that may throw up some broader issues.

Initial conclusions as to next steps or particular countermeasures to be considered should be detailed below.



**Project Background/Overview**

This research project is focused on developing machine learning and deep learning algorithms as support to other clinical tools that assist clinicians in diagnosing active vasculitis and predicting outcome. One of the aims of this research project is to choose and develop the best possible machine learning and/or deep learning techniques which will help identifying the biological structures and their relevant changes, to be used for diagnosis and prognosis in ANCA-associated vasculitis. Adopting this alternative strategy, in-house developed datasets consisting of medical images will be used to define tissue morphological changes (descriptors) that can be used as predictors of outcome in renal ANCA vasculitis. "Medical images" here means WSI's (Whole Slide Images) of the kidney tissue. Based on existing descriptors and algorithms, this project aims at defining morphological changes in renal biopsies from patients with ANCA vasculitis that are suited to automated morphometric analysis and subsequent validation using existing clinical outcome data.

To make most of AI, transfer learning will be used, which means publicly available pretrained deep learning models will be used to improve accuracy of deep learning algorithms. Deep learning models will be validated and augmented to give best possible results for given tasks, thus including also the use of publicly available datasets to augment the training dataset to be fed to the models.

The main goal of the first stage of the project is to develop a deep learning algorithm which will be able to segment glomeruli within WSI's, i.e. glomeruli will be distinguished from the rest of the kidney tissue.

For the next stage deep learning will be used to identify (segment) other relevant structures for the diagnosis and/or prognosis of ANCA-associated vasculitis Deep learning models will be used also to classify the identified structures as being healthy or unhealthy. Similar classes/subclasses might be also used to classify with a higher granularity the state of the structures.

**Comparison of process steps (simplified): [optional]**

**This allows identification of what processing is new or changed through the project:**

Step	Current	Proposed
Labelling/annotating images needed for development of deep learning algorithms	In process	
Development of deep learning algorithms for image segmentation	In process	
Validation of existing methods for image segmentation	In process	



Step	Current	Proposed

### Initial Conclusions

concerning further counter-measures or business viability [possibly tentative]

1. NA

### Compliance Checks required:

Tick	Requirement	Notes [replace guide text with response]
<input checked="" type="checkbox"/>	Does the project involve processing 'personal data' of any sort?	Histopathological images of patient's kidney tissue.
<input type="checkbox"/>	Does the project involve processing 'confidential data' of any sort?	
<b>Data Availability requirements</b>		
<input checked="" type="checkbox"/>	Does data need to be held for GCP compliance?	Yes, patients could be re-identified.
<input checked="" type="checkbox"/>	Does data need to be held to meet 'Open Data' requirements?	Yes. Results of the project could be published.
<input checked="" type="checkbox"/>	Does data need to be held to meet ICMJE requirements or commitments?	Yes. Good publication practice - reviewers might require access to the data.



GDPR Compliance Checklist – where ‘personal data’ is processed:

Tick	Requirement	Notes [replace guide text with response]
<b>Article 5: Principles compliance checks</b>		
<input checked="" type="checkbox"/>	a) Is processing lawful, fair, and transparent?	Yes. Process is confirmed by data protection committee of Medical University of Vienna.
<input checked="" type="checkbox"/>	b) Is the purpose (or purposes) of the processing clearly defined	Yes, it is defined in ethics application.
<input checked="" type="checkbox"/>	c) adequate, relevant and limited to what is necessary	Only images will be used, without any personal patient data.
<input type="checkbox"/>	d) accurate and, where necessary, kept up to date	Data (images) are accurate, there is no need to keep the data up to date.
<input type="checkbox"/>	e) kept and permits identification of data subjects for no longer than is necessary	Data is archived in Medical University of Vienna, Department of Pathology.
<input checked="" type="checkbox"/>	f) processed securely	The reference table is only kept in a safe place at the Medical University of Vienna. Data protection is ensured through the pseudonymization of patients.
<input checked="" type="checkbox"/>	2) can you demonstrate this compliance?	Can't be demonstrated.
<b>Articles 13 &amp; 14 compliance</b>		[See detailed Transparency Checklist below]
<input checked="" type="checkbox"/>	Did the data come from publicly accessible sources?	Some of the data used in development of deep learning algorithms will come from publicly accessible sources.
<input checked="" type="checkbox"/>	Are data subjects informed before processing starts for any new purpose if incompatible with original purpose where the controller wants to use data for a different purpose to the purpose for which they currently hold data	As it is already stated in ethics application which received positive vote from ethics committee of Medical University of Vienna: It is not intended to obtain patient consent. This study is done on images of archived tissue samples from patients which, probably to a significant extent, have already died. Recruitment procedures based on informed consent are therefore impractical.
<input type="checkbox"/>	Does the Privacy Notice and/or PIL cover this processing?	
<input checked="" type="checkbox"/>	What patient choices are available? Are these explained?	As it is already stated in ethics application which received positive vote from ethics committee of Medical University of Vienna: It is not practical to ask patients for consent.
<b>Articles 6 and 9: legal bases</b>		
<input checked="" type="checkbox"/>	What are legal bases under Article 6	Processing is necessary for the performance of a task carried out in the public interest - scientific research



Tick	Requirement	Notes [replace guide text with response]
<input checked="" type="checkbox"/>	What are legal bases under Article 9 (if 'special category' data)	Publication the research is in the public interest
<input checked="" type="checkbox"/>	Are Article 6 legitimate interests explained where relevant?	Explained in ethics application.
<input checked="" type="checkbox"/>	Are details of statutory obligations for Article 6 explained where relevant.	[Quote statutes or regulation] Explained in ethics application.
<input type="checkbox"/>	Is this proposed processing compatible with the declared purposes?	[Check against any privacy notices and public information]
<b>Article 89(1) research exemption</b>		
<input type="checkbox"/>	If for research, do we meet Art 89(1) data minimisation	
<b>Articles 15-23: Data Subject Rights</b>		[See detailed table below]
<input checked="" type="checkbox"/>	Do we support data subject rights?	The data (histological images) will be pseudonymized before it reaches researchers computer.
<input checked="" type="checkbox"/>	There is no use of automated decision making (e.g. profiling)	[Otherwise need at least a 'discussion note']
<b>Articles 24-43: Controller-Processor</b>		
<input checked="" type="checkbox"/>	A28 & 29: What measures are there to ensure processors comply?	Project already have positive vote from ethics committee and data protection committee of Medical University of Vienna.
<input type="checkbox"/>	A30: Is there an entry for this processing/data held in the register?	Does university has data register?
<input checked="" type="checkbox"/>	A32-34: Do we ensure appropriate security, including protection against unauthorised or unlawful processing and against accidental loss, destruction or damage, using appropriate technical or organisational measures?	The data (histological images) will be pseudonymized before it reach researchers computer. The reference table is only kept in a safe place at the Medical University of Vienna.
<input type="checkbox"/>	A37-39: Is there a DPO and have they been or will they be consulted?	[part of sign-off of the DPIA]
<b>Articles 44-50: International transfers</b>		
<input type="checkbox"/>	What form of data will be transferred to a third country or international organisation	[describe nature of data and whether identified, identifiable, de-identified or anonymous]
<input type="checkbox"/>	Are there safeguards for international transfers?	[e.g. US Privacy Shield, anonymisation, GDPR equivalence, approved contractual clauses, or BCR]
<b>Article 90: Obligations of secrecy</b>		
<input checked="" type="checkbox"/>	Do we meet medical confidentiality requirements?	Yes, the transferred data will be pseudonymized.



Data Subject Rights:

Note if supported and what process/procedure applies; if not, then describe the legal justification for not supporting this right.

<input type="checkbox"/>	To be informed: about processing, about choices, about rights, about controller	As it is already stated in ethics application, it is not practical to inform patients.
<input checked="" type="checkbox"/>	the right of access to see or receive a printed copy	This is clinical use and regulated by the hospital's documentation office. Patients can access their data.
<input checked="" type="checkbox"/>	the right to rectification – to correct any material errors in the personal data	If there are errors, they will be corrected.
<input checked="" type="checkbox"/>	the right to erasure – where appropriate, to ask that all personal data is erased	Yes.
<input checked="" type="checkbox"/>	the right to restrict processing – to ask that some or all processing ceases [see opt-out]	Yes.
<input type="checkbox"/>	the right to data portability – this only applies to data provided directly by individual	NA
<input checked="" type="checkbox"/>	the right to object to and not to be subject to automated decision-making, including profiling	Yes, with restrictions. This is a research project. A program that provides automated decision support on diagnosis, can subsequently be CE certified and used to make a diagnosis. The patient cannot object to how we make it. We do not do profiling.
<input type="checkbox"/>	Right to object to a Data Processing Authority (typically the relevant supervisory authority of each Member State)	This is a government decision or decided at university level. Anyone can take it up with either.
<input checked="" type="checkbox"/>	Where consent is the legal basis, the right to withdraw consent	Yes, withdrawing is always possible.





### Detailed Transparency Checklist<sup>33</sup>

Does privacy information provided to data subjects include:

<input type="checkbox"/>	The name and contact details of our organisation	
<input type="checkbox"/>	The name and contact details of our representative (if applicable)	
<input type="checkbox"/>	The contact details of our data protection officer (if applicable)	
<input type="checkbox"/>	The purposes of the processing	
<input type="checkbox"/>	The lawful bases for the processing	[Art6 for 'personal data' & Art9 for 'special category']
<input type="checkbox"/>	The legitimate interests for the processing (if applicable)	
<input type="checkbox"/>	The categories of personal data obtained (if the personal data is not obtained from the individual it relates to)	[for Art14]
<input type="checkbox"/>	The recipients or categories of recipients of the personal data	
<input type="checkbox"/>	The details of transfers of the personal data to any third countries or international organisations (if applicable)	
<input type="checkbox"/>	The retention periods for the personal data.	
<input type="checkbox"/>	The rights available to individuals in respect of the processing	
<input type="checkbox"/>	The right to withdraw consent (if applicable)	
<input type="checkbox"/>	The right to lodge a complaint with a supervisory authority	
<input type="checkbox"/>	The source of the personal data (if the personal data is not obtained from the individual it relates to)	[For Art14]
<input type="checkbox"/>	The details of whether individuals are under a statutory or contractual obligation to provide the personal data (if applicable, and if the personal data is collected from the individual it relates to)	
<input type="checkbox"/>	The details of the existence of automated decision-making, including profiling (if applicable)	
<input type="checkbox"/>	We provide individuals with privacy information at the time we collect their personal data from them – or where we obtain personal data from a source other than the individual it relates to, we provide them with privacy information	



<input type="checkbox"/>	within a reasonable of period of obtaining the personal data and no later than one month	
<input type="checkbox"/>	if we plan to communicate with the individual, at the latest, when the first communication takes place	
<input type="checkbox"/>	if we plan to disclose the data to someone else, at the latest, when the data is disclosed	
<input type="checkbox"/>	We provide the information in a way that is: <input type="checkbox"/> concise; <input type="checkbox"/> transparent; <input type="checkbox"/> intelligible; <input type="checkbox"/> easily accessible; and <input type="checkbox"/> uses clear and plain language.	[Describe how we check is Plain English, etc.]
<input type="checkbox"/>	When drafting the information, we: <input type="checkbox"/> undertake an information audit to find out what personal data we hold and what we do with it. <input type="checkbox"/> put ourselves in the position of the people we're collecting information about. <input type="checkbox"/> carry out user testing to evaluate how effective our privacy information is	[Note: best practice advice]
<input type="checkbox"/>	When providing our privacy information to individuals, we use a combination of appropriate techniques, such as: <input type="checkbox"/> a layered approach; <input type="checkbox"/> dashboards; <input type="checkbox"/> just-in-time notices; <input type="checkbox"/> icons; and <input type="checkbox"/> mobile and smart device functionalities.	[Note: best practice advice]



### Security & Access Control Checklist

Controls need to be appropriate to level of risk: identified special category data needs more protection against potential misuse than non-personal data.

	Data Security classification (above Official)	<input type="checkbox"/> - Official-Sensitive <input checked="" type="checkbox"/> - Secret <input type="checkbox"/> - Top Secret <input type="checkbox"/> - Public Domain
<input checked="" type="checkbox"/>	Personal Data involved [GDPR]	Yes.
<input checked="" type="checkbox"/>	Special Category of personal data involved [GDPR]	Yes.
<input type="checkbox"/>	Electronic Communications (inc. cookies) [PECR]	NA
<input type="checkbox"/>	Credit Card data	NA
<input type="checkbox"/>	Legal enforcement [LED2018]	NA
<input type="checkbox"/>	Financial data	NA
<input type="checkbox"/>	Intellectual Property (detail owner)	NA
<input type="checkbox"/>	Commercial in confidence (detail owner)	NA
	Data Location (storage or processing) (include any back-up site(s))	<input type="checkbox"/> - UK <input checked="" type="checkbox"/> - EU/EEA <input type="checkbox"/> - EU White-list <input type="checkbox"/> - USA <input type="checkbox"/> - Other:
<input checked="" type="checkbox"/>	Is data held in secure data centre?	Held at the Medical University of Vienna
<input type="checkbox"/>	Is this new supplier, location, or system?	[If so, need specific IS check; also need formal contract]
<input type="checkbox"/>	Is all user access subject to 2-factor authentication?	<input type="checkbox"/> - no control <input type="checkbox"/> - single factor (e.g. just password) <input type="checkbox"/> - 2-factor (e.g. password & fob) <input type="checkbox"/> - biometric [note: GDPR reqs] <input type="checkbox"/> - Other control:
<input type="checkbox"/>	Are there established JML procedures?	[Joiners, Movers, Leavers]
<input type="checkbox"/>	Are there checks that passwords are robust and secure enough?	[ ]
<input type="checkbox"/>	Are all administrator & user accounts routinely monitored?	[Particularly for redundant or little used accounts]
<input type="checkbox"/>	Are systems protected against malware and other attacks?	[provide details of protection software and procedures]

[Need some aspect of CIA/impact-likelihood assessment]

### Information Asset Register Checklist

<input type="checkbox"/>	Are there new IAs being created?	[provide details]
<input type="checkbox"/>	Are old IAs being retired?	[provide details]
<input type="checkbox"/>	Have IAOs & IACs been consulted?	
<input type="checkbox"/>	Has IAR been updated/amended?	[at least create project task to do so]
<input type="checkbox"/>	Data Retention classification & period	



<input type="checkbox"/>	Data retention procedure/functionality in place	
--------------------------	---	--



## Appendix A – Supervisory Authority ‘High Risk’ Check

If the DPIA shows ‘high risk’ processing which cannot be mitigated, then the DPIA should be sent to the relevant authority for review before any processing starts. Note that their review may take several weeks to process. A ‘High Risk’ assessment represents a ‘risk to the rights and freedoms of individuals’ – so may extend beyond GDPR consideration, including Human Rights.

GDPR Article 35(3) provides three examples:

- a) a systematic and extensive evaluation of personal aspects relating to natural persons which is based on automated processing, including profiling, and on which decisions are based that produce legal effects concerning the natural person or similarly significantly affect the natural person;
- b) processing on a large scale of special categories of data referred to in Article 9(1), or of personal data relating to criminal convictions and offences referred to in Article 1013; or
- c) a systematic monitoring of a publicly accessible area on a large scale

ICO cites:

1. Systematic and extensive profiling with significant effects
2. Large scale use of sensitive data [viz. ‘special category’ in GDPR terms]
3. Public monitoring

These being the same as (a)-(c) above. They further identify:

1. **New technologies:** processing involving the use of new technologies, or the novel application of existing technologies (including AI).
2. **Denial of service:** Decisions about an individual’s access to a product, service, opportunity or benefit which is based to any extent on automated decision-making (including profiling) or involves the processing of special category data.
3. **Large-scale profiling:** any profiling of individuals on a large scale.
4. **Biometrics:** any processing of biometric data.
5. **Genetic data:** any processing of genetic data, other than that processed by an individual GP or health professional for the provision of health care direct to the data subject.
6. **Data matching:** combining, comparing or matching personal data obtained from multiple sources.
7. **Invisible processing:** processing of personal data that has not been obtained direct from the data subject in circumstances where the controller considers that compliance with Article 14 would prove impossible or involve disproportionate effort.
8. **Tracking:** processing which involves tracking an individual’s geolocation or behaviour, including but not limited to the online environment.
9. **Targeting of children or other vulnerable individuals:** The use of the personal data of children or other vulnerable individuals for marketing purposes, profiling or other automated decision-making, or if you intend to offer online services directly to children.



10. **Risk of physical harm:** Where the processing is of such a nature that a personal data breach could jeopardise the [physical] health or safety of individuals.

**'High Risk' assessment using ICO criteria:**

Criterion:	Assessment	Comments
New technologies	Low	Deep learning algorithms will be used only for image processing and analysis with image segmentation as outcome.
Denial of service	NA	
Large-scale profiling	NA	
Biometrics	NA	
Genetic data	NA	
Data matching	NA	
Invisible processing	NA	
Tracking	NA	
Targeting of children or other vulnerable individuals	NA	



Criterion:	Assessment	Comments
Risk of physical harm	NA	

[The assessment can be one of N/A (not applicable), Low, Medium, or High. The comments should explain how the assessment is justified.]

Appendix B – Broad Privacy Risk Assessment:

#	Risk Description/detail	Discussion
1.	Data accuracy and timeliness	[Is data accurately recorded & kept up-to-date?]
2.	Differential treatment of patients/data subjects	[Might certain categories of people be adversely affected, e.g. children, vulnerable adults]
3.	Data Accuracy and identification	[Is the identification of individual reliable? Is there a danger of mis-attribution or incorrect linkage of data?]
4.	Holding / sharing / use of excessive data within systems	[Might too much data be held or for long? Is there a clear justification for data retention? Not 'just in case']
5.	Data held too long within systems	[Is there a clear data retention period specified and are there processes to ensure its deletion when no longer needed? Are copies tracked and deleted as well?]
6.	Excessive range of access in terms of users to personal data (consider new users/change of access privileges)	[Do more users have access than strictly necessary? Are user roles clear distinguished and reflected in the access privileges? Is there a clear process for granting and revoking access privileges?]
7.	Potential for misuse of data, unauthorised access to systems	[What are the likely threats to the data? What countermeasures are or might be applied? Is it possible for access to be granted inappropriately?]
8.	New sharing of data with other organisations, including new or change of suppliers	[Is data being shared from new data providers or with new data users? Are there new suppliers or data processors? What controls will apply?]



#	Risk Description/detail	Discussion
9.	Variable and inconsistent adoption / implementation	[How well will this system work end-to-end? How robust is it against partial adoption or system failure?]
10.	Legal compliance, particularly DP transparency requirements and support for data subject rights	[How well does this system meet legal requirements – or appear to meet legal requirements? Does it meet the ‘No surprises’ rule? What would happen if an individual requests data erasure or ceasing processing, etc.]
11.	Medical confidentiality	[Are there any addition sensitivities over confidentiality? Might specific approval (e.g. REC) be required to support this processing?]







## IG Assessment Checklist ESR13 – Profiling the autoantibody repertoire in the context of systemic vasculitis flare

### Introduction to IG Assessment process

Under the General Data Protection Regulation (GDPR), a Data Protection Impact Assessment (DPIA) is only required where proposed data processing is “likely to result in a high risk to the rights and freedoms of natural persons” (Article 35(1)). However, Article 35(3) explicitly requires one where there is ‘large-scale’ processing of ‘special category’ (e.g. healthcare) data then a DPIA is required.

One other possibility is that the data being processed is already anonymised (see Recital 26) so falls outside GDPR altogether so that no DPIA is actually required.

However, good project management and information governance suggests that there should be a general approach to risk assessment for any project or business enterprise – if only to determine whether a DPIA might be required.

Ideally, one should work from a simple initial Checklist (this document) which identifies possible areas of information risk and compliance requirements to a ‘discussion note’ which explores any issues in more depth and may help identify the necessary mitigation methods and mechanisms to offset most if not all risks. Only if risks are unmitigated or remain ‘high’ would you move to a formal DPIA report.

### The IG Assessment approach

There should be an overview of the proposed project or business change to explain what processing is envisaged as well as the purpose and intended outcome. The ‘purpose’ is important to establish the legal basis for the processing as well as ensuring that any possible mitigations or counter-measures do not undermine the main rationale for the processing.

The next step is to establish what compliance requirements may apply: GDPR, contractual or other regulatory restrictions, consent requirements, or obligations to preserve the data for legal or other reasons (including the benefit of posterity perhaps).

Once the precise range of obligations has been established, then appropriate checks can be made and recorded within the document.

The most obvious of these being GDPR compliance. There must be a ‘High Risk’ assessment (Appendix A) to determine whether the supervisory authority needs to be informed – generally, it is expected that it will not be necessary; if so, then a formal DPIA report will be needed.

Appendix B has a broader Privacy Impact Assessment that may throw up some broader issues.

Initial conclusions as to next steps or particular countermeasures to be considered should be detailed below.



### Project Background/Overview

The project has as its foundation the hypothesis that there are more autoantibodies to be identified to describe new subgroups of systemic vasculitis flare and aims to ultimately showcase and screen for and characterise potential novel autoantibodies associated with systemic vasculitis flare. In order for this to be achieved, high density spotted antigen microarray screening samples will be derived from serial visits of patients recruited to the Irish Rare Kidney Disease biobank (through and in association with Trinity College Dublin) to characterise the differences between those suffering and not suffering a flare. Validation of these results will subsequently be conducted using samples from the Czech biobank using targeted antigen suspension bead arrays. These results will then be incorporated into the overall HELICAL project to assess whether these autoantibody repertoires are influenced by environmental factors.

### Comparison of process steps (simplified): [optional]

This allows identification of what processing is new or changed through the project:

Step	Current	Proposed
There are several steps for this work. Please refer to the Attached PDF of the Powerpoint Presentation that describes these. The first transfer will be between the Rare Kidney Disease Registry and KTH. This will also involve samples sharing	These transfers do not currently exist	This will involve the sharing of various data items as per the presentation. The data required is Patient samples, indicating gender, age, stage of disease and medication, MPO + and PR3, time point and date of collection, as Genetic data or Health-related data, date of diagnosis, date of COVID diagnosis.
Sample set from RKD	Samples were transferred from RKD to KTH, May 2021	
Data will be transferred from Vienna University.	These transfers do not currently exist	The DPIA must be updated prior to this sharing and processing
Sample set from MedUni	Samples were transferred from MedUni Vienna, Nov 2021.	Exosome samples from cell lines. The samples didn't contain any personal information; no human blood sample was included.



Step	Current	Proposed
Data analysis	Raw data produced from RKD experiment will be analysed by R statistical program. (not performed yet, but in near future)	Raw data will not be shared, only outcome of the analysis will be presented to WP3 and all consortium.
Next steps will involve the sharing of samples and data from BBMRI.cz in the Czech Republic with KTH	These transfers do not currently exist. Samples are not defined yet.	The DPIA must be updated prior to this sharing.

#### Initial Conclusions

concerning further counter-measures or business viability [possibly tentative]

18. Data sharing and materials transfer agreements must be in place to cover all sharing points described above.
19. The extent of the Machine Learning must also be considered and whether it represents closed loop processing. WE will need to explore this in more detail when this step is ready to be executed and prior to this processing.
20. Whilst efforts are being made to ensure only the transfer of anonymous data we must proceed with full compliance given that there is a raised chance of re-identification should data be shared outside of secure environments and end up in the public domain through accidental disclosure.

#### Compliance Checks required:

Tick	Requirement	Notes [replace guide text with response]
<input checked="" type="checkbox"/>	Does the project involve processing 'personal data' of any sort?	Yes, the project involves obtaining access to registries and personal data of patients such as age, sex, medical history and location. Whilst there are no direct identifiers, see initial conclusion item 3 above.



Tick	Requirement	Notes [replace guide text with response]
<input checked="" type="checkbox"/>	Does the project involve processing 'confidential data' of any sort?	Yes, all data access will be considered confidential.
<b>Data Availability requirements</b>		
<input checked="" type="checkbox"/>	Does data need to be held for GCP compliance?	Yes
<input checked="" type="checkbox"/>	Does data need to be held to meet 'Open Data' requirements?	Yes, some of the data must be made available under these terms. This will need to be reviewed prior to a decision being made.
<input checked="" type="checkbox"/>	Does data need to be held to meet ICMJE requirements or commitments?	Likely – all ESRs have been encouraged to familiarise themselves with these requirements.



GDPR Compliance Checklist – where ‘personal data’ is processed:

Tick	Requirement	Notes [replace guide text with response]
<b>Article 5: Principles compliance checks</b>		
<input checked="" type="checkbox"/>	ccc) Is processing lawful, fair, and transparent?	Yes. We must review the lawful basis for processing the data. This could be either consist of the consent obtained to be part of the RKD Registry, contractual obligation or public task depending on jurisdiction
<input checked="" type="checkbox"/>	ddd) Is the purpose (or purposes) of the processing clearly defined	Yes – please see the RKD documentation at <a href="https://www.tcd.ie/medicine/thkc/research/rare.php">https://www.tcd.ie/medicine/thkc/research/rare.php</a> . It is likely that this work could prepare additional materials
<input checked="" type="checkbox"/>	eee) adequate, relevant and limited to what is necessary	Yes – data minimisation has been applied in terms of justification of limited data sets and no direct identifiers.
<input checked="" type="checkbox"/>	fff) accurate and, where necessary, kept up to date	Data are kept safe based in GDPR regulation. Data will be held within the RKD registry as per its existing governance. Exports to KTH will be governed by an appropriate Data Sharing Agreement. We will update the DPIA for the other steps when they occur.
<input checked="" type="checkbox"/>	ggg) kept and permits identification of data subjects for no longer than is necessary	Data are being kept until the end of study and will be archived in accordance with governance retention requirements as per scientific research regulations , including data shared with KTH
<input checked="" type="checkbox"/>	hhh) processed securely	Yes – within the governance regime of TCD and under the DSA with KTH, including transfers via OneDrive (where these will be encrypted and uploaded).
<input checked="" type="checkbox"/>	11) can you demonstrate this compliance?	Yes – with regards auditing provision both for TCD and KTH (covered by the DSA)
<b>Articles 13 &amp; 14 compliance</b>		
<input type="checkbox"/>	Did the data came from publicly accessible sources?	No - the data will be received from RKD biobank of Trinity College Dublin [TCD] and securely shared with KTH using OneDrive.



Tick	Requirement	Notes [replace guide text with response]
<input checked="" type="checkbox"/>	Are data subjects informed before processing starts for any new purpose if incompatible with original purpose where the controller wants to use data for a different purpose to the purpose for which they currently hold data	Depending on the type research, if another person wants to use them it is possible, but should be checked with patient's permission certificate too
<input checked="" type="checkbox"/>	Does the Privacy Notice and/or PIL cover this processing?	Yes the data use is transparent with regards the RKD details at <a href="https://www.tcd.ie/medicine/thkc/research/rare.php">https://www.tcd.ie/medicine/thkc/research/rare.php</a>
<input checked="" type="checkbox"/>	What patient choices are available? Are these explained?	Refer to the RKD PILs for details at <a href="https://www.tcd.ie/medicine/thkc/research/rare.php">https://www.tcd.ie/medicine/thkc/research/rare.php</a>
<b>Articles 6 and 9: legal bases</b>		
<input checked="" type="checkbox"/>	What are legal bases under Article 6	TBC – likely consent (Art 6 (1)(a), 9(2)(a)and (j) as this involves genetic data
<input checked="" type="checkbox"/>	What are legal bases under Article 9 (if 'special category' data)	Consent and / or Scientific Research
<input type="checkbox"/>	Are Article 6 legitimate interests explained where relevant?	N/A
<input type="checkbox"/>	Are details of statutory obligations for Article 6 explained where relevant.	Yes
<input checked="" type="checkbox"/>	Is this proposed processing compatible with the declared purposes?	Yes [We believe so but TCD and RKD must be satisfied with this in entering into the DSA with KTH.]
<b>Article 89(1) research exemption</b>		
<input checked="" type="checkbox"/>	If for research, do we meet Art 89(1) data minimisation	Yes, TBC



Tick	Requirement	Notes [replace guide text with response]
<b>Articles 15-23: Data Subject Rights</b>		
<input checked="" type="checkbox"/>	Do we support data subject rights?	This work will proceed in line with the rights and freedoms as defined within the RKD PILs, plus data is anonymised [double check with TCD]
o	There is no use of automated decision making (e.g. profiling)	TBC – there may be some as part of the work with Tissuegnostics
<b>Articles 24-43: Controller-Processor</b>		
<input checked="" type="checkbox"/>	A28 & 29: What measures are there to ensure processors comply?	There is a formal Data Sharing Agreement being developed between TCD and KTH. A similar arrangement will be needed with the other partners as the work proceeds.
<input type="checkbox"/>	A30: Is there an entry for this processing/data held in the register?	This is for RKD – this should be confirmed.
<input checked="" type="checkbox"/>	A32-34: Do we ensure appropriate security, including protection against unauthorised or unlawful processing and against accidental loss, destruction or damage, using appropriate technical or organisational measures?	This must be handled by the DSA and the review of procedures for KTH.
<input checked="" type="checkbox"/>	A37-39: Is there a DPO and have they been or will they be consulted?	Yes – TCD DPO is being informed. KTH can be as well.
<b>Articles 44-50: International transfers</b>		
	What form of data will be transferred to a third country or international organisation	None, all the collaborators are in EU. Note the possibilities of issues with regards Brexit.





Tick	Requirement	Notes [replace guide text with response]
<input checked="" type="checkbox"/>	Are there safeguards for international transfers?	Yes – by virtue of the fact that this will not occur. This DPIA must be reviewed should Third Country transfers be planned (including UK with regards Brexit).
<b>Article 90: Obligations of secrecy</b>		
<input checked="" type="checkbox"/>	Do we meet medical confidentiality requirements?	Yes – working in line with the RKD Registry Governance

#### Data Subject Rights:

Note if supported and what process/procedure applies; if not, then describe the legal justification for not supporting this right. **Please refer to the RKD PILs and other materials.**

<input type="checkbox"/>	To be informed: about processing, about choices, about rights, about controller	
<input type="checkbox"/>	the right of access to see or receive a printed copy	
<input type="checkbox"/>	the right to rectification – to correct any material errors in the personal data	
<input type="checkbox"/>	the right to erasure – where appropriate, to ask that all personal data is erased	
<input type="checkbox"/>	the right to restrict processing – to ask that some or all processing ceases [see opt-out]	
<input type="checkbox"/>	the right to data portability – this only applies to data provided directly by individual	
<input type="checkbox"/>	the right to object to and not to be subject to automated decision-making, including profiling	
<input type="checkbox"/>	Right to object to a Data Processing Authority (typically the relevant supervisory authority of each Member State)	
<input type="checkbox"/>	Where consent is the legal basis, the right to withdraw consent	



### Detailed Transparency Checklist<sup>34</sup>

Does privacy information provided to data subjects include:

<input checked="" type="checkbox"/>	The name and contact details of our organisation	KTH <a href="http://www.kth.se">www.kth.se</a> – this will be updated to include other partners as described under the first section and in the attached PDF.
<input checked="" type="checkbox"/>	The name and contact details of our representative (if applicable)	Sbayati@kth.se
<input checked="" type="checkbox"/>	The contact details of our data protection officer (if applicable)	Robin Roy rroy@kth.se
<input checked="" type="checkbox"/>	The purposes of the processing	Discovering new biomarker for ANCA Vasculitis
<input checked="" type="checkbox"/>	The lawful bases for the processing	This is likely consent to scientific research.  Please refer to the following link which explains how KTH as well as Sweden meets the relevant lawful bases <a href="https://intra.kth.se/anstallning/anstallningsvillkor/att-vara-statligt-an/behandling-av-person/dataskyddsförordningen-gdpr-1.800623">https://intra.kth.se/anstallning/anstallningsvillkor/att-vara-statligt-an/behandling-av-person/dataskyddsförordningen-gdpr-1.800623</a>
<input type="checkbox"/>	The legitimate interests for the processing (if applicable)	N/A
<input type="checkbox"/>	The categories of personal data obtained (if the personal data is not obtained from the individual it relates to)	Patient samples, indicating gender, age, stage of disease and medication, MPO + and PR3, time point and date of collection, as Genetic data or Health-related data
<input type="checkbox"/>	The recipients or categories of recipients of the personal data	KTH – to be updated when other transfers occur.
<input type="checkbox"/>	The details of transfers of the personal data to any third countries or international organisations (if applicable)	No third country is planned to receive the data.

<sup>34</sup> Taken from UK Information Commissioner's Office template



<input type="checkbox"/>	The retention periods for the personal data.	In line with the requirements of Swedish law for retaining data
<input checked="" type="checkbox"/>	The rights available to individuals in respect of the processing	A consent form and information documents are provided to patients upon donating their blood, the relevant documents are provided by RKD
<input type="checkbox"/>	The right to withdraw consent (if applicable)	As per the above.
<input type="checkbox"/>	The right to lodge a complaint with a supervisory authority	As per the above.
<input type="checkbox"/>	The source of the personal data (if the personal data is not obtained from the individual it relates to)	RKD Registry.
<input checked="" type="checkbox"/>	The details of whether individuals are under a statutory or contractual obligation to provide the personal data  (if applicable, and if the personal data is collected from the individual it relates to)	Yes
x	The details of the existence of automated decision-making, including profiling (if applicable)	As per the later steps in this work. To be reviewed/
<input checked="" type="checkbox"/>	We provide individuals with privacy information at the time we collect their personal data from them – or where we obtain personal data from a source other than the individual it relates to, we provide them with privacy information:...	As per RKD PIL.
<input checked="" type="checkbox"/>	...within a reasonable of period of obtaining the personal data and no later than one month	



<input type="checkbox"/>	...if we plan to communicate with the individual, at the latest, when the first communication takes place	No contact with the patient is allowed
<input type="checkbox"/>	...if we plan to disclose the data to someone else, at the latest, when the data is disclosed	In this case the data is anonymous and transfers should be in line with the details within the PIL.
<input type="checkbox"/>	We provide the information in a way that is: <input checked="" type="checkbox"/> concise; <input checked="" type="checkbox"/> transparent; <input checked="" type="checkbox"/> intelligible; <input checked="" type="checkbox"/> easily accessible; and <input checked="" type="checkbox"/> uses clear and plain language.	Refer to the RKD PIL
<input type="checkbox"/>	When drafting the information, we: <input checked="" type="checkbox"/> undertake an information audit to find out what personal data we hold and what we do with it. <input checked="" type="checkbox"/> put ourselves in the position of the people we're collecting information about. <input checked="" type="checkbox"/> carry out user testing to evaluate how effective our privacy information is	This is with reference to RKD's processes for handling transparency. Likely within the PILs.
<input type="checkbox"/>	When providing our privacy information to individuals, we use a combination of appropriate techniques, such as: <input type="checkbox"/> a layered approach; <input type="checkbox"/> dashboards;	This has been handled by RKD.



<input type="checkbox"/> just-in-time notices;  <input type="checkbox"/> icons; and  <input type="checkbox"/> mobile and smart device functionalities.	
--	--

### Security & Access Control Checklist

Controls need to be appropriate to level of risk: identified special category data needs more protection against potential misuse than non-personal data.

	Data Security classification (above Official)	<input type="checkbox"/> - Official-Sensitive  <input checked="" type="checkbox"/> - Secret  <input type="checkbox"/> - Top Secret  <input type="checkbox"/> - Public Domain
<input checked="" type="checkbox"/>	Personal Data involved [GDPR]	It is anonymous but with a likelihood of reidentification if leaked into the public domain.
	Special Category of personal data involved [GDPR]	Health and genetic data
<input type="checkbox"/>	Electronic Communications (inc. cookies) [PECR]	
<input type="checkbox"/>	Credit Card data	
<input type="checkbox"/>	Legal enforcement [LED2018]	
<input type="checkbox"/>	Financial data	
<input type="checkbox"/>	Intellectual Property (detail owner)	
<input type="checkbox"/>	Commercial in confidence (detail owner)	
	Data Location (storage or processing)  (include any back-up site(s))	<input type="checkbox"/> - UK  <input checked="" type="checkbox"/> - EU/EEA  <input type="checkbox"/> - EU White-list  <input type="checkbox"/> - USA



		<input type="checkbox"/> - Other:
<input checked="" type="checkbox"/>	Is data held in secure data centre?	<p>Yes.</p> <p>The data is safely kept and stored in the RKD biobank of Trinity College Dublin.</p> <p>Data will be transferred to KTH via OneDrive at the request of the TCD DPO. KTH will undertake to securely process the data as part of the DSA.</p> <p>KTH Box is a file synchronizing service that is jointly procured by colleges / universities in Sweden via SUNET. KTH Box allows you to synchronize your personal files between computers, telephones and tablets as well as share files with colleagues and external collaborators in a safe manner. However, the TCD DPO has requested the use of OneDrive</p>
<input type="checkbox"/>	Is this new supplier, location, or system?	
<input type="checkbox"/>	Is all user access subject to 2-factor authentication?	<input type="checkbox"/> - no control <input type="checkbox"/> - single factor (e.g. just password) X- 2-factor (e.g. password & fob) <input type="checkbox"/> - biometric [note: GDPR reqs] <input type="checkbox"/> - Other control:
<input type="checkbox"/>	Are there established JML procedures?	A requirement of the DSA – this will be assured by declaration of the receiving party KTH



<input type="checkbox"/>	Are there checks that passwords are robust and secure enough?	A requirement of the DSA – this will be assured by declaration of the receiving party KTH
<input type="checkbox"/>	Are all administrator & user accounts routinely monitored?	A requirement of the DSA – this will be assured by declaration of the receiving party KTH
<input type="checkbox"/>	Are systems protected against malware and other attacks?	A requirement of the DSA – this will be assured by declaration of the receiving party KTH

[Need some aspect of CIA/impact-likelihood assessment]

#### Information Asset Register Checklist

<input type="checkbox"/>	Are there new IAs being created?	Yes – within KTH
<input type="checkbox"/>	Are old IAs being retired?	No
<input type="checkbox"/>	Have IAOs & IACs been consulted?	Yes – RKD consulted as well as TCD.
<input type="checkbox"/>	Has IAR been updated/amended?	TCD and KTH must do this.
<input type="checkbox"/>	Data Retention classification & period	Scientific research (Irish and Swedish Jurisdictions)
<input type="checkbox"/>	Data retention procedure/functionality in place	TBC – as part of DSA.



## Appendix A – Supervisory Authority ‘High Risk’ Check

If the DPIA shows ‘high risk’ processing which cannot be mitigated, then the DPIA should be sent to the relevant authority for review before any processing starts. Note that their review may take several weeks to process. A ‘High Risk’ assessment represents a ‘risk to the rights and freedoms of individuals’ – so may extend beyond GDPR consideration, including Human Rights.

GDPR Article 35(3) provides three examples:

- bb) a systematic and extensive evaluation of personal aspects relating to natural persons which is based on automated processing, including profiling, and on which decisions are based that produce legal effects concerning the natural person or similarly significantly affect the natural person;
- cc) processing on a large scale of special categories of data referred to in Article 9(1), or of personal data relating to criminal convictions and offences referred to in Article 1013; or
- dd) a systematic monitoring of a publicly accessible area on a large scale

ICO cites:

- 28. Systematic and extensive profiling with significant effects
- 29. Large scale use of sensitive data [viz. ‘special category’ in GDPR terms]
- 30. Public monitoring

These being the same as (a)-(c) above. They further identify:

- 91. **New technologies:** processing involving the use of new technologies, or the novel application of existing technologies (including AI).
- 92. **Denial of service:** Decisions about an individual’s access to a product, service, opportunity or benefit which is based to any extent on automated decision-making (including profiling) or involves the processing of special category data.
- 93. **Large-scale profiling:** any profiling of individuals on a large scale.
- 94. **Biometrics:** any processing of biometric data.
- 95. **Genetic data:** any processing of genetic data, other than that processed by an individual GP or health professional for the provision of health care direct to the data subject.
- 96. **Data matching:** combining, comparing or matching personal data obtained from multiple sources.
- 97. **Invisible processing:** processing of personal data that has not been obtained direct from the data subject in circumstances where the controller considers that compliance with Article 14 would prove impossible or involve disproportionate effort.
- 98. **Tracking:** processing which involves tracking an individual’s geolocation or behaviour, including but not limited to the online environment.
- 99. **Targeting of children or other vulnerable individuals:** The use of the personal data of children or other vulnerable individuals for marketing purposes, profiling or other automated decision-making, or if you intend to offer online services directly to children.





100. **Risk of physical harm:** Where the processing is of such a nature that a personal data breach could jeopardise the [physical] health or safety of individuals.

**'High Risk' assessment using ICO criteria:**

Criterion:	Assessment	Comments
New technologies		Suspension Bead Array, Plannar Array as well as the Machine Learning processing from Tissuegnostics
Denial of service		Unlikely to be an issue.
Large-scale profiling		A few number of patients will be analysed for the initial phase of study, afterwards we will scale up the number of samples and perform greater scale of profiling -
Biometrics		None
Genetic data		Genetic data is appreciated If the sample provider has any for each patient, but in this specific disease, it is not due to genetic disorder.
Data matching		At the moment, data combination with other collaborators is not decided yet, but there will be a chance to see the correspondence of our result with MUW. We will need to revisit this
Invisible processing		
Tracking		Patients are tracked by an online mobile application only



Criterion:	Assessment	Comments
Targeting of children or other vulnerable individuals		Children are not included as part of the project but patients with COVID-19 disease are vulnerable while they suffer from Vasculitis too and some may be elderly and considered vulnerable. The RKD governance will have this covered.
Risk of physical harm		None

This likely represents low risk – whilst data is anonymous there may be a chance of data being reidentified if it is disclosed publicly. The leaking of data may cause distress to participants and impact the credibility of RKD and TCD as well as KTH.

Provided there are assurances from KTH as defined in a DSA, the likelihood of public disclosure is low.

#### Appendix B – Broad Privacy Risk Assessment:

#	Risk Description/detail	Discussion
100	Data accuracy and timeliness	They are accurate as per RKD Governance
101	Differential treatment of patients/data subjects	The study does not involve direct patient contact or treatment of patients
102	Data Accuracy and identification	All the samples are coded by the data provider and I am not able to decode them or identify patients
103	Holding / sharing / use of excessive data within [Company] systems	All of samples will be kept until the end of study and won't be shared. They will be only shared with our publication our within HELICAL
104	Data held too long within [Company] systems	Yes at the end of study physical samples will be discarded and data will be published
105	Excessive range of access in terms of users to personal data (consider new users/change of access privileges)	[Do more users have access than strictly necessary? Are user roles clear distinguished and reflected in the access privileges? Is there a clear process for granting and revoking access privileges?] This must be covered in the DSA.



#	Risk Description/detail	Discussion
106	Potential for misuse of data, unauthorised access to systems	Protections are in place within KTH and across the HELICAL ITN to ensure data is carefully used and highly protected.
107	New sharing of data with other organisations, including new or change of suppliers	The data that I will use will be shared with the ESR's collaborators who are within WP3 in HELICAL
108	Variable and inconsistent adoption / implementation	
109	Legal compliance, particularly DP transparency requirements and support for data subject rights	Data providers in Ireland, process and support their data based on GDPR and Irish law regulations
110	Medical confidentiality	The work within RKD and terms of the DSA must assure this.



## Annex C – Material Transfer Agreements

These are available on request providing the parties to the agreements consent to their being shared.



**HRCDC**  
Health Research Consent  
Declaration Committee

Teach Grattan  
67-72 Sráid an Mhóta Iochtarach  
Baile Átha Cliath 2  
D02 H63B  
Éire

Grattan House  
67-72 Lower Mount Street  
Dublin 2  
D02 H63B  
Ireland

T: 353 1 234 5000  
F: 353 1 661 2335  
E: [info@hrcdc.ie](mailto:info@hrcdc.ie)  
[www.hrcdc.ie](http://www.hrcdc.ie)

**PRIVATE AND CONFIDENTIAL**

Prof. Mark Little  
School of Medicine  
Trinity College Dublin  
Dublin 2

By email only: [mlittle@tcd.ie](mailto:mlittle@tcd.ie)  
Cc by email only: [researchDPO@tcd.ie](mailto:researchDPO@tcd.ie)

30<sup>th</sup> July 2021

Dear Mark,

**RE: Application:** "Rare Kidney Disease Registry and Bioresource"  
**Reference ID:** 19-044-AF2  
**Data Controller(s):** Trinity College Dublin  
Firalis SAS  
**Decision:** Conditional Declaration

Thank you for your application to the HRCDC seeking a consent declaration on behalf Trinity College Dublin. The HRCDC convened on 20<sup>th</sup> July 2021 and reviewed the above referenced application, accompanying documents and responses to the Secretariat queries. After careful consideration, we are pleased to inform you that the following decision was made by the HRCDC:

- The HRCDC has exercised its right under Regulation (8)(4)(b) and has made a **Conditional Declaration** that the public interest in carrying out the health research significantly outweighs the requirement of the Applicant(s) to seek explicit consent of the data subject, whose personal data is being processed for the above referenced health research study.
- The scope of the Declaration is for the following data processing activities specifically related to the above referenced health research study:

**Scope of Declaration:**

For the data processing activities being carried out in connection with the industry collaborator, Firalis SAS. Specifically, the declaration is for the continued sharing and use of data and associated samples from the RKD Registry/Bioresource with Firalis SAS, for the purpose of the HELICAL study.

- The following specific conditions have been attached to the Conditional Declaration as follows:

**Condition 1.** Firalis SAS is confirmed as a joint-data controller with TCD for the collaborative HELICAL study accessing data from the RKD Registry/Bioresource. The consent declaration will not become effective until the following actions are completed, and this condition is fully met:

- an authorised signatory on behalf of Firalis SAS must be provided on the HRCDC application form and submitted to the HRCDC. This is in line with the HRCDC processes where joint-data controllers apply for a consent declaration.
- feedback from the Firalis SAS data protection officer (DPO), or equivalent, must be provided to the HRCDC on the RKD Registry/Bioresource data protection impact assessment (DPIA), to ensure any data protection risks to data being processed by Firalis SAS is assessed, and risks mitigated against where necessary.



**HRCDC**  
Health Research Consent  
Declaration Committee

Teach Grattan  
67-72 Sráid an Mhóta Iochtarach  
Baile Átha Cliath 2  
D02 H638  
Éire

Grattan House  
67-72 Lower Mount Street  
Dublin 2  
D02 H638  
Ireland

T: 353 1 234 5000  
F: 353 1 661 2335  
E: info@hrcdc.ie  
www.hrcdc.ie

- (iii) confirmation that Firalis SAS shall support the implementation and compliance of the consent declaration jointly with Trinity College Dublin, where required.

**Condition 2.** The scope of this consent declaration is limited to the processing of personal data from the RKD Registry/Bioresource specifically for the purpose collaboration with Firalis SAS only. For the avoidance of doubt, the consent declaration does not extend to the processing of data for other industry collaborations. If required, an amendment request or new consent declaration application, as appropriate, can be submitted to the HRCDC for consideration.

**Condition 3.** The Applicant is required to confirm to the HRCDC that the contractual agreement/arrangement in place with Firalis SAS, has adequate terms and conditions with regards a joint-data controller arrangement, as required under Article 26 of GDPR. Where existing arrangements require updating, this should be carried out as soon as possible and no later than 3 months of receipt of the decision letter, with notification provided to the HRCDC.

**Condition 4.** The Applicant is required to continue efforts to re-consent participants. This includes efforts to obtain re-consent via the hospital/clinical visits as well as by alternative means, such as telephone or post, where participants are not attending the clinics. The efforts made to obtain re-consent from participants via clinics and other approaches, as well as the numbers who have re-consented is a reporting requirement of the Annual Review.

- The Declaration is made solely to the Applicant(s) who is the Data Controller and not to any other third party.
- The Declaration is made commencing 8th August 2018 and shall be valid until re-consent can be obtained, or if re-consent is not achievable, until the end of the Firalis SAS / Helical study, at which point all samples and data will be irrevocably anonymised by 31st Dec 2022. The Applicant may seek an extension to the consent declaration by way of submitting an amendment request to the HRCDC for consideration.

In addition to the decision made by the HRCDC, the following standard conditions of the Declaration shall apply:

- the Applicant must complete an Annual Review to the HRCDC on the anniversary date of this decision letter and for every year, or part year, the Declaration is valid,
- the Applicant must have any necessary contractual obligations in place,
- all activities being carried out are in compliance with the General Data Protection Regulations, the Data Protection Act 2018 and Health Research Regulations 2018, for the duration of the Declaration,
- any breaches that occur that affect the integrity of the Declaration and the protection of data subjects, must be reported to the HRCDC,
- the health research must be conducted lawfully and ethically.

**NOTE:** Failure to meet a condition, including the condition to submit an Annual Review to the HRCDC which is a statutory requirement under the Health Research Regulations (Regulation 13(1)), may lead to a revocation of the consent declaration.



**HRCDC**  
Health Research Consent  
Declaration Committee

Teach Grattan  
67-72 Sráid an Mhóta Íochtarach  
Baile Átha Cliath 2  
D02 H638  
Éire

Grattan House  
67-72 Lower Mount Street  
Dublin 2  
D02 H638  
Ireland

T: 353 1 234 5000  
F: 353 1 661 2335  
E: info@hrcdc.ie  
www.hrcdc.ie

**Note:** Lastly, the HRCDC notes that re-consent has been obtained from 235 participants and further notes the Applicant's statement that *'None of the participants refused to provide updated consent'*. The HRCDC therefore understands that the participants that re-consented to date have all actively affirmed re-consent for data processing, and are not 'non-responders'. **If this understanding is incorrect, the Applicant must provide clarification to the HRCDC.**

Please confirm acceptance of the Declaration within 30 working days of receipt of this letter, or the Declaration will lapse. Any clarifications required with respect to the decision made must be requested within the 30 day timeline.

Please notify your Data Protection Officer or equivalent authority within your organisation of this decision.

On behalf of the HRCDC and Secretariat, we wish you the very best of luck with the research study.

Kind regards,

---

Emily Vereker, PhD  
Programme Manager, Secretariat  
Health Research Consent Declaration Committee



## Annex D – Data Transfer Agreements

These are available on request providing the parties to the agreements consent to their being shared.





## Annex E – List of Legal Bases and Special Category Personal Data Justifications

ESR	Legal basis under GDPR	Further explanation
ESR 1 - Albert	<p>Art. 6(1)(e) – public interest research</p> <p>Art. 9(2)(j) – public interest research</p> <p>Art. 89(1) requirements met</p>	<p>Data falls under special category data (Art.9 GDPR). ESR is monitoring consent legal basis secondary legislation requirement.</p> <p>Art 89 requirements are met – ESR is using the minimum personal and confidential data to achieve the research goal</p>
ESR 2 – Anna	<p>Art. 6(1)(e) – task carried out in the public interest</p> <p>Art. 9(2)(j) – public interest research</p> <p>Art. 89(1) requirements met</p>	
ESR 3 - Bahareh	<p>Art. 6(1)(e) – public interest research</p> <p>Art. 9(2)(j) – public interest research</p> <p>Art. 9(2)(j) – public interest research</p> <p>Art. 89(1) requirements met</p>	
ESR 4 – Enock	<p>Art. 6(1)(a) – consent</p> <p>Art. 6(1)(f) – legitimate interest</p> <p>Art. 9(2)(a) – consent</p> <p>Art. 9(2)(j) - public interest (Archiving, research and statistics)</p>	<p>Legitimate interests were explained using a series of questions that covers the 3-step test for “legitimate interest”.</p> <p><i>Purpose test: what are UK Biobank’s legitimate interests?</i></p> <ul style="list-style-type: none"> <li>• <i>What is UK Biobank trying to achieve? Our objective is to set up and manage a major international</i></li> </ul>



		<p>research resource for health-related research that is in the public interest.</p> <ul style="list-style-type: none"> <li>• <i>Who benefits from UK Biobank's processing?</i> Patients and the wider public benefit from the advances made in the prevention, diagnosis and treatment of disease.</li> <li>• <i>How significant/important are these benefits?</i> UK Biobank is now one of the largest and most used health research resources in the world. Over 6,000 institutions are registered with us and over 1,000 health-related research applications have been approved.</li> </ul> <p><i>Necessity test: is the processing necessary for the legitimate interests?</i></p> <ul style="list-style-type: none"> <li>• <i>Is processing personal data a reasonable way to achieve the objective?</i> Without the personal data provided voluntarily by you and the other participants, UK Biobank would not exist.</li> <li>• <i>Is there another less obtrusive way to meet our purposes?</i> Your data are stored in a way that makes it is extremely difficult even for UK Biobank to re-identify you. Only a very few individuals within UK Biobank are allowed to do so (and they are strictly monitored) in order that further information about you can be added. Data provided to researchers have personal identifiers removed so that an individual participant cannot be identified. There are no circumstances in which your data can be processed in a</li> </ul>
--	--	---



		<p>manner that could have an adverse impact on you.</p> <p><i>Balancing test: UK Biobank had to weigh up the participant's interests.</i></p> <ul style="list-style-type: none"> <li>• <i>Would participants expect UK Biobank to use their data this way? Yes; this is what we set out in the information materials provided to participants and in the consent form each of them signed.</i></li> <li>• <i>How likely would a participant be to object? In UKB view, this was very unlikely. During the past 10 years since participants joined UK Biobank during 2006-10, fewer than 800 of the 500,000 participants have withdrawn from the study and asked that we delete all of their information.</i></li> </ul>
ESR 5 - Solange	<p>Art. 6(1)(a) – consent</p> <p>Art. 9(2)(a) – consent</p>	<p>The ESR has declared that an LIA (Legitimate Interest Analysis) form will be completed. The ESR has confirmed that the details of statutory obligations for Article 6 will be explained. The ESR has stated that Art. 89 research exemption is not applicable as genetic data will not be used in this project. Consent forms were distributed and obtained.</p>
ESR 6 – Alejandro	<p>Art. 6(1)(e) – task carried out in the public interest</p> <p>Art. 9(2)(j) - scientific research</p> <p>Art. 89(1) requirements met</p>	



ESR 7 - Elkyn	<p>Art. 6(1)(e) – task carried out in the public interest</p> <p>Art. 9(2)(j) - archiving/scientific research</p> <p>Art. 89(1) requirements met</p>	<p>ESR has stated that Art.89 requirements are met - all the samples that are collected in our project are treated confidentially and are assigned a unique and consecutive alphanumeric code</p>
ESR 8 - Michal	<p>Art. 6(1)(a) – consent</p> <p>Art. 9 (2)(a) – explicit consent</p> <p>Art. 89(1) requirements met</p>	
ESR 9 - Filippo	<p>Art. 6(1)(e) – task carried out in the public interest</p> <p>Art. 9(2)(j) – public interest research</p> <p>Art. 89(1) requirements met</p>	<p>Art 89 requirements are met – ESR is using data that are limited to what is necessary for the purposes for which they are processed</p>
ESR 10 - Farah	<p>Art. 6(1)(e) – task carried out in the public interest</p> <p>Art. 9(2)(j) – public interest research</p> <p>Art. 89(1) requirements met</p>	
ESR 11 - Malgorzata	<p>Art. 6(1)(a) – consent</p> <p>Art. 9(2)(j) – public interest research</p> <p>Art. 89(1) requirements met</p>	
ESR 12 – Marco	<p>Art. 6(1)(e) – task carried out in the public interest</p> <p>Art. 9(2)(j) – public interest research</p> <p>Art. 89(1) requirements met</p>	<p>Data falls under special category data (Art.9 GDPR). Art 89 requirements are met –</p> <p>ESR has provided further information in the ethics application/approval</p>



<p>ESR 13 – Shaghayegh</p>	<p>Art 6 (1)(e) - public interest research</p> <p>Art. 9(2)(j) - public interest research</p> <p>Art. 89 (1) requirements met</p>	
<p>ESR 14 - Gisela</p>	<p>Art 6 (1)(e) - public interest research</p> <p>Art. 9(2)(j) - public interest research</p> <p>Art. 89 (1) requirements met</p>	



## Annex F – Module 2 Curriculum and Transparency Workshop Materials

### HELICAL Mid Term Module – Preparatory Work

We are looking forward to running the Data Protection and Linkage Modules over June. To run the sessions, we need you to prepare a slideshow for presentation at the session on 1<sup>st</sup> June using following the details below.

Please submit these in a PowerPoint Presentation before the scheduled online training (submission deadline: 29<sup>th</sup> May 2020) on the Basecamp Online Module Group. Please be as precise as possible and be prepared to speak for 10 minutes and take question from your fellow ESRs and the module facilitators.

#### **Provide name, title of research and home and collaborating organisations**

**Description:** In a short paragraph, describe what your project is about.

**Key Benefits of your work:** in a brief paragraph, describe what problems your research is expected to solve.

**Data needs:** what data do you need to conduct your research?

**Data sources:** where is your data going to come from?

**Data analysis and use:** what will you be doing with your research?

**Data flow diagram:** Please supply a data flow diagram of your work.

**IF YOU KNOW:** What secure infrastructures are available at your home institutions? Do you have data safe havens / trusted research platforms?

**Ethics, consent and approvals:** do you have any of these or do you know where you need to go to apply for them? If so please specify what you have, where the requirements are listed and what you need to do.

**IF YOU KNOW: What are the legal bases for processing your data?** Please specify where you have taken advice, if any. Please also specify whether consent has been sought from participants and if not, why not.

**What do you want the online modules to focus on?** Please tell us what is of particular concern to you that you would like to cover.



HELICAL ITN



Ethical linking of electronic health  
data to research data to support  
research, Open Science and uphold  
FAIR principles



*This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 813545.*



HELICAL



## Facilitators Today



 <p><b>Dr. Nathan Lea</b> i-HD</p>	 <p><b>Maria Christofidou</b> i-HD</p>	 <p><b>Prof. Mark Little</b> TCD</p>	 <p><b>Julie Power</b> RITA/VIA</p>	 <p><b>Patricia Ryan</b> VIA</p>
---	---	---	---	---

 *This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 813545.*



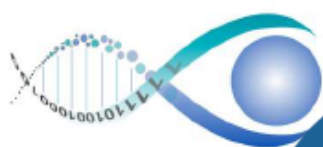


## Special Thanks

- Jessie Greene
- The Doctoral Studies Committee



*This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 813545.*




*This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 813545.*



## Overview of session

- Introductions
- Brief GDPR recap
- DPIAs and security tools introduction - how we manage risk
- Patient concerns and priorities
- FAIR Principles and Open Science
- Outcome: An overall understanding over and a recap of privacy risk management



The European Institute For  
Innovation Through Health Data



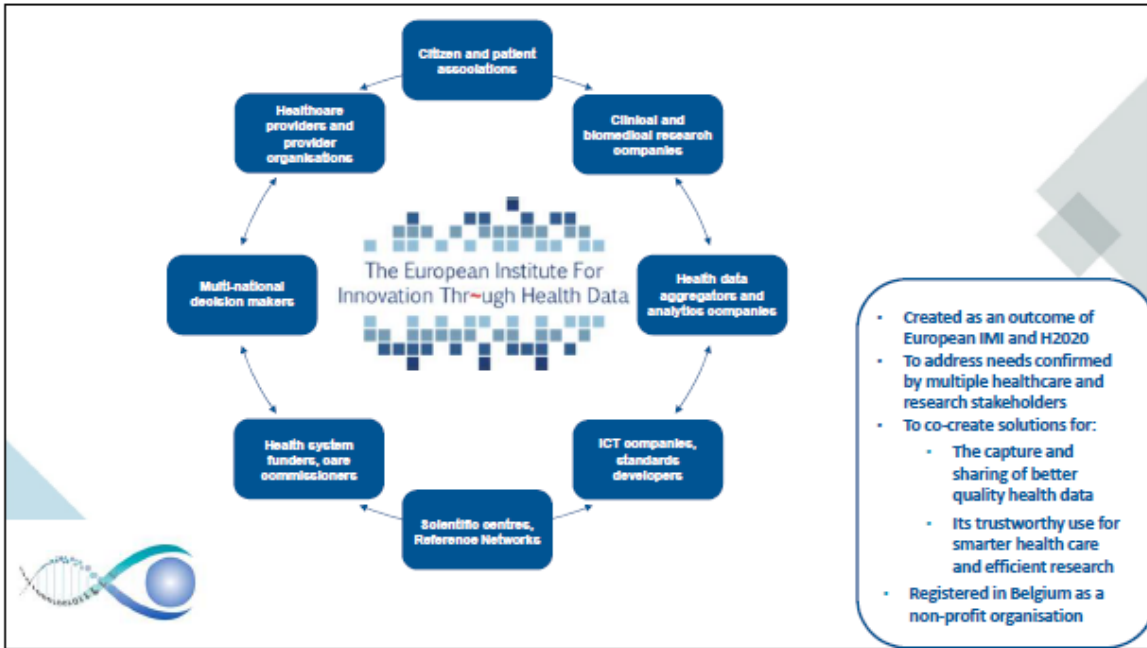
## Introduction

- COVID-19: Adapting to the situation
- Objectives
  - To become familiar with the practical considerations of Regulations like the GDPR and duties of confidence when handling and linking data
  - To appreciate how GDPR must be an enabler for robust research, supporting the Open Science agenda and serving the public interest for data driven research
  - To understand the importance of regulation and honouring the FAIR principles
  - To explore these objectives by gaining practical experience in managing regulatory compliance and working safely and securely with health data for research
  - To learn to handle regional variation in how regulations are implemented in practice
  - To learn to use tooling that helps achieve regulatory compliance including anonymisation and pseudonymisation, access controls, authorisation, encryption and Data Protection Impact Assessments (DPIAs)
  - To learn meaningful transparency in partnership with our patient awareness groups
- Online components (what we expect from you)





HELICAL



VIA Vasculitis Ireland  
AWARENESS  
RITA/VIA



## A brief recap on GDPR



- **Principles**
  - Lawful and fair purposes
  - Specific purposes - where research is no unrelated to original purpose
  - Adequate, relevant and limited
  - Accurate and kept up to date where necessary
  - Minimisation
  - Security
- **Legal Bases**
  - Consent
  - Vital Interest
  - Public Task
  - Legitimate Interest
- **Special Category**
  - Identify a legal basis
  - Satisfy at least one of 10 additional provisions
  - Consent
  - Providing health related services, public health or scientific research for reasons of public interest
  - Archiving
  - Substantial Public Interest
  - Manifestly made public by the data subject



## Data Protection Impact Assessment (DPIA)



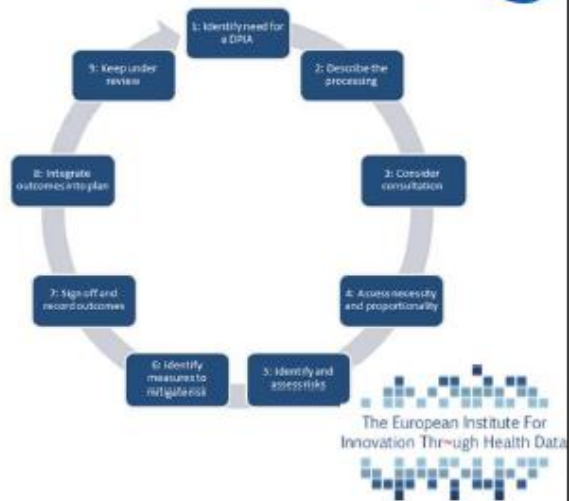
- DPIA is a process to help you identify and minimise the data protection risks of a project
- **Why we have DPIAs**
  - A risk assessment tool
  - Emerged from the old privacy model contained in the old Directive
  - This amendment has now given the model larger scope and made DPIAs mandatory – particularly for data uses that entail new technologies or sensitive information being used
- **Helps you to manage risk by:**
  - Lawfulness and consent
  - Accountability, necessity and accuracy
  - Security (data safe havens and encryption processes)
  - Transparency
- **A DPIA must:**
  - Describe the nature, scope, context and purposes of the processing
  - Assess necessity, proportionality and compliance measures
  - Identify and assess risks to individuals
  - Identify any additional measures to mitigate those risks
- **Note:** the more you input in a DPIA, the more you get out!
  - if treated as a tick-box, time is wasted as you're not including details and results aren't adequately accurate





## DPIA

- An example of a DPIA
- Once presentations have been submitted on 2<sup>nd</sup> June and following your feedback, you will take home a DPIA exercise
- Big part of GDPR compliance is transparency and transparency with patients through interactions is a vital part
- We're lucky to have the patients voice represented today to cover the important elements from a patient perspective



## Patient involvement

**VIA** Vasculitis Ireland  
AWARENESS





## Exchanges with patient representatives: takeaways

- **Topics about the GDPR and information security on which patients need better education in order to make informed decisions and in general to be empowered**
  - There is a variety of levels when it comes to the understanding of what the GDPR entails and the rights of patients that comes with it
  - Unclear to the means that personal data can be protected (anonymisation, pseudonyms etc)
- **Aspects of what patients are asked to agree to when they participate that should be better explained, or on which they should have feedback during the study**
  - Unanimous opinion for more involvement of patients in the creation of materials and overall the project set up and process
  - Favourable views for outreach/education pieces to be circulated surrounding GDPR prior to obtaining consent
- **Aspects of the research on which patient input should be more strongly included during the design stage**
  - Concern expressed towards patients not having a full/proper understanding of what is requested from them and their rights in consent forms - a general wish towards ESRs spending more time with patients when explaining this and giving the chance to come back with questions
  - To ask patients participating in surveys if they'd be willing to be contacted in the future and the means to do so
  - Keep patients in the loop as to how and when their personal data is utilised in a project as well as to the progress of said project



The FAIR Data Principles are a set of guiding principles in order to make data:

- Findable
- Accessible
- Interoperable
- Reusable

Not just for scientists

- It comes down to participants in research and having their data shared
- Striking the balance between privacy/data protection v. informing best science endeavors

Think about the open science agenda and the need for balance

- Focus on data protection as it's more complicated
- By the end of the end of the sessions we will tie this back – keeping thinking about these principles throughout
- If one positive, COVID has placed real attention on this: to ensure data is accurate and protected whilst being open/accessible



## FAIR Principles and Open Science





HELICAL



HELICAL ITN



Questions?  
Feedback?



The European Institute For  
Innovation Through Health Data



*This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 813545.*



HELICAL ITN  
ANNEX



*This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 813545.*

