FAIRVASC

FAIRVASC Registry Legal Metadata Profiles

Document date: 01/04/2021
Document version: 1
Deliverable No: 2.3

Full Name	Short Name	Number	Role
Meyer's Children's Hospital, Florence, Department of Medicine.	MEYER	1	Management Coordinator
Trinity College Dublin Trinity Translational Medicine Institute, Trinity Centre for Health Sciences ADAPT Centre for Digital Content, O'Reilly Institute, Trinity College Dublin	TCD TCD-ADAPT	2	Technical Coordinator
French Vasculitis Study Group, Cochin Hospital, APHP, University of Paris, Department of Internal Medicine	APHP	3	Partner
Lund University, Clinical Sciences, Rheumatology, Lund and Dept. of Rheumatology, Skåne University Hospital, Lund	LUND	4	Partner
University of Lübeck, Department of Rheumatology	UKSH	5	Partner
Jagiellonian University Krakow, Department of Medicine	JUK	6	Partner
General University Hospital in Prague, Department of Nephrology	GUH	7	Partner
University of Glasgow, Institute of Infection, immunity & Inflammation	GLASGOW	8	Associated research partner
Pintail Ltd	PT	9	Subcontractor
The European Institute for Innovation Through Health Data	i~HD	10	Subcontractor
Patient Advocacy Organisations (PAO)			
Vasculitis Stichting and Vasculitis International	VAS/SVI	1	PAO
Vasculitis Ireland Awareness	VIA	2	PAO
Vasculitis UK	VUK	3	PAO

D2.3 FAIRVASC Registry Legal Metadata Profiles

1 Executive Summary

Within FAIRVASC Workpackage 2 (WP2) will capture the legal and ethical environment of each registry, patient perspectives on data sharing and re-use, and also the national and international legislation and regulations that apply to FAIRVASC. D2.3 is the final in a series of deliverable describing information governance compliance across the FAIRVASC registries. In particular this deliverable provides the specific justifications around the selection of FAIRVASC data variables in line with the wider governance frameworks as identified and defined in other WP2 Deliverables. FAIRVASC teams including i~HD Glasgow, TCD ADAPT, LUND and Meyer reviewed the proposed data sets, liaised with the registries, harmonised a data model and looked to justifying data item inclusion. The result of this work is described in Appendix 6.1 which provides a high-level summary of registry data dictionary items included in FAIRVASC (as described in FAIRVASC deliverable D1.2) with their inclusion justification (as described here in D2.3).

2 Introduction & Background

2.1 Role of this Deliverable in the Project

The overall objectives of WP2 Legal and Ethical Profiling within FAIRVASC is

1. To verify that the appropriate legal basis, GDPR compliance and other permissions

for the use of data from each registry is in place

2. To create a legal, access policy, data management and regulatory profile of each

registry

3. To provide a legal framework for data sharing within and beyond FAIRVASC

Within WP2 we have reported on D2.1 FAIRVASC DPIA and Code of Conduct and D2.2 FAIRVASC Registry Legal Handbook providing an overall snapshot of information governance compliance across the registries themselves.

Deliverable D2.3 describes the overall process and provenance of the legal and governance profiles for data *within* the registry when identifying and justifying the inclusion of specific variables as listed in Deliverable 1.2.

2.2 The inputs and dependencies for the work described in deliverable

This deliverable relies upon the risk assessments and governance oversight established as part of Deliverables 2.1 and 2.2, which have informed the overall governance structure for FAIRVASC and how the Consortium approaches the management of data to achieve its purposes.

This deliverable must also be read with Deliverable 1.2 Registries and Content metadata records which describes the resulting Metadata catalogue that the particulars here support.

It relates directly to and must be read with:

- Deliverable 1.2 which lists the catalogue of data to support the research goals and outputs for FAIRVASC and the wider international Vasculitis research aims and care priorities;
- Deliverable 2.1 which describes the Data Protection Impact Assessments, risk mitigation strategies and codes of conduct for data use;
- and Deliverable 2.2, which provides the legal handbook and approvals in place across the registries and describes their applicability to data handling within the project.

The primary inputs for this work have been driven by the Data Protection Impact Assessment (DPIA) provided in Deliverable 2.1 and the wider controlling mechanisms including ethical oversight as described in Deliverable 2.2 as informed by the DPIA. The discussions and oversight of the clinical, governance and patient representatives has been essential in defining the variables that are needed and the justification for their inclusion.

These inputs have had to balance the public good in ensuring rich data can:

- support the advancement of Vasculitis research and its care;
- ensure sufficient data is provided to achieve this without undue risk to rights and freedoms or participants with regards privacy and confidentiality;
- assure the accuracy of data and thereafter findings;
- allowing the transparency of data use for public review;
- and ensure that these goals are achieved within the requirements of FAIR Principles.

2.3 The relationship between what is shown in this deliverable and the overall objectives of the project.

This deliverable provides the specific justifications around the selection of variables in line with the wider governance frameworks as identified and defined in other WP2 Deliverables. It should be treated as a living document as the scope and justifications of data items to be used will likely expand and need refinement.

3 Approach

The DPIA and wider legal and contractual reviews set the tone for the consortium deliberation and provided focus points for discourse. Guided by i~HD from the information governance and Data protection by Design and Default perspective, the WP2 teams including Glasgow, TCD ADAPT, LUND and Meyer reviewed the proposed data sets, liaised with the registries, harmonised a data model and looked to justifying data item inclusion.

This involved frequent meetings and oversight where the guiding principle was to justify data inclusion and review how the data should be shared (i.e. to harmonise across data sets as well as perturb data to maintain participant anonymity, or simply exclude data items that were not needed).

The meetings and online collaborations included clinical and wider consortium oversight and expertise.

4 Results

The results of this work are represented in Deliverable 1.2 and summarised in the table in Appendix 6.1 where the following specific principles apply:

1. Data Items must be selected on the basis of answering specific clinical research questions

This would allow a clear justification for selecting data items to serve the public interest and better understand Vasculitis and its management.

2. Question formulation and reformulation is a "living process" and must be treated as such

In justifying data item selection, evolving understanding and specific research questions is a clear output from collegiate research prioritisation and understanding results of data interrogation. This means that restriction of data items must support and not hinder research exploration freedom.

3. Balance risk of harm for sharing too much data with risk of harm from stifled research

This principle relies heavily on understanding the particulars of the DPIAs and other risk management and policy where the approach has been to ensure the safe, reliable and secure transfer of data and ensures that the patient representative community have a voice in prioritising the research and articulating concerns. The exclusion of data items is purely about ensuring best research and must be applied with that in mind, where the risks and benefits are as much about participant views as they are for research and regulators.

4. Justify, justify, justify…

Specific data items and their sharing must support research whilst maintaining participant anonymity. When providing ranges instead of discrete figures, the value of helping to assure anonymity cannot devalue the quality and nature of the research.

Likewise the use of those variables must be clearly justified while data restrictions must themselves be justified.

Note that a treatise of the interpretation of these principles is included here and in Deliverable 1.2 for each of the demographic data items and any uniquely attributable clinical data. For FAIRVASC, four 'inclusion justification' categories have been specified:

1. Essential to determine baseline characteristics for each subject
2. Essential to determine intermediate biomarkers of importance
3. Essential to determine clinical outcomes of importance for each subject
4. Essential information for justifying scientific approaches that may require differing legal justification.

As in appendix 6.1 (High level summary of registry data dictionary items and inclusion justification (D1.2 and D2.3)), different inclusion justification categories apply depending on the data in question. For example, with respect to category 1, patient demographics, which would typically include age and sex, is an example of data that are an essential component of baseline population characteristics in any epidemiological research study. Category 2 refers to 'intermediate biomarkers'. An example of this from the non-vasculitis research sphere is: in a trial of a statin medication, cholesterol blood levels can

be used as an intermediated biomarker to evaluate the efficacy of treatment, with the inference that lowering blood cholesterol presumably could lead to reduced risk of 'hard clinical endpoints' such as heart attack or death. Intermediate biomarkers in vasculitis could include blood inflammatory markers such as C-reactive Protein or autoimmune antibodies. Category 3 refers to hard clinical end points as described above such as heart attack or death. An example of additional relevance in vasculitis includes End-Stage Kidney Disease, in addition to heart attack and death. Category 4 covers consent data, this is important to gather as it may help determine whether different data analysis approaches may be required in future due to legal or ethical considerations. Notably there are no data items which the FAIRVASC consortium has deemed to be non-essential for the scientific objectives of the study. This is due to significant pre-selection of data variables that occurred at the design stage of each individual registry – all registries involved in FAIRVASC were designed to facilitate clinical research relating to vasculitis, therefore it follows that all collected data items are of high utility with respect to the research objectives of the consortium.

5 Impact & Conclusion

Through a collegiate approach that has prioritised research outputs with ensuring a transparent balancing of risks to participant privacy, FAIRVASC has commenced with a justified and rich set of data variables to begin to answer challenging questions. The Project has ensured there is a safe level of flexibility to maintain the anonymity of participants whilst protecting the evolving nature of clinical research questions. By balancing these risks, the foundation of a resource that substantially meets the needs of the public and the Vasculitis community has been set in the form of a data dictionary.

6 Appendices

6.1 High level summary of registry data dictionary items and inclusion justification (D1.2 and D2.3)

EUVAS	RKD	GEVAS	Czech	Skåne	POLVAS	UKIVAS	FVSG	Inclusion justification*
demographics	demographics	demographics	demographics	demographics	demographics	demographics	demographics	1
consent	consent	consent	-	-	consent	consent	consent	4
vasculitis diagnosis	vasculitis diagnosis	vasculitis diagnosis	vasculitis diagnosis	vasculitis diagnosis	vasculitis diagnosis	vasculitis diagnosis	vasculitis diagnosis	1
clinical history	clinical history	clinical history	clinical history	-	clinical history	clinical history	clinical history	1
anca IF	anca IF	anca IF	anca IF	anca IF	anca IF	anca IF	anca IF	1,2
anca elisa	anca elisa	anca elisa	anca elisa	anca elisa	anca elisa	anca elisa	anca elisa	1,2
bloods	bloods	bloods	bloods	bloods	bloods (minimal)	bloods	bloods	1,2
urine tests	urine tests	urine tests	urine tests	urine tests	-	urine tests	urine tests	1,2
treatment	treatment	treatment	treatment	treatment	treatment	treatment	treatment	1,3
relapse occurrence	relapse occurrence	relapse occurrence	relapse occurrence	-	relapse occurrence	relapse occurrence	relapse occurrence	1,3
complications	complications	complications	complications	complications	complications	complications	complications	1,3
mortality	mortality	mortality	mortality	mortality	mortality	mortality	mortality	3
organ involvement	organ involvement	organ involvement	organ involvement	organ involvement	organ involvement	organ involvement	organ involvement	1,2,3
bvas	bvas	bvas	bvas	bvas	-	bvas	bvas	1,2,3
vdi	-	vdi	vdi	vdi	-	vdi	vdi	1,2,3
covid	covid	-	-	-	-	covid	-	1,2,3
aavpro	-	-	-	-	-	-	-	1,2,3
QOL	QOL	-	-	-	-	-	-	1,2,3
biopsy findings	biopsy findings	biopsy findings	biopsy findings	-	-	-	biopsy findings	1,2
imaging	imaging	imaging	-	-	-	-	imaging	1,2
-	-	pregnancy	-	-	-	-	-	1,3

kidney transplantation	kidney transplantation	-	-	-	-	kidney transplantation	-	1,3

* 1 = Essential for determining baseline characteristics for each subject, 2 = Essential to determine intermediate biomarkers of importance, 3 = Essential for determining clinical outcomes of importance for each subject, 4 = Essential information for justifying scientific approaches that may require differing legal justification