H2020-MSCA-ITN-2018-813545

HELICAL

Health Data Linkage for Clinical Benefit

Deliverable 4.2

HELICAL Information Governance Policies

Introduction

The development of policies governing data use within national and international medical research projects received much needed support and clarification with the arrival of key regulatory changes. In 2018, the General Data Protection Regulation provided two key requirements for the handling of personal and sensitive (special category) data. The first was the notion of "data protection by design and default" where protection of data needed to be built in from the outset of any data intensive activity. 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 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.

There is no doubt that the HELICAL Innovative Training Network would stand to benefit from GDPR and these two particular requirements. This is clear in the objectives of the ITN – to train the next generation of researchers in the particulars of handling data, protecting data assets, whilst balancing the mandates around open science and free flows of data to support research. As with any medical research there are challenges to understanding the correct and optimal interpretation of GDPR and to ensure that research and its participants are protected. HELICAL, with its focus on the rare disease Vasculitis, represents a particularly complex area. This is an area where risks of participant identification are greater given the smaller number of cases that are recorded. HELICAL is also a multinational pursuit that must meet regulatory compliance that varies across different member states, and this is no different for derogations within GDPR and the flexibility it offers when meeting requirements around matters such as choice of legal basis and how it interplays with member state laws around ethics and research governance.

This requires consistency with how each of the HELICAL participants approach these challenges and a common approach to managing the educational requirements. Such an approach relies on developing a common policy set around data use to unpick some of the ambiguous or challenging areas of regulatory compliance where sensitive data from a smaller cohort must be processed by novel omics, machine learning and data driven research techniques. This deliverable describes the approach taken to understand the research space at play across HELICAL and to highlight the requirements for policy, achieve the educational goals of the ITN, prepare the Early Stage Researchers for a career steeped in robust and effective sensitive data handling and to assure regulatory compliance in this space.

In this Deliverable we present the challenges that HELICAL would need to address in the regulatory space, describe the approach taken to address them in line with the regulatory tools that GDPR provides, summarise the educational and engagement outreach and provide the policy items that have been developed as a result.

Understanding and meeting the challenges

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.

Firstly, given the number of sites and multiple early stage researchers focusing on various aspects of special category data for a rare disease, a tailored approach was seen to be needed in order to tackle the individual information governance issues for each of the projects. An identified issue which arose is that the clarity of regulatory requirements is difficult to determine where novel data processing techniques (with regards omics processing and potentially machine learning) where the technological setup would need to be mapped to understand the data flows and processing requirements.

Secondly, given each member state as a sovereign nation under independent regulatory jurisdiction provides its own legislation around use of data, research governance, ethics and consent amongst other areas, a careful approach was required so that each HELICAL partner would be able to match the regulatory requirements. This includes the role of consent in the regulatory framework where Ireland and the UK operate under a Common Law Duty of Confidence and Ireland has implemented its own Health Research Regulation that mandates consent is used as the lawful basis for data processing in addition to research participation consent requirements. In other member states a statutory provision exists but they rely on “public task for scientific research” as the legal basis and special category justification for data processing under GDPR where the consent requirements are met to participate in research under separate regulations.

The varied nature of the data, materials and regulatory requirements for HELICAL mandated a trusted and authoritative approach for navigating these. GDPR offers the Data Protection by Design and Default approach where data protection concerns are addressed from the outset of data processing activity. A practical tool mandated by GDPR is the Data Protection Impact Assessment (DPIA) which helps the establish data flows, overarching legislation and compliance with the GDPR principles. The end result is to ensure that appropriate policies and agreements are set up based upon an understanding of the data flows and lawful purposes of the data processing.

However, DPIAs and the overarching design and default approach are relatively new concepts, so an additional challenge was to devise an appropriate format for the DPIA and to ensure that data protection concerns were addressed within the HELICAL context. To that end, The European Institute for Innovation through Health Data (i~HD) adapted a DPIA Template that had been developed by our experts and used in the context of secondary use health data sets.

As part of the HELICAL activities the need for a common framework was identified as a means of understanding potential risks and data protection impacts. By applying data protection by design and default principles, the HELICAL Early Stage Researchers (ESRs) themselves were required to develop DPIAs for their own projects to help them understand the details of their data processing needs, understand the importance of accurate, secure, lawful and ethical data handling for scientific

research and, by making this part of the network-wide training Module 2 delivery, learn how to handle these areas in detail. 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. By working closely with the patient advocate communities through Vasculitis Ireland Awareness (VIA) who were able to join the sessions through Module 2 and provide training to the ESRs on effective public engagement.

By taking this approach HELICAL has been able to adopt an overall risk mitigation strategy through the information gathered through the DPIAs and develop a common policy which is available in Annex 1. Please refer to Deliverable 8.5 for a full treatise on the approach and particulars taken for policy development.

Conclusions

The DPIAs, Educational Materials, Transparency Materials resulting from this work are attached as appendices. A review by the Information Governance Board in March 2021 is proposed for update where appropriate.

Annex 1: Information Governance Policy for HELICAL ITN

Context

This policy describes the overarching approaches that are expected and required of Early Stage Researchers (ESRs) as they handle data within the regulatory requirements of the General Data Protection Regulation (GDPR) and legal requirements for Research Governance.

GDPR provides a basis to develop rich documentary, contractual and policy-based controls to assure compliance with its principles. This policy has therefore been developed on the findings of a series of Data Protection Impact Assessments conducted at a general project level and individual ESR data requirements to achieve their research goals.

Scope

This policy is designed to cover the HELICAL Consortium's work at a high level and has been drafted to be applicable to any ESR and their supervisory teams to work too. This policy does not override or replace any existing institutional policies that the HELICAL Teams may be working to within their host institutions. These policies will exist within the academic institutions where ESRs may be registered to conduct their PhDs, the industrial, public or not for profit organisations that employ the ESRs. The intention therefore is to ensure that this policy works alongside and in an interoperable fashion with existing requirements and institutional governance.

HELICAL will maintain a set of guidance resources outside of this policy to assist HELICAL Team Members to support ESRs and their Supervisors in where to seek advice on Member State Variations to legal bases, retention requirements and best practices and tooling for sharing data sets.

Policy Items

The clauses for this policy are grouped according to main data protection themes as established by the GDPR principles with additional guidance to seek further assistance where needed.

If this policy or practice is unclear

In the first instance the HELICAL team member should consult with their immediate supervisor. Where the supervisor is unable to assist, please refer to Nathan Lea nathan.lea@i-hd.eu and Maria Christofidou maria.christofidou@i-hd.eu at the European Institute for Innovation through Health Data (i~HD).

General Provisions and Conduct

1. At all times HELICAL Team Members will work responsibly and ethically with materials and data collected within their studies.
2. Where team members have any concerns around the security, accuracy, robustness or protection of the data they are using, they will raise it with their supervisors.
3. HELICAL Team Members will at all times use data that is sufficient to achieve their goals, and where appropriate, anonymous.

4. Where anonymous, HELICAL Team Members will nevertheless handle it as sensitive, confidential data, mindful that it came from participants' medical records and / or during their participation in medical research or registries as if it were owed a duty of confidence.
5. HELICAL Team Members will ensure they are familiar with their institutions' standard operating procedures, codes of conduct and data or materials sharing agreements that cover the handling of data.
6. At all times it remains the responsibility of the HELICAL Team Members, their supervisors and their host institutions to ensure they are aware of data breach reporting procedures.
7. HELICAL Team Members will abide by training requirements of their home institutions and any that the HELICAL Project requires of them.

Lawfulness of Purpose

1. HELICAL Team Members will work within the bounds of regulatory permission. This includes work for projects that are ethically approved, run within the limits of that approval and what would be reasonably expected with regards participant consent.
2. Where samples are being processed, this will be in line with secondary legislation within the member state where the samples are held.
3. Where there is uncertainty with regards the first two clauses, clarification from the sample and Biobank curators must be sought prior to any data or materials processing.
4. The purpose for which the samples and / or data are being processed will be to fulfil scientific research in the public interest.
5. The legal bases for the research will be dependent on the member state of the sponsoring research or industrial organisation and local advisory and the local jurisdiction directions will be followed.
6. The purpose shall be limited to that of research in the public interest where any other purpose will need to be assessed for regulatory compliance.

Accountability

1. HELICAL Team Members will ensure that they update their host institutions' information asset registers, records of processing activity and information security records (including authorised personnel).
2. HELICAL Team Members will permit and comply with any audits that their host institutions, data controller partners, and regulatory and competent authorities require in line with statutory provisions, sharing agreements and when directed to do so by the Data Protection and Research Governance Officers.
3. HELICAL Team Members will keep a record of any required deviations from their protocols or risk assessments and seek necessary approvals for these amendments where it remains the responsibility of the Team Members and their supervisors to ensure they monitor these.
4. HELICAL Team Members remain responsible for updating any Data Protection Impact Assessments to account for any changes to data processing.

Data Accuracy, Adequacy and Minimisation

1. HELICAL Team Members will clearly describe, justify and verify all data items they use for their research.
2. Where possible, only anonymous data will be used to conduct the research in line with the Anonymisation and Pseudonymisation clauses.
3. Where required, a pseudonym will be retained by the originating registry, lab or clinic that is linked back to identifiable data by the originating centre.
4. Links to identifiable data will only be used in cases where reidentification is required (say for findings of clinical significance).
5. Links to identifiable data will only be shared with the researchers where there is a lawful basis and with a favourable view of independent ethics committees.
6. HELICAL Team Members will ensure that the data they require is adequate (notwithstanding the anonymity requirement).
7. Where identifiable data is required, this will need to be justified.
8. HELICAL Team Members will ensure the validity of the data and will alert the issuing institution immediately if any error should be discovered by agreed mechanisms as part of sharing agreements.
9. HELICAL Team Members will familiarise themselves with retention requirements (usually for a period of five to twenty years after the completion of a project, depending on the research and participants involved) for their member state jurisdictions and archive data according to local requirements, removing identifiable data as appropriate to limit the amount of time participants may remain identifiable.

Anonymisation and Pseudonymisation

1. Where data sets are to be pseudonymised, HELICAL Team Members will request this of the data controller(s) that hold responsibility for identifiable data.
2. HELICAL Team Members will refrain from accessing or using the linking keys between pseudonyms and the identifiable records unless this has been authorised by an ethics committee and the DPIA has been updated and reassessed.
3. Pseudonyms or other study identity numbers or codes will retain no artefacts or components of identifiable data (i.e. initials, soundex codes, years of birth etc.)
4. Anonymous data exports will be handled in adherence to the Four Eyes Principle[1] or local member state supervisory authority guidance on anonymity and its assurance.

Transparency and FAIRness

1. HELICAL Team Members will publish detailed information leaflets for participants and the wider public.
2. HELICAL Team Members will take every opportunity to engage with patient advocate groups and public engagement events.

[1] https://ec.europa.eu/eurostat/cros/content/four-eyes-principle_en

3. HELICAL Team Members will work with patient advocate groups to ensure their work is clear, meaningful and understandable to the membership and the people they represent.
4. HELICAL Team Members will advise and update their home institutions to update their materials for transparency notices and regulatory requirements.
5. HELICAL Team Members will work to ensure that their materials and outputs abide by the FAIR principles and Open Science.
6. In any and all cases, the needs to protect data will be respected and balanced by adherence to the FAIR principles, Open Science and the interests of the public and where this is unclear, the HELICAL Team Member(s) will revert to their HELICAL colleagues, home institutions Data Protection Officers and other compliance officers (including independent ethical review) where appropriate.

Information and Cyber Security

1. HELICAL Team Members will work in compliance with their institutional information security policies and governance.
2. HELICAL Team Members will make themselves available for information security training and educational activities that are required by their host institutions, supervisors or where the HELICAL Project determines this is required.
3. HELICAL Team Members will register all data assets (including laptops, hard disks, data sets and samples) with their data protection and / or information security offices as appropriate.
4. HELICAL Team Members will abide by strong password advisory from their institutions and where this is not available, they should use passwords of 8-15 characters and mix in capital letters and symbols for their passwords.
5. HELICAL Team Members will operate in line with their institutional “clear desk” policies, ensuring that no materials are left out in the open where all materials should be locked away safely.
6. When working at home, the same care and attention must be applied as if working form office premises, and extra care should be taken with regards family members and online services such as Zoom.
7. For Consortium business, HELICAL Team Members will use the licensed HELICAL Zoom account or other licensed account form their home institutions.
8. Any credentials (usernames, passwords, tokens where appropriate) issued to the HELICAL Team Member will be used only by that Team Member and nobody else.
9. Any software used must be licensed and correctly configured by the host institution prior to use by the ESRs.
10. HELICAL Team Members will encrypt data at rest whether identifiable or anonymous using 256-bit Advanced Encryption Standard encryption where the encryption shall be portions of the local hard disk or encrypted external media.
11. HELICAL Team Members will back up their work using their host institutions’ facilities in line with their policies on backing up research materials, both personal data and otherwise.

12. Where HELICAL Team Members back up their own work, they will, host institution policy permitting, do so with encrypted media and where cloud or networked services are used for this, they must be licensed by the home institution.
13. Where files need to be transferred, this will occur using a secure service that has full encryption and logging facilities, an arrangement with the host institution sharing and / or receiving the files and which may be subject to audit by HELICAL or any of the host institutions. HELICAL recommends use of Microsoft OneDrive for file sharing. Files should not be emailed without full encryption of the email.
14. It remains the responsibility of each HELICAL Team Member to ensure that their equipment is updated regularly, protected by anti-virus software that is fully licensed by their home institutions or themselves where this is especially the case for institutions that operate a "bring your own device policy" where the HELICAL Team Members will act in accordance with their host institutions' policies and direction.