**DeCOmPRESS Data Sharing Plan**

# **A. Preamble**

The COVID-19 pandemic is an ongoing outbreak of coronavirus disease 2019 (COVID-19), caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). Researchers worldwide have quickly mobilised to learn more about disease pathogenesis, clinical course and characteristics, and to develop potential treatments and therapies. A number of academic publishers have committed to support the response to COVID-19 by enabling free and open access to the latest available research, evidence, and data. Unpublished preprints have also been uploaded to free online archives and distribution services ahead of formal peer review for publication.

This new research landscape has highlighted the importance of a comprehensive data sharing strategy for COVID-19-related projects.

The DeCOmPRESS study aims to determine whether immunosuppressant therapy for chronic autoimmune disease protects against the cytokine storm associated with COVID-19 and reduces the severity of the clinical syndrome. This work has the potential to impact treatment protocols and government health advice for the estimated 10% of people in Ireland living with an autoimmune condition. Therefore, rapid and far-reaching dissemination and of reliable data is crucial.

The DeCOmPRESS project team intend to publicly share results and processes as soon as they become available. We welcome input and collaboration from external research groups and will establish robust agreements for data access and sharing in collaboration with the TCD technology transfer office. Patient confidentiality is of paramount importance and will always be protected in accordance with Health Research Regulations. Data will only be made available when the DeCOmPRESS team are satisfied that any academic or commercial interests of our own study would not be unreasonably negatively impacted by its sharing/dissemination.

# **B. Open Data**

The DeCOmPRESS project team strongly support open access and the FAIR data principles ([see Section C](#_C._FAIR_Data_1)). We are committed to the principles of open data, data sharing, reusing data resources and research transparency. However, in view of the use of sensitive patient data, the partners cannot currently commit to an unconditional open access policy, but agree on a policy of “as open as possible, as restricted as necessary”.

Wherever possible DeCOmPRESS will create data sets that are sufficiently de-identified and/or aggregated to enable them to be published and reused. However, two important constraints need to be highlighted for transparency here. Firstly, some of the data accessed and used will be pre-existing data sources (such as the RKD biobank) that have been curated and are being held by other research projects, which will continue to have ownership of those data. Secondly, much of the novel research will be undertaken using patient data. In compliance with national and European laws, we will primarily respect conformance to data protection legislation above any wish to make research data openly accessible.

In situations where potentially useful data are under the custodianship of DeCOmPRESS and cannot be made openly accessible, we will publish a data sharing specification outlining how external research teams may formulate a data sharing request in order to access the data sets (see [Section D](#_C._FAIR_Data_1)). The members of the DeCOmPRESS research group are acutely aware of the sensitivity of personal data and the need for its protection, and have developed an appropriate Data Management Plan (see [Appendix 1](#_Appendix_1:_Data) above and also accessible at: <https://www.tcd.ie/medicine/thkc/decompress/>) to account for this. DeCOmPRESS partners will align with the policies and metadata specifications of the openAIRE pilot (<https://www.openaire.eu/>).

## B.1 Standard Metadata Describing the Data Set

Each data document generated as part of the DeCOmPRESS study will include a metadata descriptor. The Project Manager (PM) will be responsible for compiling, maintaining, and updating metadata for this study. Benefits of this approach include data storage in a common format, quicker and more intuitive querying of the data, consistent combination of diverse data sources, and ease of reporting. Datasets and metadata will be archived using openAIRE. The RKD registry (including the COVID dataset) has been registered with the European Rare Disease Registry Infrastructure (ERDRI, see: <https://eu-rd-platform.jrc.ec.europa.eu/erdri-description_en#inline-nav-1>). This creates an interoperable and machine-readable mechanism for allowing international researchers to view the data dictionary. All datasets will be uplifted into an enriched RDF data model and an ontology will be published in collaboration with the FAIRVASC project. We will adopt minimal metadata standards required for biological data as specified by the Minimum Information for Biological and Biomedical Investigations (MIBBI) standards respectively. Metadata tables will be established and compiled once the dataset is defined in detail.

# **C. FAIR Data Principles**

All data collected in the course of the DeCOmPRESS study will be retained in the secure structures as defined above. Data will be stored in accordance with FAIR principles: findability, accessibility, interoperability, and reusability. This will allow results from this work to be interoperable with international data collection initiatives and maximise the impact of the project. This work will be aligned with the FAIRVASC project, which seeks to merge multiple vasculitis registries across Europe (see [Section F.3](#_F.3_Databases)).

## C.1 Findability

* (Meta)data will be assigned a globally unique and eternally persistent identifier.
* Data will be described with rich metadata.
* (Meta)data will be registered or indexed in a searchable resource.
* Metadata will specify the data identifier.

## C.2 Accessibility

* (Meta)data will be retrievable by their identifier using a standardised communications protocol either through the ERDRI portal, the FAIRVASC user interface or directly exposed on the DeCOmPRESS website.
  + The protocol will be open, free, and universally implementable.
  + The protocol will allow for an authentication and authorisation procedure, where necessary. e.g. the FAIRVASC project will incorporate a credentialing service with password control.
* Metadata will be accessible, even when the data are no longer available.

## C3. Interoperability

* (Meta)data will use a formal, accessible, shared, and broadly applicable language for knowledge representation. In the case of DeCOmPRESS, the study data will be uplifted into the RDF data model, which is considered the standard by the GO-FAIR group.
* (Meta)data will use vocabularies and ontologies that follow FAIR principles.
* (Meta)data will include qualified references to other (meta)data.

## C.4 Reusability

* (Meta)data will have a plurality of accurate and relevant attributes.
  + (meta)data will be released with a clear and accessible data usage license.
  + (meta)data will be associated with their provenance. The proposed RDF data model uses the Prov-O ontology to describe data provenance.
  + (meta)data will meet domain-relevant community standards.

# **D. External Data Access Mechanisms**

Internal access to data documents generated during the DeCOmPRESS project are controlled as outlined in the Data Management Plan. We encourage collaboration with external research groups, and requests for data access will be managed by the Principal Investigator (PI) and PM with consent from other DeCOmPRESS team members. Data access requests should be directed to PI Prof. Mark Little ([mlittle@tcd.ie](mailto:mlittle@tcd.ie)) and PM Ms. Emma Leacy ([leacyej@tcd.ie](mailto:leacyej@tcd.ie)).

## D.1 External Research Group Requests

Requests should include the person’s name, position, institution, a brief description of the project, the intended data use, the list of fields required, and whether the project is intended to be commercialised. The email subject should be entitled “DeCOmPRESS data request –“ as well as the name of the researcher’s project. Under normal circumstances the email will be responded to within 7 days, although this may take longer under certain circumstances. Subsequent correspondence may:

* request further clarifications; or
* grant access to the relevant data where no sensitive data is requested, and no other special considerations are required; or
* otherwise will provide a formal Data Request Form (DRF). This will require that the researcher describes how the data will be used, provides justification for sharing of sensitive fields, and commits to managing the data in an appropriate way on behalf of all project members.

When a DRF is required, the PI and PM (with assent from the DeCOmPRESS study team) will be responsible for confirming the appropriateness of sharing the requested data, and reserve the right to provide only the data they themselves consider justifiable. This decision will normally be made within 30 days of receipt of the form. We expect to share the data in RDF format. Sharing of data in other formats such as comma-separated value (CSV) or extensible markup language (XML) should be stated in the DRF. Only anonymised or pseudonymised data will be shared with external sources. Data will only be made available when the DeCOmPRESS team are satisfied that any academic or commercial interests of our own study would not be negatively impacted by its sharing/dissemination.

## D.2 Data Subject Requests

Patients and healthy controls (Data Subjects) who have contributed to the DeCOmPRESS study may request access to data generated from their participation. Such requests are in compliance with the Data Subject Rights provisions detailed in Chapter 3 of the EU General Data Protection Regulation (GDPR). Data Subjects are advised to use the Trinity College Data Access Request Form (accessible at: <https://www.tcd.ie/info_compliance/foi/request/>) for this purpose. Requests can also be made by contacting the PM Ms. Emma Leacy ([leacyej@tcd.ie](mailto:leacyej@tcd.ie)). For additional information please see the DeCOmPRESS Data Management Plan.

# **E. Intellectual Property and Authorship**

The primary aim of the DeCOmPRESS study is to define the disease course of COVID-19 in immunosuppressed patients. This will inform clinical guidelines for immunocompromised patients and contribute to a wider knowledge base of how COVID-19 affects the immune system. We therefore welcome collaboration from external academic and clinical researchers. We have established an authorship committee and developed two separate policies to address the issues of authorship: one for the primary papers and a separate one for the associated papers. The authorship committee comprises: Mark Little (ML), Lina Zgaga (LZ), Michelle O'Shaughnessy (MOS), Emma Leacy (EL). For this policy and for all decisions on authorship and acknowledgements, the decision of the authorship committee is final and binding.

## E.1 Principles of Authorship

These principles are based on consideration of the following:

* 1. Compliance with the most current version of the International Committee of Medical Journal Editors (ICMJE) “Uniform requirements for manuscripts submitted to Biomedical Journals: Writing and Editing for Biomedical Journals” (www.icmje.org) requires named author to meet all 3 of the following:
* “substantial contributions to conception and design, or acquisition of data, or analysis and interpretation of data”
* Involvement with drafting of the article or “revising it critically for important intellectual content”
* final approval of version to be published
  1. To be as inclusive as feasible. However, complying with Authorship guidelines “drafting the article” or “revising it critically for important intellectual content” becomes increasingly impractical with large numbers.

## E.2 Primary Papers

Primary papers are considered to be those publications which detail the clinical and immunological disease course of COVID-19 in our patient cohort. These papers will conform to the original requirements defined in our funding application for the project. As per HRB Policy on Open Access publications (<https://hrbopenresearch.org/for-authors/article-guidelines>) resulting in whole or in part from HRB-funded research must be deposited in an open access repository. We intend to publish primary papers on the HRB Open Research platform or other similar open access journal.

Each of the main investigators (ML, LZ, MOS) will take senior authorship role in rotation, for all papers. For each paper, the other chief investigators will take 2nd, 3rd etc place in authorship list, unless there is no readily identifiable primary author amongst the group, in which case one of the main investigators will take primary authorship for the work. The primary author (with support from EL) will undertake the main task of preparing the first draft of each paper and help to oversee the editing changes with assistance from the main investigators, primarily the nominated senior author for that paper. If the primary author cannot undertake this task, or there is no suitable primary author, the PM will nominate one of the main investigators to this task with assent from the main investigators. Clinical study site team members including research nurses and Central Pathology Laboratory (CPL) technicians who have been essential to the running of DeCOmPRESS will be listed as authors, with the proviso that they fulfil the role of an author in being involved in reviewing and revising the manuscript in a timely fashion.

## E.3 Associated Papers

Authorship guidelines for associated projects will vary and reflect the nature of the specific work done. As the authorship committee, we retain the right to be authors on all of these associated projects. It is the responsibility of the main investigators to decide how to enforce this for each paper on a case by case basis; in some cases there will be a requirement for all main investigators to be included, or only 1 or 2 of them, or if there are projects where none of the main investigators have made a significant contribution, then we would acknowledge this and none of them would be authors. It is preferable (and likely) that all the associated papers will, however, be based on projects for which the 3 main investigators have provided substantial input. All sub-projects will be discussed and approved by the entire DeCOmPRESS research team before the project is undertaken to ensure appropriate use of DeCOmPRESS patient data. If these projects divert substantially from the DeCOmPRESS mission, they will require assent of the overarching RKD steering committee and approval of the SJH/TUH ethics committee.

We expect the lead investigator for each subproject to liaise with the PM and authorship committee over publication of the paper(s) and for them to propose eligibility and order of authorship. DeCOmPRESS investigators should not expect to be 1st or senior author on any of these papers, unless they are primarily responsible for that sub-project. All associated papers are expected to include investigators and/or core DeCOmPRESS team members who have made substantial contributions to that project as authors for that particular paper. The final decision regarding authorship on all associated papers will rest with the authorship committee.

The PM will be the primary point of contact for subproject investigators with another DeCOmPRESS team member (usually an authorship committee member) allocated as appropriate. These contacts will be responsible for maintaining regular communication with relevant stakeholders and ensuring the quality of associated studies/papers/abstracts. They will also provide feedback on the progress of these studies to the wider DeCOmPRESS team. All abstracts arising from associated projects should be shared with the DeCOmPRESS research team as part of the regular communication process. This should occur at least ten working days before any proposed abstract deadlines to allow sufficient time for review. We expect authorship on abstracts to follow the same rules as for the full papers; exceptions to full lists of authors on abstracts may be made in cases where author names and affiliations substantially count toward word/character limits for the abstract text. The final decision regarding authorship on all abstracts will rest with the authorship committee.

We encourage the use of contributorship with an appendix to acknowledge those involved less directly in the study, but would not expect the full contributorship list which we would apply to primary papers to necessarily apply to each associated paper. However, we do require that all those who have been significantly involved in each paper to be acknowledged appropriately. The HRB funding body will be acknowledged on all manuscripts deriving from the DeCOmPRESS project. The PM (with support from the authorship committee) will take responsibility for providing a full list of personnel who could be named as authors, contributors or appear in the appendix. It will be the responsibility of the senior author and the allocated member of the DeCOmPRESS team for that particular paper, to edit the list for use in the associated paper or papers as required e.g. if particular names or groups of names should be authors, contributors or simply appear in the appendix. The senior author will be responsible for ensuring the accuracy of these lists and also for making sure that the manuscript complies with the manuscript submission requirements of any individual journal.

# **F. Dissemination**

A fundamental requirement of the DeCOmPRESS study is the timely and open-access dissemination of study results. Our grant requirements stipulate that projects funded whole or in part from HRB-funded research must be deposited in an open access repository. We intend to publish progress reports, study documentation, and primary papers on the HRB Open Research platform (<https://hrbopenresearch.org/>) as described in [Section E.2](#_E.2_Primary_Papers). We will produce a preliminary research report at 3 months and a final report at 18 months. The PM & PI will be responsible for study reporting. Audits of data access will be a fundamental part of this reporting procedure. All research outputs will be reported in real time in open access formats and shared with the relevant stakeholders in line with the HRB Joint Statement on sharing research data and findings relevant to the novel coronavirus (COVID-19) outbreak ([https://wellcome.ac.uk/coronavirus-COVID-19/open-data](https://wellcome.ac.uk/coronavirus-covid-19/open-data)). Additional data outputs will be distributed to the appropriate stake holders as described below.

## F.1 Vasculitis Patients

Vasculitis patients who are registered/active users of “patientMpower” will be quickly informed of study outputs via the smartphone app. Additional communication with patients is facilitated through our Patient Advocate partner Julie Power. As Founder and Chairperson of Vasculitis Ireland Awareness – an all-Ireland support group for those affected by any of the Vasculitis diseases – Julie is key to allow direct communication between the research team and study recruits.

## F.2 European Vasculitis Community

Professor Little holds several chair and co-chair positions in European vasculitis societies including Vasculitis Ireland Network (VINE), the UK/Ireland Vasculitis society (UKIVAS), and the rare immune disorder European Reference Network [ERNRITA](http://rita.ern-net.eu/about-rita/governance/).). In collaboration with UKIVAS co-chair Neil Basu and the Irish/UK patient organisations have convened to form a COVID-19 taskforce. The immediate outputs were a vasculitis-focused COVID-19 data dictionary and weekly literature review, circulated to VINE and UKIVAS members. Patient clinical data and outcomes generated from the DeCOmPRESS study will be included in these weekly updates. These updates will also be shared with ERNRITA via website and newsletters, thereby reaching all European expert centres caring for patients with rare immune disorders like vasculitis. Periodic summaries of these updates will also be sent to the World Health Organisation (WHO) via their media contact team.

## F.3 Databases

We will incorporate our results into a FAIR COVID-19 dataset (designed to be interoperable with international data collection initiatives, see [Section C](#_C._FAIR_Data)) and deposition of data in an open science repository will allow ready integration with other studies to maximise the impact of this project. We will publicise results in open-source preprint formats such as medRxiv/bioRxiv and publish protocols and results in HRB-Open. Datasets and metadata will be archived using openAIRE to the appropriate standards (see [Section B.1](#_B.1_Standard_Metadata)). Datasets generated by this project will be made available to other researchers in a manner that respects patient privacy using OpenAIRE services. Appropriate dataset sharing will be achieved with support from the the European Institute for Innovation through Health Data.

## F.4 Clinical Guidelines

We will share the COVID-19 data dictionary and advocate that all researchers in Europe studying COVID-19 in vasculitis patients use this dataset so that results can be readily combined. We will inform the HPSC and NPHET and Chief Medical officer, to inform their guidance to vulnerable groups (including autoimmune disease). We will also work with HPRA to develop guidance statements for healthcare professionals on COVID-19 management in immunosuppressed patients. The DeCOmPRESS patient data summaries will be shared with the Irish healthcare community and the public health infrastructure in Ireland via the appropriate reporting channels.

## F.5 Marketing & Communications

The dedicated DeCOmPRESS study page is accessible from the THKC Renal Inflammation Group webpage and through the link: <https://www.tcd.ie/medicine/thkc/decompress/>. Updates and information relevant to the study will be shared on social media through the DeCOmPRESS Twitter page ([@DecompressStudy](https://twitter.com/DecompressStudy)). The PM is responsible for maintaining the Twitter profile.