

**National Research Data Infrastructure for
Personal Health Data – NFDI4Health**
Consortia Progress Report Part 1 **(B-1)**



Cologne, 29th September 2023

Table of contents

1	General Information	3
2	Summary	4
3	Composition of the consortium	6
A.	Acronyms and abbreviations.....	13

1 General information

- *Name of the consortium*
National Research Data Infrastructure for Personal Health Data (NFDI4Health)
- *Research domains or research methods addressed by the consortium*
Our highly transdisciplinary consortium (medicine, epidemiology, law, informatics, statistics) consists of (1) data holding organisations (high IT standards, established use and access procedures, high data protection standards) with long-term experience in the conduct of clinical and epidemiological studies as well as public health surveys, in research data management and in analysing personal health data; (2) infrastructural institutions with expertise in standards and services; and (3) institutions with a clear focus on the involvement of our user communities.

All co-applicants have outstanding expertise in education and training and are integrated into a broad network of (inter)national experts/institutions.

Related DFG review boards

205-01 Epidemiology, Medical Biometry/Statistics

205-02 Public Health, Health Services Research, Social Medicine

205-05 Nutritional Sciences

205-07 Medical Informatics and Medical Bioinformatics

205 Clinical Trials (These refer to all medical subjects listed under review board 205.)

- *URL of the consortium website and repositories used for publishing output:*
<https://www.nfdi4health.de/>

2 Summary

The overall goal of NFDI4Health is to support clinical and epidemiological researchers in sharing their data with the user community in a privacy preserving and ethical manner. In accordance with the FAIR principles, NFDI4Health creates new opportunities for data analyses that will improve population health. The main objectives of NFDI4Health are to (1) make structured health data from clinical and epidemiological studies, disease registries, administrative health databases and health reporting in Germany findable and accessible; (2) implement a framework for centralised search and access to existing decentralised infrastructures for health data; (3) facilitate data sharing, record linkage and harmonised data quality assessment; (4) enable collaborative (federated) analyses of personal health data; (5) enable the development and use of new machine-readable consent mechanisms and innovative data access services; (6) support collaboration between clinical research, epidemiology and public health; (7) promote interoperability of currently fragmented IT solutions for metadata repositories, cohort searching, data quality and harmonisation.

In order to achieve the above objectives, NFDI4Health has successfully developed a Research Data Management (RDM) strategy that meets the needs of our user communities comprising data holding organisations providing the data and data users requesting their use. Our RDM strategy simultaneously accounts for the sensitivity of personal health data. In detail:

To achieve central findability of health data, NFDI4Health has established a workflow to make sensitive data visible. This includes the publication of (a) metadata describing health datasets and their resources and of (b) metadata at the data element level, e.g., in the form of study-specific data catalogues. To this end, (1) the NFDI4Health metadata schema (MDS) for health studies, (2) publication guidelines incl. licensing guidelines, and (3) the German Central Health Study Hub (Health Study Hub) aiming for a comprehensive coverage of health data have been developed. First versions of these components were released early in response to the COVID-19 pandemic and with additional DFG funding for the NFDI4Health Task Force COVID-19. Thus, the Health Study Hub already provides a unique overview of COVID-19 studies conducted in Germany. The MDS and the Health Study Hub services are continuously being developed further, based on needs/feedback from the community. To achieve a comprehensive coverage of relevant health data, metadata from (1) clinical trials, (2) epidemiological studies, (3) an administrative health database and (4) cancer registries have been integrated into the Health Study Hub. These metadata will feed a central access point for data requests which is currently being prepared in close collaboration with the Medical Informatics Initiative (MII) to ensure a uniform, centralised and interoperable solution for all German health data infrastructures. Based on the MII software

of the Research Data Portal Health, a pilot version of the NFDI4Health central access point will become available in 2024.

Given the high sensitivity of personal health data, a federated data analysis strategy has been developed as an integral part of our RDM strategy. By implementing the technical infrastructure at the data holding organisations, we support the effort required to create, maintain and share their valuable databases. For this purpose, we provide a dedicated software application, the Local Data Hub, which provides an RDM solution that can be managed locally while connecting to central NFDI4Health services and distributed data analysis frameworks such as DataSHIELD and Personal Health Train. This overall approach has been put into practice in collaboration with all NFDI4Health task areas and in close exchange with our user communities. Here, our use cases provide blueprints for upscaling by implementing NFDI4Health services and standards for major German data sources. In order to intensify the interaction with our user communities, we have so far launched two project calls granted by our flex funds. The first call addresses the interoperability between NFDI4Health and the German Centers for Health Research infrastructure and the second call initiates Citizen Science projects in the context of NFDI4Health.

In order to realise the services mentioned above, NFDI4Health has comprehensively evaluated the data protection requirements for data findability and reusability, integrated them into the overall workflow of our use cases, and published corresponding guidelines. Methods for anonymisation, synthesisation and re-identification risk assessment have been reviewed and critically appraised for our use cases. For a broad dissemination of interoperable standards for health data, NFDI4Health has published a FAIRsharing collection. Use of these standards is supported by tools provided by NFDI4Health such as a terminology service or an annotation service. Another service is the software dataquiR to assess the quality of study data.

We also engage directly with our user communities by raising awareness and supporting capacity building for RDM and NFDI4Health-specific services: (1) We organise workshops and hands-on virtual and on-site training sessions. (2) We conduct user research on the usability and functionality of our services. (3) We are heavily involved in the Bremen interdisciplinary graduate training programme Data Train to educate a new generation of data scientists.

Overall, NFDI4Health is well integrated in national and international activities to establish the German and European Health Data Space (EHDS, GHDS). Formally mandated by the NFDI Directorate, NFDI4Health acts as a key intermediary between the NFDI and external health infrastructures. As part of our political commitment, we compiled a White Paper to inform decision makers on the needs for record linkage of personal health data in close collaboration with major actors in this field.

3 Composition of the consortium

- *Applicant institution*

Applicant institution	Location	Duration
ZB MED – Information Centre for Life Sciences	Gleueler Straße 60, 50931 Cologne	10/20 – now

- *Spokesperson*

Spokesperson	Institution, location	Duration
Prof. Dr. Juliane Fluck https://orcid.org/0000-0003-1379-7023	ZB MED, Katzenburgweg 1a, 53115 Bonn	10/20 – now

- *Co-applicant institutions*

Co-applicant institutions	Location	Duration
Leibniz Institute for Prevention Research and Epidemiology – BIPS	Achterstraße 30, 28359 Bremen	10/20 – now
Berlin Institute of Health at Charité (Charité/BIH)	Charitéplatz 1, 10117 Berlin	10/20 – now
German Institute for Human Nutrition Potsdam-Rehbrücke (DIfE)	Arthur-Scheunert-Allee 114 – 116, 14558 Nuthetal	10/20 – now
Fraunhofer-Gesellschaft: Fraunhofer Institute for Applied Information Technology (FIT); Fraunhofer Institute for Digital Medicine (MEVIS); Fraunhofer Institute for Algorithms & Scientific Computing (SCAI)	Hansastraße 27 c, 80686 Munich	10/20 – now
Heidelberg Institute for Theoretical Studies gGmbH (HITS)	Schloss-Wolfsbrunnenweg 35, 69118 Heidelberg	10/20 – now
Network of Coordinating Centres for Clinical Trials e.V. (KKSNN)	Alt-Moabit 96A, 10559 Berlin	08/22 – now
Max Delbrück Center for Molecular Medicine in the Helmholtz Association (MDC)	Robert-Rössle-Straße 10, 13125 Berlin	10/20 – now
Robert Koch Institute (RKI)	Nordufer 20, 13353 Berlin	10/20 – now
Technology, Methods, and Infrastructure for Networked Medical Research (TMF)	Charlottenstraße 42, 10117 Berlin	10/20 – now
University of Bonn (U Bonn)	Regina-Pacis-Weg 3,	10/20 – now

	53113 Bonn	
University of Augsburg (U Augsburg)	Universitätsstraße 2, 86159 Augsburg	09/22 - now
University of Bremen (U Bremen)	Bibliothekstraße 1, 28359 Bremen	10/20 – 08/22
University of Cologne (U Cologne)	Albertus-Magnus-Platz, 50923 Cologne	10/20 – now
University Hospital Cologne (UKK)	Kerpener Str. 62, 50937 Cologne	08/22 – now
University of Leipzig (U Leipzig)	Ritterstraße 26, 04109 Leipzig	10/20 – now
University of Applied Sciences Mittweida (UAS Mittweida)	Technikumplatz 17, 09648 Mittweida	10/20 – now
Georg-August University Göttingen, University Medical Center (UM Göttingen)	Von-Siebold-Straße 3, 37075 Göttingen	10/20 – now
University Medicine Greifswald (UM Greifswald)	Fleischmann Str. 8, 17475 Greifswald	10/20 – now

- Co-spokespersons

Co-spokespersons	Institution, location	Task area(s)	Duration
Prof. Dr. Wolfgang Ahrens https://orcid.org/0000-0003-3777-570X	BIPS, Bremen	TA2, 3, 5, 6	10/20 – now
Prof. Dr. Iris Pigeot https://orcid.org/0000-0001-7483-0726	BIPS, Bremen	TA1, 2, 3, 4, 5, 6	10/20 – now
Prof. Dr. Hajo Zeeb https://orcid.org/0000-0001-7509-242X	BIPS, Bremen	TA2, 3, 4, 5	10/20 – now
Prof. Dr. Fabian Prasser https://orcid.org/0000-0003-3172-3095	Charité/BIH, Berlin	TA6	10/20 – now
Prof. Dr. Sylvia Thun https://orcid.org/0000-0002-3346-6806	Charité/BIH, Berlin	TA1, 2	10/20 – now
Prof. Dr. Matthias Schulze https://orcid.org/0000-0002-0830-5277	DIfE, Nuthetal	TA1, 2, 3, 5, 6	10/20 – now
Prof. Dr. Oya Beyan https://orcid.org/0000-0001-7611-3501	Fraunhofer FIT, Aachen	TA1, 3, 6	10/20 – now

	UKK, Cologne		08/22 – now
Prof. Dr. Holger Fröhlich https://orcid.org/0000-0002-5328-1243	Fraunhofer SCAI, St. Augustin	TA1, 6	10/20 – now
Prof. Dr.-Ing. Horst Hahn https://orcid.org/0000-0001-7512-5762	Fraunhofer MEVIS, Bremen	TA1, 2, 5	10/20 – now
Dr. Martin Golebiewski https://orcid.org/0000-0002-8683-7084	HITS, Heidelberg	TA 2, 3, 4	10/20 – now
PD Dr. Wolfgang Müller https://orcid.org/0000-0002-4980-3512	HITS, Heidelberg	TA1, 2, 3, 4	10/20 – now
PD Dr. Sebastian Klammt https://orcid.org/0000-0001-7852-4769	KKSN, Berlin	TA1, 2, 4, 5	10/20 – now
Prof. Dr. Tobias Pischon https://orcid.org/0000-0003-1568-767X	MDC, Berlin	TA1, 2, 3, 5	10/20 – now
Prof. Dr. Lothar H. Wieler https://orcid.org/0000-0003-4365-3717	RKI, Berlin	TA 4, 5	10/20 – 03/23
PD Dr. Linus Grabenhenrich https://orcid.org/0000-0002-9300-6625	RKI, Berlin	TA1, 2, 3, 4, 5, 6	10/20 – now
Juliane Gehrke	TMF, Berlin	TA2, 4, 6	10/20 – now
Sebastian C. Semler https://orcid.org/0000-0001-8547-6785	TMF, Berlin	TA1, 2, 4, 6	10/20 – now
Prof. Dr. Ute Nöthlings https://orcid.org/0000-0002-5789-2252	U Bonn, Bonn	TA1, 4, 5	10/20 – now
Prof. Dr. Benedikt Buchner https://orcid.org/0000-0003-1438-5127	U Bremen, Bremen U Augsburg, Augsburg	TA1, 6	10/20 – 07/22 08/22 – now
Ralf Depping https://orcid.org/0000-0003-2537-7229	U Cologne, Cologne	TA2, 4	10/20 – now
Dr. Jens Dierkes https://orcid.org/0000-0002-0121-9261	U Cologne, Cologne	TA1, 4	10/20 – now

Dr. Hubertus Neuhausen https://orcid.org/0000-0001-5223-9177	U Cologne, Cologne	TA1, 4	10/20 – now
Dr. Oana Brosteanu https://orcid.org/0000-0002-8706-217X	U Leipzig, Leipzig	TA5	10/20 – now
Matthias Löbe https://orcid.org/0000-0002-2344-0426	U Leipzig, Leipzig	TA2, 4, 5	10/20 – now
Prof. Dr. Markus Löffler https://orcid.org/0000-0003-2513-1317	U Leipzig, Leipzig	TA1, 2, 3, 4, 5	10/20 – 03/23
Dr. Frank Meineke https://orcid.org/0000-0002-9256-7543	U Leipzig, Leipzig	TA3	10/20 – now
Prof. Dr.-Ing. Toralf Kirsten https://orcid.org/0000-0001-7117-4268	UAS Mittweida, Mittweida U Leipzig, Leipzig	TA1, 3 TA1, 2, 3, 4, 5, 6	10/20 – now 04/23 – now
Dr. Harald Kusch https://orcid.org/0000-0002-9895-2469	UM Göttingen, Göttingen	TA3, 4, 6	10/20 – now
Prof. Dr. Ulrich Sax https://orcid.org/0000-0002-8188-3495	UM Göttingen, Göttingen	TA1, 2, 3, 4, 6	10/20 – now
Prof. Dr. Carsten Oliver Schmidt https://orcid.org/0000-0001-5266-9396	UM Greifswald, Greifswald	TA1, 2, 3, 4, 5, 6	10/20 – now
Prof. Dr.-Ing. Dagmar Waltemath https://orcid.org/0000-0002-5886-5563	UM Greifswald, Greifswald	TA2, 3, 4	10/20 – now
Johannes Darms https://orcid.org/0000-0001-5809-2276	ZB MED, Cologne	TA3	10/20 – now
Birte Lindstädt https://orcid.org/0000-0002-8251-1597	ZB MED, Cologne	TA2, 4	10/20 – now

- *Participants*

Participating institutions (contact person)	Location	Duration
University of Bielefeld, Center for Biotechnology (CeBiTec) (Prof. Dr. Alfred Pühler)	Bielefeld	10/20 – now
University Hospital of Regensburg, Center for Clinical Studies Regensburg (Prof. Dr. Michael Koller)	Regensburg	10/20 – now
University of Duisburg-Essen, Center for Clinical Trials	Essen	10/20 – now

(Prof. Dr. Andreas Stang)		
University Hospital of Münster, Center for Clinical Trials (Prof. Dr. Andreas Faldum)	Münster	10/20 – now
University Hospital of Cologne, Clinical Trials Centre (Prof. Dr. Oliver A. Cornely/Alexandra Nieß)	Cologne	10/20 – now
University Hospital of Freiburg, Clinical Trials Unit (Dr. Britta Lang)	Freiburg	10/20 – now
University of Düsseldorf, Coordination Center for Clinical Trials Düsseldorf (Henrike Kolbe)	Düsseldorf	10/20 – now
University of Magdeburg, Coordination Center for Clinical Studies in Magdeburg (Dr. Antje Wiede)	Magdeburg	10/20 – now
German Institute for Standardisation (DIN) (Christoph Winterhalter)	Berlin	10/20 – now
Federal Centre of Health Education (BZgA) (Prof. Dr. Martin Dietrich)	Cologne	10/20 – now
German Association for Medical Informatics, Biometry and Epidemiology (GMDS) (Beatrix Behrendt)	Cologne	10/20 – now
German Association of Medical Faculties (MFT) (Dr. Frank Wissing)	Berlin	10/20 – now
German Cancer Research Center (DKFZ), Division of Cancer Epidemiology (Prof. Dr. Rudolf Kaaks)	Heidelberg	10/20 – now
German Consortium of Hereditary Breast and Ovarian Cancer (Prof. Dr. Rita Schmutzler)	Cologne	10/20 – now
German Data Forum (RatSWD) (Lisa Kühn)	Berlin	10/20 – now
Federal Institute for Drugs and Medical Devices (BfArM) (Prof. Dr. Karl Broich)	Bonn	10/20 – now
German Network for Evidence-based Medicine (DNEbM)	Berlin	10/20 – now
German Nutrition Society (DGE) (Dr. Kiran Virmani)	Bonn	10/20 – now

German Public Health Association (DGPH)	Bremen	10/20 – now
German Region of the International Biometric Society (IBS-DR)	Hannover	10/20 – now
German Society for Drug Utilization Research and Drug Epidemiology (GAA)	Dresden	10/20 – now
German Society for Epidemiology (DGEpi) <ul style="list-style-type: none"> AGENS – Working group on the collection and use of secondary data 	Ulm	10/20 – now
German Society for Occupational and Environmental Medicine (DGAUM)	Munich	10/20 – now
German Society of Medical Sociology (DGMS) (Christoph Kowalski)	Hannover	10/20 – now
German Society of Social Medicine and Prevention (DGSMP) (Edmund Fröhlich)	Berlin	10/20 – now
Medical Center of the Johannes Gutenberg University Mainz, Gutenberg Health Study (Prof. Dr. Philipp Wild)	Mainz	10/20 – now
Authority for Science, Research, Equality and Districts (BWFGB), Hamburg Cancer Registry (Dr. Alice Nennecke)	Hamburg	10/20 – now
University of Hamburg (UHH), Hamburg City Health Study (Prof. Dr. Stefan Blankenberg, Dr. Annika Jagodzinski)	Hamburg	10/20 – now
University Hospital Essen, Heinz Nixdorf Recall and Multigeneration Study, Institute for Medical Informatics, Biometry and Epidemiology (IMIBE) (Prof. Dr. Andreas Stang)	Essen	10/20 – now
Ruhr-University Bochum, Institute for Prevention and Occupational Medicine of the German Social Accident Insurance Institute (IPA) (Prof. Dr. Thomas Behrens)	Bochum	10/20 – now
University of Münster, Institute of Biostatistics and Clinical Research (Prof. Dr. Andreas Faldum)	Münster	10/20 – now

Helmholtz Zentrum München (HMGU), Institute of Epidemiology (Prof. Dr. Annette Peters)	Munich	10/20 – now
Martin Luther University Halle-Wittenberg, Institute of Medical Epidemiology, Biometrics and Informatics (IMEBI) (Prof. Dr. Rafael Mikolajczyk)	Halle (Saale)	10/20 – now
Otto-von-Guericke University Magdeburg, Institute of Social Medicine and Health Economics, Faculty of Medicine (Prof. Dr. Christian Apfelbacher)	Magdeburg	10/20 – now
Leibniz Institute for Resilience Research (LIR) (Prof. Dr. Klaus Lieb, Prof. Dr. Beat Lutz)	Mainz	10/20 – now
Max Rubner-Institute (MRI) (Prof. Dr. Pablo Steinberg)	Karlsruhe	10/20 – now
University of Bielefeld, Medical School OWL (Prof. Dr. Claudia Hornberg)	Bielefeld	10/20 – now
mediStatistica (Dr. Burkhard Haastert)	Neuenrade	10/20 – now
NAKO Health Study, German National Cohort (Prof. Dr. Henry Völzke)	Heidelberg	10/20 – now
OFFIS Institute for Information Technology (Prof. Dr. Susanne Boll-Westermann)	Oldenburg	10/20 – now
Carl von Ossietzky-University Oldenburg, School of Medicine and Health Sciences (Prof. Dr. Hans Gerd Nothwang)	Oldenburg	10/20 – now
Society of Epidemiological Cancer Registries in Germany (GEKID)	Lübeck	10/20 – now
University of Greifswald, University Computing Centre (Prof. Dr. Katharina Riedel)	Greifswald	10/20 – now
Hamburg University of Applied Sciences (HAW Hamburg) (Prof. Dr. Micha Teuscher)	Hamburg	01/23 – now

Participating individuals	Institution, location	Duration
Prof. Dr. Ludwig Kuntz	Seminar for Business Administration and Healthcare Management, University of Cologne, Cologne	10/20 – now
Markus Brechtel	University Hospital Cologne, Cologne	01/22 – now
Sabine Hanß	University of Göttingen, Göttingen	01/22 – now
Prof. Dr. Lothar H. Wieler	Hasso-Plattner-Institut, Brandenburg	04/22 – now
Dr. Nadine Scholten	University of Cologne, CoRe-Net (BMBF Infrastructure Network), Cologne	10/20 – now
Ingo Meyer	University of Cologne, PMV Research Group	10/20 – now
Prof. Dr. Oliver A. Cornely	University Hospital of Cologne, Clinical Trials Centre, Cologne	10/20 – now
Alexandra Nieß	University Hospital of Cologne, Clinical Trials Centre, Cologne	10/20 – now
Prof. Dr. Petra Ritter	Charité Universitätsmedizin Berlin, Brain Simulation Section at the Department Neurology, Berlin	10/20 – now
Dr. Burkhard Haastert	mediStatistica, Neuenrade	10/20 – now
Prof. Dr. Claudia Hornberg	University of Bielefeld, Medical School OWL, Bielefeld	10/20 – now
Dr. Annika Jagodzinski	University of Hamburg (UHH), Hamburg City Health Study, Hamburg	10/20 – now
Prof. Dr. Rudolf Kaaks	German Cancer Research Center (DKFZ), Division of Cancer Epidemiology, Heidelberg	10/20 – now
Prof. Dr. Ludwig Kuntz	Seminar for Business Administration and Healthcare Management, University of Cologne, Cologne	10/20 – now
Prof. Dr. Klaus Lieb	Leibniz Institute for Resilience Research (LIR), Mainz	10/20 – now
Prof. Dr. Beat Lutz	Leibniz Institute for Resilience Research (LIR), Mainz	10/20 – now

Prof. Dr. Rafael Mikolajczyk	Martin Luther University Halle-Wittenberg, Institute of Medical Epidemiology, Biometrics and Informatics (IMEBI), Medical Faculty, Halle (Saale)	10/20 – now
Prof. Dr. Hans Gerd Nothwang	Carl von Ossietzky-University Oldenburg, School of Medicine and Health Sciences, Oldenburg	10/20 – now
Prof. Dr. Alfred Pühler	University of Bielefeld, Center for Biotechnology (CeBiTec), Bielefeld	10/20 – now
Prof. Dr. Rita Schmutzler	German Consortium of Hereditary Breast and Ovarian Cancer, Cologne	10/20 – now
Prof. Dr. Pablo Steinberg	Max Rubner-Institute (MRI), Karlsruhe	10/20 – now
Prof. Dr. Philipp Wild	Medical Center of the Johannes Gutenberg University Mainz, Gutenberg Health Study, Mainz	10/20 – now
Dr. Frank Wissing	German Association of Medical Faculties (MFT), Berlin	10/20 – now
Prof. Dr. Friedemann Paul	Experimental and Clinical Research Center (ECRC), Max Delbrück Center, Berlin	01/23 – now
Dr. Stefan Hentschel	Authority for Science, Research, Equality and Districts (BWFGB), Hamburg Cancer Registry, Hamburg	10/20 – now
Dr. Steffen Heß	Federal Institute for Drugs and Medical Devices (BfArM), Bonn	10/20 – now
Rüdiger Marquardt	German Institute for Standardisation (DIN), Berlin	10/20 – now
Dr. Marie Standl	Helmholtz Zentrum München (HMGU), Institute of Epidemiology, Munich	10/20 – now
Dr. Heidrun M. Thaiss	Federal Centre of Health Education (BZgA), Cologne	10/20 – now
Prof. Dr. Johanna Eleonore Weber	University of Greifswald, Greifswald	10/20 – now

A. Acronyms and abbreviations

DFG	German Research Foundation
ECRC	Experimental and Clinical Research Center
EHDS	European Health Data Space
FAIR	Findability, accessibility, interoperability, and reusability
GHDS	German Health Data Space
IT	Information Technology
MDS	Metadata Schema
MII	Medical Informatics Initiative
NAKO	German National Cohort
NFDI	National Research Data Infrastructure
NFDI4Health	National Research Data Infrastructure for Personal Health Data
RDM	Research Data Management
TA	Task Area
USB	University and City Library of Cologne
ZKS	Clinical Trial Centre, Leipzig