

FAIR workshop

ASAP, 28 March 2023



What is FAIR & why you should care

Ronald Cornet





Findable



Accessible



Interoperable



Reusable



Findable



Accessible

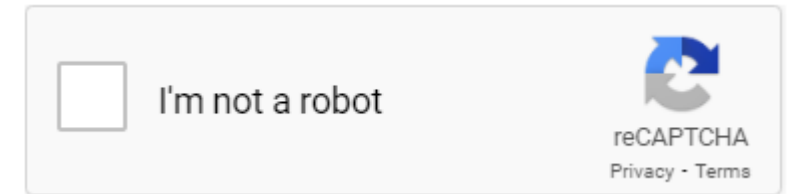


Interoperable



Reusable

For:



Schultes, Erik. (2021). Official GO FAIR Foundation icons for the Three-Point FAIRification Framework (Version 1). Zenodo. Image: doi.org/10.5281/zenodo.4678332
Creative Commons Attribution Share Alike 4.0 International



- Find & use
- Metadata
- Infrastructure
- Comm. Protocol
- Vocabularies
- Provenance



15 FAIR Principles

- Detailed description: <https://www.go-fair.org/fair-principles/>
- FAIR Principles = “What”, not “how”
- Enabling paradigm shift from data sharing to **data visiting / availability**





Image by [Clker-Free-Vector-Images](#) from [Pixabay](#)



Image by [OpenClipart-Vectors](#) from [Pixabay](#)



Why you should care

1. Strategy

Amsterdam UMC - Strategische doelstellingen voor 2021 -2025

Zorg

Wij leveren zorg van aantoonbare en persoonlijke waarde

Van iedere patiënt leren wij voor de zorg van morgen

Onze zorg is muurloos

Wij verkleinen de gezondheidsverschillen tussen Amsterdammers en in de wereld

Onderzoek

Wij gebruiken onderzoek en data om de zorg en gezondheid in de wereld te verbeteren

Met onze universiteiten onderzoeken we de interactie tussen molecuul, patiënt en populatie

Wij investeren in talentvolle teams

Wij waarderen onderzoek op maatschappelijke impact

Onderwijs & Opleiden

We leiden op tot professionals die aantoonbaar samenwerken en zorgzaam en initiatiefrijk zijn

Wij maken gebruik van de kracht van onze beide universiteiten

Wij leiden op *met de regio, voor de wereld*

Al onze mensen blijven een leven lang leren en opleiden



Wij zijn **ZORGZAAM** voor onze patiënten, studenten, collega's, voor de maatschappij en voor de wereld



Alleen SAMEN met beide universiteiten, partners, de regio en in teams kunnen wij Amsterdam UMC zijn



DATA EN DIGITALISERING maken wat wij doen steeds beter

Organisatie

Wij zijn Amsterdam UMC

Amsterdam UMC inspireert al zijn medewerkers

Onze leidinggevenden verbinden, nemen verantwoordelijkheid, tonen veranderkracht en zijn vrolijk

Duurzaamheid is de leidraad in ons denken en doen

In al ons werk gebruiken en delen we data



Why you should care

1. Strategy
2. Money
 - “we found that the annual cost of not having FAIR research data costs the European economy at least €10.2bn every year.”
[Cost of not having FAIR research data \(europa.eu\)](#)

ZonMw's procedures voor datamanagement en datastewardship zijn erop gericht om data te creëren die FAIR zijn, dat wil zeggen vindbaar, toegankelijk, koppelbaar en herbruikbaar zijn (Findable, Accessible, Interoperable and Reusable). Dit beleid is erop gericht om bij te dragen aan onderzoek van hoge kwaliteit binnen én buiten de projecten gefinancierd door ZonMw.

Lees meer over de volgende onderwerpen op onze Engelse website:

- [FAIR data en datamanagement](#)
- [Datamanagement in jouw onderzoeksproject](#)
- [Achtergrondinformatie](#)
- [Het FAIRificeren van data](#)
- [Open Science in COVID-19 onderzoek](#)

FAIR data en datamanagement is onderdeel van het ZonMw beleidsspeerpunt [Open Science – FAIR data](#)

Wij communiceren over FAIR-data met name in het Engels omdat Nederlandse onderzoekers ook in het buitenland werken en omdat buitenlandse onderzoekers aan Nederlandse onderzoeksprojecten meewerken.



Why you should care

1. Strategy
2. Money
3. Time



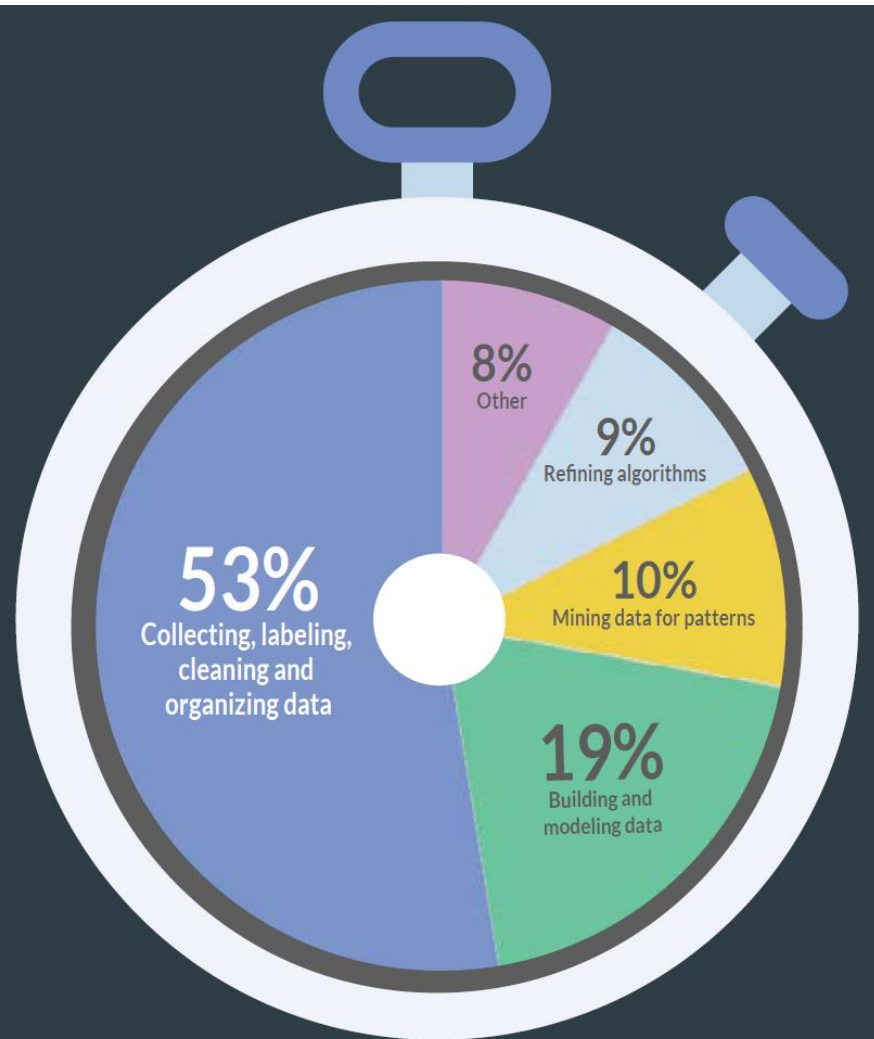
The New York Times

For Big-Data Scientists, 'Janitor Work' Is Key Hurdle to Insights

- Data scientists, according to interviews and expert estimates, **spend from 50 percent to 80 percent of their time** mired in this more mundane labor of collecting and preparing unruly digital data, before it can be explored for useful nuggets.



What activity takes up most of your time?



Data Scientist Report 2017 - CrowdFlower




Why you should care

1. Strategy
2. Money
3. Time
4. Sustainability
 - Burden for clinicians, trial nurses, patients
 - Existing data without research capacity
 - Research capacity without existing data



FAIR Research Data Management

- 
1. Reuse existing data as much as possible
 2. Reuse existing metadata as much as possible (schemas, vocabularies, data dictionaries, ...)
 3. Think about and implement making data FAIR **before you start** collecting
 4. Make existing data FAIR



<https://orcid.org/0000-0002-1704-5980>
<https://www.wikidata.org/wiki/Q44402672>
<https://www.linkedin.com/in/ronaldcornet/>
<http://www.isni.org/0000000352294699>
<https://dblp.org/pid/79/6471>
<info:eu-repo/dai/nl/292976100>

Ronald Cornet, PhD | Associate professor ; Principal Investigator ;
Principal Educator
Medical Informatics
Amsterdam Public Health Research Institute
Location AMC | J1b-115 | Meibergdreef 15, 1105 AZ Amsterdam
T: +31 20-566 52 05 | E: r.cornet@amsterdamumc.nl
kik.amc.nl/home/rcornet | www.amsterdamumc.nl | www.amc.nl



FAIR workshop

ASAP, 28 March 2023





The HELIUS study & FAIR data

Dr. Henrike Galenkamp

Dept of Public and Occupational Health

Support4Research meeting 24 January 2023





Background

- HELIUS: HEalthy LIfe in an Urban Setting
- Aim: to get insight in (the causes of) the unequal burden of diseases across ethnic groups
- Initiated in 2010 by AMC together with Public Health Service (GGD) of Amsterdam
- Cohort study among 6 ethnic groups (aged 18-70y) living in Amsterdam:

Dutch

Turkish

African Surinamese

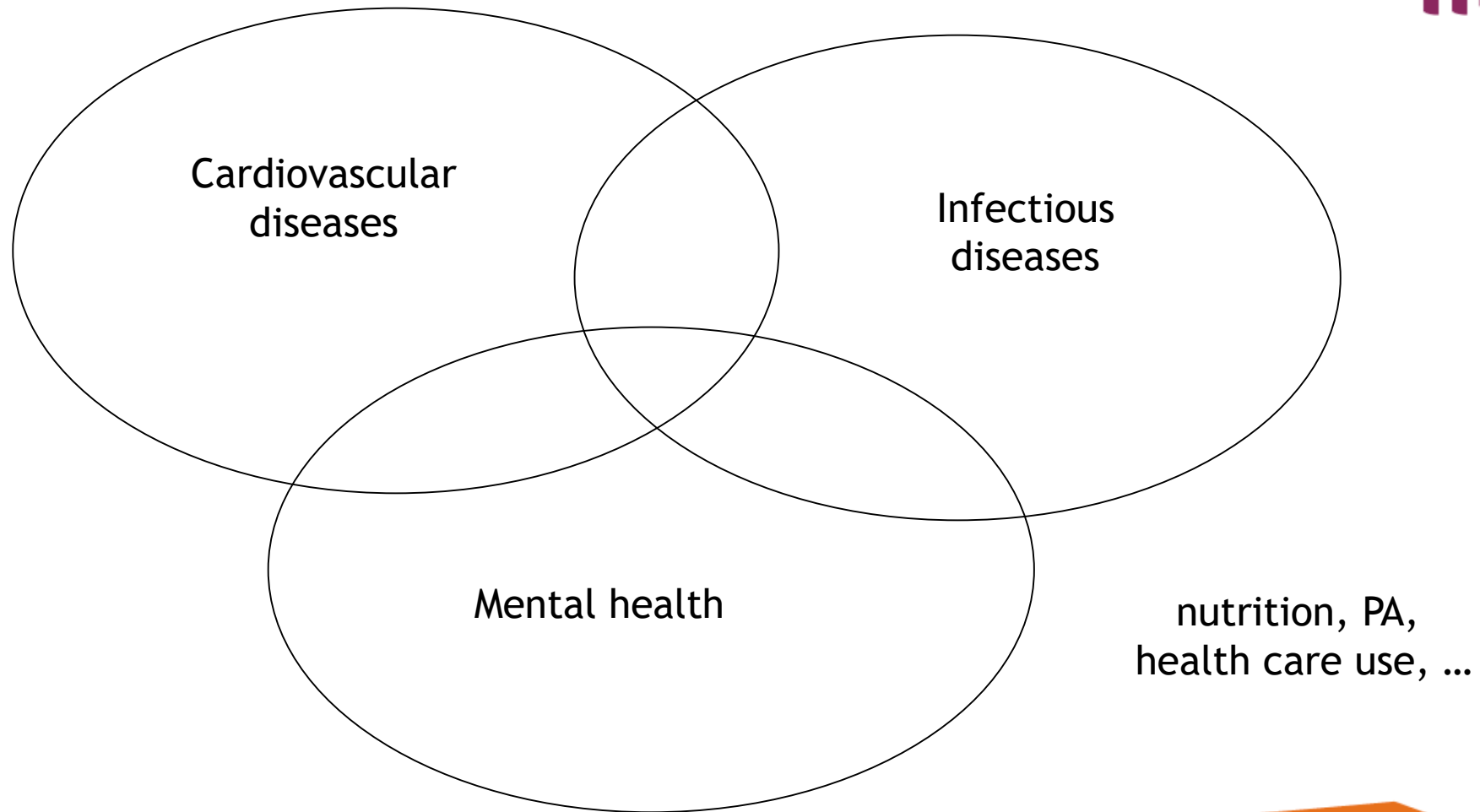
Moroccan

South-Asian Surinamese

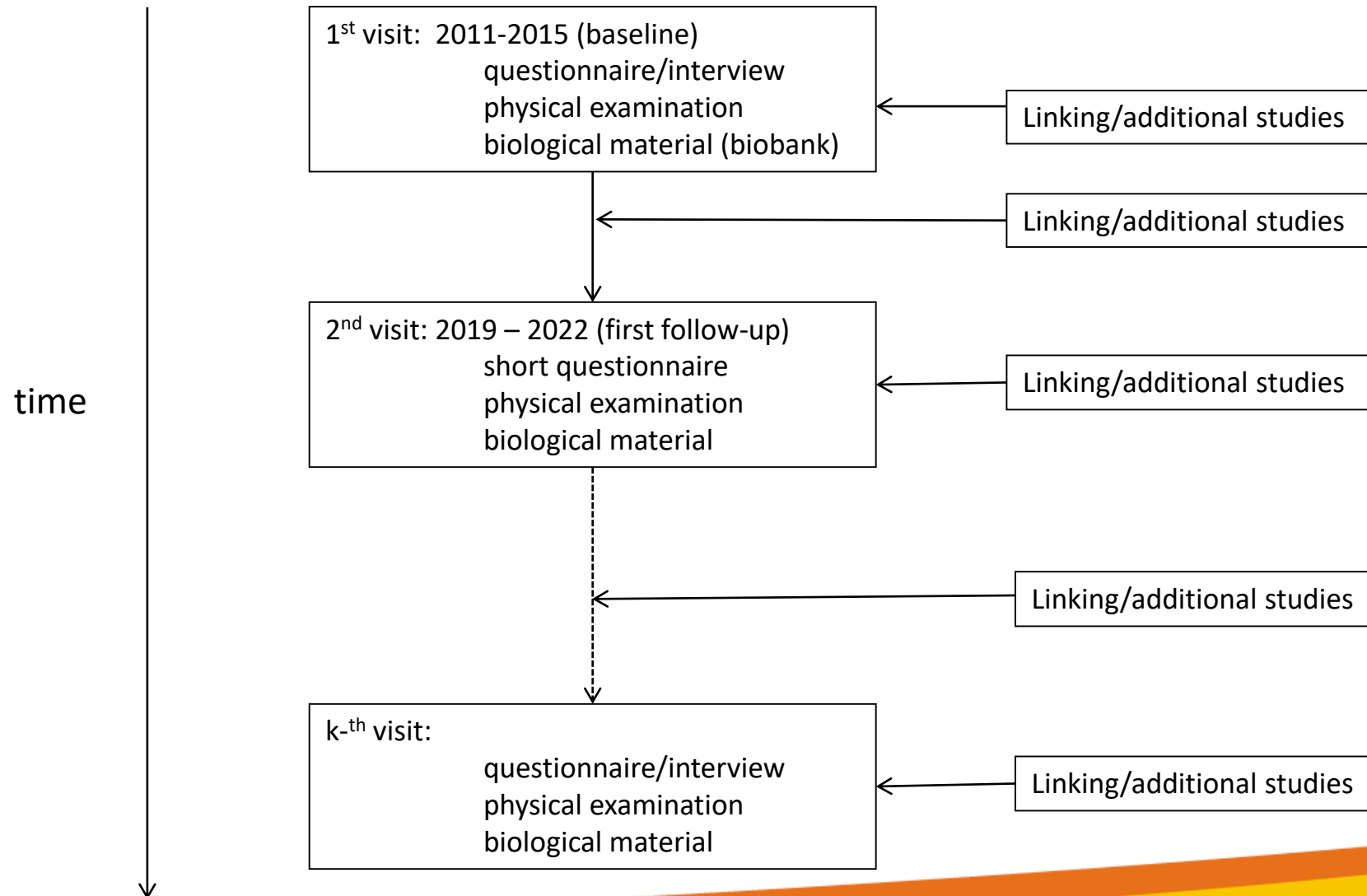
Ghanaian



Focus on three disease categories:



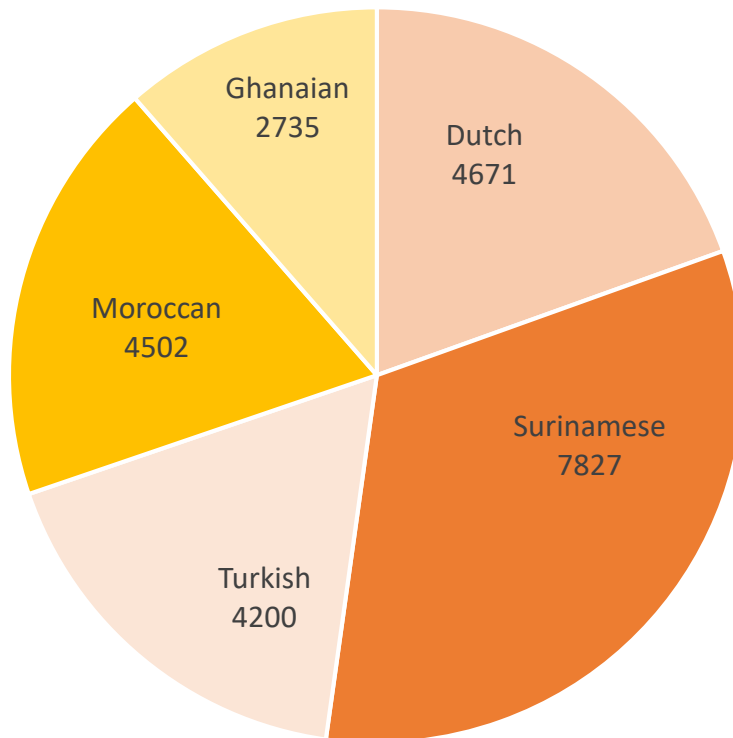
Longitudinal design





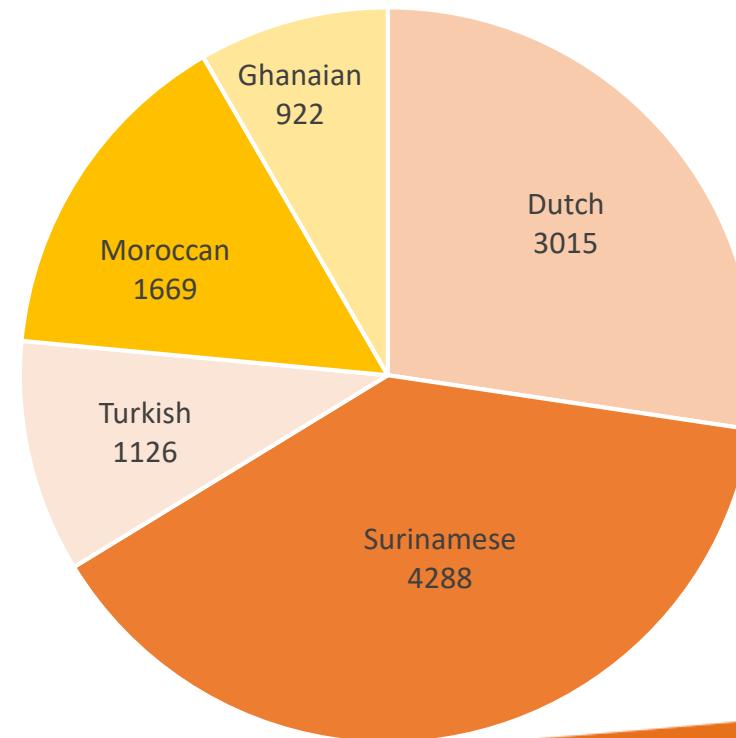
24,782

Participants finished the HELIUS-1 measurement



11,039

Participants finished the HELIUS-2 measurement





FAIR data in HELIUS

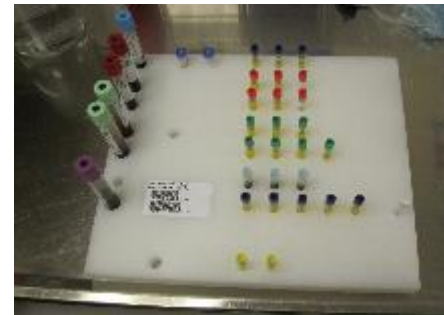
General: we strive for good quality data, that are meaningful and that can be used by researchers from different disciplines

1. Data collection
2. Data sharing



Measurements in HELIUS

1. Questionnaire/ interview
2. Physical examination
3. Biological material (lab and biobank storage)
4. Linking to registries
5. Additional studies (substudies)





Ensuring data quality



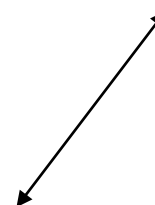
- Logistics database
- Research database
 - Applications: Castor, Limesurvey
- Quality control measures
 - Training
 - SOPs
 - Validation checks
- Use of validated measurements, existing vocabularies, scripts
- Data safety control



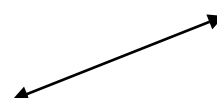
Data linkages



Statistics
Netherlands



× GGD
× Amsterdam



vektis



Zorg **TTP**
Privacy & Vertrouwen



FAIR

1. Data collection
2. Data sharing



Between 2010 and 2022: 180 scientific publications



Access to HELIUS data

We welcome researchers to use HELIUS data!

Procedures are described in HELIUS Collaboration Policy:

1. Submit a Publication proposal, including:
 - Statistical analysis plan, List of requested variables
2. Review by HELIUS board
 - Falls within scope of HELIUS, Scientific quality, Potential overlap
3. Sign Data Transfer Agreement
4. Free of charge for Amsterdam UMC researchers



EUROPEAN
GENOME-PHENOME
ARCHIVE



Archiving

- Tracking of datasets:
 - Each project has own identifier:



- Enabling:
 - Findability
 - Reproducibility
 - Transparency about results (e.g. for audits)



FAIR data – further steps

- Data sharing – Open science
 - For publishing data additional consent is needed
 - Synthetic data
- Data visibility
 - Secured environment for analysing data
 - Improve data showcase
 - Increased complexity (follow-up waves, sub studies)





Publishing metadata



- HELIUS-COVID study

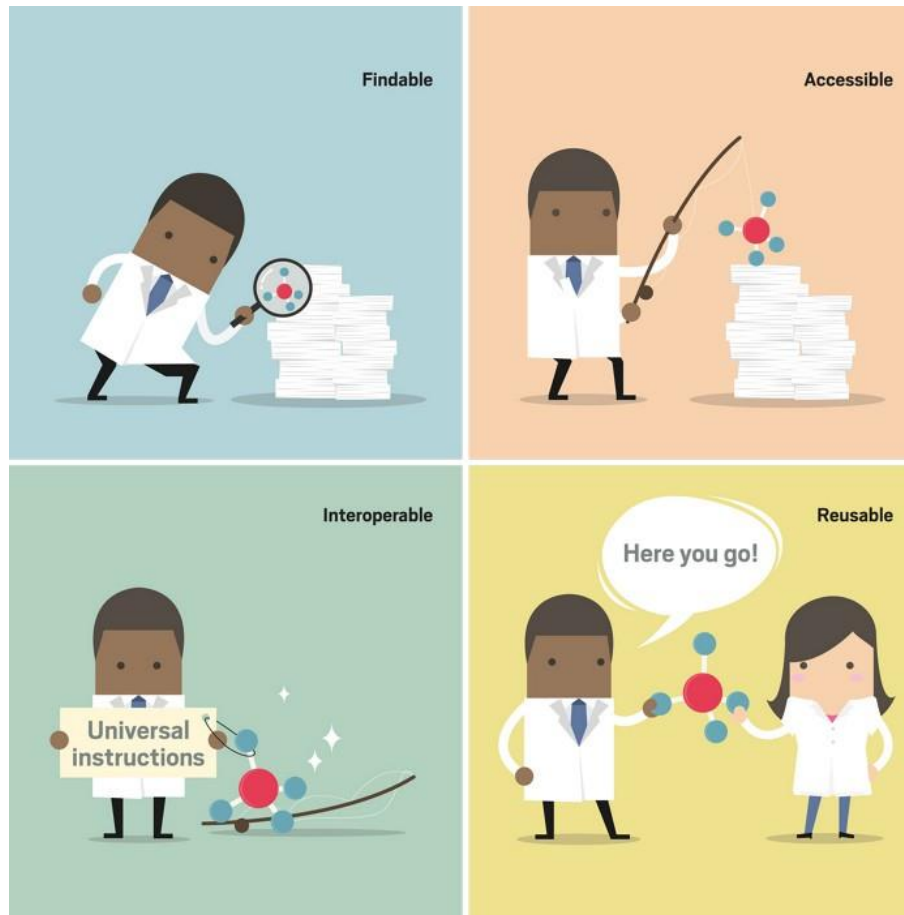


[View project](#)



General information

CEDAR admin project ID	https://repo.metadatacenter.org/template-instances/b7251b2e-730b-4cb5-80d4-2465f045fb48	📄 👁
Title	Ethnicity and COVID-19: epidemiology and control measures	
Acronym		
Websites	https://www.heliusstudy.nl/	
Principal investigator	https://orcid.org/0000-0002-0921-2232	
Data stewards	https://orcid.org/0000-0002-5904-3465 https://orcid.org/0000-0002-8963-4012	



heliuscoordinator@amsterdamumc.nl

www.Heliusstudy.nl

FAIR workshop

ASAP, 28 March 2023



How do I make my data FAIR?

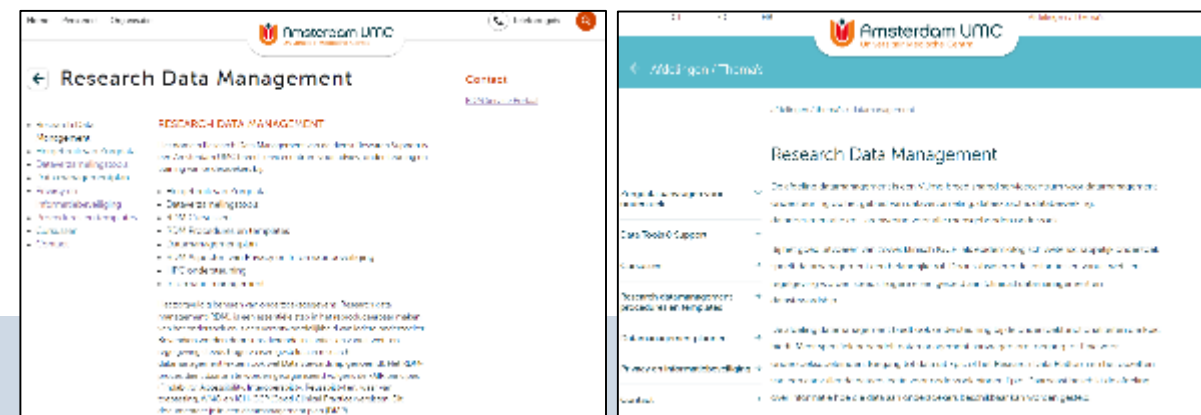
Rudy Scholte, domain head Research Data Management





RDM is here to help you!

- Research Data Management
- Compute, storage facilities
- Software management
- Consultation, information and education
- Applications for data collection, archiving, reuse
- Reuse of care data, text mining
- Intranet [AMC](#) / [VUmc](#) (search for 'RDM')
- [ServiceNow](#)





Study preparation - 1

- Follow [SOP Research Data Management](#) (K2)
Write your Data Management Plan
- Include data reuse in your informed consent
For nWMO: general consent procedure?





University Medical Centers

Research Data Management

Standard operating procedure

SOP-id	001
Version	2.0


University Medical Centers

Data management plan

F01 Data management plan v2.0 – RDM

For requirements, explanation of the topics, definitions and abbreviations, see [SOP 001 – RDM – Amsterdam UMC](#). The numbers in this DMP template refer to the numbers in the SOP 001 RDM.

The following study info provides relevant information regarding data management requirements:

Study number (ABR/METC)	
Study (acronym or short title)	
Sponsor (Verrichter/initiator/coördinerende hoofdonderzoeker)	



Study preparation - 2

- Choose your
 - application for data collection
consider reuse of existing (care) data
eCRF / questionnaire support and review
 - working / storage / compute environment



Epic + Research Data Platform

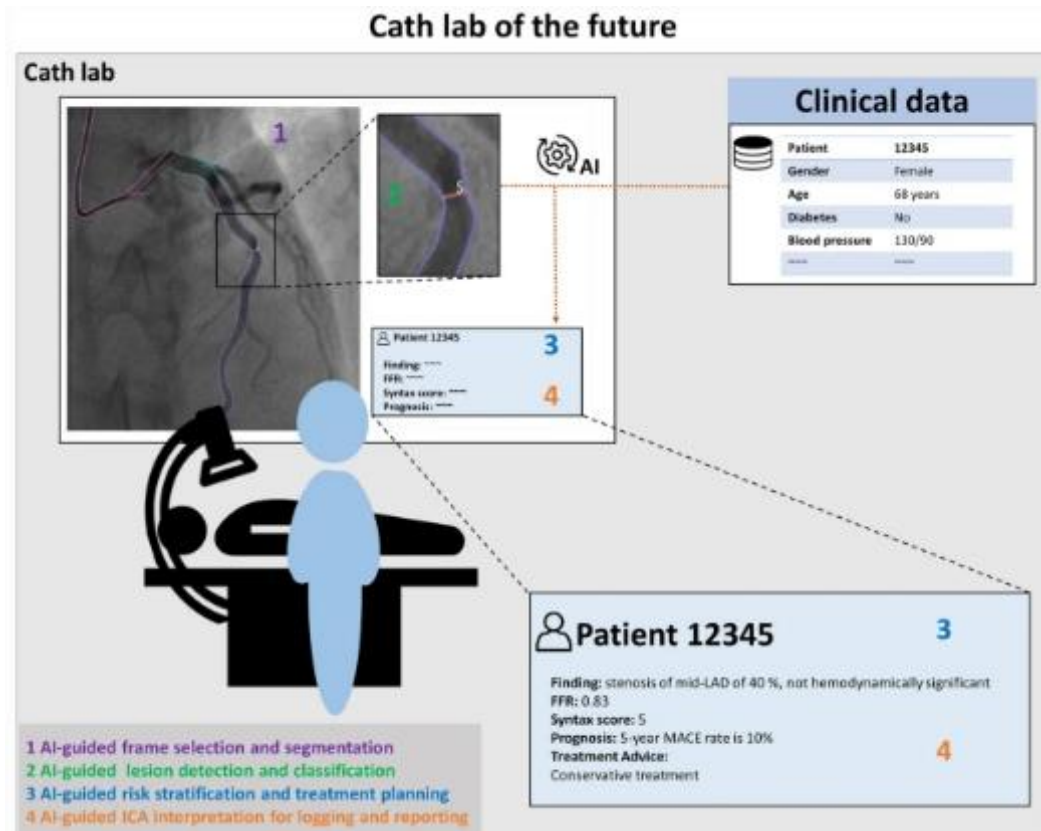


- Define your dataset according to common standards
 - SNOMED, LOINC
 - common definitions in your field
 - validated questionnaires, standard Castor pages





Reuse of existing data - RAISE (Cardiology)



- Discrete data and notes, images
- Anonymization by external partner
- Algorithm development in MyDRE





Data collection

- Quality control
 - Built-in validation checks on missings, outliers, logical combinations
- Change control
 - On data and application design
- Safe and restricted access to / transfer of data



Standard applications and tools offer these facilities



Data processing & analysis

- Have a Statistical Analysis Plan (SAP) in place
- Control your software / syntax / script

Software should be FAIR as well, especially when software is main research output

- Version control
- Descriptive comments





Writing / publishing your manuscript

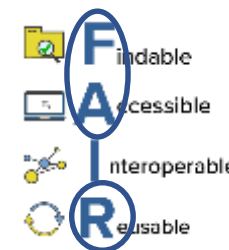
- Create a Digital Object Identifier (DOI) and add to manuscript
- Dataverse as publication platform





Data sharing and archiving

- Publish study documentation:
 - Protocol, DMP, SAP
 - Code book
 - Conditions & procedures for reuse



- Keep track of your data files
 - Metadata schema
 - File tracking, automated actions



- Archive your data in collaboration with RDM - *sustainable and for free*

- Sharing and archiving software



Data sharing - AmsterdamUMCdb (IC)

AmsterdamUMCdb

The first freely accessible intensive care database from within the European Union containing de-identified health data related to tens of thousands of European intensive care unit admissions, including demographics, vital signs, laboratory tests and medications.

AmsterdamUMCdb Leadership



Paul Elbers,
MD, PhD, EDIC
Intensivist, Amsterdam UMC



Patrick Thorat,
MD
Intensivist, Amsterdam UMC



Database content and technical aspects

AmsterdamUMCdb is a relational database consisting of 7 tables. The current version is v1.0. It is distributed as tab separated files and will be made available at the [EASY repository](#) from the Dutch Archiving and Networked Services (DANS). A detailed description of AmsterdamUMCdb and its data is available at [our GitHub repository](#).

Requesting Access

The database, although de-identified, still contains detailed information regarding the clinical care of patients, so must be treated with appropriate care and respect and cannot be shared without permission. This is outlined in the End User Licence. Those seeking to use AmsterdamUMCdb must complete the following formal steps:



More information and contact

ICT IT IR
Amsterdam UMC
Universitair Medisch Centrum

Afdelingen / Thema's

Afdelingen / Thema's > datamanagement

Research Data Management

De afdeling datamanagement, bij het VUmc, treedt als breed servicecentrum voor datamanagement, ondersteuning op het gebied van dataverzameling, dataextractie, databewaking, databestuurbaarheid en -overname van data management en procesen.

Op het gebied van zowel klinisch (GPI) als epidemiologisch wetenschappelijk onderzoek wordt datamanagement een belangrijk onderdeel. Data uit verschillende instanties en vanuit verschillende disciplines worden steeds frequenter samen gebruikt aan klinisch datamanagement en databewaking.

De afdeling datamanagement biedt ook ondersteuning bij de onderzoeksfuncties die de Edele biedt. Meer specifiek behandelt datamanagement aanvragen om toegang tot - en voor onderzoekers, toegang tot data uit het UMC, het Research Data Platform en het openbaar van de verzamelde data voor onderzoek binnen Epic. Daarnaast beschikt de afdeling over informatie hoe de data van onderzoekers beschikbaar kan worden gesteld.

Home Personeel Organisatie

Amsterdam UMC
Universitair Medisch Centrum

Research Data Management

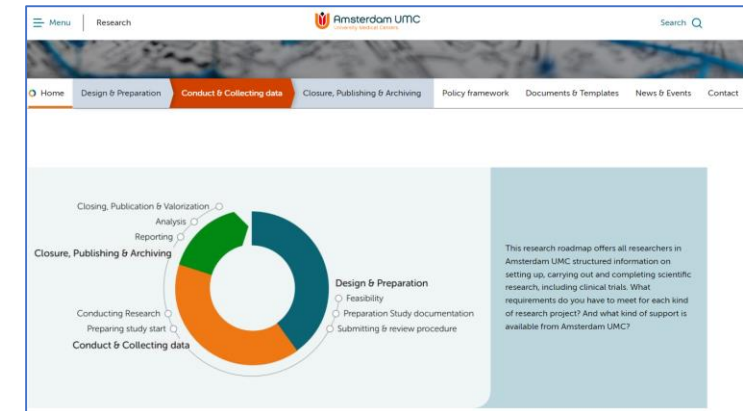
- Research Data Management
- Hergebruik van Zorgdata
- Dataverzamelings tools
- Data managementplan
- Privacy en informatiebeveiliging
- Procedures en templates
- Cursussen
- Contact

RESEARCH DATA MANAGEMENT

Het domein Research Data Management van de dienst Research Support is een Amsterdam UMC breed servicecentrum voor advies, ondersteuning en training van onderzoekers bij:

- Hergebruik van Zorgdata
- Dataverzamelings tools
- RDM Cursussen
- RDM Procedures en templates
- Datamanagementplan
- RDM Aspecten van Privacy en Informatiebeveiliging
- HPC ondersteuning
- Informatie management

Het zorgvuldig behouden van onderzoeksgegevens, Research data management (RDM), is een essentiële stap in het reproduceerbaar maken van het onderzoek en is een verantwoordelijkheid van iedere onderzoeker.



- Intranet: [VUmc](#) / [AMC](#) (search for 'RDM')
- [ServiceNow](#)
- Research website: [Research roadmap Amsterdam UMC](#) / [Research Support Amsterdam UMC](#)