

Genom- und andere Gesundheitsdaten in Kohortenstudien

Tobias Pischon

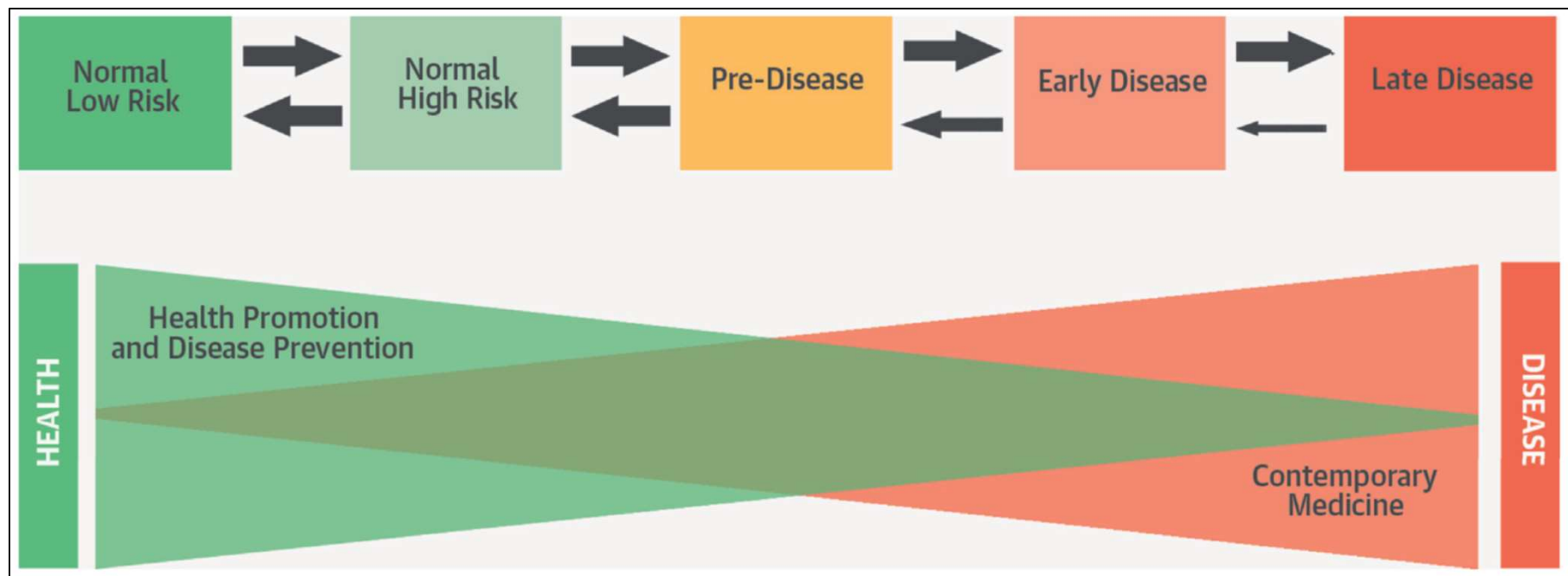
Max Delbrück Center (MDC)
NAKO e.V., Heidelberg



DAS KONTINUUM VON GESUNDHEIT ZU KRANKHEIT

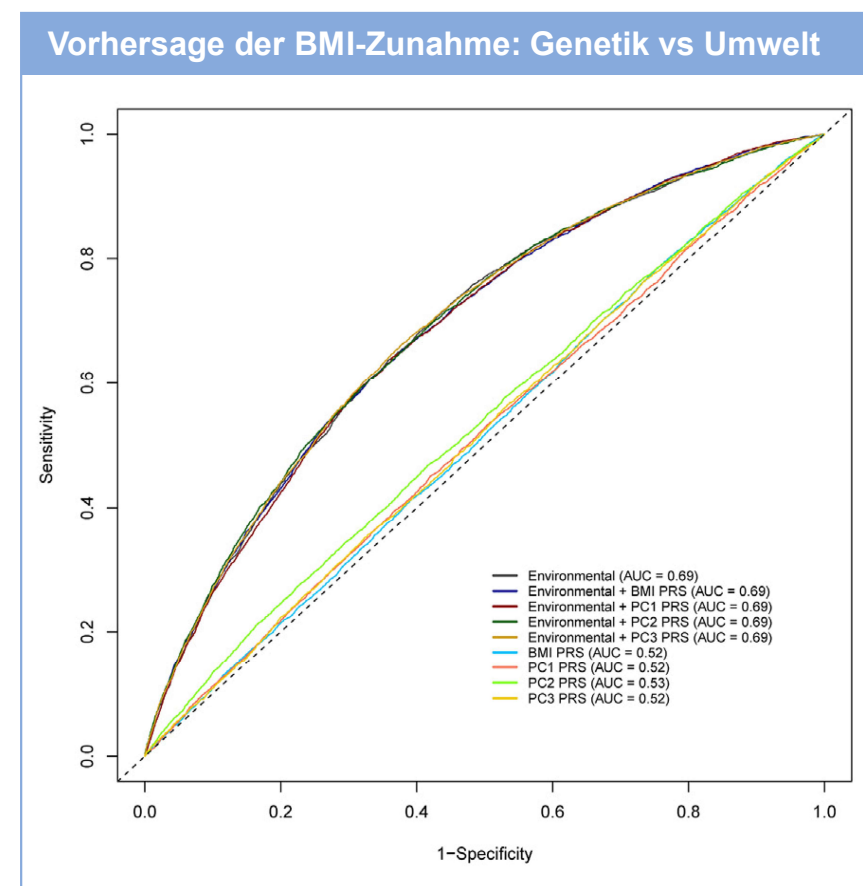
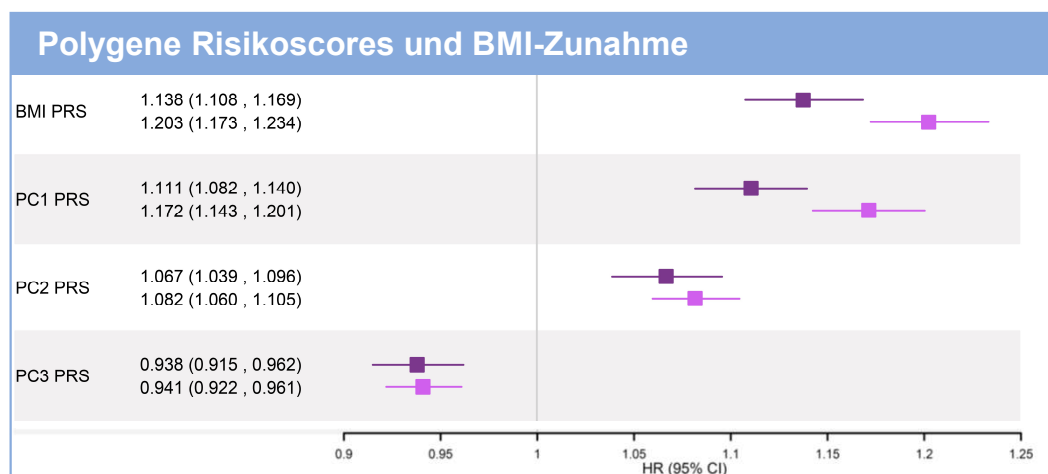
„Prognosen sind schwierig, besonders wenn sie die Zukunft betreffen.“

(Herkunft des Zitats wird verschiedenen Personen zugesprochen)



Knapper, et al. J Am Coll Cardiol. 2015

GENETIK VERSUS UMWELT: PRÄDIKTOREN FÜR ZUNAHME DES BMI UM $\geq 5\%$



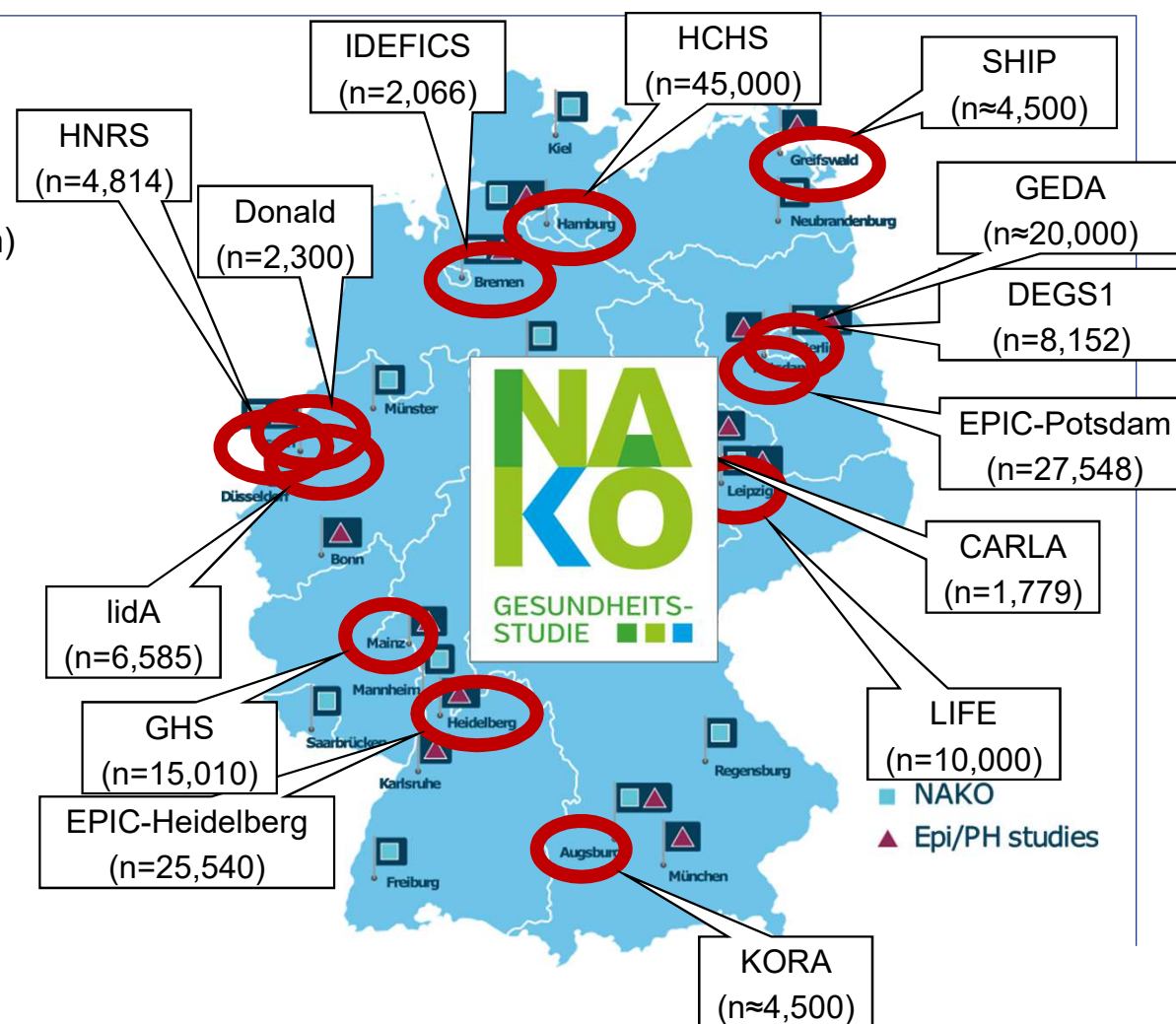
Peruchet-Noray et al., EBioMedicine 2025

EPIDEMIOLOGISCHE STUDIEN IN DEUTSCHLAND

In der Vergangenheit viele kleine bis mittelgroße Kohortenstudien in Deutschland:

- Unterschiedliche thematische Schwerpunkte (heterogen)
- Teilweise auch genetische Daten verfügbar
- Daten oft nur lokal verfügbar; lokale Antragstellung
- Unterschiedliche Daten- und Metadatenformate
- Verschiedene Datenhaltende Organisationen (DHO)

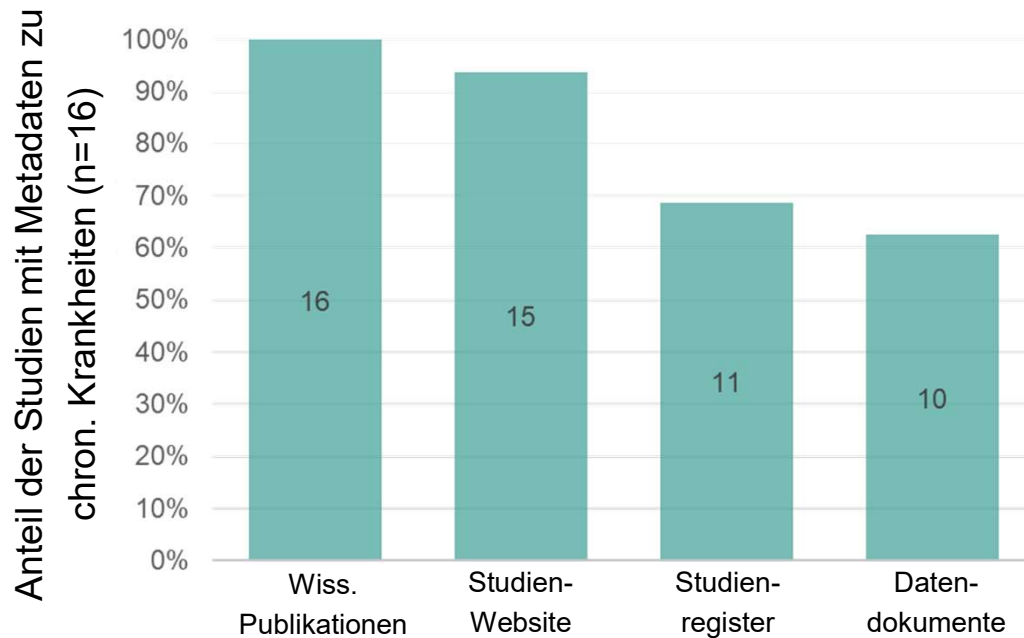
NEU: NAKO Gesundheitsstudie



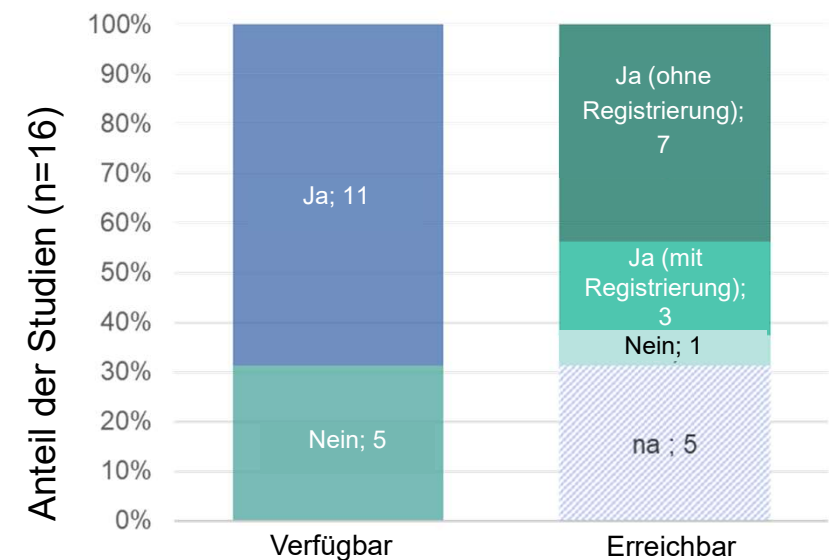
QUELLEN FÜR ZUGRIFF AUF METADATEN UND DATEN

BEOBACHTUNGSSTUDIEN ZU CHRON. KRANKHEITEN IN DEUTSCHLAND

Quellen für Metadaten



Verfügbarkeit und Erreichbarkeit von Infrastrukturen für Zugriff auf (Meta-)Daten



Übergeordnete Ziele

- Unterstützung der klinischen und epidemiologischen Forschungsgemeinschaft, ihre Daten in Übereinstimmung mit Datenschutzbestimmungen und ethischen Grundsätzen mit der Nutzergemeinschaft zu teilen
- Schaffung neuer Möglichkeiten der Datenanalyse innerhalb der NFDI im Interesse der Verbesserung der Gesundheit der Bevölkerung
- Basierend auf den FAIR-Prinzipien (Findable, Accesible, Interoperable, Reusable)



NATIONALE FORSCHUNGSDATENINFRASTRUKTUR FÜR PERSONENBEZOGENE GESUNDHEITSDATEN (NFDI4HEALTH)

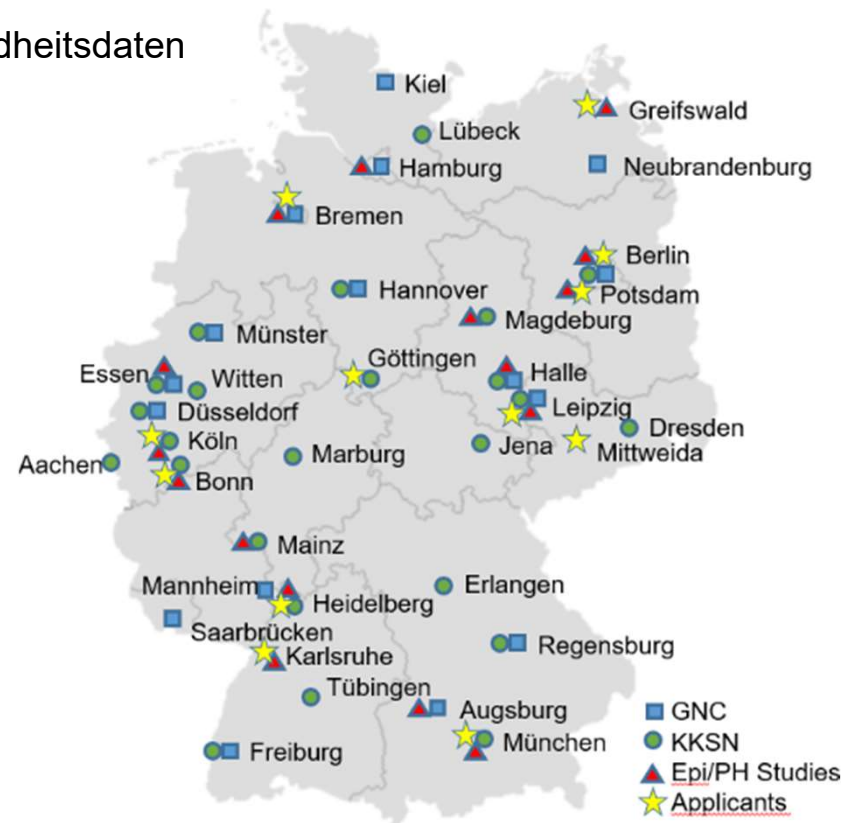
Forschungseinrichtungen mit Expertise in:

- Erhebung, Analyse und Erforschung personenbezogener Gesundheitsdaten
- Entwicklung von Standards und Methoden im Bereich Daten
- Gemeinschaftliches Engagement (Community Involvement)

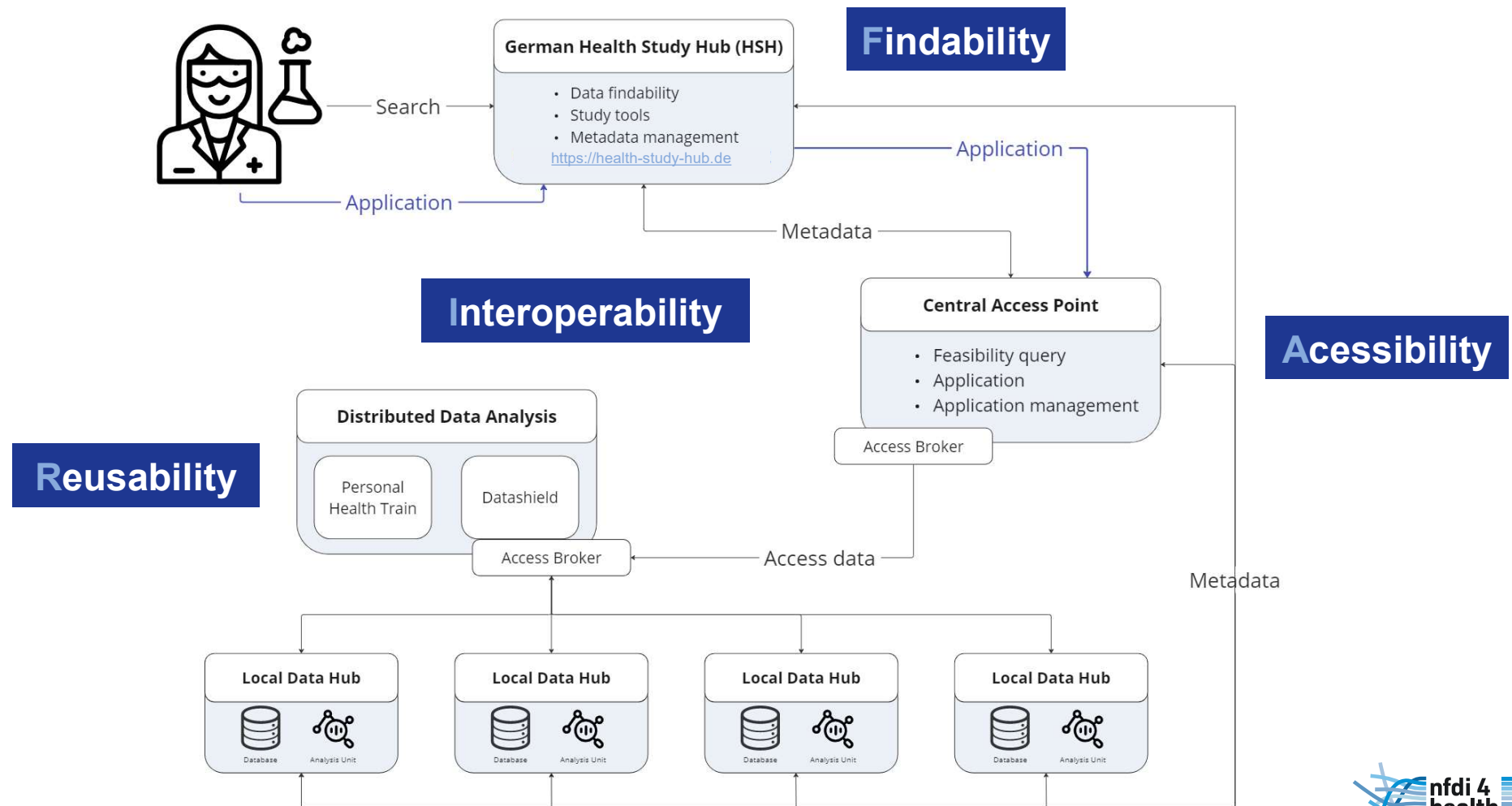


Datenhaltende Organisationen (DHOs):

- Epidemiologische und Public Health-Studien
 - 26 lokale Studien mit > 400.000 Teilnehmern
 - NAKO Gesundheitsstudie
- Klinische Studien
 - 27 universitäre Studienzentren
- Register und administrative Gesundheitsdatenbanken



NFDI4HEALTH – KONZEPT DER FAIR SERVICES



NFDI4HEALTH – HEALTH STUDY HUB FINDABILITY

- Eine zentrale Plattform für alle Studien
- Verbunden mit lokalen und übergreifenden Study hubs
- Metadaten über Studien, Datensammlungen, Fragebögen und Daten (data dictionary)
- Ggw. ca. 44.000 Einträge

The screenshot displays the NFDI4HEALTH Health Study Hub interface. The top navigation bar includes a search bar, filters for 'Explore by type of the resource' and 'Explore by collection(s)', and links for 'Help/documentation', 'Register your Data', and 'Sign in or Sign up'. The main content area shows a search results table for COVID-19 related studies. The table has columns for Name, Label, Value type, Annotations, Study, and Instrument. The search criteria on the left include Variables, Areas of Information, Source & target, Properties, Datasets, and Studies. The search results table lists various COVID-19 related studies, including those from the German Central Health Study and Mica.

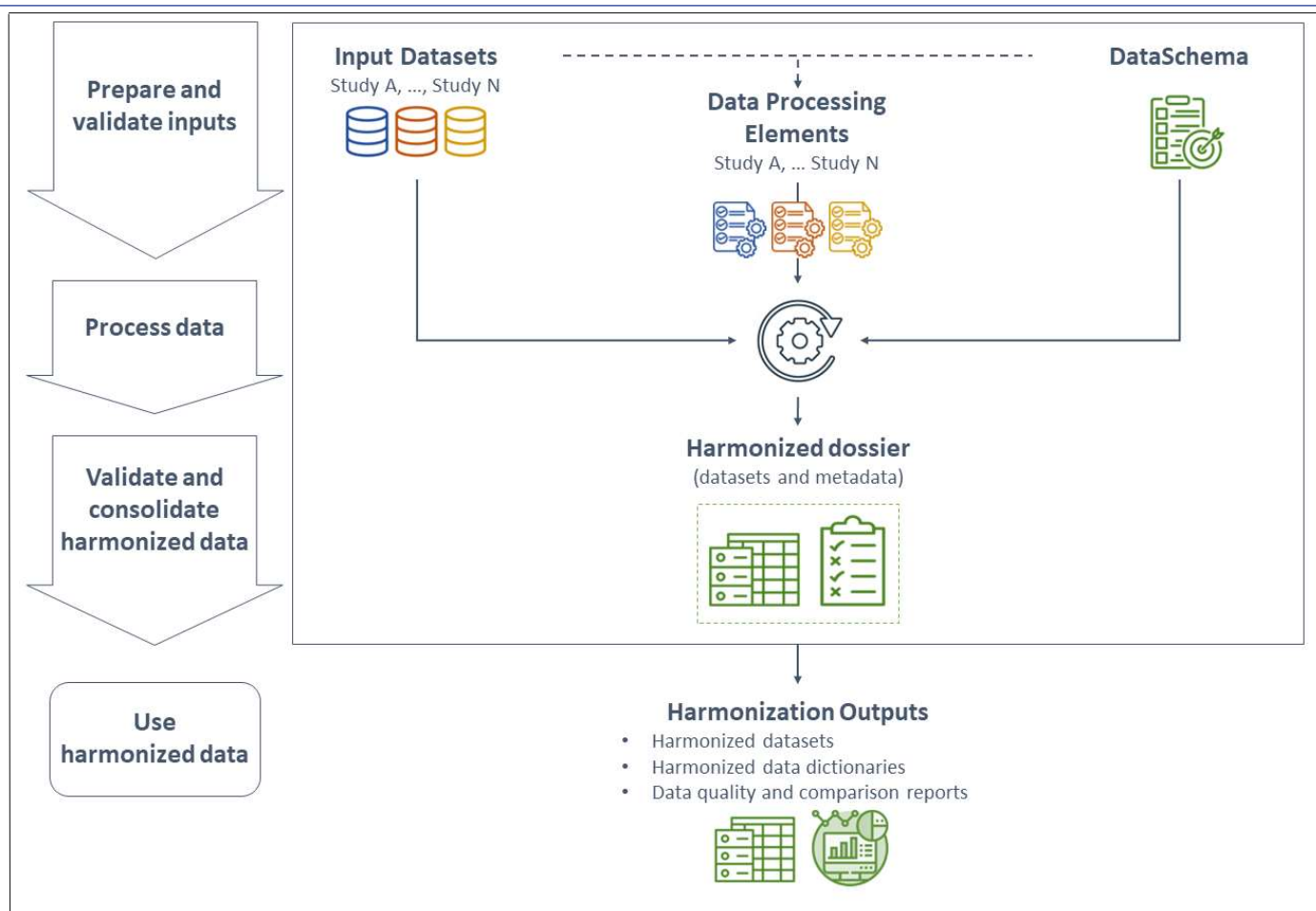
Name	Label	Value type	Annotations	Study	Instrument
q0	Datum Heute	Date	Date and time-related information	NAKO	NAKO
q1	Wie würden Sie Ihren Gesundheitszustand im Allgemeinen beschreiben?	Integer	Perception of health	NAKO	NAKO
q2	Verglichen mit der Zeit vor der Corona-Pandemie, wie würden Sie Ihren derzeitigen Gesundheitszustand beschreiben?	Integer	Perception of health	NAKO	NAKO
q3	Hatte seit dem 1. Februar 2020 jemand in Ihrem Haushalt eine Infektion der oberen Atemwege (z. B. Schnupfen oder Infektionen von Nasennebenhöhlen, Mandeln, Mittelohr, Rachen, Kehlkopf)?	Integer	Diseases of the respiratory system (J00-J99)	NAKO	NAKO
q4	Hatte seit dem 1. Februar 2020 jemand in Ihrem Haushalt eine Infektion der unteren Atemwege (z. B. Bronchitis oder Lungenentzündung)?	Integer	Diseases of the respiratory system (J00-J99)	NAKO	NAKO
q5	Waren seit dem 1. Februar 2020 Personen in Ihrem Haushalt wegen der in Frage 3 oder 4 erfassten Erkrankung(en) so krank, dass sie im Bett bleiben mussten oder ihrem normalen Alltag nicht nachgehen konnten (Hausarbeit, Hobbies etc.)?	Integer	Diseases of the respiratory system (J00-J99)	NAKO	NAKO
q6	Waren seit dem 1. Februar 2020 Personen in Ihrem Haushalt wegen der in Frage 3 oder 4 erfassten Erkrankung(en) krankgemeldet und konnten ihrem Beruf nicht nachgehen?	Integer	Diseases of the respiratory system (J00-J99)	NAKO	NAKO
q7	Waren seit dem 1. Februar 2020 Personen in Ihrem Haushalt wegen der in Frage 3 oder 4 erfassten Erkrankung(en) in stationärer Behandlung (d. h. mindestens eine Nacht im Krankenhaus)?	Integer	Hospitalizations	NAKO	NAKO
q8	Wurden Sie selbst seit dem 1. Februar 2020 in einer Arztpraxis, in einem Testzentrum oder in einem Krankenhaus einmalig oder mehrfach auf das Corona-Virus getestet?	Integer	Laboratory diagnosis interventions	NAKO	NAKO

DATENHARMONISIERUNG IN NFDI4HEALTH

INTEROPERABILITY – SEMIAUTOMATISCHE DATENHARMONISIERUNG

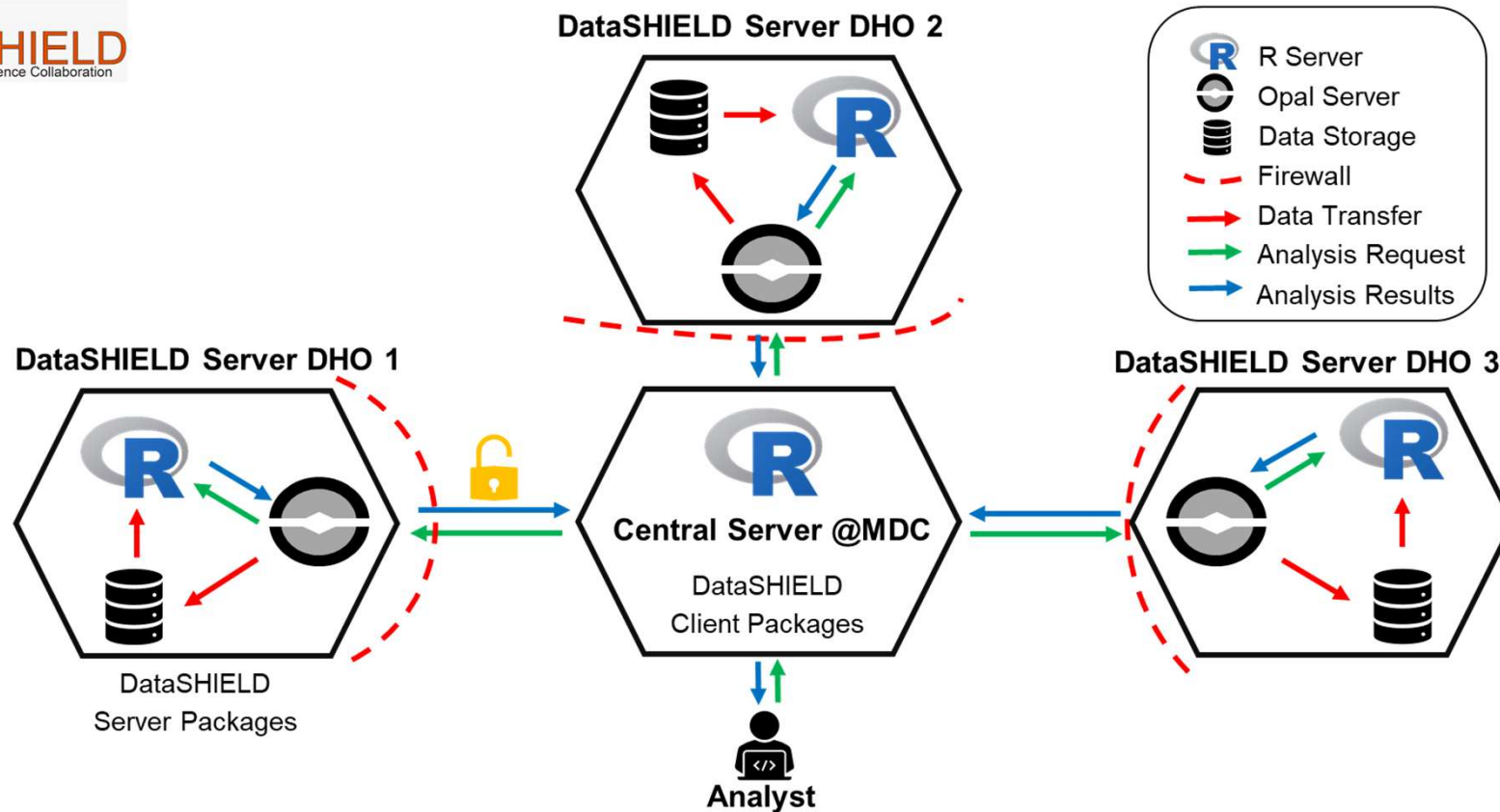


- R package Rmonize
- Developed and published by Maelstrom Research (Guillaume Fabre)



NFDI4HEALTH – FÖDERIERTE DATENANALYSE

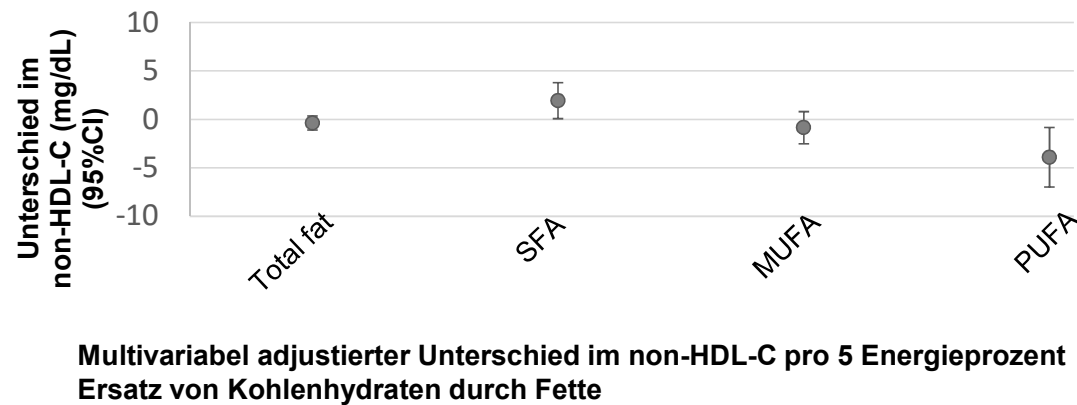
REUSABILITY



SOP for Installation and Configuration of Opal DataSHIELD in NFDI4Health:
<https://github.com/nfdi4health/opal-datashield-sop/wiki>

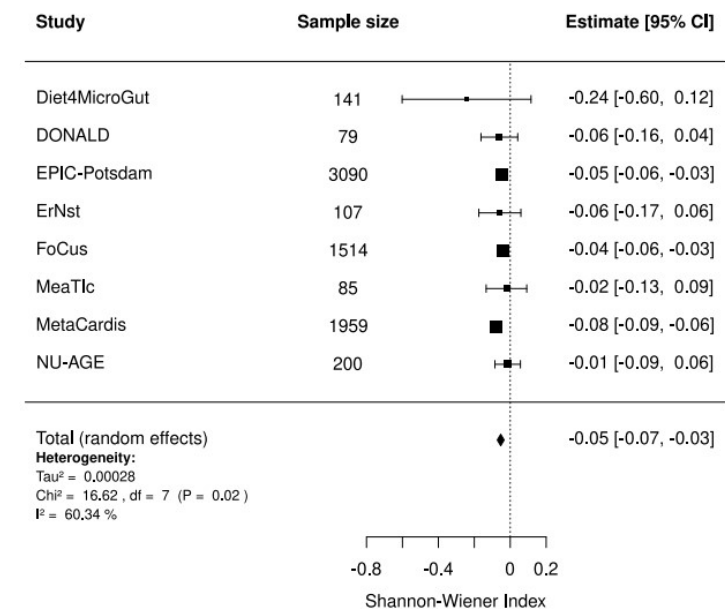
NFDI4HEALTH – BEISPIELE FÖDERIERTER DATENANALYSEN

Zusammenhang zwischen Fettzufuhr und non-HDL Cholesterin: föderierte Datenanalyse in 8 Europäischen Studien (n=5,919)



Pinart et al. Eur J Nutr 2021

Darmmikrobiom und Adipositas in 8 Europäischen Studien (n=7,656)



Unterschied in der α -diversity pro 5 kg/m² höherem BMI
(adjustiert für Alter und Geschlecht)

Schwedhelm et al. under review

NFDI4HEALTH – ANWENDERORIENTIERTE ENTWICKLUNG UND NACHHALTIGKEIT

Use Case:

Ernährungsepidemiologie

methodische Limitationen bei der Ableitung von Ernährungsmustern

Use Case:

Epidemiologie chronischer Krankheiten

Zusammenhang zwischen dem Zuckerkonsum, der Entwicklung von Adipositas und dem Auftreten chronischer Krankheiten

Harmonisierung

Ernährungsdaten
Chronische Krankheiten
Covariablen (z.B. Sozialstatus, Rauchen, etc.)

	Datenhaltende Organisation	Fortschritt		
		Metadaten-Sammlung	Datashield-installation	Daten-harmonisierung
GINIplus	HMGU			
LISA	HMGU			
KORA	HMGU			
KarMeN	MRI			
ActivE-Study	MDC			
DONALD	U Bonn			
CARLA	UM/UK Halle			
EPIC-Potsdam	DIfE			
DEGS1	RKI			
IDEFICS/I.Family	BIPS			
SHIP/SHIP Trend	UM Greifswald			
EPIC-Heidelberg	DKFZ			
NAKO Health Study	NAKO			
LIFE-Adult	U Leipzig			
Gutenberg Health Study	UM Mainz			
Hamburg City Health Study	UKE			
Heinz-Nixdorf Recall Study	IMIBE			

abgeschlossen

andauernd

in Vorbereitung



GENOM- UND ANDERE GESUNDHEITSDATEN IN KOHORTENSTUDIEN



NAKO GESUNDHEITSSTUDIE

Die größte Gesundheitsstudie in Deutschland



- Populationsbasierte, prospektive Kohortenstudie
- 205,000 Männer und Frauen im Alter von 20-69 Jahren
- 30,000 MRT Untersuchungen
- 18 Studienzentren, 5 MRT Zentren
- NAKO e.V., 26 Mitgliedseinrichtungen
- Nachbeobachtung geplant für >30 yrs
- Wiederholungsuntersuchungen nach 5 und 10 Jahren
- Gefördert durch das Bundesministerium für Forschung, Technologie und Raumfahrt, die beteiligten Bundesländer, die Helmholtz-Gemeinschaft und die Mitgliedseinrichtungen



DEUTSCHLANDS DETAILREICHSTE PLATTFORM FÜR PRÄVENTIONSFORSCHUNG



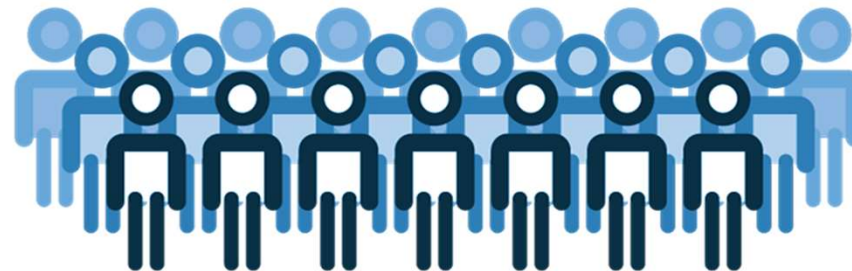
Molekulare Faktoren



Umwelt



Lebensstil



Herz-Kreislauf



Diabetes



Krebs



Lunge



Allergien



Infektionen



Mentale
Gesundheit

EINZIGARTIGE DATENSAMMLUNG

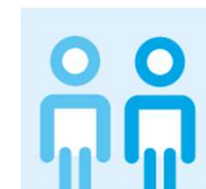
(3- bis 5-stündiges Untersuchungsprogramm pro Teilnehmerin und Teilnehmer)



STATUS JULI 2025

Untersuchung

Untersuchung	Zeitperiode	Ziel	Erreicht, N	Erreicht, %
Basisuntersuchung	2014-2019	200.000	205.415	103% ✓
Zweituntersuchung (5 J.)	2018-2024	135.000	137.922	102% ✓
Drittuntersuchung (10 J.)	2024-2028	85.000	23.716	andauernd



MRT Untersuchung

Basisuntersuchung		30.000	30.861	103% ✓
Zweituntersuchung (5 J.)		18.000	19.314	107% ✓
Drittuntersuchung (10 J.)		11.150	3.421	andauernd



Bioprobensammlung

Serum, Plasma, Buffy Coat, RNA, Urin	@ Basis-, Zweit- und Drittunters.	} 27.3 Millionen Proben ✓
Stool	@ Basis- und Drittunters.	
Saliva, nasal swaps	@ Basisuntersuchung	



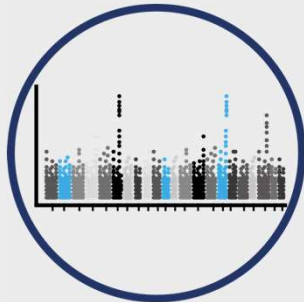
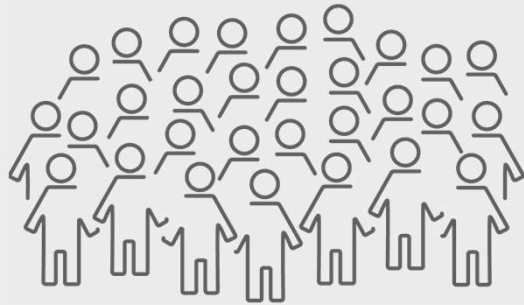
Aktive und passive Nachbeobachtung zu inzidenten Krankheiten und Mortalität

- Erster Fragebogenzyklus abgeschlossen (N=170,226) ✓
- Zweiter Fragebogenzyklus andauernd ✓
- Mortalitätsnachbeobachtung und Linkage zu Sekundärdaten andauernd ✓



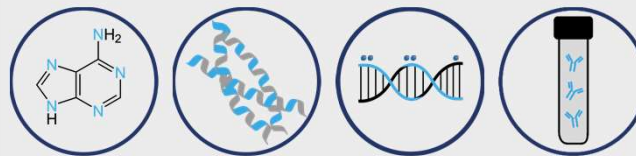
MOLEKULARE PHÄNOTYPISIERUNG IN NAKO

Alle Teilnehmer*innen



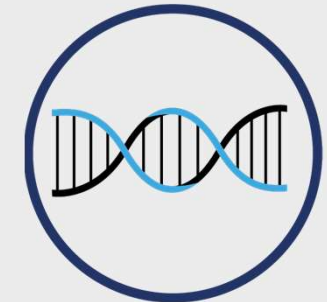
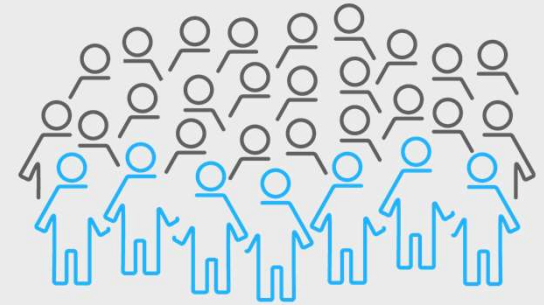
Genotypisierung

5,000 Teilnehmer*innen



Molekulare Charakterisierung

25,000 Teilnehmer*innen



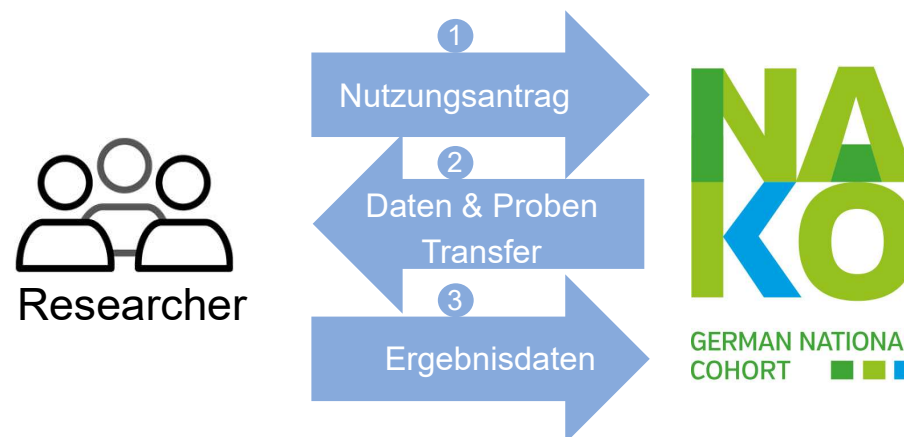
Sequenzierung

NUTZUNG VON DATEN UND PROBEN DER NAKO GESUNDHEITSSTUDIE

NAKO-Daten entdecken: transfer.nako.de

The screenshot shows the NAKO TransferHub website. At the top, there's a navigation bar with 'NAKO TransferHub > Home', 'EN', 'FAQ', and 'Logout'. The main heading is 'German National Cohort (NAKO)'. Below this, there's a grid of participant photos on the left with the NAKO logo and 'nako.de'. To the right, there's a text block explaining that the NAKO e.V.'s Transfer Unit offers scientists from research institutions in the EU and the EEA the data and biosample material from the German National Cohort (NAKO) for scientific purposes. It mentions that Use and Access Applications can be submitted via this web portal in accordance with the provisions of the 'Use and Access Policy'. Before submitting your application, please read the 'Information on use and access procedure'. Below this, there are buttons for 'Data dictionary' and 'Approved research'. To the right of the text, there's a section titled 'Important information' with links to 'NAKO study description 2022', '2014 Use and Access Policy', 'Publication Policy', 'data dictionary preview for Excel', 'Beantragung von Bioproben', 'Use and access procedure', 'Verwendung von aggregierten Daten', 'Beantragung einer Zwischenübergabe', 'Metadaten und Ergebnisdatenübermittlung', 'Korrekturgewichte', and 'Beantragung von Umweltdaten'. At the bottom, there's a section titled 'Create a new Use and Access Application for scientific evaluation' with a brief description of the application process. Below this, there are links for 'My Applications', 'My Variable Sets', 'My Subpopulations', 'User Profile', 'Contractual Partner', and 'Support Request'.

Nutzung von Daten und Proben

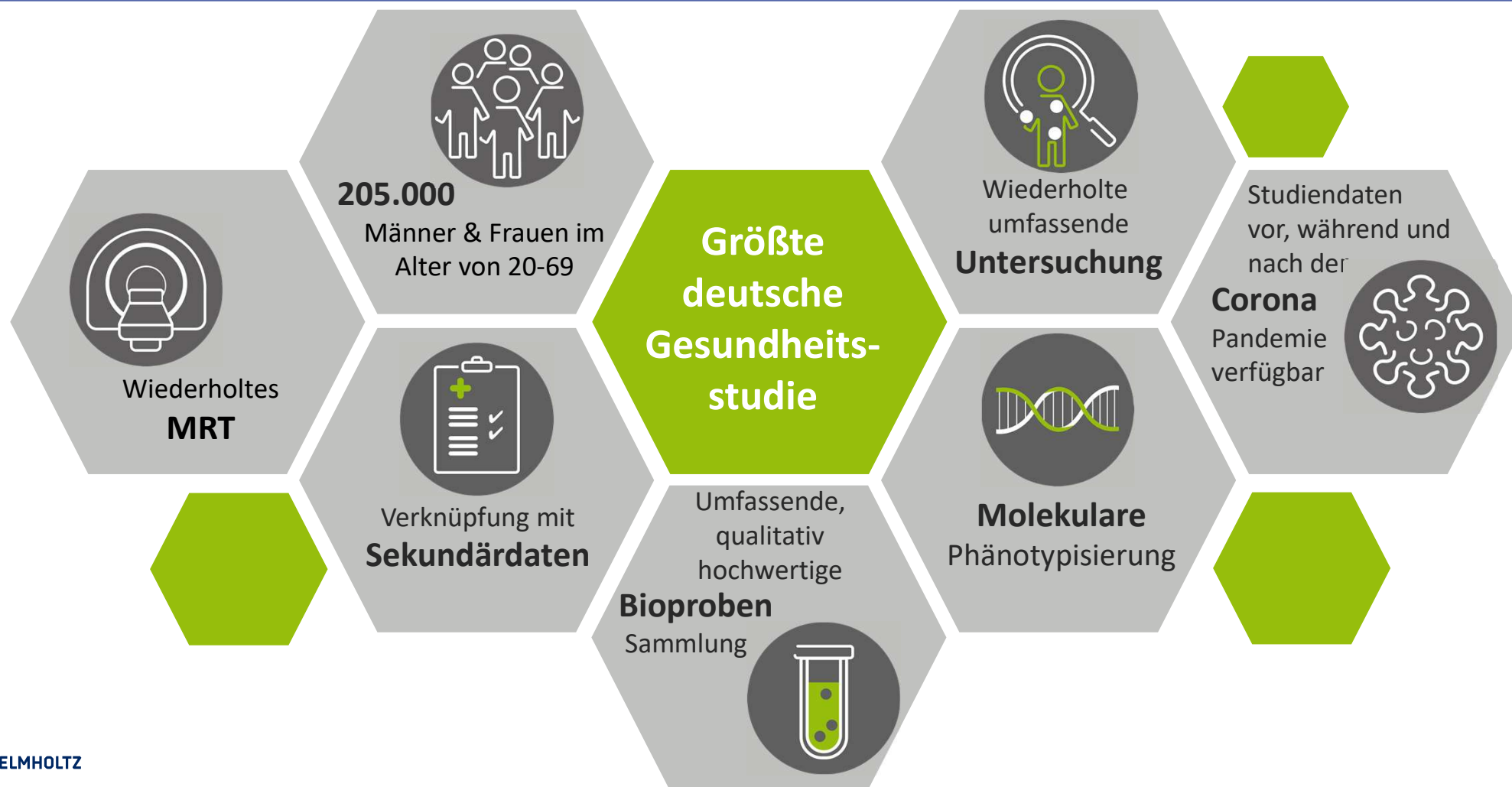


Hintergrundinformationen:

<https://nako.de/en/>

Peters, et al. Eur J Epidemiol 37, 1107–1124 (2022)

NAKO AUF EINEN BLICK



- Kohortenstudien leisten einen essentiellen Beitrag zur Präventionsforschung
- NFDI4Health unterstützt die Nutzung personenbezogener Gesundheitsdaten für wissenschaftliche Forschung nach den FAIR-Prinzipien
- Die NAKO Gesundheitsstudie ist die größte Ressource für den Bereich die ursachenbezogenen populationsbasierten Präventionsforschung in Deutschland
- Die Verfügbarkeit von Genom und anderen Gesundheitsdaten in Kohortenstudien ermöglicht Forschung im Bereich der Präzisionsprävention

Vielen Dank für Ihre Aufmerksamkeit



Gefördert durch
DFG Deutsche
Forschungsgemeinschaft



Gefördert vom Bund,
den Ländern und der
Helmholtz-Gemeinschaft

