SHARING OF SENSITIVE HEALTH DATA

— TWO PRACTICAL EXAMPLES —

Iris Pigeot

Leibniz Institute for Prevention Research and Epidemiology – BIPS

G20 DGI-2 Workshop on Recommendation II.20 “Promotion of Data Sharing”
25 March 2021
Personal Health Data

- Highly **sensitive** and specially **protected**
- Creation of **absolute anonymity** not possible
  - Deeply phenotyped
- **Living data sets** with need for continuous maintenance and update
- Data usage based on **informed consent** of participants

➤ **Data protection and ethical issues essential**
Motivation for Data Sharing

- Not possible to **fully exploit potential** of research data obtained for projects with narrow focus and limited duration
  - Opportunity to investigate research questions that may not have been foreseen at time of project inception

- Investigation of health risks and protective factors highly benefit from **large sample sizes**
  - Rare diseases
  - Small effect sizes (e.g., genetic risks)
  - Heterogeneous populations
  - Season effects (e.g., dietary and physical activity behavior)

- **Record linkage** of special interest and with particular challenges
  - Use of secondary data to reduce burden of participants and recall bias

➢ **A matter of efficient use of resources**
Overview

➢ Multi-centre cohort study

➢ 8 European countries

➢ Baseline 2007/2008:
  • 16,228 children, 2-9.9 years

➢ 1st follow-up 2009/2010:
  • 13,596 children, 2-12 years

➢ 2nd follow-up 2013/2014:
  • 9,617 children, 5-17 years

IDEFICS study

- Enhance knowledge of health effects of changing diet & altered social environment & lifestyle of children, 2-9 years, in Europe
- Develop, implement & evaluate specific intervention approaches to reduce prevalence of diet- & lifestyle-related diseases & disorders

I.Family study

- Make significant contribution to reduce burden of nutrition-related diseases
- Focus on individual and his/ her family
- Assess dynamic nature of causal factors over time and during transition into adolescence
Based on individual **cooperation agreements**

- Check whether intended usage in line with informed consent
- Final decision by steering committee

**Tailored analysis data set**

- Constrained to variables needed for specific analysis (principle of data minimization)
- Access via **Central Data Server** hosted by BIPS
  - Remote access (VPN, firewall-protected)
  - Individual password-protected user accounts
  - Access restricted to analysis data set
  - Download restricted to analysis results (summary statistics)
Huge international pooling studies
- Example: NCD Risk Factor Collaboration (NCD-RisC)
- Usually based on summary statistics
  - Unified inclusion and exclusion criteria
  - Unified calculations
German Pharmacoepidemiological Research Database


General Information

- Claims data of ~25 million insurants
  - 4 statutory health insurances: DAK, hkk, TK, AOK Bremen/Bremerhaven,

- Information on persons who have been insured with one of the providers since 2004 or later

- About 20% of German population

- As of today: yearly data delivery (most recent 2018)

- Coverage: all geographical regions of Germany

- Contains demographic information, information on hospital stays, outpatient visits and drug prescriptions

- National reference for Germany

https://www.bips-institut.de/en/research/research-infrastructures/gepard.html
Research Aims

- Monitoring of drugs and vaccines after approval
- Analysis of
  - Drug and vaccine utilization, misuse in pharmaceutical therapy of certain indications
  - (Rare) drug/vaccine risks
  - Drug/vaccine risks with long latency periods
  - Drug-drug interactions

- Often only be possible in cooperation with international consortia

- Specific challenges in international projects
  - Social data particularly protected in Germany (Social Code Book X)
  - Data may not at all be analyzed outside BIPS (no remote access)
OMOP – Common Data Model

- **Concept:**
  - Transform data from different databases into common format and representation (terminologies, vocabularies, coding schemes)
  - Perform systematic analyses using library of standard analytic routines based on the common format

[Diagram showing data transformation and analysis processes]

OMOP – Observational Medical Outcomes Partnership; graphic according to https://www.ohdsi.org/data-standardization/the-common-data-model/
Example: Safety of NSAIDs

- EU-project “Safety of non-steroidal anti-inflammatory drugs (NSAIDs)” (SOS)
  - Coordination: Erasmus University Medical Centre, Rotterdam, NL
  - 7th EU Framework Programme

- Five population-based healthcare databases from four European countries (Netherlands, Italy, Germany, UK)

- Revealed more serious cardiovascular side effects for Diclofenac than for other traditional NSAIDs

- New safety warning by the European Medicines Agency (EMA)
  - Diclofenac subject to the same safety precautions as selective COX-2-inhibitors

OUTLOOK

National Research Data Infrastructure for Personal Health Data

www.nfdi4health.de
National Research Data Infrastructure

- Initiative of German Federal and Länder governments (German Joint Science Conference)
  - Based on recommendations of German Council for Scientific Information Infrastructures (RfII)

- Aims
  - Sustainable securing, indexing and utilization of research data via regional and networked knowledge repositories
  - Establishment of research data management according to FAIR principles
  - Connection and networking to international initiatives such as European Open Science Cloud (EOSC)

- Budget: 90 m€ per year (final stage)
- Up to 30 consortia shall cover science landscape
- First nine consortia started in October 2020
NFDI4Health – Consortium

- **Lead:** ZB MED – Information Centre Life Sciences
- **Co-lead:** Leibniz Institute for Prevention Research and Epidemiology – BIPS
- **Target data:** Health data from registries, cohort studies, clinical trials, administrative health databases, epidemiological and public health studies
- **17 partners and 48 further participants**

- **Vision:** To boost the scientific exploitation of personal health data
NFDI4Health – Aims

• To enable findability of and access to structured health data
• To maintain federated framework of data holding organisations
• To enable privacy preserving exchange and linkage of personal health data
• To develop automated services (e.g., use and access, analysis tools)
• To enhance interoperability and reusability
• To promote use case oriented cooperation between research communities
There is a strong argument to be made that leaving data unshared is an impediment to the scientists of the future.
Thank you very much for your attention!

www.leibniz-bips.de/en

Contact
Prof. Dr. Iris Pigeot
Leibniz Institute for Prevention Research and Epidemiology – BIPS
Achterstraße 30
D-28359 Bremen
pigeot@leibniz-bips.de