



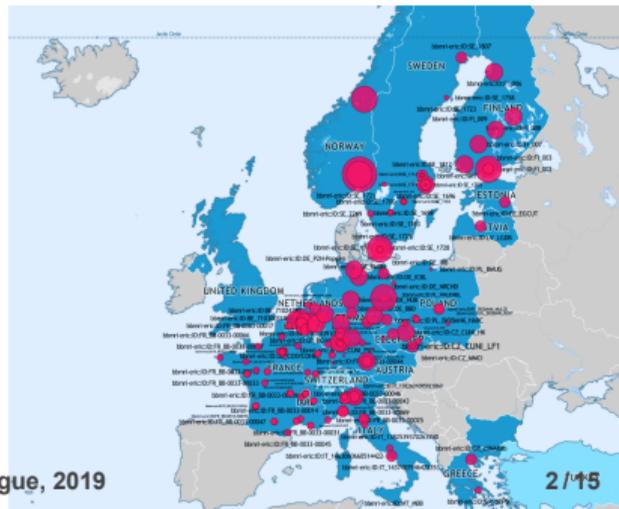
MAKING
NEW
TREAT
MENTS
POSSIBLE

**SHARING OF RESEARCH DATA IN ACADEMIC
ENVIRONMENTS – TRANSNATIONAL EXPERIENCE
FROM BBMRI-ERIC**

Assoc. Prof. RNDr. Petr Holub, Ph.D.

WHAT IS BBMRI-ERIC?

- ▶ **One of largest European RIs for medical research**
 - spans **20 member states** + IARC
 - **federated** biobanks and other resources across members states
 - makes **biological samples & medical/health data more FAIR** (findable, accessible, interoperable, reusable) while also compliant to **privacy protection** regulations
 - facilitates adoption of **quality procedures** to **improve reproducibility** of medical research



SPECIFICS OF PERSONAL RESEARCH DATA

- ▶ What it is: **personal data used for research purposes**
 - health, genetic, geo-location (for exposure), ...
- ▶ What makes **research data specific**?
 - many things are completely different compared to personal data processing for other purposes
 - research is typically **multi-center/international**
 - research needs to be **verifiable** and **reproducible**
 - pressure on publishing research data as FAIR or open
 - need to store data in research infrastructures: **quality-assured data** made available for reuse
 - even industrial research now considers data sharing/pooling: dealing with **mounting costs of collecting high-quality data**

RESEARCH DATA IN CZ UNDER GDPR

- ▶ All the public **Czech universities teamed up to develop joint recommendations** how to deal with research data under GDPR
 - <https://doi.org/10.5281/zenodo.2532860>
 - goal was to utilize flexibility given by the GDPR for research purposes (e.g., where informed consent is not needed) – i.e., **GDPR perceived beneficial**
 - procedural recommendations (e.g., event handling)
 - being updated now after adoption of national GDPR implementation – minor updates only

**Metodika aplikace GDPR na
výzkumná data v prostředí
vysokých škol v ČR**

Radim Polčák, Leoš Ševčík, Michal Koščík,
Jakub Klodwig, Petr Holub

Masarykova univerzita

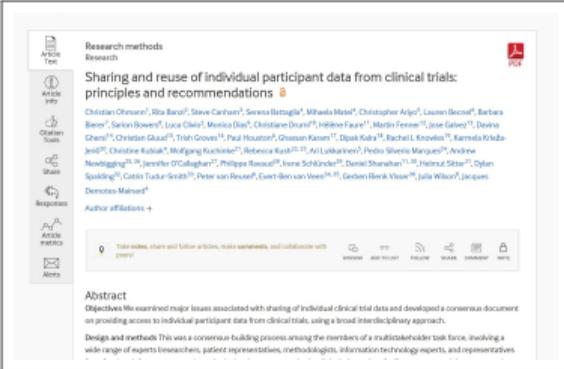


2018

SHARING DATA FROM CLINICAL TRIALS

▶ Ohmann, Christian, et al. **”Sharing and reuse of individual participant data from clinical trials: principles and recommendations.”**

- <https://bmjopen.bmj.com/content/7/12/e018647.abstract>
- outcome of CORBEL project
- 10 principles and 50 recommendations
- consent management
- protection of trial participants
- data standards, rights, types and management of access
- data management and repositories, discoverability, and metadata



The screenshot shows the top portion of a research article page. On the left is a navigation sidebar with icons for 'Article Text', 'Article Info', 'Citation Tools', 'Share', 'Responses', and 'Alerts'. The main content area is titled 'Research methods Research' and features the article title 'Sharing and reuse of individual participant data from clinical trials: principles and recommendations'. Below the title is a list of authors: Christian Ohmann, Rita Baro, Steve Carreau, Serena Battaglia, Mihaela Most, Christopher Argo, Lauren Becroft, Barbara Bener, Gordon Brown, Luca Chiar, Monica Day, Christine Dixon, Hildegarde Faur, Marlene Frenn, Jose Galvez, David Ganev, Christian Guad, Trish Green, Paul Huxford, Ghassan Karam, Dushyant K, Barbara Kowale, Samira Kribala, Christine Kubak, Mufgang Kuchel, Rebecca Kulkarni, Arif Lubbukov, Pedro Serrano Marques, Andrew Newbigging, Jennifer O'Callaghan, Philippe Rissal, Irene Schindler, Daniel Shanahan, Helmut Sitter, Dylan Spalding, Cassin Tucker-Smith, Peter van Rensen, Evert-Jan van Veen, Gordon Bank-Voss, Julia Wilson, and Jacques Demotes-Mainard. Below the author list is a search bar with the text 'Take notes, share and follow articles, make comments, and collaborate with peers!' and icons for 'SEARCH', 'ADD TO LIST', 'FOLLOW', 'SHARE', 'COMMENT', and 'MORE'. The 'Abstract' section begins with the text: 'Objectives We examined major issues associated with sharing of individual clinical trial data and developed a consensus document on providing access to individual participant data from clinical trials, using a broad interdisciplinary approach. Design and methods This was a consensus building process among the members of a multistakeholder task force, involving a wide range of experts (clinicians, patient representatives, methodologists, information technology experts and representatives'.

USING PERSONAL OR ANONYMIZED DATA?

- ▶ **Pseudonymized data is personal data**
 - identifier(s) of a pseudonym is replaced by a pseudonym
- ▶ **Anonymized data is non-personal data**
 - anonymization is not a perfect process (nothing like perfectly anonymized while still useful exists)
 - anonymization is about finding balance between damage of the data and privacy protection
 - best known anonymization technique – differential privacy – was questioned by WP29 (→ EDPB)

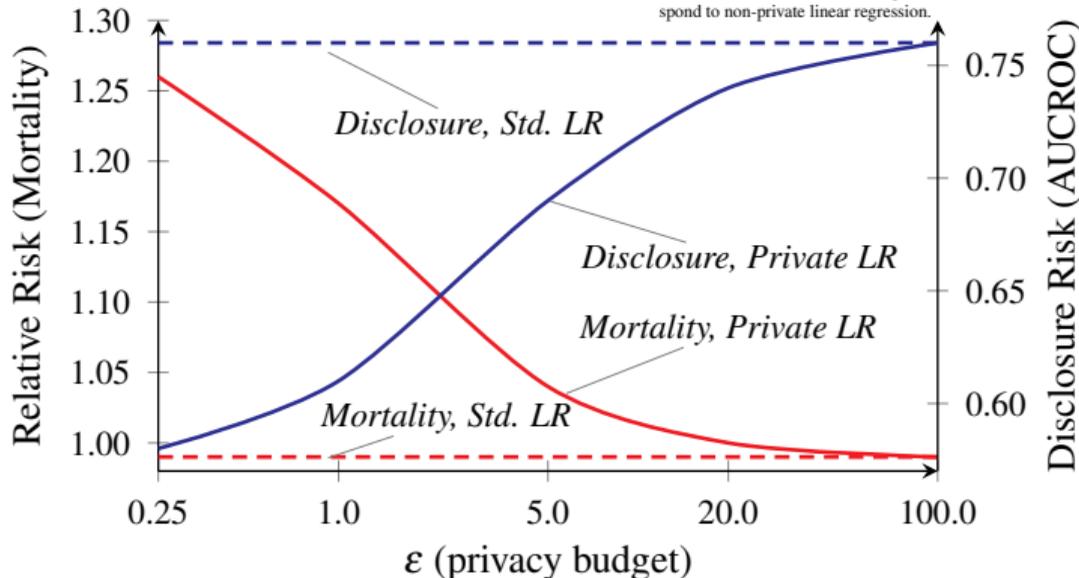
USING PERSONAL OR ANONYMIZED DATA?

- ▶ **For medical research it is almost always better to work with pseudonymized data**
 - one typically has or can obtain a legal basis – justified interest or informed consent
 - possible effect caused by data damage are highly undesirable and helps with reproducibility
 - pseudonymized data allows dealing with incidental findings

USING PERSONAL OR ANONYMIZED DATA?

Compromise of utility vs. privacy¹

Figure 1: Mortality risk (relative to current clinical practice) for, and VKORC1 genotype disclosure risk of, ϵ -differentially private linear regression (LR) used for warfarin dosing (over five values of ϵ , curves are interpolated). Dashed lines correspond to non-private linear regression.



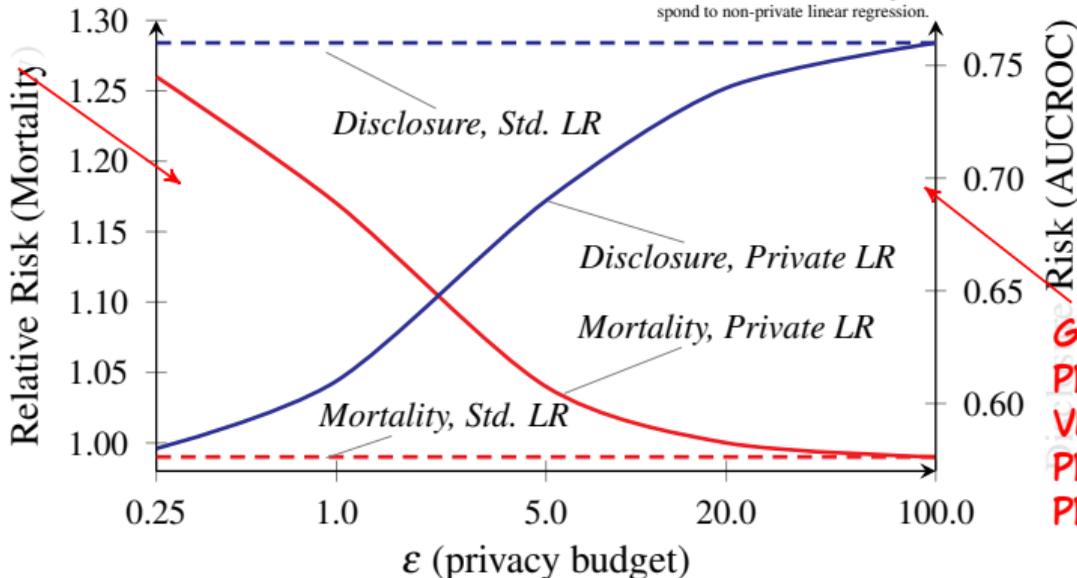
¹ Fredrikson, M., Lantz, E., Jha, S., Lin, S., Page, D., & Ristenpart, T. (2014). Privacy in Pharmacogenetics: An End-to-End Case Study of Personalized Warfarin Dosing. Proceedings of the ... USENIX Security Symposium. UNIX Security Symposium, 2014, 17–32. Retrieved from <http://www.biostat.wisc.edu/page/WarfarinUsenix2014.pdf>

GOOD PRIVACY PROTECTION, BUT TENDS TO KILL PATIENTS.

PERSONAL OR ANONYMIZED DATA?

Trade-off of utility vs. privacy¹

Figure 1: Mortality risk (relative to current clinical practice) for, and VKORC1 genotype disclosure risk of, ϵ -differentially private linear regression (LR) used for warfarin dosing (over five values of ϵ , curves are interpolated). Dashed lines correspond to non-private linear regression.



GOOD TREATMENT PREDICTION, BUT VERY POOR PRIVACY PROTECTION.

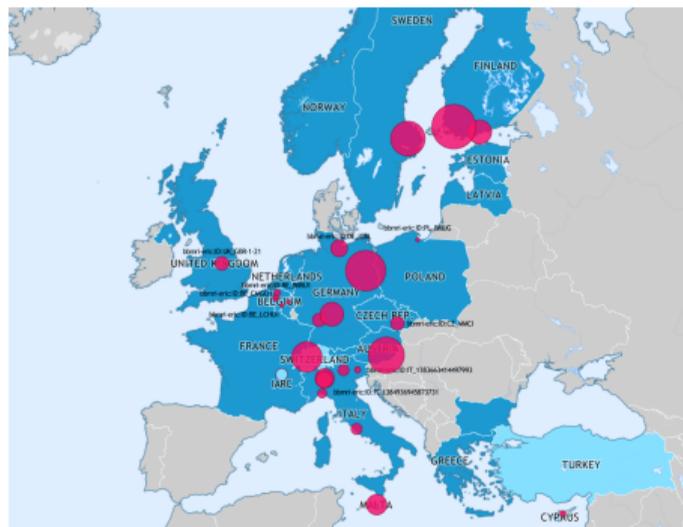
¹ Fredrikson, M., Lantz, E., Jha, S., Lin, S., Page, D., & Ristenpart, T. (2014). Privacy in Pharmacogenetics: An End-to-End Case Study of Personalized Warfarin Dosing. Proceedings of the ... USENIX Security Symposium. UNIX Security Symposium, 2014, 17–32. Retrieved from <http://www.biostat.wisc.edu/page/WarfarinUsenix2014.pdf>

CHALLENGES OF (INTERNATIONAL) SHARING

- ▶ **Problem of national GDPR derogations and additional national regulations**
- ▶ **Difficult to get Codes of Conduct under Art. 40 ready and approved**
 - e.g., 3 codes in preparation for cloud computing
 - e.g., 2 codes for medical data sharing (one organized by BBMRI-ERIC)
 - process takes years, unclear process in the beginning
 - EDPB slow to start
- ▶ **Problem with **status of full international organizations** (e.g., EBI/EMBL which is global resource of bioinformatics DBs)**
 - ERICs are much easier – defined jurisdiction of EU and hosting country.

CRC-COHORT – EXAMPLE OF THE PROCESS

- ▶ 10,000 colorectal cancer cases spread over the Europe
 - BBMRI-ERIC is coordinator, host, and custodian



CRC-COHORT – EXAMPLE OF THE PROCESS

- ▶ **10,000 colorectal cancer cases spread over the Europe**
 - examples of delicate complexities
 - Finnish national derogations and additional regulations prevent data to leave Finland permanently (also influences dbSNP, EGA, and other major European/global databases)
 - Austrian implementation of derogations for research make data handling very delicate – all accesses must be logged
 - lack of cloud computing Code complicates storage and processing – currently only national compliance (BBMRI-ERIC in AT, but storage in IT)
 - Czech implementation of GDPR is one of the least problematic

CONCLUSIONS

- ▶ **Research data is a specific topic** because of global nature of science
- ▶ **GDPR is actually positive step forward** for research data sharing – clarifies and harmonizes many things
 - performative nature of GDPR allows for good flexibility
 - gives good flexibility for research purposes, unless hampered by national derogations and/or additional regulations
- ▶ Adoption and further development of **common guidelines** needed
 - approval of Art. 40 Codes needs to accelerate
- ▶ **ERICs are a good framework for facilitating research data** sharing in Europe as demonstrated by CRC-Cohort developed by BBMRI-ERIC

MAKING
NEW
TREAT
MENTS
POSSIBLE

THANK YOU!

 contact@bbmri-eric.eu

 www.bbmri-eric.eu

 @BBMRIERIC

 [BBMRI-ERIC](https://www.linkedin.com/company/bbmri-eric)