

Zu den normativen Herausforderungen von Big Data in der Forschung

Jahrestagung des Deutschen Ethikrats, Berlin 21. Mai 2015

Prof. Dr. iur. Nils Hoppe

Centre for Ethics and Law in the Life Sciences, Leibniz Universität Hannover
www.cells.uni-hannover.de

Übersicht

- Einleitendes
- Fallstudien: **care.data** und **UK Biobank**
- Wesentliche Herausforderungen
- Lösungsstrategien
- Ausblick und abschließende Gedanken

Big Data – Big Problem?

- Welche Diskussionen entstehen überhaupt genuin *neu* durch Big Data?
- Grundlegende Fragen der Regulierung neuer Technologien:
 - Technologien als Regulierungsziele *und* als Regulierungsinstrumente;
 - Innovation als flüchtiger Regulierungsgegenstand;
 - Überregulierung als Wissenschaftshemmnis.

Fallstudien: Big Data & Forschung



NHS care.data

GP and hospital records into the Health and Social Care Information Centre Database; updated every month; opt-out only system

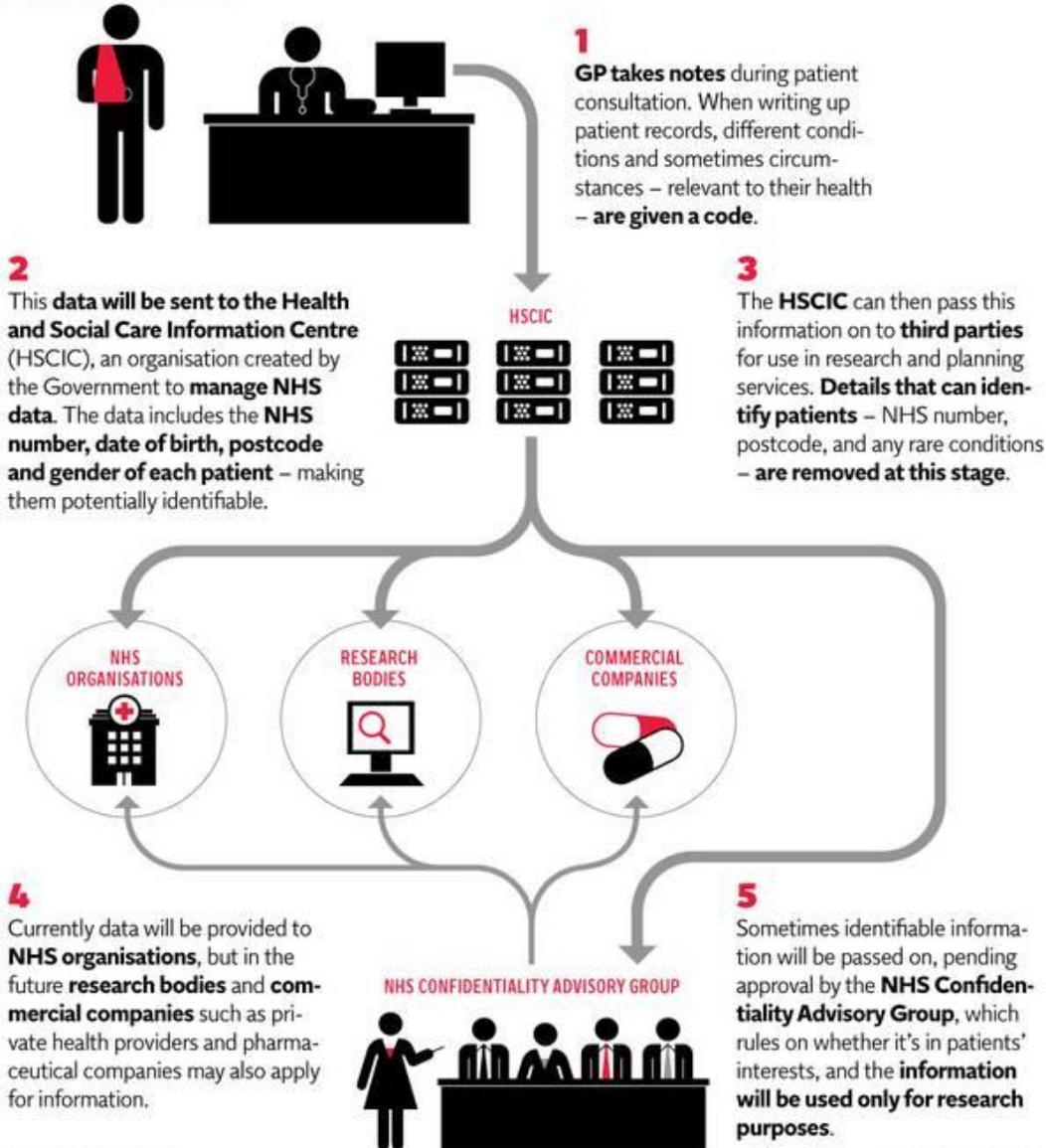
UK Biobank

500,000 participants aged 40-69; blood, urine, saliva samples; detailed health questionnaires; sustained following of individuals' health; medical imaging

NHS care.data

- Überführung von Hausarztzeichnungen, Krankenhausdaten, etc in eine zentrale Datenbank.
- Verknüpfung mit existierenden Datenbeständen, 'Matching' von Datensätzen zu konsistenten Personendatensätzen.
- Nutzung zur Qualitätskontrolle, öffentlich geförderte Forschung, privat geförderte Forschung (u.a. Pharma).

FOR THE RECORDS HOW CARE.DATA WILL WORK



GRAPHIC: JOHN BRADLEY

WORDS: CHARLIE COOPER

Quelle:

<http://www.norvichealthcare.co.uk> (14/05/15)

NHS care.data

- Grundlage für Datensammlung per Gesetz eingeführt (Health and Social Care Act 2012).
- Kritik: Mangelhafte Konsultation im Vorfeld; mangelhafte Aufklärung über Widerspruchsmöglichkeiten; unklare Datenlage; unklare Weitergabe an Dritte.
- Signifikante öffentliche Ablehnung 2014 führt zur Suspendierung des Systems.
- Neuausrichtung, Beteiligung, 'Safeguarding' bis Ende 2015.

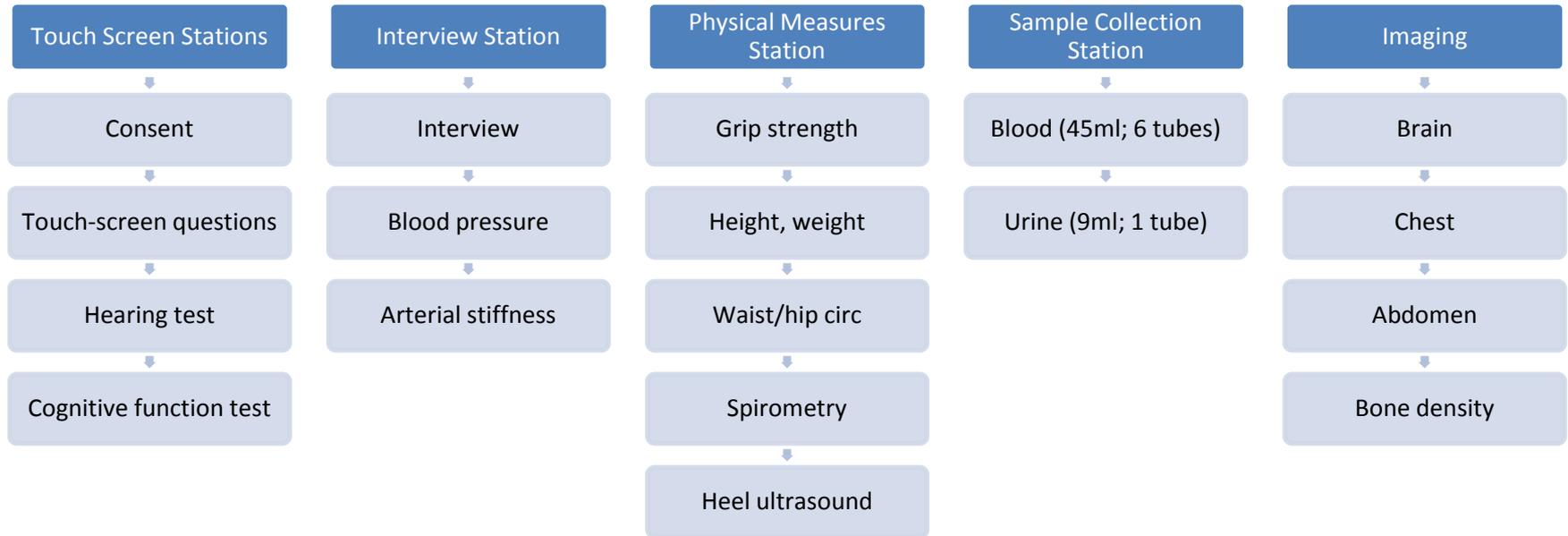
UK Biobank

- Nationale Biobankeninitiative – Sammlung von Daten von 500k Probandinnen und Probanden 2006-2010.
- Material- und Datensammlung, ‘long-term follow-up’ – dauerhafte Datenverbindung mit NHS Datensatz.
- Bestimmte Forschungsprojekte haben Zugriff auf Material und Daten: Krebserkrankungen, Herzerkrankungen, Schlaganfall, Diabetes, Arthritis, Osteoporose, Augenerkrankungen, Depression und Demenzerkrankungen.

UK Biobank

- Erhebung eigener Materialien (10 Mio Proben) und Datensätze (10 Mrd Datenpunkte).
- Verknüpfung mit Datenverarbeitung der Hausärzte.
- Regelmäßige Auffrischung der Datensätze.
- Bildgebende Verfahren
- Genotypisierung der Probandenproben durch *Affymetrix* (Axiom Biobank Genotyping Array, \approx 650k Genvariationen).

UK Biobank



- + Verknüpfung mit nationalen Datensätzen (Todesfälle; Krebsdiagnostik; Krankenhausdaten; Hausarzt Daten)
- + Sog. “Re-assessment visits”

Wesentliche Herausforderungen

- Wesentliche Herausforderungen sind Ergebnis der hohen *Datendichte* und *Granularität*, bei hoher *Unsicherheit* über Verwendungszweck und –nutzen der Daten:
 - Einwilligungsfragen, Rücknahme der Einwilligung;
 - Zufallsbefunde, Informationsmanagement;
 - (de-)Identifizierung von Daten;
 - Zugriff und Verknüpfung.

Einwilligungsfragen

- **care.data**
 - Gesetzliche Widerspruchslösung;
 - fehlende Beteiligung im Vorfeld;
 - mangelhafte Information über Rechte.
- **UK Biobank**
 - Klassische (unproblematische) Einwilligungspunkte: Körperverletzung, Datenerhebung, -speicherung;
 - problematische Einwilligungspunkte: Zukünftige Verwendung für die Forschung (klassisches Biobankendilemma).
- Generell: hochskalierte Datensammlung, -verknüpfung verändert den Charakter der Datensammlung und der möglichen Nutzungsweisen.
- Konsequenzen für die “Informiertheit” der Einwilligung?

Zufallsbefunde

- Frage des Informationsmanagements
- Grdgl. Problem des Forschungsdesigns:
 - Kommunikation oder nicht;
 - Ausschluß von Probanden und Probandinnen, die auf Kommunikation nicht verzichten wollen.
- “Risiko” eines einschlägigen Befundes steigt mit der Menge der verfügbaren Daten und der Automatisierung der Prozesse.
- Bilderkennung für bildgebende Verfahren
- Automatische Befunderhebung Teil des Forschungsdesigns (Kohortenzuordnung), negiert die Ablehnung von Pflichten aus dem nicht-klinischen Kontext.

De-Identifizierung

- Hohe Datendichte führt zu einer mangelhaften De-Identifizierung der Datensätze.
- Zwei Ansätze: *Expert Determination* und *Safe Harbor*
 - Technische Möglichkeiten ändern sich (sonst wären wir heute nicht hier!).
 - Entfernung von identifizierenden Merkmalen nach HIPAA nicht ausgerichtet auf Genomsequenzierung etc.

De-Identifizierung

- Zusätzlich: UK Biobank will inzwischen in bestimmten Fällen Feedback geben (z.B. bildgebende Verfahren).
- Eine vollständige de-Identifizierung von Informationen führt zu einem Problem bei der Mitteilung von Zufallsbefunden.
- Auskunftspflichten werden durch vollumfängliche de-Identifizierung stark erschwert.

Zugriff

- Insbesondere Befürchtungen einer forensischen Nutzung der großen Datenmengen.
- Verknüpfung von care.data mit bestehenden Regierungsdatenbanken vorgesehen (z.B. Führerscheinstelle); Recherche in Gesamtdatensatz nicht ausgeschlossen.
- Seit Marper v UK – klar, dass eine forensische Datenbank mit Genotyp-Daten existiert.

Lösungsstrategien

- **Einwilligung**
 - Radikale Lösung: Widerspruchslösung
 - Flucht nach vorn – dynamic consent
- **Zufallsbefunde**
 - Ausschluss von Probanden/Probandinnen mit Informationswunsch.
 - Zweiwegesystem für Informationsfluss zum Hausarzt.
 - Ggf. gesetzliche Lösung, Rechtfertigungsgrund für Nichtmitteilung.
- **De-Identifizierung**
 - Neudiskussion unter Zuhilfenahme von Terminologie aus der Sicherheitstechnik; mehr Datensicherheit, weniger Datenschutz.
 - Diskussion von Technologie, die Daten mit einem Apoptoseparameter ausstattet.

Lösungsstrategien

- **Zugriff**
 - **care.data:**
 - Dem Schicksal ergeben;
 - keine Bemühungen, die gesetzliche Grundlage um z.B. eine besondere Vertraulichkeit oder Zeugnisverweigerungsrecht zu erweitern;
 - neidische Blicke auf die deutsche Diskussion (!) um ein Biobankgeheimnis.
 - **UK Biobank**
 - Intensive Diskussionen im EGC;
 - klares Erkennen des Vertrauensschadenrisikos für die Ressource;
 - Einrichtung eines Verteidigungsfonds für forensische Anfragen (private Rechtsetzung?).

Ausblick

- Regulierungsansätze und technische Machbarkeit: *Gartenzaun-Effekt*.
- Hohe Mengen an Daten waren auch mal ein Schutz: *DDOS, Informations-Filibuster*.
- Big Data sollte ggf. ein Impuls sein:
 - Grundlegend über Kontrolle von Daten zu sprechen;
 - technische Machbarkeit von Teilhabe auszuschöpfen;
 - negative Effekte zu regulieren und nicht jede Verarbeitung von Informationen.

Prof. Dr. iur. Nils Hoppe

LEIBNIZ UNIVERSITÄT HANNOVER

Centre for Ethics and Law in the Life Sciences

Am Klagesmarkt 14-17

30159 Hannover

www.cells.uni-hannover.de

nils.hoppe@cells.uni-hannover.de