



MOSAIC

# MOSAIC: Zentrales Datenmanagement

## als Methode zur Verbesserung der Nutzbarkeit medizinischer Forschungsdaten

Martin Bialke

Institut für Community Medicine (Abt. VC)

Universitätsmedizin Greifswald

25. Sept. 2013

- MOSAIC: a **modular systematic approach to implement a central data management**"
- DFG-Ausschreibung „Informationsinfrastrukturen für Forschungsdaten“
- Laufzeit: 3 Jahre, Projektstart im Aug. 2012
- Zentrales Datenmanagement mit Fokus auf Kohortenstudien im klinisch-epidemiologischen Kontext
  
- Ziele:
  - Prozesse, Abläufe und Techniken generalisieren
  - Aufbau eines modularen Unterstützungssystems
  - Anleitungen, Werkzeuge, Musterlösungen begleiten sämtliche Phasen des Datenlebenszyklus
  - Zielgruppe: Forschungsgruppe mit IT-Personal



- Erfassung und Verfügbarmachung med. Daten im Studien- und Versorgungskontext
  - Heterogene, verteilte Quellsysteme
  - vielfältige Datentypen und -strukturen
  - Metadaten zur Dokumentation des Entstehungskontexts und zur Sicherstellung der Interpretierbarkeit
- fachlicher und technischer Umfang eines Zentralen Datenmanagements variiert je Studienszenario
  - Gerätetypen, Schnittstellen
  - Rechtliche und ethische Rahmenbedingungen
  - Erfassungsformulare und Informed Consent
  - Integrität und Qualität



- Besondere Schutzwürdigkeit der med. Daten erfordert strukturiertes, reproduzierbares Vorgehen
- Gesucht: anpassbare Architektur, um
  - Werkzeuge, Methoden und Lösungen nachnutzbar zu gestalten,
  - Methoden und Abläufe zu standardisieren
  - und so Vergleichbarkeit erhobener Daten zu unterstützen.

# Methoden – Erfahrungen und Vorarbeiten



MOSAIC

- epidemiologische Forschungsprojekte mit Zentralem Datenmanagement am ICM VC der Universitätsmedizin Greifswald (Auszug)



**DZHK**

DEUTSCHES ZENTRUM FÜR  
HERZ-KREISLAUF-FORSCHUNG E.V.



**SHIP**

Study of Health in Pomerania



**DelpHi-MV**



ZENTRALES KLINISCHES  
**KREBSREGISTER**  
Mecklenburg-Vorpommern



**GANI MED**

Greifswald Approach to Individualized Medicine



Community  
Medicine  
im Neugeborenenalter

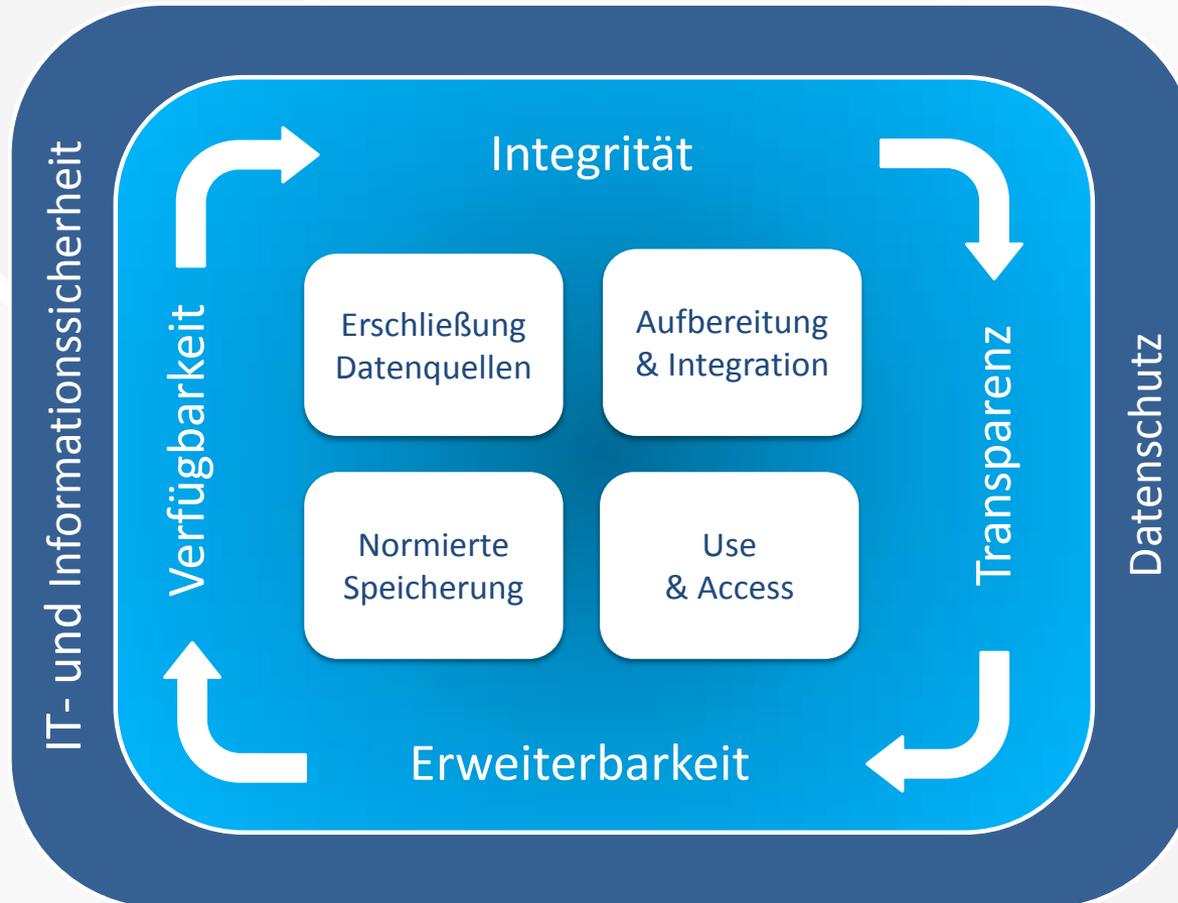


**DZNE**

Deutsches Zentrum für  
Neurodegenerative Erkrankungen  
in der Helmholtz-Gemeinschaft



**Nationale Kohorte e.V.**



© Universitätsmedizin Greifswald 2013

## Zentrales Datenmanagement



- Zentrales Datenmanagement:
  - umfasst sämtliche Phasen des Datenlebenszyklus
  - Reihe organisatorischer / technischer Aufgaben, methodischer und konzeptioneller Prozesse
- Nutzungsvoraussetzungen:
  - ZDM bereits in Studiendesign einplanen
  - IT-Personal zur fachlichen Unterstützung der Forscherteams





- Zentrales Datenmanagement mit Fokus auf Primär- und Metadaten der medizinischen Forschung
  - > Ausschluss von Sekundärdaten und Bioproben
- Folge sind datenschutzrechtliche und ethische Forderungen:
  - Eineindeutige Identifikation von Probanden/Patienten
  - Verwaltung von Einverständnissen, Ermächtigungen und Widerruf
  - Pseudonymisierung / Anonymisierung von Daten (und Bioproben)



- Unterstützungssystem als Sammlung von Werkzeugen und Lösungsstrategien





\* = aktuell in Bearbeitung

- Konkret (Auszug):
  - Planungswerkzeug\*
  - Template-Generator zur Erstellung von Datenschutzkonzepten (für multizentrische Studien)\*
  - Pseudonym-, Consent- und ID-Management\* als autarke Dienste
  - generisches Datenmodell
  - Konzepte zur Umsetzung des Datenmodells in DBMS und Data Warehouse
  - Strategien zur Sicherung und Archivierung des Datenbestandes
  - Evaluierung von Schnittstellenwerkzeugen
  - Methoden und Beispiellösungen für Qualitätssicherung
  - Workflows und Templates zur Datenbereitstellung und Rückführung



- Verbesserung der Nachnutzbarkeit durch
  - Systematische Kapselung von Verantwortlichkeiten
  - Generalisierung bestehender Lösungen und Ansätze
  - Standardisierung von Methoden und Abläufen
  - Gewährleistung IT-Sicherheit und Datenschutz



- Erfolg des Konzepts
  - einheitliche Datenerfassung heterogener Quellen
  - Datenspeicherung in homogenem Data Repository
  - Standardisierte Bereitstellung vergleichbarer Daten
- Aussicht
  - Entwicklung zusätzlicher Werkzeuge und weiterführende Evaluierung bisheriger Ergebnisse

Literatur:

**Meyer J, Ostrzinski S, Fredrich D, Havemann C, Krafczyk J, Hoffmann W.** Efficient data management in a large-scale epidemiology research project. *Comput Methods Programs Biomed.* 2012, Bd. 107, 3.

**Sariyar M, et al.** A practical framework for data management processes and their evaluation in population-based medical registries. *Informatics for Health and Social care.* 2013, Bd. 38, 2.

**Fraser HSF, et al.** Adaptation of a web-based, open source electronic medical record system platform to support a large study of tuberculosis epidemiology. *BMC Medical Informatics and Decision Making.* 12, 2012, Vol. 125.

# Vielen Dank für ihre Aufmerksamkeit



MOSAIC



Institut für Community Medicine  
Abt. Versorgungsepidemiologie und Community Health  
Ellernholzstr. 1/2  
17487 Greifswald

## Konzeption und Entwicklung

Martin Bialke

Tel.: 03834 867580

E-Mail.: [martin.bialke@uni-greifswald.de](mailto:martin.bialke@uni-greifswald.de)

Jens Piegsa

Tel.: 03834 867721

E-Mail.: [jens.piegsa@uni-greifswald.de](mailto:jens.piegsa@uni-greifswald.de)

Weitere Informationen unter

**[mosaic-greifswald.de](http://mosaic-greifswald.de)**