

# Anmerkungen zu einem angemessenen datenschutzrechtlichen Rahmen für medizinische Forschung und statistische Evaluierung

Waltraut Kotschy

MedInfo Symposium  
27. Sept. 2007, Graz

# Anwendungsbereich des Datenschutzes

- Nur auf ‘**personenbezogene Daten**’, d.s. Daten, die sich auf **identifizierte oder identifizierbare Menschen** beziehen = höchstpersönliches Recht, daher **nur für lebende Menschen**

Spezielle Probleme:

- Daten über Verstorbene: kein Datenschutz, aber Schutz von Persönlichkeitsrechten z.B. nach § 16 ABGB
  - Daten über Ungeborene → ??
  - Medizinische Daten sind oft zugleich Daten mehrerer Betroffener, nämlich auch anderer Familienmitglieder
- ‘**Anonyme Daten**’ beziehen sich auf nicht identifizierbare Individuen → nicht Gegenstand von Datenschutz

# Ergebnis medizinischer Forschung

- Grundsätzlich nie personenbezogen
- Worin besteht dann eigentlich das **Datenschutzproblem?**

Auf dem Weg bis zum Ergebnis

# Ausgangspunkt medizinischer Forschung?

- Substrat der Forschung sind **Echtdaten** –  
Quelle von Echtdaten sind  
Krankengeschichten  
Ärztl. Dokumentation  
Patientenbefragung  
etc.



Daten, die meist in personenbezogener Form gespeichert sind

# Was sind „personenbezogene Daten“?

- § 4 Z 1 DSGVO 2000: Daten über  
identifizierte oder  
identifizierbare Personen

- Wann ist eine Person „identifiziert“?

Es gibt keine gesetzliche Definition –

üblicherweise: ausreichende Anzahl der traditionellen  
Identifikatoren wie **Name**, Geburtsdatum, etc.

- Wann ist eine Person „identifizierbar“?  
???

# Wann sind Personen 'identifizierbar'?

z.B. bei 

ID-NUMMER
-----------

Inhaltsdaten
--------------

  
nur wenn Zugang zur Konkordanzliste besteht

Geb.Datum, BERUF, PLZ, GESCHLECHT
-----------------------------------

Seltene Krankheit
-------------------

Abbild einer Person
---------------------

DNA-PROFIL
------------

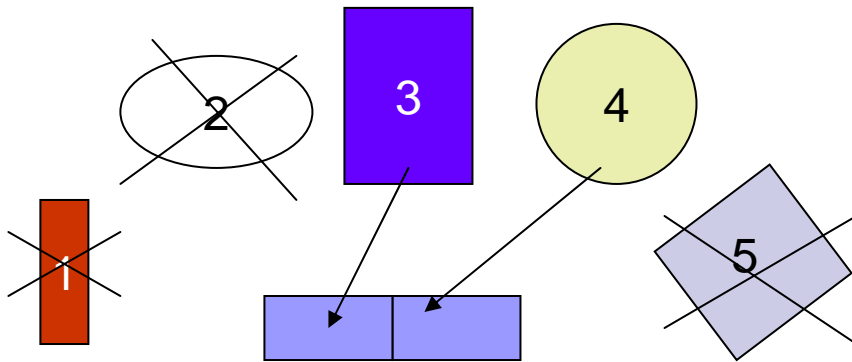
Identifizierbarkeit hängt davon ab, wie wahrscheinlich es ist, dass jemand die notwendige Zusatzinformation zur Identifizierung erlangen kann

# Wann sind Daten „anonymisiert“?

Wenn **keine Identifikatoren** vorliegen **UND**  
der **Inhalt keine Elemente** enthält, die den  
Betroffenen identifizierbar machen  
= **echt anonymisiert**

# Warum verwendet man nicht immer anonymisierte Daten?

## ■ Datenermittlung:



Aus **verschiedenen Quellen**  
oder  
für **verschiedene Zeitpunkte**,  
aber  
für **dasselbe Individuum**

**➔** daher sind **anonymisierte Daten oft nicht brauchbar** – die Datensätze brauchen einen Identifikator, d.i. üblicherweise der Name des Patienten

# Wissenschaftsprivileg

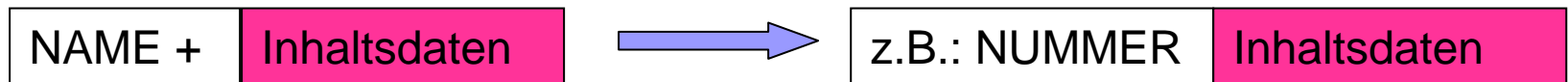
- Gemäß Art. 6 (1) b der Datenschutz-Richtlinie 95/46 ist die **Weiterverwendung von personenbezogenen Daten für die wissenschaftliche Forschung** und für Statistik „**nicht** als **unvereinbar** anzusehen, vorausgesetzt dass die Mitgliedstaaten geeignete **begleitende Schutzmaßnahmen** vorsehen“.
- **§ 46 DSG 2000 ist die österr. Umsetzung**

# § 46 Abs. 5 DSGVO 2000

- Pflicht, personenbezogene Daten, die für Zwecke wissenschaftlicher Forschung verwendet werden, so bald als möglich zu **anonymisieren** oder zumindest zu **pseudonymisieren** - am besten noch vor der Übermittlung der Daten an den Forscher  
= begleitende Schutzmaßnahme im Sinne des Art. 6 (1) b der RL 95/46

# Pseudonymisierung – was ist das?

- Die Identifikatoren werden mittels eines Algorithmus transformiert:



- **Pseudonymisierung** =  
**Individualisierung ohne Identifikation**
- Die Art des Algorithmus und die Größe der Identifikatorenbasis ist maßgeblich für die Frage, ob Re-Identifikation möglich ist oder nicht:  
Einwegverschlüsselung über eine große Menge von Identifikatoren (vielstellige Identifikatoren) erzeugt Pseudonyme, die ziemlich sicher gegen Re-Identifizierung sind.

# Ermittlungsprivileg für pseudonymisierte Daten

- Pseudonymisierte Daten dürfen **ohne datenschutzrechtliche Beschränkungen** übermittelt und weiterverwendet werden
- **Voraussetzung:**
  - Die Daten wurden **VOR** der Übermittlung pseudonymisiert – wenn erst der Empfänger pseudonymisiert, gilt das Privileg nicht („indirekt personenbezogene Daten“)
- **Wer übernimmt den Aufwand der Pseudonymisierung?**
  - Fehlen von professionellen Pseudonymisierungsstellen ist ein wesentliches praktisches Problem

# Ermittlung nicht-pseudonymisierter Daten?

- Weiterverwendung von Daten aus der Behandlung für Forschungszwecke:

nur beim selben Auftraggeber erlaubt  
(Med.Universitäten !)

 in diesem Fall ist Zweckwechsel unbeachtlich

- Zustimmung der Betroffenen
- Genehmigung der Datenschutzkommission
  - Einholung der Zustimmung aller Betroffenen unmöglich oder unverhältnismäßiger Aufwand
  - Wichtiges öffentliches Interesse
  - Fachliche Befähigung

# Pseudonymisierungstechniken

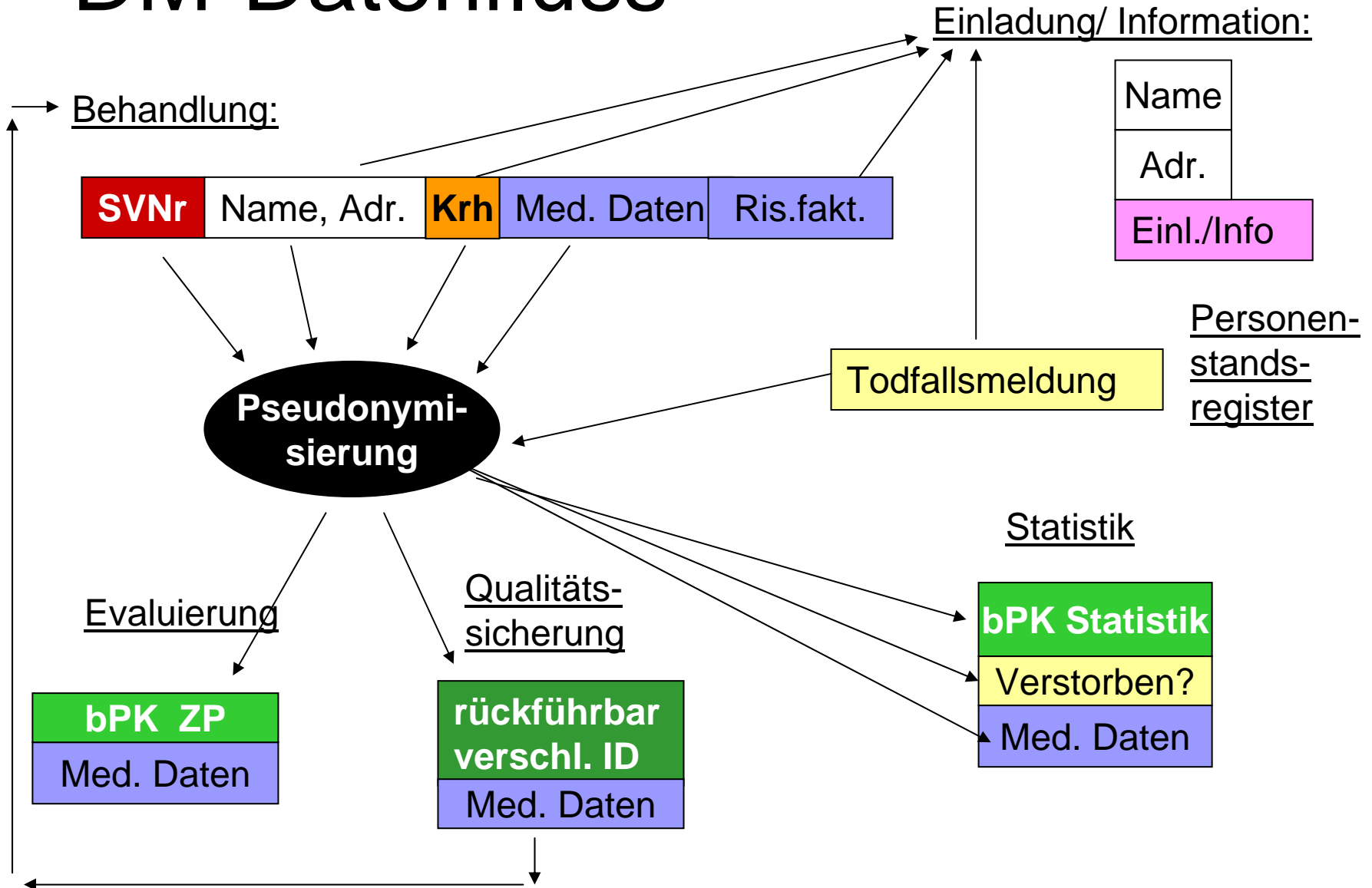
- Österr. E-Government-System bietet bereits einsatzfähige Instrumente hierfür:
  - Bereichsspezifische Personenkennzeichen (bPK)
- Wesentlich ist die Unterscheidung nach
  - Rückführbar (umkehrbarer) Pseudonymisierung und
  - Nicht rückführbarer Pseudonymisierung
    - durch Einwegverschlüsselung der Identifikatoren

# Module in Disease Management-Systemen

Verlangen unterschiedliche Grade des Personenbezugs:

- **Behandlung:** Name, SVNr.....
- **Einladung** zur regelmäßigen Kontrolle: Name, Adresse
- **Information** an den Patienten außerhalb des Behandlungsgesprächs: Name, Adresse
- **Qualitätssicherung:** rückführbar pseudonymisierte Identifikation des Patienten
- **Evaluierung:** möglicherweise anonymisierte Daten ausreichend
- **Statistik** Register (z.B. Tumorregister): nicht rückführbar pseudonymisierte Patientendaten
- etc.

# DM-Datenfluss



# Wo soll Verschlüsselung erfolgen?

- Beim Arzt ?

Probleme sind Arztwechsel,  
Schlüsselmanagement,  
unterschiedliche Verschlüsselungsmethoden

- Bei einer unabhängigen Pseudonymisierungsstelle:

Vorteil wäre eine nachhaltiges, professionelles Pseudonymisierungsmanagement mit sicherer Schlüsselverwaltung

# Pseudonymisierungsstelle als Drehscheibe

- Arzt schickt an geeignete Pseudonymisierungsstelle
  - **Identitätsdaten** „offen“ zwecks Erstellung der unterschiedlichen Pseudonyme
  - **medizinische Daten verschlüsselt**, sodass sie nur von den eigentlichen Empfängern gelesen werden können
    - (der gesamte Datensatz wird symmetrisch verschlüsselt und der Entschlüsselungsschlüssel wird asymmetrisch mit dem öffentlichen Schlüssel des jeweiligen Empfängers verschlüsselt)
- Pseudonymisierungsstelle schickt
  - **Unverschlüsselte Identitätsdaten an Einladesystem**
  - **pseudonyme Identitätsdaten** mit den **verschlüsselten medizinischen Daten** an die anderen Empfänger
  - könnte zusätzliche Dienste erbringen (Anfrage bei der Personenstandsbehörde über Todesfälle, bPK-Errechnung für Weiterleitung der Daten an Statistik....)