

# GPS - Gute Praxis Sekundärdatenanalyse

## Arbeitsgruppe Erhebung und Nutzung von Sekundärdaten (AGENS) der Deutschen Gesellschaft für Sozialmedizin und Prävention (DGSMP)

**Kurzfassung:** Die wissenschaftliche Nutzung von Sekundärdaten, vorrangig von Routinedaten der gesetzlichen Krankenversicherung, hat in den letzten Jahren kontinuierlich zugenommen. Die Arbeitsgruppe „Erhebung und Nutzung von Sekundärdaten“ (AGENS) der Deutschen Gesellschaft für Sozialmedizin und Prävention (DGSMP) ergriff deshalb die Initiative zur Formulierung von Qualitätsstandards für die Sekundärdatenanalyse. Als Ausgangsbasis dienten hierbei die vorhandenen Empfehlungen „Gute Epidemiologische Praxis“ (GEP), die durch AGENS in einem mehrstufigen Konsensprozess an die spezifischen Erfordernisse der Sekundärdatenanalyse angepasst wurden. Die Leitlinie „Gute Praxis Sekundärdatenanalyse“ (GPS) wurde am 15. Jan. 2005 verabschiedet.

Die GPS bestehen aus 10 Leitlinien, die in erläuternde Kommentare und einzelne Empfehlungen untergliedert sind. Neben der Etablierung von Standards für die Durchführung von Sekundärdatenanalysen stellen die GPS eine Grundlage für vertraglichen Absprachen zwischen Datenherra und Forscher dar.

Als Zielgruppe der GPS sind vor allem sozialmedizinisch und versorgungsepidemiologisch tätige Forscher angesprochen. Die AGENS verpflichtet sich zu einer kontinuierlichen Pflege und Fortentwicklung der GPS.

Version 1.00

Verabschiedet am 15. Januar 2005

Stand: 31. Januar 2005

Überarbeitung bis spätestens Januar 2007

An der Formulierung der GPS haben folgende Mitglieder der AGENS mitgewirkt: Susanne Bisson (Hannover), Wolfgang Bödecker (Essen), Siegfried Geyer (Hannover), Thomas Grobe (Hannover), Werner Hofmann (Hildesheim), Peter Ihle (Köln), Jürgen John (Neuherberg), Hedy Kerek-Bodden (Berlin), Susanne Moebus (Essen), Marina Mosafer (Frankfurt), Rolf Müller (Bremen), Matthias Perleth (Berlin), Burgi Riens (Göttingen), Ingrid Schubert (Köln), Daniel Sieg (Bielefeld), Philipp Storz (Berlin), Enno Swart (Magdeburg)

## Zielsetzung und Vorgehensweise

Die Nutzung von Sekundärdaten, vorrangig von Routinedaten der gesetzlichen Krankenversicherung, hat in den letzten Jahren kontinuierlich zugenommen. Möglichkeiten aber auch Grenzen der Nutzung derartiger Daten für Zwecke der (Versorgungs-)forschung, die primär zu anderen Zwecken erhoben wurden, werden kontrovers diskutiert. Zusätzlich zeigt der in das Sozialgesetzbuch eingefügte § 303a-f (Datentransparenz) einen neuen gesetzlichen Rahmen zur Nutzung von Daten der GKV auf.

Die Arbeitsgruppe "Erhebung und Nutzung von Sekundärdaten (AGENS)" der Deutschen Gesellschaft für Sozialmedizin und Prävention (DGSMP) sieht daher die Notwendigkeit, Qualitätsstandards für die Sekundärdatenanalyse zu formulieren und durch sie auf die Forschung anleitend zu wirken. Mit den Empfehlungen für eine „Gute Praxis – good practice – Sekundärdatenanalyse“ (GPS) verfolgt AGENS vorrangig zwei Ziele: Zum einen sollen allgemein akzeptierte Standards für die Durchführung von Sekundärdatenanalysen formuliert werden, zum anderen soll die GPS eine Grundlage für die notwendigen vertraglichen Absprachen zwischen Datenherra (Primärnutzer) und Forscher (als Zweitnutzer) darstellen. Darüber hinaus verpflichtet sich AGENS, die GPS zu pflegen und fortzuschreiben, um sie an veränderte Rahmenbedingungen anzupassen.

Ausgangspunkt für die Erarbeitung von Empfehlungen bilden die von einer Arbeitsgruppe der Deutschen Arbeitsgemeinschaft für Epidemiologie (da epi), der Dachorganisation der in Deutschland tätigen Epidemiologen, im Jahr 2000 veröffentlichten und im Sommer 2004 erstmalig revidierten Grundsätze „Guter Epidemiologischer Praxis“ (GEP) ([www.daepi.de](http://www.daepi.de)).

Zur Erarbeitung einer auf die spezifischen Anforderungen der Sekundärdatenanalyse bezogenen "Gute Praxis Sekundärdatenanalyse" wurde innerhalb der AGENS eine Redaktionsgruppe gebildet. Im ersten Schritt erfolgte eine Überprüfung der GEP hinsichtlich ihrer Übertragbarkeit auf die Sekundärdatenanalyse. Die Durchsicht der GEP zeigte, dass neben der grundsätzlichen Struktur einzelne Leitlinien und Empfehlungen nahezu unverändert übernommen werden konnten, andere wurden umformuliert oder durch entsprechende Kommentare an die spezifischen Erfordernisse der Sekundärdatenanalyse angepasst.

Die Adaption der GEP für eine "Gute Praxis Sekundärdatenanalyse" (GPS) erfolgte anschließend in einem mehrstufigen Prozess und wurde Januar 2005 nach mehrmaligen AG-Sitzungen im Konsens aller AGENS-Mitglieder verabschiedet.

Zielgruppe der GPS sind neben Primärnutzern wie den Sozialversicherungsträgern sozialmedizinisch und versorgungsepidemiologisch tätige Forscher und die Nutzer sekundäranalytischer Forschungsergebnisse angesprochen.

## Begriffsdefinitionen

Zum besseren Verständnis werden die zentralen Begriffe der GPS einleitend definiert.

**Sekundärdatenanalyse** bezeichnet die Analyse von Daten im Rahmen wissenschaftlicher Untersuchungen ohne direkten Bezug zum primären Erhebungsanlass.

**Primärnutzer:** Unter diesem Begriff werden im Rahmen der „Good Practice Sekundärdatenanalyse“ diejenigen Einrichtungen verstanden, die die Daten, die später für sekundäre Analysen genutzt werden, für ihre Aufgaben (primär) erheben, speichern und nutzen. Im Bereich der Gesetzlichen Krankenversicherung, aus dem bislang die meisten wissenschaftlichen Arbeiten der AGENS hervorgingen, sind dies vor allem die Krankenkassen oder Kassenärztlichen Vereinigungen, die versichertenbezogene (medizinische) Daten für administrative Aufgaben speichern.

**Forscher:** Unter diesem Begriff werden im Rahmen der GPS diejenigen Stellen verstanden, die Daten von Primärnutzern erhalten und zur Bearbeitung von Forschungsfragen unabhängig vom formalen primären Erhebungszweck aufbereiten und auswerten.

**Personenbezogene Daten:** Unter personenbezogenen Daten sind solche Informationen zu verstehen, die einer einzelnen Person als Beobachtungseinheit zugeordnet werden können. Aus datenschutzrechtlicher Sicht ist es hierbei unerheblich, ob diese Zuordnung zu einer Person *direkt* (durch personenidentifizierende Angaben), oder über ein *Pseudonym* (durch Depseudonymisierung) möglich ist.

**Pseudonymisierung:** Definition nach BDSG §3 (6a): *Pseudonymisieren ist das Ersetzen des Namens und anderer Identifikationsmerkmale durch ein Kennzeichen zu dem Zweck, die Bestimmung des Betroffenen auszuschließen oder wesentlich zu erschweren.*

Bei der Pseudonymisierung von personenbezogenen Daten werden die direkt personenidentifizierenden Daten (z. B. Name, Vorname, Telefonnummer, Sozialversicherungsnummer, Personalausweisnummer) aus den Daten entfernt und durch eindeutige Kennzeichen (z. B. eine Identifikationsnummer) ersetzt. Gängige Verfahren sind hierbei die Anwendung von Algorithmen oder Referenzlisten. Pseudonymisierte Daten gelten im datenschutzrechtlichen Sinne weiterhin als personenbezogen, da sie durch Rückrechnung wieder auf die Person bezogen werden können. Von der Pseudonymisierung ist insbesondere dann Gebrauch zu machen, wenn bei Nachziehungen oder Folgeerhebungen die personenbezogenen Daten immer wieder über dasselbe Pseudonym bereits pseudonymisierten Daten zugeordnet werden sollen.

**Anonymisierung:** Definition nach BDSG §3 (6a) BDSG: *Anonymisieren ist das Verändern personenbezogener Daten derart, dass die Einzelangaben über persönliche oder sachliche Verhältnisse nicht mehr oder nur mit einem unverhältnismäßig großen Aufwand an Zeit, Kosten und Arbeitskraft einer bestimmten oder bestimmaren natürlichen Person zugeordnet werden können.*

Unter den Begriff Anonymisierung fallen Maßnahmen, welche in einem Datenkörper die Identifikation einer Person wirksam verhindern. Dies kann wie bei der Pseudonymisierung das Ersetzen der personenidentifizierenden Daten durch eine Identifikationsnummer sein. Allerdings darf dieses Ersetzen nicht umkehrbar sein. Dies wird beispielsweise erreicht, indem der für die Anonymisierung verwendete Schlüssel oder die Referenzliste vernichtet wird. Häufig werden bestimmte Merkmale mittels Klassifikation (Bildung von Altersgruppen, Kürzen von Postleitzahlen u.a.) bei gleichzeitiger Löschung der unklassifizierten Angaben verändert. Auch innerhalb eines vollständig anonymisierten Datenkörpers können verschiedenen Angaben, z. B. im Längsschnitt, auf eine Beobachtungseinheit (ein Patient, ein Versicherter) durch Vergabe einer eindeutigen Kennziffer bezogen werden, aus der aber nicht auf die natürliche Person zurückgeschlossen werden kann. Ano-

nymisierte Daten fallen nicht mehr unter die datenschutzrechtlichen Bestimmungen über den Umgang mit personenbezogenen Daten.

## **LL 01 Ethik**

**Die Sekundärdatenanalyse muss wie alle wissenschaftliche Forschung im Einklang mit ethischen Prinzipien erfolgen und Menschenwürde sowie Menschenrechte respektieren.**

Ethische Prinzipien werden in allgemeinen Menschen- und Bürgerrechten formuliert. Ethische Prinzipien sind auch dann einzuhalten, wenn eine rechtliche Verpflichtung hierzu nicht besteht.

## **LL 02 Forschungsfrage**

**Die allgemeine wissenschaftliche Forderung nach Klarheit und Präzision bei der Formulierung der Forschungsfrage gilt auch für jede Sekundärdatenanalyse.**

Dazu bedarf es einer Begründung der Forschungsfrage auf der Basis verfügbaren Wissens, sowohl bei deskriptiven wie explorativen oder hypothesengenerierende Studien ein. In der Forschungsfrage muss die Zielpopulation einschließlich räumlicher und zeitlicher Abgrenzungen benannt sein, auf die sich die Ergebnisse der Sekundärdatenanalyse beziehen sollen.

### *Empfehlung 02.01 - Auswertungsplan*

Die Sekundärdatenanalyse bedarf einer an den verfügbaren Daten orientierten Planung der Auswertungsstrategie. Diese muss die Messgenauigkeit und Vollständigkeit (hinsichtlich vorhandener Merkmale sowie hinsichtlich möglicher Störfaktoren und Wechselwirkungen) der Daten berücksichtigen. Die im Rahmen einer Sekundärdatenanalyse zu überprüfenden Hypothesen müssen vor Beginn der Untersuchung formuliert werden, ebenso die bei diesen Prüfungen anzulegenden Entscheidungskriterien.

## **LL 03 Studienkonzept**

**Grundlage der Sekundärdatenanalyse ist ein detailliertes und verbindliches Studienkonzept, in dem die Studiencharakteristika und Abläufe schriftlich festgelegt werden.**

Die Sekundärdatenanalyse sollte sich auf eine Studienpopulation beziehen, deren Auswahl auf einer kritischen Beurteilung von Erhebungszweck, Qualität, Reliabilität, Validität und Verallgemeinerbarkeit der verwendeten Daten beruht.

### *Empfehlung 03.01 - Studiendesign*

Das Studiendesign soll dargelegt und begründet werden.

#### *Empfehlung 03.02 - Datenbasis*

Die Auswahl der Studienpopulation, der Auswertungsmerkmale und der Beobachtungszeiträume sollen dargelegt werden. Die Eignung des Datenkörpers als Basis für Auswertungen im Sinne des Studienkonzeptes ist zu begründen. Insbesondere bei Auswertungen zu seltenen Ereignissen oder bei kleineren Zielpopulationen sollten quantitative Abschätzungen zur Bestimmung der erforderlichen Populationsgrößen durchgeführt werden (vgl. Empfehlung 04.04).

#### *Empfehlung 03.03 - Ressourcen*

Für die Studie sind ausreichende zeitliche und personelle Ressourcen vorzusehen. Dies gilt gleichermaßen für die Datenbereitstellung, Aufarbeitung, Auswertung und Darstellung der Ergebnisse sowie für die erforderliche Kommunikation und Absprachen innerhalb und zwischen den beteiligten Stellen.

#### *Empfehlung 03.04 - Operationshandbuch*

Ergänzend zum Studienkonzept sollten in einem Operationshandbuch sämtliche organisatorischen Festlegungen zur Vorbereitung und Durchführung der Studie sowie deren schrittweise Abarbeitung dokumentiert werden.

## **LL 04 Datenaufbereitung**

### **Die Datenübernahme, -aufbereitung und -pflege der von den datenliefernden Stellen an den Forscher übermittelten Ausgangsdaten sowie Plausibilitätsprüfungen sind zu dokumentieren.**

Die Notwendigkeit der Dokumentation erstreckt sich auf sämtliche der Datenanalyse vorgeschalteten Prozesse der Datenaufbereitung und -prüfung. Die Dokumentation in Form eines Datenaufbereitungsprotokolls sollte die vollständige und nachvollziehbare Beschreibung des Datenhaltungssystems, der übermittelten Daten (Datum der Erhebung, Datum der Lieferung, Anzahl der Datensätze, verwendetes Übertragungsformat, verwendete Kodier- und Referenzlisten, Pseudonymisierungs- und Anonymisierungsschritte etc.) und die nachgehenden Transformationen bis zur Generierung eines Auswertedatensatzes umfassen.

#### *Empfehlung 04.01 – Datenerhebung und -übermittlung*

Die Grundlagen und Wege der Daten von ihrer primären Erfassung bis zur Bereitstellung für die Forschergruppe sollen dargestellt sein. Dies beinhaltet eine Darstellung des Verwendungszwecks der Datensammlung, der Erfassungsregeln und die Überprüfung der Konstanz von Erfassungsregeln (wie z.B. Diagnose- oder Prozedurenkodierung) und der gesetzlichen Rahmenbedingungen über die berücksichtigte Erhebungsperiode.

#### *Empfehlung 04.02 - Pretest*

Im Vorfeld der Datenbereitstellung sollte die Möglichkeit einer Bereitstellung von Probedaten mit einer reduzierten Zahl von Beobachtungen erwogen werden.

#### *Empfehlung 04.03 - Ausgangsdatensatz*

Der vom Datenherren übermittelte Ausgangsdatensatz soll in unveränderter Form über den gesamten Zeitraum der Studiendurchführung verfügbar sein.

#### *Empfehlung 04.04 - Datenbeschreibung*

Der Umfang und die Struktur der bereitgestellten und verwendeten Daten sollen dokumentiert werden. Hinweise auf fehlende Datensätze und deren Häufigkeit sowie die Feststellung redundanter Datensätze sind – insbesondere in Bezug auf die Zielpopulation (Ausschöpfungsquote) – gleichfalls zu dokumentieren (vgl. Empfehlung 03.02).

#### *Empfehlung 04.05 - Variablendefinition*

Die verwendeten Variablen müssen mit Erläuterungen zu ihrer inhaltlichen Bedeutung dokumentiert sein, ebenso alle als zulässig definierten Ausprägungen der Variablen, ggf. inklusive Angaben zu Gültigkeitszeiträumen.

#### *Empfehlung 04.06 - Datenqualität*

Die Reliabilität und Validität der Daten sollten auf der Basis verfügbarer Informationen überprüft werden. Die externe Validierung kritischer Merkmale im Rahmen von Primärerhebungen, z. B. an Subpopulationen, stellt dabei ein ausgewiesenes Qualitätsmerkmal dar.

#### *Empfehlung 04.07 - Plausibilitätskontrollen*

Vor und während der Sekundärdatenanalyse sind Plausibilitätskontrollen durchzuführen. Sie erfolgen prinzipiell auf der Grundlage des Ausgangsdatensatzes. Vor eventuell erforderlichen Korrekturen sollte in der Regel eine Rücksprache mit dem Datenherren erfolgen. Alle Ergänzungen und Änderungen von Variablenwerten sind vollständig und schriftlich zu dokumentieren.

#### *Empfehlung 04.08 – Operationalisierung - abgeleitete Merkmale*

Die Sekundärdatenanalyse umfasst in der Regel die Bildung von Merkmalen und Merkmalkategorien auf der Basis mehrerer primär bereitgestellter Variablen und deren Ausprägungen. Die Bildung abgeleiteter Merkmale ist nachvollziehbar zu dokumentieren und sollte sich ggf. an vorhandenen Standards orientieren, um eine Vergleichbarkeit der Ergebnisse zu erleichtern. Ebenso sollte die Bildung neuer Variablen vollständig dokumentiert werden.

#### *Empfehlung 04.09 – Auswertungsdatensatz*

Der nach Plausibilitätsprüfung und Datentransformation überarbeitete Datensatz ist als Auswertungsdatensatz zu kennzeichnen und unabhängig vom Ausgangsdatensatz zu speichern und zu sichern.

#### *Empfehlung 04.10 - Datenaufbewahrung*

Eine nach den jeweils gültigen Standards vorzusehende Aufbewahrungsfrist von Ausgangsdaten- und Auswertungsdatensatz sowie der Datenaufbereitungsprotokolle beim Forscher ist vertraglich zu regeln.

## LL 05 Datenanalyse

### **Die Auswertung soll unter Verwendung adäquater Methoden erfolgen.**

Bei der Auswertung ist den wissenschaftlichen Gepflogenheiten zu folgen, die die Überprüfbarkeit der Studienergebnisse in angemessener Form vorsehen. Die Regelung der Datenaufbewahrungsfrist nach Empfehlung 04.10 hat dies zu berücksichtigen. Die einzelnen Auswertungsschritte sollten nachvollziehbar sein und hinsichtlich möglicher Rückwirkungen auf die Selektion der Studienpopulation, die Operationalisierung der Forschungsfrage, der Reichweite der Ergebnisse, der Entscheidungskriterien bei der Hypothesenprüfung usw. kritisch analysiert werden. Strategien zur Verbesserung der Übertragbarkeit sind bereits vor Studienbeginn zu erwägen, wenn sie die Auswahl von Merkmalen oder die Selektion der Population beeinflussen.

#### *Empfehlung 05.01 – Analysemethoden*

Die Datenauswertung muss auf dem aktuellen Stand des epidemiologischen, statistischen oder methodischen Wissens erfolgen und hinsichtlich der allgemeinen Auswertungsstrategie (Stichprobenbildung, Modellierungen usw.) sowie der einzelnen Auswertungsschritte nachvollziehbar begründet werden. Begründete und notwendige Änderungen des ursprünglichen Auswertungsplanes sind zu dokumentieren.

#### *Empfehlung 05.02 – Personelle Verantwortung*

Eine für die Datenanalysen verantwortliche Person muss vor Beginn der Arbeit benannt und in die rechtlichen und organisatorischen Voraussetzungen der Datensammlung eingewiesen werden. Dies gilt auch für weitere an den Datenanalysen beteiligte Personen, wenn das Verständnis der Datenerhebung für die sachgerechte Durchführung der Analyse notwendig ist.

## LL 06 Qualitätssicherung

### **Vor dem Beginn und während der Sekundärdatenanalyse sind Qualitätssicherungsmaßnahmen durchzuführen.**

Aufgrund der Vielfalt der potenziell einer Sekundäranalyse zugrunde liegenden Datenquellen, ihrer unterschiedlichen ursprünglichen Verwendungszwecke, rechtlichen Grundlagen, Datenherren sowie Übertragungswege muss Transparenz bzgl. der Datenentstehung und -übermittlung geschaffen werden, um Auswertungsprozesse nachvollziehbar zu gestalten. Die Qualitätssicherung setzt an verschiedenen Punkten der Datenentstehung und -übermittlung an, um valide Auswertungsgrundlagen zu schaffen. Im einzelnen umfasst die Qualitätssicherung die Prüfung der Datenintegrität, Plausibilitätskontrollen und die Festlegung personeller Zuständigkeiten (vgl. Empfehlungen 04.06, 04.07, 05.02). Ebenso erstreckt sich die Qualitätssicherung auf die in Leitlinie 04 angesprochene Dokumentation der Datenaufbereitung.

## LL 07    **Datenschutz**

### **Bei der Planung und Durchführung von Sekundärdatenanalysen sind die geltenden Datenschutzvorschriften zu beachten.**

Alle Personen, die im Rahmen eines Forschungsprojektes Umgang mit personenbezogenen Daten haben, müssen über Inhalte, Reichweite und Möglichkeiten der einschlägigen gesetzlichen Bestimmungen informiert sein.

Bei der Forschung mit personenbezogenen Daten müssen das Recht des Einzelnen auf informationelle Selbstbestimmung und zugleich das Recht auf Freiheit von Wissenschaft und Forschung geachtet werden.

#### *Empfehlung 07.01 – Zusammenarbeit mit Datenschutzbeauftragten*

Bereits bei der Planung einer Sekundärdatenanalyse sind die zuständigen Datenschutzbeauftragten zu informieren, um mögliche Hemmnisse frühzeitig zu erfassen und einvernehmlich beseitigen zu können

#### *Empfehlung 07.02 – Zweck der Datenbereitstellung*

Der Zweck (im datenschutzrechtlichen Sinne) der Datenbereitstellung ist das Erreichen des Forschungsziels (siehe LL 02) und ist schriftlich zu fixieren.

#### *Empfehlung 07.03 – Pseudonymisieren und Anonymisieren*

Von den im Bundesdatenschutzgesetz aufgeführten Möglichkeiten der Pseudonymisierung und Anonymisierung ist Gebrauch zu machen (§ 3a BDSG Datenvermeidung und Datensparsamkeit). Hierbei ist auch die Einbindung eines Datentreuhänders zu prüfen.

#### *Empfehlung 07.04 – De-Pseudonymisierung und Re-Identifikation*

In den vertraglichen Rahmenbedingungen ist schriftlich festzulegen, ob und wenn ja, in welchen Fällen eine De-Pseudonymisierung möglich sein soll. Bei der Analyse ist eine unzulässige Re-Identifikation durch geeignete Mittel (technisch und vertraglich) zu verhindern.

#### *Empfehlung 07.05 – Weitergabe personenbezogener Daten an Dritte*

Die Weitergabe personenbezogener Daten erfolgt in der Regel nur durch den Primärnutzer.

#### *Empfehlung 07.06 – Personenbezogenes Datenlinkage mit externen Datenquellen*

Jedes nicht explizit vorgesehene *personenbezogene* Datenlinkage mit externen Datenquellen bedarf der Einhaltung datenschutzrechtlicher Vorschriften.

#### *Empfehlung 07.07 – Verantwortlicher für den Datenschutz*

Bei jeder Sekundärdatenanalyse sind die nationalen und internationalen Standards der Datensicherheit und des Datenschutzes einzuhalten. Innerhalb einer Forschungseinheit ist eine für die Datenverarbeitung verantwortliche Person zu benennen, die die Einhaltung dieser Standards zu überwachen hat. Diese Person ist entsprechend ihrer Aufgaben zu qualifizieren.

#### *Empfehlung 07.08 - Löschfristen*

Wenn die für die Sekundärdatenanalyse bereitgestellten Daten nach Erreichen des Studienzieles aus datenschutzrechtlichen Gründen zu löschen oder zu anonymisieren sind,



muss dies in Einklang mit den nach Empfehlung 04.10 geforderten Aufbewahrung von Ausgangs- und Auswertungsdatensatz erfolgen. Ebenso muss die nach Leitlinie 05 geforderte Möglichkeit zur Überprüfung der Ergebnisse aus einer Sekundärnutzung bei der Festlegung von evtl. Löschrufen mitbedacht werden.

## **LL 08 Vertragliche Rahmenbedingungen**

**Die Sekundärdatenanalyse setzt einen Vertrag voraus, in dem die rechtlichen Rahmenbedingungen definiert sind.**

Sekundärdatenanalysen beruhen in der Regel auf einer Kooperation zwischen Primärnutzer und Forscher. Aufgrund der gesetzlichen Auflagen zum Umgang mit Sozialdaten und anderen personenbezogenen Daten sind rechtswirksame Vereinbarungen der Kooperationspartner anzustreben, in denen die rechtlichen und finanziellen Rahmenbedingungen der intendierten Sekundärdatennutzung festgelegt werden.

### *Empfehlung 08.01 – Vertragsinhalte*

Die Rechte und Pflichten der Kooperationspartner sind vertraglich zu regeln. Hierzu gehören insbesondere Datenbereitstellung, Nutzungsrechte sowie die zu erbringenden Leistungen.

### *Empfehlung 08.02 – Verwendung der Ergebnisse*

Der Forscher sollte die Ergebnisse im Rahmen der vereinbarten Sekundärdatenanalyse für Forschung und Lehre nutzen können. In der Vertragsgestaltung sollten die Veröffentlichungsrechte der Auftragnehmer und –geber geregelt werden; es sollte darauf gedrungen werden, dass ein Zustimmungsvorbehaltspflicht des Auftraggebers zur Veröffentlichung von Ergebnissen nicht an die Inhalte der Veröffentlichung gebunden sein darf.

## **LL09 Interpretation**

**Die Interpretation der während der Sekundärdatenanalyse erzielten Ergebnisse und der daraus abgeleiteten Schlussfolgerungen fällt in die Zuständigkeit und Verantwortung des Forscher und sollte unabhängig vom Datenherrn erfolgen.**

Bezüglich der Einordnung der eigenen Vorgehensweisen und Erkenntnisse in den Rahmen vorhandener Evidenz sind allgemein verbreitete wissenschaftliche Standards einzuhalten. Einschränkungen bzgl. der Übertragbarkeit der Studienergebnisse auf andere Populationen oder Zeiträume als die betrachteten sind zu diskutieren.

**Ergebnisse von Sekundärdatenanalysen, die in gesundheitsrelevante Maßnahmen umgesetzt werden sollen, sollten mit den betroffenen Bevölkerungsgruppen angemessen diskutiert werden. In diese Kommunikation sollte die interessierte Öffentlichkeit einbezogen werden.**

Sekundärdatenanalysen können die Bewertung von Strukturen und Leistungen des Gesundheitssystems oder die Umsetzung und Evaluation gesundheitsrelevanter Maßnahmen zum Gegenstand haben. Ergibt sich nach dem fachlichen Urteil der Forscher aus den Ergebnissen einer Sekundärdatenanalyse die Notwendigkeit von Konsequenzen, sollen diese, beispielsweise in Form einer Empfehlung, explizit formuliert werden können. Dabei müssen sich die Forscher auch für eine verständliche Kommunikation mit Nicht-Fachleuten verantwortlich fühlen. Die Forscher können darüber hinaus fachlich begründete Empfehlungen an den Datenherrn zur öffentlichen Bereitstellung von Informationen formulieren und an der technischen Umsetzung mitwirken.

*Empfehlung 10.01 – Unabhängigkeit des Datennutzers*

In Unabhängigkeit ihrer Forschungstätigkeit sollten Forscher die Möglichkeit besitzen, sich bzgl. der praktischen Konsequenzen ihrer Analysen für die betroffenen Bevölkerungsgruppen direkt und ohne Abstimmung mit dem Datenherrn zu äußern. Formale Aspekte der direkten Kommunikation der Zweitnutzer mit der Öffentlichkeit sollten vertraglich geregelt werden.

*Empfehlung 10.02 – Transparenz der Methoden*

Die in einer Sekundärdatenanalyse verwandten Methoden sollen an geeigneter Stelle veröffentlicht und interessierten Forschern im Sinne der Methodentransparenz zugänglich gemacht werden. Die Weitergabe oder Publikation von Auswertungsroutinen, analytischen Verfahren und Berichtsformaten sollte unabhängig vom Datenherrn erfolgen können.