

# Statistische Analysen im Rahmen der Sekundären Datennutzung

Stand: 11. Januar 2018 / Gesine Schäfer-Reimers, Dr. Steffen Heß

Im Rahmen der Sekundären Datennutzung soll gemäß 8. Kapitel 1. Abschnitt § 4 Abs. 4 VerFO dem Antragsteller oder der Antragstellerin die Möglichkeit gegeben werden, „auf Basis der Datensatzbeschreibung die konkreten Auswertungsziele [zu] benennen und den konkreten Auswertungsplan ein[z]ureichen und in dem Antrag (Anlage I) entsprechende Angaben zu den hierfür benötigten Daten sowie ihrer Auswertung [zu] machen.“

Dazu wird für jedes Erfassungsjahr für den Datenbestand zur Sekundären Datennutzung eine Datensatzbeschreibung als Excel-Datei zur Verfügung gestellt, in der dem Antragsteller zu jedem Erfassungsmodul eines QS-Verfahrens Angaben zu den Datenfeldern übermittelt werden, von der ausgehend die Formulierung einer Rechenvorschrift ermöglicht wird. Sie wird durch die technische Spezifikation ergänzt<sup>1</sup>. Die Datensatzbeschreibung enthält diese Informationen (Feldbezeichner stehen in Klammern):

- Modul/Teildatensatz
  - Modulname und Teildatensatz-Kennzeichnung
  - Technischer Bezeichner (Modul) eines Datensatzes (z. B. „09/1“), der in der Datenbank als Bezeichner dient.
  - Bogenbezeichnung<sup>2</sup> (Patienten-, Eingriffs- und Prozedurebene), durch einen Doppelpunkt abgetrennt vom Modulbezeichner (z. B. „M“ für „Mutter“).
  - Ein Erfassungsmodul besteht mindestens aus einem Patientenbogen (i. d. R. „B“ für „Basis“).
  - Einige Module besitzen weitere Bögen auf bis zu drei Hierarchieebenen<sup>3</sup>.
- QS-Verfahren
  - Textuelle Bezeichnung des QS-Verfahrens (z. B. „Herzschrittmacher-Implantation“)
  - Die Bezeichnung bezieht sich auf das Erfassungsmodul.

---

<sup>1</sup> Eine Spezifikation ist eine detaillierte Beschreibung von Anforderungen für QS-Erfassungsinstrumente, die auf der [Website des IQTIG](#) veröffentlicht wird. In der QS-Basispezifikation für Leistungserbringer ist die technische Realisierung der Anforderungen u. a. in Form von Dokumentationsbögen, Datenbanken für die Auslösung (QS-Filter) und Erfassung der Dokumentation (QS-Dok), Ausfüllhinweisen und Anwenderinformationen hinterlegt.

<sup>2</sup> In der Tabelle „Bogen“ der QS-Dok Datenbank wird beschrieben, welche Erfassungsbögen es in welchen QS-Verfahren gibt.

<sup>3</sup> In der QS-Dok Datenbank finden sich die Hierarchieebenen in der Abfrage „Teildatensätze“ wieder.

- Zeile
  - Zeilen- bzw. Feldnummer des Erfassungsbogens
  - Er ist auf den Dokumentationsbögen<sup>4</sup> der Spezifikation ebenfalls angegeben
- Feldname
  - Technischer Name des Feldes aus der QS-Dokumentation
  - Einige Feldinhalte werden nicht direkt exportiert, sondern als berechnete Werte in einem Ersatzfeld zur Weiterleitung an die Bundesauswertungsstelle übergeben. Die Art der Berechnung für die Ersatzfelder können der Spezifikation<sup>5</sup> entnommen werden.
- Bezeichnung
  - Bezeichnung des Feldes
  - Die Bezeichnung ist auf dem Dokumentationsbogen der Spezifikation abgebildet
  - bei berechneten Feldern ist hier die Bezeichnung des zugrundeliegenden Feldes auf dem Dokumentationsbogen der Spezifikation angegeben
- ergänzende Bezeichnung
  - Ergänzende Bezeichnungen zur Interpretation des Feldes
- Muss/Kann
  - Einstufung des Feldes als Pflicht- (M) oder Kann-Feld (K)
  - Felder mit der Kennzeichnung „M“ müssen ausgefüllt werden.
  - Bei Feldern mit der Kennzeichnung „K“ handelt es sich entweder um Felder, die nicht ausgefüllt werden müssen, oder um Felder, die von anderen abhängig sind und in dem Kontext dann entweder verpflichtend sind und optional ausgefüllt werden können.
  - Der Kontext zum Ausfüllen eines mit „K“ gekennzeichneten Feldes ergibt sich aus den Plausibilitätsregeln der Spezifikation<sup>6</sup>.
- Max Feldlänge
  - Maximale Feldlänge eines Feldes
  - Bei Listenelementen (s. Erläuterung weiter unten) bezieht sich die Angabe auf die Länge eines Listenfeldes.
- Nachkommastellen
  - Zu erfassende Nachkommastellen bei numerischen Feldtypen (Nachkommastellen)
- Schlüssel
  - Referenzen zu Schlüsseln und Schlüssellisten, die in der Erfassung zugelassen sind.

---

<sup>4</sup> Muster der Dokumentationsbögen sind als Teil der Spezifikation in einem gleichnamigen Verzeichnis abgelegt.

<sup>5</sup> In der Abfrage „ExportfelderFürEinModul“ der QS-Dok Datenbank wird der Zusammenhang zwischen Erfassungsfeld, Ersatzfeld und Exportfeld beschrieben, und in der Tabelle „Ersatzfelder“ ist die einem Ersatzfeld zugrundeliegende Berechnungsvorschrift hinterlegt.

<sup>6</sup> In der Tabelle „Regeln“ der QS-Dok Datenbank werden die Abhängigkeiten beschrieben.

- In den Datensätzen sind die Schlüsselcodes hinterlegt. Die Beschreibungen der Codes können in der Spezifikation<sup>7</sup> oder in den Dokumentationsbögen abgelesen werden.
- Es gibt rund 14.000 Schlüssel, die in einem oder mehreren Erfassungsmodulen zum Tragen kommen. Sie sind in der Spezifikation<sup>8</sup> hinterlegt.
- Zudem gibt es noch 140 Schlüssellisten, die rund 16000 ICD- oder OPS-Schlüssel für bestimmte Diagnosen und Prozeduren umschreiben, die für ein Erfassungsmodul relevant sind.
- Bei bestimmten Feldern ist ein beliebiger Code aus dem DIMDI Katalog möglich. Diese Felder werden durch den Eintrag „OPS301“ in der Spezifikation gekennzeichnet.
- **Basistyp**
  - Einordnung der Felder in Feldtypen bzw. die Art des Schlüssels
  - „berechnetes Feld“ kennzeichnet ein Feld, das gemäß der Spezifikation<sup>9</sup> aus dem in „Bezeichnung“ angegebenen Feld berechnet oder verschlüsselt wurde
  - Die Ausprägungen der Schlüsseltypen erfolgt in der Spezifikation<sup>10</sup>.
- **Listenelemente**
  - Anzahl der erlaubten Feldwerte bei Mehrfachnennungen
  - Mehrfachnennungen sind z. B. bei den Entlassdiagnosen erlaubt; der Wert in der Spalte Listenelemente gibt an, wie viele Codes maximal in das Feld eingetragen werden können.
  - Jeder Eintrag wird in einem eigenen Datenfeld abgelegt.
- **Einheit**
  - Einheit des Erfassungswertes
  - Eine Einheit ist nur bei Feldern angegeben, wo sie zur Klärung des einzugebenden Wertes erforderlich sind.
  - Schlüsselwerte, Schlüssellisten, Datumsangaben und boolesche Werte haben keine Einheit
- **format Anweisung**
  - Format-Anweisungen der Feldwerte
  - Sie legen fest, welche Ausprägungen ein Feldwert annehmen kann.

### **Erläuterungen zur Verwendung Datenfeldbeschreibung**

Die Datenfeldbeschreibung dient zur Orientierung und legt den Umfang der Datenfelder pro QS-Verfahren dar. Sie stellt einen Extrakt aus der Spezifikationsdatenbank dar und enthält die Datenfelder, die für die Sekundäre Datennutzung herangezogen werden können, und eine Beschreibung ihrer Ausprägung. Eine flache Feldbeschreibung wie diese Datenfeldbeschreibung

---

<sup>7</sup> In der Tabelle „Feld“ der QS-Dok Datenbank werden die Schlüssel zu den Datenfeldern hinterlegt, die mit der Tabelle „SchlüsselWert“ verknüpft werden kann, um die Schlüsselcodes im Detail zu ermitteln.

<sup>8</sup> In der Tabelle „SchlüsselWert“ der QS-Dok Datenbank sind die Werte der einzelnen Schlüssel hinterlegt.

<sup>9</sup> In der Tabelle „BasisTyp“ der QS-Dok Datenbank sind die Formate der einzelnen Schlüssel hinterlegt.

<sup>10</sup> In der Tabelle „BasisTyp“ der QS-Dok Datenbank sind die Formate der einzelnen Schlüssel hinterlegt.

kann nicht die Komplexität der Abhängigkeiten, Regeln und relationalen Zusammenhänge der Datenfelder und ihrer Inhalte abbilden. Hierfür wird auf die Spezifikation verwiesen. Die Spezifikation steht erfassungsjahrbezogen auf der [Website des IQTIG](#) zum Download zur Verfügung. Für die Ermittlung der Zusammenhänge und Feldwerte sind folgende Dateien relevant:

- Die MS Access Datenbank zur QS-Dokumentation (z. B. 2016\_BASIS\_QSDOK\_V07.mdb), kurz QS-Dok Datenbank. Sie enthält einen Großteil weiterer Informationen zu den Angaben in dieser Datensatzbeschreibung. Der Name dieser Datei setzt sich aus dem Erfassungsjahr, der Kennzeichnung „BASIS\_QSDOK“ und der Versionsnummer zusammen. Die höchste Versionsnummer bezieht sich auf die Version des Datensatzes, der für die Sekundäre Datennutzung zur Verfügung steht. Die für die sekundäre Datennutzung relevanten Tabellen werden im Verlauf dieses Dokumentes kontextbezogen angegeben.
- Die MS Access Datenbank zur Definition der Auslösung einer Dokumentation (z. B. 2016\_QSF\_V05.mdb), kurz QS-Filter Datenbank. Der Name dieser Datei setzt sich aus dem Erfassungsjahr, der Kennzeichnung „QSF“ und der Versionsnummer zusammen. Die höchste Versionsnummer bezieht sich auf die Version des Datensatzes, der für die Sekundäre Datennutzung zur Verfügung steht. Die für die sekundäre Datennutzung relevanten Tabellen werden im Verlauf dieses Dokumentes kontextbezogen angegeben.
- Die Dokumentationsbögen, die den Erfassungsdatsatz veranschaulichen, als PDF. Sie liegen in einem entsprechend gekennzeichneten Verzeichnis. Ihr Name entspricht dem Modulnamen aus der Datenfeldbeschreibung.
- Die Ausfüllhinweise, die Erläuterungen zur Dokumentation enthalten, als HTML Dateien. Sie liegen in einem entsprechend gekennzeichneten Verzeichnis. Ihr Name entspricht dem Modulnamen aus der Datenfeldbeschreibung.
- Die Anwenderinformationen, die die Codes für eine Auslösung zur Dokumentation enthalten. Sie liegen als HTML-Dokument vor und tragen den technischen Modulnamen. Sie können im Rahmen der Sekundären Datennutzung dazu dienen, die Grundgesamtheit der Patienten eines Datentopfes detailliert zu ermitteln.

Es wird empfohlen, die Dokumentationsbögen aus der Spezifikation heranzuziehen, um die Feld- und Bogenabhängigkeiten zu visualisieren. Für die Formulierung von Rechenvorschriften ist es notwendig, zusätzlich zu der Datenfeldbeschreibung noch weitere Dokumente aus der Spezifikation des Erfassungsjahres heranzuziehen. Hier ist auf die MS Access Datenbanken zu verweisen, in der die Schlüssel der einzelnen Datenfelder und Schlüssellisten für die QS-Dokumentation aufgeführt sind. Zudem geben Plausibilitätsregeln<sup>11</sup> Auskunft über die Regelungen<sup>12</sup> zu MUSS und KANN Feldern bei kombinierten Bogenfeldern (Feldgruppen), bei denen das Ausfüllen eines Datenfeldes von den Werten eines anderen Feldes gesteuert wird.

---

<sup>11</sup> In der Abfrage „Plausibilitätsregeln“ oder „PlausibilitätsregelnFürEinModul“ der QS-Dok Datenbank sind die Regeln dargelegt.

<sup>12</sup> In der Tabelle „Regeln“ der QS-Dok Datenbank sind die fachlichen Zusammenhänge dargelegt.

Der Antragsteller oder die Antragstellerin müssen zur Formulierung von konkreten Auswertungszielen und dem konkreten Auswertungsplan gute Vorkenntnisse im Umgang mit Datenbanken und der Formulierung von Rechenregeln haben. Mit diesen Vorkenntnissen ist es ihm/ihr möglich, aus der MS Access-Datenbank für die QS-Dokumentation zu jedem QS-Verfahren die relevanten Feld-Schlüssel oder Schlüssellisten der in der Datensatzbeschreibung genannten Datenfelder zu ermitteln. Die Spezifikation enthält zudem zahlreiche weiterführende Angaben zur externen Qualitätssicherung nach §§ 136 ff SGB V wie Auslösekriterien für die Dokumentationspflicht, Ausfüllhinweise zu den Datenfeldern, Anwenderhinweise mit Beschreibungen zu den Kodelisten u. v. m., die bei der Entwicklung von Auswertungszielen hilfreich sein können.

### Struktur der Datenbank mit den QS-Daten für die Sekundäre Datennutzung

Für die sekundäre Datennutzung stehen die Datenpools der Bundesauswertungen der Erfassungsjahre zur Verfügung. Diese Datenpools stehen erst zur Verfügung, wenn die Bundesauswertungen vom Gemeinsamen Bundesausschuss veröffentlicht worden sind. Die Datenpools sind unveränderlich.

Für die Formulierung von Rechenvorschriften ist zu beachten, dass die Daten zu Datensätzen mit einer Teildatensatzstruktur<sup>13</sup> in der Datenbank für die Sekundäre Datennutzung als zusammengefasster Datensatz pro kleinster Teildatensatzeinheit zusammengefasst wurden. Am Beispiel des QS-Verfahrens Geburtshilfe soll dies exemplarisch dargelegt werden. Hier gibt es einen Basisdatensatz (Daten zur Mutter) und einen Teildatensatz (Angaben zu den Kindern). Pro Kind wird ein Teildatensatz angelegt. Bei einer Mutter mit Zwillingen erfolgt die Ablage der Daten wie folgt:

Datenfelder Mutter	Datenfelder Kind 1
Datenfelder Mutter	Datenfelder Kind 2

Bei Datensätzen mit drei Hierarchieebenen erweitert sich der Aufbau entsprechend:

Datenfelder Patient	Datenfelder Prozedur 1	Datenfelder Eingriff 1
Datenfelder Patient	Datenfelder Prozedur 1	Datenfelder Eingriff 2
Datenfelder Patient	Datenfelder Prozedur 2	Datenfelder Eingriff I
Datenfelder Patient	Datenfelder Prozedur 2	Datenfelder Eingriff II

<sup>13</sup> In der QS-Dok Datenbank finden sich die Abhängigkeiten in der Abfrage „Teildatensätze“ wieder.

## **Hinweise zu Datenfeldern**

Zusätzlich zu den bereits im Feld Bezeichnung angegebenen Erklärungen gilt für folgende Felder:

„versichertenidgkv“ enthält ein Versichertenpseudonym für gesetzlich Versicherte.

„IKNRKH“ enthält u.a. eine Regionskennung und kann daher zur Bestimmung des Bundeslandes genutzt werden. Im stationären Bereich werden 16 Bundesländer berücksichtigt, im vertragsärztlichen Bereich 17 Kassenärztliche Vereinigung (Nordrhein und Westfalen-Lippe sind hier getrennt).

## **Datenschutzrechtliche Maßnahmen**

Der Antragsteller bzw. die Antragstellerin erhalten aus datenschutzrechtlichen Gründen lediglich die Ergebnisse der von ihm/ihr eingereichten Rechenvorschriften ausgehändigt. Er hat keinen Zugriff auf die Datenbank mit den Daten.

Der Zweck der Sekundären Datennutzung besteht in der Möglichkeit, Analysen zu Daten auch über mehrere Jahre hinweg vorzunehmen. Auf Grund dieser Zweckbestimmung sind keine Löschfristen zu beachten. Da Patientenidentifikatoren ausschließlich als Anonyme hinterlegt sind, wird den Löschfristen der Richtlinien zu QS-Verfahren mit patientenidentifizierenden Daten Genüge getan.

In der Datenbank zur Sekundären Datennutzung werden bei den Indirekten Verfahren der QSKH die Leistungserbringer i. d. R. über Leistungserbringerpseudonyme identifiziert. Bei den Indirekten Verfahren liegen IK-Nummern vor. In den QS-Verfahren der Qesü sind die IK-Nummern des stationären Sektors und die Betriebsstättennummern des ambulanten Sektors in Form von Pseudonymen abgelegt. Diese sind spezifisch für bestimmte QS-Verfahren. Patientenidentifikatoren sind ausschließlich als Anonyme hinterlegt.

Antragstellern werden in den aggregierten Ergebnissen nur pseudonymisierte Leistungserbringerinformationen übergeben. Ein einzelner Leistungserbringer kann die Ergebnisse seiner Einrichtung als solche gekennzeichnet erhalten, sofern hierfür die IK-Nummer im Datensatz vorliegt. Eine Zuordnung zur Art der Einrichtung kann aufgrund fehlender Informationen im Datensatz nicht vorgenommen werden.

## **Datenumfang und Qualitätsergebnisse**

Als Institut nach §137a SGB V verarbeitet das IQTIG seit dem Erfassungsjahr 2015 die für Bundesauswertungen relevanten Daten und stellt diese für Auswertungen zur Verfügung. Die Fallzahlen zu den einzelnen QS-Verfahren ändern sich jährlich und sind den Bundesauswertungen auf der Website des IQTIG zu entnehmen. Daten früherer Jahre (2009 bis 2014) stehen aus dem Datenpool von AQUA auch zur Verfügung.

Zur Beurteilung von Indikatorergebnissen im Zeitverlauf kann eine Liste aller QI-Ergebnisse seit dem Erfassungsjahr 2009 zur Verfügung gestellt werden.