

Modul III: »SQL für epidemiologische Sekundärdatenanalysen: Schlüsselstabellen – der Schlüssel zum Analyseerfolg«

Für die Datenhaltung von GKV-Routinedaten zum Zwecke der Versorgungsforschung kommen je nach Anforderung und Eignung unterschiedliche Konzepte zum Einsatz. Eines davon, welches sich vor allem bei sehr umfangreichen Datenerhebungen anbietet, ist die Speicherung in einer relationalen Datenbank und die Aufbereitung und Auswertung der Daten mit Hilfe der Standard-Abfragesprache SQL (Structured Query Language, seit 1986 ANSI-Standard).

Hier setzt der hier angebotene Einführungs- und Vertiefungskurs an: Er richtet sich an Epidemiologen und andere Professionen, die einen Einstieg in SQL-Abfragetechniken suchen bzw. an solche SQL-Anwender, die ihr Wissen vertiefen wollen.

Ausgehend von einem typischen Auswertungsszenario zur Indikation „Diabetes mellitus“ werden zentrale Berichtsformate wie die Bestimmung von Bezugspopulation, die Berechnung von Prävalenz und Inzidenz nach Alters- und Geschlechtsgruppen oder von ausgewählten medizinischen Inanspruchnahmeparametern (z. B. Art der medikamentösen Therapie, durchschnittliche Anzahl von Krankenhausaufenthalten) von den Kursteilnehmern unter Anleitung des Seminarleiters selbstständig erstellt und mithilfe von SQL-Skripten umgesetzt. Hierbei geht es vor allem darum, das SQL-typische Sprachkonzept der relationalen Tabellenverknüpfung, der so genannten *joins* (inner join, outer join, cross join) zielführend einzusetzen. In diesem Zusammenhang kommt auch dem Einsatz von *Schlüsselstabellen* eine entscheidende Bedeutung zu, da über solche Tabellen die Rahmenbedingungen einer Auswertung festgelegt werden können. Diese *flexible zentrale Skalierung* gestattet es zudem, unterschiedliche Szenarien durchzuspielen und im Sinne einer Sensitivitätsanalyse einzusetzen. Darüber hinaus ist es damit möglich, Auswertungstemplates durch Ändern oder Austauschen der Schlüsselstabellen ressourcenschonend für weitere Auswertungen, z. B. andere Indikationen, zu nutzen.

Die Auswertung in der Testdatenbank erfolgt auf PCs und als Beamerpräsentation anhand eines Übungsskripts. Die Kursteilnehmer erhalten eine Anleitung, die Auswertungsumgebung und Testdatenbank auf einem eigenen Rechner zu installieren. Die Kursteilnehmer sind nach erfolgreicher Teilnahme in der Lage, SQL-Skripte zu erstellen, zu modifizieren und anzuwenden.

Spezielle Voraussetzungen: PC-Grundkenntnisse, insbesondere mit dem Betriebssystem Windows. Grundkenntnisse in SQL nicht erforderlich aber hilfreich [1].

Kursdauer: 2 x 4 Stunden

Verantwortlich für das Modul: Peter Ihle, Universität zu Köln, PMV forschungsgruppe, Herderstraße 52, 50931 Köln, Telefon +49/0 - 221 478 6548, Peter.Ihle@uk-koeln.de

Literatur:

[1] SQL Grundlagen und Datenbankdesign. Handbuch des RRZN – Regionales Rechenzentrum für Niedersachsen 3. bis 5. Auflage (inhaltsidentisch)