

Enno Swart  
Peter Ihle  
Holger Gothe  
David Matusiewicz

Herausgeber

# Routinedaten im Gesundheitswesen

**Handbuch Sekundärdatenanalyse:  
Grundlagen, Methoden und Perspektiven**

2., vollständig überarbeitete und erweiterte Auflage

Verlag Hans Huber

Lektorat: Dr. Klaus Reinhardt  
Herstellung: Daniel Berger  
Bearbeitung: Ulrike Boos, Freiburg  
Umschlaggestaltung: Claude Borer, Basel  
Druckvorstufe: punktgenau gmbh, Bühl  
Druck und buchbinderische Verarbeitung: Hubert & Co., Göttingen  
Printed in Germany

*Bibliografische Information der Deutschen Nationalbibliothek*

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.



*Anregungen und Zuschriften bitte an:*

Verlag Hans Huber  
Lektorat Medizin/Gesundheit  
Länggass-Strasse 76  
CH-3000 Bern 9  
Tel: 0041 (0)31 300 4500  
[verlag@hanshuber.com](mailto:verlag@hanshuber.com)  
[www.verlag-hanshuber.com](http://www.verlag-hanshuber.com)

2. Auflage 2014

© 2005/2014 by Verlag Hans Huber, Hogrefe AG, Bern

(E-Book-ISBN [PDF] 978-3-456-95435-6)

ISBN 978-3-456-85435-9

# Inhalt

Vorwort .....	12
Vorwort der ersten Auflage	
Sekundärdatenanalyse: Aufgaben und Ziele ( <i>Enno Swart, Peter Ihle</i> ) .....	16
AGENS – eine schrecklich routinierte Familie .....	19
( <i>David Matusiewicz, Holger Gothe, Peter Ihle, Enno Swart</i> )	
<b>I. Versorgungsbereiche im Spiegel der Gesetzlichen Krankenversicherung .....</b>	<b>27</b>
1. Stammdaten und Versichertenhistorien ( <i>Thomas G. Grobe, Peter Ihle</i> ) .....	28
1.1 Einleitung .....	28
1.2 Versichertengruppen .....	29
1.3 Versichertenbezogene Merkmale .....	30
1.4 Nutzung für wissenschaftliche Fragestellungen .....	34
1.5 Auswahl von Bezugspopulationen .....	35
2. Analysen zur Sterblichkeit ( <i>Rembrandt Scholz, Sebastian Sauer, Rolf Müller</i> ) .....	38
2.1 Hintergrund .....	38
2.2 Untersuchungsgegenstand: Sterberaten .....	39
2.3 Datengrundlage, methodisches Vorgehen und Selektion .....	39
2.4 Sterberaten der Deutschen Rentenversicherung und der Gesetzlichen Krankenversicherung versus Bevölkerungsdaten der Human Mortality Database .....	40
2.5 Ausblick .....	42
3. Ambulante ärztliche Versorgung ( <i>Thomas G. Grobe, Hendrik Dräther</i> ) .....	43
3.1 Hintergrund – Abgrenzung – gesetzliche Grundlagen .....	44
3.2 Datenerhebungswege – Datenhalter .....	47
3.3 Datenstruktur – Datenumfang – Kennwerte .....	48
3.4 Historische Verfügbarkeit – wesentliche Veränderungen .....	56
3.5 Beispielauswertungen .....	57
3.6 Resümee: Möglichkeiten und Grenzen .....	60
4. Dokumentation der Disease-Management-Programme ( <i>Dirk Horenkamp-Sonntag, Roland Linder, Dagmar Köppel, Daniel Wildner</i> ) .....	63
4.1 Hintergrund und Historie .....	64
4.2 Datenentstehung .....	65
4.3 Datenstruktur .....	67
4.4 Historische Verfügbarkeit .....	70
4.5 Evaluation .....	71
4.6 Daten-Qualität der Dokumentationsdaten .....	71
4.7 Ausblick: Möglichkeiten und Grenzen .....	72
5. Arzneimittelverordnungen ( <i>Helmut Schröder</i> ) .....	74
5.1 Einführung .....	74
5.2 Was sind Arzneimittel? .....	75

## 6 Inhalt

---

5.3	Wie kommen die Arzneimitteldaten zu den Krankenkassen? .....	76
5.4	Analysen mit Arzneimitteldaten .....	79
5.5	Kassenartenübergreifende Arzneimitteldaten .....	82
5.6	Kassenspezifische Arzneimitteldaten .....	83
5.7	Fazit .....	85
6.	Heilmittel ( <i>Andrea Waltersbacher</i> ) .....	88
6.1	Hintergrund .....	88
6.2	Wie entstehen die Heilmitteldatenpools? .....	90
6.3	Heilmitteldatenpools, Kennwerte und Veröffentlichungen von Heilmitteldaten .....	92
6.4	Ausblick: Möglichkeiten und Grenzen .....	102
7.	Zahnärztliche Versorgung ( <i>Rugzan Hussein, Andrea Hartmann, Thomas Schäfer, Eva-Maria Bitzer</i> ) .....	104
7.1	Hintergrund – Abgrenzung – gesetzliche Grundlagen .....	105
7.2	Datenerhebungswege – Datenhalter .....	106
7.3	Ablauf der elektronischen Datenübermittlung .....	106
7.4	Formulare, die vor der elektronischen Datenübermittlung eingesetzt wurden .....	108
7.5	Datenstruktur – Datenumfang .....	112
7.6	Kennwerte .....	112
7.7	Publikationen und bevölkerungsbezogene Studien in Deutschland – Referenzergebnisse .	116
7.8	Resümee: Möglichkeiten und Grenzen .....	118
8.	Krankenhausbehandlung ( <i>Thomas G. Grobe, Ulrike Nimptsch, Jörg Friedrich</i> ) .....	121
8.1	Hintergrund – Abgrenzung – gesetzliche Grundlagen .....	122
8.2	Datenerhebungswege – Datenhalter .....	126
8.3	Datenstruktur – Datenumfang – Kennwerte .....	128
8.4	Auswertung und Kennwerte .....	135
8.5	Beispielauswertungen .....	141
8.6	Historische Verfügbarkeit – wesentliche Veränderungen .....	143
8.7	Resümee .....	144
9.	Arbeitsunfähigkeit ( <i>Markus Meyer</i> ) .....	146
9.1	Hintergrund und gesetzliche Rahmenbedingungen .....	146
9.2	Entstehung der Arbeitsunfähigkeitsdaten und Datenhaltung .....	147
9.3	Datenstruktur: Welche Daten stehen zur Verfügung? .....	148
9.4	Arbeitsunfähigkeitsdaten in der Gesundheitsberichterstattung und Epidemiologie .....	149
9.5	Kennwerte .....	150
9.6	Möglichkeiten und Grenzen .....	158
9.7	Resümee .....	159
<b>II. Versorgungsbereiche im Spiegel anderer Sozialversicherungsträger .....</b>		<b>161</b>
10.	Pflegeleistungen nach Sozialgesetzbuch XI ( <i>Rolf Müller, Heinz Rothgang, Rainer Unger</i> ) .....	162
10.1	Hintergrund .....	163
10.2	Datenentstehung in der Pflegekasse .....	166
10.3	Daten bei den Pflegekassen .....	167
10.4	Historische Verfügbarkeit – wesentliche Veränderungen .....	169
10.5	Amtliche Statistiken zur Pflegeversicherung .....	170

10.6	Resümee	171
10.7	Ausblick	173
11.	Routinedaten zur Rehabilitation durch die Träger der Sozialversicherung ( <i>Anja Bestmann, Ute Polak, Ingmar Hansen</i> )	176
11.1	Hintergrund	176
11.2	Datenentstehung in der Rehabilitation	179
11.3	Datenstruktur, -aufbau und Datenumfang, Besonderheiten	181
11.4	Resümee: Möglichkeiten und Grenzen	186
12.	Rekonstruktion von Erwerbsverläufen ( <i>Anja Burghardt, Anita Tisch, Silke Tophoven</i> )	192
12.1	Einleitung	192
12.2	Sekundärdaten der Arbeitsverwaltung	193
12.3	Die Verwendung von erwerbsbiografischen Informationen aus den Sekundärdaten der Bundesagentur für Arbeit und des Instituts für Arbeitsmarkt- und Berufsforschung	196
12.4	Die Verknüpfung von administrativen Erwerbsbiografiedaten mit Primärdaten der lidA-(leben in der Arbeit)Kohortenstudie zu Arbeit und Gesundheit	197
12.5	Die Sequenzmusteranalyse als Beispiel der Abbildung erwerbsbiografischer Verläufe	198
12.6	Zusammenfassung	200
13.	Daten zu Erwerbsminderungsrenten ( <i>Ralf K. Himmelreicher</i> )	203
13.1	Hintergrund	203
13.2	Datenentstehung	204
13.3	Datenstruktur	204
13.4	Resümee	210
<b>III.</b>	<b>Ergänzende Routinedaten mit Gesundheitsbezug</b>	<b>213</b>
14.	Gesundheitsbefragungen und Aggregatdaten ( <i>Jelena Jaunzeme, Yvonne Marx, Enno Swart, Siegfried Geyer</i> )	214
14.1	Einleitung	215
14.2	Gesundheitsmonitoring des Robert Koch-Instituts	215
14.3	Weitere gesundheitsrelevante Surveys	217
14.4	Bevölkerungsumfragen mit gesundheitsrelevanten Themen	218
14.5	Nutzung von Aggregatdaten	219
14.6	Resümee	221
15.	Möglichkeiten und Grenzen von Befragungsdaten und Daten gesetzlicher Krankenversicherungen ( <i>Siegfried Geyer, Jelena Jaunzeme</i> )	223
15.1	Einführung	224
15.2	Befragungsdaten	224
15.3	Daten gesetzlicher Krankenversicherungen	227
15.4	Zusammenfassende Beurteilung	230
16.	Registerdaten ( <i>Birga Maier, Holger Gothe, Joachim Kieschke</i> )	234
16.1	Einführung	234
16.2	Was ist ein Register?	234
16.3	Welche Register gibt es?	235
16.4	Epidemiologisches Krebsregister Niedersachsen (EKN) als Beispiel für ein Krankheitsregister	237

16.5	Berliner Herzinfarktregister als Beispiel für ein Qualitätsregister	239
16.6	Welche Vorzüge haben Register?	241
16.7	Welche Grenzen und Probleme weisen Register auf?	241
16.8	Zusammenfassung und Ausblick	243
17.	Daten der Privaten Krankenversicherung (PKV) ( <i>Holger Gothe, Anne-Dorothee Köster</i> )	245
17.1	Das System der Privaten Krankenversicherung in Deutschland	245
17.2	Hintergründe für die Datenlage in der PKV	246
17.3	Versorgungsdaten der PKV	247
17.4	Stellenwert der Nutzung von Versorgungsdaten der PKV	251
18.	Leistungserbringerbezogene Merkmale ( <i>Ulrike Nimptsch</i> )	254
18.1	Hintergrund	254
18.2	Merkmale zur Abgrenzung von Leistungserbringern	255
18.3	Zuordnung von leistungserbringerbezogenen Merkmalen	258
19.	Routinedaten im Ausland ( <i>Holger Gothe</i> )	260
19.1	Hintergrund	260
19.2	Arztbasierte Datensätze	261
19.3	Administrative Datensätze	263
19.4	Fazit	265
<b>IV.</b>	<b>Datenzugang und Datenvalidierung</b>	<b>269</b>
20.	Zugang zu Routinedaten ( <i>Ulrike Nimptsch, Anja Bestmann, Michael Erhart, Stefan Dudey, Yvonne Marx, Joachim Saam, Michael Schopen, Helmut Schröder, Enno Swart</i> )	270
20.1	Einleitung	270
20.2	Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder	271
20.3	Reha-Statistik der deutschen Rentenversicherung	275
20.4	Informationssystem Versorgungsdaten (Datentransparenz)	278
20.5	KV-Daten-Stichprobe des Zentralinstituts für die kassenärztliche Versorgung	279
20.6	Versorgungsdaten der Versicherten bei gesetzlichen Krankenversicherungen	283
21.	Datenschutzrechtliche Aspekte bei der Nutzung von Routinedaten ( <i>Stefanie March, Angela Rauch, Stefan Bender, Peter Ihle</i> )	291
21.1	Rechtlicher Hintergrund	291
21.2	Organisatorisches Vorgehen	294
21.3	Zwei Beispielstudien unter datenschutzrechtlichen Gesichtspunkten	300
21.4	Fazit	301
22.	Datenhaltung und -analyse aus technischer Sicht ( <i>Peter Ihle, David Matusiewicz, Thomas G. Grobe, Gisbert W. Selke</i> )	304
22.1	Einleitung	304
22.2	Datenhaltung	305
22.3	Datenverknüpfung	306
22.4	Datenzugang und Datenübermittlung	306
22.5	Datenvolumen	308
22.6	Software	309
22.7	Fazit	313

23. Prüfung der Datenqualität und Validität von GKV-Routinedaten ( <i>Dirk Horenkamp-Sonntag, Roland Linder, Fabian Wenzel, Bettina Gerste, Peter Ihle</i> )	314
23.1 Warum müssen Daten geprüft werden?	315
23.2 Welche Möglichkeiten der Prüfung gibt es?	317
23.3 Anwendungsbeispiele zur Plausibilität	319
23.4 Anwendungsbeispiele zur internen Validierung	322
23.5 Anwendungsbeispiele zur externen Validierung	324
23.6 Prüfstrategien	326
23.7 Fazit zur Prüfung der Daten-Qualität	329
24. Klassifikationssysteme ( <i>Anja Bestmann, Carsten Telschow, Michael Stegmann</i> )	331
24.1 Hintergrund	331
24.2 Internationale statistische Klassifikation der Krankheiten und verwandten Gesundheitsprobleme (ICD-10)	332
24.3 Operationen- und Prozedurenschlüssel (OPS)	334
24.4 Einheitlicher Bewertungsmaßstab (EBM)	335
24.5 Anatomisch-therapeutisch-chemisches System (ATC)	338
24.6 Pharmazentralnummer (PZN)	340
24.7 Die Neugestaltung des Tätigkeitsschlüssels	342
24.8 Institutionskennzeichen (IK)	343
25. Datenlinkage ( <i>Stefanie March, Christoph Stallmann, Enno Swart</i> )	347
25.1 Allgemeine Anforderungen an das Datenlinkage	348
25.2 Ausgewählte Beispiele des Datenlinkages	350
<b>V. Routinedaten als Informationsquelle für Morbiditätsschätzungen</b>	357
26. Krankheitsereignis: Operationalisierung und Falldefinition ( <i>Ingrid Schubert, Ingrid Köster</i> )	358
26.1 Hintergrund	359
26.2 Operationalisierung von Krankheiten in der Versorgungsforschung und Pharmakovigilanz: Stellenwert der Diagnosedaten	360
26.3 Szenarien der Diagnosegenerierung im ambulanten Sektor	361
26.4 Interne Diagnosevalidierung und Falldefinition	363
26.5 Externe Diagnosevalidierung/Plausibilisierung	365
26.6 Fazit	366
27. Bestimmung von Behandlungs- und Erkrankungsperioden in Routinedaten ( <i>Sascha Abbas, Peter Ihle</i> )	369
27.1 Hintergrund und Datengrundlage	369
27.2 Behandlungs- versus Erkrankungszeitraum	370
27.3 Prävalenz- und Inzidenzbestimmung – praktische Umsetzung in den Daten	371
27.4 Persistenz der Erkrankung und Diagnosemuster in Routinedaten	374
27.5 Fazit	374
28. Die Population unter Risiko bei Prävalenz- und Inzidenzschätzungen – Nennerkonzepte ( <i>Christoph Wagner</i> )	376
28.1 Hintergrund	377
28.2 Besonderheiten von Versichertenstammdaten für eine Nennerdefinition	378
28.3 Zeitliche Verfügbarkeit von Leistungsdaten	381

28.4	Verfahren zur Selektion einer Population unter Risiko	381
28.5	Nennerkonzepte für Daten der Datentransparenzverordnung (§ 303a–e SGB V)	385
28.6	Fazit	386
29.	Fallbeispiele für Prävalenz- und Inzidenzschätzungen ( <i>Christoph Wagner, Dirk Horenkamp-Sonntag, Thomas Grobe, Roland Linder</i> )	389
29.1	Einleitung	389
29.2	Neugeborenen-Inzidenz für Mukoviszidose – ein cleverer Spezialfall	391
29.3	Extrapolation der Lebenszeitprävalenz der Aufmerksamkeitsdefizit-/Hyperaktivitäts- störung	393
29.4	Wie umgehen mit der Erkrankungsdunkelziffer bei Osteoporose?	396
29.5	Fazit	397
<b>VI.</b>	<b>Spezielle Methoden, Instrumente und Anwendungen</b>	<b>401</b>
30.	Matching-Verfahren ( <i>Sarah Mostardt, Gerald Lux, Helmut Dahl, David Matusiewicz, Janine Biermann</i> )	402
30.1	Hintergrund	402
30.2	Verfügbarkeit von Matching-Variablen in Sekundärdaten	403
30.3	Matching-Verfahren	404
30.4	Herausforderungen und Probleme beim Matching	406
30.5	Fazit und Ausblick	410
31.	Risikoadjustierung und Komorbiditäten ( <i>Gerald Lux, Janine Biermann, Helmut Dahl, David Matusiewicz, Sarah Mostardt, Ulrike Nimptsch, Jürgen Wasem, Anke Walendzik</i> )	411
31.1	Hintergrund	411
31.2	Berücksichtigungsfähige Variablen für eine Risikoadjustierung	412
31.3	Komorbiditätsbasierte Risikoindizes	415
31.4	Methodische Herausforderungen der Risikoadjustierung	416
31.5	Anwendungsfelder der Risikoadjustierung	417
31.6	Fazit	421
32.	Bias und Confounding ( <i>Raffaella Matteucci Gothe, Barbara Buchberger</i> )	424
32.1	Einführung	424
32.2	Definition und Beispiele von Bias und Confounding	426
32.3	Maßnahmen zur Vermeidung/Adjustierung von Bias und Confounding	428
32.4	Zusammenfassung	432
33.	Routinedaten für kleinräumige Analysen ( <i>Janett Powietzka, Enno Swart</i> )	435
33.1	Vorbemerkungen	435
33.2	Regionalisierte Daten der Gesetzlichen Krankenversicherungen	436
33.3	Versorgungsatlas	437
33.4	Raumordnungsdaten des Bundesinstituts für Bau-, Stadt- und Raumforschung	438
33.5	Regionalisierte Daten der amtlichen Statistik	439
33.6	Mikrozensus	440
33.7	Survey-Daten des Robert Koch-Instituts	441
33.8	Das Sozioökonomische Panel	441
33.9	Fazit	442

34. Erfahrungen aus Sozialmedizin und Gesundheitsökonomie ( <i>Janine Biermann, Sarah Mostardt, Gerald Lux, Jürgen Wasem, Helmut Dahl, David Matusiewicz</i> ) . . . . .	446
34.1 Hintergrund . . . . .	446
34.2 Ausgewählte Problemfelder und Erfahrungen im Rahmen von Routinedatenanalysen. . . . .	447
34.3 Fazit . . . . .	458
35. Nutzung von Sekundärdaten in der gesetzlichen Qualitätssicherung ( <i>Günther Heller, Björn Borge, Joachim Szecsenyi</i> ) . . . . .	460
35.1 Einführung . . . . .	460
35.2 Sekundärdaten und Qualitätssicherung? . . . . .	461
35.3 Gesetzliche Grundlagen . . . . .	462
35.4 Entwicklung eines sekundärdatenbasierten Qualitätssicherungsverfahrens im Leistungsbereich Dekubitusprophylaxe . . . . .	463
35.5 Nutzung von stationären Abrechnungsdaten im Rahmen der Veröffentlichung der Ergebnisqualität für Perinatalzentren der Level 1 und Level 2 «Phase B» . . . . .	467
35.6 Nutzung von Sekundärdaten zur Weiterentwicklung des Leistungsbereichs Cholezystektomie . . . . .	468
35.7 Resümee und Ausblick . . . . .	470
36. Vorschlag für eine standardisierte Berichterstattung von Sekundärdatenanalysen ( <i>Jochen Schmitt, Enno Swart</i> ) . . . . .	474
36.1 Vorbemerkungen . . . . .	477
36.2 Zielsetzung und Vorgehen . . . . .	478
36.3 Bewertung der STROBE-Kriterien und Empfehlungen für deren Ergänzung . . . . .	478
36.4 Diskussion . . . . .	480
<b>Anhang</b> . . . . .	487
Gute Praxis Sekundärdatenanalyse (GPS): Leitlinien und Empfehlungen (3. Fassung 2012) (Arbeitsgruppe Erhebung und Nutzung von Sekundärdaten (AGENS) der Deutschen Gesellschaft für Sozialmedizin und Prävention (DGSM) und der Deutschen Gesellschaft für Epidemiologie (DGEpi)) . . . . .	488
Abkürzungsverzeichnis . . . . .	498
Glossar . . . . .	504
Autorenverzeichnis . . . . .	516
Adressen der Herausgeber . . . . .	526
Sachregister . . . . .	527