

Institut für
Rehabilitationsmedizinische Forschung
an der Universität Ulm



Beeinflusst die Arbeitslosenquote die (Wieder-)Eingliederung in das Erwerbsleben nach medizinischer Rehabilitation ?

Rainer Kaluscha¹, Silke Jankowiak¹, Jakob Holstiege¹, Gert Krischak^{1,2}

1. Institut für Rehabilitationsmedizinische Forschung an der Universität Ulm
2. Federseeklinik Bad Buchau, Abt. Orthopädie und Unfallchirurgie

Das Institut ...

- forscht auf den Gebieten *Rehabilitation* und *Prävention* in Ulm und um Ulm herum
- wurde 1995 als Stiftung des bürgerlichen Rechts gegründet und per Kooperationsvertrag der Universität Ulm angegliedert
- (Zu-)Stifter:
 - ❖ Deutsche Rentenversicherung Baden-Württemberg
 - ❖ Waldburg-Zeil-Kliniken
 - ❖ Universitätsklinikum Ulm
 - ❖ Therapiezentrum Federsee, Bad Buchau
 - ❖ Nanz Medico

www.uni-ulm.de/institute/reha-forschung/



Routinedaten am Institut

Dr. biol.hum. Dipl.-Inform. Rainer Kaluscha

- Versorgungsforschung mit Routinedaten ist einer der Schwerpunkte am Institut
- seit 1999 Nutzung von anonymisierten Routinedaten der damaligen Landesversicherungsanstalt Württemberg (Forschungsdatenbank „Patientenkonto“)
- Pilotnutzer beim Forschungsdatenzentrum der Rentenversicherung (FDZ-RV)
- Evaluation von Modellprojekten unter Einbeziehung von Routinedaten (insbesondere mittel- und langfristige Beitragszahlungen in die Sozialsysteme)



Fragestellung

„Welche Faktoren beeinflussen die berufliche Wiedereingliederung eines Rehabilitanden ?“

- Patientenbezogene (Diagnose, Alter, Geschlecht, ...)
- Behandlungsbezogene (Akut, Reha, Nachsorge)
- Umfeld (Betrieb, Arbeitsmarkt, ...)

→ Hat die Arbeitslosenquote einen Einfluss ?



Methode

Datenbasis:

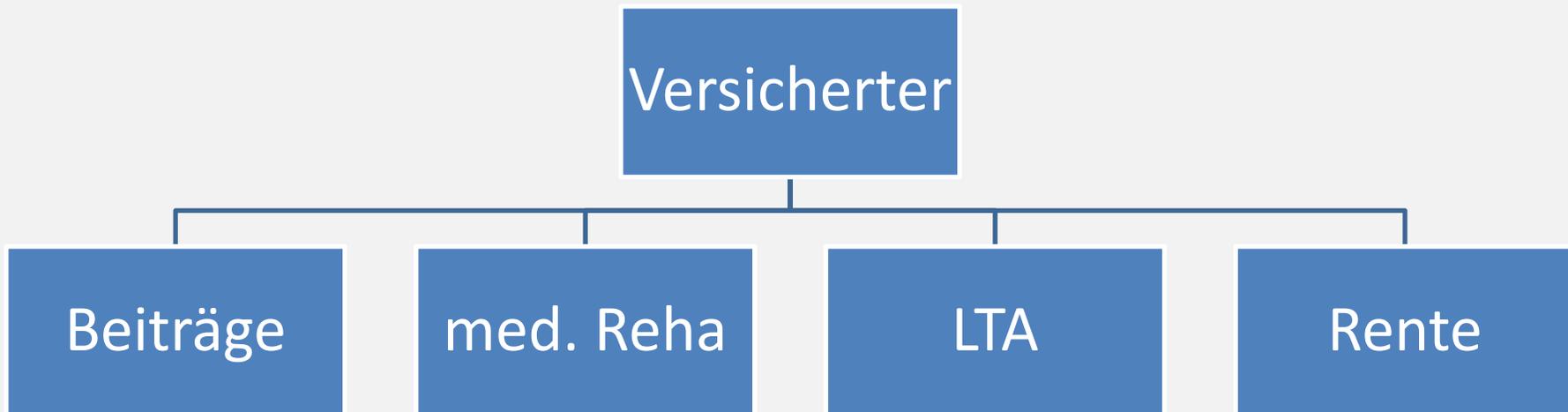
- Scientific Use File „Abgeschlossene Rehabilitation im Versicherungsverlauf 2002 – 2009“ des FDZ-RV ¹ (Reha-Längsschnittdatensatz SUFRSDLV09B)
- monatliche Arbeitslosenquoten der Agentur für Arbeit ² (Bundes- und Landesquoten)
- Verknüpfung SUF und Arbeitslosenquoten über Bundesland und Monat / Jahr möglich

Modell: logistische Regression zur Vorhersage des Ausscheidens aus dem Erwerbsleben

1. Forschungsdatenzentrum der Rentenversicherung, vgl. Vortrag Anja Bestmann und www.fdz-rv.de
2. Vgl. statistik.arbeitsagentur.de/



Daten-Struktur



Zeitfenster: 8 Jahre (Beiträge bis zu 11 Jahre)

Population: Antragsteller auf med. Reha, LTA, Rente

CAVE: bei Analysen Berücksichtigung etlicher sozialrechtlicher Besonderheiten erforderlich !



Stichprobe „Berufstätige“

aus Reha-Längsschnittdatensatz des FDZ-RV:

- im Jahr vor der med. Reha
mind. 6 Monate versicherungspflichtig gearbeitet
mit mind. 6.000,- € Entgelt
- Erkrankung aus den ICD-10-Kapiteln „M“, „F“, „I“, „C“, „J“
(die fünf größten Gruppen in der Rehabilitation)
- Bundesland bekannt (d.h. Wohnort Inland)
- Reha in den Jahren 2003 - 2007

→ Stichprobenumfang: **412.155** Fälle



Zielgröße: Verbleib im Erwerbsleben

„Kandidaten“ für Operationalisierung im SUF:

jahresweise:

- **versicherungspflichtiges Entgelt**
- Tage versicherungspflichtiger Beschäftigung
- Berufsklassenschlüssel

monatlich:

- Beitragsart (aus Beschäftigung, Sozialleistungen etc.)
- Anrechnungszeiten (Erziehung, Pflege etc.)



Operationalisierung der Zielgröße

binäre Zielgröße: „draußen“
(d.h. mit hoher Wahrscheinlichkeit aus dem Erwerbsleben ausgeschieden),

- wenn **sozialversicherungspflichtiges Entgelt** im *Folgejahr* der Reha **weniger als 5%** des *Vorjahres* der Reha beträgt

→ **87.569** Fälle (**21.25%**) sind „draußen“

→ logistische Regression (SAS 9.3, Proc Logistic)

→ Prädiktoren / potentielle Confounder ?



Prädiktoren / Confounder SUF 1

- Geschlecht
- Erkrankungsgruppe (ICD 1-stellig)
- Bundesland
- Berufsklasse nach Blossfeld
- Stellung im Beruf
- Bildung
- zumutbarer zeitlicher Umfang (< 3h, 3-6h, > 6h):
 - aktuelle Tätigkeit
 - allgemeiner Arbeitsmarkt



Prädiktoren / Confounder SUF 2

- Entlassung arbeitsfähig
- Reha: stationär / ambulant
- Alter (Jahre)
- Entgelt im Vorjahr der Reha (in 1.000 €)
- 12 potentielle Confounder im SUF verfügbar
→ $2^{12} = 4.096$ Modellvarianten



Confounder Arbeitsmarkt

Zeitpunkt / Quote	Bund absolut	Land absolut	Land standardisiert*
Reha-Monat	1. Variante	2. Variante	3. Variante
12 Monate danach	4. Variante	5. Variante	6. Variante
Durchschn. 12 Monate	7. Variante	8. Variante	9. Variante

** Standardisierung der Landesquote auf den Durchschnitt des jeweiligen Bundeslandes im Zeitraum 1999 - 2009*

10. Variante: ohne Arbeitsmarktindikator

- 4.096 Modellvarianten * 10 Arbeitsmarktindikatoren
40.960 Modellvarianten insgesamt (Full Model Selection)
(vollständige Angaben zu 385.937 Fällen verfügbar)



Technisches

- Datenbank: Oracle 10g
- Statistik: SAS 9.3 (Proc Logistic)
- Computeserver: 16 CPU-Kerne, 256 GB RAM
- SuSE Linux Enterprise 11
- Batch-Files: Shell-Scripts
- Unix-Utilities: grep, cut, sort, ...
- Perl
- deutliche Beschleunigung bei Full Model Selection durch Parallelisierung (16 SAS-Jobs) möglich ¹

1. vgl. Beitrag Kaluscha, Jankowiak, Krischak zur KSFE 2013: www.ksfe2013.de



Modellauswahl

Cave:

- Variablenselektion hat Besonderheiten bei großen Datensätzen
- Multikollinearitätsprüfung ist hier z.T. schwierig, da auch zeitlich variable Abhängigkeiten zwischen *kategorialen* (Bundesland) und *numerischen* (Arbeitslosenquoten) Confoundern vorhanden sind
- ferner subtile Abhängigkeiten zwischen Variablen, z.B. *zeitlicher Umfang aktuelle Tätigkeit* kann nicht größer als *zeitlicher Umfang allgemeiner Arbeitsmarkt* sein



Modellauswahl

Kriterien:

- Modellgüte (insbesondere c-Statistik, AIC)
- inhaltliche Überlegungen

- c-Statistik der Modellvarianten liegt zwischen 0.502 und 0.812
- d.h. beste Modelle sagen Ausscheiden aus Erwerbsleben in 81,2% der Fälle korrekt voraus



Ergebnisse Modell 4095-9

Ausgewähltes Modell: alle zwölf SUF-Variablen und standardisierter
12-Monatsdurchschnitt der Landesquote

Typ-3-Effektanalyse			
Effekt	DF	Waldsches Chi-Quadrat	Pr > ChiSq
Geschlecht	1	878.5261	<.0001
Erkrankung	4	3995.5164	<.0001
Bundesland	15	420.4269	<.0001
Beruf nach Blossfeld	13	407.4606	<.0001
Stellung im Beruf	6	1494.9869	<.0001
Bildung	6	502.3995	<.0001



Ergebnisse Modell 4095-9

Typ-3-Effektanalyse			
Effekt	DF	Waldsches Chi-Quadrat	Pr > ChiSq
zeitl. Leistungsfähigkeit letzte Tätigkeit	2	15212.1538	<.0001
zeitl. Leistungsfähigkeit allgemeiner Arbeitsmarkt	2	49.7606	<.0001
Entlassung arbeitsfähig	1	8570.4368	<.0001
Ambulant / Stationär	1	2.2136	0.1368
Alter	1	1924.1841	<.0001
Entgelt Vorjahr (in 1.000 €)	1	14877.2921	<.0001



Ergebnisse Modell 4095-9

Odds-Ratio-Schätzer			
<i>Effekt</i>	<i>Punktschätzer</i>	<i>95% Waldsche Konfidenzgrenzen</i>	
Männer zu Frauen	1.474	1.436	1.512
ICD10: C vs M	0.621	0.596	0.648
ICD10: F vs M	1.985	1.935	2.037
ICD10: I vs M	0.850	0.824	0.877
ICD10: J vs M	0.922	0.855	0.994

nicht dargestellt: Bundesland, Berufsgruppe nach Blossfeld, Stellung im Beruf



Ergebnisse Modell 4095-9

Odds-Ratio-Schätzer			
Effekt	Punktschätzer	95% Waldsche Konfidenzgrenzen	
Zeitl. Umfang <i>aktuelle Tätigkeit</i> : < 3h zu >6h	5.115	4.978	5.255
Zeitl. Umfang <i>aktuelle Tätigkeit</i> : 3-6h zu >6h	2.733	2.647	2.823
Zeitl. Umfang <i>allg. Arbeitsmarkt</i> : < 3h zu >6h	1.103	1.016	1.197
Zeitl. Umfang <i>allg. Arbeitsmarkt</i> : 3-6h zu >6h	0.779	0.723	0.839
Entlassung arbeitsfähig	0.373	0.365	0.381



Ergebnisse Modell 4095-9

Odds-Ratio-Schätzer			
<i>Effekt</i>	<i>Punktschätzer</i>	<i>95% Waldsche Konfidenzgrenzen</i>	
Alter	1.026	1.025	1.027
Entgelt Vorjahr (in 1.000 €)	0.999	0.999	0.999
Arbeitslosenquote Land: Durchschnitt 12 Monate, standardisiert	2.282	2.125	2.451



Diskussion

- Arbeitslosenquote ist nur ein Aspekt des Arbeitsmarktes
- Arbeitslosigkeit **nach** Reha ist besserer Prädiktor als Arbeitslosigkeit **während** Reha
- Bundesquote und standardisierte Landesquote sind Landesquote überlegen
- zumutbarer zeitlicher Umfang der letzten Tätigkeit wichtiger als der auf dem allgemeinen Arbeitsmarkt (höheres Odds Ratio)
- zumutbarer zeitlicher Umfang auf dem allgemeinen Arbeitsmarkt von 3-6h wirkt anscheinend protektiv, d.h. Teilrente plus Teilzeittätigkeit statt Komplettausstieg ?



Diskussion

- Ist die inhaltliche Interpretation der Variable „Bundesland“, wenn
 - nur Bundesland
 - Bundesland und Bundesquote
 - Bundesland und Landesquoteim Modell, immer gleich oder unterschiedlich ?
- *zumutbarer zeitlicher Umfang auf dem allgemeinen Arbeitsmarkt von 3-6h wirkt*
 - protektiv, wenn auch *zumutbarer zeitlicher Umfang bei der aktuellen Tätigkeit* im Modell ist
 - risiko-erhöhend, wenn nur *zumutbarer zeitlicher Umfang auf dem allgemeinen Arbeitsmarkt* im Modell



Fazit

- Modellierung kann durch Hinzunahme externer Information aus öffentlichen Datenquellen präziser werden
- sorgfältige Analysen der möglichen Modellvarianten und Prüfung auf Robustheit der Ergebnisse erforderlich
- Arbeitsmarkt ist wichtige Einflussgröße auf berufliche Wiedereingliederung nach einer med. Rehabilitation ...
- und muss daher als potentieller Confounder berücksichtigt werden, z.B. bei Länder- oder Einrichtungsvergleichen



Vielen Dank für Ihre Aufmerksamkeit!

Institut für
Rehabilitationsmedizinische Forschung
an der Universität Ulm



Kontakt: rainer.kaluscha@uni-ulm.de

www.uni-ulm.de/institute/reha-forschung