

4 Daten und Methoden

Im Folgenden werden zuerst die Datenquellen besprochen sowie deren Potenziale und Probleme diskutiert. Als nächstes wird ausführlich auf die abhängigen und unabhängigen Variablen eingegangen. Danach wird die Auswahl der multivariaten statistischen Methoden – negativ-binomiale und multinomiale Regressionsmodelle – begründet und es werden die wesentlichen Aspekte der Methoden vorgestellt.

4.1 Daten

Für die Analysen werden drei verschiedene Datenquellen des Schweizerischen Bundesamtes für Statistik (BFS) verwendet:

- (1) Die „Swiss National Cohort“ (SNC) ist eine anonyme Verlinkung zwischen Daten aus der Volkszählung von 1990 und 2000 sowie der Todesursachenstatistik unter Berücksichtigung von Zu- und Auswanderung in die Schweiz. Diese Zensusdaten wurden zwischen 1850 und 2000 alle zehn Jahre erhoben und beinhalten Informationen über Bildung, Haushaltssituation sowie familiäre Situation. 1990 und 2000 wurde ausserdem das genaue Geburtsdatum erhoben, was die Verlinkung mit der Todesursachenstatistik ermöglichte (Bopp et al. 2009).
- (2) Die Medizinische Statistik der Krankenhäuser (MedStat) erfasst jährlich die Daten aller Hospitalisierungen in Schweizer Krankenhäusern.¹⁵ Darin ist jedes Krankenhaus bzw. jede Klinik mit eingeschlossen. Neben medizinischen Informationen wie Ein- und Austritt in ein Spital, Diagnosen und Behandlungen werden auch soziodemografische Angaben über Alter, Geschlecht oder Wohnregion erhoben (vgl. Bopp et al. 2011).
- (3) Die Statistik der sozialmedizinischen Institutionen (SOMED) ist eine administrative Statistik über die Tätigkeiten von Betrieben, die sich um Betagte

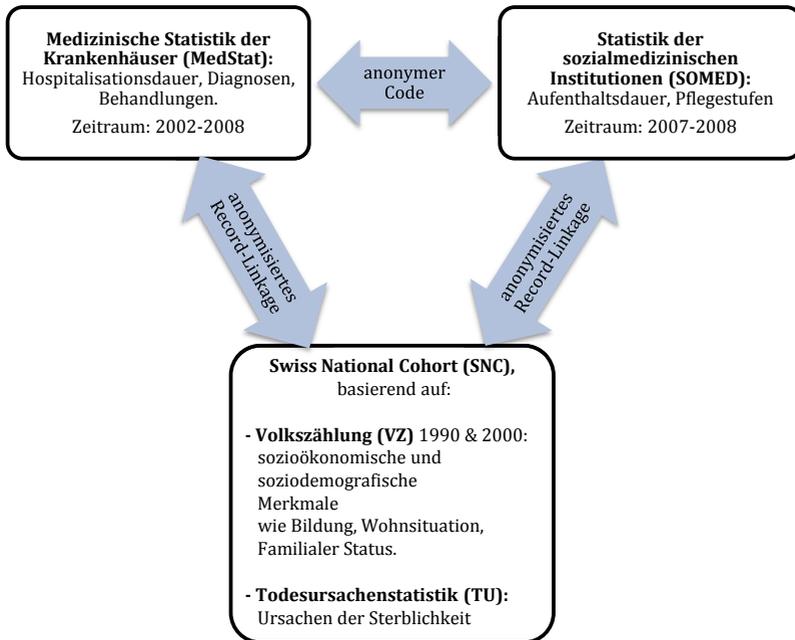
¹⁵ Weitere Informationen zur MedStat:
http://www.bfs.admin.ch/bfs/portal/de/index/infothek/erhebungen_quellen/blank/blank/mkh/01.html [Zugriff: 3.2.2015].

oder Behinderte kümmern.¹⁶ Alle diese Institutionen erstatten jährlich Bericht über ihre erbrachten Leistungen wie etwa Pflege und betreute Patienten.

Alle drei Datenquellen sind obligatorische Vollerhebungen und wurden primär für administrative Zwecke erhoben. Aus Datenschutzgründen gibt es in der Schweiz keine gemeinsame personale Identifikationsnummer, die es erlauben würde, die Personen auf individueller Ebene zu verlinken. Für die Analysen werden aber gleichzeitig Informationen aus allen drei Datenquellen benötigt. Eine anonyme Nummer ist nur verfügbar für die Verlinkung zwischen der MedStat und der SOMED, während die Verlinkung mit der SNC anhand übereinstimmender Merkmale wie Wohnort oder Geburts- und Sterbedatum erfolgt. In der MedStat und der SOMED sind aber die Informationen über das genaue Sterbedatum sowie den Geburtsjahrgang nur für in Spitälern bzw. in Heimen Verstorbene verfügbar. Im Endeffekt ist dadurch nur für Verstorbene eine erfolgreiche Verlinkung aller drei Quellen möglich (Zellweger & Bopp 2010, 2011). In Abbildung 4.1 ist die Verlinkung der drei Quellen grafisch dargestellt.

¹⁶ Weitere Informationen zur SOMED:
http://www.bfs.admin.ch/bfs/portal/de/index/infothek/erhebungen_quellen/blank/blank/ssmi/01.html [Zugriff: 3.2.2015].

Abbildung 4.1: Drei Datenquellen und deren Verlinkung



(Quelle: Eigene Darstellung)

Die jeweiligen Verlinkungsraten unterscheiden sich zwischen den einzelnen Quellen. Die Verlinkung zwischen der SNC und der Medstat war für 98.9% der Einträge¹⁷ möglich, dabei stimmen 95.1% der Einträge vollständig überein und 3.8% grösstenteils. Dies betrifft beispielsweise Fälle mit identischem Geschlecht, Geburtsdatum und Wohnregion (etwa 600 Regionen in der gesamten Schweiz), aber mit kleineren Differenzen beim genauen Sterbedatum. Die Datenqualität im Zeitraum 2002-2008 ist relativ homogen (Zellweger & Bopp 2010). Die Verlinkung der SNC mit der SOMED ist einig schwieriger, weil diese Datenquelle erst seit 2006 in dieser Form existiert und die Qualität erst ab 2007 sinnvolle Analy-

¹⁷ Unter Einträge wird dabei jeder einzelne Spital- oder Heimaufenthalt verstanden und nicht Personen. Eine einzelne Person kann theoretisch beliebig viele Einträge haben. Innerhalb der MedStat oder SOMED ist jeder Eintrag in jedem Fall eindeutig den einzelnen Personen innerhalb der MedStat bzw. SOMED zugewiesen.

sen zulässt. Dennoch existieren in der Form des Eintrittsdatums und der Pflegestufen wichtige Information bereits vor 2007; denn wenn eine Person etwa im Jahr 2008 verstorben ist, aber bereits seit 2002 in einem Heim wohnte, so ist diese Information in dem Eintrag von 2008 enthalten, weil das Ein- und Austrittsdatum bei allen Einträgen angegeben ist. Gleiches gilt für die Pflegestufen, die theoretisch während des gleichen Aufenthaltes beliebig häufig wechseln können, leider aber von den Heimen nur unzuverlässig ausgefüllt wurden (vgl. Bayer-Oglesby & Höpflinger 2010). Die Verlinkungsrate für 2007 und 2008 beträgt insgesamt 92%, davon können 65.1% direkt anhand vollständiger Informationen verlinkt werden und 25.9% anhand grösstenteils übereinstimmender Merkmale (Zellweger & Bopp 2011).

Dieses Verlinkungsverfahren ermöglicht es in der Schweiz erstmals, Zensusdaten mit Spital- und Heimdaten am Lebensende zu verbinden, wodurch Analysen über die Versorgungslage am Lebensende überhaupt erst möglich werden. Und weil es sich um Vollerhebungen handelt, existieren einerseits keine Probleme mit der Repräsentativität der Daten und andererseits sind - aufgrund der hohen Fallzahlen - spezifische Analysen mit kleineren Gruppen möglich. Weitere Stärken der verwendeten Daten sind die hohe Qualität der SNC- und MedStat-Datenquellen sowie die für administrative Daten umfangreiche Anzahl an horizontalen und vertikalen Statusindikatoren. Letztlich ermöglicht das innovative Verlinkungsverfahren, spezifische Analysen über soziale Ungleichheit von Heim- und Spitalaufenthalten zu rechnen, die man bis anhin in der Schweiz so nicht durchführen konnte.

Demgegenüber stehen auch einige Schwächen. Der grösste Nachteil der Daten ist sicherlich die Beschränkung auf Verstorbene, was aufgrund der Verlinkung der Datenquellen leider nicht anders möglich ist und beispielsweise keine Vergleiche zwischen Überlebenden und Verstorbenen zulässt. Weil aber genaue Informationen über die Institutionsbiografie vor dem Tod zur Verfügung stehen, werden retrospektive Analysen über Ungleichheit und Gesundheit auch vor dem Tod möglich. Des Weiteren sind viele soziodemografische Angaben wie etwa das Bildungsniveau nicht zum Todeszeitpunkt verfügbar. Ein zusätzliches Problem ist die teilweise verbesserungswürdige Datenqualität der Heimdaten, insbesondere bei einigen spezifischen Angaben wie etwa den Pflegestufen oder den genauen Todesdaten (vgl. Bayer-Oglesby & Höpflinger 2010). Die genauen Todesdaten konnten aber mit einem Abgleichen mit der Todesursachenstatistik korrigiert

werden (vgl. Zellweger & Bopp 2011).¹⁸ Auch konnten bei beiden Verlinkungen (SOMED und SNC sowie MedStat und SNC) nicht alle Einträge anhand vollständig übereinstimmender Kriterien verlinkt werden, was zu möglichen Fehlzusweisungen bzw. Verzerrungen führen kann, etwa wenn einige Fälle aus systematischen Gründen fehlten. Dieses Problem ist bei der Verlinkung der SOMED mit der SNC aufgrund der geringeren Verlinkungsrate stärker ausgeprägt als bei der Verlinkung zwischen der MedStat und der SNC.

Die Datengrundlage in der Schweiz ist verglichen mit anderen Ländern wie etwa Schweden oder Norwegen, die über zentrale Bevölkerungsregister mit zugewiesenen anonymen Codes für jedes Individuum verfügen und dadurch nahezu beliebig Daten zwischen den einzelnen Quellen verknüpfen können, sicherlich nicht ganz optimal. Weiterhin ist das föderalistische System der Schweiz nicht hilfreich, schweizweit gültige Parameter wie etwa homogene Pflegestufen zu berechnen. Auf der anderen Seite überrascht - gerade bei nicht ganz optimal zu operationalisierenden Variablen wie beispielsweise den Pflegestufen - die Konsistenz der Resultate.

Sämtliche Berechnungen in allen Stichproben werden anhand von Verstorbenen ab 65 Jahren (d. h. vor 1942 geborene Personen) in den Jahren 2007 und 2008 durchgeführt.¹⁹

4.2 Operationalisierung

Für die empirischen Analysen werden drei multivariate Modelle, jeweils separat für Männer und Frauen berechnet. Dazu werden drei verschiedene abhängige Variablen und drei - je nach Forschungsfrage - unterschiedliche Untersuchungspopulationen und teilweise auch unterschiedliche unabhängige Variablen verwendet.

Heimaufenthaltsdauer

Als Indikator der Heimaufenthaltsdauer wird bei allen Personen die Länge des letzten Heimaufenthaltes vor dem Tod in Tagen in sämtlichen Alters- und Pflege-

¹⁸ In der SOMED war das Austrittsdatum, welches bei einem Todesfall mit dem Sterbedatum identisch sein muss, in einigen Fällen vermutlich das administrativ benutzte Austrittsdatum, welches logischerweise immer nach dem effektiven Sterbedatum lag.

¹⁹ Aus pragmatischen Gründen wird bei Frauen und Männern das gleiche Rentenalter verwendet, obwohl bei Frauen das ordentliche Rentenalter in der Schweiz gegenwärtig bei 64 Jahren festgelegt ist.

heimen mit anschliessendem Tod im Heim benutzt. Dabei werden nur Personen ab mindestens 65 Jahren, die gleichzeitig auch frühestens mit 65 Jahren in ein Heim eingetreten sind, analysiert. Die Studienpopulation beträgt $N=11'486$ Männer und $N=24'252$ Frauen.

Für die Heimanalysen wird mit dem Alter zum Eintrittszeitpunkt gerechnet. Der höchste Bildungsabschluss wird aus den SNC-Daten in ISCED-97 Stufen, angepasst an die Schweiz²⁰, umgerechnet: Keine oder tiefe Bildung (ISCED 0-2), mittlere Bildung (ISCED 3-4) und hohe Bildung (ISCED 5-6). Ebenfalls aus den SNC-Daten stammt die Wohnungs- und Hausbesitzvariable (ja/nein).

Die Todesursachen, wiederum aus der SNC, werden anhand von ICD-10²¹ Kodierungen aufgrund von Angaben in der Todesursachenstatistik operationalisiert: Krebs (ICD 10: C00-C99), koronare Herzkrankheiten (I20-I25), chronisch obstruktive Lungenerkrankungen (COPD) (J40-J47), Demenz (F01, F03, G30)²², Schlaganfall (I60-I69) sowie eine Restgruppe mit allen anderen Todesursachen.

Die Pflegestufen sind in der SOMED enthalten, leider wurden die Angaben von den Heimen teilweise mangelhaft ausgefüllt (vgl. Bayer-Oglesby & Höpflinger 2010). Des Weiteren gibt es schweizweit neun unterschiedliche Einstufungssysteme, die teilweise zwischen und innerhalb der Kantone variieren und die man nur anhand der Anzahl Pflegeminuten pro Tag in eine landesweit nutzbare Variable umrechnen kann. Auch mögliche Pflegestufenwechsel sind nicht immer zuverlässig erfasst, weshalb in dieser Arbeit die erste erfasste Pflegestufe des letzten Heimaufenthaltes verwendet wird. Schlussendlich wurden vier Stufen gebildet: Fehlende Angabe bzw. unbekannt, keine Pflege notwendig oder wenig pflegebedürftig (bis 40 Minuten pro Tag), mittel pflegebedürftig (41 bis 80 Minuten pro Tag) sowie stark pflegebedürftig (mehr als 81 Minuten pro Tag). Dabei dient das BESA-4-Stufen System als Grundlage, welches besonders in einigen deutsch-

²⁰ Weitere Informationen über die ISCED-97-Kodierung in der Schweiz:

<http://www.bfs.admin.ch/bfs/portal/de/index/infothek/nomenklaturen/blank/blank/isced/01.html> [Zugriff: 5.2.2015].

²¹ Es wird die im deutschsprachigen Raum üblicherweise verwendete ICD-10-GM-Version benutzt. Informationen unter: <https://www.dimdi.de/static/de/klassi/icd-10-gm/> [Zugriff: 5.2.2015].

²² Die Alzheimer-Demenz stellt mit über 50% aller Fälle im hohen Alter die häufigste Form von Demenzerkrankungen dar. Die Ursachen dieser Krankheit sind erst teilweise erforscht, in der Diagnose, Behandlung und Betreuung sind jedoch in jüngerer Zeit klare Fortschritte zu erkennen (Höpflinger & Hugentobler 2003). Gemäss der obigen Definition von Todesursachen ist ausserdem davon auszugehen, dass Demenz verhältnismässig oft vorkommt.

sprachigen Kantonen verbreitet ist.²³ Obwohl die so zusammengestellte Variable besser als nichts ist, stellt sie eine Schwachstelle in diesem Modell dar.

Die Multimorbidität wird mit Daten aus der MedStat berechnet und es werden insgesamt 129 chronische Krankheiten anhand der ICPC-2-Klassifikation berücksichtigt (vgl. O'Halloran et al. 2004). Dazu werden nur Spitalaufenthalte im Zeitraum 2-6 Jahren bzw. 730-2190 Tagen vor dem Tod der jeweiligen Person berücksichtigt, weil andernfalls die Prävalenzraten viel zu hoch ausfallen würden. Gleichzeitig muss Multimorbidität, d. h. mindestens zwei chronische Krankheiten gleichzeitig, während des gleichen damaligen Spitalaufenthaltes diagnostiziert worden sein. Danach werden drei Stufen gebildet: Keine Abklärung bzw. kein Spitalaufenthalt im besagten Zeitraum, keine Multimorbidität und Multimorbidität. Zur Kontrolle von Spitalaufenthalten wird noch eine zusätzliche Variable von Spitalaufenthalten (ja/nein) in den letzten 365 Tagen vor dem Tod berücksichtigt.

Beim Zivilstand war es möglich, allfällige Änderungen wie etwa eine Verwitwung seit der Volkszählung im Jahr 2000 zu berücksichtigen, wodurch man den Zivilstand beim Heimeintritt berechnen konnte (ledig, verheiratet, verwitwet, geschieden). Als weitere familiäre Variable aus der SNC wird noch untersucht, ob die Personen Kinder hatten (ja/nein).

Bei der kulturell-kontextuelle Variable Nationalität, die aus der SNC stammt, wird zwischen Schweizern und Nicht-Schweizern unterschieden. Die Wohnregion wird anhand der Kantone gebildet, zweisprachige Kantone werden dabei der Mehrheit von Sprechern zugewiesen. Beispielsweise werden die Kantone Wallis und Fribourg der französischsprachigen Schweiz zugeteilt.²⁴ Die italienischsprachige Schweiz besteht nur aus dem Kanton Tessin. Als letzte unabhängige Variable wird die Heimdichte bzw. Heimbettverfügbarkeit pro 100 Einwohner ab 65 Jahren in einer bestimmten Region verwendet. Die total 106 Regionen wurden dabei vom BFS gebildet mit dem Ziel, möglichst homogen und vergleichbar zu sein.²⁵ In Tabelle 4.1 sind die Variablen sowie ihre Verteilungen für die Analyse der Heimaufenthaltsdauer abgebildet.

²³ Eine Übersicht über die BESA-Stufen liefert die Informationsseite des Walliser Gesundheitsobservatoriums: <http://www.ovs.ch/gesundheit/besa.html> [Zugriff: 5.2.2015].

²⁴ Informationen über die genaue Verteilung der Sprachen bzw. Sprachregionen: http://www.bfs.admin.ch/bfs/portal/de/index/regionen/thematische_karten/maps/bevoelkerung/sprachen_religionen.html [Zugriff: 5.2.2015]

²⁵ Weitere Informationen über die 106 BFS-Regionen: <http://www.bfs.admin.ch/bfs/portal/de/index/infothek/nomenklaturen/blank/blank/msreg/01.html> [Zugriff: 5.2.2015].

Tabelle 4.1: Variablen Heimaufenthaltsdauer

	Männer		Frauen	
	N		N	
Total N	11'486	100 %	24'253	100 %
Individuelle Faktoren				
Alter (zum Eintrittszeitpunkt)				
≤75	1'475	12.8 %	1'862	7.7 %
>75-≤80	1'745	15.2 %	3'100	12.8 %
>80-≤85	2'964	25.8 %	6'171	25.5 %
>85-≤90	3'185	27.7 %	7'204	29.7 %
>90	2'117	18.4 %	5'915	24.4 %
Bildung				
tief	3'599	31.3 %	12'740	52.5 %
mittel	4'168	36.3 %	6'089	25.1 %
hoch	1'615	14.1 %	661	2.7 %
unbekannt	2'104	18.3 %	4'762	19.6 %
Haus- oder Wohnungseigentümer				
kein Eigentümer	6'784	59.1 %	16'815	69.3 %
Eigentümer	4'702	40.9 %	7'437	30.7 %
Todesursache				
Krebs	2'119	18.5 %	2'794	11.5 %
Herzinfarkt	1'804	15.7 %	3'834	15.8 %
COPD	566	4.9 %	551	2.3 %
Demenz	1'718	15.0 %	4'322	17.8 %
Schlaganfall	1'013	8.8 %	2'308	9.5 %
Rest	4'266	37.1 %	10'443	43.1 %
Pflegestufe				
tief	3'082	26.8 %	6'954	28.7 %
mittel	3'114	27.1 %	6'966	28.7 %
hoch	4'781	41.6 %	9'555	39.4 %
unbekannt	509	4.4 %	777	3.2 %
Multimorbidität				
Nein	5'289	46.1 %	10'195	42.0 %
Ja	4'528	39.4 %	8'498	35.0 %
nicht diagnostiziert	1'669	14.5 %	5'559	22.9 %

...Fortsetzung von Tabelle 4.1

Hospitalisation in den letzten

365 Tagen

Nein	6'035	52.5 %	16'421	67.7 %
Ja	5'451	47.5 %	7'831	32.3 %

Familiäre Faktoren

Zivilstand (zum Eintrittszeitpunkt)

verheiratet	6'098	53.1 %	3'824	15.8 %
ledig	1'213	10.6 %	2'960	12.2 %
verwitwet	3'499	30.5 %	15'904	65.6 %
geschieden	676	5.9 %	1'564	6.5 %

Kinder

Nein	2'074	18.1 %	4'802	19.8 %
Ja	8'441	73.5 %	17'025	70.2 %
unbekannt	971	8.5 %	2'425	10.0 %

Kulturell-kontextuelle Faktoren

Nationalität

Schweiz	10'983	95.62 %	23'510	96.94 %
Ausland	503	4.38 %	742	3.06 %

Sprachregion

Deutschschweiz	8'770	76.35 %	17'811	73.44 %
Westschweiz	2'279	19.84 %	5'296	21.84 %
Tessin	437	3.8 %	1'145	4.72 %

Heimbettdichte*: Mittelwert	7.1	7.0
------------------------------------	-----	-----

Datenquelle: MEDSTAT/SOMED/TU/SNC 2007/2008. Basis: Eigene Berechnungen.

*Durchschnittliche Anzahl verfügbarer Betten in Alters- und Pflegeheimen pro 100 mindestens 65 Jahre alten Bewohnern nach Region (Total 106 Regionen)

Spitalaufenthaltsdauer

Bei der Spitalaufenthaltsdauer wird die Aufenthaltsdauer in Tagen von sämtlichen stationären Spitalaufenthalten im letzten Lebensjahr bzw. in den letzten 365 Tagen vor dem Tod berücksichtigt.²⁶ Dabei werden nur Personen ab 65 Jahren

²⁶ Bei einem überschneidenden Spitalaufenthalt im Zeitraum vor einem Jahr vor dem Tod werden erst die Tage ab dem Stichtag vor 365 Tagen vor dem Tod gezählt, wodurch das theoretische Maximum 365 Tage beträgt. Als ein Tag gilt dabei immer nur ein Aufenthalt mit Übernachtung im Spital.

untersucht, die mindestens eine Hospitalisierung²⁷ im gleichen Zeitraum hatten und dadurch auch mit einem Spitalaufenthalt bzw. einer Spitalbehandlung konfrontiert wurden. Aufgrund mangelnder Daten bzw. Verlinkungsmöglichkeiten und der Homogenität der Untersuchungspopulation werden dabei nur Personen untersucht, die letztlich auch im Spital verstorben sind. Die Studienpopulation beträgt N=18'993 Männer und N=16'605 Frauen.

Die unabhängigen Variablen werden grösstenteils von den Heimanalysen übernommen. Beim Alter wird mit dem genauen Alter beim Todeszeitpunkt gerechnet. Bildungsniveau, Wohneigentum, Todesursachen, Multimorbidität, Kinder sowie die Nationalität, Sprachregionen und die Heimdichte werden identisch operationalisiert wie bei den Heimanalysen. Der Zivilstand wird aus der Todesursachenstatistik übernommen. In Tabelle 4.2 sind die Variablen sowie ihre Verteilungen für die Analyse der Heimaufenthaltsdauer abgebildet.

Tabelle 4.2: Variablen Spitalaufenthaltsdauer

	Männer		Frauen	
	N		N	
Total N	18'993	100 %	16'605	100 %
Individuelle Faktoren				
Alter (zum Todeszeitpunkt)				
≤75	5'833	30.7 %	3'756	22.6 %
>75-≤80	4'012	21.1 %	2'948	17.8 %
>80-≤85	4'378	23.1 %	4'001	24.1 %
>85-≤90	3'190	16.8 %	3'441	20.7 %
>90	1'580	8.3 %	2'459	14.8 %
Bildung				
tief	5'426	28.6 %	8'200	49.4 %
mittel	7'624	40.1 %	4'785	28.8 %
hoch	3'307	17.4 %	716	4.3 %
unbekannt	2'636	13.9 %	2'904	17.5 %
Haus- oder Wohnungseigentümer				
kein Eigentümer	10'201	53.7 %	10'021	60.4 %
Eigentümer	8'792	46.3 %	6'584	39.7 %

²⁷ Unter Hospitalisierung wird jeder Spitaleintritt unabhängig von der Dauer verstanden.

...Fortsetzung von Tabelle 4.2

Todesursache

Krebs	7'243	38.1 %	5'390	32.5 %
Herzinfarkt	2'544	13.4 %	1'899	11.4 %
COPD	735	3.9 %	446	2.7 %
Demenz	233	1.2 %	220	1.3 %
Schlaganfall	1'270	6.7 %	1'580	9.5 %
Rest	6'968	36.7 %	7'070	42.6 %

Multimorbidität

Nein	4'429	23.3 %	3'821	23.0 %
Ja	6'637	34.9 %	5'521	33.3 %
nicht diagnostiziert	7'927	41.7 %	7'263	43.7 %

Familiäre Faktoren

Zivilstand (zum Todeszeitpunkt)

verheiratet	13'119	69.1 %	4'814	29.0 %
ledig	1'261	6.6 %	1'583	9.5 %
verwitwet	3'433	18.1 %	8'862	53.4 %
geschieden	1'180	6.2 %	1'346	8.1 %

Kinder

Nein	2'851	15.0 %	2'818	17.0 %
Ja	15'043	79.2 %	12'597	75.9 %
unbekannt	1'099	5.8 %	1'190	7.2 %

Kulturell-kontextuelle Faktoren

Nationalität

Schweiz	17'199	90.6 %	15'609	94.0 %
Ausland	1'794	9.5 %	996	6.0 %

Sprachregion

Deutschschweiz	12'674	66.7 %	11'367	68.5 %
Westschweiz	5'351	28.2 %	4'394	26.5 %
Tessin	968	5.1 %	844	5.1 %

Heimbettdichte*: Mittelwert

6.6

6.7

Datenquelle: MEDSTAT/SOMED/TU/SNC 2007/2008. Basis: Eigene Berechnungen.

*Durchschnittliche Anzahl verfügbarer Betten in Alters- und Pflegeheimen pro 100 mindestens 65 Jahre alten Bewohnern nach Region (Total 106 Regionen)

Unterschiedliche Sterbeorte

Anhand der theoretischen Überlegungen werden beim Sterbeort vier Kategorien gebildet: Alters- oder Pflegeheim, Spital kurz (Aufenthalt von maximal 48 Stunden bzw. 2 Tagen), Spital lang (Aufenthalt ab 3 Tagen) und zu Hause. Die Dauer der Spitalaufenthalte bezieht sich dabei immer auf den letzten Aufenthalt im gleichen Spital. Die zu Hause Verstorbenen werden anhand der Restgruppe der nicht in Institutionen Verstorbenen gebildet, bei denen davon ausgegangen wird, dass in dieser Altersgruppe der grösste Teil zu Hause verstorben ist. Anhand der Angaben aus der Volkszählung ist ersichtlich, wer bereits im Jahre 2000 in einem Alters- oder Pflegeheim lebte (sogenannte Kollektivhaushalte). Dieser Typus wird von vornherein ausgeschlossen, weil er eine spezifische Bevölkerungsgruppe darstellt, die einerseits eine durchschnittlich schlechtere Gesundheit als die Durchschnittsbevölkerung aufweist und andererseits mit sehr grosser Wahrscheinlichkeit im gleichen Heim bzw. in der gleichen Institution sterben wird. In dieser Arbeit interessieren aber primär Menschen, die damals noch in Privathaushalten lebten und erst später in eine Institution wechselten. Die Studienpopulation beträgt $N=39'798$ Männer und $45'331$ Frauen. Wie bei den Heim- und Spitalaufenthaltsanalysen werden auch hier Verstorbene ab 65 Jahren analysiert.

Viele unabhängige Variablen können wiederum von den Heim- und Spitalaufenthaltsanalysen übernommen werden. Das Alter interessiert dabei nur als Kontrollvariable und muss nicht interpretiert werden können; um eventuelle Abweichungen von der Linearität erfassen zu können werden deshalb kubische Splines mit Knoten bei 69, 78, 84, 89 und 96 Jahren modelliert. Weil die Fallzahlen durchwegs höher sind als bei den Spital- und Heimanalysen wird das Bildungsniveau noch feiner unterteilt: Anstatt nur eine Gruppe für solche mit keiner oder tiefer Bildung wird jeweils eine Gruppe mit keiner Bildung (ISCED 0) und tiefer Bildung (ISCED 1-2) gebildet. Die Wohnbesitzvariable und die Todesursachen werden identisch operationalisiert wie bei den Heim- und Spitalaufenthaltsanalysen. Zusätzlich zum Zivilstand aus der Todesursachenstatistik wird eine Variable mit der Anzahl Zimmer pro Person in der Privatwohnung im Jahr 2000 aus der SNC berechnet.²⁸ Dabei wird zwischen $0 \leq 1.5$, $1.5 < 2.5$ sowie > 2.5 Zimmer pro Person differenziert. Bei den Kindern wird zwischen keinen, ein bis zwei sowie drei oder mehr unterschieden. Die kulturell-kontextuellen Variablen Nationalität, Sprachregion und Heimdichte sind wiederum identisch wie oben beschrieben. In

²⁸ Zu diesem Zeitpunkt lebten ja alle noch in Privathaushalten, weil diejenigen, die früher in Kollektivhaushalten lebten, von vornherein ausgeschlossen wurden.

Tabelle 4.3 sind die Variablen sowie ihre Verteilungen (Spaltenprozent) für die Analyse der unterschiedlichen Sterbeorte abgebildet.

Tabelle 4.3: Variablen Sterbeort

Männer:	zu Hause	Spital kurz	Spital lang	Alters- und Pflege- heim
Total N	11'008	4'634	13'723	10'433
Individuelle Faktoren				
Alter				
≤75	29.8	30.3	30.9	9.4 %
>75-≤80	20.0	20.3	21.4	12.9 %
>80-≤85	21.9	23.0	23.3	23.4 %
>85-≤90	17.6	18.1	16.3	29.6 %
>90	10.8	8.3	8.2	24.8 %
Bildung				
kein Abschluss	2.9	3.2	3.1	3.2 %
tief	25.9	25.1	25.2	26.8 %
mittel	40.1	41.3	40.3	37.2 %
hoch	18.6	16.7	17.9	14.7 %
unbekannt	12.5	13.7	13.5	18.1 %
Haus- oder Wohnungseigentümer				
kein Eigentümer	47.7	53.5	55.1	57.9 %
Eigentümer	52.3	46.5	44.9	42.1 %
Todesursache				
Krebs	28.1	25.7	41.5	18.3 %
Herzinfarkt	25.9	19.8	11.2	15.7 %
COPD	4.3	4.3	3.7	4.8 %
Demenz	3.0	0.4	1.5	15.1 %
Schlaganfall	3.4	7.5	6.4	8.8 %
Übrige	35.2	42.3	35.7	37.2 %

...Fortsetzung von Tabelle 4.3

Familiäre Faktoren				
Zivilstand				
verheiratet	68.9	69.6	70.5	49.9 %
ledig	7.1	6.3	5.7	8.4 %
verwitwet	18.5	18.6	17.7	36.0 %
geschieden	5.5	5.5	6.1	5.7 %
Kinder				
keine Kinder	13.8	14.1	14.6	16.2 %
1-≤2 Kinder	40.1	42.9	44.3	38.5 %
≥3 Kinder	40.6	37.4	35.7	37.5 %
unbekannt	5.5	5.5	5.4	7.8 %
Haushaltssituation				
Einpersonenhaushalt	18.2	17.2	17.5	28.7 %
Mehrpersonenhaushalt	81.8	82.8	82.5	71.3 %
Anzahl Zimmer pro Person				
0 - ≤1.5 Zimmer	35.0	38.1	38.1	36.4 %
>1.5 - ≤2.5 Zimmer	42.7	41.7	41.8	37.2 %
>2.5 Zimmer	22.3	20.2	20.0	26.5 %
Kulturell-kontextuelle Faktoren				
Nationalität				
Schweizer	93.6	91.5	90.1	95.5 %
Ausländer	6.4	8.5	9.9	4.5 %
Sprachregion				
Deutschschweiz	77.5	70.8	65.3	76.4 %
Westschweiz	18.8	24.3	29.3	19.7 %
Tessin	3.7	4.8	5.3	4.0 %
Heimbettdichte*: Mittelwert				
	6.8	6.7	6.6	7.1

...Fortsetzung von Tabelle 4.3

Frauen:

	zu Hause	Spital kurz	Spital lang	Alters- und Pfle- heim
Total N	9'267	3'900	11'992	20'172

Individuelle Faktoren

Alter

≤75	18.5	19.0	24.0	4.8 %
>75-≤80	14.4	16.2	18.7	8.3 %
>80-≤85	21.3	26.3	23.5	19.5 %
>85-≤90	22.4	23.2	19.8	29.3 %
>90	23.4	15.3	14.0	38.0 %

Bildung

kein Abschluss	4.7	4.3	4.6	5.0 %
tief	44.0	45.5	44.5	46.5 %
mittel	29.3	28.3	29.3	25.8 %
hoch	4.2	3.7	4.6	2.9 %
unbekannt	17.8	18.2	17.0	19.8 %

Haus- oder Wohnungseigentümer

kein Eigentümer	56.2	61.4	62.0	66.7 %
Eigentümer	43.8	38.6	38.0	33.3 %

Todesursache

Krebs	20.3	16.6	37.1	11.8 %
Herzinfarkt	20.5	16.8	9.5	15.6 %
COPD	2.8	3.3	2.4	2.3 %
Demenz	6.2	0.6	1.7	17.4 %
Schlaganfall	5.6	11.3	8.9	9.7 %
Übrige	44.6	51.4	40.4	43.3 %

Familiale Faktoren

Zivilstand

verheiratet	24.4	28.4	30.3	12.2 %
ledig	9.8	8.8	8.9	10.7 %
verwitwet	58.5	55.4	52.5	70.5 %
geschieden	7.2	7.3	8.3	6.5 %

...Fortsetzung von Tabelle 4.3

Kinder

keine Kinder	16.3	15.8	16.8	18.2 %
1-≤2 Kinder	38.3	41.0	42.9	38.9 %
≥3 Kinder	38.0	35.3	33.8	34.1 %
unbekannt	7.4	7.9	6.4	8.7 %

Haushaltssituation

Einpersonenhaushalt	52.2	50.0	49.2	65.2 %
Mehrpersonenhaushalt	47.8	50.0	50.8	34.8 %

Anzahl Zimmer pro Person

0 - ≤1.5 Zimmer	23.7	24.7	26.1	22.1 %
>1.5 - ≤2.5 Zimmer	32.9	35.6	34.0	31.1 %
>2.5 Zimmer	43.4	39.7	39.9	46.8 %

Kulturell-kontextuelle Faktoren

Nationalität

Schweiz	95.5	94.8	93.6	96.9 %
Ausland	4.5	5.2	6.4	3.1 %

Sprachregion

Deutschschweiz	76.6	71.3	67.3	73.1 %
Westschweiz	19.9	23.5	27.5	22.3 %
Tessin	3.6	5.2	5.2	4.6 %

Heimbettdichte*: Mittel-

wert	6.9	6.8	6.6	7.0
------	-----	-----	-----	-----

Datenquelle: MEDSTAT/SOMED/TU/SNC 2007/2008. Basis: Eigene Berechnungen.

*Durchschnittliche Anzahl verfügbarer Betten in Alters- und Pflegeheimen pro 100 mindestens 65 Jahre alten Bewohnern nach Region (Total 106 Regionen)

4.3 Methoden

Für die empirischen Analysen werden die drei abhängigen Variablen Heimaufenthalte, Spitalaufenthalte sowie unterschiedliche Sterbeorte in eigenen Modellen überprüft. Eine empirische Analyse muss dabei jeweils die individuellen, familialen sowie kulturell-kontextuellen Determinanten erfassen und deren Einfluss unter Kontrolle aller anderen im Modell enthaltenen Variablen berechnen. Die beiden Variablen Heim- und Spitalaufenthaltsdauer weisen ein metrisches Skalenniveau auf und beinhalten nur positive ganze Zahlen. Demgegenüber be-

steht die Variable mit den unterschiedlichen Sterbeorten aus vier Ausprägungen, wodurch die Variable als nominalskaliert betrachtet wird (vgl. Diekmann 2005).

Sämtliche Berechnungen werden in STATA, Version 12, durchgeführt (Stata-Corp LP, College Station, Texas, 2011). Zur deskriptiven Analyse werden je nach Skalenniveau Durchschnitte bzw. Anteile berechnet. Als Mass für die Stärke des Zusammenhangs zweier stetiger Variablen werden bei metrisch skalierten Variablen der Korrelationskoeffizient nach Pearson und bei kategorialen Variablen χ^2 -Tests berechnet.

Heim- und Spitalaufenthalte

Bei der Analyse von Heim- und Spitalaufenthaltsdauern muss besonders die extrem rechtsschiefe Verteilung berücksichtigt werden. Sowohl bei den Spital- als auch bei den Heimaufenthalten weist etwa der Unterschied zwischen Median und arithmetischem Mittelwert darauf hin, dass viele Aufenthalte nur kurz sind und es verhältnismässig wenige lange Aufenthalte gibt. Die Null kommt in den Daten verhältnismässig häufig vor. Bei Heim- und Spitaldauern sind dies Personen, die gleich am Eintrittstag in einem Heim oder Spital verstorben sind. Um diese Daten optimal zu analysieren gibt es zwei Möglichkeiten. Man kann entweder multiple lineare Regressionsmodelle oder multiple Regressionsmodelle für Zähldaten rechnen. Wie eine Studie anhand von Testberechnungen herausfand, liefert eine lineare Regression insbesondere bei grossen Stichproben erstaunlich valide Ergebnisse, auch wenn die Verteilung der abhängigen Variable stark von der Normalverteilung abweicht (vgl. Lumley et al. 2002). Ein lineares Modell basiert aber, wenn es für Zähldaten verwendet wird, auf falschen Annahmen: Zähldaten sind nur positive ganze Zahlen, wohingegen normalverteilte Daten stetig sind, also jeden beliebigen Zwischenwert annehmen können. Letztlich ist deshalb ein für Zähldaten optimiertes Regressionsmodell besser geeignet.

Das am häufigsten verwendete Zähldatenmodell ist die Poisson-Regression, welches mittlerweile auch in einigen Studien angewendet wurde (vgl. Berlin et al. 2014; Coxe et al. 2009; Evans et al. 2014; Vittinghoff et al. 2005). Bei der Poisson-Regression wird als abhängige Variable die logarithmierte Rate der interessierenden Grösse y_i als lineare Funktion der unabhängigen Variablen x_i dargestellt:

$$(1) \quad \ln(y_i) = \sum_{j=1}^k \beta_j x_{ji}$$

Dabei steht j für die jeweilige unabhängige Variable, und i repräsentiert die jeweilige Person in den Daten. Die geschätzten Regressionskoeffizienten β_j können

interpretiert werden als die additive Veränderung des logarithmierten Erwartungswerts bei einer Änderung der unabhängigen Variable um eine Einheit. Das Verständnis wird erleichtert, wenn man die Exponentialfunktion auf die geschätzten Koeffizienten anwendet, weil sie dann als multiplikative Veränderung des Erwartungswerts selbst interpretiert werden kann.

Bei Anwendung der Poisson-Regression wird vorausgesetzt, dass die Varianz und das arithmetische Mittel der abhängigen Variable etwa gleich gross sind. Wird diese Annahme verletzt, so können die Standardfehler unterschätzt werden, was zu mehr signifikanten Prädiktoren führen würde, die tatsächlich gar nicht signifikant sind. Diese sogenannte Überdispersion ist in der Praxis häufig, und auch im vorliegenden Fall liegt eine deutliche Überdispersion sowohl bei den Heim- als auch bei den Spitaltagen vor. Um dieses Problem zu lösen, ist ein negativ-binomiales Regressionsmodell gut geeignet, weil es auf einerseits auf der Poisson-Regression beruht und andererseits das Problem der Überdispersion in Poisson-Modellen löst (Hilbe 2007).²⁹ Vereinfacht dargestellt basiert dieses Regressionsmodell auf einer so genannten negativen Binomial-Verteilung anstatt der Poisson-Verteilung. Dabei handelt es sich um eine Verteilung, die eine deutlich grössere Varianz hat, so dass diese Eigenschaft der Daten mit modelliert werden kann.

Bei negativ-binomialen Regressionsmodellen sind besonders drei Aspekte zu berücksichtigen. Als erstes ist die Methode - auch wenn sie in Statistiklehrbüchern (vgl. Hilbe 2007) als Vorgehensweise für Daten mit Überdispersion empfohlen wird - in der epidemiologischen und sozialwissenschaftlichen Forschungspraxis kaum verbreitet, was das Verständnis und die Interpretation der Modelle erschweren kann. Zweitens können, wie oben bereits dargestellt, die geschätzten Regressionskoeffizienten kaum direkt interpretiert werden, sondern müssen zunächst mit der Exponentialfunktion transformiert werden. In dieser Arbeit werden stets die bereits transformierten Koeffizienten, hier „Incidence Risk Ratios“ (IRR) genannt, angegeben, die als multiplikative Veränderung der jeweiligen Rate zu interpretieren sind. Zudem werden die 95%-Konfidenzintervalle angegeben, weil diese direkt als mögliche Werte dieser Faktoren verstanden werden können, wohingegen andere Angaben wie z-Werte oder Standardfehler deutlich schwieriger zu interpretieren sind.

²⁹ Für weitergehende Informationen über negativ-binomiale Regressionsmodelle wird empfohlen, das gut lesbare Lehrbuch von Hilbe (2007) zu studieren. Eine Kurzübersicht über die Methode und deren Anwendung in STATA liefert auch die Homepage des Institutes für digitale Forschung und Ausbildung der Universität Los Angeles: <http://www.ats.ucla.edu/stat/stata/dae/nbreg.htm> [Zugriff: 6.2.2015].

Als letztes ist noch die Modellwahl zu beachten. Grundsätzlich wurden in dieser Arbeit die Variablen aus rein inhaltlichen Erwägungen in die Modelle aufgenommen. In einigen Fällen, wo unklar war, ob eine Variable tatsächlich in ein Modell aufgenommen werden soll, wurden das AIC (Akaike-Informationskriterium) und BIC (Bayesianisches Informationskriterium) berechnet, wie im Lehrbuch von Hilbe (2007) vorgeschlagen. Diese Informationskriterien kombinieren ein Mass für die Güte der Modellanpassung mit einem Bestrafungsterm für Modellkomplexität, wobei diese Strafe beim BIC stärker gewichtet wird, wodurch das BIC in der Regel weniger komplexe Modelle bevorzugt. Häufig sind die Ergebnisse von AIC und BIC aber vergleichbar. Es ist zu beachten, dass ein kleinerer Wert dieser Informationskriterien das jeweils bessere Modell auszeichnet. Zur Auswahl des besten Modells wurden in dieser Arbeit die Modelle schrittweise gerechnet und bezüglich AIC und BIC verglichen, etwa indem temporär ausgewählte Interaktionsterme ins Modell integriert wurden. Dabei wurde überprüft, ob sich das Modell im Vergleich zum Modell ohne diese Terme verbessert hat. Letztlich wurden die inhaltlich und statistisch optimalsten Modelle ausgewählt. Weil in den Modellen viele kategoriale Variablen enthalten sind, wurden zusätzlich Likelihood-Quotienten Tests durchgeführt, die einen gemeinsamen p-Wert für einzelne unabhängige Variable berechnen, wodurch Variablen als Ganzes eingeschätzt werden können.

Unterschiedliche Sterbeorte

Bei den Sterbeortanalysen liegt eine abhängige Variable mit nominalem Skalenniveau und vier Ausprägungen vor. In der Praxis werden dabei meist die Kategorien so zusammengelegt, dass letztlich eine binäre Variable in einem logistischen Regressionsmodell analysiert werden kann. Für die Sterbeortanalysen ist dies aber nicht möglich, denn erstens können die Ausprägungen nicht zusammengelegt werden und zweitens verliert man dadurch immer auch viel Informationsgehalt. Deshalb werden die vier Kategorien beibehalten und in einem multinomialen Regressionsmodell analysiert. Dieses Verfahren stellt ein erweitertes logistisches Regressionsmodell dar, bei dem die abhängige Variable mehr als zwei Ausprägungen haben kann. Dabei wird eine Ausprägung als Referenzkategorie festgelegt, und die Resultate der anderen Ausprägungen zeigen dann die jeweiligen Unterschiede zur Referenzkategorie auf. Die Formel des Modells sieht folgendermassen aus (j =Anzahl Ausprägungen der abhängigen Variable):

$$\text{pr}(y_i = 1 | \mathbf{x}_i) = \frac{1}{1 + \sum_{j=2}^J \exp(x_{ij}\beta_j)} \quad \text{für } m=1$$

$$(2) \quad \text{pr}(y_i = m | \mathbf{x}_i) = \frac{\exp(x_{im}\beta_m)}{1 + \sum_{j=2}^J \exp(x_{ij}\beta_j)} \quad \text{für } m > 1$$

Das Hauptproblem von multinomialen Logit-Modellen besteht in der Interpretation der Koeffizienten, weil beispielsweise die Vorzeichen der geschätzten Regressionskoeffizienten schwierig direkt zu interpretieren sind (Kohler & Kreuter 2008:300 ff.; Zimmer & House 2003).³⁰ Diesem Problem wird mit zwei Lösungsansätzen begegnet. Erstens werden anstatt β -Koeffizienten Relative Risk Ratios (RRR) angegeben, die sehr ähnlich wie die von der logistischen Regression bekannten Odds Ratios interpretiert werden können (vgl. Hosmer & Lemeshow 2000).³¹ Daneben werden wiederum die 95%-Konfidenzintervalle angegeben, weil auch bei multinomialen Logit-Modellen z-Werte oder Standardfehler für Nicht-Statistiker schwer zu interpretieren sind. Um das Verständnis zu erleichtern werden ausserdem für einige der besonders wichtigen unabhängigen Variablen wie Bildung, Haus- oder Wohnungsbesitz, Zivilstand und Sprachregion grafische Wahrscheinlichkeitsplots gezeigt. Dadurch erhält man eine sehr präzise Analyse der sich ändernden Wahrscheinlichkeit einer Ausprägung unter gleichzeitiger Berücksichtigung der jeweils anderen drei Ausprägungen. Ein weiteres Problem von multinomialen logistischen Regressionsmodellen ist, ähnlich wie bei den negativ-binomialen Modellen, ihre ebenfalls geringe Verbreitung in der sozialwissenschaftlichen Forschungspraxis sowie insbesondere die uneinheitliche Art und Weise, wie dem Interpretationsproblem von Koeffizienten begegnet wird.

Zur Modellwahl wurde wie bei den Heim- und Spitalanalysen auf möglichst geringe AIC und BIC Werte innerhalb des Modells geachtet, und schlussendlich wurden die sowohl aus statistischer als auch aus inhaltlicher Sicht besten Modelle ausgewählt. Wiederum wurden Likelihood-Quotienten Tests berechnet.

³⁰ Es gibt in den wissenschaftlichen Journals keinen gemeinsamen Nenner bezüglich Interpretation von β -Koeffizienten in multinomialen Logit-Modellen. In einigen Studien werden Koeffizienten direkt interpretiert (vgl. Reich et al. 2013); andere Studien weisen auf das Problem der Interpretation hin und berechnen grafische Hilfsmittel (vgl. Zimmer & House 2003). Letztlich ist eine zusätzliche grafische Darstellung in jedem Fall hilfreich. Aus pragmatischen Gründen wurde die Auswahl auf einige besonders wichtige unabhängigen Variablen beschränkt.

³¹ Eine gute und praxisbezogene Kurzübersicht über multinomiale Regressionsmodelle im Allgemeinen sowie spezifisch über Umrechnungen von β -Koeffizienten in Relative Risk Ratios (RRR) in STATA liefert die Homepage des Institutes für digitale Forschung und Ausbildung der Universität Los Angeles: <http://www.ats.ucla.edu/stat/stata/dae/mlogit.htm> [Zugriff: 6.2.2015].