Der TMI auf dem Prüfstand: Güte und Probleme der Kostenvorhersage anhand einfach kodierbarer Patientenmorbidität

Joachim E. Fischer[a,b], *Michael Peltenburg*[c,d], *René Kühne*[e], *Corinne Achermann*[e] *und die zmed-Studiengruppe*[c]

a Institut für Public Health, Medizinische Fakultät Mannheim, Universität Heidelberg, Deutschland

b HealthVision GmbH, Zug

c zmed Ärztegesellschaft, Zürich

d Einheit für Hausarztmedizin, Universität Zürich

e Helsana Versicherungen AG, Abt. Versorgungsforschung, Zürich

Interessenverbindungen
Keiner der Autoren deklariert finanzielle Interessen an den eingesetzten Instrumenten. J. Fischer war zum Zeitpunkt der Untersuchung Geschäftsführer der HealthVision GmbH. J. Fischer ist Gesellschafter der HealthVision GmbH. Die Untersuchung wurde finanziert von der Helsana Versicherungen AG.

Korrespondenz:
Prof. Dr. med. Joachim E. Fischer
Institut für Public Health,
Sozial- und Präventivmedizin
Ludolf-Krehl-Strasse 7–11
D-68167 Mannheim
Tel. +49 621 383 9910
Fax +49 621 383 9920

jfischer@medma.uni-heidelberg.de

Einleitung

Versicherer, Ärzte und Versicherte erhoffen sich von neuen Versicherungsmodellen eine verbesserte Versorgung bei gleichbleibenden oder geringeren Kosten. Ziel dieser Versicherungsmodelle ist, finanzielle Anreize für eine optimierte Versorgung zu schaffen. Bei den bekannteren Hausarztsystemen ist der versorgende Primärarzt die steuernde Schlüsselperson respektive der «Gatekeeper» zu kostenträchtigen Spezialuntersuchungen und -behandlungen. Ärztenetzwerke bieten eine Chance, innerhalb des Netzwerks Behandlungsabläufe so zu optimieren, dass Ressourcen bestmöglich zum Wohl der Patienten eingesetzt werden, ohne jeweils das maximal Machbare oder Vergütbare auszuschöpfen. Dazu bedarf es finanzieller Anreizsysteme innerhalb der bestehenden gesetzlichen Rahmenbedingungen, die Ärzte für optimierte Behandlung und nicht für maximale Leistungsausweitung honorieren. Mögliche derartige Anreizsysteme sind Versicherungsmodelle auf Budgetbasis, d. h. mit ärztlicher *Budgetmitverantwortung*. Die amerikanischen Erfahrungen zeigen indes, dass Anbieter derartiger Versicherungspläne eher gesündere und kostengünstigere Patienten aufnehmen, als dass sie die Behandlungsabläufe der Patienten optimieren [1]. Versicherungsmodelle auf Budgetbasis können unter diesen Bedingungen sogar höhere Kosten als die klassische Abrechnung aufgrund erbrachter Leistungen verursachen [2, 3].

Das zentrale Steuerungsinstrument von *budgetbasierten* Versicherungsmodellen ist daher die korrekte Vorhersage der zu erwartenden Kosten. Das Vorhersagemodell sollte optimierte Versorgung «belohnen» [4, 5]. Ein erweitertes Vorhersagemodell der Duke University berücksichtigt nicht nur die Diagnosen, sondern auch den Schweregrad der Erkrankung, die gesundheitsbezogene Lebensqualität der Patienten sowie zusätzliche Indikatoren. Diese umfassen Lebensumstände wie etwa Schulbildung, soziale Unterstützung und Gesundheitsverhalten. Das Modell erzielt in einem städtischen Umfeld eine Varianzaufklärung von fast 30% [6, 7]. Der Kodierauf-

Indice de morbidité thurgovien sur la sellette: fiabilité et problèmes du pronostic des coûts mis en évidence par un système de codage simple

Les systèmes de rémunération pourvus d'incitations financières visant une optimisation de la couverture en soins doivent être dotés d'outils simples permettant d'équilibrer les paramètres de morbidité. L'indice de morbidité thurgovien (IMT), validé pour les réseaux de médecins dans les régions rurales, constitue un tel outil. Sur la base d'un échantillon de 1903 patients provenant de 27 cabinets médicaux de la ville de Zurich, la présente étude analyse la valeur de pronostic de l'IMT dans un environnement urbain avec une proportion élevée de patients présentant des problèmes sociaux. Les auteurs comparent l'IMT avec un indicateur de morbidité répandu au niveau international (l'indice de gravité DUSOI) et avec un codage des cas en 51 catégories de diagnostic basé sur la classification internationale des soins primaires (CISP-2). Tant l'IMT (r = 0,54) que l'indice DUSOI (r = 0,51), aisé à établir, sont en bonne corrélation avec les coûts intervenus au cours d'une même année. Les deux indicateurs améliorent le pronostic des coûts par rapport au modèle fondé sur l'âge, le sexe et la franchise choisie. L'analyse fournit en outre des pistes en vue de compléter l'IMT et de pouvoir en recommander l'usage dans les modèles de rémunération: les compléments proposés consistent en un remaniement des définitions et en une intégration des catégories de diagnostic CISP-2 significatives en termes de coûts.

wand für sehr gute Vorhersagemodelle ist indes in der Regel hoch und in der Praxis nur schwer zu leisten. Eine Arbeitsgruppe der Duke University entwickelte daher einen vereinfachten Morbiditätsindikator, den DUSOI-Score, der die Symptome und Komplikationen der vergangenen Woche sowie die Behandelbarkeit und die Prognose der Erkrankung berücksichtigt [8]. Einen anderen Ansatz wählte eine Arbeitsgruppe von Ärztinnen und Ärzten aus dem Thurgau, die vor sieben Jahren den Thurgauer Morbiditätsindikator (TMI) entwickelten [9].

Der TMI berücksichtigt einerseits die Anzahl der akuten Konsultationsgründe und andererseits den Schweregrad aller chronischen Erkrankungen in einem siebenstufigen System von 0 = «gesund» bis 6 = «schwerst polymorbid». Bei einer kürzlich publizierten Studie mit Versicherten eines Ärztenetzwerkes im Raum Winterthur/Andelfingen und Daten aus 46 Arztpraxen wurde der TMI prospektiv an 1925 Patienten mit chronischen Erkrankungen erhoben. Dieses Vorhersagemodell aus TMI, Alter und Geschlecht sagt 32 % der Varianz der im gleichen Jahr entstehenden Nettokosten vorher, im Gegensatz zu 18 % aus Alter und Geschlecht allein [10]. Bisher fehlen indes Arbeiten, die den TMI hinsichtlich seiner Vorhersagegüte zu anderen Morbiditätsindikatoren oder Diagnosegruppen in Bezug setzen und die den TMI hinsichtlich der Trennschärfe und Reliabilität seiner einzelnen Definitionen überprüfbar machen. Ferner fehlen Daten, ob der TMI auch in einem Umfeld städtischer Patienten mit höherem Anteil sozialer Randgruppen anwendbar ist.

Das Ziel der vorliegenden Studie ist, prospektiv in einem städtischen Ärztenetzwerk den TMI parallel zu Diagnosegruppen nach dem ICPC-2 und anderen Morbiditätsindikatoren zur Kostenvorhersage einzusetzen. In dieser Arbeit berichten wir über das Studiendesign sowie über die Ergebnisse der Kostenvorhersage im gleichen Jahr wie der Erhebung der Prognosefaktoren.

Methoden

Patienten und teilnehmende Praxen

Die Studie wurde als prospektive Kohortenstudie vom 15. Mai bis zum 31. Dezember 2005 in 30 der zmed angeschlossen Praxen mit Grundversorgungscharakter durchgeführt. Teilnahmeberechtigt waren alle Patienten der Helsana-Gruppe*, die im genannten Studienzeitraum Kontakt mit einer der teilnehmenden Arztpraxen hatten, eingeschlossen waren auch Kontakte im Rahmen des ärztlichen Notfalldienstes oder von Notfallkonsultationen. Ausgeschlossen waren Asylsuchende** sowie Kinder unter 16 Jahren. Alle Teilnehmer stimmten der Untersuchung schriftlich zu. Die Studie wurde von der zuständigen Subkommission der Kantonalen Ethikkommission genehmigt. Im Verlauf schieden drei Praxen aus der Studie aus (ein Praxisarzt hatte grundsätzliche Einwände gegen das Studiendesign, zwei Praxen konnten wegen Engpässen bei der Praxisassistenz die Rekrutierung nicht bewältigen).

Studienablauf und erhobene Variablen

Patientendaten und mögliche Prognosefaktoren wurden mittels Arzt- und Patientenfragebogen erhoben. Die Arztfragebögen wurden einmal je Abrechnungsperiode ausgefüllt, die Patientenfragebögen einmal oder nach besonderen Veränderungen (z. B. Neudiagnose Tumorleiden).

Der Patientenfragebogen enthält den SF12-2 zur Erfassung der gesundheitsbezogenen Lebensqualität [11], ferner Angaben zu Familienstand, Sprache, Schulbildung, Lebensumständen, sozialer Unterstützung bei gesundheitlichen Problemen, Berufstätigkeit, körperlicher Aktivität, besonderen Lebensereignissen, Versicherungsmodus, Arztpraxis sowie zu psychosozialem Stress. Letztere Items wurden aus der Interhart-Studie zur Evaluation von Risikofaktoren bei koronarer Herzkrankheit abgeleitet [12].

Kernstück des Arztfragebogens ist eine aus dem ICPC-2 abgeleitete Liste mit 51 Diagnosen, zu der die Ärzte den Konsultationsgrund (akute Erkrankung oder Follow-up) und den Grad der chronischen Erkrankung nach dem Thurgauer Morbiditätsindikator erfassten. Folgende Kategorien waren vorgesehen: a) Status nach Ereignis bzw. nach TMI-Definition derzeit gesund, b) präklinische chronische Erkrankung, c) leichte bis mittelschwere chronische Erkrankung und d) schwere chronische Erkrankung [9]. Zur Verbesserung der Datengüte enthält das Datenblatt auf einer Ausklappseite die TMI-Definitionen und Fallbeispiele zu möglichen Kodierungsproblemen. Zur Minimierung möglicher Kodierfehler wurde der TMI-Gesamtscore sowohl für die akuten als auch die chronischen Erkrankungen nach dem TMI-Algorithmus aus den Arztangaben zu den einzelnen Diagnosegruppen errechnet (Abb. 1) [10].

Zusätzlich enthält das Kodierblatt den DUSOI-Morbiditätsscore, bezogen auf die jeweilige ICPC-2-Diagnosegruppe [8]. Der DUSOI-Score beurteilt nach einer fünfstufigen Likert-Skala (0–4) eine Erkrankung nach der Schwere der Symptome in der vergangenen Woche, der Komplikationen in der vergangenen Woche, der Prognose ohne Behandlung und der Behandlungs-

* Helsana, Progrès, Sansan und Avanex.

** Die Grundversicherung für Asylsuchende ist in einem speziellen, mit dem Kanton vereinbarten Gatekeepingvertrag geregelt.

bedürftigkeit bzw. dem zu erwartenden Behandlungserfolg. Der DUSOI-Gesamtscore (0–16) wurde ebenfalls vom Auswertungsprogramm errechnet. Der Arztfragebogen enthält ferner Items zu Konsultationszeitpunkt, Konsultationsanlass (Notfall, geplante Behandlung), bekannten sozialen Problemen, Gründen für das Nichtausfüllen des Patientenbogens und Fragen zum sozialen Umfeld.

Alle teilnehmenden Ärzte wurden zu Beginn der Studie während eines Nachmittages geschult. Im Verlauf der Studie erhielten sie zu zwei Terminen insgesamt 25 Fallvignetten (10 + 15) für eine schriftliche Kodierung. Aufgrund dieser Fallvignetten wurde die Reliabilität zwischen verschiedenen Untersuchern ermittelt. Anhand der eingesandten Teilnehmerabschnitte ermittelte die Versicherung die Kostendaten der Versicherten für die Jahre 2003 bis 2005, aufgeschlüsselt nach Kosten für ärztliche Leistungen, Leistungen von Apotheken, Leistungen von unterstützenden Heilberufen und Pflege sowie Leistungen für Hospitalisationen.

Statistische Analyse

Einfache Skalenunterschiede zwischen einzelnen Teilnehmergruppen werden mittels nichtparametrischer Verfahren (Wilcoxon-Test), kategorische Variablen mittels Chi-Quadrat-Test oder Fishers exaktem Test verglichen. Für einfache Korrelationsmodelle wird der nichtparametrische Korrelationskoeffizient (Spearmans Rho) errechnet. Für Vorhersagemodelle mit mehreren Variablen ist es notwendig, die stark schief verteilten Daten (z. B. Kosten mit vielen niedrigen Werten und wenig hohen Werten) durch Logarithmierung der Normalverteilung anzunähern. Da der Wert 0 nicht logarithmiert werden kann, wird für Personen ohne Inanspruchnahme von Leistungen ein Betrag von Fr. 50.– als Jahreskosten eingesetzt. Andere Variablen mit nichtnormaler Verteilung (insbesondere TMI) werden durch Dummy-Kodierung in kategorische Daten überführt. Durch die Dummy-Kodierung errechnet das statistische Modell für jede einzelne Stufe des TMI separate Koeffizienten. Nur so lässt sich im Rechenmodell überprüfen, ob die Kostensteigerung etwa von TMI 2 nach TMI 3 gleich gross ist wie die Steigerung von TMI 4 nach TMI 5. Auf diese Weise lassen sich auch fehlende Unterschiede zwischen TMI-Kategorien identifizieren (TMI 5 vs. TMI 6).

Für die statistische Auswertung in dieser Arbeit wird als Referenz ein Modell aus Alter, Geschlecht, Kosten des Jahres 2004 (Vorjahreskosten) sowie Franchise gewählt. Dieses Referenzmodell wird verglichen mit fünf anderen Modellen, deren Variablen in Tabelle 1 aufgeführt sind. Alle Analysen werden in SAS (Version 8.2) oder SPSS (Version 14) durchgeführt. Eine Irrtumswahrscheinlichkeit von weniger als 5% ($p < 0{,}05$) gilt als Nachweis einer statistischen Signifikanz.

Ergebnisse

Patienten

Zwischen Mai und Dezember 2005 rekrutierten 27 teilnehmende Praxen insgesamt 1967 Patienten (Durchschnittsalter 52,6 Jahre, 58% Frauen). Im Mittel wurden je Praxis 73 Patienten rekrutiert, aus 6 Praxen kamen jeweils weniger als 40 Teilnehmer, 6 Praxen trugen mehr als 100 Patienten zum Datensatz bei. Durch Patienten mit

Abbildung 1
Auswertungsalgorithmus für den Thurgauer Morbiditätsindikator.

TMI-Kategorie	Bezeichnung	Anzahl präklinischer Erkrankungen	Anzahl leichter bis mittelschwerer chronischer Erkrankungen	Anzahl schwerer chronischer Erkrankungen
0	gesund	0	0	0
1	präklinisch	Ja	0	0
2	chronisch leicht a		1 bis 2	0
3	chronisch leicht b		3 und mehr	0
4	chronisch schwer a		0 bis 2	1
5	chronisch schwer b		3 und mehr	1
6	schwerst polymorbid			2 und mehr

Als chronische Erkrankung gelten im Prinzip alle Erkrankungen von mehr als einem Jahr Dauer sowie gewisse als chronische Erkrankung definierte Zustände (z. B. BMI >30). Als schwere chronische Erkrankungen gelten chronische Erkrankungen mit Dekompensation, Instabilität, schwerem Schub, rascher Progredienz, schwer eingeschränkter Organfunktion oder mit schwerer sozialer Beeinträchtigung. Ferner gelten als schwere chronische Erkrankungen maligne Tumorerkrankungen unter Behandlung sowie medizinisch als schwer definierte Systemerkrankungen. Als «präklinisch» gelten Zustände, die Anlass zur ärztlichen Überwachung sind, aber noch nicht als Erkrankung diagnostiziert wurden (z. B. ärztlich angezeigte Kontrolle des Blutdrucks ohne Diagnose Hypertonie).

Konsultationen in mehreren Abrechnungsperioden ergaben sich insgesamt 2880 ausgefüllte Arztbögen und 2203 Patientenbögen. Vollständige Datensätze aus Arztbogen und Versicherungsdaten liegen für 1903 Patienten vor. Die häufigsten Diagnosen sind mit je über 200 Fällen Diabetes Typ II, Angst und Panikstörungen, chronischer Substanzmissbrauch, andere neurologischen Erkrankungen, Beschwerden des Rückens, Herz-Kreislauf-Symptomatik, Bluthochdruck, Erkrankungen des Verdauungssystems sowie HIV-Infektionen. Tabelle 2 gibt eine Übersicht über weitere Patientendaten.

Morbiditätsindikatoren

Bei 1593 Patienten (82%) wurden insgesamt 3919 leichte bis mittelschwere sowie 516 schwere chronische Erkrankungen kodiert (2,85 pro Patient). 374 Patienten (18%) hatten keine chronischen Erkrankungen, sondern konsultierten die Praxen allein aufgrund akuter Erkrankungen oder zur Vorsorge. Damit sind in diesem Patientenkollektiv gegenüber der versicherten Bevölkerung und im Vergleich zu den früheren TMI-Studien die Patienten mit mittelschweren und schweren chronischen Erkrankungen stärker repräsentiert, insbesondere auch Patienten mit chronischem Substanzmissbrauch und mit HIV. Der TMI korreliert hoch mit dem DUSOI-Score als externem Morbiditätsindikator (Spearman Rho $r = 0{,}75$, $p < 0{,}001$).

Die Übereinstimmung der Kodierung der Fallvignetten mit der als «Goldstandard» angesehen Beurteilung unterscheidet sich zwischen den 25 von den Ärzten bearbeiteten Fällen stark. Die Ärzte stimmen im Durchschnitt zu 63% mit der Musterkodierung überein, bei einfach zu kodierenden Fällen bis zu 90%. Beim «schwierigsten Fall» liegt die exakte Übereinstimmung nur bei 14,5%, wird eine Abweichung um maximal eine Kategorie akzeptiert, liegt die «Trefferquote» bei 29%. Die häufigsten und bedeutsamsten Abweichungen von der richtigen Lösung ergeben sich bei den Fallvignetten zwischen den TMI-Kategorien 3 und 5.

Der TMI zeigt bei den 491 Patienten mit mindestens zwei Arztbögen eine gute Re-Test-Reliabilität (Intraclass-Korrelations-Koeffizient = 0,72). Abbildung 2 zeigt die Verteilung der höchsten Kategorie des Thurgauer Morbiditätsindikators sowohl hinsichtlich der akuten Erkrankungen als auch der chronischen Erkrankungen, ferner die mittleren Bruttokosten je TMI-Kategorie.

Kosten und Kostenentwicklung

Tabelle 4 schlüsselt die Kosten auf. Der überwiegende Anteil der Kosten entfällt auf Hospitalisationen (37%), Arztkonsultationen (27%) und Vergütungen für Medikamente (20%). Von 2004 (Mittel Fr. 7176.–) auf 2005 (Mittel Fr. 7726.–) steigen die Kosten um 7,6%. Hier überlagern sich zwei Effekte: steigende Kosten durch Voranschreiten chronischer Erkrankungen und die allgemeine Kostenzunahme im Gesundheitswesen.

Tabelle 1
Multivariable Rechenmodelle zur Vorhersage der Kosten 2005.

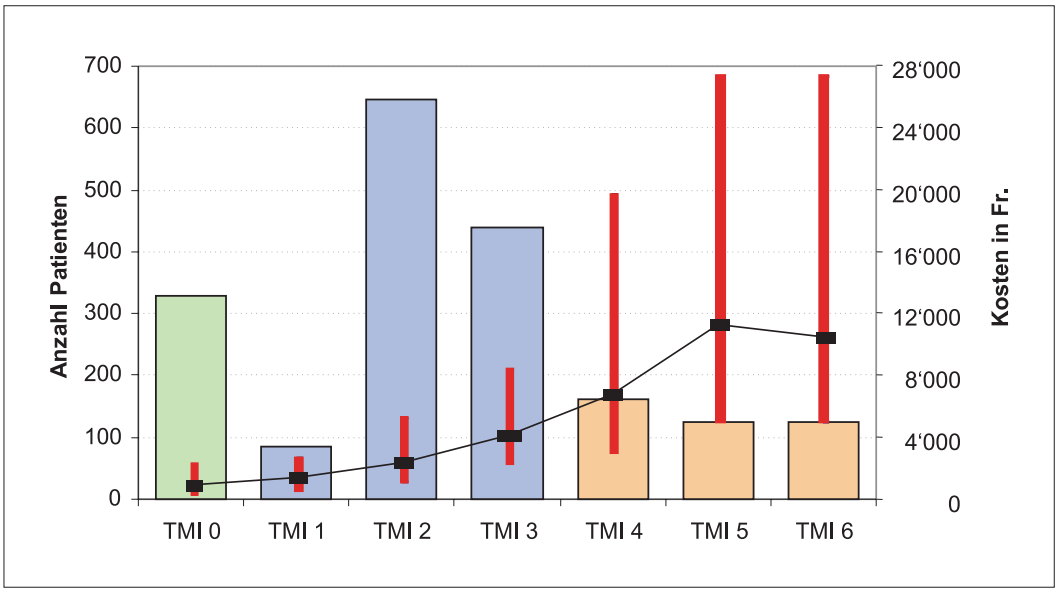
	Referenz	Modell 1	Modell 2	Modell 3	Modell 4	Modell 5	Modell 6
Alter und Geschlecht							
Kosten 2004							
Gewählte Franchise							
TMI							
DUSOI-Score							
Diagnosegruppen ICPC-2							
Soziale Daten Patientenbogen*							
Soziale Daten Arztbogen*							
Lebensqualität SF-12							
Stress und Lebensereignisse							
Aufgeklärte Varianz (R^2)	0,48	0,24	0,37	0,43	0,45	0,42**	0,59**
Vergleich zu Referenzmodell*	–	50%	77%	90%	94%	88%	123%

Gelb markiert sind die Prognosefaktoren. Bei den sozialen Ereignissen wurden jeweils die drei am meisten zur Varianzaufklärung beitragenden Prognosefaktoren eingesetzt.

* Vergleich zu Referenzmodell = Aufgeklärte Varianz/0,48; ausgedrückt in Prozent.

** Vollständige Daten nur für n = 758 Patienten, für diese Patienten gilt für Modell 2 $R^2 = 0{,}36$.

Abbildung 2
Verteilung der Kategorien des Thurgauer Morbiditätsindikators (n = 1967).



Die Säulen zeigen die Anzahl Patienten. Die vertikalen Balken zeigen bezogen auf die rechte Skala den Median der Kosten der jeweiligen TMI-Kategorie an sowie als Mass der Streuung den Bereich von der 25. bis zur 75. Perzentile. Die Kosten unterscheiden sich zwischen TMI 5 und TMI 6 nicht (p = 0,83).

Tabelle 2
Patienten (n = 1967).

	Anzahl*	Wert oder Prozent
Versicherungsdaten		
Alter (Jahre, ± Standardabweichung)	1967	52,6 ±20,4
Geschlecht (n, % Frauen)	1967	58%
Franchise (bis 300/500/1000/1500/über 2000)	1967	60/22/3/9/6%
Arztbogen		
TMI	1903	2,5 ±1,7
DUSOI-Score	1903	5,7 ±4,5
Arzt schätzt Kosten >Fr. 20 000 (%)	1903	174 (9,1%)
Schulbildung (n, Grundschule/Sek./Matur/Studium)	810	23/52/10/15%
Lebensumstände (n, allein/Familie/Institution)	1066	37/57/6%
Lebensunterhalt (n, Einkommen/Rente/IV/Arbeitslos/Sozialhilfe)	1083	45/38/7/4/6%
Problem Armut (n, unbekannt/keine/leicht/mittel/schwer)	1361	9/65/14/8/3%
Soziale Integration (n, sehr gut/mittel/gering)	1416	67/26/7%
Arztbogen: Compliance (n, sehr gut/mittel/gering)	1421	67/28/5%
Patientenbogen		
SF-12 körperliche Skala (n, Mittelwert, ±Standardabweichung)**	956	47,1 ±10,7
SF-12 psychische Skala (n, Mittelwert, ±Standardabweichung)**	956	48,3 ±11,3
Sport pro Woche (n, keinen/tägl. spazieren/bis 1 h/bis 2 h)	966	36/29/16/18%
Sprachbeherrschung Deutsch (n, Muttersprache oder sehr gut/gut/ nur mit Mühe/kaum)	1222	90,6/7,7/1,5/0,2%
Vertrauen in den Arzt (n, sehr gut/gut/mässig/etwas/kennt Probleme kaum)	1186	50/37/6/3,4/3,5%

* Anzahl gibt für kontinuierliche Daten (Alter, Skalenwerte) die Anzahl von vollständigen Datensätzen für diese Variable an, für kategorische Daten die absolute Anzahl Patienten.

** Der SF-12 ist auf einen Mittelwert von 50 und eine Standardabweichung von 10 bei gesunden Personen im Alter von 30 bis 50 geeicht.

Tabelle 3
Einfache Korrelationsmatrix.

Volltext	log BVL 2005	Nicht–Spital 2005	log BVL 2004	Alter	Geschlecht	Franchise	TMI	DUSOI	Kosten > 20	Armut	Soziokult Probl.	HIV/AIDS	Hypertension	Substanzmissb.	Soziale Integr.	Compliance	Einkommen	SF–12 körp.	SF–12 psych.	Dauer Sport	Sprachbeherr.
Bruttoversicherungsleistung 2005																					
Summe Kosten 2005 ohne Spital	**0.92**																				
Summe Kosten BVL Vorjahr	**0.68**	**0.70**																			
Alterskategorie	**0.37**	**0.38**	**0.33**																		
Geschlecht (0 = M, 1 = W)	0.09	0.10	0.12	0.10																	
Franchise	**–0.29**	**–0.31**	**–0.28**	**–0.21**	–0.07																
TMI	**0.54**	**0.54**	**0.46**	**0.40**		**–0.29**															
DUSOI Score	**0.51**	**0.51**	**0.40**	**0.32**		**–0.25**	**0.75**														
Arzt schätzt Kosten >Fr. 20000	**0.40**	**0.35**	**0.28**	0.11	–0.06	–0.11	**0.34**	**0.30**													
Soziale Probleme Armut	0.06		0.05	–0.11		–0.09	0.16	0.14													
Soziokulturelle Probleme	0.06		0.06	–0.12		–0.11	0.14	0.12	0.05	**0.85**											
Diagnose HIV/Aids	**0.22**	**0.23**	**0.23**	–0.09	–0.08	–0.10	0.09		**0.30**												
Diagnose Hypertension	0.18	0.19	0.17	**0.45**	0.06	–0.10	**0.34**	**0.23**	0.09	–0.05	–0.06	–0.09									
Diagnose Substanzmissbrauch	0.17	0.15	0.15	–0.11	–0.14	–0.11	**0.28**	**0.31**	0.13	0.16	0.17	0.11	–0.05								
Arztbogen: soziale Integration	0.07	0.06	0.06	–0.07		–0.09	0.18	0.18		**0.69**	**0.73**			**0.22**							
Arztbogen: schlechte Compliance						–0.07	0.15	0.15		**0.62**	**0.65**		0.05	0.20	**0.84**						
Arztbogen: niedriges/kein Einkommen	0.12	0.12	0.11	0.08	0.05	–0.08	0.14	0.11		**0.60**	**0.58**		0.05	0.10	**0.59**	**0.55**					
SF-12 körperliche Skala	**–0.39**	**–0.39**	**–0.34**	**–0.34**	–0.14	**0.24**	**–0.42**	**–0.40**	–0.15	–0.08	–0.10		–0.16	–0.09	–0.11	–0.11	–0.14				
SF-12 psychische Skala	–0.14	–0.16	–0.14	0.22	–0.06	0.13	–0.13	–0.17		–0.19	–0.19	–0.07	0.12	–0.16	–0.20	–0.15	–0.10				
Dauer Sport pro Woche	–0.09	–0.10	–0.10	–0.20	–0.06	0.17	–0.20	–0.16	–0.10	–0.11	–0.15		–0.12	–0.09	–0.13	–0.11	–0.11	**0.21**	0.09		
Sprachbeherrschung Deutsch					0.06	–0.09			–0.05	0.11	0.16	–0.06			0.09			–0.17	–0.14	–0.13	
Vertrauen in den Arzt	**–0.26**	**–0.27**	–0.20	**–0.23**		0.17	**–0.24**	**–0.24**	–0.13				–0.10	–0.10				0.19		0.10	–0.16

Die Tabelle zeigt den Spearman-Korrelationskoeffizienten für Zusammenhänge mit einem Signifikanzniveau p <0,01. Korrelationen mit r >0,2 sind fett gedruckt.

EMH Schwabe Editores Medicorum Helveticorum

Einfache Korrelationen zwischen TMI, Kosten und anderen Prognosefaktoren

Tabelle 3 gibt eine Übersicht über die einfachen Zusammenhänge zwischen TMI, DUSOI sowie Kosten und den wichtigsten anderen Prognosefaktoren. Zwischen den logarithmierten Kosten des Jahres 2005 und dem für das gleiche Jahr erhobenen chronischen Anteil des TMI besteht eine Korrelation von r = 0,54, für den DUSOI-Score gilt r = 0,51. Einzig die Kosten des Vorjahres zeigen mit r = 0,68 einen stärkeren Zusammenhang. Die Frage, ob die Gesamtkosten Fr. 20 000.– übersteigen werden, zeigt einen vergleichbaren Zusammenhang wie Alter (r = 0,37) oder wie die körperliche gesundheitsbezogene Lebensqualität nach dem Patientenfragebogen (n = 956, r = 0,39). Bedeutsame weitere Informationen liefert die Diagnose HIV (r = 0,22), die Diagnose Bluthochdruck (r = 0,18), die gewählte Franchise (r = –0,29) und die Auskunft der Patienten, wie gut der Arzt ihre Probleme kennt (r = 0,26; je besser mit den Problemen vertraut, desto höher die Kosten). Die letztgenannte Zahl weist auf die beschränkte Aussagekraft von einfachen Korrelationen hin: In der Regel werden Patienten mit komplexen und kostenträchtigeren Erkrankungen Ärzte aufsuchen, die mit ihren Problemen auch besser vertraut sind.

Vergleich der Vorhersagemodelle (Tabelle 1)

Der TMI verbessert das einfache Modell aus Alter, Geschlecht und gewählter Franchise sowie Versicherungsmodus (24 %) zu einer Varianzaufklärung von 37 %. Informationen über die wichtigsten Diagnosen sowie der DUSOI-Score steigern die aufgeklärte Varianz auf 43 %. Ein Modell nur aus Informationen des Patientenfragebogens erzielt eine Varianzaufklärung von 42 % (SF-12 körperliche und psychische Summenskalen, Sport, soziodemographische Variablen, Belastungen und Stress, Versicherungsinformation). Jedoch verbessern Patienteninformationen das aus dem Arztfragebogen abgeleitete Modell nicht mehr wesentlich (zusätzliche Varianzaufklärung unter 5 %); gleiches gilt für Angaben im Arztbogen zu sozialen Problemen. Das maximale Vorhersagemodell (Tab. 1, Modell 6) erklärt 59 % der Varianz der Kosten im gleichen Jahr und enthält: Versicherungsdaten (Alter, Franchise, Vorjahreskosten), Daten aus dem Arztbogen (TMI, DUSOI, Diagnosen HIV, Hauterkrankungen, Fettstoffwechselstörungen) sowie Daten aus dem Patientenbogen (SF-12 körperliche und psychische Summenskala, soziale Unterstützung). Der Anteil der Kategorie «akute Erkrankungen» des TMI bringt nur 0,1–0,5 % zusätzliche Varianzaufklärung.

Tabelle 4

Kosten und Kostenentwicklung.

Die Tabelle zeigt die durchschnittlichen Bruttoversicherungsleistungen pro Studienpatient und Kategorie Leistungserbringer für die Jahre 2004 und 2005 sowie den Anteil an den Gesamtbruttokosten und die relative Steigerung in Prozent, bezogen auf 2004.

	2005			2004
Leistungserbringer	**Kosten in Fr.**	**Anteil in %**	**Steigerung in % zu 2004***	**Kosten in Fr.**
Spitäler	2634	37,1 %	7,6 %	2449
Ärzte	1940	27,3 %	5,8 %	1834
Apotheken	1451	20,4 %	10,5 %	1313
Pflegeheime	330	4,7 %	53,6 %	215
SPITEX-Organisationen	199	2,8 %	31,4 %	151
PhysiotherapeutInnen	141	2,0 %	14,3 %	123
Übrige Rechnungssteller	132	1,9 %	38,2 %	96
Laboratorien	131	1,9 %	17,1 %	112
Zahnärzte	34	0,5 %	–2,0 %	35
Abgabestellen MiGeL	32	0,5 %	–10,3 %	35
Chiropraktoren	15	0,2 %	–	18
Nichtärztliche Psychotherapeuten	14	0,2 %	–	4
Transport-/Rettungsunternehmen	14	0,2 %	–	20
Sonstige	36	0,6 %	–	32

* Steigerung nur ermittelt für Kostengruppen mit einem Anteil von 0,5 % oder höher.

Diskussion

In dieser prospektiven Studie wird anhand eines Kollektivs von 1967 Patienten aus 27 Praxen in einem städtischen Umfeld der Thurgauer Morbiditätsindikator mit anderen möglichen Daten zur Kostenvorhersage verglichen. Parallel zum TMI werden Diagnosegruppen nach der ICPC-2-Liste kodiert, analog zum gegenwärtig in der Schweiz diskutierten Vorschlag für einen erweiterten TMI. Das wesentliche positive Ergebnis ist, dass sich mit einer Kombination aus Alter, Geschlecht, gewählter Franchise, TMI der Kategorien «chronische Erkrankungen» und wenigen Informationen über die Diagnosegruppen eine gute bis sehr gute Vorhersage der im Erhebungsjahr entstandenen Versorgungskosten erzielen lässt. Das zweite wesentliche positive Ergebnis ist, dass der in dieser Studie durch einen verbesserten Erfassungsbogen modifizierte TMI gleich gute Kostenvorhersagen im heterogenen Patientenkollektiv einer städtischen Grundversorgung ermöglicht wie der ursprüngliche TMI im Kontext von vorwiegend ländlichen und kleinstädtischen Versorgungsnetzwerken.

Die kritischen Ergebnisse dieser Studie lassen sich wie folgt zusammenfassen: Die Kostenvorhersage durch den deutlich einfacheren DUSOI-Score ist fast ebenso gut wie diejenige durch den TMI, beide Instrumente korrelieren untereinander hoch. Dabei fokussiert der DUSOI auf den einfacher zu erhebenden Zustand der vergangenen Woche sowie auf die Prognose und die Behandelbarkeit. Was bedeutet dieser Befund? Entweder schöpft der TMI, wie die anderen Instrumente auch, die maximal mögliche Kostenvorhersage durch Morbiditätsindikatoren aus [2]. Oder aber der TMI enthält noch Potential für Verbesserungen. Die Stärke der vorliegenden Studie gegenüber allen vorangehenden publizierten Arbeiten zum TMI ist, dass der TMI hier nicht wie eine «Black box» eingesetzt wurde, sondern die diagnosebezogenen Bestandteile des TMI erfasst wurden. Damit besteht die Möglichkeit, das Zustandekommen einer TMI-Kodierung zurückzuverfolgen.

Wo liegen bei einer solchen Analyse Ansätze für eine Verbesserung des TMI? Auffällig ist, wie selten bei den chronischen Erkrankungen der Zustand «präklinisch» bzw. die Kategorie TMI 1 kodiert wurde. Abhilfemöglichkeiten bestünden in einer Abänderung der Definitionen, beispielsweise mit dem Ziel, eine Kategorie zu schaffen, die leichte chronische Erkrankungen (z.B. gut eingestellter Hypertonus) und präklinische Zustände zusammenfasst und damit eine bessere Unterscheidung innerhalb der leichten bis mittelschweren Erkrankungen ermöglicht. Auch bei der Unterscheidung zwischen mittelschweren und schweren Erkrankungen bestehen latente, in der Definition des TMI festgelegte Kodierschwierigkeiten, die die Heterogenität der Daten unnötig erhöhen und sich in der schlechten Übereinstimmung mit der Musterauswertung bei einigen der Fallvignetten manifestierten. Obschon die Definitionen des TMI klare und nachvollziehbare Kriterien für die Unterscheidung mittelschwer zu schwer festlegen und wir ausführliche Hilfestellungen auf den Kodierblättern anbrachten, erweist sich die Kodierung von TMI = 3 anstatt TMI = 5 oder umgekehrt als die häufigste Fehlerquelle bei den Fallvignetten. Möglichkeiten zur Optimierung bestünden in einer Veränderung des Auswertungsalgorithmus, insbesondere bei Kombination des TMI mit Diagnosegruppen nach ICPC-2 und mit Medikamentengruppen.

Eignet sich der TMI als Instrument in Versorgungsmodellen, bei denen Morbiditätsindikatoren für die Berechnung von Behandlungsbudgets herangezogen werden? Zunächst lässt sich die Frage positiv entscheiden: Der TMI oder andere Morbiditätsindikatoren verbessern die Kostenvorhersage in relevantem Ausmass. Kritisch ist jedoch die im vorangegangenen Abschnitt angesprochene Kodierunsicherheit zu bewerten, insbesondere zwischen den TMI-Kategorien 3 und 5. Der derzeitige TMI bietet keine interne Handhabe zum Überprüfen der Eingruppierung, sondern kann nur anhand von externer Begutachtung und anhand konkreter Einzelfälle manuell überprüft werden. Würde über den derzeit bestehenden TMI eine Anpassung der Vergütung in Versicherungsmodellen mit Budgetmitverantwortung vorgenommen, so liesse sich eine Tendenz zur Überkodierung von der TMI-Kategorie 3 nach Kategorie 5, wie auch eine Unterkodierung (TMI 3 statt TMI 5) ohne externe Daten kaum überprüfen. Abbildung 1 zeigt, dass zwischen den zu erwartenden Kosten bei TMI 3 und TMI 5 bedeutsame Unterschiede bestehen. Bereits eine Über- oder Unterkodierung von nur 15% aller TMI-3- bzw. TMI-5-Patienten – die sich selbst durch externe Daten retrospektiv kaum nachweisen lassen dürfte, würde nach den Häufigkeits- und Kostenverteilungen in Abbildung 1 die Budgetvorhersagen um etwa 10% verändern. Somit wäre ein vornehmlich TMI-gesteuertes Versicherungsmodell nicht davor geschützt, die Vergütung über eine höhere Kodierung der Patienten zu optimieren als über die schwieriger zu erzielende Optimierung der Versorgung. Hier ist eine Nacharbeitung der TMI-Definitionen ange-

zeigt und eine Kopplung mit extern überprüfbaren Daten, etwa ICPC-2-Klassifikation und Medikamentendaten.

Auf einem anderen Blatt steht die gesundheitspolitische Bewertung: Bei Einsatz für Budgetvorhersagen würden sich in einem hinreichend grossen Ärztekollektiv (etwa der Anzahl der teilnehmenden Praxen an dieser Studie) die Fehlkodierungen über Praxen hinweg zufällig verteilen und ausgleichen. Dies setzt allerdings Stabilität der Vorhersagealgorithmen voraus und verlangt, dass keiner der beteiligten Ärzte bewusste oder unbewusste Überkodierungen vornimmt – und das über längere Zeit hinweg. Zielführender, als derart hehre Kodierungsmoral zu fordern, erscheint den Autoren, zunächst das Instrument zu optimieren.

Liefert die jetzige Studie Ansätze für die Entwicklung eines *TMI-plus*, der präziser und überprüfbarer ist? Die Autoren dieser Arbeit sind der Ansicht, eine Kombination des TMI mit einer Auswahl an Diagnosekategorien verbunden mit einer Ergänzung für besonders kostenintensive Medikamente wäre ein Ausgangspunkt für ein stabileres Instrument. Die TMI-Kodierung müsste dabei innerhalb der Diagnosegruppen vorgenommen werden. Ein so erweiterter *TMI-plus* kommt den derzeit im Gesundheitswesen diskutierten Vorschlägen nahe.

Limitationen

Die vorliegende Studie hat verschiedene zu diskutierende Einschränkungen. Zunächst bedeutet der in dieser Studie eingesetzte TMI-Erfassungsbogen für die Ärzte einen erheblichen zusätzlichen Kodieraufwand. Ein praxistauglicher *TMI-plus* müsste daher auf nicht mehr als 15 bis 20 Diagnosekategorien und wenige zusätzliche Fragen beschränkt sein. In dieser Studie wurden zudem nur diejenigen Versicherten aufgenommen, die während der Untersuchungszeit die teilnehmenden Praxen aufsuchten. Innerhalb der Praxen unterschied sich die Rekrutierungsquote erheblich. Daher sind die Daten aus der vorliegenden Studienpopulation nicht ohne zusätzliche Gewichtung auf die gesamte Bevölkerung generalisierbar. Wie nach jeder prognostischen Studie sollten die hier ermittelten Koeffizienten in einem nächsten Schritt auf die Vorhersagegüte in einer unabhängigen Population überprüft werden. Da bei Versicherungsmodellen mit Budgetmitverantwortung die aktuellen Morbiditätsdaten des Versichertenkollektivs herangezogen werden, um eine Zielgrösse für die Kosten im Folgejahr zu ermitteln, muss die Vorhersagegüte der jetzt im Jahr 2005 erhobenen Daten auf die im Jahr 2006 entstandenen Kosten erst noch ermittelt werden. Die dafür erforderlichen Zahlen stehen erst im Verlaufe 2007 zur Verfügung. Eine weitere Limitation ist der Verzicht auf die Frage nach dem Rauchen.

Generell muss an dieser Stelle die Problematik von Rechenmodellen mit logarithmierten Kostenwerten angesprochen werden. Aus statistischen Gründen ist die mathematische Umformung zwingend. Die vorhergesagte Kosteneinheit Logstufe unterscheidet sich jedoch bei der Rücktransformation in reale Franken ganz beträchtlich: Ausgehend vom geometrischen Mittel von Fr. 2864.– bedeutet eine Kostenzunahme um zwei Logstufen eine Steigerung um Fr. 18 300.–, eine Abnahme um zwei Logstufen eine Minderung um Fr. 2480.–. Anders ausgedrückt: Die Unsicherheit der Vorhersage nimmt bei hohen Kosten exponentiell zu. Daher sollte für Budgetmodelle, die auf den hier vorgestellten oder ähnlichen Rechenmodellen basieren, ein Kostendach pro Versicherten und Jahr vorgesehen werden – etwa bei Fr. 30 000.–.

Zusammenfassung

Die vorliegende Arbeit zeigt, dass der TMI in einem städtischen Kollektiv eine ähnlich gute Kostenvorhersage für die im gleichen Jahr anfallenden Gesundheitskosten ermöglicht wie in den bereits vorliegenden Untersuchungen zum TMI aus mehr ländlichen Regionen. Die jetzige Arbeit liefert Ansatzpunkte für notwendige und sinnvolle Ergänzungen des TMI, bevor ein Einsatz für Budgetversicherungsmodelle empfohlen werden kann. Dazu gehört eine mögliche Revision der TMI-Definitionen und des TMI-Auswertungsalgorithmus sowie eine Ergänzung um ICPC-2-Diagnosekategorien und andere Informationen, wie etwa die gewählte Franchise und die Medikamentengruppen. Der TMI zeigt eine gleichwertige bis bessere Vorhersagekraft als der in den USA entwickelte DUSOI-Score. Der jetzige TMI bietet damit eine gute Ausgangsbasis für die Entwicklung morbiditätsadjustierter Vertragsmodelle oder für einen morbiditätsadjustierten Risikoausgleich zwischen Versicherungen.

Danksagung

Ohne den Einsatz folgender Mitglieder der zmed-Studiengruppe wäre die Arbeit nicht zustande gekommen: Dr. Beda Basler, Dr. Reinhard Baumann, Dr. Hansjörg Bumbacher, Dr. Michael Burkhard, Dr. Peter Christen, Dr. Beat Coradi, Dr. Katja Deseoe, Dr. Thomas Frey, Dr. Andreas Graf, Dr. Philipp Herzog, Dr. Rainer Hurni, Dr. René Jaccard, Dr. Peter Meerwein, Dr. Martin Meili,

Dr. Markus Morf, Dr. Brigitte Müller, Dr. Thomas Müller, Dr. Eva Oberle Nager, Dr. Christoph Schuppli, Dr. André Seidenberg, Dr. Eva Sturzenegger, Dr. Thomas Vögeli, Dr. Luzius von Rechenberg, Dr. Thomas Walser, Dr. Jörg Wälti, Dr. Marco Zoller, Dr. Franziska Züst Lanz; alle Zürich.

Literatur

1 Friedman B, Devers K, Hellinger F, Fraser I. Carve outs and related models of contracting for specialty care: framework and highlights of a workshop. Am J Manag Care. 1998;4 Suppl: SP11-21; discussion SP23-16.

2 Iezzoni LI, Ayanian JZ, Bates DW, Burstin HR. Paying more fairly for Medicare capitated care. N Engl J Med. 1998;339:1933-8.

3 Greenwald LM, Esposito A, Ingber MJ, Levy JM. Risk Adjustment for the Medicare program: lessons learned from research and demonstrations. Inquiry. 1998;35:193-209.

4 Hellinger FJ. Regulating the financial incentives facing physicians in managed care plans. Am J Manag Care. 1998;4:663-74.

5 Morgan RO, Virnig BA, DeVito CA, Persily NA. The Medicare-HMO revolving door – the healthy go in and the sick go out. N Engl J Med. 1997; 337:169-75.

6 Parkerson GR, Jr., Harrell FE, Jr., Hammond WE, Wang XQ. Characteristics of adult primary care patients as predictors of future health services charges. Med Care. 2001;39:1170-81.

7 Parkerson GR, Jr., Hammond WE, Michener JL, Yarnall KS, Johnson JL. Risk classification of adult primary care patients by self-reported quality of life. Med Care. 2005;43:189-93.

8 Okkes IM, Veldhuis M, Lamberts H. Severity of episodes of care assessed by family physicians and patients: the DUSOI/WONCA as an extension of the International Classification of Primary Care (ICPC). Fam Pract. 2002;19:350-6.

9 Bührer A, Zaugg PY. Der Thurgauer Morbiditätsindikator (TMI) – ein wichtiger Schritt auf dem Weg zu besseren Pauschalabgeltungsmodellen. Schweiz Ärztezeitung. 2003;84:264-7.

10 Spycher, S. Der TMI – ein Indikator, der zählt. Schweiz Ärztezeitung. 2006;87:1038-45.

11 Gandek B, Ware JE, Aaronson NK, Apolone G, Bjorner JB, Brazier JE, Bullinger M, et al. Cross-validation of item selection and scoring for the SF-12 Health Survey in nine countries: results from the IQOLA Project. International Quality of Life Assessment. J Clin Epidemiol. 1998;51:1171-8.

12 Yusuf S, Hawken S, Ounpuu S, Dans T, Avezum A, Lanas F, et al. Effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries (the INTERHEART study): case-control study. Lancet. 2004;364:937-52.