

Public-Health-Studien unter der Lupe

Irreführende Janusstudien als Grundlage schwerwiegender Steuerungsentscheide

Der Artikel untersucht anhand der Studie «Geographic variation in the costs of ambulatory care in Switzerland» von André Busato et al. exemplarisch das Problem der fraglichen Qualität von Public-Health-Studien der Schweiz und die daraus folgenden Implikationen auf grundlegende Entscheidungen des Gesundheitswesens.

Michel Romanens^a,
Flavian Kurth^b

a Dr. med., Präsident Verein
Ethik und Medizin Schweiz
(VEMS)

b Sekretär VEMS

Ein Kommentar von
André Busato und Thomas
Rosemann findet sich am
Ende dieses Beitrags.

Richtlinien, an die sich nicht alle halten

Im Gesundheitswesen zahlen die Nachfrager (die Patienten) nicht direkt für die bezogenen Dienstleistungen, sondern indirekt über ihre Versicherungsbeiträge. In einem solchen System (third-party payment) kommt es naturgemäss sowohl zu einer Überproduktion als auch einem Überkonsum, weshalb ohne regulierende Steuerungsmassnahmen die Gefahr einer unkontrollierten Kostenausweitung besteht. Die Grundlagen für diese Steuerungsmassnahmen bilden Public-Health-Studien, deren Qualität folglich die Wirksamkeit der Massnahmen bestimmt. Ist diese mangelhaft, drohen nicht nur die gewünschten Effekte auszubleiben, es kann auch zu genau gegenteiligen Effekten mit Folgen nicht nur für die Kostenentwicklung, sondern auch für die Volksgesundheit und den sozialen Frieden kommen. Wie die Qualität solcher Studien definiert wird, regeln die «Grundzüge guter epidemiologischer Praxis» [1] von Public Health Schweiz (www.public-health.ch), ein Standard bei der Planung, Durchführung, Analyse und Beurteilung der Studienresultate. Nun würde man vermuten, dass sich die Public-Health-Institute der Schweiz an diese selbstauferlegten Richtlinien halten. Man vermutet aber falsch, wie an einem exemplarischen Beispiel besonders gravierender Art aufgezeigt werden soll.

Eine Grundlage, die keine ist

Die 2011 publizierte Studie von André Busato, Pius Matter et al. der Universität Bern mit dem Titel «Geographic variation in the costs of ambulatory care in Switzerland» [2] untersucht den Zusammenhang von Ärztedichte und Überarztung und kommt zum Schluss, dass eine höhere Verfügbarkeit medizinischer Leistungen tendenziell zu mehr unangebrachten Behandlungen führe, und zwar vor allem bei den Spezialärzten. Die Implikationen dieses Schlusses sind Massnahmen wie Zulassungsstopps und die Stärkung der Hausarztmedizin. Das Volk sah das anders, es lehnte die Managed-Care-Vorlage mit 78% Neinstimmen unerwartet deutlich ab. Aus Angst, den Arzt

Des études trompeuses à la base de graves décisions de régulation

Sur la base de l'étude «Geographic variation in the costs of ambulatory care in Switzerland» d'André Busato, Pius Matter et al. de l'Université de Berne, le présent article analyse la question de la faible qualité des études de santé publique réalisées en Suisse et leur implication dans la prise de décisions fondamentales pour le système de santé.

Les auteurs plaident pour la création d'un organe indépendant chargé d'évaluer ces études, d'une chaire d'enseignement et de recherche sur l'adéquation des prestations médicales ainsi que d'un système permettant de vérifier l'adéquation de ces dernières et non plus uniquement leur économicité.

sonst nicht mehr selber wählen zu können, analysierten die Befürworter nach der Abstimmung, die Panikmache der Gegner habe eben gewirkt. Vielleicht war die Bevölkerung aber auch einfach schlauer und sah die proklamierte These in zu argem Widerspruch zur eigenen Empirie: Wieso sollen denn drei Termine beim Hausarzt, bis dieser endlich zum Spezialisten weiterverweist, weniger kosten als einer und gleich beim Spezialisten? Wie konnte es passieren, dass unsere Regierung jahrelang an einer Vorlage arbeitet, die so komplett am Volkswillen vorbeizieht? Die Antwort ist erschreckend: Sie stützte sich dabei auf den Typus obiger exemplarischer Studie, und diese Studien missachten häufig mehrere Auflagen der «Grundzüge guter epidemiologischer Praxis» [1].

Der Verein Ethik und Medizin Schweiz (VEMS) hat die Busato-Studie bereits im November 2011 analysiert und folgende gravierenden Fehler darin entdeckt: a) die mangelhafte Qualität der Datengrund-

Korrespondenz:
Dr. med. Michel Romanens
Innere Medizin
und Kardiologie FMH
Verein Ethik und Medizin
Schweiz (VEMS)
Ziegelfeldstrasse 1
CH-4600 Olten
info[at]kardiolab.ch
www.vems.ch

lage von santésuisse (nicht alle Kassen liefern ihre Daten, Selbstzahler sind nicht erfasst), b) die mangelhafte Risikostrukturadjustierung (Faktoren, die Auswirkungen auf den tatsächlichen Versorgungsbedarf haben, wie Alter und Morbidität, wurden nicht berücksichtigt), und c) der Migrationsbias wurde nicht eingerechnet (Verzerrungen durch Patienten, die sich nicht am Wohnort behandeln lassen).

Es ist klar, dass sich vor diesem Hintergrund der Aussagegehalt der Studie so stark relativiert, dass mit gleichem Recht hätte geschlossen werden können, zwischen Krankheitslast und Arztdichte bestehe kein Zusammenhang. Interessanterweise räumt die Studie das auch ein, proklamiert dann aber doch das Gegenteil, ohne zu sagen, weshalb. Der VEMS nennt solche Studien Janusstudien, weil sie zweideutig sind und deshalb nicht Antworten liefern, sich dafür aber leicht missbrauchen lassen. Wir schickten unsere Kritik an den Autor der Studie, warten aber bis heute auf eine Antwort.

Gutachten schafft Klarheit

Als Professor Thomas Rosemann, Ordinarius am Institut für Hausarztmedizin Zürich, sich in der Nachrichtenendung 10 vor 10 vom 18. Juni 2012 [3] zur Äusserung hinreissen liess, Kardiologen hätten Anreize, unnötige Untersuchungen durchzuführen, und suggerierte, dass sie dies, wenn die Praxis nicht voll sei, oftmals auch täten, um mehr Geld zu verdienen, intervenierten wir brieflich [4]. In seiner E-

Mail-Antwort vom 21. Juni 2012 räumte Professor Rosemann erstaunlicherweise ein [10]: «... Forschungsgelder für eine Versorgungsforschung, die solche Zusammenhänge untersucht, fehlten – wohl weil niemand an der entsprechenden Transparenz interessiert ist ...».

Und was – bitteschön? – soll denn der Auftrag eines Professors sein, wenn nicht der, forschend und lehrend Klarheit und Transparenz zu schaffen? Wir sahen dringenden Handlungsbedarf und beauftragten Prof. Dr. Jürgen Wasem (Lehrstuhl für Medizinmanagement Universität Duisburg-Essen) und die ForBiG GmbH mit der Ausarbeitung einer neutralen Beurteilung der besagten Studie. Das erstellte Kurzgutachten [5] liegt nun vor und bestätigt die vom VEMS monierten Mängel. Auch weist es auf die Arbeit von Dr. Dominik von Stillfried und Thomas Czihal vom Februar 2012 hin [6], welche die groben Fehler der Busato-Studie nicht begangen hat und zu genau gegenteiligen Schlüssen kommt: Zwischen Krankheitslast und Arztdichte besteht kein Zusammenhang. Die Inanspruchnahme stationärer und ambulanter Versorgung ist höher an Wohnorten mit höherer Krankheitslast. Regionale Unterschiede der Arztdichte werden ohne Berücksichtigung der Mitversorgungsfunktion deutlich überschätzt. Entgegen der landläufigen Meinung führt eine höhere Arztdichte im ambulanten Bereich zudem zu einer statistisch signifikant niedrigeren Hospitalisationsrate ($p < 0,0001$) und zu einer geringeren Sterblichkeit ($p < 0,0001$),

¹ Einschränkung: randomisierte Studien beobachten häufig ein hochselektioniertes Krankengut und sind somit nur beschränkt auf den klinischen Alltag übertragbar. Aus randomisierten Studien abgeleitete Richtlinien können deshalb nicht über den Status einer Empfehlung hinausgehen und die Beachtung solcher Richtlinien nicht grundsätzlich als beweisfähiges Qualitätsmerkmal ärztlichen Handelns verstanden werden. Die Zweckmässigkeit eines Indikationsentscheids, basierend auf Level-I-Studien, begründet zudem im Einzelfall der Arzt-Patienten-Interaktion nicht per se einen Handlungszwang, da die Wahrscheinlichkeit des Therapieeffektes nicht uniform, sondern abhängig vom Schweregrad der zu behandelnden Krankheit ist. Dies bedeutet auch, dass die HTA Disziplin CEA (comparative effectiveness assessment) hier noch mit ungelösten statistischen Problemen zu kämpfen hat (siehe auch: <http://physicianprofiling.ch/HTAPretestProbabilityUnresolvedBMJ2012.pdf>).

Typen von Studien

Randomisierte Studie: Probanden erhalten auf zufälliger Basis den Wirkstoff und werden mit jenen verglichen, die an dessen Stelle ein Placebo erhielten. Das Ursachen-Wirkungs-Prinzip lässt sich gut untersuchen, vorausgesetzt, die Teilnehmer sind gutvergleichbar und die Teilnehmer gegenüber dem Wirkstoff verblindet.

Beobachtungsstudie: Aus beobachteten Unterschieden werden Schlüsse gezogen, wobei der Erklärungsgehalt davon abhängt, wie gut der Ursachen-Wirkungs-Zusammenhang untersucht wurde.

Randomisierte Studien sind für die Pharmazeutische Industrie Pflicht, im Bereich der Public Health aber oft zu teuer. Hier haben wir drei Typen von Beobachtungsstudien:

Typ A, Audit-Studien (Zweckmässigkeitsforschung): Vermutete Ursachen werden auf der Ebene der Krankengeschichten systematisch untersucht, um die in Studien vom Typ B vermuteten Erklärungsvariablen zu validieren.

Typ B, Mittelwertvergleichsstudien mit Einschluss von Erklärungsvariablen: Unterschiede werden erfasst, um mögliche Erklärungsvariablen zu finden, wie dies beispielsweise beim BERNCA-Index geschieht.

Typ C, Mittelwertvergleichsstudien ohne ausreichenden Einschluss von Erklärungsvariablen: Janusstudien ohne

wirklichen Erklärungsgehalt, anfällig für Manipulation und gefährlich, weil sie dazu benutzt werden können, Handlungen zu motivieren, die gerade gegenteilige Effekte haben.

Daraus ergibt sich der Evidenzgrad einer Studie:

Level I: Mehrere randomisierte Studien haben den Ursachen-Wirkungs-Zusammenhang bewiesen, es gibt wenig Zweifel daran.

Level IIa: Es gibt noch zu wenige Beobachtungsdaten, die den Ursachen-Wirkungs-Zusammenhang belegen, von diesem auszugehen, ist aber wahrscheinlich nicht falsch.

Level IIb: Es gibt keine randomisierten Studien, aber die Mehrheit der Experten ist sich einig darin, dass der Ursachen-Wirkungs-Zusammenhang überwiegt.

Level IIc: Es gibt noch zu wenige randomisierte Studien, die den Ursachen-Wirkungs-Zusammenhang belegen, aber doch schon so viele, dass mit zu überprüfender Wahrscheinlichkeit von diesem ausgegangen werden kann (Hypothesenstudien).

Level III: Der Ursachen-Wirkungs-Zusammenhang ist nicht belegt, Handlungen, die sich auf ihn stützen, bringen nicht die gewünschte Wirkung oder schaden gar.

womit auch aufgezeigt wird, wo effektiv grosses Sparpotential vorliegt. Diese Dokumente sind inzwischen dem Autor der Studie, Professor Rosemann, dem Ethikrat für Statistik, BAG und EDI zugänglich gemacht. Geändert hat dies allerdings nichts, die Studie wird weiterhin dazu missbraucht, Argumente für zweifelhafte, aber politisch opportune, weil oberflächlich plausible Steuerungsentscheide zu liefern.

Wir fordern eine neutrale Stelle für die Qualitätsbewertung von Public-Health-Studien.

Blindflug auf wackeligen Füßen

Von Mark Twain stammt das Bonmot: «Die Vorurteile eines Professors nennt man Theorie.» Es ist aber klar, dass die Politik auf praktische, empirisch belegte Grundlagen angewiesen ist. Die Versorgungsforschung ist eine grundlegende Wissenschaft, deren Qualität erstrangig zu sein hat und die den Auflagen an die Wissenschaftlichkeit unbedingt und in jeder Hinsicht genügen muss: Unvoreingenommenheit, Transparenz, intellektuelle Redlichkeit.

Dies ist bei der Busato-Studie eindeutig nicht gegeben. Wenn in einer so wichtigen Grundlagenstudie so fundamentale Fehler begangen werden, ohne dass irgendwelche regulativen Kräfte und Instanzen diese Arbeit gebührend überprüft und richtig beurteilt haben, dann können Steuerungsentscheide nicht mehr sein als ein Experimentieren auf wackeligen Füßen. Wozu dies führt, zeigen ebenfalls obige Studie und ihre Folgen: Seit der Aufhebung des Zulassungsstopps im Jahr 2012 ist es zu einer sprunghaften Zunahme potentiell geplanter Spezialarztpraxen gekommen. Der Busato-Studie gemäss wäre jeder weitere Anstieg der ambulanten Gesundheitskosten auf diese neuzugelassenen Spezialärzte zurückzuführen, was bereits wieder den Ruf nach einem Zulassungsstopp laut werden liess. Mit der Gefahr nicht nur einer Versorgungsknappheit, sondern auch einer Kostenausweitung. Denn seit dem letzten Zulassungsstopp ist eine enorme Zunahme der ambulanten Spitalleistungen zu beobachten. Diese machten 2009 bereits 34% der Gesamtkosten aus gegenüber 21% im Jahr 1996 (Quelle Bundesamt für Statistik). Offensichtlich ist es also während des Zulassungsstopps zu einer Verlagerung spezialärztlicher Behandlungen aus den Praxen in die Ambulatorien der Spitäler gekommen. Und dort arbeitet der Spezialist eben gerade nicht zweckmässiger – noch schlimmer: Er wird diesbezüglich nicht einmal kontrolliert, denn die Leistungen der in den ambulanten Abteilungen der Spitäler tätigen Spezialisten unterliegen nicht den Wirtschaftlichkeitsverfahren der santésuisse. Dies bedeutet, dass im Jahr 2009 34% der Rechnungen für OKP-Leistungen keiner gesetzlich vorgeschriebenen

Kontrolle durch die Kassen unterzogen wurden. Kommt es nun zu einem erneuten Zulassungsstopp für Spezialärzte, so würden noch mehr Spezialisten in den ambulanten Bereich der Spitäler abwandern, was unter Umständen zu einer weiteren Kostenausweitung führt.

Es stellt sich doch die Frage, ob die ambulanten Spitalleistungen nach der Einführung von DRG nicht zunehmend zur Quersubventionierung der Spitäler im stationären Sektor verwendet werden. Wäre dies der Fall, würde daraus ein Umsatzdruck auf die ambulant tätigen Spitalärzte resultieren, mit dem Anreiz, unangebrachte Leistungen zu erbringen. Auch ist zu bedenken, dass die hohe Verfügbarkeit der Spitalinfrastruktur zu einer Mengenausweitung verleitet, wie sie in der eigenen Praxis nicht existiert. Drittens ist zu berücksichtigen, dass die Spitäler für Spitalambulatorien und ambulant erbrachte Leistungen wie Röntgenuntersuchungen einen höheren Taxpunktwert verrechnen können, wodurch die Kosten gegenüber dem ambulanten Sektor nochmals um geschätzte 5–10% höher sind. Viertens schliesslich besteht für das Spital ein unleugbarer Anreiz, den ambulanten Sektor für die Rekrutierung von Patienten für den stationären Bereich zu benutzen, was die Kostenlast erneut erhöht. Der VEMS hat hierzu ein Thesenpapier [7] verfasst, zu dem die santésuisse allerdings ihre Stellungnahme ebenso verweigert hat wie die Antwort auf die Frage, ob diese Leistungen auch zukünftig nicht auf ihre Wirtschaftlichkeit hin überprüft werden.

Zweckmässigkeitsforschung und -beurteilung

Die Busato-Studie zeigt exemplarisch, wie in unserem Land eine einzige unprofessionell erstellte Arbeit eine ganze Kaskade unzweckmässiger Handlungen anzustossen vermag. Besonders stossend dabei ist, dass die Resultate solcher Studien nicht als das verkauft werden, was sie sind: Hypothesen. Es ist nachgerade fahrlässig, wenn Ignazio Cassis sagt:

«Das Konzept der Public Health entspringt einer strengen Denkweise, die auf wissenschaftlichen Me-



© complize/photocase.com

Janusstudien machen zweideutige Aussagen, sind deshalb irreführend und können manipulativ missbraucht werden.

Public Health» – also der «auf Wirksamkeitsnachweisen beruhenden Public Health» – lässt keinen Raum für Improvisationen oder Gefühlentscheidungen und umfasst die Entwicklung, Umsetzung und Bewertung wirksamer Programme und Politiken durch Anwendung wissenschaftlicher Überlegungen und systematische Nutzung von Daten und Programmiermodellen».

Das ist Wunschdenken. Die heute verfügbaren Public-Health-Studien der Schweiz sind grösstenteils reine Mittelwertvergleichsstudien vom Typ C und kommen deshalb nicht über den Evidenzgrad Level III hinaus (siehe Kasten). Sie sind aus diesem Grund nicht nur unnützlich, sondern gefährlich. Mittelwertvergleichsstudien unter Einschluss von Erklärungsvariablen vom Typ B wie beispielsweise der BERNCA-Index können Evidenzgrad IIc erreichen. Mit Audit-Studien vom Typ A ist ein Evidenzgrad IIb erreichbar, und dies sollte das Ziel der Public-Health-Forschung der Schweiz sein.

Zu unseren Schlüssen kommt auch das vom Schweizerischen Nationalfonds finanzierte Projekt FP45, wenn die Autoren schreiben: «Die regionalen Unterschiede der Inanspruchnahme von Gesundheitsleistungen lassen vermuten, dass ein Teil der Unterschiede bei den kurativen Leistungen nicht ausschliesslich medizinische Gründe – wie unterschiedliche Häufigkeit von Beschwerden, Krankheiten und Unfällen – haben. Sie können auch durch das Angebot (Ärztedichte) bzw. ein regional unterschiedliches Inanspruchnahmeverhalten mitbeeinflusst sein. Somit sind sowohl Massnahmen aufseits Patientenschaft als auch Ärzteschaft zu postulieren. Aber auch die Forschung ist weiterhin gefordert. Die auffälligen regionalen Unterschiede der Inanspruchnahme sollten vertieft untersucht werden, wenn möglich auf der Basis klinisch abgesicherter Daten.»

Unsere Forderungen

Der VEMS fordert deshalb dreierlei: die Schaffung einer neutralen Stelle für die Qualitätsbewertung von Public-Health-Studien, die Schaffung eines Lehrstuhls für Zweckmässigkeitsforschung und die Schaffung eines Dysfunction Boards [8], um auffällige Ärzte durch die Ärzteschaft selbst zu beurteilen; damit hätte man nicht nur ein Instrument, ärztliche Dysfunktion zu ahnden, es liessen sich so auch auf der Audit-Ebene und via Peer Reviews jene begründenden Erklärungsvariablen finden, die für aussagekräftige Studien unabdingbar sind. Diese Forderungen haben wir Bundesrat Alain Berset mit Brief vom 18. September 2012 mitgeteilt. Sie finden Sie im Detail auf www.vems.ch/public-health sowie im VEMS-Manifest Zweckmässigkeitsforschung [9].

Referenzen

- 1 www.public-health.ch/logicio/client/publichealth/file/EGEP_de.pdf
- 2 Busato A, Matter P, Künzi B, Goodman DC. Geographic variation of consultation cost for ambulatory care in Switzerland. *J Health Serv Res Policy*. 2012; 17(1):18–23. doi: 10.1258/jhsrp.2011.010056. Epub 2011 18. Oktober.
- 3 www.videoportal.sf.tv/video?id=056a574a-cee9-43e0-86ca-056eda732e9f
- 4 <http://physicianprofiling.ch/VEMSKorrespondenzRosemann062012.pdf>
- 5 <http://physicianprofiling.ch/GutachtenWasemBusato092012.pdf>
- 6 http://www.kbv.de/media/pdf/120229_KBVMesse_Dr_vonStillfried.pdf
- 7 www.physicianprofiling.ch/VEMSThesenSpitalkosten122012.pdf
- 8 <http://www.saez.ch/docs/saez/2012/36/fr/BMS-00705.pdf>
- 9 www.physicianprofiling.ch/VEMSManifestZweckmaessigkeit122012.pdf
- 10 www.physicianprofiling.ch/RosemannT2012.pdf

Stellungnahme von André Busato und Thomas Rosemann zur Kritik an ihrer Studie

Herr Romanens unterstellt unserer Studie wissenschaftliche Unredlichkeit, scheut aber selbst nicht davor zurück, uns (T. R.) fragmentarisch und falsch zu zitieren und mit einem als «Kurzgutachten» bezeichneten Werk zu argumentieren, das auf falschen Zahlen beruht und jeden Anspruch an Wissenschaftlichkeit vermissen lässt. Das Niveau der Stellungnahme spricht für sich und lässt für uns eine seriöse, objektive, wissenschaftliche Auseinandersetzung unmöglich erscheinen.

Unter dem Deckmantel der Ethik sollen hier vielmehr politische Partikularinteressen artikuliert werden. Angebotsinduzierte Nachfrage und unangemessene Variation in der Leistungserbringung wird in der internationalen Literatur in vielen Ländern beschrieben. Wir würden dem Autor empfehlen, diese Literatur zur Kenntnis zu nehmen und valide zu belegen, warum die Schweiz das einzige industrialisierte Land ist, in dem diese Effekte nicht auftreten. In Deutschland etwa hat die angebotsgetriggerte Leis-

tungsausweitung nach Einführung der DRGs bereits zu einer heftigen Diskussion darüber geführt, ob Ärzten überhaupt noch die Indikationsstellung zu komplexen und teuren Interventionen zugestehen ist. Dies wäre ein dramatischer Paradigmenwechsel. Genau dies gilt es zu verhindern. Das wiederum geht nicht durch Negieren, sondern durch angemessene Massnahmen, die belegen, dass das in die Ärzteschaft gesetzte Vertrauen im Hinblick auf die Objektivität ihrer diagnostischen und therapeutischen Entscheidungen gerechtfertigt ist.

*Prof. Dr. med. André Busato, Msc.,
Institut für Sozial- und Präventivmedizin Bern,
Institut für Hausarztmedizin IHAM der Universität Zürich*

Prof. Dr. med. Thomas Rosemann PhD, IHAM der Universität Zürich