

Stellungnahme zu
Methodische Weiterentwicklung der Wirtschaftlichkeitsprüfung

Methodenbericht

B,B,S. Volkswirtschaftliche Beratung AG

Auftraggeber:

Verein Ethik und Medizin Schweiz

c/o Dr. med. Michel Romanens
CH-4600 Olten

Angefertigt durch: Dr. Walter Warmuth

Zossen, 14.03.2016

Inhaltsverzeichnis

1.	Vorbemerkung	3
2.	Fundamentalfehler bei der „Modellierungen“	3
3.	Ökonometrische Modellierung – Weiterentwicklung des Unsinn?	6
4.	Zusammenfassung	8

1. Vorbemerkung

Der Autor dieser Stellungnahme erklärt, dass für ihn das im Folgenden Dargestellte keinen Interessenkonflikten unterliegt. Die Darstellungen beruhen auf eigenen wissenschaftlichen Publikationen, für die er keine Zuwendungen aus öffentlichen Mitteln oder Projektfördermitteln von Unternehmen oder Verbänden der privaten Wirtschaft erhalten hat. Der Autor war und ist Mitglied in wissenschaftlichen Fachgesellschaften, er hatte oder hat dort keine Ämter der Leitung oder Selbstverwaltung inne.

2. Fundamentalfehler bei der „Modellierungen“

In der Einleitung (1. des Methodenberichtes) verweisen die Autoren darauf, dass das derzeitige zur „statistischen Wirtschaftlichkeitsprüfung der ambulanten Leistungserbringer“ verwendete Verfahren für „eine flexible Weiterentwicklung ungeeignet“ ist.

Bisher als auch in dem anstehenden „Methodenbericht“ wird unterstellt, dass aus einer „statistischen“ Wirtschaftlichkeitsprüfung auf einen (individuellen) u. U. „teuren“ (Praxis-) Stil geschlossen werden kann.

Das mag bei der Betrachtung der Kosten von Routineuntersuchungen von Kraftfahrzeugen gelingen, auf der Basis von Behandlungskosten ist das unmöglich (**Fundamentalfehler der Gesundheitsökonom Nummer 1**).

Aus statistischen Einflussfaktoren auf den individuellen Praxisstil zu schließen, ist ein Widerspruch in sich: Nehmen wir hypothetisch an, eine Operationsmethode ist in 80 % der Fälle spätfolgenlos erfolgreich, in 20 % der Fälle verstirbt der Patient. In 80 % der Fälle liegt der individuelle Erfolgswert (100 %) statistisch über dem Mittel – 80 %, in 20 % der Fälle (Erfolgswert 0 %) unter dem Mittelwert. Erst bei letztendlich unendlich vielen Operationen in einer einzelnen Praxis könnte aus der erreichten Erfolgswahrscheinlichkeit ein Schluss gewagt werden. Wohl kaum könnte ein Zweifel am Praxisstil begründet werden, denn die neuen unendlich vielen durchgeführten Operationen begründen auch neue statistische Prävalenzen, Inzidenzen und statistisch mittlere Erfolgswerte.

Es verwundert, dass Gesundheitsökonomien sich wenig oder nie mit der Effektivität von Arztpraxen befassen, „Effizienz-Untersuchungen“ werden von Gesundheitsökonomien in der Regel vom Effekt losgelöst betrachtet. In diesem Methodenbericht spielt der Effekt einer Betreuung durch Ärzte überhaupt keine Rolle. Von den Autoren des Methodenberichtes ist es gewissenlos und auch aus Sicht der Sozialwissenschaften unwissenschaftlich, Ineffizienz allein an hohen Honorarforderungen der Ärzte festzumachen. Eine Behandlung ist effizienter als eine andere Behandlung, wenn sie bei einem vergleichbaren Effekt kostengünstiger ausfällt. In dem obigen fiktiven Beispiel könnten die Kosten bei den erfolgreich operierten Patienten untereinander verglichen werden, ebenso könnten auch nur die Operationen mit den verstorbenen Patienten verglichen werden – vorausgesetzt, dass es sich jeweils um auch nur in der Anamnese, in ihren Symptomen und in dem Grad der Überzeugtheit des Arztes, im Rahmen seiner Diagnostik mit einer Therapie beginnen zu können, miteinander vergleichbare Patienten handelt.

Eine Untersuchung von Effizienz ohne Sicherstellung eines ähnlichen Effektes ist nicht möglich.

Teuer wird zudem in Beziehung zu „überdurchschnittlich“ gebracht. Ist der Durchschnitt durch eine Mengenausweitung verkleinerbar, dann führt diese „gesundheitsökonomische“ Bewertungsweise zu summarischen Mehrkosten: (u. U. unnötig) mehr relativ kostenstabile Routineuntersuchungen (z. B. bei Durchgangsärzten) werden nicht „erkannt“. Handelt es sich um „normale“ Facharztgruppen mit Heavy-tail Kostenverteilungen, dann führen zusätzliche Honorarforderungen für an nicht so sehr bedürftigen Patienten durchgeführte Behandlungen zur Senkung des Mittelwertes und zu summarisch höheren Kosten (**Fundamentalfehler der Gesundheitsökonom Nummer 2**).

Zur weiteren Klarstellung – bisherige Verfahren fußten auf einer (mathematischen) Regressionsanalyse. Die Anwendung war an die Erfüllung von Bedingungen geknüpft.

Dieses Modellieren bedeutet:

- I. Die abhängige Variable (Kosten) wird als normalverteilt modelliert.¹
- II. Für die abhängige normalverteilte Variable wird als Modell eine lineare Funktion unabhängiger Variabler und ein zufälliger Fehler, der für alle anderen „Einflüsse“ steht, gewählt.
- III. Die zufälligen Fehler müssen sich im Mittel aufheben.
- IV. Die unabhängigen Variablen sind nicht zufällig, eine neue Beobachtung zeigt in Bezug auf diese Variablen immer das gleiche Ergebnis.
- V. Die unabhängigen Variablen sind linear unabhängig.
- VI. Schwankungen sind homogen, die Streuungen sind gleich.
- VII. Schwankungen sind nicht autokorreliert, Schwankungen verschiedener Beobachtungen sind unkorreliert.

In gesundheitsökonomischen Abhandlungen wird die Gültigkeit der Bedingungen zur Anwendung eines solchen mathematischen Modells nicht untersucht (**Fundamentalfehler der Gesundheitsökonom Nummer 3**).

¹ Bei allen Verteilungen von Behandlungskosten handelt es sich um Heavy-tail Verteilungen, hohe Kosten sind hier „normal“. Ein gutes Modell stellt die Pareto-Verteilung dar. Real sind mit 80 % der Behandlungen rund 20 % der Kosten gedeckt, im 95 % Konfidenzintervall liegen 65 % der Kosten, auf wenige nicht korrekt als Hochkostenfälle [aus gesundheitsökonomischer Sicht sind sie es ganz sicher] bezeichnete Fälle entfallen noch 35 % der Kosten.

Die Anwendbarkeit ist an die Erfüllung dieser Bedingungen geknüpft. Bei der alleinigen Nichterfüllung von VI. kann das falsche Modell dennoch (etwas) nützlich sein. Die Nichterfüllung von I. ist jedoch fatal, Ableitungen aus einer dennoch vorgenommenen Modellierung als Regressionsanalyse sind Unsinn.

Im vorliegenden Methodenbericht wird auf die ANOVA²-Korrektur von Roth & Stahel (2005) verwiesen. Die ANOVA Korrektur wird im Methodenbericht als „bisheriges Verfahren“ angeführt. Es sagt nur etwas über die Kosten aus, die um einen von höheren Kosten bereinigten Mittelwert schwanken und sich dabei ausgleichen. Im Rahmen eines Beharrens auf einer „statistischen Durchschnittsmethode“ stellt die ANOVA Korrektur eine Veränderung dar, die, dort wo sie anwendbar ist, weder gut noch schlecht ist. Die bedingte Anwendbarkeit bezieht sich nur auf den sich ausgleichenden Kostenbereich. Es wird an keiner Stelle untersucht, ob eine Varianzanalyse überhaupt anwendbar ist.

Die Methode gründet sich auf eine Zerlegung der Varianz in als identisch postulierten Schwankungen um die zunächst noch unterschiedlichen Mittelwerte und die durch die Unterschiede der Mittelwerte bedingten Schwankungen. Die unterschiedlichen Mittelwerte werden zu gleichen Mittelwerten „korrigiert“. Methodisch ist das im gesetzten Rahmen soweit in Ordnung. Nun ist die Kostenverteilung aber weit entfernt davon, symmetrisch verteilt oder gar normalverteilt zu sein. Wenige hohe Kosten in Arztreferenzkonstellationen beeinflussen sehr stark den Mittelwert, die Varianzen sind ebenfalls hochsensibel betroffen. Von identischen Varianzen auszugehen, ist nicht lauter.

Für einzelne Facharztgruppen ist die abhängige Variable (Kosten) als nahezu symmetrisch verteilt (eine Unterschreitung von mittleren Kosten wird durch adäquate Überschreitungen kompensiert) anzusehen – mathematisch ist das im Einzugsbereich des Zentralen Grenzwertsatzes erklärbar. Eine Verwendung eines Regressionsmodells ausschließlich für diese Facharztgruppen wäre möglich. Für das Ziel, einer Überarztung zu begegnen, ist das jedoch unbrauchbar. In diesen Facharztgruppen kommen die sonst üblichen wenigen hohen Kosten vermutlich nicht vor. Die Ausrichtung am Mittelwert lässt die Überarztung durch eine Mengenausweitung als nicht auffällig erscheinen. Bei der Verwendung von anzahlmäßigen Indizes blockiert der Fundamentalfehler der Gesundheitsökonom Nummer 1, es kann nicht von einer statistischen Wirtschaftlichkeit auf eine individuelle Wirtschaftlichkeit geschlossen werden. Auch aus dem Wissen heraus, dass Mengenausweitungen zu einem Fallen der Mittelwerte führen, sind Kandidaten Arztpraxen mit einer Überarztung gut durch andere Arztpraxen maskiert.

Da die Gesundheitskosten in der Mehrzahl der Facharztgruppen aber Heavy-tail Verteilungen genügen, handelt es sich bei den im Methodenbericht verkündeten Weiterentwicklungen um die Weiterentwicklung von Unsinn.

² ANOVA (von analysis of variance) Varianzanalyse

3. Ökonometrische Modellierung – Weiterentwicklung des Unsinn?

Der als „Weiterentwicklung“ ausgegebene ökonometrische Modellrahmen verringert die Grundfehler der Modellierung „Wirtschaftlichkeitsprüfung der ambulanten Leistungserbringer“ nicht:

- a) Es wird weiterhin behauptet, aus statistisch ermittelten Kostengrößen auf die Unwirtschaftlichkeit von konkreten Arztpraxen schließen zu können.
- b) Auch wenn die Untersuchung der Effektivität im sozialwissenschaftlichen Kontext letztendlich auf die Formel „koste es, was es wolle“ hinausläuft, – es ist sogar für Sozialwissenschaftler unwissenschaftlich, die Effizienz von Arztpraxen ohne Betrachtung der Effektivität – wie in diesem Methodenbericht – zu postulieren.
- c) Der Aufklärung, worin sich Überarztung vermutlich erkennen lässt, wird in dem Methodenbericht überhaupt nicht nachgegangen.
- d) Ein fachlicher Bezug zum Gesundheitswesen der Schweiz wird von den Autoren des Methodenberichtes nicht hergestellt:

Worte, die auf Verwaltungsdaten schließen lassen, sind Bruttoleistung nach Leistungsart. Als Panel wird die zweidimensionale Struktur der Gruppen von Erkrankten (eine Dimension Geschlecht-Alter-Gruppe) und Gruppen von Ärzten (eine Dimension Facharzt-Kanton-Gruppe) genutzt. Hiermit erschöpft sich schon der Bezug auf Daten im Methodenbericht „Methodische Weiterentwicklung der Wirtschaftlichkeitsprüfung“.

Die Methodendarstellung erfolgt lehrmethodisch, sie ist vermutlich einer Methodendarstellung entlehnt, wo Kosten den Charakter von Kosten verschiedener Generika für Aspirin haben. Vermutlich um einen Hinweis auf das kurative Wirken von Ärzten zu geben, taucht im Methodenbericht eine virtuelle Matrix von zusätzlich „berücksichtigten Variablen“, die „die Morbidität oder das Patientenverhalten abbilden“, auf. Ebenfalls gibt es einen Vektor, der „zusätzlich ... Variablen subsumiert, welche nur auf der Arztebene variieren“. Inhaltlich gibt es zu den beiden letztgenannten Strukturen keine Erläuterungen. Aus dem Hinweis im Methodenbericht, dass es sich dabei „sowohl“ um Variablen handeln kann, die „diskret als auch kontinuierlich“ sind, kann geschlussfolgert werden, dass diesem Teil der „Methode“ noch die Ausarbeitung fehlt.

Als mathematisches Modell wird eine Regressionsanalyse verwandt (im Abschnitt 2.3 des Methodenberichtes als multivariate Regressionsanalyse bezeichnet, in 2.3.1. wird noch angenommen, dass nicht beobachtete Einflüsse [also alle Beobachtungen außer Zugehörigkeit zur Geschlecht-Alter-Gruppe und Zugehörigkeit zur Facharzt-Kanton-Gruppe] „nicht von den erklärenden Variablen abhängen“). In der Methodenstudie wird nicht versucht, die Anwendbarkeit dieser Regressionsanalyse zu belegen. Dadurch, dass Kosten „der ambulanten Leistungserbringer“ in der Regel einer Heavy-tail Verteilung unterliegen, ist die Modellierung von diesen Kosten im Rahmen einer Regressionsanalyse eigentlich Un-

sinn³. Es muss als Hohn angesehen werden, den Statistikern in Gesundheitswesen und in der Medizin zu erklären, „dass Statistikexperten aus verschiedenen Disziplinen die Methode nachvollziehen können“. Auch das Argument, dass die „Regressionsanalyse ... eine weit verbreitete Methode ... in vielen empirischen Wissenschaften“ ist, rettet den Methodenbericht nicht⁴.

Im Abschnitt „3. Erweiterungen mit zusätzlichen Variablen“ wird mit einem Modellgüte-Maß (angepasstes R^2) operiert. Dieses Maß hat nur Sinn für den Vergleich von Regressionsmodellen, hier: welcher Unsinn hat eine bessere Güte.

Auf dem Weg, gute Schätzungen in unsinnigen Modellen zu generieren, gelingt es den Autoren des Methodenberichtes, jeder Facharztgruppe letztendlich in manchen Fällen nützliche Mittelwerte zuzuordnen. Für manche Facharztgruppen wäre eine separate Spezifizierung auch im Rahmen eines dafür separaten Regressionsmodells möglich. Hier wäre auch die wesentliche Annahme vertretbar, dass nicht beobachtete Einflüsse [also alle Beobachtungen außer Zugehörigkeit zur Geschlecht-Alter-Gruppe und Zugehörigkeit zur Facharzt-Kanton-Gruppe] „nicht von den erklärenden Variablen abhängen“. Im Rahmen des Gesamtmodells wurde die Auswirkung dieses Trainings im Rahmen der „ersten Stufe“ (2.3.1.) nicht untersucht, letztendlich ist aber eine Aufteilung in Facharztbereiche mit sich einerseits gut symmetrisch kompensierenden Kosten und die anderen Facharztbereiche kontraproduktiv.

In den wenigen nun gut modellierten Facharztgruppen kann von einer „Weiterentwicklung“ nicht ausgegangen werden, z. B. ist das Problem einer Überarztung nicht mehr erkennbar, weil kaum wesentliche Abweichungen von Mittelwerten auszumachen sind. Die Überarztung wäre an einem u. U. geringeren Mittelwert erkennbar – entstanden durch ein übermäßiges Aufgreifen auch von nicht so „bedürftigen Patienten“. Hier ist die Überarztung durch eine Mengenausweitung mit summarischen höheren Kosten erkennbar. Alle Mittelwert-Methoden, bisherige Modelle und die Methoden des Methodenberichts, verfehlen hier das verkündete Ziel.

Die Spezifizierung in einer Facharztgruppe mit nahezu normalverteilten Kosten führt zu einer Separierung, das Ziel einer Wirtschaftlichkeitsprüfung wird jedoch hier mit dem derzeitigen und im Methodenbericht benannten Instrumentarium überhaupt nicht erreicht. Den anderen Facharztgruppen wird aber nun Masse entzogen, die verbleibenden hohen Kosten (dort zwar selten aber eigentlich als normal auftretend anzusehen) suggerieren im Lichte der Mittelwertmethoden eine Überarztung – sie sind es aber in ihrem Umfang ganz sicher nicht. In den Facharztgruppen mit den wirklichen vereinzelt höheren Kosten wäre Heavy-tail noch mehr schwanzlastig, eine multivariate Regressionsanalyse ist nicht anwendbar – noch mehr Unsinn.

³ Die Argumentation, „zu einer massiven Verkürzung der Rechenzeit im Vergleich zum bisherigen Verfahren“ beizutragen, kann der Autor dieser Stellungnahme nicht positiv aufnehmen. Die Autoren des Methodenberichtes hätten besser daran getan, nach Methoden zu suchen, die in Gesundheitswesen konkret bei ambulanten Leistungserbringern anwendbar und u. U. nützlich sind. Was wurde denn überhaupt „gerechnet“, mit welchen Daten wurde „gerechnet“, wozu und warum wurde denn überhaupt „gerechnet“?

⁴ Diese Argumentation hat die Gedankenschwere, einem Chirurgen als Werkzeug einen Hammer anzupfehlen, da ein Hammer in vielen Berufen benutzt wird.

4. Zusammenfassung

Der Methodenbericht „Methodische Weiterentwicklung der Wirtschaftlichkeitsprüfung“ ist in Gesundheitswesen und zur Bewertung der Wirtschaftlichkeit ambulanter Leistungserbringer unbrauchbar. Die Begründung dieser Aussage ist ausführlich im Abschnitt 3 dieser Stellungnahme dargestellt.

Die Methoden entbehren des Bezuges zum Gegenstand „Wirtschaftlichkeitsprüfung der ambulanten Leistungserbringer“. Als ökonometrische Methoden der Effizienzprüfung heben die Methoden von jeder Art Effektivitätsbezug ab – das ist in empirischen Sozialwissenschaften unwissenschaftlich: die Bewertung wird zwar als regelgebunden dargestellt, die Bewertungen stellen aber keine systematischen und transparenten Bewertungen des Gegenstandes Wirtschaftlichkeit ambulanter Leistungserbringer dar. In Bezug auf die Qualität der Bewertungen der Wirtschaftlichkeit ambulanter Leistungserbringer werden im Methodenbericht alle 27 Einzelstandards des Standards der Schweizerischen Evaluationsgesellschaft SEVAL – zur Nützlichkeit, Durchführbarkeit, Korrektheit (in Deutschland – DeGEval – heißen diese inhaltlich gleichen Einzelstandards Fairness) und Genauigkeit verletzt.

Ambulante Leistungserbringer behandeln ihre Schutzbefohlenen nach völlig anderen Kriterien, als das im Methodenbericht angenommen wurde. Ihre Leistung wird an Patienten erbracht; Anamnese, Symptomatik, Diagnostik, ... steuern Therapien, steuern die Leistung – dafür beanspruchen Ärzte ein Honorar.



Zossen, 14.03.2016

Dr. Walter Warmuth