Zipaper

4 2014

Grundsätze für die Wahl einer Standardisierungsmethode

Peter von der Lippe

www.zi.de

Zentralinstitut für die
kassenärztliche Versorgung
in der Bundesrepublik Deutschland
Herbert-Lewin-Platz 3
10623 Berlin

E-Mail: mleibner@zi.de
Tel. +49 30 4005 2450
Fax +49 30 4005 2490

ISSN 2199-1480 (online)

Grundsätze für die Wahl einer Standardisierungsmethode

Prof. Dr. Peter von der Lippe[1]

1 Emeritus, Fachgebiet: Statistik, Schwerpunkte: Wirtschafts- und Sozialstatistik, deskriptive Statistik (insbesondere Preisindextheorie) und Stichprobentheorie, Universität Duisburg-Essen

Inhaltsverzeichnis

1. Anlass und Gegenstand der folgenden Ausführungen

In seinem Papier "Standardisierung des Überschusses von Arztpraxen im niedergelassenen Bereich - Vergleichende Analysen des ZI-Praxis-Panel auf Grundlage der Erhebungswellen 2010 bis 2012" (kurz ZI-Papier) diskutierte Martin Kohler "Verbesserungen im Standar-disierungsverfahren" und er entwickelte Vorschläge für eine neue Standardisierungsmethode. Das Papier schließt mit der Feststellung: "Allerdings fehlen bisher noch Kriterien, die eine Empfehlung zugunsten eines spezifischen Standardisierungsansatzes begründen könnten."(S. 39).

Das lässt den Verdacht aufkommen, dass hier das Pferd von hinten aufgezäumt wird; denn man sollte m. E. zuerst "Kriterien, die eine Empfehlung … begründen könnten" entwickeln und dann auf der Basis dieser Kriterien alternative Standardisierungsmethoden diskutieren[1].

Dies, aber auch ein zweiter Verdacht, dass wir uns nämlich beim Thema "Standardisierung" vielleicht im Kreise drehen könnten (davon mehr in Abschn. 5) gab den Anlass zu der vorliegenden Betrachtung, in der ich fünf Gütekriterien einer Standardisierung vorschlage.

Ich stütze mich dabei auf drei früher von mir verfasste Texte:

1. Standardisierung der Einnahmen einer Arztpraxis Methoden der Honorarumrechnung einer Arztpraxis auf Einnahmen einer "Normpraxis", die ausschließlich EBM-Leistungen (für GKV Patienten) in Vollzeit erbringt, Heft 191 (2011) der IBES Diskus-sionsbeiträge aus dem FB Wirtschaftswissenschaften Univ. Duisburg-Essen (Campus Essen) November 2011; kurz: λ-Papier,
2. die unveröffentlichte Arbeit (intern für das ZI verfasst) "Methoden zur Schätzung des Aufschlagfaktors α"; Sept. 2011; kurz α-Papier, und (die weit weniger wichtige veröf-fentlichte Arbeit)
3. Standardisierung des Überschusses (der Einnahmen und der Kosten) bei Arztpraxen Berücksichtigung evtl. unterschiedlicher Kosten bei der Behandlung von PKV und GKV Patienten. Heft 197 (2012) der IBES Diskussionsbeiträge aus dem FB Wirt-schaftswissenschaften Univ. Duisburg-Essen (Campus Essen)[2] kurz Kosten-Papier.

sowie auf zwei weitere interne (nicht veröffentlichte) Papiere zu Entwürfen des Kohler Papiers, nämlich Notiz 1 (4 Seiten vom 9.2.2014) und Notiz 2 (17 Seiten v. 30.5.2014). Die von mir vorgeschlagene Standardisierungsmethode ist v. a. im λ-Papier Schritt für Schritt entwickelt und begründet worden. Ich halte es aber für geboten, trotzdem im Folgenden (u. a. auch im Anhang zu diesem Papier) noch einmal auf einige der dort vorgetragenen grundsätzlichen Überlegungen einzugehen.

2. Zielsetzung der Standardisierung und Definition der Normpraxis (NP)

Die Anpassung des Orientierungswertes und der EBM Bewertungsrelationen verlangt eine Umrechnung der im ZiPP erhobenen Daten über Einnahmen (Umsätze) E_i und Reinerträge einer erhobenen konkreten Arztpraxis i (i = 1,…,n), mit einem Privatpatientenanteil von $p_i > 0$ und einer Arbeitszeit A_i in einen (geschätzten, fiktiven) Erlös, den eine entsprechende, den Verhandlungen über Arzthonorare zugrunde zu legende "Normpraxis" (NP) hätte. Diese fiktive NP ist dadurch gekennzeichnet, dass sie ausschließlich EBM-Leistungen (für GKV Patienten) in Vollzeit erbringt, also

1. nur GKV Patienten versorgt ($p_i = 0$) und damit Einnahmen (E) nur von der GKV hat so

1 Es ist mir auch ein Rätsel, wie man einerseits von "Optimierung" (Kap. 7 des Kohler-Papiers) sprechen kann, andererseits aber das Fehlen von Gütekriterien beklagen kann.

2 Das ist praktisch eine Fortführung des -Papiers.

dass gilt $E_i = E_{Gi}$ und dies

2. in einer jährliche Arbeitszeit von A* = 140.148 Minuten (= 2.335,8 Stunden) erzielt.

Es ist also den Besonderheiten $p_i \neq 0$ und $A_i \neq A^*$ der Praxis i Rechnung zu tragen. Wie sinnvoll eine solche Standardisierung ist, und was man sich von einer entsprechenden Umrechnung verspricht, ist nicht Gegenstand der folgenden Betrachtung. Was aber schon gleich hier deutlich werden sollte ist, dass die Definition NP mit $p_i = 0$ und $\kappa_i = A_i/A^* = 1$ Dreh- und Angelpunkt der Standardisierung ist. Das führt zum ersten Gütekriterium:

1. Im Vordergrund aller Betrachtungen zum Standardisierungsverfahren sollte die Frage stehen: wie (mit welchen Merkmalen, in welcher Differenzierung) ist die NP definiert und simuliert man mit der Rechnung auch tatsächlich *diese* so definierte NP?

Ein zweites Kriterium ergibt sich daraus, dass die Standardisierung der Einnahmen E_i der Praxis i kaum jemals unumstritten sein wird, weil i.d.R. unausgesprochen stets der Verdacht mitschwingen wird, dass sich hier die Ärzte mit dubiosen Tricks "arm rechnen" wollen. Die Konsequenz daraus kann m. E. nur darin bestehen, dass streng beachtet wird:

2. Das Standardisierungsverfahren sollte

- *allein* mit den *im ZiPP erhobenen Daten* der Praxen (*keine* Bezugnahme auf *externe* Parameter[3], weil solchen Größen oft andere Definitionen zugrundeliegen) und

- *nur* mit reinen *Definitionsgleichungen* operieren (kein Rückgriff auf Modellbetrach-tungen, z. B. einer Regressionsrechnung[4], weil diese nie ohne Annahmen auskommen, die evtl. nicht erfüllt sind und nur mehr oder weniger gut den Daten angepasst sein können, was zu Diskussionen verleiten dürfte, die vermeidbar sind.

Nimmt man auf Größen Bezug, die von anderen Autoren geschätzt wurden wird man sich immer mit Problemen der Vergleichbarkeit von zugrundeliegenden Definitionen und mit entsprechenden Diskussionen über methodische Probleme der Arbeiten dieser Autoren auseinan-dersetzen müssen, auch wenn diese in der Fachwelt noch so angesehen sein mögen.

Im Fall der Regressionsrechnung kann man stets darüber diskutieren welche Regressoren (erklärende Variablen) "richtig" sind, ob z. B. die Annahme der Linearität zutrifft, und ob auch die für die Methode der kleinsten Quadrate notwendigen Annahmen über die Störgrößen (Zufallsvariablen U_i) gelten, was keineswegs selbstverständlich ist (auch wenn einige, wie z. B. die Homoskedastizität getestet werden können)[5]. Es ist stets das Problem der Fehlspezifikation zu gewärtigen, wenn wichtige Regressoren im Modell nicht erscheinen (omitted variables) oder Regressoren mit der Störgröße korrelieren (endogeneity bias).

Man sollte solche Diskussion wenn möglich vermeiden. Sie machen ein Standardisierungsverfahren, das auf externe und mit Modellbetrachtungen gewonnene Größen Bezug nimmt unnötig angreifbar (was bedauerlich ist, zumal ein Standardisierungsverfahren, wie das von mir vorgeschlagene einen solchen Rückgriff auch gar nicht erforderlich macht).

Darüber hinaus haben wir bei allen Übernahmen von bestimmten Konstanten aber auch noch folgende zentrale Probleme (siehe Kriterien 3 und 4)

3 wie etwa der Aufschlagsfaktor α = 2,3 von Wasem.

4 egal ob mit ZiPP- Daten oder – noch bedenklicher – mit anderen Daten.

5 Solche Tests (ob die Annahmen eines Regressionsmodells gelten) unterbleiben jedoch meistens in der Praxis.

- wie diese Größen in die Gleichung eingebaut werden sollen, mit der die Umrechnung der konkreten Einnahmen Ei einer Praxis in die standardisierten Einnahmen dieser Praxis i erfolgen soll (die Umrechnung symbolisieren wir mit $\mathrm{E}_{\mathrm{i}} \rightarrow \hat{\mathrm{E}}_{\mathrm{i}}$).

- ob bei der Standardisierung überhaupt mit solchen praxisübergreifenden (für alle n Praxen oder alle Praxen einer Gruppe, z. B. einer Fachgruppe oder Region geltenden) Größen gerechnet werden sollte statt ausschließlich mit praxisspezifischen (praxisindividuellen) Größen (ich gehe darauf in Abschn. 7 mit Kriterium 4 ein)

Auf die eine Umrechnung $\mathrm{E}_{\mathrm{i}} \rightarrow \hat{\mathrm{E}}_{\mathrm{i}}$ ermöglichende Gleichung kommen wir im Folgenden wiederholt zurück. Sie ist zentral für ein Standardisierungsverfahren und sie sollte deshalb explizit angegeben und sorgfältig begründet werden; denn ohne sie weiß man nicht, was tat-sächlich gerechnet wird bzw. wurde und wie dies zu interpretieren ist.

3. Einige implizite Kriterien im ZI-Papier

Die Betrachtungen von M. Kohler gehen implizit von der Vorstellung aus, dass es gut für das Standardisierungsverfahren sei, wenn die Bestimmungsgrößen des Verfahrens[6], wie E_i, A_i, p_i oder auch κ_i und λ_i und die Ergebnisse des Verfahrens "stabil"[7] sind, aber ich kann keine Begründung hierfür finden. Die Frage, wie "stabil" oder volatil Größen, die zur Standardisierung herangezogen werden (oder auch die Ergebnisse der Standardisierung) sind kann kein Kriterium für die Sinnhaftigkeit eines Rechenverfahrens sein oder dafür, dass evtl. andere, weniger volatile Größen zur Standardisierung herangezogen werden sollten.

Ein Standardisierungsverfahren ist nicht gut, weil sich die empirischen Ergebnisse im Zeitablauf nicht unterscheiden und es ist nicht schlecht, wenn sie sich im Zeitablauf stark unterscheiden.

Die Frage "kann man das Standardisierungsverfahren verbessern und wie kann man das?" betrifft ein *konzeptionelles* Problem (Was soll gemessen werden? Wie ist etwas definiert? Berechne ich mit meinen Formeln auch das, was mit der Standardisierung intendiert ist?). Es ist nicht eine Frage danach, welche empirischen Ergebnisse man mit den konkret im ZiPP erhobenen Daten gewonnen hat und wie sich diese im Zeitablauf verändern.

Ob und wie sich die empirischen Ergebnisse für E_i, κ_i, λ_i usw. in den Befragungswellen unter-scheiden steht auf einer ganz anderen Ebene als die konzeptionellen Probleme (also die Frage der Validität der Messung: für diese ist nach Kriterium 1 allein maßgebend ob sie im Einklang mit der vorgegebenen Definition der NP stehen).

Es heißt im Kohler Papier, die von ihm vorgeschlagenen Methoden hätten den Vorteil, dass sie "nicht auf die vulnerable Relation von gruppenspezifischen Patientenzahlen und –erlösen angewiesen sind. Dabei zeigen sich stabilere Ergebnisse im Zeitverlauf" (Seite 39). Wenn mit vulnerabel "volatil" gemeint ist, wird hier m. E. ein nicht sachgerechtes Kriterium angewendet. Ob eine statistische Methode gut oder schlecht ist, entscheidet sich nicht danach, ob ihre Ergebnisse stabil oder volatil sind, sondern danach, wie sie begründet wird.

Ergebnisorientiertes Denken ist bei Entscheidungen über Methoden nicht angebracht, schon gar nicht, wenn die Präferenz für "Stabilität" nicht kritisch hinterfragt wird. Dass sich eine Größe im Zeitablauf verändert ist als solches ja nicht überraschend und besagt nichts, wenn man keine Gründe dafür angibt, warum sie sich vielleicht besser nicht verändern sollte.

---

6 Sie werden erst im folgenden Abschnitt 4 im Einzelnen vorgestellt.

7 Es soll an dieser Stelle noch nicht problematisiert werden, wie Kohler die "Stabilität" der Ergebnisse definiert (nämlich mit Korrelationskoeffizienten). Wir kommen darauf erst in Abschn. 10 zu sprechen.

4. Erste Einführung in die Standardsierung mit κ_i und λ_i

Es dürfte sinnvoll sein, die Standardisierung nach dem von mir vorgeschlagenen Verfahren zunächst in zwei Schritte zu zerlegen, weil auch zwei Aspekten der Definition der NP Rechnung zu tragen ist, nämlich

- der Arbeitszeit (die konkrete Arbeitszeit A_i der Praxis i kann von der Arbeitszeit A* der NP abweichen, weshalb $\kappa_i = A_i/A^* \neq 1$ sein kann) $E_i \rightarrow \hat{E}_i^Z = E_i/\kappa_i$ [8] und

- der Struktur der Patienten ($E_i \rightarrow \hat{E}_i^S = E_i/\lambda_i$).

Es ist danach die Zeit- oder Auslastungsstandardisierung einerseits und die [Versicherten-] Strukturstandardisierung andererseits zu unterscheiden, auch wenn mit $\ddot{E}_i = E_i/\kappa_i\lambda_i$ die beiden Standardisierungen zusammen vorgenommen werden können. Wichtig ist allein, dass so (nach Kriterium 1) die (rechnerische) Simulation einer NP (also einer reinen GKV-Praxis mit Arbeitszeit A*) gelingt und dass auch begründet werden kann, warum

(1) $\hat{E}_i = E_i/\kappa_i\lambda_i$

in der Tat die Gesamteinnahmen dieser NP wären.

Die gesamten Einnahmen E_i einer Praxis i setzen sich zusammen aus Honoraren für die Behandlung von GKV versicherten Patienten (E_{Gi}) und anderen Einnahmen $E_{Pi} = E_i - E_{Gi}$,

(2) $E_i = E_{Gi} + E_{Pi}$.

Es ist eine Frage nach der korrekten Definition der NP, ob bei E_P alle Einnahmen außer E_G zu nehmen sind, oder ob EP nur Einnahmen aus der Behandlung von Privatpatienten umfassen sollte, so dass dann E_i in Gl. 2 zuerst einmal um sonstigen Einnahmen bereinigt werden muss.

Mir scheint die erste Definition korrekt zu sein[9], weil die NP ja durch eine *Gesamteinnahme* definiert ist, die man hätte, wenn man nur Einnahmen aus der Behandlung von GKV-Patienten hätte[10], und keine sonstigen Einnahmen wie IGeL, BG- oder Privatpatienten, Gut-achten, Leistungen für Krankenhäuser usw. hätte. Der Arzt muss allein von seinen GKV Ein-nahmen leben können[11].

Gilt dagegen die zweite Definition wäre für E_i in Gl. 2 eine um entsprechende Positionen gekürzte Einnahme einzusetzen und E_{Pi} enthielte dann nicht mehr die sonstigen Einnahmen (weder von der GKV noch von der PKV). In jedem Fall ist (2) eine reine Definitionsglei-chung, d. h. sie gilt immer und kann nie "falsch" sein weil sie aus den zugrundeliegenden De-finitionen folgt. Das gilt auch für

(2a) $N_i = N_{Gi} + N_{Pi}$,

wonach sich die N_i Patienten der Praxis i zusammensetzen aus N_{Gi} GKV- Patienten und NPi PKV-Patienten (oder andere "Nicht-GKV-Patienten"), so dass auch gilt

---

8 Die hinter dieser Gleichung mit κ steckende Logik hat stets weniger Probleme bereitet als das Verständnis von λ. Wir brauchen also noch nicht an dieser Stelle (sondern erst im Anhang) auf den Sinn hinter κ einzugehen.

9 Wenn das der Fall ist kann m.E. die mit Gl. 4 definierte Größe α_i nicht als Fallwertrelation bezeichnet werden.

10 Die NP ist m.E. so definiert, dass es auf die Einnahmen E (= Umsätze) ankommt, die man allein mit GKV Patienten erzielt, und die standardisierten Einnahmen nicht nur keine PKV Honorare umfassen, sondern auch keine sonstigen Einnahmen (die evtl. gar nicht mit der Behandlung von Patienten zusammenhängen).

11 Dem Konzept der NP haftet daher etwas vom Alimentationsprinzip des Beamtenrechts an: mangels marktmä-ßiger Bewertungsmöglichkeiten der Leistung des Beamten (wir wollen nicht, dass der Beamte sein Amt "ver-marktet") muss das Prinzip sein, dass ihm quasi ein "standesgemäßer" Lebensunterhalt möglich sein sollte.

(2b) $p_i = N_{Pi}/N_i$ und $1 - p_i = N_{Gi}/N_i$.

Man kann nun E_i nach Gl. 2 in Verbindung mit den Patientenzahlen (Gl. 2a) bringen und erhält zwei Einkommens- bzw. Umsatzbestandteile[12]

(3) $E_i = N_i(1-p_i)e_{Gi} + N_i p_i \alpha_i = N_i e_{Gi}[(1-p_i) + p_i\alpha_i] = N_i e_{Gi}[1+(\alpha_i-1)p_i] = N_i e_{Gi}\lambda_i$.

Wenn hierin mit $e_{Gi} = E_{Gi}/N_{Gi}$ und $e_{Pi} = E_{pi}/N_{pi}$ quasi pro-Kopf-Einnahmen[13] definiert sind und für die ominöse Größe α_i gilt

(4) $\alpha_i = e_P / e_G$

ist immer noch keine, evtl. der Wirklichkeit nicht gerecht werdende Annahme im Spiel[14]. Daraus erhält man die (im Anhang noch einmal erläuterte) Definition für λ_i

(5) $\lambda_i = 1 + (\alpha_i - 1)p_i$

Hätte die Praxis i nur GKV Patienten (p_i wäre dann p_i =0) und würde sie (wie mit ihren bisher schon vorhandenen GKV Patienten) pro GKV Patient eGi erhalten, dann wäre der Umsatz dieser Praxis als einer der Praxis i entsprechenden (bei weiterhin N_i Patienten) NP

(6) $\hat{E}_i^S = N_i e_G = E_i/\lambda_i$

Nach Gl. 1 gehen also folgende praxisspezifische (jeweils für die Praxis i geltende) Größen in die Betrachtung ein: κ_i und λ_i, wobei für λ_i die Größen p_i und α_i bestimmend sind. Und die Standardisierung beruht

- *ausschließlich* auf *praxisspezifische* (praxisindividuelle) *Größen*[15] κ_i und λ_i nicht auf irgendwelche aggregierte (praxisübergreifende, für mehrere oder alle Praxen geltende konstante) Größen, etwa entsprechende gewogene oder ungewogene Mittelwerte $\overline{\kappa}$ oder $\overline{\lambda}$, die sich aus den ZiPP-Daten errechnen lassen und

- auf *Definitionsgleichungen* mit denen rechnerisch der Übergang $E_i \rightarrow \hat{E}_i^S$ vollständig aufgezeigt und begründet wird. Wir kommen damit zu einem dritten und den wohl wichtigsten Kriterium für eine Standardisierungsmethode

3 Alle Schritte im Rahmen eines Standardisierungsverfahrens sollten explizit in einem vollständigen System von Gleichungen dargestellt werden, das den Berechnungs- und Begründungszusammenhang des Verfahrens liefert. Ein Standardisierungsverfahren darf keine "black box" sein, in der mit den eingegebenen Zahlen (d. h. den Daten der erhobenen Praxen) etwas Undurchsichtiges geschieht.

Es sollte z. B. nicht nur klar sein, wie α_i mit λ_i zusammenhängt, sondern auch warum αi und nicht irgendeine andere Größe α (anstelle von α_i in λ_i) im Rahmen der Gleichung $\hat{E}_i^S = E_i/\lambda_i$ den Übergang $E_i \rightarrow \hat{E}_i^S$ erlaubt. Der Berechnungs- und Begründungszusammenhang der λ-Methode ist mit den Gleichungen (2) bis (6) hier vollständig wiedergegeben. Im Anhang wer-den noch einige weitere Begründungen und Implikationen angeführt. Nur wenn ein solcher Zusammenhang in Form von aufeinander bezogenen Gleichungen vollständig dargestellt wird, lässt sich ein Standardisierungsverfahren beurteilen, vor allem auch im Hinblick auf das

12 Das ist dann Gl. 11 in dem λ-Papier.

13 Man kann natürlich auch die Relation $e_{Pi} = E_{pi}/N_{pi}$ bilden, auch wenn Epi unter anderem auch Einnahmen um-fasst, die gar nicht auf die Behandlung von Privatpatienten zurückzuführen sind.

14 Ganz im Unterschied z.B. zur Gl. 8 in Abschn. 6 unten.

15 Man erkennt das daran, dass sie ein Subskript (tiefgestellter "Index") i haben.

wichtige Kriterium 1, ob nämlich das Rechenergebnis tatsächlich als die Gesamteinnahme der NP (so wie es die Definition der NP verlangt) zu verstehen ist.

So einen Berechnungs- und Begründungszusammenhang mit allen hierfür erforderlichen Gleichungen für ein Standardisierungsverfahren habe ich im Kohler-Papier an keiner Stelle gefunden.

Das hier geforderte Gleichungssystem (wie Gl. 1 bis 6) muss auch zeigen, welche Praxisdaten das Standardisierungsverfahren für welche Berechnung benötigt werden. Wir brauchen beispielsweise A_i für κ_i und Patientenzahlen und Einnahmen nach GKV und Nicht-GKV gegliedert für λ_i. Die Rolle von κ_i und auch p_i (im Rahmen von λ_i) der konkreten Praxis i ist klar. Der absolute trouble maker ist hier nur α_i (als Element in λ_i). Diese Größe beschäftigt uns seit Jahren. So wird z. B. immer wieder gefragt,

- wie das α_i in Gl. 4 und 5 mit Wasems[16] α zusammenhängt, und
- ob man nicht die α_i (also $\alpha_1, \alpha_2, \dots, \alpha_n$) in Gl. 5 durch ein mittleres α oder durch einen per Regressionsrechnung bestimmtes α ersetzen könnte oder sollte, das dann für alle n Praxen gelten soll (vgl. hierzu auch Abschn. 7)

Es sollte eigentlich klar sein, dass in beiden Fällen dann nicht mehr der obige Berechnungs- und Begründungszusammenhang gilt und dass dann auch das ganze Verfahren neu durchdacht werden muss. Würde das ZI die α_i durch eine Konstante α ersetzen, dann würde es eigentlich gar nicht mehr mit der λ-Methode rechnen (und Gl. 6 nicht mehr gelten) und es wäre von Grund auf neu zu überlegen, was die so durchgeführte Berechnung eigentlich aussagt[17].

Es gibt auch viele Gründe dafür, dass nicht zu erwarten ist, dass man noch das Gleiche erhält wenn man die n Größen α_i durch eine ganz anders definierte Größe α aus einer Studie mit ganz anderen Daten und einer ganz anderen Zielsetzung ersetzt[18].

Auf Unterschiede zwischen dem α_i in Gl. 4 und Wasems α bin ich im Anhang zum λ-Papier eingegangen. Was αi beinhaltet hängt davon ab, welche Einnahmen unter E_P und E_G fallen. Entscheidend für e_P (im Verhältnis zu e_G) sind nicht nur Vergütungsunterschiede für vergleichbare Abrechnungspositionen (nach EBM und GOÄ), die zu quantifizieren das primäre Ziel von WGMW war, sondern auch von PKV und GKV geleistete Zahlungen, die nicht zum sog. "standardisierten Leistungsniveau" (d. h. zu den vergleichbaren Leistungen) gehören, oder die auch von den entsprechenden Patienten selbst und nicht von ihrer Versicherung bezahlt wurden.

Enthält E_P alle Einnahmen außer E_G, dann wird α_i auch bestimmt von Einnahmen, die gar nicht mit der Behandlung von Patienten zusammenhängen. Man kann deshalb wohl auch nicht α_i mit der Fallwertrelation gleichsetzen.

Da e_G und e_P aus Umsätzen (Produkt von Preis und Menge) hergeleitet wird, ist auch die Mengenkomponente entscheidend, während es WGMW vor allem darum ging, die reine Preiskomponente (quasi für einen einheitlichen Warenkorb) herauszuarbeiten. Dagegen wird "unser" α_i auch davon bestimmt, wie oft bzw. selten Privatpatienten im Vergleich zu Kassenpatienten zum Arzt gehen und welche evtl. andere abrechnungsfähige Behandlungen sie dabei erfahren. Bei so vielen Unterschieden ist es eigentlich abwegig, die n verschiedenen α_i durch ein

16 Vgl. Anke Walendzik, Stefan Greß, Maral Manougian und Jürgen Wasem (kurz WGMW), Vergütungsunter-schiede im ärztlichen Bereich zwischen PKV und GKV auf Basis des standardisierten Leistungsniveaus der GKV und Modelle der Vergütungsangleichung, Diskussionsbeiträge aus dem Fachbereich Wirtschaftswissenschaften Universität Duisburg-Essen, Campus Essen, Nr. 165, Febr. 2008. Die Arbeit verfolgte ein ganz anderes Ziel als das der Standardisierung, so dass auch ein ganz anderes α im Fokus war.

17 Neben dieser sehr offensichtlichen Konsequenz sind auch andere Dinge zu beachten, auf die ich v.a. in Abschn. 8 eingehen werde.

18 Dazu mehr in Abschn. 7. Wer empfiehlt, die α_i durch ein externes α zu ersetzen, sollte sich auch die Mühe machen (im Sinne von Kriterium 3), zu zeigen, wie sich dadurch die Ergebnisse der Standardisierung verändern.

externes α aus der Arbeit von WGMW, in der es gar nicht um die Aufgabe der Standardisierung ging, substituieren zu wollen.

5. Regressionsrechnungen

5.1 Déjà vu

Im ZI-Papier findet man auf S. 22 unter dem Titel "Regressionsanalytischer Ansatz" die Gleichung

(7) $$\frac{E}{Z} = \beta_0 + \beta_1 U_{Ap} + \sum \beta_F d_f + u_i \,,$$

Ohne den offenbar Dummy Regressoren betreffenden Teil $\sum \beta_F d_f$ wird aus (7)

(7a) $$\frac{E}{Z} = \beta_0 + \beta_1 U_{Ap} + u_i$$

und im ZI-Papier konzentriert sich die Aufmerksamkeit dann auf die Größe

(7b) $$\tilde{\alpha} = \frac{\beta_0}{\beta_0 + \beta_1}$$

(genau genommen müssten es natürlich die geschätzten Regressionskoeffizienten sein, also $\tilde{\alpha} = \left(\hat{\beta}_0 + \hat{\beta}_1\right) / \hat{\beta}_0$), die als *das* praxisübergreifende α interpretiert wird; denn dann ist $\hat{\beta}_0$ quasi die Produktivität (Effizienz) der Arztminute (oder Arztstunde, je nachdem, wie die Zeit in $Z_i = A_i$ definiert ist) einer NP, die nur GKV Patienten hat (weil dann $p_i = 0$ und damit auch $U_{Ap} = 0$ ist), und $\hat{\beta}_0 + \hat{\beta}_1$ ist dann die Produktivität bei einer reinen PKV-Praxis.

5.2 Alte und neue Kritik am "regressionsanalytischen Ansatz"

Schon im α Papier von 2011, dann aber wieder in meiner Notiz 2 (2014) habe ich auf folgende Punkte hingewiesen,

1. Wenn $\hat{\beta}_0$ die Produktivität der Arztminute einer NP ist und diese NP definitionsgemäß einen Arbeitseinsatz von A*=140.148 Minuten hat, dann ist die Zielgröße der Standardisierung, also der absolute (in € gemessene) Umsatz der NP gegeben mit

   (7c) $$A^* \cdot \hat{\beta}_0 \,.$$

   Das ist aber eine Konstant im Unterschied zu $\hat{E}_i^S = N_i e_{Gi}$, was nach Gl. 6, bzw. 1 (wenn dort κ_i = 1 ist) gerade keine Konstante ist, weil N_i und e_{Gi} keine Konstanten sind.

   Eine Standardisierung, die bei jeder Praxis i (ganz unabhängig davon, wie groß E_i, N_i, α_i, p_i oder auch die Arbeitszeit A_i und damit κ_i sind) immer zum gleichen konstanten Ergebnis $A^* \hat{\beta}_0$ führt kann nicht sehr sinnvoll sein.

   Ganz abgesehen davon: wie rechtfertigt man, dass $A^* \hat{\beta}_0$ tatsächlich der gewünschte Um-satz der NP ist, wo doch in der Definition der NP nur Bezug genommen wird auf die Ar-beitszeit A*, nicht aber auf eine bestimmte Höhe $\hat{\beta}_0$ der Arbeitsproduktivität?

2. Von der Regressionsfunktion der Gl. 7a interessieren nur zwei Punkte, nämlich $U_{Ap} = 0$ und $U_{Ap} = 1$. Wie die Kurve zwischen den beiden "Endpunkten" verläuft ist egal. Es könnte auch ein nichtlinearer Zusammenhang sein.

3. Es gibt gute Gründe, die dafür sprechen, dass die Regressionsfunktion (Gl. 7, bzw. 7a) schlecht spezifiziert ist. Es ist praktisch eine Produktionsfunktion. Aber woher weiß man, dass E/Z abhängig ist von U_{Ap} und nicht vielleicht sehr viel mehr von anderen Regressoren, die vergessen wurden (omitted variables), dass noch dazu der Zusammenhang *linear* ist und dass der Regressor U_{Ap} nicht mit der Störgröße (u_i in Gl. 7) korreliert ist?

Im α Papier von 2011 habe ich auch gezeigt, dass sich die Annahmen hinter Gl. 7 nicht vertragen damit, was implizit bei der Zeitstandardisierung (Gl. 11) vorausgesetzt wird, dass dann nämlich E/Z zufällig um eine Konstante schwanken sollte. Aber das ist im Rahmen dieses Ansatzes vielleicht auch gar nicht relevant, weil es ohnehin auf κ_i in Gl. 7a nicht ankommt und Zeit- und Strukturstandardisierung offenbar in einem Aufwasch gemacht wird

5.3 Generelle Probleme mit dem Regressionsansatz

Nicht nur das alles ist seit Herbst 2011 unbeachtet geblieben, auch naheliegende Argumente, die ganz generell gegen den Regressionsansatz sprechen hat man offenbar nicht gesehen:

1. Hat man eine Konstante, wie etwa $\tilde{\alpha}$ bestimmt, so ist zu fragen, wie man mit dem so gewonnenen weiterrechnet, um die Umsätze Ei zu standardisieren. Man kann ja nicht einfach alle E_i (also $E_1, E_2, \ldots, E_n$) durch das gleiche $\tilde{\alpha}$ dividieren, d. h. es fehlt der mit dem Kriterium 3 geforderte vollständige Berechnungs- und Begründungszusammenhang in Form von entsprechenden Gleichungen. Bei keinem der in Abschn. 7 vorgestellten Ansätze des Kohler Papiers, bei denen ja i.d.R. mit praxisübergreifenden Konstanten wie $\tilde{\alpha}$ etc. findet man bei hierzu viel Konkretes (eine Gleichung zur Gewinnung von $\hat{E}_i$).

   Wo sind die Gleichungen, in denen $\tilde{\alpha}$ oder ähnliche Größen auftauchen und mit denen die E_i umgerechnet werden zu $\hat{E}_i$ und welche Größen der konkreten Praxis spielen dann in den Gleichungen zum Übergang $E_i \rightarrow \hat{E}_i$ eine Rolle? Kommt es in ihnen neben auch auf A_i (Arbeitszeit) oder pi der konkreten Praxis i an, oder spielen solche Größen nur für die Schätzung eines Parameters eine Rolle, danach aber nicht mehr?

2. Weil man alle diese praxisspezifischen Größen wie Ai, pi oder auch Ni und ui (= UAp bei Kohler) auch zur Schätzung bestimmter Konstanten benutzen kann (so wie U_{Ap} in Gl. 7a zur Schätzung von ß0, ß1, und damit von $\tilde{\alpha}$) könnte man sich sehr viele weitere unterschiedliche Regressionsgleichung ausdenken. Warum sollte E/Z (in meiner Notation E_i/A_i) nicht z. B. von der Praxisgröße (Patientenzahl) N_i oder von der strukturbeschreibenden Größe p_i, statt von U_{Ap} abhängen wie in Gl. 7a.

   Anders gesagt: bedient man sich der Regressionsrechnung um das Problem der Standardisierung zu lösen, kann man es evtl. endlos weiter mit irgendwelchen neuen mehr oder weniger sinnvoll erscheinenden Regressionsgleichungen versuchen (dazu mehr unten in Abschn. 9). Arbeitet man dagegen soweit wie möglich nur mit Definitionsgleichungen hat das Spiel durchaus bald einmal sein Ende und der Kreativität und Phantasie sind sehr viel engere Grenzen gesetzt.

3. Wie berücksichtigt man den Umstand, dass die Regression je nach Datenlage ein *unterschiedlich großes Bestimmtheitsmaß* r^2 bedeuten mag (das Modell also mehr oder weniger gut angepasst ist; wir operieren ja nicht mehr wie bei dem bisherigen Standardisie-

rungsverfahren allein mit Definitionsgleichungen, die immer gelten)?

Soll man den Regressionsansatz nur dann verfolgen wenn r^2 hinreichend groß ist (was heißt "hinreichend"?) oder soll man ihn nicht wählen wenn dies nicht der Fall und ab einem wie kleinen r^2 (wenn $r^2 < 0{,}3$ oder erst wenn $r^2 < 0{,}2$ ist?) sollte man auf den Regressionsansatz verzichten?

6. Der "Hebesatz"vorschlag und Mittelwerte der α_i

6.1 Zwei Verfahren und zwei Hebesätze

Hebesätze kommen im ZI-Papier an zwei Stellen vor.

1. Auf S. 22f heißt es, in die Hebesatzberechnungen mit der abhängigen Variable y "Umsatz pro Arzt" gingen die Regressoren

   - Anteil Privateinahmen und
   - die Jahresarbeitszeiten aller Ärzte
   - sowie fünf Dummy Regressoren ein.

2. In einem offensichtlich anderen Sinn ist auch vom Hebesatz im Rahmen der "arbeitszeitba-sierten Strukturstandardisierung" die Rede. Es heißt dort (S. 25): in den Hebesatz der Praxis (offenbar anders als eben eine praxisspezifische Größe), nennen wir ihn H_i "fließen … folgende … Merkmale ein"

   - Wochenarbeitszeit der Ärzte, gegliedert nach Patientengruppen (erhoben auf Arzt- und Inhaberebene) offenbar t_{PKV} und t_{GKV} mit $ß = t_{PKV}/t_{GKV}$ in den folgenden Gleichungen,
   - Anzahl der GKV-Patienten (erhoben auf Praxisebene) AGKV (bei mir N_{Gi})
   - GKV-Einnahmen aus dem Finanzteil (Praxisebene) U_{GKV} (bei mir E_{Gi})sowie
   - vollständige Angaben zu den weiteren Einkommensbestandteilen (Praxisebene)

Wenn letztere erhoben werden und somit "einfließen" fragt es sich, warum in den folgenden Gleichungen viel Mühe darauf verwendet wird, eine Größe $\hat{U}_{zGKV}$ (was als Umsatz zusätzlich zu GKV Umsätzen definiert ist) und $\hat{A}_{zGKV}$ zu schätzen (weil Patienten der Praxis i, die bisher Privatpatienten waren in der entsprechenden NP rechnerisch zu GKV werden). Aus den vier Gleichungen auf S. 25 folgt

$$(8) \qquad H_i = \frac{\hat{U}_{ges}}{U_{ges}} = 1 - \frac{U_{PKV}}{U_{ges}} + \frac{\beta U_{GKV}}{U_{ges}},$$

was inhaltlich zu interpretieren und zu begründen nicht ganz einfach sein dürfte. Aus Gl. 8 und der verbalen Erläuterung folgt, dass H_i (nach unserer Notation $\hat{E}_i / E_i$) so etwas wie $1/\kappa_i\lambda_i$ sein müsste. Dazu passt auch die verbale Beschreibung (S. 23), die sich jedoch auf den – wohl kaum gleichen – Hebesatz unter 1) bezieht "Der berichtete Hebesatz drückt das Verhältnis zwischen der geschätzten Erlössituation einer "reinen" GKV-Praxis zu den tatsäch-lich erzielten Erlösen der Praxis aus." (Im gleich darauf folgenden Satz heißt es "Das entspricht der relativen Differenz aus erzieltem Jahresüberschuss und standardisiertem Jahresüberschuss, bezogen auf den Letzteren", was eher die Definition $\left(U_{ges} - \hat{U}_{ges}\right)/\hat{U}_{ges}$ vermuten lässt. Wenn der erste Satz gilt und auch angesichts Gl. 8 überrascht es allerdings, dass man in Tab. 8 Werte findet, die größer als 1 sind, und z. T. größer als 5 oder 6 bei den Neurologen (Ausreißer?) und der Hebesatz in einem Fall sogar negativ ist.

Dies soll hier nicht weiter vertieft werden, zumal unser Ziel hier nicht eine Kritik des ZI-Papiers ist, sondern die Entwicklung von Gütekriterien für ein Standardisierungsverfahren. Es zeigt sich hier, wie wichtig das Kriterium 2 ist: ohne einen exakten und vollständigen Berech-nungs- und Begründungszusammenhang lässt sich ein Standardisierungsverfahren nicht beur-teilen und oft noch nicht einmal verstehen. Kriterium 2 verlangt, dass klar ist, mit welcher Formel bei einer neuen Methode die Umrechnung $E_i \rightarrow \hat{E}_i$ erfolgt, und wenn dann offenbar nicht mehr $\hat{E}_i = E_i / \kappa_i \lambda_i$ gilt, sollte auch klar sein was κ_i und λ_i entbehrlich macht oder sogar eine Verbesserung gegenüber κ_i und λ_i darstellt?

6.2 Mittelwerte: Was ist *das* "ZI-alpha" im Unterschied zu dem alpha von Wasem?

In Tab. 28 auf S. 28 findet man Zahlen über *die* im ZiPP verwendeten Aufschlagfaktoren α (im Vergleich zum Aufschlagfaktor nach Wasem). Das wirft die Fragen auf, wie *der* (offenbar praxisübergreifende) "ZiPP Aufschlagfaktor" α aus den praxisindividuellen α_i bestimmt wur-de. Wenn dies durch eine Mittelung der α_i erfolgt ist, ist zu beachten, dass es viele gleicher-maßen sinnvolle Arten gibt, von den α_i zu dem aggregierten α zu gelangen.

In meinem α-Papier (S. 1-8) habe ich herausgearbeitet, dass es neben dem ungewogenen Mittel $\overline{\alpha} = \sum_i \alpha_i / n$ noch mindestens sechs andere gewogene Mittelwerte der α_i gibt, die alle sehr viel sinnvoller zu interpretieren sind als $\overline{\alpha}$. Mögliche Gewichte sind z. B. $N_i/\Sigma N_i$ (große Praxen – gemessen an der Patientenzahl – fallen mehr ins Gewicht als kleine Praxen) oder auch $p_i/\Sigma pi$ usw. Das soll hier nicht vertieft werden, weil solche aggregierte Größen zwar als solche eine große Bedeutung haben mögen, für ein Standardisierungsverfahren aber nicht erforderlich sind, oder gar – wie jetzt zu zeigen ist – kontraproduktiv wären.

7. Die Standardisierung muss allein auf praxisindividuelle Größen abstellen

Das folgende Beispiel soll zeigen, dass es nicht darum gehen kann, *welches* konstante α (ein Mittelwert oder eine mit einer Regression geschätzte Konstante) unterschiedliche α_i (i = 1,…,n) substituieren sollte, sondern dass man überhaupt nicht mit solchen Konstanten (wel-che auch immer) anstelle von α_i operieren sollte. Es zeigt zugleich, welche Überlegung hinter der Strukturstandardisierung mit λ_i steht. Dazu nehmen wir zwei Praxen an:

i	NP	NG	EPi	EGi	Ei	ai	li
1	2	8	2*22 = 44	8*10= 80	124	22/10 = 2,2	1,24
2	2	8	2*24 = 48	8*10= 80	128	24/10 = 2,4	1,28

und rechnen einmal mit den beiden unterschiedlichen α_i (und λ_i) und dann mit einem für bei-de Praxen gleichermaßen extern vorgegebenen $\overline{\alpha}$ von 2,3 und λ = 1+(2,3 – 1)0,2 = 1,26). Mit den praxisindividuellen λ_i von 1,24 und 1,28 erhält man wegen (Gl. 6) für beide Praxen den gleichen Wert $\hat{E}_1^S$ = $\hat{E}_2^S$ = 100, aber mit $\overline{\alpha}$ = 2,3 und λ = 1,26 sind die standardisierten Einnahmen aber unterschiedlich, nämlich

bei Praxis 1 $\hat{E}_1^S$ = 124/1,26 = 98,41 und bei Praxis 2 $\hat{E}_2^S$ = 128/1,26 = 101,59.

Die Frage ist nun: Ist es fair, wenn die Strukturstandardisierung bei Praxis 1, die wegen (pro PKV Patient) relativ gering vergüteter Nicht-GKV-Einnahmen e_P = 22 einen geringeren Gesamterlös E_1 = 124 < E_2 = 128 erzielt hat auch zu einer geringeren standardisierte Einnahmen ($\hat{E}_1^S$ = 98,41 < $\hat{E}_2^S$ = 101,59) führt?

Im Beispiel ist E_1 = 124 ja vielleicht allein deshalb kleiner als E_2 = 128, weil Praxis 1 bei gleich vielen, nämlich 2 Privatpatienten (besser: Nicht-GKV-Patienten) weniger sonstige Einnahmen (PKV und andere Einnahmen) erzielt hat als Praxis 2. Warum sollten deshalb auch die standardisierten Einnahmen bei Praxis 1 geringer sein als bei Praxis 2, wo doch ansonsten die beiden Praxen hinsichtlich aller Merkmale der NP gleich sind?

Was sind aber die Merkmale der NP? Es ist nur p = 0 gefordert, dass es also nur GKV Patienten gibt, nicht aber wie viele es sind. Daraus allein ergibt sich noch nicht eine (fiktive) Einnahme. Entscheidend ist auch wie viele GKV Patienten die Praxen haben: wenn $N_1=N_{G1}$ = 10 und $N_2=N_{G2}$ = 12 wäre, dann wäre es gerechtfertigt, dass $\hat{E}_1^S < \hat{E}_2^S$.

Aber auch dann wenn, wie in dem Beispiel N_1 = N_2 = 10 ist, ergibt sich daraus noch keine Einnahme (kein Umsatz). Es muss auch angenommen werden, wie viel eine Praxis pro GKV Patient im Schnitt verdient. Bei der Praxis i ist das e_{Gi} und die Vorgabe ist ja "was wäre der Umsatz E, wenn *unter sonst gleichen Umständen* der Privatpatientenanteil p = 0 wäre?"

Dem entspricht Gl. 6 als Ergebnis der λ-Methode, nämlich $\hat{E}_i^S = N_i e_G$. Im Beispiel gibt es überhaupt keinen Grund warum nicht $\hat{E}_1^S$ = $\hat{E}_2^S$ = 100 sein sollte, wo doch die beiden Praxen

- gleich groß sind (N_1 = N_2 = 10)
- gleich strukturiert sind (Privatpatientenanteil p_i jeweils 20%) und sie auch
- für jeden GKV Patienten einen gleich großen Betrag e_{G1} = e_{G2} = 10 erlösen;

denn schließlich ist ja die NP durch die Einnahme N_{eG} bestimmt und nicht durch sonstige Einnahmen E_P, oder dadurch, wie ertragreich diese gemessen an E_P/N_P = e_P sind. Das Verhältnis α_i = e_{Pi}/e_{Gi} interessiert nur für E_i, aber nicht für $\hat{E}_i^S$. Wenn es keine Nicht-GKV Patienten gibt, kann man auch nicht mehr im Schnitt e_{Pi} = $\alpha_i e_{Gi}$ für jeden von ihnen verdienen.

Die NP ist nicht durch ein bestimmtes e_P und damit auch ein bestimmtes α gekennzeichnet, sondern durch einen Anteil p der Privatpatienten von p = 0. Ein konkreter Wert von e_P und damit auch von α > 1 ist nicht Teil der Definition der NP.

Die Berücksichtigung von $e_{Gi} \neq e_{Gj}$ in $\hat{E}_i^S = N_i e_G$ dagegen ist wie die Berücksichtigung unterschiedlicher Produktivitäten oder unterschiedlicher Arbeitszeiten zu sehen. So wie sich Unterschiede $A_i \neq A_j$ bei der Zeitstandardisierung auswirken, so auch jetzt Unterschiede bei e_G. Ist e_{Gi} der Praxis i um 20% höher als e_{Gj} der Praxis j (e_{Gi} =1,2e_{Gj}) so sind doch auch N_{Gi} = 10 Patienten in i mit N_{Gj} = 12 Patienten in j gleichzusetzen.

Daraus folgt: allein die Rechnung mit den individuellen λi Werten, die 124/1,24 = 128/1,28 = 100 ergibt, ist zu rechtfertigen, nicht das ungleiche Ergebnis, das man mit einem irgendwie aggregierten oder durch eine Regressionsrechnung ermittelten für beide Praxen gleichen α erhält. Es ist damit als viertes Kriterium zu fordern:

4 Bei der Standardisierung und Umrechnung $E_i \rightarrow \hat{E}_i$ sollte man mit praxisindividuellen und nicht mit aggregierten (für alle Praxen gleichen) Größen rechnen (also mit n Größen α_i statt mit nur einer, wie immer geschätzten oder gemittelten Konstanten α).

In gleicher Weise haben wir auch bei der Zeitstandardisierung über κ_i in $\hat{E}_i^Z = E_i / \kappa_i$ die Bezugnahme auf die praxisindividuelle (praxisspezifische) Arbeitszeit A_i und nicht auf ein konstantes $\overline{\kappa} = \overline{A}/A^*$ mit dem man dann $\tilde{E}_i^Z = E_i / \overline{\kappa}$ erhielte und damit Abweichungen $\hat{E}_i^Z \neq \tilde{E}_i^Z$ weil $A_i \neq \overline{A}$ ist. Sollen diese Abweichungen einfach unter den Tisch fallen? Sie sollten es nicht, wo doch der Sinn der Zeitstandardisierung gerade der ist, dass die Abweichung der konkreten Arbeitszeit A_i der Praxis i von der Konstanten A* nicht unter den Tisch fallen, sondern (über ein $\kappa_i \neq 1$) berücksichtigt werden sollen.

Um noch einmal auf die Probleme mit α_i im Rahmen der λ-Methode zurückzukommen: Die Bezugnahme auf α in $\lambda_i = 1 + (\alpha_i - 1)pi$ mag Vorteile für die Interpretation von λ haben, aber man kann sie auch ganz vermeiden, gerade weil α zu so vielen Diskussionen Anlass gibt; denn es gilt auch

$$\lambda_i = \frac{1-p_i}{1-u_i} = \frac{\text{Anteil GKV - Patienten}}{\text{Umsatzantel GKV}} \quad \text{und} \quad \alpha_i = \frac{u_i/(1-u_i)}{p_i/(1-p_i)},$$

und es kann nicht oft genug betont werden, dass Aufschlagfaktor (α) und Standardisierung zwei ganz verschiedene Themen sind. Schon im λ-Papier (2011) habe ich betont:

"Für die Implementierung der λ-Methode in der Praxis ist es *wichtig, dass man die Stan-dardisierung der Praxiseinahmen E_i mit dem für diese Praxis i individuellen λ_i durchführt*. Es sollte unbedingt beachtet werden, dass man nicht von E_i zu $\hat{E}_i^S$ gelangt, wenn man nicht λ_i sondern einen Mittelwert $\overline{\lambda}$ verwendet." (S. 22)

8. Anlässe zur Revision eines Standardisierungsverfahrens

Man könnte meinen, dass ich als "Erfinder" der λ-Methode diese um jeden Preis verteidige und schon allein deshalb alle alternativen Methoden kritisiere. Der Eindruck wäre falsch. Es gibt Anlässe, über ein anderes Vorgehen bei der Standardisierung nachzudenken, wenn

1) die gestellte Aufgabe verändert/erweitert wird, und wenn man

2) die für die Methode erforderlichen Daten nicht oder nicht in der erforderlichen Qualität hat.

zu 1: Bei einer Vortragsveranstaltung des ZI (20.11.2012) wurde z. B. argumentiert, dass bei der Bestimmung des Jahresüberschusses auch berücksichtigt werden sollte, dass die Behandlung von GKV und die von PKV Patienten unterschiedliche Kosten verursacht. *Wenn* das bei den Vorgaben über die NP zu berücksichtigen verlangt wird (was m. E, nicht der Fall ist) *dann* müsste auch das Standardisierungsverfahren entsprechend geändert werden.

Das wäre durch eine Erweiterung des ausschließlich auf Definitionsgleichungen aufbauenden "Basismodells" aus dem die λ-Methode abgeleitet wurde jederzeit möglich und ich habe hierzu entsprechende Überlegungen in meinem Kosten-Papier (2012) angestellt.

Wenn aber die Definition NP solche Differenzierungen nicht verlangt, wäre eine entsprechende Änderung der λ-Methode kontraproduktiv, weil jede "Verfeinerung" ja auch wieder problematisiert werden kann und hierfür auch wieder Daten geliefert werden müssen, die man vielleicht durch eine Befragung gar nicht erhalten kann.

zu 2: Auf diesen Punkt komme ich in Abschnitt 8.2 zurück.

8.1 Definition der Normpraxis NP

Ich habe aus gutem Grund das Kriterium 1 an die erste Stelle gesetzt. Wie das Standardisierungsverfahren aussehen sollte ist weitgehend eine Frage nach der Definition der Normpraxis NP und damit nach der evtl. noch weiter zu konkretisierenden Aufgabenstellung der Standardisierung.

1. Wenn die Definition der NP Bezug nimmt auf eine jährliche *Arbeitszeit von allen Ärzten* in einer Praxis in Höhe von 2.335,8 Stunden (140.148 Minuten), dann ist K auch genau diese so definierte Arbeitszeit zugrunde zu legen. Es wäre keine "Verfeinerung" oder "Weiterentwicklung" des Standardisierungsverfahren, wenn man mit einem K operieren würde, bei dem z. B. abweichend hiervon die Arbeitszeit der Inhaber herausgenommen wird, oder nur die Arbeitszeit der Inhaber betrachtet wird oder z. B. nur die Arbeitszeit an Patienten und nicht die gesamte Arbeitszeit zählt.

2. Wenn die NP so definiert ist, dass es auf die *Einnahmen* E (= Umsätze) ankommt, die man allein mit GKV Patienten erzielt (also eine *Gesamteinnahmen* die man hätte, wenn man nicht nur keine Privatpatienten, sondern auch keine sonstige Einnahmen wie IGeL, BG, Gutachten, Leistungen für Krankenhäuser usw.) hätte, dann muss in α auch E_P im Sinne von $E_P = E - E_G$ eingehen wobei E_G die Einnahmen mit GKV Patienten sind und E_P *alle* sonstigen Einnahmen umfasst (also nicht nur Einnahmen aus der Behandlung von Privatpatienten) und dann kann man nicht mit einem ganz anderes definierten α operieren, etwa mit einer Fallwertrelati-on (bei der – wie mir scheint – Einnahmen, die nicht auf die Behandlung von Patienten zu-rückzuführen sind nicht relevant sind).

Selbst wenn man es – aus welchen Gründen auch immer – für geboten hält, drei statt zwei Einkommenskategorien zu unterscheiden sollte man sich fragen, ob es noch Sinn macht, nur eine von ihnen durch ein λ zu dividieren, wo doch das λ in Gl. 5 so definiert ist, dass das *gesamte* Einkommen durch λ zu dividieren ist: was ergibt etwa $E_{GKV} + E_{PKV}/\lambda + E_{konstant}$?

3. Die Definition der NP verlangt m. W. nur eine Differenzierung in N_G GKV Patienten und $N_P = N - N_G$ andere Patienten, deren Anzahl bei der NP null sein sollte. Daraus ergibt sich, dass es nicht zielführend sein kann bei den Nicht-GKV Patienten auch nach BG/UV, PKV und Sonstigen zu unterscheiden. Es ist nicht einzusehen, dass die NP besser durch 0 + 0 + 0 statt durch 0 Nicht-GKV Patienten charakterisiert wird.

Dem steht nicht entgegen, dass die Definition der NP selbst problematisch sein mag und vielleicht auch als zu grob und undifferenziert oder zu vage und wenig exakt kritisiert werden kann. Es mag vielleicht auch gute Gründe geben, nicht auf den Jahresüberschuss *insgesamt* von 105.572 € abzustellen sondern auf den entsprechend geringeren Jahresüberschuss, den man allein aus der Behandlung von Patienten erzielt.

Aber das ist ein anderes Thema als das Thema "Standardisierung" und man sollte klar zwischen einer Weiterentwicklung der Definition der NP und einer Weiterentwicklung des Standardisierungsverfahrens unterscheiden.

Es ist insbesondere nicht zu erkennen, wie aus einer größeren oder geringeren Korrelation zwischen x_t und x_t-1 (etwa Patientenzahlen in Erhebungswelle t und in Erhebungswelle t-1 folgt, dass es geboten ist, in Zukunft bei der Standardisierung innerhalb von x weniger oder mehr zu differenzieren. Maßgeblich für Differenzierungen im Standardisierungsverfahren ist nach Kriterium 1 allein, was die Definition der NP verlangt und nicht die zeitliche Veränderung.

8.2 Datenprobleme

Auch Datenprobleme können Anlass geben, ein Standardisierungsverfahren zu überdenken und ggfls. entsprechend zu modifizieren. Es könnte z. B. sein, dass Arbeitszeiten, wie t_{PKV} und t_{GKV} besser zu erheben sind als Daten zur Struktur der Patienten, wie p_i und $1\text{-}p_i$ wie (we-niger Nichtbeantwortung, von Befragten weniger missverstanden und deshalb eher korrekt beantwortet, geringerer Stichprobenfehler usw.). Aber selbst wenn das so wäre, was mir zu glauben schwer fällt, wäre der Vorteil schnell verspielt, wenn man diese Daten in zweifelhafte Schätzgleichungen einbaut und es nicht gelingt, mit t_{PKV} und t_{GKV} einen Berechnungs- und überzeugenden Begründungszusammenhang herzustellen.

Auf der gleichen Ebene (was ist besser oder schlechter im Rahmen eines Standardisierungsverfahrens verwertbar) liegt auch die Wahl zwischen erhobenen und geschätzten Größen. Er-setzt man eine erhobene Größe wie α_i durch einen Mittelwert $\overline{\alpha}$ oder einen durch eine Regression bestimmten Wert $\hat{\alpha}$ so ist ja immer zu fragen, wie groß die Streuung der α_i ist und damit wie "repräsentativ" $\overline{\alpha}$ ist, bzw. wie groß das Bestimmtheitsmaß r^2 bei der Schätzung von $\hat{\alpha}$ ist.

Das Problem, dass die Anwendung einer Methode evtl. schwer zu erhebende Merkmale verlangt ist sehr ernst zu nehmen, aber man kann es i.d.R. nicht dadurch lösen, dass man zu einem Verfahren wechselt, das stattdessen auf problematische Schätzungen angewiesen ist.

Gleichwohl sind die folgenden drei Fragen zur Operationalisierung von Messkonzepten und Realisierbarkeit der Datenbeschaffung bei jeder Entscheidung über eine Standardisierungsmethode von ganz grundsätzlicher Bedeutung:

1. Liefert das ZiPP Daten auch für genau die in der Definition der NP implizierten Konzepte oder gibt es hier Abweichungen

2. Gibt es Gründe, die in der Befragungstechnik liegen, weshalb man evtl. mit der Standardisierung nicht genau das mit der Definition der NP gewünschte Ziel erreicht?

3. Kann ein anderes Standardisierungsverfahren evtl. dem Ziel der Standardisierung besser dienen weil es Variablen verwendet, für die Daten aus dem ZiPP leichter und besser verfügbar sind?

Wie gut eine Erhebung welche Sachverhalte wiedergibt und wie gut die Qualität (Repräsentativität, Validität) einer geschätzten Größe ist, muss von Fall zu Fall entschieden werden. Es lässt sich hier also kein allgemeines Qualitätskriterium formulieren.

9. Axiome

Dass es viele immer wieder neue Lösungsvorschläge gibt und der Eindruck entsteht, dass man nie zu einem Schluss kommt und sich vielleicht sogar im Kreise dreht kommt ist in der Wissenschaft nicht selten vor. In der Ma-thematik und Statistik ist das fast immer ein Indiz dafür, dass eine Aufgabenstellung noch nicht exakt genug beschrieben und vielleicht auch noch nicht ganz richtig verstanden ist. Wie kommt man dann weiter?

Abgesehen von dem Versuch, die Aufgabenstellung zu präzisieren also in unserem Fall die NP exakter zu definieren, was eine bleibende, hier an verschiedenen Stellen eingeforderte Aufgabe ist, und alles noch einmal gründlich zu durchdenken scheint es mir vorrangig zu sein, Standardisierungsverfahren an "Axiome" zu messen. Diese Verfahren sind letzten Endes auch nur Rechenverfahren, wie z. B. auch die Berechnung eines Mittelwerts oder eines Preisindexes.

Statistiker kennen viele Mittelwerte, nicht nur das arithmetische, geometrische und harmonische Mittel, auch den Median, Modus etc., oder ausgefallenere Mittel wie ein log mean, power mean etc. und sie haben viele hundert Formeln für die Berechnung eines Preisindexes eingeführt über die man endlos streiten kann und es ist nicht zu erwarten, dass sich irgend-wann einmal alle auf die ultimative Formel einigen können. Also braucht man Maßstäbe, mit denen man eindeutig entscheiden kann, ob ein Rechenverfahren "sinnvoll" ist, wobei aber "meaningful", wie es in der entsprechenden Literatur heißt, nicht sonderlich klar ist. Es muss also etwas sein, was eindeutig erfüllt oder nicht erfüllt ist. Das sind "Axiome".

Ein Beispiel ist: Ein Mittelwert M sollte (wie das Wort "Mitte" ja schon sagt) nicht kleiner als der kleinste und nicht größer als der größte einzelne x-Wert sein. Für ihn muss also gelten

(9) $x_{min} \leq M \leq x_{max}$

Wäre z. B. $x_1 = x_{min} < x_2 < x_3 \dots < x_n = x_{max}$, so wäre leicht zu zeigen, dass das arithmetische Mittel $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ als eine erste Möglichkeit für M dieses Axiom erfüllt. Eine andere Formel, etwa $M_2 = (x_1 - 2x_2 - 3x_3)/6$ würde es aber nicht erfüllen, weil $M_2 < x_{min}$ wäre. M_2 wäre also anders als $M_1 = \bar{x}$ keine "sinnvolle" Formel für einen Mittelwert. Und so könnte man viele andere Formeln für Mittelwerte M_3, M_4, … auch untersuchen. Es ist zu erwarten, dass gemes-sen an diesem Axiom als dem einzigen Axiom eine ganze Reihe Formeln akzeptabel wären (also braucht man meist mehrere Axiome, die sich natürlich nicht widersprechen oder redun-dant sein dürfen), aber diese Herangehensweise reicht zumindest aus, um mehr oder weniger unsinnige Lösungsvorschläge begründet abzulehnen.

Ein zweites und letztes Beispiel wären Formeln für einen Preisindex. Wir geben gleich zwei Axiome an, bei denen es einleuchtend ist, dass sie beide erfüllt sein sollten:

1. Identität. wenn sich kein Preis von 0 bis t verändert hat soll der Preisindex auch 1 betragen, also 100% anzeigen, und

2. Monotonie: er sollte nicht 1 sein, sondern kleiner (bzw. größer) als 1 sein, wenn *ein* Preis ceteris paribus (bei Konstanz aller anderen Preise) gesunken (bzw. gestiegen) ist.

Wichtig ist, dass ein Axiom exakt formuliert werden kann. Im Falle der Identität ist das einfach: gilt $p_{it}/p_{i0} = 1$ für alle i, dann muss auch der Preisindex $P_{0t} = f(p_{1t}/p_{10}, \dots, p_{nt}/p_{n0}) = 1$ sein.

Früher sprach man im Deutschen auch schon mal – wie in diesem Papier – von "Kriterien", was man jetzt "Axiome" nennt. Der Unterschied ist weitgehend der, dass Axiome exakt als Gleichungen formuliert sind, so dass man eindeutig entscheiden kann, ob sie erfüllt sind oder nicht, während Kriterien auch "nur" rein verbal beschriebene Anforderungen an eine gute Methode sein können.

In Abschn. 7 wurde bereits ein Kriterium angewendet, das man auch Axiom nennen könnte

A1: wenn $N_i = N_j$ und eGi = e_{Gj} sollte auch $\hat{E}_i^S = \hat{E}_j^S$ sein

und im Anhang zu diesem Papier findet sich die Aussage "Ist z. B. die Arbeitszeit einer Praxis i um 20% höher als die Normarbeitszeit A* (also κ_i = 1,2), dann darf die Einnahme E_i auch um 20% höher sein als die der NP."

A2: wenn $A_i = \theta A^*$ dann $E_i = \theta\hat{E}_i^Z$

Man kann darüber streiten, ob solche Axiome nicht zu einschränkend sind. So könnte z. B. kein anderer Wert als θ = κ_i das Axiom **A2** erfüllen, denn aus $E_i = \theta \hat{E}_i^Z$ folgt ja $\hat{E}_i^Z = E_i / \theta$ und damit auch θ = A_i/A* = κ_i. Welche Axiome man aufstellen will und wie tolerant sie sein sollen in dem Sinne, dass sie verschiedene Lösungen zulassen, ist natürlich eine inhaltliche Frage; denn es geht ja darum, welche Rechenergebnisse man als "sinnvoll" akzeptieren möch-te. Ob die mit der "Simulation der NP" beschrieben Aufgabenstellung dafür klar genug ist, sei dahingestellt. Es dürfte nicht nur schwierig sein, sich weitere Tests auszudenken, die nicht zu restriktiv sind (wie z. B. **A2** aber auch **A1**), sondern auch sich darauf als besonders "wichtige" oder "essentielle" Forderungen zu einigen. Die axiomatische Herangehensweise verlangt auch sehr große Vertrautheit mit dem anstehenden Messproblem, aber sie zwingt uns auch zu mehr Exaktheit und könnte helfen, die unbefriedigende Situation von immer neuen Messmethoden und Regressionsfunktionen zu überwinden:

> "Man kann sich die verschiedensten Koeffizienten ausdenken oder diverse Regressionsfunktionen schätzen. Das allein kann es nicht sein, schon allein deshalb nicht, weil aus dieser Vorgehensweise keine *Kriterien dafür* zu gewinnen sind, *welche Koeffizienten man* aus dem prinzipiell beliebig großen Angebot an verschiedenen Koeffizienten *auswählen soll*. Solche Kriterien können nur aus einem Begründungszusammenhang gewonnen werden. Plausibilität der Ergebnisse empirischer Anwendungen ist für sich genommen kein ausreichendes Kriterium. (λ-Papier, S. 35)

Man könnte auch sagen: die Aufgabe von Funktion ist es, solche Kriterien zu formulieren. Es war früher auch üblich, von "Tests" statt von Axiomen zu sprechen. Im Prinzip liefern die Zahlen für die beiden fiktiven Praxen in Abschn. 7 ein Material, um einen solchen "Test" A1 (hier synonym zu "Axiom A1" durchzuführen. Man kann sich alle möglichen Standardisierungsverfahren ausdenken, aber es könnten zumindest schon einmal einige als unbrauchbar ausscheiden, wenn man sie auf die Zahlen der beiden Praxen in Abschn. 7 anwendet und dann *nicht* $\hat{E}_1^S = \hat{E}_2^S$, sondern $\hat{E}_1^S \neq \hat{E}_2^S$ erhält.

5 Zur Beurteilung von Standardisierungsverfahren sollte man Axiome postulieren (exakt formulierte Anforderungen an das Rechenergebnis) und die Verfahren mit fiktiven Zahlenbeispielen (Tests) durchrechnen.

10. Analyse von Veränderungen mit Korrelationskoeffizienten

Bei einem Standardisierungsverfahren interessieren natürlich empirische Fakten darüber, wie sich dessen Bestimmungsgrößen (wie λ, K etc.) und Ergebnisse im Zeitablauf verändern. Da diese aber, wie in Abschn. 3 erklärt, nicht zentral für die Entscheidung über ein Standardisierungsverfahren sein sollten wiederhole ich im Folgenden deshalb nur kurz einige andernorts ausführlicher gemachte kritische Anmerkungen zum ZI-Papier.

Für Fragen nach Stabilität vs. Veränderung oder Instabilität (und deren Ursachen) sind Korrelationskoeffizienten r_x (Korrelation zwischen x_t und x_t-1) ") ungeeignet, weil

1. es keine einfachen Formeln für die Zusammenhänge zwischen r_x, und entsprechend definierten Korrelationen r_y, r_z usw. gibt (es wäre z. B. erheblich einfacher, mit Wachstumsfaktoren $f_{x,i} = x_{t,i} / x_{t-1,i}$ zu operieren)

2. die Annahme einer lineare Beziehung x_t = $ß_0$ + $ß_{1xt-1}$ + u_t (u: Störgröße), also einer Autoregression (erster Ordnung) der Variable x meist nicht zu begründen ist,

3. man z. B. bei Wachstumsfaktoren $f_{x,i}$ *zwei* zusammenfassende Größen hat,

- Mittelwert $\bar{f}_x$ (über i = 1,…,n Praxen gemittelte $f_{x,i}$) *und*
- die Varianz der $f_{x,i}$, definiert als $\mathrm{var}(f_x) = \frac{1}{n}\sum_i \left(f_{x,i} - \bar{f}_x\right)^2$,

(wobei die Varianz für alle analytischen Bemühungen die eigentlich interessante Größe ist) bei Korrelationskoeffizienten aber nur ein zusammenfassendes Maß, nämlich den Korrelationskoeffizient r_x (ein Maß, in das die x_t und x_{t-1} Werte aller n Praxen einfließen). Die var(f_x) kann mit Regressions- und Varianzanalyse additiv in einen "erklärten" und einen "residualen" Teil zerlegt werden. Man könnte z. B. mit einer Regressionsgleichung die n Werte der zu erklärenden Variable (des Regressanden) mit einem Regressor wie "erklären" aber es gibt keine Regressionsgleichung wenn links (auf der Seite des Regressanden) eine Größe ry steht statt n Größen $f_{y,i}$. Es gibt nichts der Varianzzerlegung Analoges wenn man Veränderungen mit Korrelationskoeffizienten r_x statt mit Wachstumsfaktoren $f_{y,i}$, $f_{x,i}$ etc. misst.

Zu1: Es heißt z. B. die GKV Fallwerte (dies sei x_t und x_{t-1}) korrelierten mit r_x = 0,877 und sie seien deshalb relativ stabil, während die PKV Fallwerte (dies sei y_t und y_{t-1}) wegen r_y = 0,628 weniger stabil seien. Ich wüsste nicht, wie man mit solchen Angaben auf die Korrelation (eigentlich Autokorrelation) der Fallwertrelation, die mit $z_t = y_t/x_t$ (und z_{t-1} entsprechend) definiert ist, also die Korrelation zwischen z_t und zt-1 schließen soll. Sie beträgt r_z = 0,83. Aber wie hängen r_x, r_y und r_z zusammen? Mit welcher Formel erhält man 0,83 aus 0,877 und 0,628 wo doch bei der Berechnung von 0,83 genau die gleichen Daten benutzt wurden, mit denen man auch 0,877 und 0,628 berechnet hatte? Rechnet man mit Wachstumsfaktoren ist einfach $f_{z,i} = f_{y,i}/f_{x,i}$.

Anhang 1

Zur Rechtfertigung und Herleitung des Standardisierungsverfahren mit K und λ

Ich habe wiederholt die hinter der Standardisierung mit κ_i und λ_i stehende "Logik" erläutert. Es erscheint mir sehr wichtig, die außerordentliche Schlichtheit der zugrundeliegenden Überlegung hier noch einmal ganz klar zu machen.

1. Zeitstandardisierung

κ_i ist definiert als Verhältnis der Arbeitszeit A_i (Jahresarbeitszeit in Minuten) einer Praxis i zur "Standardarbeitszeit" von A* = 140.148 Minuten, also $\kappa_i = A_i/A^*$. Da es um die Standardisierung der (Gesamt)Einnahmen E_i geht brauchen wir eine *Beziehung zwischen A_i und E_i*. Hierzu *nehmen wir einen linearen Zusammenhang* an (also quasi keinen abnehmenden, sondern einen konstanten Grenzertrag der Arbeitszeit A [das ist die einzige Annahme; ganz ohne geht es also auch hier nicht]), so dass man erhält

(10) $E_i = \gamma_Z + \beta_Z A_i$

Bemerkung zu Gl. 10

Man beachte dass Gl. 10 keine Regressionsgleichung ist, das wäre sie nur dann wenn

(11) $E_i = \gamma_Z + \beta_Z A_i + \varepsilon_i$

ist mit der Störgröße ε_i. In Gl. 11 sind γ_Z und $ß_Z$ Konstante im Unterschied zu γ_{Zi} und $ß_{Zi}$ d

Man beachte, dass Gl. 10 keine Regressionsgleichung ist.

Weil bei einem Arbeitseinsatz von Null Minuten (A_i = 0) auch die Einnahme E Null sein sollte folgt aus Gl. 10

(12) $ß_{Zi} = E_i/A_i$ und für die Normpraxis (NP) mit der Normarbeitszeit A*

(13) $\beta_Z = \hat{E}_i^Z / A^*$.
Es liegt nahe, die beiden Quotienten $ß_{Zi}$ gleichzusetzen und man erhält dann wenn $\kappa_i \neq 1$ (nur dann besteht ja ein "Bedarf" an Zeitstandardisierung)

(14) $\frac{E_i}{\kappa_i} = \beta_Z \frac{A_i A^*}{A_i} = \beta_Z A^* = \hat{E}_i^Z$ weil ja $\frac{1}{\kappa_i} = \frac{A^*}{A_i}$.
Man beachte: E_i wird durch das *praxisindividuelle* κ_i dividiert *nicht durch einen aggregierten* Parameter $\overline{\kappa} = \overline{A}/A^*$. Ist $\kappa_i > \overline{\kappa}$ (gleichbedeutend $A_i > \overline{A}$) oder speziell $\kappa_i = 1{,}2 \cdot \overline{\kappa}$ dann ist für $\hat{E}^Z = E_i/\overline{\kappa}$ auch größer als $\hat{E}_i^Z$ gem. Gl. 12. Es ist richtig mit dem individuellen (für die Praxis i geltenden) κ_i statt mit $\overline{\kappa} = K_j$ zu rechnen, denn Gleichheit $\hat{E}_i^Z = \hat{E}^Z = \hat{E}_j^Z$ wäre nur dann gerechtfertigt, wenn auch $E_i = 1{,}2E_j$. Das ergibt sich aus der "Logik" der Zeitstandardisierung, die beschrieben werden kann mit der Maxime:

Ist z. B. die Arbeitszeit einer Praxis i um 20% höher als die Normarbeitszeit A*, also $\kappa_i = 1{,}2$, dann darf die Einnahme Ei auch um 20% höher sein als die der NP, also $E_i = 1{,}2 \cdot \hat{E}_i^Z$.

Allein das ist der hinter der Zeitstandardisierung stehende Gedanke. Aus ihm folgt auch, dass mit praxisspezifischen Größe, wie κ_i statt mit festen praxisübergreifenden Größen wie $\overline{\kappa}$ gerechnet werden sollte. Bei der im Folgenden behandelten Strukturstandardisierung ist die Si-tuation im Grunde ganz ähnlich, aber es scheint schwieriger zu sein

- zu zeigen, was hinter α_i und λ_i steckt (weil nicht eine Größe, wie κ_i, sondern zwei Größen α_i und p_i für λ_i bestimmend sind), und
- warum auch hier mit praxisspezifischen Größen α_i und $\lambda_i = 1+(\alpha_i -1)p_i$ gerechnet werden sollte und nicht mit $1+(\overline{\alpha}-1)p_i$ oder gar $1+(\overline{\alpha}-1)\overline{p}$.

2. Strukturstandardisierung

Den Gleichungen $\hat{E}_i^S = E_i/\lambda_i$ und $\lambda_i = 1+(\alpha_i - 1)p_i$ liegt folgender Gedanke zugrunde:

Ei kann nur dann größer sein als $\hat{E}_i^S$ (oder $\lambda_i > 1$ sein) wenn die einzelne reale Praxis i

1. einen Privatpatientenanteil p_i von über Null hat ($p_i > 0$) und
2. das Verhältnis der Einnahmen $e_{pi} = E_{Pi}/N_{Pi}$ zu $e_{Gi} = E_{Gi}/N_{Gi}$ also $\alpha_i > 1$ ist.

Es gibt also zwei Bestimmungsgrößen für λ, nämlich α_i und pi

1. Anteil der Privatpatienten an den Patienten (p_i)

Kein Bedarf an Standardisierung ist gegeben, wenn die reale Praxis schon wegen pi = 0 gleich der Normpraxis ist. Ist nämlich $p_i = 0$, dann ist auch λ, = 1, und zwar auch dann wenn $\alpha_i > 1$ ist. Ist aber $p_i > 0$ fragt es sich: um wie viel kann bzw. darf E_i größer sein als $\hat{E}_i^S$? Das hängt natürlich davon ab, wie groß α_i ist und das folgt aus der Formel $\lambda_i = 1+(\alpha_i - 1)p_i$.

Bei gegebenen $p_i > 0$ kommt es also darauf an, um wie viel α_i größer ist als 1. Es ist also leicht einzusehen, dass es auf *beide* Größen ankommt, auf $p_i > 0$ *und* $\alpha_i > 1$. Der Zusammenhang zwischen E_i und p_i ist wegen $E_i = \lambda_i \hat{E}_i^S = \hat{E}_i^S[1+(\alpha_i - 1)p_i] = \hat{E}_i^S + \hat{E}_i^S(\alpha - 1)p_i$ auch nicht einfach linear (eine Gerade). Er wäre es auch dann nicht, wenn α_i eine Konstante wäre, also $\alpha_i = \alpha$, weil $\hat{E}_i^S$ (die Einnahme, die gerechtfertigt ist, wenn man keine Privatpatienten hat) keine Konstante ist.

2. "Pro-Kopf" Einnahmerelation (α_i)

Es ist auch unmittelbar einsichtig, dass $\alpha_i = 1$ keine Veranlassung gäbe, eine Standardisierung bezüglich der Versichertenstruktur vorzunehmen. Denn $\alpha_i = 1$ bedeutet ja nach Gl. 4, dass die pro-Patient Einnahmen e_P aus der Behandlung von PKV Patienten *und* auch aus anderen Tätigkeiten (es gilt ja nach der Definition $E_P = E - E_G$) nicht größer sind als die pro-Patient Einnahmen e_G aus der Behandlung von N_G GKV Patienten.

Nur das in Gl. 4 definierte von α_i liegt der Herleitung von λ_i und der Formel $\hat{E}^S = E_i / \lambda_i$ zugrunde und wegen $E_P = E - E_G$ (so dass E_P auch Einnahmen enthält die gar nicht mit der Behandlung von Privatpatienten zusammenhängen) ist auch klar, dass sich α_i auch dann ändern kann, wenn sich die entsprechenden Gebührenordnungen EBM und GOÄ gar nicht ändern. Man könnte viele ganz anders definierte Größen α_i konstruieren, aber dann müsste man auch mit einem anders definierten λ_i rechnen und zeigen (mit entsprechenden Definitionsgleichungen nach Art von Gl. 1 bis 6), dass man auch mit diesem anders definierten λ_i die korrekte Standardisierung ( ) durchführen kann.

Anhang 2

Die fünf Kriterien zur Beurteilung von Standardisierungsverfahren

1. Definition der Normpraxis (NP) als Maßstab

   Maßgeblich für ein Standardisierungsverfahren ist die Frage, mit welchen Merkmalen, und in welcher Differenzierung die NP definiert ist und ob man mit der Rechnung auch tatsächlich *diese* so definierte NP simuliert.

2. Beschränkung auf erhobene Daten

   Das Standardisierungsverfahren sollte
   - *allein* mit den *im ZiPP erhobenen Daten* der Praxen und
   - *nur* auf der Basis reiner *Definitionsgleichungen* operieren.

3. Keine black box

   Alle Schritte im Rahmen eines Standardisierungsverfahrens sind explizit in einen vollständigen System von Gleichungen darzustellen. Das Gleichungssystem liefert den <u>Berechnungs- und Begründungszusammenhang</u> des Verfahrens.

4. Keine (externe oder geschätzte) aggregierte Größen

   Bei der Standardisierung und Umrechnung E_i $\hat{E}_i$ sollte man mit praxisindividuellen und nicht mit aggregierten (für alle Praxen gleichen) Größen rechnen (z. B. mit n Größen α_i statt mit einer, wie immer bestimmten Konstante α).

5. Axiome und Tests (Beispielrechnungen)

   Zur Beurteilung von Standardisierungsverfahren sollte man Axiome postulieren (exakt formulierte Anforderungen an das Rechenergebnis) und die Verfahren mit fiktiven Zahlenbeispielen (Tests) durchrechnen.