La recherche concomitante, nécessaire et envisagée à l'occasion de l'introduction de SwissDRG, devra être effectuée de manière sérieuse pour être crédible. Par leur formation et leur expérience, les médecins ont des exigences scientifiques élevées, toute bonne pratique devant être basée sur l'évidence selon des critères rigoureux et validés. Tout naturellement, ils attendent de la recherche concomitante qu'elle suive les mêmes règles: l'article du Dr Romanens illustre bien ce propos.
Cependant, il faudra faire la distinction entre ce qui sera simplement évalué selon les prescriptions de la LAMal et ce qui fera l'objet d'une véritable recherche. Pour des raisons de temps et de coût, tous les sujets intéressants ne pourront pas être examinés selon un protocole de recherche complexe. Il faudra définir des priorités, probablement dans le domaine de la qualité. D'autre part, le corps médical doit conserver des ressources pour participer activement à la création d'une structure tarifaire SwissDRG qui reflète au mieux la réalité hospitalière de notre pays. En effet, s'il est adéquat d'en contrôler les effets de manière pertinente, il est tout aussi nécessaire de veiller à la construction d'un système cohérent.

Dr Pierre-François Cuénoud, Membre du Comité Central de la FMH, Responsable du domaine SwissDRG

Begleitforschung SwissDRG: Aufruf zu einem nationalen Konsens

Michel Romanens[a], *Franz Ackermann*[b], *Bernhard Hofmeier*[c]

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

b Dr. med., Vorstandsmitglied Stiftung Vasculäres Risiko (VARIFO)

c Dr. med. Vorstandsmitglied Verein Ethik und Medizin Schweiz (VEMS.CH)

Korrespondenz:
Dr. med. Michel Romanens
Ziegelfeldstr. 1
CH-4600 Olten
Tel. 062 212 44 10
Fax 062 212 44 30

info@kardiolab.ch

www.physicianprofiling.ch

Recherche concomitante des SwissDRG: l'appel à un consensus national

La recherche dont l'objectif est de recenser les effets négatifs du financement des hôpitaux lors de l'introduction des DRG et des nouveaux points de référence visant à comparer les hôpitaux en termes de performance devrait dès le début faire partie intégrante de cette recherche et appeler à un consensus national. Cela permettrait d'empêcher que le monitorage crée des incitatifs; ces derniers pourraient, par le biais de mécanismes de sélection des risques, conduire à une baisse de la qualité et de la sécurité des soins en Suisse. Il convient de mentionner deux points importants: l'absence d'une large discussion au sujet de la confusion dans la recherche et l'absence de validation des DRG au moyen de groupes de coûts pharmaceutiques (PCG).

Hintergrund

Vor dem Hintergrund der politisch verordneten Prämienexplosion sind die Kosten im Gesundheitswesen für viele Schweizerinnen und Schweizer zu einem finanziellen Problem geworden. Dies erhöht auch politisch betrachtet das Risiko, dass die Durchsetzung alter Forderungen zur Kostensenkung im Gesundheitswesen durchsetzbar geworden ist, insbesondere betreffend Kosten im Primary Care (Globalbudget) und Spitalschliessungen (politisch gewollte Reduktion der Baserate im DRG-System).

Diese einseitig auf Kostenreduktion ausgerichtete Budgetverantwortung für medizinisch begründete Interventionen entspricht einer falsch verstandenen «Ökonomisierung» der medizinischen Versorgung der Schweizerischen Bevölkerung. Dies birgt erhebliche Risiken, insbesondere betreffend Verminderung der Versorgungsqualität, Verminderung der Patientensicherheit, vermehrte versteckte oder offene Rationierung notwendiger medizinischer Leistungen oder Behandlungsverweigerung von Patientinnen und Patienten mit hohen Kostenrisiken.

Eine nie dagewesene Flut von Ratschlägen zum Monitoring des Gesundheitswesens wurde in den letzten Jahren von Gesundheitsökonomen, medizinischen Fachgesellschaften, Epidemiologen, Ethikern und der FMH publiziert [3–5, 6–12]. Dabei kann ohne Weiteres festgestellt werden, dass die Instrumente zur Erfüllung des Monitoringanspruchs mehrheitlich nicht validiert sind.

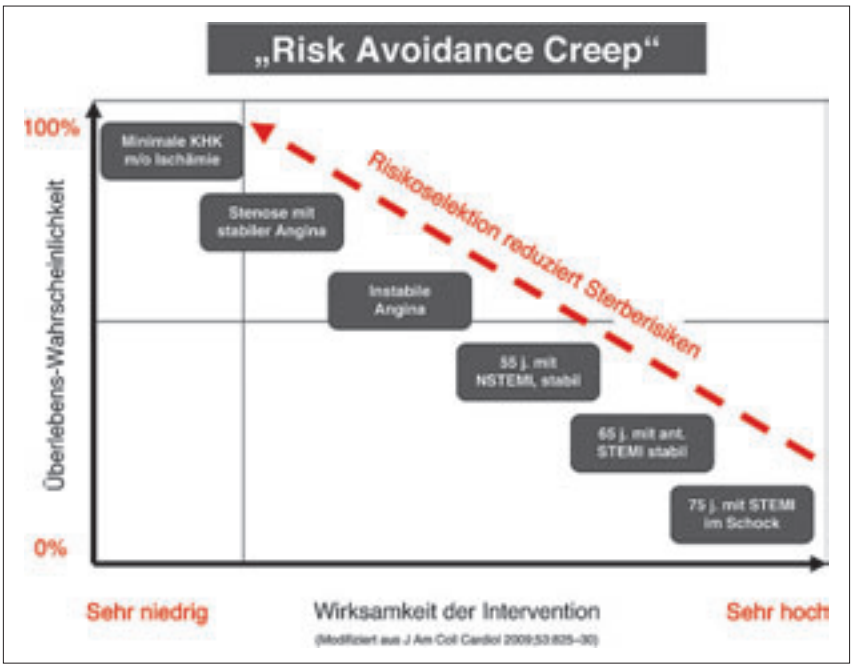
Das erklärte Ziel, Missstände als Folge von Globalbudgets und SwissDRG künftig – und notabene ohne Verankerung dieser Aktivitäten im Krankenversicherungsgesetz – aufzudecken, könnte jedoch weit verfehlt werden. Es ist sogar wahrscheinlich, dass die Monitoring-Instrumente – insbesondere unter der simplen Verwendung von Mittelwertbeobachtungen und meist nicht validierten und arbiträren Korrekturfaktoren – kontraproduktive Wirkung haben und zu einer Risikoselektion zum Schaden der Patientinnen und Patienten führen, z. B. mit dem Mechanismus des in der Abbildung gezeigten «Risk Avoidance Creep»: Möglichkeit der Risikoselektion in Abhängigkeit von vermuteten Sterberisiken (Level of Evidence: IIb).

Im ökonomischen Diskurs über Ressourcen im Gesundheitswesen ist jedoch der Einbezug von Kosteneffizienz -, Effektivitäts-, Qualitäts- und Gerechtigkeitsanalysen notwendig [6]. Damit schafft die Gesundheitsökonomie professionelle und vor allem auch sachliche Diskussionsgrundlagen für politische Entscheide.

Davon abzugrenzen ist eine qualitativ ungenügende Gesundheitsökonomie, insbesondere wenn diese den publizierten Richtlinien (Hannoveraner Konsens Gesundheitsökonomie 2007, Good Epidemiological Practice Konsens 2004) nicht genügen.

Das zentrale Problem ist somit, dass die Monitoring-Methoden die Grundsatzfrage der Ursachen für beobachtbare Unterschiede ungenügend beleuchten und damit Raum lassen für unzulässige Schlussfolgerungen. Internationale Konsensuspapiere in Gesundheitsökonomie und Epidemiologie fordern deshalb u. a. [12]: «Bereits bei der Planung einer Studie sollten Massnahmen zur Abwehr von Biases ergriffen werden, die durch Selektion, Confounding (s. Glossar) etc. entstehen können. Dazu zählen zum Beispiel das Matching oder eine Einschränkung der Variabilität von Störfaktoren oder aber die Erfassung von Informationen, die zu der Kontrolle von Confounding erforderlich ist. Zur Abschätzung der Auswirkungen von Messfehlern auf das Studienergebnis können zusätzliche Erhebungen zur Durchführung von Sensitivitätsanalysen geplant werden.»

Das Problem des Confounding wird in den bisher zugänglichen Schriften zum Monitoring negativer Auswirkungen der SwissDRG aus den Kreisen der Ökonomie, Epidemiologie und Ethik nicht erwähnt [8, 11].



«Risk Avoidance Creep»: Möglichkeit der Risikoselektion in Abhängigkeit von vermuteten Sterberisiken.

Ruf nach nationalem Monitoring-Konsens im Gesundheitswesen

Eine Publikation inkorrekter Daten oder eine irreführende Interpretation – insbesondere von Mittelwert-Statistiken – sind ethisch nicht vertretbar. Ein Monitoring-Konsens ist zurzeit in der Schweiz inexistent, insbesondere scheint die Gefahr, welche von Mittelwertvergleichen ohne gleichzeitige Confounding-Forschung ausgeht, ungenügend erkannt. Das Positionspapier für die Begleitforschung aus Anlass der Einführung von SwissDRG der FMH belegt dies deutlich [8]. Grundsätzlich sollte die Validität einer Evidenz gemäss

Glossar

Confounder

Unter einem Confounder (deutsch: «Störfaktor») versteht man innerhalb von epidemiologischen Studien einen nicht erfassten und daher nicht kontrollierbaren Störfaktor, der mit einem Faktor unter Beobachtung (Exposition) in Beziehung steht. Ein Confounder ist eine Variable, die das Auftreten eines Risikofaktors (z. B. Krankheit) und die beobachtete Exposition mitbestimmt. Untersucht man den Zusammenhang zwischen Nikotinkonsum und Leberzirrhose oder hepatozellulärem Karzinom, kann man eine signifikante Korrelation feststellen. Es besteht jedoch kein kausaler Zusammenhang, Nikotinkonsum führt nicht zur Leberzirrhose. Vielmehr sind viele Raucher auch Trinker und es besteht ein kausaler Zusammenhang zwischen Alkoholkonsum und Leberzirrhose. Der Alkoholkonsum ist also der Confounder, Nikotinkonsum ist dagegen nur ein Indikator für Leberzirrhose (statistisch ausgedrückt: Typ 1 Fehler).

R^2

Das Bestimmtheitsmass (Abk. R^2 oder B, auch Determinationskoeffizient) ist ein Mass der Statistik für den erklärten Anteil der Variabilität (Varianz) einer abhängigen Variablen Y durch ein statistisches Modell. Beispiel: Um die Gesundheitskostenausgaben einer Person vorherzusagen, können verschiedene Variablen benutzt werden. Gemäss Beck [1] erklären 0 Variablen 0 % der Kosten, die Variablen «Alter und Geschlecht» 11 % (R2 = 0,11), die Variablen «Alter, Geschlecht, Hospitalisation im Vorjahr» 21 % und die Variablen «Alter, Geschlecht, Hospitalisation im Vorjahr und pharmaceutical cost groups (PCG)» 30 % der Kosten. Somit fehlen noch die Confounders (Variablen), welche die 70 % der nicht erklärten Kosten beleuchten. Die weltweit durchgeführte Interheart Study [2] zeigte demgegenüber, dass 9 kardiovaskuläre Risikofaktoren (Rauchen, Lipide, Diabetes, Übergewicht, Diät, Bewegung, Alkoholkonsum und psychosoziale Faktoren) 90 % des Risikos für Herzinfarkt erklären. Hier ist die Wahrscheinlichkeit für das Vorhandensein eines Confounders gering.

internationalen Kriterien geprüft werden (s. Tabelle). Dies gestattet, Aussagen zur Validität einer Monitoringvariable genau zu definieren. Damit kann die Gefahr eingegrenzt werden, dass Klasse-IIc-Aussagen zu Klasse-Ia-Aussagen umgewandelt werden. In einer Publikation zu Monitoringresultaten müsste somit zu jeder Aussage der Evidenzgrad erwähnt werden, bevor politische Massnahmen eingeleitet werden.

Forderungen an eine qualitativ hochstehende Durchführung des SwissDRG Monitorings und des Benchmarkings unter den Spitälern:

– Aus der Sicht der Patientensicherheit ist das Positionspapier Begleitforschung DRG der FMH zu wenig ausgereift. Die zur Qualitätskontrolle vorgesehenen Instrumente sind zu einseitig auf harte Endpunkte anhand von Raten ausgerichtet, entsprechen nicht den heutigen Qualitätsstandards gesundheitsökonomischer und epidemiologischer Methodik und schaffen Anreize, die dazu führen werden, dass hohe Risiken nicht mehr behandelt werden aus Angst vor der Generierung einer institutionsbezogenen erhöhten Mortalitätsrate pro Eingriff.
– Für die Definition von Qualitätsindikatoren im stationären Bereich sind die vorab gesundheitsökonomischen Standards zu definieren, zu erforschen und zu begründen. Dabei kann es nicht genügen, im Benchmarking zwischen den Spitälern eine gewichtete «Manhattandistanz» anhand von (zu?) stark gewichteten Casemix-Indices zu verwenden, da damit die Leistungsorientierung der Spitäler zu wenig die Struktur bedingten Kosten einer Institution berücksichtigen wird [13]. Vor dem Abbau von 100 Spitälern in der Schweiz wäre zudem eine Bedarfsanalyse sinnvoll, die Fehlbelegungen von Patienten in Akutspitälern untersuchen würde (ähnlich einer Vollerhebung von 1983 im Kanton Basel-Stadt [14]. Umso mehr, als im OECD-Vergleich 2007 die Schweiz eine weit unterdurchschnittliche Zahl an Akutbetten aufweist (3,4/1000), die zudem überdurchschnittlich belegt sind (87 %) [15]. Mangel an intellektuellem Leistungswillen ist für die Bewältigung der anstehenden Probleme ungeeignet («Zur Beschreibung der Leistungen würden eigentlich auch Indikatoren zu deren Qualität gehören. Leider ist es aufgrund der verfügbaren Bundesstatistiken noch nicht möglich, derartige Indikatoren abzuleiten, weshalb wir auf diese Dimension verzichten und ersatzweise davon ausgehen, dass das Qualitätsniveau konstant ist.»[13].
– Die Hauptträgerin der Begleitforschung kann nicht die SwissDRG AG sein (Interessenkonflikt).
– Die Hauptstossrichtung des Qualitätsmonitorings sollte unter Einschluss der pharmaceutical cost groups auch im stationären Bereich erfolgen. Es ist eine Kontrolle (Validierung) der DRG mit Hilfe der PCG notwendig.
– Es sollte eine nationale Taskforce «Patientensicherheit und DRG» zur Qualitätskontrolle der DRG und der Spitalleistungen ernannt werden. Diese sollte weitere Gremien berücksichtigen, wie z. B. die SAMW, die Arbeitsgruppe um Yves Eggli, Lausanne, das Zürcher Hortonzentrum unter der Leitung von Prof. J. Steurer und eventuell den Verein Physician Profiling, den Verein Outcome, Verein Spitalbenchmark, die FMH, die Swiss DRG, die H+, den Bund/BAG, die Kantone u. a. m. Die Aufgaben der Taskforce sind:
  a. Definition eines nationalen Standards für die Monitoring-Methoden unter Einbezug einer «Level of evidence» berücksichtigenden Klassifizierung;

Tabelle

Schätzung der Sicherheit (Präzision) einer Aussage (modifiziert nach Circulation. 2009;119:1977–2016).

	Klasse I Nutzen >>> Risiken.	**Klasse IIa** Nutzen >> Risiken.	**Klasse IIb** Nutzen ≥ Risiken. Weitere Studien notwendig.	**Klasse III** Risiko ≥ Nutzen.
Level A: mehrere Populationen untersucht; Daten stammen aus mehreren randomisierten klinischen Studien oder Metaanalysen.	Empfehlung: Massnahme ist nützlich/effektiv. Evidenz: ausreichend	Empfehlung: Massnahme spricht für nützlich/effektiv. Evidenz: teilweise widersprüchlich	Empfehlung: Massnahme ist wenig etabliert. Evidenz: mehrere Studien mit erheblicher Widersprüchlichkeit	Empfehlung: Massnahme ist nicht nützlich/effektiv, evtl. gefährlich. Evidenz: ausreichend
Level B: wenige Populationen untersucht. Daten stammen nur aus einer Studie.	Empfehlung: Massnahme ist nützlich/effektiv. Evidenz: nur wenige Studien	Empfehlung: Massnahme spricht für nützlich/effektiv. Evidenz: widersprüchliche Studienlage	Empfehlung: Massnahme ist wenig etabliert. Evidenz: widersprüchliche Studienlage	Empfehlung: Massnahme ist nicht nützlich/effektiv, evtl. gefährlich. Evidenz: wenig ausreichend
Level C: sehr wenige Studien. Nur Daten aus einzelnen Berichten und Konsensus von Experten.	Empfehlung: Massnahme ist nützlich/effektiv. Evidenz: nur Expertenmeinungen	Empfehlung: Massnahme spricht für nützlich/effektiv. Evidenz: teils divergierende Expertenmeinung.	Empfehlung: Massnahme ist wenig etabliert. Evidenz: nur divergierende Expertenmeinungen	Empfehlung: Massnahme ist nicht nützlich/effektiv, evtl. gefährlich. Evidenz: nur Expertenmeinung

b. Daraus abzuleiten und zu definieren: wissenschaftlich validierte Messinstrumente unter höchstmöglichen Qualitätsstandards der Messmethode selbst, speziell unter Berücksichtigung des Confounding Problems;
c. Einbezug von PCG in die Qualitätssicherung auf stationärer und ambulanter Ebene zur Kontrolle der manipulierbaren DRG; damit Erfassung von «Risk avoidance creep»-Effekten und (erzwungenen?) Fehl-Codierungen von DRG in den Spitälern;
d. Ausarbeitung und Validierung nationaler Standards zur Qualitätskontrolle bis 2017;
e. Erlass gesetzgeberischer Massnahmen auf Ebene KVG zur einstweiligen Verhinderung von Fehlentwicklungsmöglichkeiten, die sich aus den DRG ab 2012 ergeben, insbesondere unter Berücksichtigung, dass die Finanzierung von Leistungen in «dubio pro reo» bis auf weiteres auch gesetzlich festgelegt ist und dass bis 2017 politisch motivierte (künstliche) Senkungen der Baserate gesetzlich verboten sind (Stichwort: Verhinderung von ungerechtfertigtem Spitalsterben).
f. Überprüfung der Auswirkungen von DRG bis 2017 und gesetzliche Verankerung dieses Auftrages. Anschliessend weitere gesetzgeberische Anpassungen.
g. Allgemeine Zugänglichkeit der Datengrundlage zur Erforschung und Generierung von validierbaren Spitalvergleichen, Baserate-Erzeugnissen und Monitoring-Resultaten.
h. Validierung von Benchmark-Methoden zwecks Vergleich von Spitälern untereinander, z. B. anhand eines Tageszensus [16].

Schlussfolgerungen

Das heutige qualitativ hochstehende schweizerische Gesundheitswesen erfährt zurzeit einen Umbauprozess, dessen Auswirkungen auf die Versorgungssicherheit unbekannt sind. Der Ruf nach einem Monitoring zur Erfassung negativer Auswirkungen dieses Prozesses ist mehr als berechtigt. Die bisher präsentierte Methodik des Monitorings und des Benchmarkings zwischen den Spitälern ist zu wenig breit abgestützt, teilweise geheimniskrämerisch und berücksichtigt zudem in ungenügender Weise das methodische Problem des Confoundings, also der eigentlichen Ursachenforschung für ermittelte Performance-Unterschiede. Ohne diese Grundlagenforschung können die Resultate des Monitorings für einen Abbau der medizinischen Versorgungsqualität und der Patientensicherheit missbraucht werden. Sämtliche Studien sollten eine Schätzung des Erklärungsgehalts der Daten (R^2, s. Glossar) berechnen und daraus ableiten, ob die Realität in den Aussagen mit genügend Anstand berücksichtigt bzw. gemessen wurde.

Literatur

1 http://physicianprofiling.ch/HealthEconomicsPCG-ArticelBeckWorkingPaper2006.pdf

2 Yusuf S, Hawken S, Ounpuu S, on behalf of the INTERHEART Study Investigators. Effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries (the INTERHEART study): case-control study. Lancet. 2004;364:937–952.

3 Sterberaten BAG 2009 http://physicianprofiling.ch/SterberatenBAGOriginal042009.pdf

4 Sterberaten BAG 2009, Beilage http://physicianprofiling.ch/SterberatenBAGQ_IndikBAG_Spez_V11_d.pdf

5 Sterberaten Kardiologie http://physicianprofiling.ch/PublicHealthHazardsMittelwerteSterberatenJACC032009.pdf

6 Hannoveraner Konsens zur Gesundheitsökonomie 2007 http://physicianprofiling.ch/PHYHannoveranerKonsensGesundheitsoekonomie2007.pdf

7 Schwank A. Gesundheit kostet. In: Soziale Medizin; 2009 (9). http://physicianprofiling.ch/Gesundheit-HatIhrenPreisDrSchwankSozialeMedizin092009.pdf

8 SwissDRG. Monitoring für eine nachhaltige Gesundheitspolitik erforderlich aus der Sicht der FMH. http://physicianprofiling.ch/SwissDRGFMH092009.pdf ; http://physicianprofiling.ch/SwissDRGKonzeptBegleitforschungFMH02009.pdf; aktualisierte Version im Internet: www.fmh.ch/tarife/swissdrg/position_fmh.html

9 SwissDRG: Monitoring aus ethischer Sicht. http://physicianprofiling.ch/SwissDRGKriterienEthik092009.pdf

10 SwissDRG. Vergleichszahlen können falsch sein, USZ hat nachgerechnet. http://physicianprofiling.ch/SwissDRGStatistikProblemeLueschwer092009.pdf

11 Qualität Datenerhebung der SAMW. Bericht. http://physicianprofiling.ch/QualitaetDatenerhebungSAMW2009.pdf

12 Richtlinien für die «Good Epidemiological Practice», 2004. http://physicianprofiling.ch/GEPVollstaendig2004.pdf

13 Benchmark Making zwischen den Spitälern, ein Konzept des BAG, BFS. Preisüberwachung. http://physicianprofiling.ch/SpitalBenchmarkMethodeUndBeispieleBFS2006.pdf

14 Pflegefälle und Akutbetten: Ergebnisse eines Tageszensus. http://physicianprofiling.ch/RomanensTageszensusApril1983.pdf

15 http://physicianprofiling.ch/KocherSAeZ112009.pdf

16 www.physicianprofiling.ch/FlatRateControl1.pdf