

Weiterentwicklung der Risikoadjustierung für die German Inpatient Quality Indicators

Daniel Schwarzkopf, Raphael Graf und Ulrike Nimptsch

Im Indikatorensystem der German Inpatient Quality Indicators (G-IQI), das auf Krankenhausabrechnungsdaten beruht, wird für ausgewählte Sterblichkeitsindikatoren eine Standardisierung nach Alter und Geschlecht angewandt, um die Indikatorergebnisse der Krankenhäuser unabhängig von einer ggf. abweichenden Fallzusammensetzung bezüglich der berücksichtigten Risikofaktoren bewerten zu können (Mansky u. Nimptsch 2014). In den letzten Jahren entstand bei den IQM Mitgliedskrankenhäusern das Anliegen, sich intensiver mit der Risikoadjustierung auseinanderzusetzen. Die Thematik wurde ergänzend dazu durch die teils mediale Nutzung der veröffentlichten Qualitätsergebnisse für externe Qualitätsvergleiche in den Fokus gerückt. Auch entstand der Wunsch einiger Mitgliedskrankenhäuser, die patientenseitigen Risikofaktoren stärker in der Berechnung der G-IQI zu berücksichtigen. Seitens IQM hat man sich entschieden, nach einer Evaluation, eine weitergehende Adjustierung zur Berechnung von ausgewählten Qualitätsindikatoren einzuführen. Der vorliegende Beitrag gibt einen Überblick über den aktuellen Stand der Weiterentwicklung, reflektiert kritisch die Limitationen risikoadjustierter Qualitätsindikatoren und leitet Empfehlungen für die Nutzung dieser Indikatoren durch die IQM Mitgliedskrankenhäuser sowie die Weiterentwicklung der Methodik durch die IQM ab.

6.1 Aktueller Stand der Entwicklung einer erweiterten Risikoadjustierung

Um einen Überblick zu aktuell angewandten Modellen und Methoden zu erhalten, hat die IQM am 1. Februar 2019 ein Hearing für Fachexperten organisiert. Dabei wurden, mit Fokus

auf die G-IQI, verschiedene Modelle und Verfahren vorgestellt (Initiative Qualitätsmedizin 2019). Als Folge des Hearings hat die IQM eine Arbeitsgruppe „Risikoadjustierung“ gegründet, zu welcher ausgewählte externe Expertinnen und Experten eingeladen wurden. Ziel der Arbeitsgruppe war es, Indikatoren zu identifi-

zieren, für die als Alternative ein erweitertes Risikomodell eingesetzt werden kann. Die Ergebnisse und Empfehlungen der Arbeitsgruppe wurden veröffentlicht (Schwarzkopf et al. 2021). Interessierte IQM Mitgliedskrankenhäuser wurden frühzeitig in die Projektarbeit eingebunden. Dadurch konnten erste Berechnungen anhand von Krankenhausabrechnungsdaten mit erweiterten Risikoadjustierungsmodellen für ausgewählte Indikatoren in den Daten von über 170 IQM Mitgliedskrankenhäusern durchgeführt werden (Schwarzkopf et al. 2021). Die teilnehmenden Kliniken haben dabei frühzeitig einen Ausblick auf mögliche krankenhausespezifische Auswirkungen und Änderungen erhalten, die mit dem Einsatz alternativer Risikoadjustierungsmodelle einhergehen. Diese erweiterte Risikoadjustierung wurde nun in der IQM Halbjahresauswertung der Daten 2021 erstmals für alle Mitgliedskrankenhäuser parallel zur bisherigen Methodik eingesetzt. Für zukünftige Auswertungen ist eine Weiterentwicklung der alternativen Modelle im Hinblick auf den spezifischen Anwendungszweck der Qualitätsmessung vorgesehen. Zudem wird geprüft, ob erweiterte Risikoadjustierungsmodelle für weitere Indikatoren infrage kommen.

6.2 Überblick zur alten und neuen Methodik

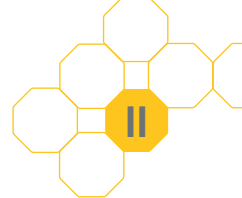
Beide Verfahren beruhen auf dem Prinzip der indirekten Standardisierung. Dabei wird jedem Behandlungsfall aufgrund seiner Eigenschaften das gemäß dem Bundesdurchschnitt zu erwartende Sterberisiko zugewiesen. Die dafür verwendeten Referenzzahlen werden mit allen in Deutschland verfügbaren Abrechnungsdaten aus der *Diagnosis-related-groups* (DRG)-Statistik beim statistischen Bundesamt abgefragt (Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder 2019). Bei der bisherigen Methodik werden dabei nur die Merkmale Alter und Geschlecht berücksichtigt, während die erweiterten Risikomodelle darüber hinaus

weitere Merkmale einbeziehen. Konkret erfolgt die bisherige Risikoadjustierung auf Basis von 5-Jahres-Altersgruppen sowie dem Geschlecht (Nimptsch u. Mansky 2020). Jedem Behandlungsfall wird als erwartetes Sterberisiko die Sterblichkeit zugewiesen, welche in der DRG-Statistik in der entsprechenden Alters- und Geschlechtsgruppe für diesen Indikationsbereich beobachtet wurde. Für die neuen erweiterten Risikomodelle funktioniert dies nach einem ähnlichen Prinzip, wobei in die Ermittlung der erwarteten Sterblichkeit mehr Faktoren eingehen und das mathematische Verfahren komplexer ist. Durch Aufsummierung dieser Risiken über die Behandlungsfälle i eines Krankenhauses j wird die erwartete Anzahl von Verstorbenen in einem Indikator bestimmt und mit der tatsächlich beobachteten Anzahl der Todesfälle ins Verhältnis gesetzt. Der Quotient von beobachteter und erwarteter Sterblichkeit ist das standardisierte Sterblichkeitsverhältnis (*standardized mortality ratio*, SMR):

$$SMR_j = \frac{\sum O_{ij}}{\sum E_{ij}},$$

O (observed) bezeichnet den beobachteten Status eines Falles (mit 0 = überlebt, 1 = verstorben) und E das erwartete Sterberisiko (als Wahrscheinlichkeit zwischen 0 und 1). Eine SMR unter 1 bedeutet, dass die Sterblichkeit in einem Krankenhaus – unabhängig von den in der Risikoadjustierung berücksichtigten Faktoren – geringer als im Bundesdurchschnitt ist, während eine SMR über 1 eine höhere Krankenhaussterblichkeit im Vergleich zum Bundesdurchschnitt anzeigt. In dem nachfolgenden Beispiel (s. Abb. 1) wird dargestellt, wie anhand der indirekten Standardisierung ein Erwartungswert in einem bestimmten Indikator ermittelt wird.

Für fünf Indikatoren wurden indikatorspezifische erweiterte Risikomodelle angewendet (Sterblichkeit bei 1.1 Herzinfarkt, 2.1 Herzinsuffizienz, 9.3 Hirninfarkt, 14.2 Pneumonie, 21.311 Kolonresektion). Die verwendeten Modelle wurden bereits in vorangegangenen Arbeiten, jedoch nicht für den Zweck der Qualitäts-



Altersgruppe	20-24	20-24	25-29	25-29	30-34	30-34	35-39	35-39	...	90-94	90-94	95-xx	95-xx
Geschlecht	m	w	m	w	m	w	m	w	...	m	w	m	w
Anteil Todesfälle	2,510%	2,990%	2,510%	2,990%	2,510%	2,990%	2,510%	2,990%	...	21,710%	23,180%	29,210%	28,470%

(a) Aus der DRG-Statistik 2018 abgeleitete Referenzwerte für die Sterblichkeit nach Geschlecht und 5-jahres-Altersgruppe

Krankenhaus XY					
Fall	Altersgruppe	Geschl.	Zähler (O)	Nenner (N)	Erwartungswert (E)
00001	25-29	m	1	1	0,0251
00002	25-29	w	0	1	0,0299
00003	25-29	m	0	1	0,0251
00004	30-34	m	0	1	0,0251
00005	90-94	w	0	1	0,2318
00006	25-29	w	0	1	0,0299
00007	20-24	m	0	1	0,0251
00008	95-xx	m	0	1	0,2921
00009	95-xx	m	0	1	0,2921
00010	90-94	w	0	1	0,2318
00011	90-94	w	0	1	0,2318
00012	25-29	m	0	1	0,0251
00013	25-29	w	0	1	0,0299
00014	95-xx	w	1	1	0,2847
...					
Summe			2	14	1,7795

(b) Beispielhafte Darstellung der fallindividuellen Zuweisung von Erwartungswerten auf Basis der Referenzwerte (a).

Kennzahl	Ergebnis	Berechnung
Zähler (O-Summe Zähler)	2	
Nenner (N-Summe Nenner)	14	
Erw. Zähler (E-Summe Erwartungswert)	1,7795	
Rate (O / N)	0,14285714	$2/14 \cdot 100$
Erwartete Rate (E / N)	0,12710714	$1,7795/14 \cdot 100$
SMR (O / E)	1,12391121	$2/1,7795$

(c) Berechnung der üblicherweise berechneten Kennzahlen auf Basis der beispielhaft dargestellten Fälle (b).

Abb. 1 Beispiel für die Ermittlung des standardisierten Sterblichkeitsverhältnisses (SMR)

messung, sondern zur Untersuchung von Mengen-Ergebnis-Beziehungen entwickelt (Bolczek et al. 2020; Nimptsch u. Mansky 2017). Berücksichtigt werden je nach Indikator unterschiedliche Begleiterkrankungen, ggf. die Erkrankungsschwere, der Zuverlegungsstatus sowie das Alter und Geschlecht. Ein Überblick zu den Indikatoren, den Variablen sowie der auf Bundesebene auf Basis der DRG-Statistik des Jahres 2019 ermittelten Odds Ratios aller eingesetzten Modelle wird in Tabelle 1 dargestellt.

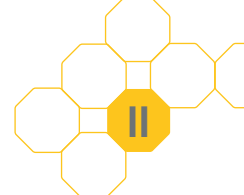
6.3 Beispielhafte Darstellung der Indikatoren mit beiden Risikomodellen für IQM Mitgliedskrankenhäuser

Damit die Mitgliedskrankenhäuser einen Überblick zu den indikatorindividuellen Auswirkungen der neuen Risikoadjustierung erhalten können, werden die fünf betroffenen Indikatoren ab der Halbjahresauswertung H1 2021 in zweifacher Ausführung berichtet. Die Auswertungen werden um jeweils einen zusätzlichen Indikator ergänzt. Dieser ist mit „R“ für Risikoadjustierung gekennzeichnet und beinhaltet neben den Informationen zum Zähler und Nenner auch die SMR sowie den Erwartungswert auf Basis der neuen Modelle (s. Tab. 2).

Tab. 1 Definition der indicatorspezifischen Risikoadjustierungsmodelle und zugehörige Odds Ratios, DRG-Statistik 2019

Variablen	Definition	Indikator (G-IQI 5.3), Sterblichkeit bei				
		1.1 Herzinfarkt	2.1 Herzinsuffizienz	9.3 Hirninfarkt	14.2 Pneumonie	21.311 Kolonresektion
Alter in Jahren		1,07 [1,07–1,07]	1,05 [1,05–1,05]	1,07 [1,07–1,07]	1,06 [1,06–1,06]	1,07 [1,06–1,07]
Weibliches Geschlecht		1,10 [1,06–1,14]	0,81 [0,80–0,83]	1,11 [1,08–1,15]	0,86 [0,83–0,89]	0,72 [0,64–0,82]
Zuverlegung aus anderem Krankenhaus	Aufnahmeanlass A, V	0,98 [0,93–1,04]	1,60 [1,53–1,67]	1,23 [1,18–1,29]	–	2,34 [1,70–3,22]
Herzrhythmusstörung	ND I442 I48 Z450 Z950	–	0,86 [0,84–0,88]	–	1,04 [1,00–1,07]	1,58 [1,38–1,82]
Herzinsuffizienz/ Kardiomyopathie	ND I50 I110 I130 I132 I420 I426 I427 I428 I429	–	–	1,55 [1,48–1,61]	1,39 [1,34–1,43]	2,03 [1,75–2,35]
Chronische ischämi- sche Herzkrankheit	ND I25	–	0,76 [0,74–0,78]	1,05 [1,00–1,10]	0,92 [0,88–0,95]	0,98 [0,84–1,15]
Hypertonie (ohne Herz- oder Niereninsuffizienz)	ND I10 I119 I129 I139 I15	0,43 [0,41–0,44]	0,76 [0,74–0,78]	0,55 [0,53–0,57]	0,63 [0,61–0,65]	0,70 [0,61–0,79]
Aorten-/ Mitralklappenvitien	ND I340 I342 I350 I351 I352 I050 I051 I052 I060 I061 I062 Q231 Q232 Q233	0,70 [0,66–0,74]	0,72 [0,70–0,74]	0,61 [0,56–0,66]	0,79 [0,74–0,85]	0,81 [0,61–1,07]
Atherosklerose der Extremitätenarterien	ND I702	–	1,18 [1,13–1,22]	1,31 [1,21–1,41]	1,21 [1,13–1,30]	1,97 [1,52–2,56]
Chronische Lungenerkrankung	ND J41 J42 J44 J45 J47	0,98 [0,91–1,04]	1,01 [0,98–1,04]	1,16 [1,09–1,24]	–	1,37 [1,14–1,64]
Chronische Lebererkrankung	ND B18 I864 I982 K70 K73 K74 K760 K761 K765 K766 K767 Q446 Q447	1,08 [0,93–1,24]	1,40 [1,34–1,47]	1,12 [0,97–1,30]	1,05 [0,96–1,15]	1,58 [1,23–2,03]
Schwere Nierenerkrankung oder chronische Niereninsuffizienz	ND I120 I131 I132 N03 N04 N05 N07 N08 N11 N12 N14 N15 N16 N18 N19 Z992	–	0,98 [0,96–1,01]	1,00 [0,96–1,04]	1,01 [0,97–1,04]	1,21 [1,04–1,40]
Chronische Niereninsuffizienz	ND I120 I131 I132 N18 N19 Z992	0,94 [0,90–0,98]	–	–	–	–
Diabetes mellitus	ND E10 E11 E12 E13 E14	1,01 [0,97–1,06]	0,98 [0,96–1,00]	1,07 [1,04–1,11]	1,04 [1,00–1,07]	1,13 [0,98–1,30]
Adipositas	ND E66	–	0,67 [0,64–0,70]	0,77 [0,70–0,84]	0,64 [0,58–0,70]	0,80 [0,64–1,00]
Kachexie/ Mangelernährung	ND R64 R634 E43 E44	–	1,50 [1,40–1,60]	1,33 [1,20–1,48]	1,51 [1,41–1,62]	1,10 [0,90–1,36]

Ausgewiesen werden Odds Ratios mit 95% Konfidenzintervall (nicht signifikante Werte in grauen Feldern). HD: Hauptdiagnose; ND: Nebendiagnose



Variablen	Definition	Indikator (G-IQI 5.3), Sterblichkeit bei				
		1.1 Herzinfarkt	2.1 Herzinsuffizienz	9.3 Hirninfarkt	14.2 Pneumonie	21.311 Kolonresektion
Gerinnungsstörung	ND D66 D67 D680 D681 D682 D684 D685 D686 D688 D689 D691 D693 D694	–	4,23 [3,97–4,51]	3,42 [3,03–3,87]	3,53 [3,19–3,90]	5,15 [4,42–6,00]
Bösartige Neubildung	HD oder ND C00-C97	1,55 [1,38–1,73]	1,67 [1,59–1,75]	2,57 [2,40–2,76]	–	–
STEMI	HD I210 I211 I212 I213	1,58 [1,52–1,64]	–	–	–	–
Rezidivierender AMI	HD I22	2,73 [1,90–3,91]	–	–	–	–
Kardiogener Schock	ND R570	25,23 [24,12–26,39]	–	–	–	–
Folgen einer zerebro- vaskulären Krankheit	ND I69	–	1,46 [1,39–1,54]	–	–	–
Linksherzinsuffizienz NYHA Stadium 4	HD oder ND I5014	–	3,22 [3,15–3,29]	–	–	–
Metastasierende Neubildung	HD oder ND C77 C78 C79	–	–	–	–	1,41 [1,24–1,60]
Erweiterte Kolonresektion	OPS 54540 54541 54542 54543 54544 54545 54546 5501 5502 5437 5436 5454 x 5454 y	–	–	–	–	1,93 [1,65–2,27]
Fläche unter der Kurve (c-Statistik)		0,847	0,730	0,735	0,712	0,812

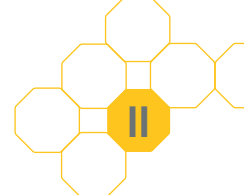
Ausgewiesen werden Odds Ratios mit 95% Konfidenzintervall (nicht signifikante Werte in grauen Feldern). HD: Hauptdiagnose; ND: Nebendiagnose

Tab. 2 Beispielhafte Darstellung der G-IQI mit den verschiedenen Risikomodellen

G-IQI Version 5.3	Referenzwerte		IQM Gesamt 2021						Krankenhaus 1				
Indikatoren	Zielwert/ Indikatortyp (Hinweise)	Referenz- wert	SMR/ Ratio	Rate	Erwar- tete Rate	Zähler	Nenner	SMR/ Ratio	Rate	Erwar- tete Rate	Zähler	Nen- ner	
01 – Herzinfarkt													
01.1 – Hauptdiagnose Herzinfarkt (Alter > 19), Anteil Todesfälle, aufgetreten	unter Erwartungswert	7,9%	1,07	8,42%	7,84%	6.868	81.591	1,23	9,73%	7,89%	85	870	
01.1 – Hauptdiagnose Herzinfarkt (Alter > 19) <i>ohne COVID-19</i> , Anteil Todesfälle, aufgetreten	unter Erwartungswert	7,9%	1,05	8,24%	7,84%	6.643	80.577	1,00	8,86%	8,89%	49	549	
01.1R – Hauptdiagnose Herzinfarkt (Alter > 19), Anteil Todesfälle, aufgetreten (differenziertes Risikomodell)	unter Erwartungswert	7,9%	1,01	8,42%	8,34%	6.868	81.591	0,98	9,73%	9,97%	85	870	

II Aktuelle Schwerpunkte, Projekte und Kampagnen

G-IQI Version 5.3	Referenzwerte		IQM Gesamt 2021					Krankenhaus 1				
Indikatoren	Zielwert/ Indikatortyp (Hinweise)	Referenz- wert	SMR/ Ratio	Rate	Erwar- tete Rate	Zähler	Nenner	SMR/ Ratio	Rate	Erwar- tete Rate	Zähler	Nen- ner
01.1R – Hauptdiagnose Herzinfarkt (Alter > 19) <i>ohne COVID-19</i> , Anteil Todesfälle, aufgetreten (differenziertes Risikomodell)	unter Erwartungswert	7,9%	0,99	8,24%	8,31%	6.643	80.577	0,81	8,86%	10,97%	49	549
02 – Herzinsuffizienz												
02.1 – Hauptdiagnose Herzinsuffizienz (Alter > 19), Anteil Todesfälle, aufgetreten	unter Erwartungswert	7,8%	1,03	8,14%	7,90%	14.115	173.427	0,86	5,51%	6,44%	25	454
02.1 – Hauptdiagnose Herzinsuffizienz (Alter > 19) <i>ohne COVID-19</i> , Anteil Todesfälle, aufgetreten	unter Erwartungswert	7,8%	1,00	7,90%	7,89%	13.487	170.790	0,98	5,45%	5,58%	24	441
02.1R – Hauptdiagnose Herzinsuffizienz (Alter > 19), Anteil Todesfälle, aufgetreten (differenziertes Risikomodell)	unter Erwartungswert	7,8%	1,18	8,14%	6,87%	14.115	173.426	0,97	5,51%	5,68%	25	454
02.1R – Hauptdiagnose Herzinsuffizienz (Alter > 19) <i>ohne COVID-19</i> , Anteil Todesfälle, aufgetreten (differenziertes Risikomodell)	unter Erwartungswert	7,8%	1,15	7,90%	6,85%	13.487	170.789	1,22	5,45%	4,46%	24	441
09 – Schlaganfall												
09.3 – Hirninfarkt (ICD I63, Alter > 19), Anteil Todesfälle, aufgetreten	unter Erwartungswert	6,5%	1,12	7,26%	6,50%	7.150	98.495	1,38	8,46%	6,15%	67	786
09.3 – Hirninfarkt (ICD I63, Alter > 19) <i>ohne COVID-19</i> , Anteil Todesfälle, aufgetreten	unter Erwartungswert	6,5%	1,08	7,05%	6,50%	6.832	96.937	1,45	8,40%	5,81%	63	751
09.3R – Hirninfarkt (ICD I63, Alter > 19), Anteil Todesfälle, aufgetreten (differenziertes Risikomodell)	unter Erwartungswert	6,5%	1,14	7,26%	6,35%	7.150	98.495	1,13	8,46%	7,52%	67	786
09.3R – Hirninfarkt (ICD I63, Alter > 19) <i>ohne COVID-19</i> , Anteil Todesfälle, aufgetreten (differenziertes Risikomodell)	unter Erwartungswert	6,5%	1,11	7,05%	6,33%	6.832	96.937	0,95	8,40%	8,81%	63	751
14 – Lungenentzündung (Pneumonie)												
14.2 – Pneumonie ohne Zuverlegungen, Tumor, Mukoviszidose, Alter > 19, Anteil Todesfälle	unter Erwartungswert	9,1%	2,09	15,85%	7,60%	18.450	116.369	1,70	12,91%	7,61%	63	489
14.2 – Pneumonie ohne Zuverlegungen, Tumor, Mukoviszidose, Alter > 19 <i>ohne COVID-19</i> , Anteil Todesfälle	unter Erwartungswert	9,1%	1,27	11,71%	9,25%	5.129	43.788	1,07	9,10%	8,50%	15	160



G-IQI Version 5.3	Referenzwerte		IQM Gesamt 2021					Krankenhaus 1				
Indikatoren	Zielwert/ Indikatortyp (Hinweise)	Referenz- wert	SMR/ Ratio	Rate	Erwar- tete Rate	Zähler	Nenner	SMR/ Ratio	Rate	Erwar- tete Rate	Zähler	Nen- ner
14.2R – Pneumonie ohne Zuverlegungen, Tumor, Mukoviszidose, Alter > 19, Anteil Todesfälle (differenziertes Risikomodell)	unter Erwartungswert	9,1%	2,07	15,85%	7,66%	18.450	116.369	1,10	12,91%	11,69%	63	489
14.2R – Pneumonie ohne Zuverlegungen, Tumor, Mukoviszidose, Alter > 19 ohne COVID-19, Anteil Todesfälle (differenziertes Risikomodell)	unter Erwartungswert	9,1%	1,23	11,71%	9,49%	5.129	43.788	0,88	9,10%	10,30%	15	160
21 – Erkrankungen von Dickdarm und Enddarm												
21.311 – davon Kolonresektionen bei kolorekt. Karzinom ohne kompl. Diagnose, Anteil Todesfälle	< 4,0%	4,0%		3,86%		387	10.030		2,58%		4	155
21.311 – davon Kolonresektionen bei kolorekt. Karzinom ohne kompl. Diagnose ohne COVID-19, Anteil Todesfälle	< 4,0%	4,0%		3,80%		377	9.916		2,68%		4	149
21.311R – davon Kolonresektionen bei kolorekt. Karzinom ohne kompl. Diagnose, Anteil Todesfälle (differenziertes Risikomodell)	unter Erwartungswert	4,0%	0,97	3,86%	3,99%	387	10.030	0,86	2,58%	3,01%	4	155
21.311R – davon Kolonresektionen bei kolorekt. Karzinom ohne kompl. Diagnose ohne COVID-19, Anteil Todesfälle (differenziertes Risikomodell)	unter Erwartungswert	4,0%	0,96	3,80%	3,96%	377	9.916	0,82	2,68%	3,28%	4	149

Dargestellt werden die 5 Indikatoren jeweils in der Variante mit und ohne COVID-19 Fälle, dem Risikomodell nach Alter und Geschlecht sowie dem erweiterten Risikomodell mit Ergebnissen aus der 2. Halbjahresauswertung 2021 für „IQM Gesamt“. Beispielhaft werden daneben Daten für „Krankenhaus 1“ dargestellt.

Lesebeispiel für die Variante inkl. COVID-19 Fällen: Im Indikator Hauptdiagnose Herzinfarkt beträgt der Nenner für das beispielhaft dargestellte Krankenhaus 1 $N = 870$. Der Zähler schließt $Z = 85$ Fälle vom Nenner ein. Die beobachtete Rate beträgt $O = 9,73\%$. Die erwartete Rate im bisherigen Risikomodell beträgt $E = 7,89\%$ und nach dem neuen Risikomodell $E = 9,97\%$. Daraus ergibt sich im alten Modell eine SMR von 1,23 und im neuen Modell von 0,98.

6.4 Bewertung und Diskussion einer erweiterten Risikoadjustierung

Im Folgenden sollen die Limitationen einer Risikoadjustierung auf Basis von Abrechnungsdaten diskutiert werden, um damit konkrete Empfehlungen für die Nutzung der Indikatorergebnisse mit erweiterter Risikoadjustierung durch die IQM Mitgliedskrankenhäuser abzuleiten. Die folgenden Ausführungen beruhen

Großteils auf der Publikation zur Erprobung der erweiterten Risikoadjustierung, welche in der *Zeitschrift für Evidenz, Fortbildung und Qualität im Gesundheitswesen* erschienen ist (Schwarzkopf et al. 2021), und zitieren diese teils wörtlich. Da es sich um eine Open-Access-Publikation einer eigenen Arbeit handelt, wird zur Verbesserung der Lesbarkeit auf die gesonderte Kennzeichnung von Zitaten aus dieser Publikation verzichtet.

6.4.1 Auswahl der „richtigen“ Risikofaktoren in Krankenhausabrechnungsdaten

Die Nutzung von Krankenhausabrechnungsdaten für die Qualitätsmessung kennzeichnet den pragmatischen Ansatz von IQM. Die Korrektheit und Vollständigkeit der Kodierung sind jedoch unsicher und unterliegen Erlösanreizen; zudem kann die Kodierqualität auch zwischen Krankenhäusern variieren (Iezzoni 2009). Dies wirkt sich auf die Kodierung möglicher Risikofaktoren aus, da im Rahmen des DRG-Systems ein Anreiz zur Kodierung nur so lange gegeben ist, wie diese zu einem höheren Erlös führen kann. Die Berücksichtigung spezifischer Risikofaktoren könnte einen Anreiz schaffen, häufiger (und ggf. real gar nicht vorliegende) Risikofaktoren in Abrechnungsdaten zu kodieren, um in den risikoadjustierten Indikatoren besser abzuschneiden.



Als Risikofaktoren sollten nur vorbestehende patientenseitige Merkmale verwendet werden und nicht solche, welche selbst durch die Qualität der Versorgung beeinflusst sein könnten (Iezzoni 2009; Krumholz et al. 2006a).

Eine Adjustierung nach (potenziellen) Komplikationen ist im Rahmen einer Qualitätsmessung nicht zielführend, da so eventuelle Qualitätsprobleme „wegadjustiert“ würden (Mansky u. Nimptsch 2014). Ein Beispiel wäre die Nebendiagnose einer Lungenembolie bei kolorektaler Resektion, welche als Risikofaktor vorbestehend sein kann, jedoch postoperativ als Komplikation auftreten könnte. In einer US-amerikanischen Untersuchung der zeitlichen Zuordnung schwerer Sekundärdiagnosen bei chirurgischen Fällen mittels Aktensichtung zeigte sich, dass tiefe Venenthrombosen oder Lungenembolien nur in 22% der Fälle bereits bei Aufnahme bestanden (Lawthers et al. 2000). Eine Berücksichtigung ungeeigneter Variablen als Risikofaktoren würde zwar die statistische

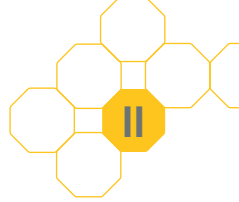
Vorhersagegüte eines Risikomodells erhöhen, zugleich aber die Bewertbarkeit der Qualität einschränken (Iezzoni 2009). Während einige Länder Abrechnungsdaten systematisch für die Nutzung zur Qualitätssicherung verbessert haben – etwa durch die Einführung einer *Present-on-admission*-Kodierung (Iezzoni 2009; Zhan et al. 2007), gibt es in Deutschland bisher (trotz wiederholter Bemühungen, u. a. der IQM) keine entsprechenden Bestrebungen aufseiten der regulatorischen Instanzen.

Neben der angesprochenen Frage der Kodierqualität der Abrechnungsdaten stellt auch deren mangelnder klinischer Gehalt eine Einschränkung ihrer Nutzbarkeit für Qualitätssicherungszwecke dar (Iezzoni 2009).



Zahlreiche klinisch relevante Risikofaktoren sind über Abrechnungsdaten nicht erfassbar.

So ist beispielsweise der Bewusstseinszustand bei Aufnahme, der einen relevanten Risikofaktor für das Versterben bei Schlaganfall darstellt (Bray et al. 2014), nicht in der ICD-10-GM kodierbar und somit nicht in Krankenhausabrechnungsdaten abbildbar. Amerikanische Studien, die die SMR auf Basis von Abrechnungsdaten mit solchen auf Basis von elektronischen Patientenakten abglichen, zeigten jedoch sehr hohe Korrelationen, was die grundsätzliche Eignung auch von Abrechnungsdaten zur Nutzung für risikoadjustierte Versorgervergleiche nahelegt (Krumholz et al. 2006b, 2006c). Eine weitere Einschränkung der Krankenhausabrechnungsdaten ist das Fehlen von Vor- und Nachbeobachtungszeiträumen, da nur Daten eines einzigen Krankenhausaufenthaltes verfügbar sind. Dies schränkt die Identifikation vorbestehender Risikofaktoren wie chronischer Begleiterkrankungen oder vorangegangener Krankheitsereignisse ein. Zudem können Qualitätsergebnisse durch den Einfluss von Verlegungen systematisch verzerrt sein, da deutsche Krankenhausabrechnungsdaten keine Endpunkte nach standardisierten Zeiträumen (etwa 90-Tage-Sterblichkeit) beinhalten. Die-



sem Problem kann durch die Kontrolle des Zuverlegungsstatus im Risikomodell begegnet werden. Hierbei kann jedoch nicht unterschieden werden, ob es sich um Aufwärtsverlegungen kritisch kranker Patienten oder Anschlussbehandlungen nach der kritischen Phase handelt, was zu spezifischen Verzerrungen führen kann. Damit könnte die Bewertung in Krankenhäusern, in denen eine dieser Konstellationen überwiegt, verzerrt werden. Möglicherweise ist eine Stratifizierung nach dem Zuverlegungsstatus – wie sie in G-IQI für die Indikatoren Herzinfarkt, Herzinsuffizienz und Hirninfarkt bereits umgesetzt wird – hier sinnvoller.

Auch muss bei der Interpretation von Indikatoregebnissen die mögliche Ungenauigkeit bei kleinen Fallzahlen berücksichtigt werden. Je kleiner die Anzahl behandelter Patienten in einem Indikationsbereich, desto höher ist die Wahrscheinlichkeit rein zufälliger Abweichungen der Indikatoregebnisse von der Erwartung (z.B. vom SMR-Wert von 1). Durch die Berücksichtigung von Konfidenzintervallen, die eine Bewertung der Präzision des Indikatoregebnisses ermöglichen, kann dieser Zufallsfehler eingeschätzt werden. Sofern die Indikatoren – wie bei IQM – als Aufgreifkriterium für weitergehende Fallanalysen mit dem Ziel der kontinuierlichen Qualitätsverbesserung verwendet werden sollen, kann der Verzicht auf Konfidenzintervalle sinnvoll sein. Damit wird vermieden, dass statistisch nicht signifikante, aber möglicherweise qualitätsrelevante Auffälligkeiten bei Kliniken mit kleiner Fallzahl gar keiner Analyse unterzogen werden. Für die (öffentliche) Bewertung und Vergleiche von Qualitätsindikatoren sollte die Verwendung von Konfidenzintervallen durch die IQM geprüft werden.



Es sollte vermieden werden, dass rein zufällige Abweichungen bzw. Unterschiede als Anzeichen mangelnder Versorgungsqualität fehlinterpretiert werden.

6.4.2 Besonderheiten bei der Ermittlung von Referenzzahlen in Zusammenhang mit COVID-19

Die Ermittlung der Referenzzahlen aus der DRG-Statistik der beim statistischen Bundesamt verfügbaren Abrechnungsdaten erfolgt immer mit einem zeitlichen Verzug von mindestens zwei Jahren. Dadurch werden aktuell bei der Ermittlung von Referenzzahlen keine Patienten mit COVID-19 berücksichtigt. Das kann dazu führen, dass die Verteilung patientenseitiger Risikofaktoren bei den seit Beginn der Pandemie behandelten Patienten nicht vergleichbar ist mit der aus dem Referenzjahr 2019. Das zeigt sich besonders bei dem Indikator G-IQI 14.2 – *Pneumonie ohne Zuverlegungen, Tumor, Mukoviszidose, Alter > 19, Anteil Todesfälle*. Bei beiden Arten der Risikoadjustierung liegt die SMR für die gesamte IQM über 2, was auf einen deutlich veränderten Case-Mix hindeutet.

Eine Möglichkeit für eine transparente Darstellung des COVID-19-Effektes besteht darin, dass die Indikatoren zusätzlich in einer Variante unter Ausschluss aller COVID-19-Patienten berichtet werden. Auch diese zusätzliche Variante kann mit dem erweiterten Risikomodell berechnet werden. Dadurch passt das Referenzjahr (prä-COVID-19) besser zu dem Case-Mix.

6.5 Empfehlungen für die Interpretation und Nutzung der erweiterten Risikoadjustierung

Der Aufwand von Mortalitäts- und Morbiditäts (M&M)-Konferenzen und Peer Review Verfahren ist hoch und es können nur ausgewählte Fälle im Rahmen dieser Prozesse detailliert analysiert werden. Für die konkrete Arbeit mit den betroffenen Qualitätsindikatoren ergeben sich, durch die erweiterte Risikoadjustierung, neue Möglichkeiten der Fallselektion. Ziel bleibt es, in der Fallanalyse Verbesserungspotenziale in der Versorgung zu identifizieren. Bisher wurden bei der Fallauswahl häufig die Patienten aufsteigend nach dem Alter berücksichtigt, für

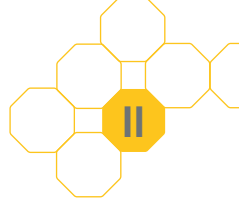
die ein unerwünschtes Ereignis im Sinne des Indikators dokumentiert wurde. Durch die erweiterte Risikoadjustierung können durch die aufsteigende Sortierung nach dem Erwartungswert zusätzliche Risikofaktoren berücksichtigt werden und somit andere Fälle aufgegriffen werden. Dadurch könnte der Anteil von Fällen, bei welchen mögliche Versorgungsdefizite identifiziert werden können, erhöht werden (höherer positiver prädiktiver Wert). Zugleich zeigen Erfahrungen aus Peer Reviews jedoch, dass auch Fälle mit hohen vorbestehenden Risiken wertvolle Informationen – z.B. in Zusammenhang mit der Indikationsstellung oder dem perioperativen Management – liefern können (Burgard 2013). Der Nutzen der Risikoadjustierung im Rahmen der Fallauswahl muss daher indikatorspezifisch und nach der konkreten Zielstellung der Fallanalyse kritisch abgewogen werden.

6.6 Empfehlungen für die Weiterentwicklung der Risikoadjustierung in IQM

Das Expertengremium zur Risikoadjustierung hat konkrete methodische Empfehlungen zur Entwicklung einer erweiterten Risikoadjustierung sowie zum Forschungsbedarf zur Prüfung und Verbesserung der Validität der Datengrundlage formuliert (Schwarzkopf et al. 2021). Die vorgestellten spezifischen Risikoadjustierungsmodelle entsprechen selbst nicht vollumfänglich diesen Empfehlungen, da sie auf vorbestehenden Arbeiten mit anderen Zielsetzungen beruhen (Bolczek et al. 2020; Nimptsch u. Mansky 2017). Auch wenn diese Modelle nun in der G-IQI-Auswertung für die Mitgliedskrankenhäuser in einem ersten Schritt zur Anwendung kommen, muss die erweiterte Risikoadjustierung im G-IQI-System in Zukunft systematisch weiterentwickelt und angepasst werden.

Empfehlungen der Expertengruppe zur Weiterentwicklung der erweiterten Risikoadjustierung (adaptiert nach Schwarzkopf et al. 2021)

- Risikomodelle sollten systematisch im Hinblick auf die Anforderungen einer Qualitätsmessung entwickelt und statistisch validiert werden.
- Für die optimale Interpretierbarkeit, auch von Veränderungen über die Zeit, sollten beobachtete Sterblichkeit, auf Basis der Risikofaktoren erwartete Sterblichkeit und risikoadjustierte Sterblichkeit in den Qualitätsberichten dargestellt werden.
- Veränderungen in der Kodierung von Risikofaktoren in den IQM Mitgliedskrankenhäusern nach der Einführung einer entsprechenden Risikoadjustierung sollten zentral validiert und interpretiert werden.
- Für IQM-interne Aufgreifkriterien zur kontinuierlichen Analyse und Verbesserung der Qualität sollten auch zukünftig keine Konfidenzintervalle Verwendung finden.
- Um fehlerhafte Interpretationen von Qualitätsindikatoren sowie Rankings in der Öffentlichkeit vorzubeugen, sollten in Zukunft für sensible Indikatoren in der Veröffentlichung ggf. Konfidenzintervalle angegeben werden.
- Sofern in Indikatoren aufgrund hoher Fallzahlen vermehrt signifikante, aber klinisch nicht relevante Abweichungen vom Erwartungswert der SMR von 1 auftreten, kann als alternatives Kriterium das signifikante Überschreiten eines kritischen Schwellenwertes zur Abbildung klinischer Relevanz in Betracht gezogen werden.
- Der Nutzen der Risikoadjustierung im Aufgreifkriterium für Fallanalysen und Peer Reviews muss indikatorspezifisch abgewogen und empirisch erprobt werden.
- Im Rahmen der IQM sollten Forschungsprojekte zur Prüfung der Validität der Krankenhausabrechnungsdaten für Qualitätssicherung angestoßen und befördert werden.
- Auf Basis von Prüfungen der Validität der einzelnen G-IQI sollte die Veröffentlichungsfähigkeit jeweils bewertet und einzelne Indikatoren sollten ggf. von der Veröffentlichungspflicht ausgenommen werden.
- Für externe Vergleiche von Qualitätsergebnissen einzelner Krankenhäuser sollten dafür konzipierte Qualitätsmessungen auf Basis besser geeigneter



Daten herangezogen werden, wie beispielsweise das QSR-Verfahren des WiDO.

- Die IQM sollte sich weiterhin für die Verbesserung der Nutzbarkeit von Krankenhausabrechnungsdaten für Qualitätsverbesserung einsetzen, vor allem durch die Forderung der Einführung einer „Present on admission“-Kennzeichnung.

Eine Weiterentwicklung der Risikoadjustierung in der Qualitätsmessung mit G-IQI kann nicht die grundsätzlichen Einschränkungen beheben, denen die Krankenhausabrechnungsdaten bei der Nutzung zur Qualitätssicherung unterliegen. Während die Bedeutung der Qualität der Gesundheitsversorgung und Patientensicherheit in der Gesellschaft sowie das Bedürfnis nach Transparenz zunehmen, steigt parallel auch die Belastung des klinischen Personals durch die für die Qualitätssicherung zusätzlich notwendigen Dokumentationen. Die Nutzung von Abrechnungsdaten kann hier eine Lösung bieten. Dies setzt jedoch die systematische Verbesserung der Nutzbarkeit dieser Daten für die Qualitätssicherung durch Gesetzgeber und Selbstverwaltung voraus. Die IQM kann und sollte hierbei als Innovator und Impulsgeber agieren.

Literatur

- Bolczek C, Nimptsch U, Möckel M, Mansky T (2020) Versorgungsstrukturen und Mengen-Ergebnis-Beziehung beim akuten Herzinfarkt – Verlaufsbetrachtung der deutschlandweiten Krankenhausabrechnungsdaten von 2005 bis 2015. *Gesundheitswesen* 82(10), 777–785
- Bray BD, Campbell J, Cloud GC, Hoffman A et al. (2014) Derivation and external validation of a case mix model for the standardized reporting of 30-day stroke mortality rates. *Stroke* 45(11), 3374–3380
- Burgard G (2013) Low-risk-Operationen bei High-risk-Patienten. In: Krahwinkel W, Meier-Hellmann A, Zacher J (Hrsg.) *Peer Review – sicher ist besser*. 49–56. Medizinisch Wissenschaftliche Verlagsgesellschaft Berlin
- Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder (2019) DRG-Statistik 2019. DOI: 10.21242/23141.2019.00.00.1.1.1
- Iezzoni LI (2009) Risk adjustment. In: Smith PC, Mossialos E, Panagoulas I, Leatherman S (Hrsg.) *Performance measurement for health system improvement: Experiences, challenges and prospects*. 251–285. Cambridge University Press
- Initiative Qualitätsmedizin (IQM) (2019) IQM Hearing „Qualitätsmessung in der Medizin – Risikoadjustierung und Risikoadjustierungsmodelle“. URL: https://www.initiative-qualitaetsmedizin.de/fileadmin/downloads/presse/Abstracts_IQM_Hearing_gesamt_WEB_20190221.pdf (abgerufen am 14.07.2022)
- Krumholz HM, Brindis RG, Brush JE, Cohen DJ et al. (2006a) Standards for statistical models used for public reporting of health outcomes – An American Heart Association scientific statement from the quality of care and outcomes research interdisciplinary writing group. Cosponsored by the Council on Epidemiology and Prevention and the Stroke Council. Endorsed by the American College of Cardiology Foundation. *Circulation* 113(3), 456–462
- Krumholz HM, Wang Y, Mattera JA, Wang YF et al. (2006b) An administrative claims model suitable for profiling hospital performance based on 30-day mortality rates among patients with an acute myocardial infarction. *Circulation* 113(13), 1683–1692
- Krumholz HM, Wang Y, Mattera JA, Wang YF et al. (2006c) An administrative claims model suitable for profiling hospital performance based on 30-day mortality rates among patients with heart failure. *Circulation* 113(13), 1693–1701
- Lawthers AG, McCarthy EP, Davis RB, Peterson LE et al. (2000) Identification of In-Hospital Complications From Claims Data: Is It Valid? *Med Care* 38(8), 785–795
- Mansky T, Nimptsch U (2014) Medizinische Qualitätsmessung im Krankenhaus – Worauf kommt es an? *Zeitschrift für Evidenz, Fortbildung und Qualität im Gesundheitswesen* 108(8), 487–494
- Nimptsch U, Mansky T (2017) Hospital volume and mortality for 25 types of inpatient treatment in German hospitals: Observational study using complete national data from 2009 to 2014. *BMJ Open* 7(9). DOI: 10.1136/bmjopen-2017-016184
- Nimptsch U, Mansky T (2020) G-IQI | German Inpatient Quality Indicators Version 5.2 – Bundesreferenzwerte für das Auswertungsjahr 2018. Universitätsverlag der Technischen Universität Berlin
- Schwarzkopf D, Nimptsch U, Graf R, Schmitt J et al. (2021) Möglichkeiten und Grenzen der Risikoadjustierung von Qualitätsindikatoren aus Krankenhausabrechnungsdaten – ein Werkstattbericht. *Zeitschrift für Evidenz, Fortbildung und Qualität im Gesundheitswesen* 163, 1–12
- Zhan C, Elixhauser A, Friedman B, Houchens R et al. (2007) Modifying DRG-PPS to include only diagnoses present on admission: financial implications and challenges. *Med Care* 45(4), 288–291



Dr. phil. Daniel Schwarzkopf, Dipl.-Psych.

Nach dem Studium der Psychologie seit 2010 Tätigkeit in der Versorgungsforschung am Universitätsklinikum Jena mit den Schwerpunkten Qualitätsmessung und -verbesserung der Sepsisversorgung sowie interprofessionelle Zusammenarbeit im Krankenhaus. Seit 2016 Entwicklung und Untersuchung der Validität risikoadjustierter Qualitätsindikatoren im Rahmen des Deutschen Qualitätsbündnisses Sepsis (DQS). Promotion zum Dr. phil. und Erwerb des Zertifikates Epidemiologie in der Medizin (2018). Seit 2019 Leitung des Koordinationsbüros des Deutschen Qualitätsbündnisses Sepsis (DQS).



Raphael Graf

Raphael Graf ist seit 2009 bei 3M Health Information Systems tätig. Aktuell betreut er als Product Owner den Grouper- und Business Intelligence Bereich für die Schweiz. In Deutschland begleitet er Projekte zur Umsetzung von Risikomodellen auf Basis von Routinedaten sowie regelmäßige Analysen zum aktuellen Pandemiegeschehen.

Zuvor war er für verschiedenen Benchmarking-Projekte tätig. Sein Health Care Management (M.Sc.) Studium mit dem Schwerpunkt „Forschung mit Routinedaten“ absolvierte er an der Hochschule Niederrhein.

Dr. Ulrike Nimptsch

Ulrike Nimptsch ist examinierte Krankenschwester und studierte Pflegemanagement sowie Gesundheitswissenschaften (Public Health) mit Schwerpunkt Epidemiologie. Von 2004 bis 2010 war sie Referentin für Qualitätsmanagement und Medizincontrolling bei den HELIOS Kliniken. Von 2010 bis 2018 war sie tätig als Wissenschaftliche Mitarbeiterin an der Technischen Universität Berlin im Fachgebiet Strukturentwicklung und Qualitätsmanagement im Gesundheitswesen, seit 2018 ist Ulrike Nimptsch Wissenschaftliche Mitarbeiterin im Fachgebiet Management im Gesundheitswesen. Ihre Forschungsschwerpunkte sind Krankenhausversorgung, Qualitätsmessung und die Analyse administrativer Daten.