

# 4

## Was wird gemessen? Qualität vs. statistische Auffälligkeit

Oliver Mayer und Vanessa Herrmann

Die Initiative Qualitätsmedizin (IQM) stellt ihren Mitgliedskrankenhäusern mit den German Inpatient Quality Indicators (G-IQI) das derzeit umfassendste, auf Routinedaten basierende Indikatorenset zur krankheits-spezifischen Abbildung von Qualitäts- und Leistungskennzahlen in den Akutkrankenhäusern zur Verfügung. Die zur Berechnung der G-IQI-Kennzahlen herangezogenen Parameter wie Alter, Geschlecht, ICD, OPS und Entlassungsgrund der stationär versorgten Patienten werden aufwandsarm den Krankenhausabrechnungsdaten gemäß § 21 KHEntgG entnommen. Zur kontinuierlichen Verbesserung der medizinischen Versorgungsqualität werden vor allem die Zielwertindikatoren mit ihrer risikoadjustierten Sterblichkeit gemessen an der Standardized Mortality Ratio (SMR) herangezogen.

### 4.1 Von der statistischen Auffälligkeit zur Qualitätsverbesserung – die IQM Systematik in Kürze

Statistisch auffällige Kennzahlen der G-IQI fungieren bei IQM als Aufgreifkriterien zur kontinuierlichen Prozessverbesserung und als Suchmechanismus für potenziell verbesserungsbedürftige Behandlungsabläufe. Die G-IQI tragen somit zur Ergebnisverbesserung im internen Qualitätsmanagement bei, indem sie in einem derartigen kennzahlengestützten QM-Ansatz als Aufgreifmechanismus für Fallanalysen, in Form von trägerübergreifenden Peer Reviews

oder M&M-Konferenzen, dienen. Insbesondere durch die fallbasierten Prozessanalysen im Rahmen eines Peer Reviews wird der gesamte Behandlungsprozess Schritt für Schritt geprüft, um konkrete Verbesserungspotenziale zu detektieren und darauf beruhende gezielte qualitätsverbessernde Maßnahmen einzuleiten.

Dieser kontinuierliche Verbesserungsprozess (KVP) von organisatorischen Abläufen, Prozessen und Strukturen in der Krankenversorgung kommt allen Patienten in Form einer verbesserten Versorgungs- und Behandlungsqualität zugute (Mayer 2011).

## 4.2 Analyse aller Peer Reviews des Jahres 2019

Um diese kurzen inhaltlichen Ausführungen zu prüfen, zu bewerten und bestenfalls zu bestätigen und zu verifizieren, wurde der Eingangsfrage „Was wird gemessen? Qualität vs. statistische Auffälligkeit“ mittels einer Analyse von 191 Peer Reviews aus dem Jahr 2019 nachgegangen, um sie abschließend zu beantworten.

Anhand einer Darstellung der Ergebnisse der zentralen Elemente „Verbesserungspotenzial“ und „Kodierung“ (s. Kap. 4.3 und 4.4) im Rahmen dieser Analyse der Peer Reviews des Jahres 2019 lassen sich die relevantesten und hilfreichsten Bausteine zur Beantwortung der Frage aufzeigen.

## 4.3 Verbesserungspotenzial

In den  $n = 191$  im Jahr 2019 durchgeführten Peer Reviews konnten die Peerteams im Durchschnitt 58,5% und die besuchten Kliniken 40,7% Verbesserungspotenziale ableiten. Dieses Ergebnis lässt sich zudem hinsichtlich der verpflichtenden Peer Reviews aufgrund einer statistischen Auffälligkeit ( $n = 176$ ) und den Pilot-Verfahren ( $n = 15$ ) unterscheiden. Hierbei detektierten die

Peerteams Verbesserungspotenziale von 60,4% (stat. Auffälligkeit) bzw. 36,1% (Pilot), während die Zahlen bei den besuchten Kliniken bei 42,4% (stat. Auffälligkeit) bzw. 20,6% (Pilot) lagen. An dieser Stelle sollen die Ergebnisse von Krahwinkel et al. 2011 sowie Mansky et al. 2013 Erwähnung finden, welche in ihren Ergebnisdarstellungen seinerzeit bereits bei 20% bzw. 30% bis 80% der bei Peer Reviews analysierten Behandlungsfälle und durchgesehenen Akten klare Verbesserungsmöglichkeiten aufzeigten.

Besonders interessant und erwähnenswert ist, dass sich aus der Perspektive der Peerteams (s. Abb. 1) bei allen aufgrund einer stat. Auffälligkeit durchgeführten Peer Reviews ( $n = 176$ ) im Jahr 2019 Verbesserungspotenziale ableiten ließen. Bei  $n = 8$  Peer Reviews wurde sogar ein Verbesserungspotenzial von 100% festgestellt.

Hingegen wurde aus Sicht der besuchten Kliniken bei  $n = 15$  Peer Reviews kein einziges Verbesserungspotenzial gesehen.

Diese Resultate verdeutlichen eindrücklich, dass eine statistische Auffälligkeit selbst nach den Fallanalysen unter den Experten unterschiedlich wahrgenommen wird, und bestätigen damit zugleich, dass das Indikatorergebnis als reiner Zahlenwert nicht zwangsläufig mit guter oder schlechter Qualität gleichzusetzen ist.

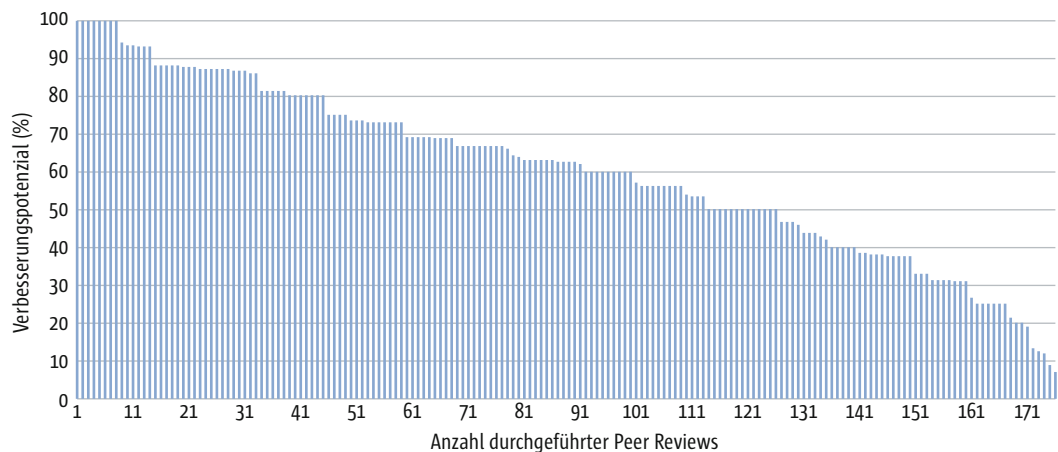
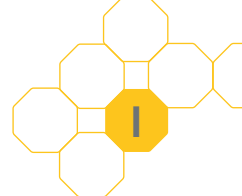


Abb. 1 Verbesserungspotenziale Peerteams



**! Die Ergebnisqualität kann nur auf Fallebene beurteilt werden, da einzig und allein hier die medizinischen Daten sowie deren Dokumentation inkl. Kodierung geprüft werden können.**

Gerade trotz dieser unterschiedlichen Sichtweisen und Blickwinkel ist es nicht nur möglich, dass bereits durch die Analyse grundlegende Probleme Anlass für Veränderungen und Optimierungen geben, sondern dass auch kleine Verbesserungen identifiziert werden können. Allein dadurch wird die Qualität gesteigert, auch wenn sie sich bereits auf einem sehr hohen Ausgangsniveau befindet.

4.4 Kodierung

Bei der Fallanalyse im Rahmen eines Peer Reviews finden nicht nur die medizinisch-organisatorischen Abläufe, Prozesse und Strukturen Berücksichtigung, sondern auch die medizinische Dokumentation inkl. deren Kodierung. Dies ist insofern von Bedeutung, als auch eine falsche Hauptkodierung (Hauptdiagnose oder -prozedur) die Ursache einer statistischen Auffälligkeit sein kann.

In der Auswertung der Ergebnisse aller Peer Reviews aus dem Jahr 2019 ist diesbezüglich vor allem ein Unterschied zwischen den verpflichtenden Peer Reviews aufgrund einer statistischen Auffälligkeit und den Pilot-Verfahren zu erkennen, was Tabelle 1 (Perspektive der Peerteams) verdeutlicht.

Fehlkodierungen traten bei den durch eine statistische Auffälligkeit ausgelösten Peer Reviews mit 55,1% gegenüber den Pilot-Verfahren mit 40,0% deutlich gehäuft auf. Auch wenn die Kodierqualität nicht im Fokus eines Peer Reviews steht, können Kodierproblematiken ein Grund dafür sein, warum Fälle einem G-IQI-Indikator zugewiesen oder diesem abgesprochen werden und warum dieser Qualitätsindikator folglich statistisch auffällig werden kann, unabhängig von der sich dahinter verbergenden medizinischen Behandlungsqualität. Die Frage der Kodierqualität wird im kollegialen Dialog des IQM Peer Reviews, jedoch nur nachrangig diskutiert und findet sich häufig nicht in den Protokollen wieder.

4.5 Annäherung von Qualität und statistischer Auffälligkeit und deren Limitation

Bevor die zentrale Frage dieses Beitrags in Kapitel 4.6 final und eindeutig beantwortet wird, soll in den nächsten Kapiteln 4.5.1 bis 4.5.3 ein Ausblick über eine mögliche künftige Annäherung der zwei in der Frage polarisierend gegenüberstehenden Elemente Qualität vs. statistische Auffälligkeit und deren Limitation gegeben werden.

Tab. 1 Auswertung dokumentierter Kodierfehler (Perspektive Peerteams) (IQM 2019)

	alle	stat. Auffälligkeit	Pilot
Anzahl Peer Reviews (n)	191	176	15
Anzahl Peer Reviews ohne Kodierfehler (n/%)	88/46,1	79/44,9	9/60,0
Anzahl Peer Reviews mit Kodierfehler (n/%)	103/53,9	97/55,1	6/40,0
Fehlkodierungen in Prozent pro Peer Review (Mittelwert)	7,5	8,0	1,8
Fehlkodierungen in Prozent pro Peer Review (Median)	2,5	4	0

#### 4.5.1 Routinedaten

Routinedaten kommen auch bei anderen freiwilligen Initiativen zur Anwendung. Zum Beispiel werden die verschiedenen Sepsisindikatoren beim Deutschen Qualitätsbündnis Sepsis (DQS) ebenso auf der Grundlage von § 21 KHEntgG Abrechnungsdaten berechnet.

Die Grenzen dieser Datensätze, die primär der Abrechnung stationärer Fälle dienen, müssen stets mitberücksichtigt werden, da die Einzigartigkeit eines jeden Patienten hiermit nicht komplett wiedergegeben werden kann. Nichtsdestotrotz überwiegen die Vorteile dieser Datensätze, die schnell und ohne großen Aufwand in der Datenbeschaffung zur Verfügung gestellt werden können.

So liefern die darauf basierend berechneten Qualitätskennzahlen dem Anwender sehr gute Anhaltspunkte („Aufgreifkriterium“) darüber, wo ein Verbesserungspotenzial aufgrund der statistischen Auffälligkeiten bestehen könnte.

#### 4.5.2 Big Data

Optimierungspotenziale hinsichtlich der im vorigen Kapitel 4.5.1 beschriebenen Datenqualität ergeben sich in Zukunft beispielsweise durch die Nutzung von Big Data oder die Anwendung von Künstlicher Intelligenz (KI).



**Je mehr patientenindividuelle Merkmale für eine Qualitätsberechnung und deren Ausweisung anhand von Indikatoren zur Verfügung stehen, desto besser lässt sich die individuelle Versorgungsqualität der Patienten abbilden. Die erweiterte Datenbasis, als Voraussetzung für die Indikator-Berechnungen, hat somit einen direkten positiven Einfluss auf die Detektion möglicher Aufgreifkriterien.**

#### 4.5.3 Risikoadjustierung

Im IQM Pilotprojekt Risikoadjustierung wurden die Möglichkeiten und Grenzen für eine erweiterte und modifizierte Risikoadjustierung von Qualitätsindikatoren aus Krankenhausabrechnungsdaten evaluiert (Schwarzkopf et al. 2021).



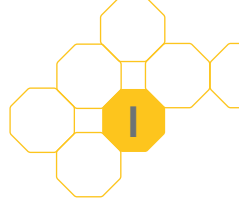
**Durch eine komplexere Risikoadjustierung als die herkömmliche über Alter und Geschlecht kann der Einfluss von Confoundern möglicherweise klein gehalten werden und die patientenindividuellen Einflussfaktoren können präziser in die Berechnung der Indikatoren einfließen.**

#### 4.6 Fazit und Beantwortung der Ausgangsfrage

Unabhängig davon, ob sich hinter einer primär statistischen Auffälligkeit auch eine schlechte medizinische Versorgungsqualität verbirgt, sollte die Chance genutzt werden, Analysen durchzuführen, um Verbesserungsmaßnahmen abzuleiten und diese umzusetzen. Im Zuge dessen werden im Rahmen von Peer Reviews die Fälle sehr oft gut analysiert, besprochen und diskutiert sowie dokumentiert. Die Autoren gehen davon aus, dass im Nachgang immer noch zu wenig Verbesserungsmaßnahmen zur Umsetzung gelangen. Jedoch ist insbesondere der ergebnisorientierte Ansatz mit der Ableitung konkreter Maßnahmen und deren Umsetzung für eine Verbesserung von entscheidender Bedeutung.



- Ziel ist es die Ergebnisqualität langfristig zu verbessern, wobei die Kontinuität der erforderlichen Maßnahmen ebenfalls ausschlaggebend ist.



Die Fragestellung „Was wird gemessen? Qualität vs. statistische Auffälligkeit“ kann abschließend folgendermaßen beantwortet werden:

- Primär wird eine statistische Auffälligkeit gemessen. Diese dient als Aufgreifkriterium für weiterführende Analysen. Hierzu eignen sich M&M-Konferenzen und im besonderen Maße die Fallanalysen im Rahmen eines Peer Reviews. Der hierbei stattfindende interprofessionelle und interdisziplinäre Dialog auf Augenhöhe, außerhalb des klinischen „Normalbetriebs“, bringt nicht nur medizinische, sondern z.B. auch strukturelle Verbesserungsmöglichkeiten zum Vorschein, deren Umsetzung sich schlussendlich auf die künftige Qualität auswirkt.
- Diese Umsetzung wirklich geeigneter Maßnahmen, die langfristig einen Nutzen haben und daher auch die Qualität der Versorgung im Ganzen fördern, ist die eigentliche Leistung, die im Rahmen dieses ganzen Prozesses zu erbringen ist.
- Die initiale statistische Auffälligkeit ist somit der Ausgangspunkt, um im Rahmen des

KVP-Prozesses die Qualität zu sichern und zu erhöhen.

### Literatur

- IQM (2019) Auswertung dokumentierter Kodierfehler (Perspektive Peerteams) [interne Quelle]
- Krahwinkel W, Rink O, Liebetrau M, Günther M, Schuler E, Kühlen R (2011) 10 Jahre Peer Review – Verbesserung der medizinischen Behandlung durch Qualitätsindikatoren aus Routinedaten. Dtsch Med Wochenschr 136, 2083–8
- Mansky T, Nimptsch U, Winklmaier C, Hellerhoff F (2013) G-IQI | German Inpatient Quality Indicators Version 4.0. Band 1: Erläuterungen und Definitionshandbuch für das Datenjahr 2012. Universitätsverlag der TU Berlin
- Mayer O (2011) Organisationsentwicklung im Klinikum Esslingen (KE) zur Generierung von Wachstum. In: Faix W, Djalali A, Horne A, Keck G, Kisgen S, Mezger P, Sailer J (Hrsg.) Management von Wachstum und Globalisierung. Best Practice Bd. 4. 963–998. Steinbeis-Edition Stuttgart
- Schwarzkopf D, Nimptsch U, Graf R, Schmitt J, Zacher J, Kühlen R (2021) Möglichkeiten und Grenzen der Risikoadjustierung von Qualitätsindikatoren aus Krankenhausabrechnungsdaten – ein Werkstattbericht. Zeitschrift für Evidenz, Fortbildung und Qualität im Gesundheitswesen 163, 1–12

### Dr. med. Oliver Mayer, MBA

Nach dem Studium der Humanmedizin in Tübingen Ausbildung zum Facharzt für Allgemeinmedizin. 2010 Master of Business Administration (MBA) in General Management. Von 2010 bis 2012 nebenberufliche Lehrtätigkeit als Projektdozent an der School of International Business and Entrepreneurship (SIBE). Zusatzbezeichnung Ärztliches Qualitätsmanagement in 2011. Seit 2017 Zertifizierter Klinischer Risikomanager nach ONR 49003:2014 und seit 2021 nach ÖNORM 4903:2021. Seit 2011 Arbeit im Klinikmanagement mit den Themenschwerpunkten Qualitäts- und Risikomanagement. Seit 2017 Leiter der Stabsstelle Qualitäts- und Risikomanagement am Universitätsklinikum Ulm. Seit 2017 Mitglied des Fachausschusses Peer Review.



### Vanessa Herrmann, M.Sc.

Bachelorstudium „Informationsmanagement im Gesundheitswesen“ an der Hochschule Ulm/Neu Ulm. Anschließend Masterstudium „Medical Process Management“ an der Friedrich-Alexander-Universität Erlangen-Nürnberg. Seit 2019 Mitarbeiterin der Stabsstelle Qualitäts- und Risikomanagement am Universitätsklinikum Ulm. Mitglied im Fachausschuss Indikatoren.

