



Universität St.Gallen

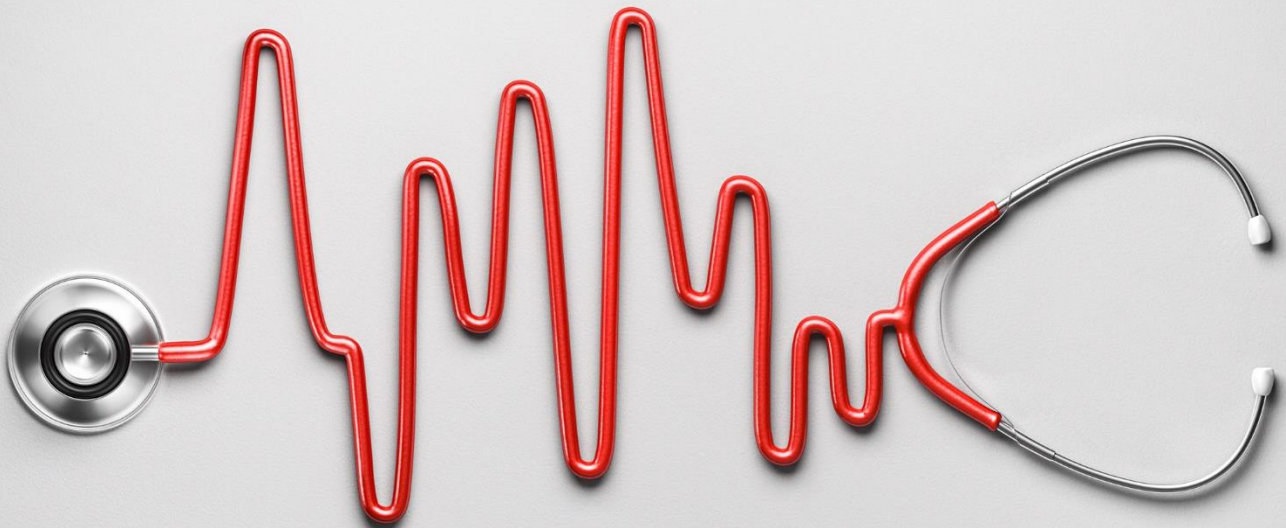
Verwendung medizinischer Ergebnisqualität in der Spitalplanung

Kurzgutachten

Dr. Justus Vogel, Dr. David Kuklinski, Prof. Dr. Alexander Geissler

Lehrstuhl Management im Gesundheitswesen, School of Medicine (MED-HSG), Universität St. Gallen

St.Gallen, 08. Mai 2023



Inhaltsverzeichnis

| | |
|--|----|
| Abkürzungsverzeichnis | II |
| 1 Einleitung | 1 |
| 2 Bisherige Initiativen | 2 |
| 2.1 CH-IQIs | 2 |
| 2.2 ANQ | 3 |
| 2.3 Ziele beider Qualitätstransparenzinitiativen | 3 |
| 3 Methodische Bedingungen | 4 |
| 4 Verwendbarkeit bestehender Indikatoren | 6 |
| 4.1 CH-IQIs | 6 |
| 4.2 ANQ-Indikatoren | 10 |
| 5 Zusammenfassende Empfehlung | 14 |

Abkürzungsverzeichnis

| | |
|-------------|--|
| ANQ | Nationaler Verein für Qualitätsentwicklung in Spitälern und Kliniken |
| BAG | Bundesamt für Gesundheit |
| BfS | Bundesamt für Statistik |
| CH-IQI..... | Swiss Inpatient Quality Indicators |
| CHOP | Schweizerische Operationsklassifikation |
| G-IQI..... | German Inpatient Quality Indicators |
| ICD..... | International Classification of Diseases |
| IVHSM..... | Interkantonalen Vereinbarung zur Hochspezialisierten Medizin |

1 Einleitung

Die Spitalplanung in der Schweiz beschränkt sich bis anhin auf Vorgaben der Struktur- und Prozessqualität. Die Verwendung von Ergebnisqualität in der Spitalplanung wird bislang nicht verfolgt – abgesehen der Leistungsgruppen, die Teil der Interkantonalen Vereinbarung zur Hochspezialisierten Medizin (IVHSM) sind. Ansonsten besteht lediglich die Verpflichtung, an den Erhebungen des Nationalen Vereins für Qualitätsentwicklung in Spitälern und Kliniken (ANQ) teilzunehmen. Im Folgenden wird erörtert, inwiefern und in welcher Weise Ergebnisqualität zukünftig direkt in der Spitalplanung (Akutsomatik) berücksichtigt werden könnte.

Zu diesem Zweck wird zuerst in Kürze auf zwei bestehende Initiativen zur Ergebnisqualitätsmessung in der Schweiz, namentlich die Swiss Inpatient Quality Indicators (CH-IQI) und den ANQ, eingegangen. Anschliessend werden drei methodische Bedingungen zur Verwendung von Ergebnisqualität in der Spitalplanung diskutiert. Schliesslich wird in fünf Schritten geprüft, inwiefern sich die CH-IQIs bzw. ANQ-Indikatoren als leistungsgruppenspezifische Anforderungen in der Spitalplanung (Akutsomatik) eignen könnten. Abschliessend folgt eine zusammenfassende Empfehlung.

Es gilt zu beachten, dass es sich beim vorliegenden Dokument um ein Kurzgutachten handelt. Das bedeutet, dass bündig und zielgerichtet argumentiert wird und keine allgemeine wissenschaftliche Herleitung anhand von Quellen erfolgt. Als Links im Dokument eingefügte Quellen sind als Hinweise an die besonders interessierte Leserschaft, die sich vertiefend mit wissenschaftlichen Auswertungen zu einem bestimmten Thema auseinandersetzen möchte, zu verstehen.

2 Bisherige Initiativen

In der Schweiz bestehen mehrere Initiativen zur Messung und Veröffentlichung von Ergebnisqualität von akutsomatischen Spitälern. Im Kontext der Spitalplanung und im Rahmen dieses Kurzgutachtens werden die CH-IQIs und ANQ-Indikatoren vorgestellt.

2.1 CH-IQIs

Die CH-IQIs werden durch das Bundesamt für Statistik (BfS) aufbereitet und durch das Bundesamt für Gesundheit (BAG) zur Verfügung gestellt (siehe [Link](#)). Die Indikatoren stammen ursprünglich aus den USA, wurden später in Deutschland als German Inpatient Quality Indicators (G-IQI) weiterentwickelt und in die Schweiz überführt. Sie basieren grundsätzlich auf akutstationären Abrechnungsdaten, die routinemässig in jedem Spital erhoben werden. Wie der Name nahelegt, beschränken sie sich auf die Betrachtung des Behandlungsergebnisses während des akutstationären Aufenthalts, also im Spital selbst. Spätere, post-stationäre unerwünschte Ereignisse, die z.B. im Laufe der Nachbehandlung auftreten, werden durch stationäre Abrechnungsdaten nicht erfasst und werden somit nicht durch die CH-IQIs gemessen.

Die CH-IQIs umfassen 65 Krankheitsbilder bzw. Behandlungen mit je mind. einem und bis zu 32 Indikatoren. Die Indikatoren bestehen aus Fallzahlen, Mortalität, Anteilswerten wie z.B. der Anteil an Kaiserschnittgeburten an allen Geburten, Verlegungsraten, Verweildauerangaben oder Spezialinformationen (z.B. Indikator «C.1.3.S Mittelwert des Barthel-Index bei Aufnahme» im Bereich der geriatrischen Frührehabilitation). Angaben zur Mortalität sind in der Regel durch Berücksichtigung des Alters und Geschlechts risikoadjustiert. Ausserdem werden Mortalitätsindikatoren einerseits für die Gesamtstichprobe und andererseits oftmals zusätzlich stratifiziert nach Altersgruppen berechnet. Auf der Webseite des BAG lässt sich ein Vergleich der Ergebnisse von gleichzeitig maximal fünf Indikatoren zwischen maximal fünf Spitälern durchführen (siehe [Link](#) zum Vergleichsportal). Darüber hinaus ist ein Download der Ergebnisse pro Spital auf der BAG-Webseite möglich.

Ziel der Berechnung und Veröffentlichung der CH-IQIs ist v.a. die Ermöglichung eines Benchmarkings der Spitäler durch Analyse der zur Verfügung gestellten Rohdaten. Da dies für Laien recht aufwendig ist und einiges an Fachwissen erfordert, werden solche Benchmarkings durch Gesundheitsbehörden, Universitäten und nicht zuletzt von den Spitälern selbst durchgeführt. Dadurch bietet sich den Spitälern die Möglichkeit, Handlungsfelder zur Verbesserung der eigenen Qualität zu identifizieren und Ziele hierfür abzuleiten (siehe auch [Moser et al., 2022](#)). Das Vergleichsportal mag auch von Patienten genutzt werden, um sich vor der Wahl für ein Spital über die Qualität des Spitals zu informieren. Jedoch ist eine Verwendung des Vergleichsportals durch breite Bevölkerungsschichten schwer vorstellbar, da die Vergleichsfunktionalität recht rudimentär ist, die Bedeutung und Interpretation der Indikatoren für medizinische Laien nicht intuitiv ist und die Darstellung der Ergebnisse im Tabellenformat die Verständlichkeit nicht unterstützt (siehe auch [von Ahlen et al., 2022](#)).

2.2 ANQ

Der ANQ berechnet und veröffentlicht – neben z.B. Indikatoren zur Patientenzufriedenheit – eine Reihe an Ergebnisqualitätsindikatoren für die Akutsomatik¹ (siehe Link; ausserdem werden die Ergebnisse den Mitgliedern des ANQs in Berichtsformat und in Zukunft auch als Dashboard zur Verfügung gestellt). Die ANQ-Auswertungen haben vordergründig zum Ziel, Spitälern ein objektives Benchmarking zu ermöglichen. Darüber hinaus können Patienten die Informationen nutzen, um eine informierte Spitalwahl zu tätigen. Auch sollen die Qualitätsinformationen Kostenträgern, d.h. Versicherern und Kantonen, als Quelle für Tarifverhandlungen bzw. Leistungsvereinbarungen und zum Qualitätsmonitoring dienen.

Der ANQ veröffentlicht Ergebnisqualitätsindikatoren auf Spitalstandortebene zu postoperativen Wundinfektionen, Sturz und Dekubitus, potenziell vermeidbaren Rehospitalisierungen und Ergebnisse aus dem Schweizerischen-Implantats-Register (SIRIS) für Wirbelsäule (bisher noch keine Ergebnisse publiziert), Hüfte und Knie (z.B. 2-JahresRevisionsraten für Hüft- und Knie totalprothesen bei primärer Arthrose). Wo sinnvoll, sind Ergebnisqualitätsindikatoren risikoadjustiert. Zur Risikoadjustierung werden neben Alter und Geschlecht eine Reihe an krankheits- bzw. behandlungsspezifischen Einflussfaktoren wie z.B. Ko-Morbidität (Diagnosen), der Body-Mass-Index oder ein Score zur Gehfähigkeit genutzt.

2.3 Ziele beider Qualitätstransparenzinitiativen

Beide Initiativen messen Ergebnisqualität mit einer Reihe an Indikatoren für medizinisch spezifische, klar definierte Prozeduren bzw. Krankheitsbilder (z.B. Herzinfarkt oder postoperative Wundinfektionsrate für Chlozystektomien). Ziel beider Initiativen ist die Herstellung von Qualitätstransparenz zur Verbesserung der Behandlungsqualität für ebendiese Prozeduren bzw. Krankheitsbilder.

Denn Transparenz kann – wenn die Qualitätsinformationen für Patientinnen und Patienten zugänglich und verständlich sind – zu informierten Entscheidungen für oder gegen ein Spital führen (siehe auch [Kuklinski et al., 2021](#)), was einen direkten Anreiz zur Qualitätsverbesserung für Spitäler darstellt. Neben Patientinnen und Patienten sprechen beide Initiativen also auch Spitäler an, die durch Benchmarks angeregt werden sollen, ihre Behandlungsstrukturen und -prozesse zu hinterfragen, Ziele abzuleiten und nach Qualitätsverbesserung zu streben. Nicht zuletzt können Kantone und Versicherungen die Qualitätsinformationen für Zwecke des Monitorings und Versorgungssteuerung nutzen.

¹ Der ANQ berechnet und veröffentlicht ausserdem Qualitätsindikatoren für Rehakliniken und psychiatrische Einrichtungen / Spitäler. In diesem Kurzgutachten fokussieren wir uns jedoch auf die Planung der akutsomatischen Spitalversorgung.

3 Methodische Bedingungen

In der Spitalplanung werden die Leistungsaufträge der Spitäler festgelegt. Spitalplanung stellt damit einerseits ein direktes Instrument zur Beeinflussung bzw. Zentralisierung von Patientenströmen und damit potenziell verbundenen Qualitätsgewinnen dar. Andererseits wird bei Nicht-Vergabe bzw. Entzug eines Leistungsauftrags das Leistungsangebot eines Spitals direkt und dauerhaft eingeschränkt. Damit ein solcher Eingriff in das Leistungsangebot gerechtfertigt ist und Patientenströme tatsächlich qualitätsorientiert geleitet werden, müssen bei der Verwendung von Ergebnisqualität mehrere (methodische) Bedingungen erfüllt sein.

Erstens müssen die verwendeten Indikatoren tatsächlich aussagekräftig für die Qualität des zu bewertenden Spitals bzw. der zu bewertenden Leistungsgruppe sein. Dies bedeutet, dass (i) die zur Berechnung des Indikators eingeschlossene Patientengruppe zum überwiegenden Teil der Patientengruppe eines Spitals bzw. einer Leistungsgruppe entspricht. Ausserdem muss (ii) im Einzelfall bewertet werden, ob ein einzelner Indikator für die Bewertung eines Spitals bzw. einer Leistungsgruppe ausreicht oder ob mehrere Indikatoren separat oder gemeinsam als Index herangezogen werden sollten. Im Falle von leistungsgruppenspezifischen Anforderungen ist zu beachten, dass (iii) Ergebnisqualitätsindikatoren nur für ausreichend homogene Leistungsgruppen genutzt werden können. Denn nur, wenn die Prozeduren einer Leistungsgruppe medizinisch sehr ähnlich sind, kann ein Indikator aussagekräftig für das Ergebnis der medizinischen Behandlung sein. Sobald verschiedene Operationen, Behandlungen, Prozeduren und diagnostische Leistungen für unterschiedliche Krankheitsbilder in einer Leistungsgruppe subsummiert sind, kann nicht mehr eindeutig zurückgeführt werden, zum Ergebnis welcher Leistung eine Aussage getroffen werden soll. Ein Index-Score aus mehreren Indikatoren könnte für heterogene Leistungsgruppen herangezogen werden. Jedoch ist die Umsetzung schwierig, da dann eine Gewichtung der Indikatoren notwendig ist. Solche Gewichtungen sind jedoch kaum «objektiv» vorzunehmen, sondern bleiben in der Regel arbiträr.

Zweitens ist eine ausreichende Fallzahl zur Berechnung des Indikators bzw. der Indikatoren, die für eine Leistungsgruppe herangezogen werden sollen, pro Spital essenziell. Für Spitäler, die nur vereinzelt Fälle versorgen, ist methodisch gesehen der Zufallseinfluss zu gross, weshalb die Berechnung von Indikatoren für solche Spitäler mit großer Vorsicht zu geniessen ist. Dann bliebe jedoch festzulegen, ob sich Spitäler mit zu geringen Fallzahlen zur robusten Schätzung der Ergebnisqualität überhaupt auf einen Leistungsauftrag bewerben dürften. Behelfsmässig könnten zur Berechnung eines Indikators die Fälle mehrerer Jahre herangezogen werden. Hierbei ist zu beachten, dass die Fälle jedoch nicht zu weit in der Vergangenheit liegen, da sich in der Zwischenzeit Veränderungen u.a. in der Struktur- und Prozessqualität ergeben haben könnten, die die Ergebnisqualität beeinflussen. Ein Zwei- bis Dreijahreszeitraum könnte vor dem Hintergrund dieser Überlegungen vertretbar sein. Jedoch könnten für hochkomplexe Leistungsgruppen auch dann die Fallzahlen mehrerer Spitäler nicht für eine stabile Berechnung ausreichen. Der ANQ stellt aufgrund der Limitation geringer Fallzahlen die Indikatorwerte pro Spital korrekterweise als Grafik mit Konfidenzintervallen bzw. dem Signifikanzniveau dar (siehe z.B. [NNIS-Adjustierte Infektionsrate für Appendektomien Erwachsene, älter 15 Jahre](#)). Wie vom ANQ veranschaulicht und oftmals empirisch zu beobachten, sind die Konfidenzintervalle bei Spitaler mit kleinen Fallzahlen

sehr breit, d.h. die Irrtumswahrscheinlichkeit, dass der berechnete Wert dem tatsächlichen «Qualitätswert» entspricht, ist relativ hoch.

Drittens müssen die berechneten Werte der verwendeten Indikatoren auf Spitalebene tatsächlich die Qualität der Versorgung ausdrücken. Das heisst, dass der Einfluss von Faktoren ausserhalb der Kontrolle eines Spitals methodisch adäquat berücksichtigt werden muss. Neben den personellen, medizinisch-technischen und baulichen Strukturen, Prozessadhärenz und personeller Qualifikation auf Spitalseite wird Ergebnisqualität v.a. durch patientenseitige Risikofaktoren beeinflusst. Je nach Berechnungsart und Indikator wird entweder gar nicht, ausschliesslich für Alter bzw. Altersgruppe und Geschlecht oder auch für weitere Faktoren wie Ko-Morbidität adjustiert.

Abschliessend sei noch auf drei allgemeine Aspekte eingegangen, die es bei einer Verwendung von Ergebnisqualität für die Spitalplanung zu beachten gilt. Erstens sollten die Ergebnisse der Indikatoren über mehrere Jahre (z.B. drei Jahre) gemittelt werden, da ein Indikatorwert in einem Jahr auch bei ausreichender Fallzahl ausserordentlichen (Zufalls-) Faktoren unterliegen könnte. Zweitens können Indikatoren, die Abrechnungsdaten nutzen, wie z.B. CH-IQI und zu grossen Teilen der ANQ, immer nur mit einem Zeitverzug, normalerweise von zwei Jahren, berechnet werden. Beispielsweise könnte sich eine Spitalplanung 2023, die ab 2024 bindend würde, auf Indikatoren von 2020/21 und älter stützen. Eine Mittelung mehrerer Jahre kann dieses Manko teilweise ausgleichen, aktuelle und kürzliche potenzielle Qualitätsverbesserungen können jedoch nicht quantitativ berücksichtigt werden. Drittens ist der Beobachtungszeitraum, den ein Indikator einschliesst, von grosser Wichtigkeit für seine Aussagekraft und Relevanz für die zu bewertende Leistung bzw. Leistungsgruppe. So betrachten z.B. die CH-IQIs nur die Zeit des Spitalaufenthalts, was für manche Leistungen angemessen sein kann (z.B. Mortalität bei Herzinfarkt), für andere Leistungen jedoch ein zu kurzer Zeitraum wäre und zu Fehlanreizen führen könnte (vgl. Kapitel 4.1).

4 Verwendbarkeit bestehender Indikatoren

Eine Einschätzung, ob ein Ergebnisqualitätsindikator als leistungsgruppenspezifische Anforderung zielgerecht verwendet werden kann, muss aus unserer Sicht und aufbauend auf Kapitel 3 in fünf Prüfschritten erfolgen.

Zunächst sollte geprüft werden, (1) welche Leistungsgruppen homogen genug sind, um von einem bzw. mehreren Ergebnisqualitätsindikatoren beschrieben zu werden. Anschliessend müssen (2) die zur Bewertung der Ergebnisqualität der Leistungsgruppe passenden Indikatoren ausgewählt werden. Schliesslich muss (3) im Detail auf Ebene der Diagnose- und/ oder Prozedurencodes geprüft werden, ob die durch den (die) Indikator(en) eingeschlossene Patientengruppe(n) mit der Leistungsgruppe übereinstimmt. Durch diese ersten drei Prüfschritte ist geklärt, ob es generell möglich ist, die Qualität einer Leistungsgruppe durch bestehende Indikatoren ausreichend zu beschreiben. Sollte dies nicht der Fall sein, also kein «medizinischer Match», d.h. keine ausreichend grosse Überschneidung der betrachteten Patientengruppen, vorliegen, können die Indikatoren nur eingeschränkt oder gar nicht verwendet werden.

Darüber hinaus schliessen sich zwei weitere Schritte an, die die methodische Verlässlichkeit der Indikatoren prüfen. Und zwar muss geprüft werden, wie (4) mit der Limitation kleiner Fallzahlen umgegangen und (5) ob und wie eine Adjustierung patientenseitiger Risiken für das Behandlungsergebnis vorgenommen wird.

Im Folgenden werden diese fünf Prüfschritte anhand beispielhafter Leistungsgruppen und Indikatoren für beide im Kapitel 2 vorgestellten Initiativen durchgeführt und eine abschliessende Bewertung vorgenommen. In die abschliessende Bewertung fliessen ausserdem die Überlegungen zu den drei allgemein zu beachtenden methodischen Aspekte wo relevant mit ein.

4.1 CH-IQIs

(1) Prüfung der Homogenität

Als beispielhafte Leistungsgruppe wird die Leistungsgruppe «KAR3 Interventionelle Kardiologie (Koronareingriffe)» herangezogen. Die Leistungsgruppe scheint «auf den ersten Blick» recht homogen, schliesst jedoch Patienten mehrerer kardiologischer Diagnosegruppen ein, die interventionell behandelt werden bzw. werden könnten. Die Definition erfolgt auf Ebene von ICD 4- bzw. 5-Stellern und CHOP 5- bzw. 7-Stellern. Als vereinfachte Veranschaulichung seien einige Auszüge auf Ebene der ICD 3-Steller und CHOP 4- bis 6-Steller genannt²:

² Es ist zu beachten, dass es sich um eine vereinfachte Darstellung auf ICD 3-Steller und CHOP 4- bis 6-Steller handelt, also nicht alle detaillierten Sub-Codes Teil der Leistungsgruppe sind. Ausserdem ist der Algorithmus zur Zuweisung der Leistungsgruppe zu einem Fall höchst komplex,

1. Auszug der Diagnosegruppen (ICD 3-Steller, Auszug)
 - I20: Angina pectoris
 - I21: Akuter Myokardinfarkt
 - I22: Rezidivierender Myokardinfarkt
 - I23: Bestimmte akute Komplikationen nach akutem Myokardinfarkt
 - I24: Sonstige akute ischämische Herzkrankheit
 - I25: Chronische ischämische Herzkrankheit
 - I71: Aortenaneurysma und -dissektion
2. Auszug der Prozedurengruppen (CHOP 4- bis 6-Steller, Auszug)
 - Z00.66.2: Koronarangioplastik [PTCA], nach Art der verwendeten Ballons
 - Z00.66.3: Koronarangioplastik [PTCA], nach Verfahren
 - Z37.22: Linksherzkatheter
 - Z37.23: Kombiniertes Rechts- und Linksherzkatheter
 - Z37.25.10: Geschlossene [perkutane] [Nadel-] Biopsie des Myokards
 - Z37.28: Intrakardiale Echokardiographie [IKE]
 - Z37.2A.1: Intravaskuläre Bildgebung der Koronargefäße
 - Z37.2A.2: Invasive intravaskuläre Koronardiagnostik mittels Katheter

Es zeigt sich also, dass unterschiedliche Krankheitsbilder sowohl akuter als auch chronischer Natur sowie verschiedene diagnostische und therapeutische Leistungen in der Leistungsgruppe subsummiert sind. Entsprechend breit müsste die Definition der zu verwendenden Indikatoren sein.

(2) Selektion von CH-IQIs

Sucht man nun im CH-IQI-Definitionsbuch (siehe [Link](#)) nach passenden Ergebnisqualitätsindikatoren, kommen «auf den ersten Blick» eine Reihe an Indikatoren für diese Leistungsgruppe in Frage:

- A.1.1.M HD Herzinfarkt (Alter >19), Mortalität
- A.1.8.M HD Herzinfarkt (Alter >19), Zuverlegungen, Mortalität
- A.3.2.M Katheter der Koronargefäße bei HD Herzinfarkt ohne OP am Herzen (Alter >19), Mortalität
- A.3.5.M Therapeutische Katheter der Koronargefäße ohne HD Herzinfarkt ohne OP am Herzen (Alter >19), Mortalität
- A.3.8.M Therapeutische Katheter der Koronargefäße ohne HD Herzinfarkt ohne weitere OP (Alter >19), Mortalität

(3) Vergleich der Patientengruppen der Leistungsgruppe und der CH-IQIs

Tatsächlich betrachten die Indikatoren A.1.1., A.1.8. und A.3.2. nur Patienten mit Hauptdiagnose Herzinfarkt, definiert durch die ICD-Codes I21 oder I22, und die Indikatoren A.3.5. und A.3.8.

d.h. Fällen mit den hier beispielhaft aufgeführten Codes könnte u.U. auch eine andere Leistungsgruppe zugewiesen werden.

eine Vielzahl (spezifischer) kardiologischer Diagnosen. Ähnliches gilt für die Auswahl der CHOP-Codes der Indikatoren des Kapitels A.3.

Die Prüfung der Schritte (1), (2) und (3) zeigt bereits, dass selbst für eher homogene, medizinisch «anspruchsvolle» Leistungsgruppen eine Beschreibung mit einzelnen Indikatoren nicht ohne weiteres möglich ist, da die zugrundeliegenden ICD- und CHOP-Codes nur teilweise übereinstimmen.

(4) Umgang mit kleinen Fallzahlen

Generell werden Fallzahl CH-IQIs nur berichtet, wenn mindestens ein Fall behandelt wurde. Darüber hinaus werden Raten nur berechnet, wenn mindestens zehn Fälle für die Berechnung herangezogen werden können. Dies ist aus methodischer Sicht richtig. Die oben beschriebene Herausforderung in der Spitalplanung, wie mit Bewerbungen von Spitälern mit kleiner Fallzahl umgegangen werden soll, für die aufgrund der geringen Fallzahl keine Indikatoren berechnet werden können, bliebe jedoch zu klären. Gleichzeitig müsste geklärt werden, wie Werte einzelner Jahre zu interpretieren sind, insbesondere für Spitäler mit geringen Fallzahlen.

(5) Einschätzung zur Risikoadjustierung

Hinsichtlich der Risikoadjustierung der CH-IQIs bleibt schliesslich festzustellen, dass zwar eine basale Adjustierung um Alter und Geschlecht vorgenommen wird, diese jedoch für den Eingriff in das Leistungsangebot der Spitäler durch die Spitalplanung nicht ausreichend erscheint. Zumindest Ko-Morbidität und bestenfalls noch weitere Faktoren wie Body-Mass-Index sollten hinzugezogen werden. Für eine Reihe an CH-IQIs, beispielsweise auch für den Indikator «A.1.1.M HD Herzinfarkt (Alter >19), Mortalität», wird zusätzlich zur Risikoadjustierung eine Stratifizierung nach Altersgruppen durchgeführt (Indikatoren A.1.2.M bis A.1.5.M). Durch diese Stratifizierung werden wertvolle Detailinformationen zu Risiken spezifischerer Patientengruppen gewonnen. Jedoch können diese Indikatoren wiederum noch weniger mit einer Leistungsgruppe gematched werden, da in einer Leistungsgruppe alle behandelten Altersgruppen inkludiert sind. So können alterstratifizierte CH-IQIs nur Qualitätsinformationen für Untergruppen der Leistungsgruppe liefern. Basierend hierauf ein Spital von der Leistungserbringung auszuschliessen, d.h. den Leistungsauftrag nicht zu vergeben, erscheint schwierig zu rechtfertigen.

Abschliessende Bewertung

Das hier in den Prüfschritten (1) bis (5) diskutierte Beispiel «KAR3 Interventionelle Kardiologie (Koronareingriffe)» scheint repräsentativ für die Mehrheit der Leistungsgruppen. Der in Schritten (1) bis (3) herausgearbeitete «Mismatch» der Patientengruppen gründet auf der unterschiedlichen Methodik zur Definition der Patientengruppen und verschiedener Zielsetzung der SPLG-Systematik und der CH-IQIs. Die SPLG-Definitionen passen «besser» zu den CH-IQIs für Leistungsgruppen der IVHSM wie etwa «VIS1.1 Grosse Pankreaseingriffe». Da die Anforderungen für diese Gruppen jedoch interkantonal geregelt sind (und teilweise bereits Ergebnisqualität beinhalten), sind (weitere) kantonale Anforderungen in der Spitalplanung nicht angezeigt bzw. – zumindest unter Verwendung von CH-IQIs – nicht weiter nötig. Insgesamt scheinen CH-IQIs also nicht das passende Instrument zur Inkludierung von Ergebnisqualität in der Spitalplanung. Um dieses qualitative Analyseergebnis zum «medizinischen Match» quantitativ zu bestätigen, könnten Daten der Medizinischen Statistik genutzt werden. So könnte ausgewertet werden, wie hoch die Fallzahl für einen bzw. mehrere Indikatoren, die zur Bewertung einer Leistungsgruppe herangezogen werden sollen, ist und welchen Anteil diese Fälle an der Gesamtfallzahl der

Leistungsgruppe hat. Sobald ein überwiegender Teil der Fälle der Leistungsgruppe abgebildet ist, wäre das Problem des «Mismatch» ggf. ausreichend behoben.

Erschwerend kommt jedoch hinzu, dass Prüfschritt (5) gezeigt hat, dass die Risikoadjustierung der CH-IQIs im Kontext der Spitalplanung kritisch zu sehen ist. Eine Adjustierung nach Alter und Geschlecht mag für andere Zwecke wie das Benchmarking eigener Leistungen in derselben Gruppe von Spitälern (z.B. ein Universitätsspital vergleicht sich mit anderen Universitätsspitalern) oder Monitoring von besonders negativ auffälligen Ausreißern ausreichen. Für den starken Eingriff der Spitalplanung in das Leistungsangebot der Spitäler erscheint die Risikoadjustierung jedoch zu basal.

Neben dem fragwürdigen «medizinischen Match» und der zu basalen Risikoadjustierung spricht gegen eine Verwendung von CH-IQIs, dass das nach Alter und Geschlecht adjustierte Mortalitätsratio im Spital der einzige «eindeutige» Ergebnisqualitätsindikator der CH-IQIs ist. Alle weiteren Indikatoren sind Surrogatparameter wie Fallzahlen, Verweildauer, Verlegungsraten und Anteilsraten (z.B. Kaiserschnitttrate), die auf gute bzw. schlechte Ergebnisqualität hindeuten können, jedoch nicht zwingend für ein einzelnes Spital kausal mit der Ergebnisqualität des Spitals verbunden sein müssen.

Darüber hinaus kann Mortalität im Spital die Qualität einer Leistung nur dann hinreichend gut beschreiben, wenn (i) sie relevant für die Leistung ist und (ii) Qualitätsvariation transparent machen kann. Ersteres kann datenseitig durch deskriptive Kennzahlen, z.B. eine Auswertung der Verteilung von Mortalitätsraten je Spital mittels eines Boxplots, geprüft werden. So wird transparent, ob Todesfälle während des Spitalaufenthalts tatsächlich regelhaft beobachtet werden (z.B. Herzinfarkt oder Schlaganfall) oder ob es sich eher um Einzelfälle und Sonderereignisse handelt (z.B. bei Erstimplantation Knie). Zusätzlich ist medizinisch-sachlogisch zu bewerten, ob für die jeweilige Leistungsgruppe Patienten in kritischem Zustand tatsächlich am ersten Tag im Spital bzw. während des Spitalaufenthalts versterben oder nicht. Sollte eine Entlassung nach Hause oder die Verlegung auf eine (externe) Palliativstation oder in ein Hospiz im Regelfall durchführbar sein, würden Todesfälle durch den Indikator nicht beobachtet werden. Eine Fokussierung auf Mortalität im Spital könnte sogar dazu führen, dass Patienten mit geringen/ ohne Heilungschance vermehrt früher entlassen werden, um den Indikatorwert nicht zu belasten. Aus diesem Grund messen Initiativen wie die Qualitätssicherung mit Routinedaten der in Deutschland aktiven «Allgemeinen Ortskrankenkassen (AOK)» Mortalität inkl. eines post-stationären Zeitraums (z.B. 30 oder 90 Tage, vgl. [Link](#)). Im Übrigen ist die hier beschriebene Limitation nicht ausschliesslich für Mortalitätsindikatoren relevant, sondern muss für alle Ergebnisqualitätsindikatoren beachtet werden. Hinsichtlich (ii) kann datenseitig geprüft werden, wie die Varianz der Indikatorwerte ist, d.h. es wird geprüft, ob durch den Indikator tatsächlich eine bestehende Qualitätsvariation transparent gemacht wird oder nicht. Denn wenn Spitäler Todesfälle verhindern können, dann muss es Spitäler geben, die wesentlich besser (bzw. schlechter) sind als andere. Sollte die Varianz niedrig sind, könnte Mortalität entweder nicht relevant für die zu bewertende Leistung sein, d.h. einfach zu selten vorkommen (vgl. (i)) oder Mortalität hängt unzureichend von Aktivitäten des Spitals ab oder alle Spitäler sind sehr gut in der Verhinderung der Mortalität. In jedem Fall ist die Aussagekraft des Indikators beschränkt.

4.2 ANQ-Indikatoren

Die ANQ-Indikatoren werden anhand derselben Prüfschritte wie oben diskutiert. Hierbei wird jedoch auf alle ANQ-Indikatoren allgemein eingegangen und nicht auf einzelne Beispiele. Entsprechend werden Schritte (1) bis (3) zusammengefasst betrachtet.

(1) Homogenität, (2) ANQ-Indikatoren und (3) Vergleich der Patientengruppen

Hinsichtlich der Homogenität der Leistungsgruppen gilt das obige Analyseergebnis: Generell scheinen auf den ersten Blick eine Reihe an Leistungsgruppen homogen genug, um durch Ergebnisqualitätsindikatoren hinreichend bewertet werden zu können. Jedoch stellt man bei genauerer Analyse fest, dass auch bei relativ homogenen Leistungsgruppen mehrere Indikatoren nötig sind, da die (im Falle der CH-IQIs) durch Indikatoren adressierte Patientengruppen in der Regel wesentlich spezifischer sind. Alternativ müssen Indikatoren gesucht werden, die generelle Rückschlüsse auf die Ergebnisqualität zulassen. Schliesslich sei wiederholt, dass Gruppen der IVHSM am spezifischsten und damit am besten geeignet für Ergebnisqualitätsindikatoren sind, jedoch durch die kantonale Spitalplanung nicht (weiter) reguliert werden können bzw. müssen.

Der ANQ misst bzw. veröffentlicht für die Akutsomatik folgende Indikatoren:

1. Nosokomiale Dekubitusprävalenz (bis 2019)
2. Stürze im Spital (bis 2019)
3. Potenziell vermeidbare Rehospitalisationen innerhalb von 30 Tagen (bis 2019)
4. Postoperative Wundinfektionsraten
5. Adjustierte 2-Jahres-Revisionsrate primäre Hüfttotalprothese bei primärer Arthrose (analog für Knie)

Alle Indikatoren werden auf Spitalebene gemessen und veröffentlicht. Die ersten drei Indikatoren umfassen bzw. stehen für alle Fälle, die ein Spital behandelt. Hierdurch scheinen diese Indikatoren allgemein passend für alle Leistungsgruppen, da keine Patientengruppen ausgeschlossen werden. Gleichzeitig kann ein Indikator auf Spitalebene keine leistungsgruppenspezifische Anforderung darstellen. Der ANQ müsste also prinzipiell diese Indikatoren auf Leistungsgruppenebene weitergeben (dürfen), um eine leistungsgruppenspezifische Verwendung zu ermöglichen. Da die Kantone Mitglieder des ANQ sind, scheint dies aber keine Hürde darzustellen.

Betrachtet man nun noch die Aussagekraft der ersten zwei Indikatoren für die Ergebnisqualität einer Leistungsgruppe, so muss festgestellt werden, dass die Indikatoren «nosokomiale Dekubitusprävalenz» und «Stürze im Spital» weniger gut geeignet sind, um die Qualität der Versorgung in einer spezifischen Leistungsgruppe zu bewerten. Tatsächlich können durch diese Indikatoren allgemeine Aussagen über die Pflege- und post-operativen Prozesse des Spitals abgeleitet werden, jedoch weniger über die Behandlung sowie Pflege- und post-operativen Prozessen von Fällen einer spezifischen Leistungsgruppe.

Demgegenüber können mittels des Indikators «potenziell vermeidbare Rehospitalisationen innerhalb von 30 Tagen» Aussagen über internistische, über interventionelle und auch über chirurgische Leistungsgruppen getroffen werden. Ein erneuter Spitalaufenthalt ist immer ein negatives Qualitätssignal: (i) Bei diagnostischen Leistungen kann darauf geschlossen werden, dass nicht die richtige Diagnose gestellt wurde und (ii) bei interventionellen bzw. chirurgischen Leistungen

wurde das Behandlungsziel verfehlt, z.B. wurden nicht alle Divertikel bei einer Divertikulitis entfernt oder es ist post-stationär eine Komplikation aufgetreten, die einen erneuten Spitalaufenthalt erfordert. Ausserdem stellt der ANQ methodisch sicher, dass es sich um eine zum Zeitpunkt des Austritts nicht vorhersehbare Wiederaufnahme handelt, die durch eine bei Austritt bereits bekannte Erkrankung ausgelöst wurde.³

Hinsichtlich des Indikators «postoperative Wundinfektion», misst und veröffentlicht der ANQ für folgende Eingriffsarten bei Erwachsenen entsprechende Raten:

1. Viszeralchirurgische Eingriffe

- Magenbypassoperation
- Colonchirurgie (Dickdarmchirurgie)
- Appendektomie (Blinddarmentfernung)
- Cholezystektomie (Gallenblasenentfernung)
- Rektumoperation (Enddarmoperation)

2. Orthopädische Eingriffe

- Elektive (zeitlich planbare) Hüftgelenksprothese
- Elektive (zeitlich planbare) Kniegelenksprothese
- Laminektomie mit und ohne Implantat (Wirbelsäulenchirurgie)

3. Herz-/Gefässchirurgische Eingriffe

- Gefässchirurgische Eingriffe an Arterien der unteren Extremitäten
- Herzchirurgie

4. Gynäkologische Eingriffe

- Hysterektomie (Gebärmutterentfernung)
- Sectio caesarea (Kaiserschnitt)

Zur Einteilung der Patientengruppen, die den verschiedenen Eingriffsarten des Indikators «postoperative Wundinfektion» zugrunde liegen, nutzt der ANQ bzw. Swisnoso CHOP-Codes (siehe [Link](#)). Die Definitionen weisen einen hohen Detailgrad auf und sind gleichzeitig umfangreich. Ein Matching mit den SPLG-Definitionen muss also sehr sorgfältig erfolgen.

Es seien zwei Beispiele für potenzielle Überschneidungen genannt. Eine grosse Überschneidung ist z.B. (1) für die Endoprothetik der Hüfte und des Knies gegeben. Die SPLGs «BEW7.1 Erstprothesen Hüfte» und «BEW7.2 Erstprothesen Knie» schliessen jedoch auch das Entfernen von Endoprothesen (-komponenten) ein und berücksichtigen andererseits einzelne detaillierte CHOP-Codes nicht, die in den Indikatoren eingeschlossen werden. Weitere hohe Überschneidungen liegen (2) für die Indikatoren von Herz- und gefässchirurgischen Eingriffen und die SPLGs «HER1

³ Tatsächlich wird in Fachkreisen jedoch bereits länger diskutiert, wie gut dies methodisch mit den zur Verfügung stehenden Daten sichergestellt werden kann. Deshalb wurde die Methodik für einen Indikator zur Wiederaufnahme überarbeitet und in Kürze soll ein neuer Indikator «ungeplante Rehospitalisationen» veröffentlicht und für Benchmarking und Monitoring genutzt werden.

Einfache Herzchirurgie», «HER1.1 Herzchirurgie und Gefässeingriffe mit Herzlungenmaschine» und «HER1.1.1 Koronarchirurgie (CABG)» vor. Jedoch bestehen auch hier mehrere Unterschiede, die eine Berücksichtigung der Indikatoren für die Spitalplanung erschweren.

Ähnliches lässt sich bzgl. des Indikators «Adjustierte 2-Jahres-Revisionsrate primäre Hüfttotalprothese bei primärer Arthrose» und dem analogen Indikator für das Kniegelenk feststellen. Die zugrundeliegenden Patientengruppen hat grosse Schnittmengen zu SPLGs «BEW7.1 Erstprothesen Hüfte» und «BEW7.2 Erstprothesen Knie» bzw. «BEW7.1.1 Wechseloperationen Hüftprothesen» und «BEW7.2.1 Wechseloperationen Knieprothesen». Jedoch berücksichtigen die Indikatordefinitionen ebenfalls die Diagnose «primäre Arthrose» und es werden Fälle zwei Jahre nachverfolgt und datenseitig verknüpft. Zwar ist die Diagnose der primären Arthrose zweifellos der wichtigste Behandlungsgrund auch für die Fälle der genannten SPLGs, führt jedoch zusätzlich zu (kleineren) Unterschieden in den CHOP-Code Definitionen zu weiterer Ungenauigkeit. Abschliessend ist zu beachten, dass der Indikator nicht die Qualität der Revisionen, sondern die der Erstimplantationen beschreibt und folglich nur für diese als Anforderung zählen könnte.

(4) Umgang mit kleinen Fallzahlen

Die Limitation kleiner Fallzahlen wird durch den ANQ auf mehrere Arten adressiert. Es seien zwei Beispiele zum Umgang mit kleinen Fallzahlen genannt: (1) Es werden 95%- und teilweise auch 99.8%-Konfidenzintervalle (bzw. 5%-Signifikanzniveaus) angegeben, die bei kleinen Fallzahlen sehr gross sind und mit höherer Fallzahl stark abnehmen. Die Unsicherheit der Messergebnisse mit kleinen Fallzahlen wird also transparent dargestellt. Darüber hinaus erklärt der ANQ ausführlich, wie die Ergebnisse innerhalb der Konfidenzintervalle zu interpretieren sind und, dass nur Ergebnisse jenseits der 95%- bzw. 99.8%-Konfidenzintervalle (bzw. jenseits des 5%-Signifikanzniveaus) sichere Qualitätsaussagen bieten. Ausserdem wird (2) z.B. bei der Risikoadjustierung von Rehospitalisationen eine Mindestfallzahl von fünf Fällen für Kombinationen von Risikofaktoren vorgegeben.

Allgemein müsste jedoch geklärt werden, wie mit Bewerbungen von Spitälern umzugehen ist, die keine bzw. nur sehr geringe Fallzahlen erbringen und für die keine Indikatoren berechnet werden können.

(5) Einschätzung zur Risikoadjustierung

Der ANQ bzw. die mit der Auswertung betreuten Partner nutzen eine grosse Zahl an Patientencharakteristika, um eine umfassende Risikoadjustierung vorzunehmen. Hierzu berechnet der ANQ bzw. die mit der Auswertung betreuten Partner eigene Vorhersagemodelle, z.B. mittels hierarchischer logistischer Regressionen. Insgesamt scheint diese Art der Risikoadjustierung – gemessen an den verfügbaren Daten und den derzeit in der Literatur diskutierten Methoden – auch für stärkere Eingriffe in die Leistungserbringung wie durch die Spitalplanung als adäquat.

Abschliessende Bewertung

Prüfschritte (1) bis (3) haben gezeigt, dass zwei der fünf gemessenen Indikatoren zu allgemein sind, um spezifische Qualitätsaussagen über einzelne Leistungsgruppen zu treffen (Sturz und Dekubitus). Für die spezifischeren Indikatoren zu «postoperativen Wundinfektionen» und «Adjustierte 2-Jahres-Revisionsrate» muss im Einzelfall bewertet werden, ob ein ausreichend hoher Match für die zu bewertende Leistungsgruppe vorliegt (s. hierzu auch die Überlegungen in Kapitel 4.1 zu einer quantitativen Prüfung des medizinischen Matches). Sollte nur die postoperative

Wundinfektion herangezogen werden, muss im Einzelfall bewertet werden, ob damit die Qualität der zu bewertenden Leistungsgruppe ausreichend umfänglich beschrieben werden kann, also ob die Relevanz hoch genug ist (s. hierzu auch die Diskussion zur Mortalität im Spital in Kapitel 4.1). Der Indikator «potenziell vermeidbare Rehospitalisationen innerhalb von 30 Tagen» bzw. dessen Weiterentwicklung scheinen aussagekräftig für eine Reihe an Leistungsgruppen.

Prüfschritte (4) und (5) haben gezeigt, dass ein Grossteil der ANQ-Indikatoren den methodischen Voraussetzungen zu kleinen Fallzahlen und Risikoadjustierung für die Spitalplanung entspricht (Ausnahme: Sturz/Dekubitus). Im Übrigen schliesst die «Adjustierte 2-Jahres-Revisionsrate» einen relativ langen post-stationären Zeitraum mit ein, was die Aussagekraft und Relevanz des Indikators wesentlich erhöht.

5 Zusammenfassende Empfehlung

Die Analyse hat gezeigt, dass CH-IQIs Qualitätsaussagen über spezifische Patientengruppen treffen, während (auch relativ homogene) Leistungsgruppen grössere Patientengruppen umfassen. Ausserdem erscheint die Risikoadjustierung der CH-IQIs unzureichend für die Spitalplanung und der Umgang mit und die transparente Interpretation kleiner Fallzahlen müsste gestärkt werden. Des Weiteren lassen CH-IQIs nur eine recht eindimensionale direkte Bewertung der Qualität über die Auswertung von Mortalität im Spital zu. CH-IQIs mögen ein wichtiges Instrument in der Qualitätsmessung, für Benchmarking und für die Qualitätstransparenz sein, sollten aber aufgrund obiger Argumente in ihrer bestehenden Form keine Anwendung in der Spitalplanung als leistungsgruppenspezifische Anforderungen finden.

ANQ-Indikatoren scheinen in Einzelfällen besser geeignet als leistungsgruppenspezifische Anforderungen. Jedoch müssten für leistungsspezifische Indikatoren wie «postoperative Wundinfektionen» oder für «Adjustierte 2-Jahres-Revisionsraten» Kompromisse aufgrund nicht deckungsgleicher Patientengruppen der Indikatoren und der SPLGs gemacht werden. Ergänzend bzw. alternativ könnten weniger leistungsspezifische Indikatoren wie «potenziell vermeidbare Rehospitalisationen innerhalb von 30 Tagen» bzw. deren Weiterentwicklung genutzt werden. Inwiefern die Aussagekraft dieses allgemeinen Indikators den starken Eingriff in das Leistungsangebot der Spitäler durch die Spitalplanung für eine bestimmte Leistungsgruppe rechtfertigt, bleibt im Einzelfall zu bewerten. Alternativ wäre denkbar, ANQ-Indikatoren nicht als leistungsspezifische Anforderung, sondern für qualitätsorientierte Auswahlentscheidungen bei einer «Überbewerbung» von Spitalern für einzelne Leistungsgruppen zu nutzen.