



Institut für Qualitätssicherung und
Transparenz im Gesundheitswesen

Datenauswertungen zu Mindestmengen bei TAVI

Abschlussbericht

Erstellt im Auftrag des
Gemeinsamen Bundesausschusses

Stand: 30. September 2022

Impressum

Thema:

Datenauswertungen zu Mindestmengen bei TAVI. Abschlussbericht

Ansprechpartnerinnen und Ansprechpartner:

Teresa Thomas, Dr. Maurilio Gutzeit, PD Dr. Günther Heller

Auftraggeber:

Gemeinsamer Bundesausschuss

Datum des Auftrags:

19. Mai 2022

Datum der Abgabe:

30. September 2022

Herausgeber:

IQTIG – Institut für Qualitätssicherung
und Transparenz im Gesundheitswesen

Katharina-Heinroth-Ufer 1
10787 Berlin

Telefon: (030) 58 58 26-0
Telefax: (030) 58 58 26-999

info@iqtig.org

<https://www.iqtig.org>

Kurzzusammenfassung

Hintergrund

Der Gemeinsame Bundesausschuss legt für verschiedene ausgewählte planbare Leistungen Mindestmengen fest. Voraussetzung hierfür ist ein nachgewiesener Volume-Outcome-Zusammenhang, d. h., die Qualität der Behandlungsergebnisse ist von der erbrachten Menge im entsprechenden Leistungsbereich abhängig. Für den Bereich der kathetergestützten Aortenklappenimplantationen (Transcatheter aortic-valve implantation – TAVI) wurde im Jahr 2020 ein entsprechendes Beratungsverfahren zur Festlegung der Mindestmenge eingeleitet. Eine Literaturrecherche zum Zusammenhang zwischen Leistungsmenge und Qualität des Behandlungsergebnisses bei der Durchführung von TAVIs wurde bereits vom Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen durchgeführt. Dabei konnte ein Volume-Outcome-Zusammenhang für das Versterben im Krankenhaus in Bezug auf Fälle pro Krankenhausstandort, aber auch ein Zusammenhang zu den jährlichen Operationen pro Arzt und Ärztin belegt werden.

Auftrag und Auftragsverständnis

Mit dem „Beschluss des Gemeinsamen Bundesausschusses über eine Beauftragung des IQTIG mit der Auswertung von esQS-Daten hinsichtlich Volume-Outcome-Beziehungen bei TAVI“ vom 19. Mai 2022 wurde das IQTIG dazu veranlasst zu überprüfen, ob ein Zusammenhang zwischen der Zahl der kathetergestützten Aortenklappenimplantationen und Outcome-Parametern (Ergebnisindikatoren) der externen stationären Qualitätssicherung (esQS) besteht. Verwendet werden Daten der Jahre 2018 und 2019 aus dem QS-Verfahren *Aortenklappenchirurgie, isoliert (HCH-AORT)*. Sofern ein Volume-Outcome-Zusammenhang erkennbar ist, soll in einem zweiten Schritt geprüft werden, ob sich mögliche Empfehlungen für Mindestmengen ableiten lassen.

Methodisches Vorgehen

Das IQTIG nutzt zur Analyse der Volume-Outcome-Zusammenhänge jeweils ein „logistisches additives gemischtes Regressionsmodell“ auf Ebene der einzelnen Fälle. Dies erlaubt u. a. die gleichzeitige Berücksichtigung unterschiedlicher Einflussfaktoren auf die Wahrscheinlichkeit für das Eintreten des interessierenden Outcomes: ggf. patientenseitige Risikofaktoren, die Fallzahl sowie weitere Standorteinflüsse, die von der Fallzahl unabhängig sind.

Der für diesen Bericht besonders relevante Einfluss der Fallzahl wird dabei glatt und flexibel modelliert. Die Analyse ist insofern ergebnisoffen und realistisch.

Die Vorgehensweise ermöglicht es, alle genannten Einflüsse getrennt voneinander darzustellen und sinnvoll in ihrer Stärke zu vergleichen.

Basierend auf einem geschätzten Volume-Outcome-Zusammenhang werden weiterhin in retrospektiver Weise Effektmaße für hypothetische Mindestmengen bestimmt, darunter die sogenannte Übersterblichkeit und „number needed to treat“.

Ergebnisse und Empfehlungen

Für das Outcome Sterblichkeit kann ein statistisch signifikanter Volume-Outcome-Zusammenhang nachgewiesen werden. Dieser ist monoton fallend, d. h., höhere Fallzahlen sind mit besseren Behandlungsergebnissen assoziiert. Die Simulation von Mindestmengen veranschaulicht zugleich, dass der Volume-Outcome-Zusammenhang nicht sehr stark ausgeprägt ist: Eine relevante Verbesserung der Sterblichkeit in der betrachteten Grundgesamtheit bedarf tendenziell einer hohen Mindestmenge. Dieses Ergebnis muss allerdings vor dem Hintergrund betrachtet werden, dass die Auswertungen rein retrospektiv sind (ohne Folgenabschätzungen) und dass der geschätzte Effekt hoher Mindestmengen mit hoher statistischer Unsicherheit behaftet ist.

Bei den anderen betrachteten Outcomes, intraprozedurale Komplikationen sowie Gefäßkomplikationen, zeigt sich ein qualitativ anderes Ergebnis. Die Analyse ergibt in beiden Fällen keinen statistisch signifikanten oder relevanten Volume-Outcome-Zusammenhang.

Fazit und Ausblick

Für die risikoadjustierte Sterblichkeit besteht ein Volume-Outcome-Zusammenhang, dabei ist kein Schwellenwert ableitbar. Für die intraprozedurale Komplikation sowie die Gefäßkomplikation lassen sich mit den vorliegenden Daten keine Volume-Outcome-Zusammenhänge feststellen.

Inhaltsverzeichnis

Kurzzusammenfassung.....	3
Tabellenverzeichnis.....	6
Abbildungsverzeichnis.....	7
Abkürzungsverzeichnis.....	8
1 Einleitung.....	10
1.1 Hintergrund.....	10
1.2 Auftrag und Auftragsverständnis.....	11
1.3 Beschreibung der Datenbasis.....	12
2 Methodisches Vorgehen.....	17
2.1 Schätzung eines Volume-Outcome-Zusammenhangs.....	17
2.2 Schätzung des Einflusses einer hypothetischen Mindestmenge auf die Anzahl an interessierenden Ereignissen (Übersterblichkeit).....	19
2.3 Mathematische Annahmen und Festlegungen, Software.....	22
3 Ergebnisse.....	24
3.1 Geschätzte Volume-Outcome-Zusammenhänge.....	24
3.1.1 Sterblichkeit.....	24
3.1.2 Intraprozedurale Komplikation.....	29
3.1.3 Gefäßkomplikation.....	30
3.2 Geschätzte Übersterblichkeit basierend auf hypothetischen Mindestmengen.....	33
3.3 Ergänzungen.....	38
4 Diskussion und Empfehlungen.....	41
5 Fazit und Ausblick.....	43
Literatur.....	44

Tabellenverzeichnis

Tabelle 1: Einfache Häufigkeit der interessierenden Outcomes.....	14
Tabelle 2: Überschneidung der Outcomes Sterblichkeit und Intraprozedurale Komplikation...	15
Tabelle 3: Überschneidung der Outcomes Sterblichkeit und Gefäßkomplikation	15
Tabelle 4: Überschneidung der Outcomes Intraprozedurale Komplikation und Gefäßkomplikation.....	16
Tabelle 5: Zusammenhang zwischen den drei Outcomes in Form von Odds Ratios	16
Tabelle 6: Kategorielle Risikofaktoren zu Sterblichkeit.....	24
Tabelle 7: Kennzahlen der LE-Einflüsse und Modellgüte bei Sterblichkeit	27
Tabelle 8: Geschätzte Odds Ratios bezüglich Variation der Fallzahl.....	28
Tabelle 9: Kennzahlen der LE-Einflüsse und Modellgüte bei intraprozeduraler Komplikation ..	29
Tabelle 10: Kategorielle Risikofaktoren zu Gefäßkomplikation	30
Tabelle 11: Kennzahlen der LE-Einflüsse und Modellgüte bei Gefäßkomplikation	32
Tabelle 12: Übersterblichkeit nach Szenario 1 in Zahlen mit 95%-Konfidenzintervallen	36
Tabelle 13: Übersterblichkeit nach Szenario 2 in Zahlen mit 95%-Konfidenzintervallen	37

Abbildungsverzeichnis

Abbildung 1: Verteilung der mittleren jährlichen Fallzahl nach Standorten	13
Abbildung 2: Lorenzkurve zur Verteilung der Fallzahl	14
Abbildung 3: Effekt des Alters in der Risikoadjustierung zu Sterblichkeit (logistisch)	26
Abbildung 4: Effekt des Alters in der Risikoadjustierung zu Sterblichkeit	26
Abbildung 5: Geschätzter Fallzahleffekt zu Sterblichkeit.....	27
Abbildung 6: Geschätzter Fallzahleffekt zu intraprozeduraler Komplikation	30
Abbildung 7: Effekt des Alters in der Risikoadjustierung zu Gefäßkomplikation (logistisch)	31
Abbildung 8: Effekt des Alters in der Risikoadjustierung zu Gefäßkomplikation	32
Abbildung 9: Geschätzter Fallzahleffekt zu Gefäßkomplikation	33
Abbildung 10: Von einer hypothetischen Mindestmenge betroffene Fälle 2019	34
Abbildung 11: Absolute und mittlere Übersterblichkeit zu hypothetischen Mindestmengen...	34
Abbildung 12: Mittlere Übersterblichkeit unter den umverteilten Fällen zu hypothetischen Mindestmengen.....	35
Abbildung 13: Absolute und mittlere Übersterblichkeit zu hypothetischen Mindestmengen, gesamter Bereich	39
Abbildung 14: Mittlere Übersterblichkeit unter den umverteilten Fällen zu hypothetischen Mindestmengen, gesamter Bereich.....	40

Abkürzungsverzeichnis

Abkürzung/Symbol	Bedeutung
ATT	average treatment effect on the treated
ATE	average treatment effect
AUC	Area under the curve
DeQS-RL	Richtlinie zur datengestützten einrichtungsübergreifenden Qualitätssicherung
esQS	externe stationäre Qualitätssicherung
G-BA	Gemeinsamer Bundesausschuss
IQTIG	Institut für Qualitätssicherung und Transparenz im Gesundheitswesen
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
LE	Leistungserbringer (hier gleichbedeutend mit Standort)
LVEF	Linksventrikuläre Ejektionsfraktion
MHI-RL	Richtlinie zu minimalinvasiven Herzklappeninterventionen. Richtlinie über Maßnahmen zur Qualitätssicherung bei der Durchführung von minimalinvasiven Herzklappeninterventionen
Mm-R	Mindestmengenregelung
MOR	Median-Odds-Ratio
NNT	number needed to treat
QIDB	Qualitätsindikatorendatenbank
QS	Qualitätssicherung
QSKH-RL	Richtlinie über Maßnahmen der Qualitätssicherung in Krankenhäusern
TAVI	Transcatheter aortic-valve implantation (kathetergestützte Aortenklappenimplantation)

Mathematische Notation

logit	Logit-Transformation. $\text{logit}(x) = \log(x/(1-x))$ für $0 < x < 1$
y_{ij}^k	Binäres Outcome für Fall j aus Standort i in Jahr k
π_{ij}^k	Wahrscheinlichkeit für das Auftreten des interessierenden Outcomes bei Fall j aus Standort i in Jahr k
β_0^k	Intercept für das Jahr k

\mathbf{x}_{ij}^k	Vektor der kategoriellen patientenseitigen Risikofaktoren (in den Daten gegeben)
$\boldsymbol{\beta}$	Koeffizientenvektor mit den Effekten der patientenseitigen Risikofaktoren
f_{Alter}	Alterseffekt (glatte Funktion)
a_{ij}^k	Patientenalter (in den Daten gegeben)
f_{FZ}	Fallzahleffekt (glatte Funktion)
n_i	Durchschnittliche jährliche Fallzahl des Standorts
u_i	Nicht fallzahlassoziiertes Standorteffekt (random intercept)
M	(Hypothetische) Mindestmenge
$r_{\text{abs}}(M)$	Absolute Übersterblichkeit gegebener Mindestmenge M
$r_{\text{mittel}}(M)$	Mittlere Übersterblichkeit gegebener Mindestmenge M
$r_{\text{umv}}(M)$	Mittlere Übersterblichkeit unter den umverteilten Fällen gegebener Mindestmenge M
$\mathcal{N}(0, \tau^2)$	Normalverteilung mit Erwartungswert 0 und Varianz $\tau^2 > 0$
$\Phi^{-1}(3/4)$	75. Perzentil der Standardnormalverteilung

1 Einleitung

1.1 Hintergrund

In Deutschland sind die geltenden Mindestmengen durch den Gemeinsamen Bundesausschuss (G-BA) in den Mindestmengenregelungen (Mm-R)¹ festgeschrieben. Für ausgewählte planbare stationäre Leistungen, z. B. die Verfahren Komplexe Eingriffe am Organsystem Pankreas für Erwachsene oder Chirurgische Behandlung des Brustkrebses (Mamma-Ca-Chirurgie), bei denen ein Volume-Outcome-Zusammenhang nachgewiesen wurde, wird in der Mm-R die Höhe der jeweiligen jährlichen Mindestmenge je Ärztin oder Arzt und/oder Krankenhausstandort geregelt. Im Jahr 2020 hat der G-BA ein Beratungsverfahren zur Festlegung der Mindestmengen für die Durchführung von kathetergestützten Aortenklappenimplantationen (Transcatheter aortic-valve implantation – TAVI) eingeleitet (G-BA 2020a). Dabei wurde als Arbeitshypothese auf Grundlage der vorhandenen Literatur und der Empfehlungen von Fachgesellschaften für weitere Beratungen von der Vorsitzenden eine Kombination aus einer Mindestmenge von 150 Eingriffen pro Krankenhausstandort und von 65 Eingriffen pro Operateurin oder Operateur vorgeschlagen (G-BA 2020b).

Daraufhin hat das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) im Auftrag des G-BA eine Literaturrecherche zum Zusammenhang zwischen Leistungsmenge und Qualität des Behandlungsergebnisses bei der Durchführung von TAVIs durchgeführt. Als Ergebnis konnte festgehalten werden, dass für das Versterben im Krankenhaus ein Volume-Outcome-Zusammenhang existiert in Bezug auf Fälle pro Krankenhausstandort, außerdem auch ein Zusammenhang zu den jährlichen Operationen pro Arzt und Ärztin. In Bezug auf Morbidität wurde bei Blutungen und Beatmung > 48 Stunden ein Volume-Outcome-Zusammenhang in Bezug auf Fälle pro Krankenhausstandort nachgewiesen (IQWiG 2021).

Darüber hinaus werden in der Richtlinie zu minimalinvasiven Herzklappeninterventionen (MHI-RL) die Struktur- und Prozessqualität geregelt, die zu einer Optimierung der Ergebnisqualität der Behandlung beitragen sollen. Die entsprechenden Krankenhäuser müssen hierfür jährlich „Checklisten für das Nachweisverfahren zur Erfüllung von Qualitätsanforderungen an die Durchführung von minimalinvasiven Herzklappeninterventionen gemäß § 136 Absatz 1 Satz 1 Nummer 2 für nach § 108 SGB V zugelassene Krankenhäuser“ ausfüllen (MHI-RL)². Das Institut für Qualitätssicherung und Transparenz im Gesundheitswesen (IQTIG) hat die MHI-RL umfassend

¹ Regelungen des Gemeinsamen Bundesausschusses gemäß § 136b Absatz 1 Satz 1 Nummer 2 SGB V für nach § 108 SGB V zugelassene Krankenhäuser (Mindestmengenregelungen, Mm-R). In der 1. Neufassung vom 21. März 2006, zuletzt geändert am 16. Juni 2022, in Kraft getreten am 15. Juli 2022. URL: <https://www.g-ba.de/richtlinien/5/> (abgerufen am 05.09.2022).

² Richtlinie des Gemeinsamen Bundesausschusses über Maßnahmen zur Qualitätssicherung bei der Durchführung von minimalinvasiven Herzklappeninterventionen gemäß § 136 Absatz 1 Satz 1 Nummer 2 für nach § 108 SGB V zugelassene Krankenhäuser (Richtlinie zu minimalinvasiven Herzklappeninterventionen/MHI-RL). In der Fassung vom 22. Januar 2015, zuletzt geändert am 21. April 2022, in Kraft getreten am 1. April 2022. URL: <https://www.g-ba.de/richtlinien/84/> (abgerufen am 05.09.2022).

evaluiert und die Ergebnisse in einem Abschlussbericht veröffentlicht (IQTIG 2022a). Im Zusammenhang mit potenziellen Mindestmengen ist dabei zu erwähnen, dass sich seit Einführung der MHI-RL einerseits die Anzahl der Leistungserbringer reduziert und andererseits die durchschnittliche Fallzahl pro Leistungserbringer erhöht hat.

1.2 Auftrag und Auftragsverständnis

Das IQTIG wurde am 19. Mai 2022 beauftragt, Ergebnisdaten der externen stationären Qualitätssicherung (esQS) mit Bezug zum Leistungsbereich Kathetergestützte Aortenklappenimplantationen (TAVI) auszuwerten (G-BA 2022). Hierzu werden die Daten des QS-Verfahrens *Aortenklappenchirurgie, isoliert (HCH-AORT)*³ gemäß der Richtlinie über Maßnahmen der Qualitätssicherung in Krankenhäusern (QSKH-RL) verwendet. Hauptgegenstand der Auswertungen ist die Analyse, ob zwischen der Zahl der kathetergestützten Aortenklappenimplantationen und Outcome-Parametern der esQS (Ergebnisindikatoren) ein Zusammenhang besteht. Zusätzlich soll geprüft werden, ob bei einem bestehenden Volume-Outcome-Zusammenhang Empfehlungen für mögliche Mindestmengen abgeleitet werden können.

Für die Analysen werden die Daten der Erfassungsjahre 2018 und 2019 verwendet und die folgenden Ergebnisindikatoren untersucht:

- QI-ID 51916: „Intraprozedurale Komplikationen“
- QI-ID 52007: „Gefäßkomplikationen“
- QI-ID 12168: „Verhältnis der beobachteten zur erwarteten Rate (O/E) an Todesfällen“

Das IQTIG hat bereits zwei Berichte bez. der Untersuchungen von Volume-Outcome-Zusammenhängen verfasst („Volumen-Outcome-Beziehungen bei Revisionseingriffen in der Knieendoprothetik“ (IQTIG 2019) und „Datenauswertung zu Mindestmengen in der Versorgung von Frühgeborenen mit einem Aufnahmegegewicht unter 1.250 g“ (IQTIG 2020b) und eine entsprechende Methodik entwickelt, die im vorliegenden Bericht ebenfalls Anwendung findet.

Der Bezugspunkt Operateurin oder Operateur wurde in verschiedenen Studien aufgegriffen und es wurde in Bezug auf Versterben im Krankenhaus ein Volume-Outcome-Zusammenhang belegt. In den zur Verfügung stehenden Daten kann keine Zuordnung zur Operateurin oder zum Operateur vorgenommen werden, sodass diese Untersuchung nicht durch den G-BA beauftragt wurde. Wie in Abschnitt 1.1 erwähnt, wurde eine umfassende Literaturrecherche zu dieser Thematik bereits vom IQWiG durchgeführt. Somit wurde vom IQTIG keine erneute Recherche vorgenommen.

Am 4. Mai 2022 wurde das IQTIG zusätzlich beauftragt, Datenanalysen für den Leistungsbereich Kathetergestützte Aortenklappenimplantationen (Transcatheter aortic-valve implantation – TAVI) zur Folgenabschätzung im Rahmen der Beratungen zu Mindestmengen durchzuführen.⁴

³ Das QS-Verfahren existiert seit 2008. Informationen zum Verfahren unter der QSKH-RL sind hier einzusehen: <https://igtig.org/qs-verfahren/hch-aort/>. Mit Umstellung auf die Richtlinie zur datengestützten einrichtungsübergreifenden Qualitätssicherung (DeQS-RL) sind die Informationen hier einsehbar: <https://igtig.org/qs-verfahren/qs-kchk/>.

⁴ Beschluss des Gemeinsamen Bundesausschusses über eine Beauftragung des IQTIG mit der Durchführung von Datenanalysen für den Leistungsbereich der kathetergestützten Aortenklappenimplantationen

Die Abgabe dieses Berichts erfolgte beim G-BA am 31. Juli 2022 (IQTIG 2022b). Die Ergebnisse der Folgenabschätzungen und Veränderungen der Fahrzeiten und Wegstrecken werden im vorliegenden Bericht nicht berücksichtigt. Ebenso werden keine regionalen Analysen vorgenommen.

1.3 Beschreibung der Datenbasis

Die im Rahmen des QS-Verfahrens *HCH-AORT (isolierte Aortenklappenchirurgie)* erhobenen QS-Daten der Erfassungsjahre 2018 und 2019 bilden die Datengrundlage für diesen Bericht. Die Daten liegen in einer einheitlichen Spezifikation vor, sodass insbesondere alle Standorte über die beiden Erfassungsjahre hinweg identifizierbar sind.

Die drei zu untersuchenden Outcomes sind:

- Sterblichkeit (i. S. v. QI-ID 12168),
- Intraprozedurale Komplikation (i. S. v. QI-ID 51916),
- Gefäßkomplikation (i. S. v. QI-ID 52007).

Diese werden im Folgenden auch synonym als interessierende Ereignisse oder interessierende Outcomes bezeichnet.

Im Rahmen der Qualitätssicherung wurden Eingriffe mit folgenden Eigenschaften als angemessene Grundgesamtheit für die zugehörigen Qualitätsindikatoren und dementsprechend auch für diesen Bericht festgelegt:

- kathetergestützte Aortenklappeneingriffe,
- keine Koronarchirurgie,
- jeweils erster Eingriff des Aufenthalts.

Der Datensatz 2018 enthält 20.974 und der Datensatz 2019 enthält 24.305 solcher Eingriffe (insgesamt 45.279), die an 88 unterschiedlichen Standorten durchgeführt wurden.

(Transcatheter aortic-valve implantation – TAVI) zur Folgenabschätzung im Rahmen von Beratungen zu Mindestmengen auf Grundlage von § 136b Abs. 1 Satz 1 Nr. 2, Abs. 3 SGB V. Vom 4. Mai 2022.

Verteilung der Fallzahlen

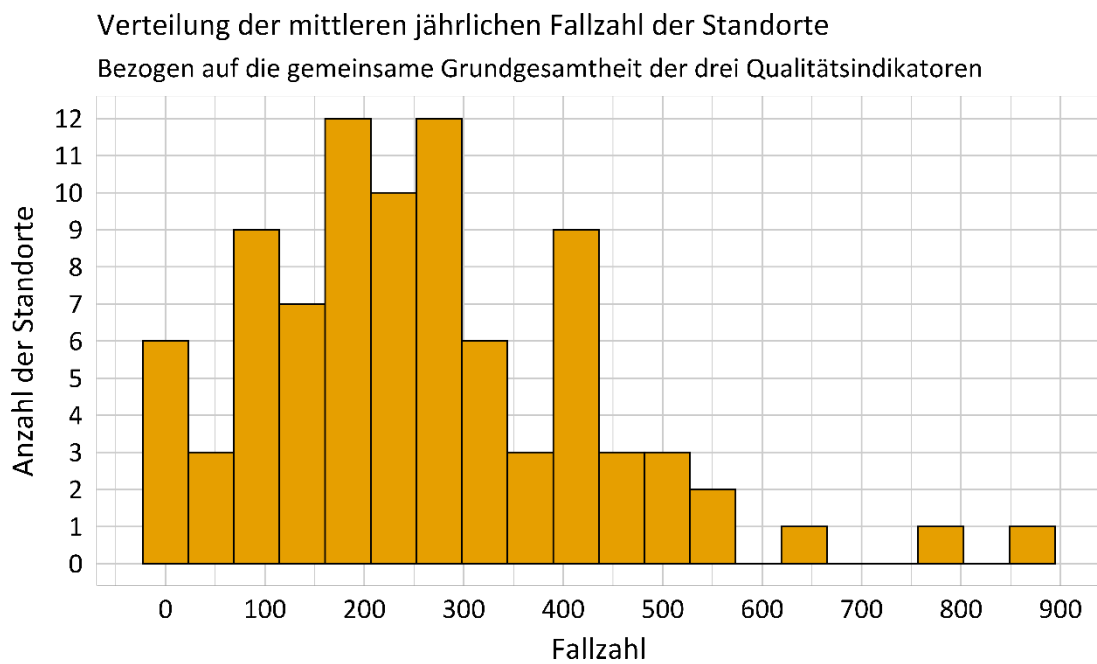


Abbildung 1: Verteilung der mittleren jährlichen Fallzahl nach Standorten

Aus Abbildung 1 ist eine Konzentration der mittleren jährlichen Fallzahlen im Bereich von etwa 100 bis 300 ersichtlich, wobei 3 Standorte mit besonders vielen Fällen herausstechen (zugehörige Werte: 639,5 sowie 761,5 und 872,5). Einen Eindruck von der Gleichmäßigkeit bzw. Ungleichmäßigkeit der Fallzahlverteilung kann auch eine sogenannte Lorenzkurve wie in Abbildung 2 geben:

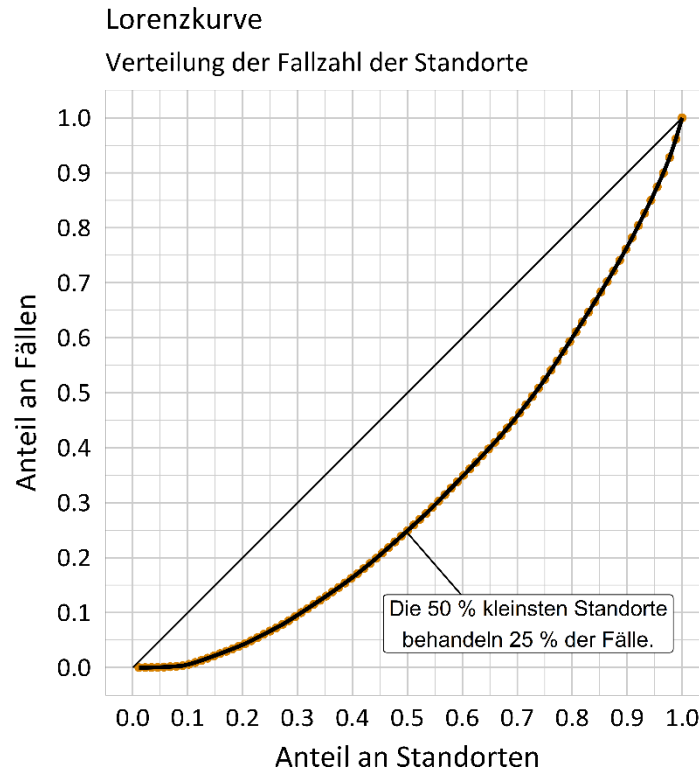


Abbildung 2: Lorenzkurve zur Verteilung der Fallzahl

Hierzu werden die Standorte der Größe nach sortiert (im Sinne der Fallzahl) und es wird jeweils ermittelt, wie viele Fälle von den x % kleinsten Standorten behandelt werden. Im theoretischen Sonderfall, dass alle Standorte gleich viele Fälle behandeln, entspräche die Kurve genau der Winkelhalbierenden.

Häufigkeit der Outcomes

Unter den 45.279 betrachteten Eingriffen hatten die drei interessierenden Outcomes schließlich die in Tabelle 1 dargestellten Häufigkeiten:

Tabelle 1: Einfache Häufigkeit der interessierenden Outcomes

Outcome	Absolute Häufigkeit	Relative Häufigkeit (95%-KI)
Sterblichkeit	1.184	2,61 % (2,47 %–2,77 %)
Intraprozedurale Komplikation	995	2,20 % (2,07 %–2,34 %)
Gefäßkomplikation	743	1,64 % (1,53 %–1,76 %)

Bei dem Qualitätsindikator zur Sterblichkeit handelt es sich im Gegensatz zu den anderen beiden um einen risikoadjustierten Qualitätsindikator. Dies wird in den folgenden Abschnitten noch thematisiert.

Bei drei zu untersuchenden Outcomes stellt sich auch die Frage, wie stark sie sich in ihrem Auftreten überschneiden bzw. voneinander abhängen, insbesondere da die drei Volume-Outcome-Zusammenhänge unabhängig voneinander untersucht werden.

Zunächst wird in Tabelle 2 ein Überblick zur Überschneidung zwischen Sterblichkeit und Intraprozeduraler Komplikation in Tabelle 2 Form einer Kreuztabelle dargestellt:

Tabelle 2: Überschneidung der Outcomes Sterblichkeit und Intraprozedurale Komplikation

Tod	Intraprozedurale Komplikation		Gesamt
	Nein	Ja	
Nein	43.346 (95,73 %)	749 (1,65 %)	44.095 (97,39 %)
Ja	938 (2,07 %)	246 (0,54 %)	1.184 (2,61 %)
Gesamt	44.284 (97,80 %)	995 (2,20 %)	45.279 (100,00 %)

Hierbei zeigt sich in absoluten Zahlen keine sehr starke Überschneidung, da unter den 1.184 verstorbenen Patientinnen und Patienten lediglich 246 zusätzlich die Komplikation aufwiesen (20,77 %) bzw. unter den 995 Patientinnen und Patienten mit Komplikation 246 verstorben sind (24,72 %). Allerdings lässt sich beim Vergleich verschiedener Gruppen ein sehr starker Zusammenhang hinsichtlich der relativen Häufigkeiten feststellen: Während die Sterblichkeit unter den Patientinnen und Patienten mit intraprozeduraler Komplikation wie bereits erwähnt bei 24,72 % liegt, beträgt sie unter den Patientinnen und Patienten ohne intraprozedurale Komplikation nur 2,12 % (938/44.284). Als Odds Ratio formuliert besteht zwischen diesen Häufigkeiten das Verhältnis 15,2.

Die Kreuztabellen für die anderen beiden paarweisen Kombinationen (Tabelle 3 und Tabelle 4) lauten wie folgt:

Tabelle 3: Überschneidung der Outcomes Sterblichkeit und Gefäßkomplikation

Tod	Gefäßkomplikation		Gesamt
	Nein	Ja	
Nein	43.469 (96,00 %)	626 (1,38 %)	44.095 (97,39 %)
Ja	1.067 (2,36 %)	117 (0,26 %)	1.184 (2,61 %)
Gesamt	44.536 (98,36 %)	743 (1,64 %)	45.279 (100,00 %)

Tabelle 4: Überschneidung der Outcomes Intraprozedurale Komplikation und Gefäßkomplikation

Intraprozedurale Komplikation	Gefäßkomplikation		Gesamt
	Nein	Ja	
Nein	43.590 (96,27 %)	694 (1,53 %)	44.284 (97,80 %)
Ja	946 (2,09 %)	49 (0,11 %)	995 (2,20 %)
Gesamt	44.536 (98,36 %)	743 (1,64 %)	45.279 (100,00 %)

Bei diesen Paarungen zeigt sich grundsätzlich ein ähnlicher Zusammenhang, wenn auch etwas weniger ausgeprägt. Tabelle 5 zeigt die zugehörigen Odds Ratios im Überblick, wobei zur Ergänzung die beiden Komplikationsoutcomes auch einmal zusammengefasst wurden:

Tabelle 5: Zusammenhang zwischen den drei Outcomes in Form von Odds Ratios

Verglichene Risiken	Odds Ratio	95%-KI
Sterblichkeit bei vorhandener vs. nicht vorhandener intraprozeduraler Komplikation	15,18	[12,96–17,77]
Sterblichkeit bei vorhandener vs. nicht vorhandener Gefäßkomplikation	7,61	[6,19–9,36]
Intraprozedurale Komplikationen bei vorhandener vs. nicht vorhandener Gefäßkomplikation	3,25	[2,42–4,38]
Sterblichkeit bei vorhandener vs. nicht vorhandener Gefäß- oder intraprozeduraler Komplikation	12,70	[11,07–14,57]

2 Methodisches Vorgehen

Die statistischen Auswertungen in diesem Bericht werden grundsätzlich auf Basis der gleichen Prinzipien durchgeführt, die bereits in IQTIG (2019) und in IQTIG (2020b) beschrieben sind. Im Folgenden werden noch einmal die wesentlichen Ziele und die methodischen Entscheidungen des IQTIG in kompakter Form erläutert sowie in den Kontext der vorliegenden Anwendung gesetzt. Eine erneute ausführliche Abgrenzung zu alternativen Methoden oder eine tiefere Diskussion mathematischer Aspekte ist aufgrund der beiden oben zitierten Berichte allerdings verzichtbar.

Während sich die ersten beiden Abschnitte 2.1 und 2.2 auf die grundsätzlichen Ideen konzentrieren, greift der letzte Abschnitt 2.3 kurz formale und technische Aspekte auf.

2.1 Schätzung eines Volume-Outcome-Zusammenhangs

Dieser Abschnitt widmet sich der Frage, wie untersucht werden kann, ob aus statistischer Sicht ein Zusammenhang zwischen der Anzahl der an einem Standort behandelten Fälle (im Folgenden Fallzahl) und dem Auftreten des jeweiligen interessierenden Ereignisses besteht sowie welche Gestalt dieser Zusammenhang ggf. hat.

Begriff der Fallzahl und Operationalisierung in der Analyse

Die Fallzahl bezieht sich hier wie üblich auf den Zeitraum eines Jahres und auf die in Abschnitt 1.3 beschriebene Grundgesamtheit pro Standort.

Während die Fallzahl keinen direkten Einfluss auf die Behandlungsergebnisse haben kann, hängt sie eng mit bestimmten dafür relevanten Standorteigenschaften zusammen, insbesondere mit der Erfahrung des Personals und der technischen Ausstattung am Standort (Wetzel 2006). Zugleich liegen derartige strukturelle Informationen über die Standorte nicht in den QS-Daten vor. Insofern ist die Fallzahl ein Surrogatparameter für diese fallzahlassoziierten Standorteigenschaften und ein etwaig statistisch vorhandener Zusammenhang der Fallzahl mit den Behandlungsergebnissen bezieht sich insbesondere auf diese Eigenschaften.

Da in der Analyse die Daten aus nur zwei Erfassungsjahren verwendet werden und solche Standorteigenschaften innerhalb dieses Zeitraums relativ stabil sind, bietet sich zunächst pro Standort die Verwendung der durchschnittlichen Fallzahl über diese beiden Jahre an. Zur Gewährleistung, dass diese Wahl für die wesentlichen Schlussfolgerungen unerheblich ist, werden ergänzend alternative Varianten geprüft, siehe Abschnitt 3.3.

Statistisches Inferenzmodell

Die so definierte Fallzahl fließt nun in ein statistisches Inferenzmodell mit ein, auf Basis dessen alle interessierenden Effekte geschätzt werden. Wesentliche in der Modellbildung berücksichtigte Prinzipien und Annahmen sind:

- Da es sich bei den betrachteten interessierenden Ereignissen um binäre Outcomes handelt (z. B. verstorben/nicht verstorben), bietet sich ein logistisches Regressionsmodell an, das die Wahrscheinlichkeit für das Auftreten des interessierenden Ereignisses beschreibt.
- Das höchste Maß an Präzision bez. der geschätzten Effekte und der zugehörigen statistischen Unsicherheit bietet eine Modellierung auf Ebene der einzelnen Fälle.
- Die Daten weisen insofern eine Clusterstruktur auf, dass die Outcomes von Fällen, die am selben Standort behandelt wurden, vermutlich korreliert sind.
- Die Wahrscheinlichkeit für das Auftreten des interessierenden Ereignisses hängt im Allgemeinen von Einflussgrößen auf unterschiedlichen Ebenen ab.
 - Patienteneigenschaften: Hierzu können etwa das Geschlecht und das Alter zählen, siehe Abschnitt 3.1 für die konkret vorliegende Anwendung.
 - Standorteigenschaften: Hierzu zählt insbesondere, jedoch nicht ausschließlich, die Fallzahl. Um den hauptsächlich interessierenden Einfluss der Fallzahl möglichst bereinigt schätzen zu können, werden im Modell explizit fallzahlassoziierte von nicht fallzahlassoziierten Einflüssen unterschieden.
- Die konkrete Modellierung des Fallzahleffekts spielt eine besondere Rolle. Eine realistische Annahme besteht darin, dass sich der Einfluss der Fallzahl nur leicht verändert, wenn sich die Fallzahl nur leicht oder sogar nur um eins verändert; insbesondere sollte der Einfluss an keiner Stelle einen Sprung aufweisen. Dementsprechend betrachtet das IQTIG die Fallzahl als stetige Einflussgröße und modelliert den Zusammenhang zwischen Volume und Outcome als glatt (Grouven et al. 2008). Speziell wird ein flexibler, ergebnisoffener Ansatz genutzt, der grundsätzlich monotone oder auch nicht monotone Zusammenhänge finden kann. Der Abschnitt 2.3 unten bespricht die genutzte Software und die konkrete Implementation.
- Anstelle einer solchen glatten Modellierung verwenden viele Arbeiten eine diskrete Aufteilung des Fallzahlbereichs in Fallzahlgruppen (z. B. Varaganam et al. (2015)), was zwangsläufig zu einem sprunghaften und daher nicht plausiblen Einfluss führt. Zudem bergen Fallzahlgruppen die Schwierigkeit, ihre Grenzen, d. h. die potenziellen Sprungstellen, festlegen zu müssen.
- Wie bereits angesprochen gehen von den Standorten auch nicht (oder zumindest schwächere) fallzahlassoziierte Einflüsse auf das Behandlungsergebnis aus, vorstellbar sind hier etwa Qualifikation und Motivation des Personals. Auch Informationen über derartige strukturelle Standorteigenschaften liegen in den QS-Daten nicht vor. Um diesen Einflüssen bei der Analyse dennoch Rechnung zu tragen und sie vom Effekt der Fallzahl abzugrenzen bzw. damit vergleichen zu können, eignen sich sogenannte „random intercepts“. Es handelt sich dabei um Standorteffekte, die sich ähnlich wie patientenseitige Einflüsse und der Fallzahleffekt auf die Wahrscheinlichkeit für das interessierende Ereignis auswirken, allerdings weiteren technischen Annahmen unterliegen (siehe Abschnitt 2.3). Ein Vorteil der Verwendung von random intercepts besteht weiterhin darin, dass sie eine Berücksichtigung der oben genannten Clusterstruktur sicherstellen.

Die Modellformel auf Ebene der einzelnen Fälle nimmt schließlich die Form

$$\text{logit}(\pi_{ij}^k) = \beta_0^k + (\mathbf{x}_{ij}^k)^\top \boldsymbol{\beta} + f_{\text{Alter}}(a_{ij}^k) + f_{\text{FZ}}(n_i) + u_i \quad (1)$$

an.⁵

Da lediglich ein zweijähriger Zeitraum betrachtet wird, liegt den Modellkomponenten die vereinfachende Annahme zugrunde, dass sich die patientenseitigen Effekte und die beiden standortseitigen Effekte zwischen den beiden Jahren nicht relevant unterscheiden. Daher fehlt ihnen der Index k für das Jahr. Globale Veränderungen, etwa allgemeiner medizinischer Fortschritt, werden ggf. durch den jahresspezifischen Intercept β_0^k erfasst. Da die Annahme eines konstanten Fallzahleffekts von größerer Bedeutung ist, wurde die diesbezügliche Robustheit der Ergebnisse geprüft, siehe Abschnitt 3.3.

Wichtige Ergebnisse der Analyse im Überblick

- Schätzer (einschließlich Unsicherheit) für alle unbekanntes Modellkomponenten, d. h. alle unbekanntes zu ermittelnden Bestandteile von Formel (1).
- Dies schließt insbesondere die Schätzung des Fallzahleffekts f_{FZ} mit einem Konfidenzband ein sowie einen statistischen Test zur Frage, ob die Fallzahl einen Einfluss auf die Behandlungsergebnisse hat.
- Die random intercepts geben Aufschluss darüber, wie stark der Einfluss der Standorte auf die Behandlungsergebnisse abgesehen von der Fallzahl ist. Eine zugehörige Kennzahl, die sogenannte Median-Odds-Ratio (MOR), erlaubt den direkten Vergleich mit dem Einfluss der Fallzahl, siehe auch hierzu Abschnitt 2.3.
- Modellgütekriterien: Als gängige Konzepte zur Einschätzung von Modellen werden in diesem Bericht die AUC (Area under the curve) und die Kalibrierung betrachtet, jeweils „out of sample“ basierend auf „cross-validation“, siehe etwa Agresti (2013). Die AUC ist ein Maß dafür, ob Prognosen basierend auf Modell (1) zur Vorhersage des interessierenden Outcomes geeignet sind. Die Zahl liegt zwischen 0,5 und 1 und ist besser, je höher sie ist. Bei der Kalibrierung wird geprüft, wie gut die basierend auf Modell (1) prognostizierten Wahrscheinlichkeiten der empirischen Häufigkeit entsprechen. Bei geringer Modellgüte könnte zwischen Prognose und Empirie eine ungewollte Verschiebung oder Skalierung bestehen. Beides kann statistisch getestet werden und hohe p-Werte sprechen für ein angemessenes Modell.

2.2 Schätzung des Einflusses einer hypothetischen Mindestmenge auf die Anzahl an interessierenden Ereignissen (Übersterblichkeit)

Anhand der in Abschnitt 2.1 beschriebenen Methodik lässt sich ein geschätzter Fallzahleffekt ermitteln, der den alleinigen Einfluss der Fallzahl auf das Behandlungsergebnis beschreibt, bereinigt um patientenseitige oder andere standortseitige Einflüsse. Sollte sich dabei zum Beispiel zeigen, dass höhere Fallzahlen mit signifikant besseren Behandlungsergebnissen assoziiert sind,

⁵ Die Bestandteile dieser und folgender Formeln sind im Abkürzungsverzeichnis erläutert.

ergibt die Erwägung von Mindestmengen grundsätzlich Sinn. Zur konkreten Einschätzung, inwiefern eine Mindestmenge die Behandlungsqualität insgesamt potenziell verbessern kann, eignen sich wiederum absolute Effektmaße wie die „Übersterblichkeit/Anzahl der potenziell vermeidbaren Todesfälle (excess mortality)“ sowie die „number needed to treat (NNT)“ besser (siehe Rothman et al. (2008), Heller (2018), Heller et al. (2022)).

Diese Kenngrößen lassen sich für beliebige hypothetische Mindestmengen anhand der gemäß Abschnitt 2.1 ermittelten Effekte ableiten, d. h. innerhalb des vorliegenden Datensatzes berechnen. Insofern handelt es sich bei den hier vorgestellten absoluten Effektmaßen weniger um Vorhersagen der Folgen von künftigen Mindestmengen, sondern eher um alternative Möglichkeiten, den gefundenen Volume-Outcome-Zusammenhang darzustellen. Diese Darstellungsweise berücksichtigt dabei explizit den Case-Mix (also die Verteilung der patientenseitigen Risikofaktoren), während der Fallzahleffekt (Abbildung 5) wie bereits erwähnt bewusst um diesen bereinigt ist.

Ausdrücklich nicht berücksichtigt wird in dieser retrospektiven Simulation die Dynamik in der Entwicklung der Fallzahlen, also welche Standorte durch eine Umverteilung potenziell wie viele Fälle zusätzlich behandeln. Dies ist Gegenstand sogenannter Folgenabschätzungen.

Ansatz zur Bestimmung der Übersterblichkeit

Unter Berücksichtigung der grundsätzlich zeitversetzten Wirkung von Mindestmengen (die Fallzahl eines Standorts in Zeitraum A ist maßgeblich dafür, ob der Standort in einem späteren Zeitraum B Fälle behandelt) beruhen die Berechnungen auf folgendem Prinzip: Ob ein Standort die jeweilige Mindestmenge einhält, wird anhand der Fallzahl 2018 entschieden und die Konsequenzen der Mindestmenge werden dann auf dem Datensatz 2019 bestimmt.

Die grundsätzliche Idee besteht darin zu simulieren, dass alle Standorte in 2019 bereits eine gegebene Mindestmenge M eingehalten haben. Zentrales Hilfsmittel hierzu ist die Formel (1): Die Volume-Outcome-Analyse gemäß Abschnitt 2.1 liefert zunächst geschätzte Effekte (markiert durch Dachakzent), die eine Berechnung der Wahrscheinlichkeit des interessierenden Outcomes für jeden Fall erlauben:

$$\text{logit}(\hat{\pi}_{ij}) = \widehat{\beta}_0 + \mathbf{x}_{ij}^T \widehat{\boldsymbol{\beta}} + \hat{f}_{\text{Alter}}(a_{ij}) + \hat{f}_{\text{FZ}}(n_i) + \hat{u}_i. \quad (2)$$

Der Index für das Jahr kann an dieser Stelle entfallen, da die absoluten Effektmaße nur auf dem Datensatz 2019 bestimmt werden.

Eine Simulation der Mindestmenge entspricht bezogen auf diese Formel der Erhöhung der Fallzahl n_i auf einen gewissen Wert $n_i' \geq M$ für diejenigen Fälle, die in einem Standort mit zu geringer Fallzahl behandelt wurden. Für alle anderen Fälle ändert sich nichts (d. h. $n_i' = n_i$).

Für die Wahl der neuen Fallzahl n_i' werden in diesem Bericht zwei Szenarien betrachtet. Eine Änderung erfolgt nur bei denjenigen Standorten, die die Mindestmenge nicht eingehalten haben, die anderen bleiben unverändert.

- **Szenario 1:** $n_i' = M$, d. h., die Fallzahl steigt auf den kleinsten möglichen Wert. Im Fall eines monoton fallenden Volume-Outcome-Zusammenhangs führt dies auch zu einer Schätzung des kleinsten möglichen Effekts der Einführung einer Mindestmenge.
- **Szenario 2:** n_i' ist der Median der Fallzahlen derjenigen Standorte (berechnet auf Fallebene), die bereits die Mindestmenge eingehalten haben. Hierdurch wird die Verlegung in bestehende größere Standorte simuliert.

In der Simulation beträgt die Wahrscheinlichkeit für ein interessierendes Ereignis

$$\text{logit}(\hat{\pi}'_{ij}) = \hat{\beta}_0 + \mathbf{x}_{ij}^T \hat{\boldsymbol{\beta}} + \hat{f}_{\text{Alter}}(a_{ij}) + \hat{f}_{\text{FZ}}(n_i') + \hat{u}_i. \quad (3)$$

Wie in den Szenarien beschrieben, besteht der Unterschied zur ursprünglichen Wahrscheinlichkeit (2) nur in der Fallzahl. Die absolute Übersterblichkeit ergibt sich als Differenz aus der Anzahl der beobachteten interessierenden Ereignisse und der Prognose für diese Anzahl basierend auf der Umverteilung, wobei I die Anzahl der Standorte ist:

$$r_{\text{abs}}(M) = \sum_{i=1}^I \sum_{j=1}^{n_i} (\hat{\pi}_{ij} - \hat{\pi}'_{ij}). \quad (4)$$

(Aus mathematischen Gründen entspricht die Summe der $\hat{\pi}_{ij}$ genau der Anzahl der beobachteten interessierenden Ereignisse.)

Diese Zahl lässt sich nun zu verschiedenen Grundgesamtheiten in Bezug setzen. Zum einen zur Gesamtzahl an Fällen, um die mittlere Übersterblichkeit („average treatment effect, ATE“) zu erhalten:

$$r_{\text{mittel}}(M) = \frac{r_{\text{abs}}(M)}{n} \quad \text{mit} \quad n = \sum_{i=1}^I n_i, \quad (5)$$

Zum anderen zur Anzahl der Fälle, die von der Mindestmenge und damit auch von der Veränderung der Prognose betroffen sind; so erhält man die mittlere Übersterblichkeit unter den umverteilten Fällen („average treatment effect on the treated, ATT“):

$$r_{\text{umv}}(M) = \frac{r_{\text{abs}}(M)}{n_{\text{umv}}(M)} \quad \text{mit} \quad n_{\text{umv}}(M) = \sum_{i: n_i \neq n_i'}^I n_i. \quad (6)$$

Die letztere Kennzahl kann wiederum direkt in eine „number needed to treat“ umgerechnet werden: Wenn der Anteil an interessierenden Ereignissen unter den umverteilten Fällen um x Prozentpunkte sinkt, bedeutet dies umgekehrt, dass die Umverteilung von $100/x$ Fällen zur Vermeidung eines interessierenden Ereignisses führt. Dementsprechend ist

$$\text{NNT}(M) = \frac{1}{r_{\text{umv}}(M)}. \quad (7)$$

Basierend auf dem ursprünglich geschätzten Modell lassen sich mithilfe der Deltamethode auch Konfidenzintervalle für diese absoluten Effektmaße bestimmen (Agresti 2013).

2.3 Mathematische Annahmen und Festlegungen, Software

Bedingte Unabhängigkeit der Outcomes

Bei Modell (1) handelt es sich um ein konditionales Modell, das auf der Annahme der bedingten Unabhängigkeit der Outcomes basiert: Bezeichnet man das jeweilige binäre Outcome als y_{ij}^k (Wert 1 für Auftreten, Wert 0 für Nichtauftreten des Ereignisses), so gilt

$$y_{ij}^k \mid \mathbf{x}_{ij}^k, a_{ij}^k, n_i, u_i \stackrel{\text{unabhängig}}{\sim} \text{Ber}(\pi_{ij}^k),$$

wobei der senkrechte Strich die Bedingtheit auf alle Kovariablen ausdrückt (Fahrmeir et al. 2013). Dabei üben die unterschiedlichen Kovariablen $\mathbf{x}_{ij}^k, a_{ij}^k, n_i, u_i$ auch unabhängig voneinander Einfluss auf die Wahrscheinlichkeit π_{ij}^k aus, da das Modell keine Interaktionen enthält.

Random intercepts

Wie in der Literatur üblich (Agresti 2013), werden die random intercepts als unabhängig und normalverteilt angenommen:

$$u_i \stackrel{\text{unabhängig}}{\sim} \mathcal{N}(0, \tau^2).$$

Die zugehörige (unbekannte) Varianz $\tau^2 > 0$ quantifiziert dabei die Stärke des nicht fallzahl-assozierten Einflusses der Standorte auf das Outcome. Um eine leichter interpretierbare Größe zu erhalten, eignet sich die Umrechnung der Varianz in die MOR (Larsen et al. 2000):

$$\text{MOR}(\tau) = \exp\left(-\sqrt{2} \cdot \tau \cdot \Phi^{-1}(3/4)\right).$$

Dabei bezeichnet $\Phi^{-1}(3/4)$ das 75. Perzentil der Standardnormalverteilung. Der MOR liegt stets zwischen 0 und 1, wobei ein höherer Wert einem schwächeren Einfluss des nicht fallzahl-assozierten Einflusses der Standorte entspricht.

Eine Interpretation der MOR lautet wie folgt: Man betrachte zwei Fälle mit identischen patientenseitigen Risikofaktoren im gleichen Jahr in zwei zufällig ausgewählten Standorten mit identischer Fallzahl. Dann sind die Odds für das Auftreten des interessierenden Outcomes bei dem Fall im Median um den Faktor $\text{MOR}(\tau)$ besser, der im Standort mit dem günstigeren random intercept behandelt wurde.

Während also bei der MOR ausschließlich der nicht fallzahlassoziierte Standorteffekt variiert, können analog zum Vergleich Odds Ratios für den Fall bestimmt werden, dass nur die Fallzahl variiert, siehe Abschnitt 3.1.1 für ein konkretes Zahlenbeispiel.

Glatte Effekte

Im Rahmen von logistischer Regression (mit random intercepts) stellen pönalisierte Splines eine sehr gut geeignete Möglichkeit dar, glatte Effekte zu modellieren (siehe Fahrmeir et al. (2013)). Die glatte Anpassung an die Daten wird dabei durch eine geschickte Kombination von Polynomen erreicht und zugleich sorgt eine Bestrafung (Pönalisierung) von starker Krümmung für die Vermeidung von Overfitting. Die konkret für diesen Bericht genutzten „thin plate regression splines“ (Wood 2003) benötigen dazu nur die Wahl eines einzigen Parameters – den Grad des Bestrafungsterms, der allerdings den natürlichen Wert 2 besitzt. Die Robustheit der Ergebnisse in Bezug auf diesen Parameter und auch in Bezug auf alternative Spline-Typen wird in Abschnitt 3.3 besprochen.

Ein glatter Effekt kann jeweils einschließlich eines simultanen Konfidenzbands geschätzt und abgebildet werden (Marra und Wood 2012). Hinzu kommt ein statistischer Test der Hypothese, dass kein Effekt besteht (Wood 2013); dieser ermöglicht z. B. die Einschätzung, ob es einen statistisch signifikanten Fallzahleffekt gibt.

Technische Umsetzung

Das IQTIG hat alle Analysen mit der Software R durchgeführt (R Core Team 2022) und dort insbesondere das Paket mgcv genutzt (Wood 2022). Dieses ermöglicht u. a. die gemeinsame, pönalisierte Maximum-Likelihood-Schätzung (mit Quantifizierung von Unsicherheit) aller Modellkomponenten und bietet zudem eine Reihe von alternativen Spezifikationen für die glatten Effekte an.

3 Ergebnisse

3.1 Geschätzte Volume-Outcome-Zusammenhänge

3.1.1 Sterblichkeit

Einfluss der Patienteneigenschaften

Im Rahmen der üblichen Qualitätssicherung wurde für den zugrunde liegenden Qualitätsindikator für die Qualitätsindikatorendatenbank (QIDB) 2019 ein Risikoadjustierungsmodell entwickelt (IQTIG 2020a). Die dort berücksichtigten patientenseitigen Risikofaktoren stellen auch im Rahmen der Volume-Outcome-Analyse eine angemessene Auswahl dar. Die folgende Tabelle 6 benennt diese Faktoren und gibt die Effekte an, die sich basierend auf Modell (1) mit der in Abschnitt 2.1 definierten Fallzahl und der in Abschnitt 1.3 beschriebenen Datenbasis ergeben.

Tabelle 6: Kategorielle Risikofaktoren zu Sterblichkeit

Faktor	logistischer Koeffizient	95%-KI (Koeffizient)	Odds Ratio	95%-KI (Odds Ratio)
Intercept				
2018	-4,700777	[-4,878; -4,524]	-	-
2019	-4,735542	[-4,909; -4,562]	-	-
Geschlecht weiblich	0,228641	[0,104; 0,354]	1,257	[1,109; 1,424]
Body-Mass-Index (BMI) < 22	0,325547	[0,161; 0,490]	1,385	[1,175; 1,632]
Herzinsuffizienz NYHA IV	0,742774	[0,584; 0,901]	2,102	[1,794; 2,462]
Inotrope i. v. (präoperativ)	0,855451	[0,552; 1,159]	2,352	[1,737; 3,186]
Kardiogener Schock innerhalb der letzten 48 Stunden	0,640952	[0,406; 0,876]	1,898	[1,501; 2,401]
Reanimation innerhalb der letzten 48 Stunden	1,424269	[0,967; 1,881]	4,155	[2,631; 6,561]
Pulmonale Hypertonie > 55 mmHg	0,316732	[0,160; 0,474]	1,373	[1,173; 1,606]

Faktor	logistischer Koeffizient	95%-KI (Koeffizient)	Odds Ratio	95%-KI (Odds Ratio)
Herzrhythmus: Vorhofflimmern oder andere Herzrhythmusstörungen	0,256232	[0,134; 0,379]	1,292	[1,143; 1,461]
LVEF				
unter 20	0,706671	[0,398; 1,015]	2,027	[1,490; 2,759]
zwischen 21 und 30	0,657183	[0,454; 0,860]	1,929	[1,575; 2,363]
zwischen 31 und 50	0,351372	[0,215; 0,488]	1,421	[1,239; 1,629]
Arterielle Gefäß- erkrankung	0,490618	[0,365; 0,617]	1,633	[1,440; 1,853]
Lungenerkrankung: COPD oder andere	0,232496	[0,091; 0,374]	1,262	[1,096; 1,453]
Präoperative Nierenersatztherapie oder präoperativer Kreatininwert > 2,3 mg/dl	0,847769	[0,684; 1,012]	2,334	[1,981; 2,751]
mechanische Kreislaufunterstützung				
ja, IABP	1,314762	[0,244; 2,385]	3,724	[1,276; 10,864]
ja, andere	1,630583	[0,925; 2,336]	5,107	[2,522; 10,339]

Für den Effekt des Alters wurde – im Wesentlichen aus den gleichen Gründen wie bei der Fallzahl – auf eine glatte Modellierung zurückgegriffen. Der zugehörige Effekt ist in Abbildung 3 dargestellt. Die abgetragenen Effekte können mit den Werten in Tabelle 6 verglichen werden.

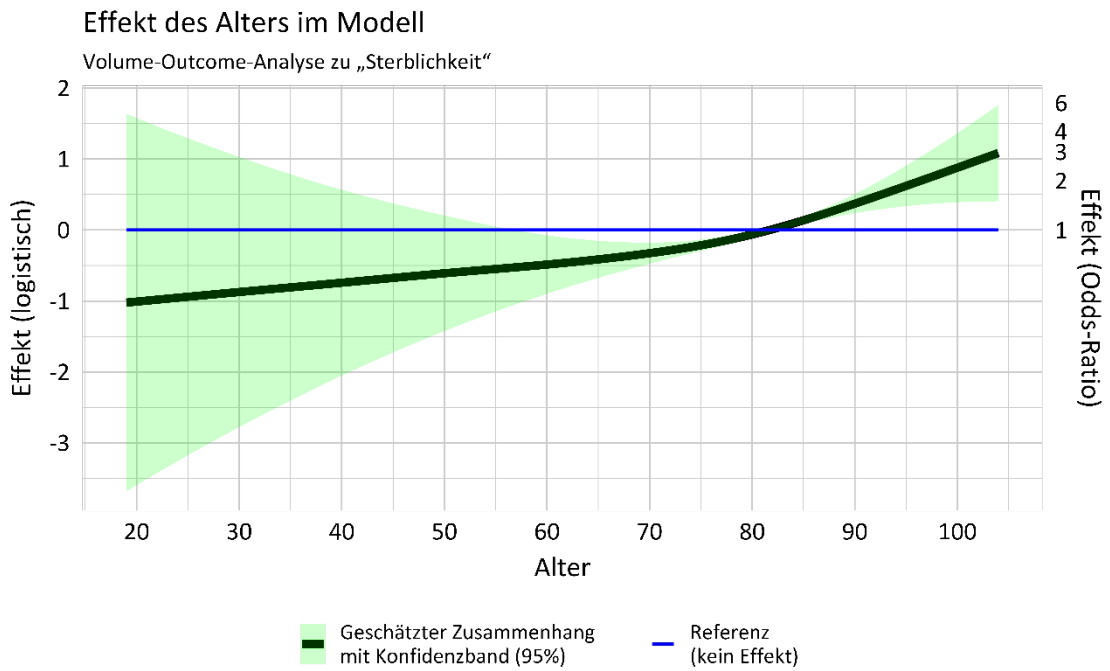


Abbildung 3: Effekt des Alters in der Risikoadjustierung zu Sterblichkeit (logistisch)

Zur besseren Veranschaulichung zeigt Abbildung 4 einen bereits auf Wahrscheinlichkeitsebene transformierten Effekt.

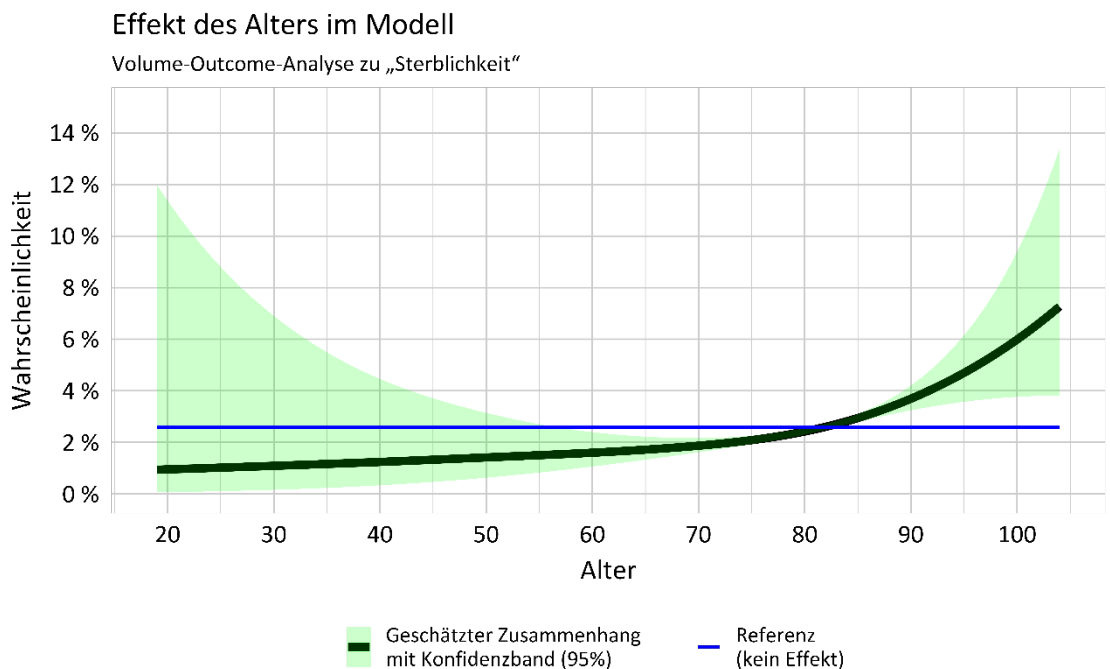


Abbildung 4: Effekt des Alters in der Risikoadjustierung zu Sterblichkeit

Die relative Häufigkeit an Todesfällen, d. h. die durchschnittliche Wahrscheinlichkeit des Versterbens, beträgt etwa 2,61 % (siehe Tabelle 1), markiert durch die blaue Linie. Die individuelle Wahrscheinlichkeit des Versterbens für einen Fall hängt allerdings vom Alter ab und die

schwarze Linie beschreibt, welche Abweichung vom Durchschnittswert sich dadurch ergibt. Erwartungsgemäß steigt die Wahrscheinlichkeit mit dem Alter.

Einfluss der Standorteigenschaften

Beim Outcome Sterblichkeit zeigen sich ein statistisch signifikanter Volume-Outcome-Zusammenhang und zudem ein großer Einfluss der nicht fallzahlassoziierten Standorteigenschaften. Tabelle 7 enthält die Kennzahlen zu diesen beiden Aspekten.

Tabelle 7: Kennzahlen der LE-Einflüsse und Modellgüte bei Sterblichkeit

p-Wert zum Fallzahleffekt	0,000021
MOR der LE-Effekte mit 95%-KI	0,727; [0,656; 0,786]
AUC des Modells (out of sample)	0,729
p-Werte zum Test auf Kalibrierung des Modells (out of sample)	Verschiebung: 0,537 Skalierung: 0,469

Die statistische Signifikanz beider Einflüsse leitet sich daraus ab, dass sowohl der p-Wert deutlich unter den üblichen Niveaus von 5 % oder 1 % liegt als auch das 95%-Konfidenzintervall um die MOR nicht die 1 einschließt.

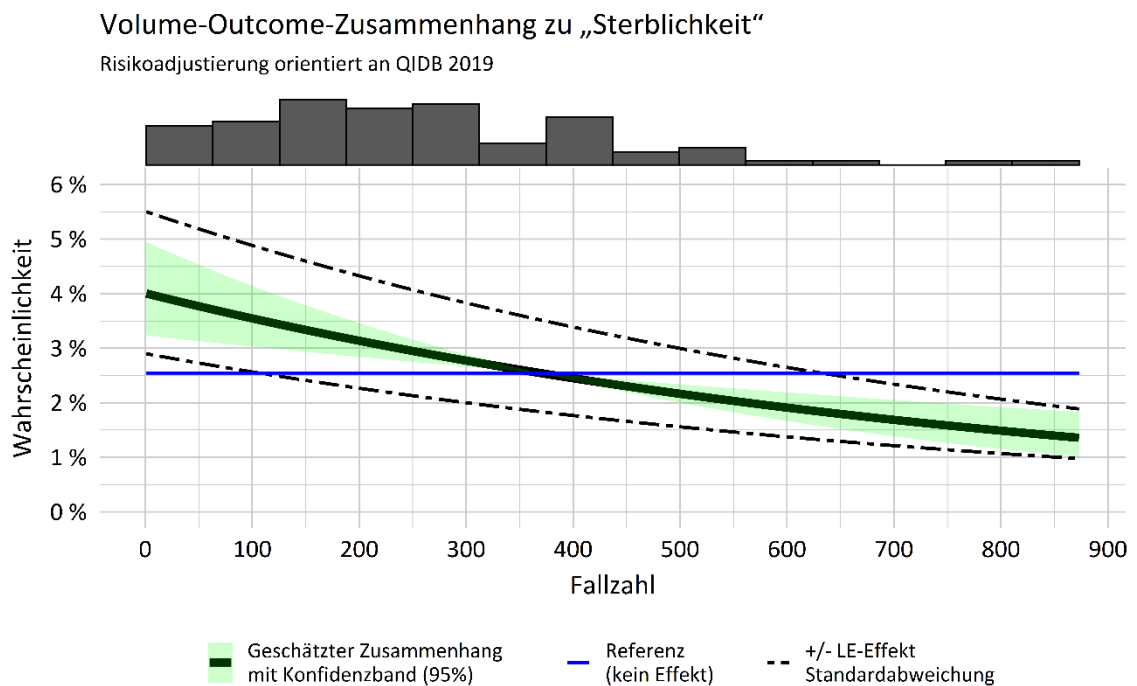


Abbildung 5: Geschätzter Fallzahleffekt zu Sterblichkeit

Abbildung 5 lässt sich sehr ähnlich zu Abbildung 4 interpretieren: Die blaue Linie markiert wieder die durchschnittliche Sterblichkeit von etwa 2,61 % und die schwarze Linie zeigt, inwiefern die Fallzahl des behandelnden Standortes diese Wahrscheinlichkeit beeinflusst, wenn ausschließlich die Fallzahl variiert wird und alle anderen patientenseitigen oder standortspezifischen Eigenschaften bestehen bleiben. Grundsätzlich fällt diese Wahrscheinlichkeit monoton mit der Fallzahl und dieser Zusammenhang ist statistisch signifikant – der geringe p-Wert für den Fallzahleffekt entspricht in der Abbildung der Tatsache, dass das Konfidenzband um den geschätzten Fallzahleffekt die Referenzlinie nicht vollständig überdeckt. Ein wichtiger Hinweis besteht hierbei darin, dass die Wahl einer Referenzwahrscheinlichkeit nur zum Zwecke einer angemessenen Darstellung notwendig ist; für die Signifikanz oder die Monotonie des Effekts spielt dies keine Rolle.

Beim Ziehen von Schlussfolgerungen aus diesem Ergebnis muss beachtet werden, dass ein statistisch signifikanter, monoton fallender Zusammenhang nicht zwangsläufig auch ein besonders starker oder relevanter Zusammenhang ist: Es braucht etwa den beachtlichen Sprung von 250 Fällen auf 550 Fälle, damit die Wahrscheinlichkeit um einen Prozentpunkt, von etwa 3 % auf etwa 2 %, sinkt. Weitere Beispiele, auch in absoluten Zahlen, folgen in Abschnitt 3.2.

Neben einer Darstellung des Einflusses der Fallzahl auf die Wahrscheinlichkeit bietet sich auch eine Berechnung von Odds Ratios an, um die Stärke des Einflusses darzustellen. Die folgende Tabelle 8 enthält solche Odds Ratios.

Tabelle 8: Geschätzte Odds Ratios bezüglich Variation der Fallzahl

Fallzahl 1/2	1	50	100	200	400
1	1	0,9397	0,8818	0,7766	0,6024
50	–	1	0,9385	0,8265	0,6411
100	–	–	1	0,8807	0,6831
200	–	–	–	1	0,7756
400	–	–	–	–	1

Zur Interpretation: Man betrachte einen beliebigen Fall und variiert wieder ausschließlich die Fallzahl des behandelnden Standorts. Dann bewirkt eine Erhöhung der Fallzahl von zum Beispiel 50 auf etwa 200 eine Verbesserung der Odds für das Versterben um den Faktor 0,8265. Diese Werte lassen sich unmittelbar zur MOR in Beziehung setzen: Der zufällige Wechsel zu einem Standort mit gleicher Fallzahl, aber günstigeren, nicht fallzahlassoziierten Eigenschaften bewirkt im Median eine Verbesserung der Odds um den stärkeren Faktor 0,727. In Abbildung 5 repräsentieren die gestrichelten Linien die Stärke dieses Einflusses. Die unterschiedlichen Odds Ratios zeigen, dass die nicht fallzahlassoziierten Standorteigenschaften einen im Vergleich zur Fallzahl starken Einfluss auf die Wahrscheinlichkeit des Versterbens besitzen.

3.1.2 Intraprozedurale Komplikation

Der QI 51916 ist nicht risikoadjustiert und die Neuentwicklung eines Risikoadjustierungsmodells ist im Rahmen dieses Auftrags nicht vorgesehen. Zumindest eine einfache Überprüfung von Risikofaktoren wurde dennoch durchgeführt, siehe Abschnitt 3.3.

Die Modellformel, die der Volume-Outcome-Analyse zugrunde liegt, nimmt dementsprechend eine einfachere Form an:

$$\text{logit}(\pi_{ij}^k) = \beta_0^k + f_{FZ}(n_i) + u_i. \quad (8)$$

Das Fehlen patientenseitiger Risikofaktoren hat zur Folge, dass die geschätzten Standorteffekte (d. h. Fallzahleffekt und random intercept) nicht bezüglich des Case-Mixes der Standorte bereinigt sind. Für die geschätzten Standorteffekte sind unterschiedliche Konsequenzen denkbar: Falls zum Beispiel die Schwere der Fälle stark mit der Fallzahl assoziiert ist, könnte der geschätzte Fallzahleffekt tatsächlich eher die patientenseitigen als die Fallzahleinflüsse widerspiegeln. Schwanken hingegen die patientenseitigen Risiken stark und unabhängig von der Fallzahl, könnte ein etwaig vorhandener Fallzahleffekt mit übermäßiger Unsicherheit behaftet sein und dadurch nicht entdeckt werden.

Vor diesem Hintergrund ergeben sich hier die in Tabelle 9 dargestellten Kennzahlen.

Tabelle 9: Kennzahlen der LE-Einflüsse und Modellgüte bei intraprozeduraler Komplikation

p-Wert zum Fallzahleffekt	0,961611
MOR der LE-Effekte mit 95%-KI	0,596; [0,518; 0,666]
AUC des Modells (out of sample)	0,621
p-Werte zum Test auf Kalibrierung des Modells (out of sample)	Verschiebung: 0,276 Skalierung: 0,477

Das zentrale Ergebnis ist der sehr hohe p-Wert für den Fallzahleffekt: Tatsächlich ist der geschätzte Fallzahleffekt bei diesem Outcome weder signifikant noch relevant. Ähnlich wie bei der Sterblichkeit erscheint der Einfluss der nicht fallzahlassoziierten Standorteigenschaften als groß. Die sehr geringe AUC ist hier allein deshalb zu erwarten, weil Prognosen auf Basis von Modell (8) praktisch ausschließlich vom behandelnden Standort abhängen und insofern sehr grob sind.

In Abbildung 6 äußert sich der sehr hohe p-Wert für den Fallzahleffekt dadurch, dass der Fallzahleffekt praktisch konstant ist:

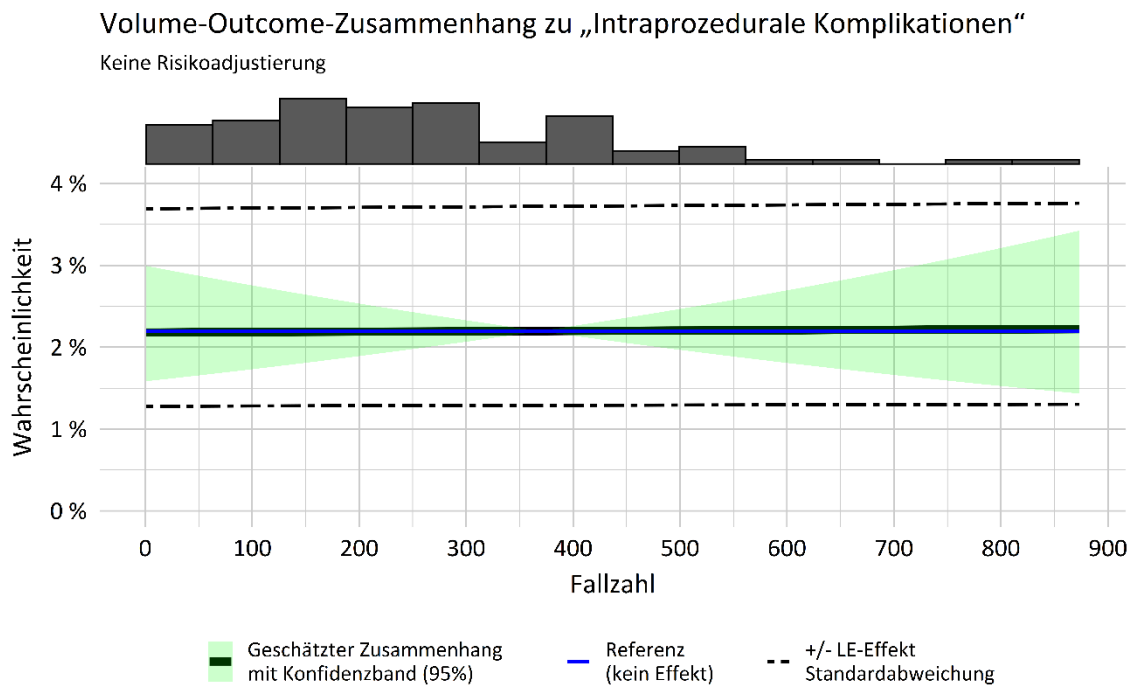


Abbildung 6: Geschätzter Fallzahleffekt zu intraprozeduraler Komplikation

3.1.3 Gefäßkomplikation

Auch der zu diesem Outcome gehörige QI 52007 ist nicht risikoadjustiert. Eine Abgrenzung zwischen patienten- und standortseitigen Einflüssen erscheint aus medizinisch-fachlicher Perspektive hier aber stärker angezeigt ist als im vorigen Abschnitt 3.1.2. Daher wurden bei einer einfachen Literaturrecherche im Rahmen dieses Auftrags einige patientenseitige Risikofaktoren identifiziert und in das Modell aufgenommen (z. B. Scarsini et al. (2019), Potluri et al (2020), Ueshima et al. (2020), Gonska et al. (2021), Mach et al. (2021)), sodass es die Form (1) annimmt.

Zunächst in Tabelle 10 ein Überblick über die kategoriellen patientenseitigen Risikofaktoren.

Tabelle 10: Kategorielle Risikofaktoren zu Gefäßkomplikation

Faktor	logistischer Koeffizient	95%-KI (Koeffizient)	Odds Ratio	95%-KI (Odds Ratio)
Intercept				
2018	-4,780349	[-5,028; -4,533]	-	-
2019	-4,959387	[-5,208; -4,711]	-	-
Geschlecht weiblich	0,653683	[0,499; 0,809]	1,923	[1,646; 2,245]
Body-Mass-Index (BMI) < 22	0,279302	[0,077; 0,482]	1,322	[1,080; 1,619]

Faktor	logistischer Koeffizient	95%-KI (Koeffizient)	Odds Ratio	95%-KI (Odds Ratio)
Pulmonale Hypertonie > 55 mmHg	0,211918	[0,000; 0,424]	1,236	[1,000; 1,528]
Arterielle Gefäß-erkrankung	0,454314	[0,298; 0,611]	1,575	[1,347; 1,842]

Ähnlich wie bei der Betrachtung der Sterblichkeit geht das Alter als stetige Variable mit glatterm Effekt in das Modell ein. Allerdings ist der Alterseffekt hier nicht signifikant, wie in Abbildung 7 bzw. Abbildung 8 ersichtlich.

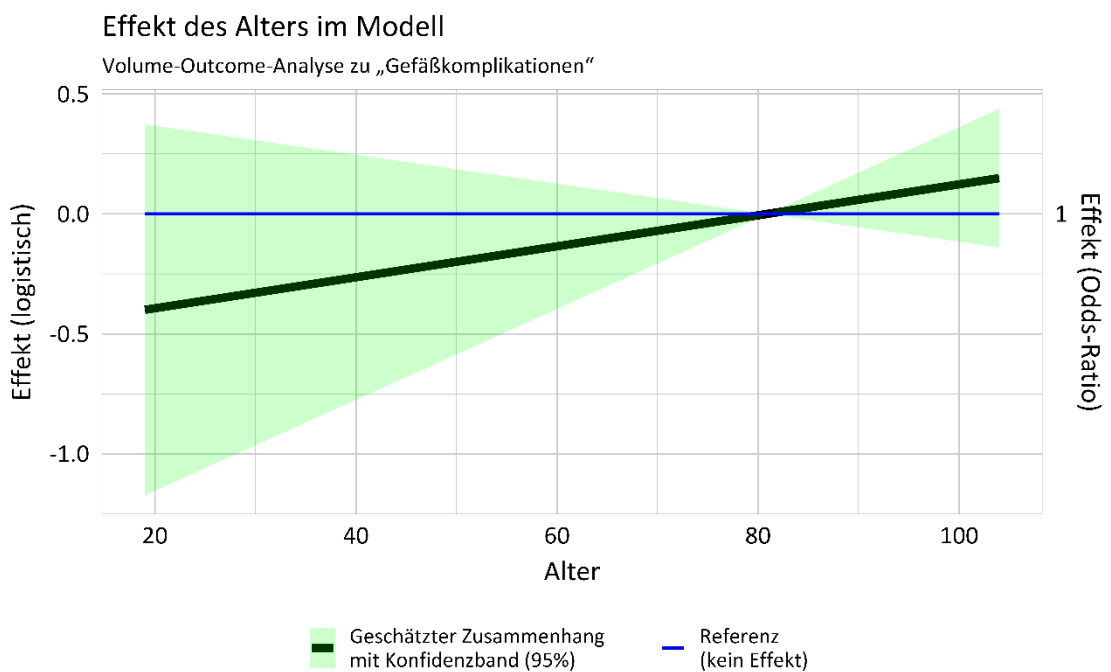


Abbildung 7: Effekt des Alters in der Risikoadjustierung zu Gefäßkomplikation (logistisch)

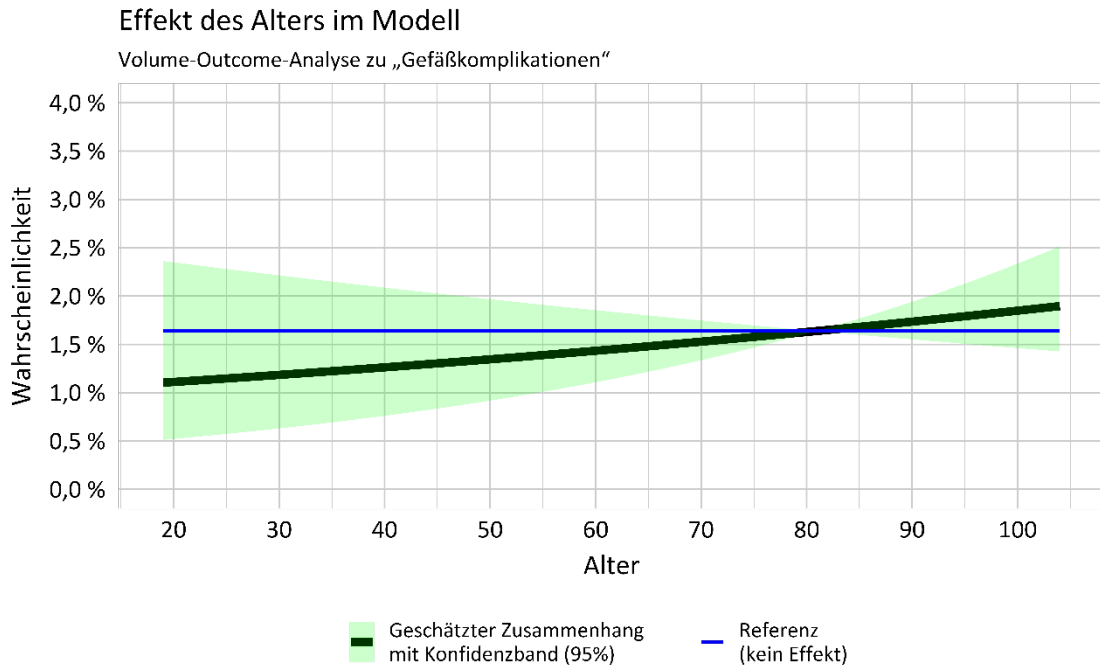


Abbildung 8: Effekt des Alters in der Risikoadjustierung zu Gefäßkomplikation

Hinsichtlich der standortspezifischen Effekte erhält man die Kennzahlen in Tabelle 11.

Tabelle 11: Kennzahlen der LE-Einflüsse und Modellgüte bei Gefäßkomplikation

p-Wert zum Fallzahleffekt	0,595544
MOR der LE-Effekte mit 95%-KI	0,496; [0,417; 0,570]
AUC des Modells (out of sample)	0,696
p-Werte zum Test auf Kalibrierung des Modells (out of sample)	Verschiebung: 0,481 Skalierung: 0,676

Die grundsätzlichen Schlussfolgerungen entsprechen denen zur intraprozeduralen Komplikation (Abschnitt 3.1.2): Während kein signifikanter Fallzahleffekt festzustellen ist, zeigt sich ein starker Einfluss der nicht fallzahlassoziierten Standorteinflüsse. In Abbildung 9 drückt sich das Ergebnis zum Fallzahleffekt wiederum dadurch aus, dass das Konfidenzband die Referenzlinie überdeckt.

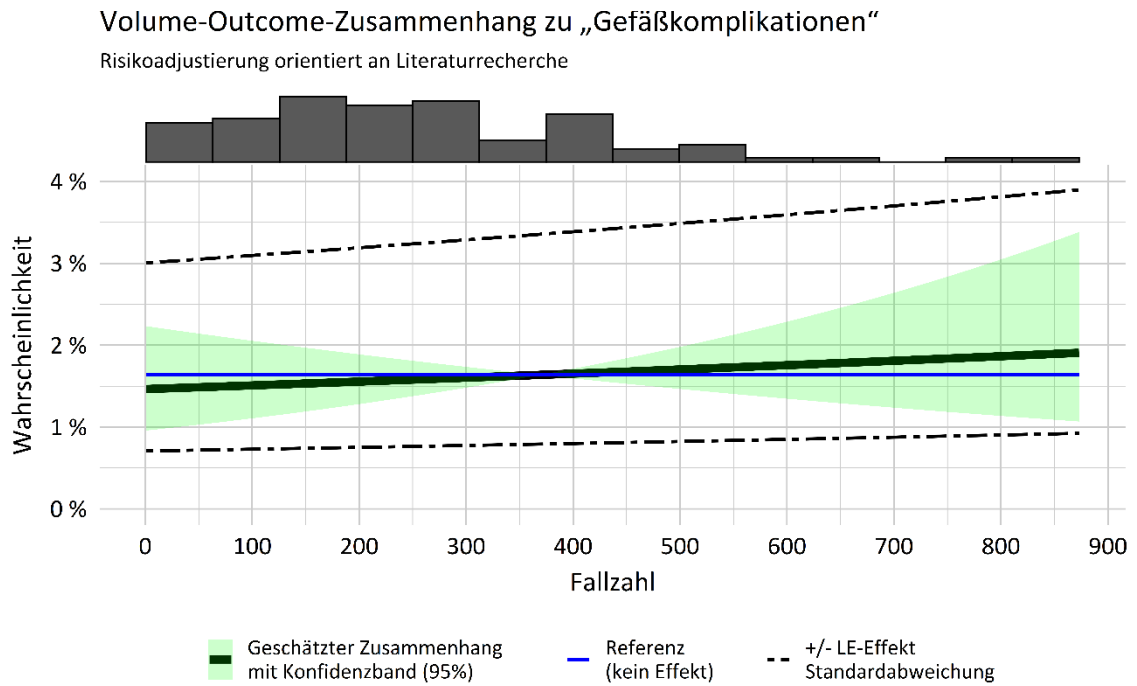


Abbildung 9: Geschätzter Fallzahleffekt zu Gefäßkomplikation

3.2 Geschätzte Übersterblichkeit basierend auf hypothetischen Mindestmengen

Im Sinne der in Kapitel 2 beschriebenen Methodik können Mindestmengen nur dann Auswirkungen auf die Behandlungsqualität haben, wenn ein (von Null verschiedener) Volume-Outcome-Zusammenhang besteht. Dementsprechend widmet sich dieser Abschnitt nur dem Outcome Sterblichkeit.

Abbildung 10 soll zunächst unabhängig vom Outcome darstellen, wie viele Fälle bzw. welcher Anteil von Fällen im Jahr 2019 jeweils von einer Mindestmenge betroffen gewesen wäre, das heißt potenziell hätte umverlegt werden müssen.

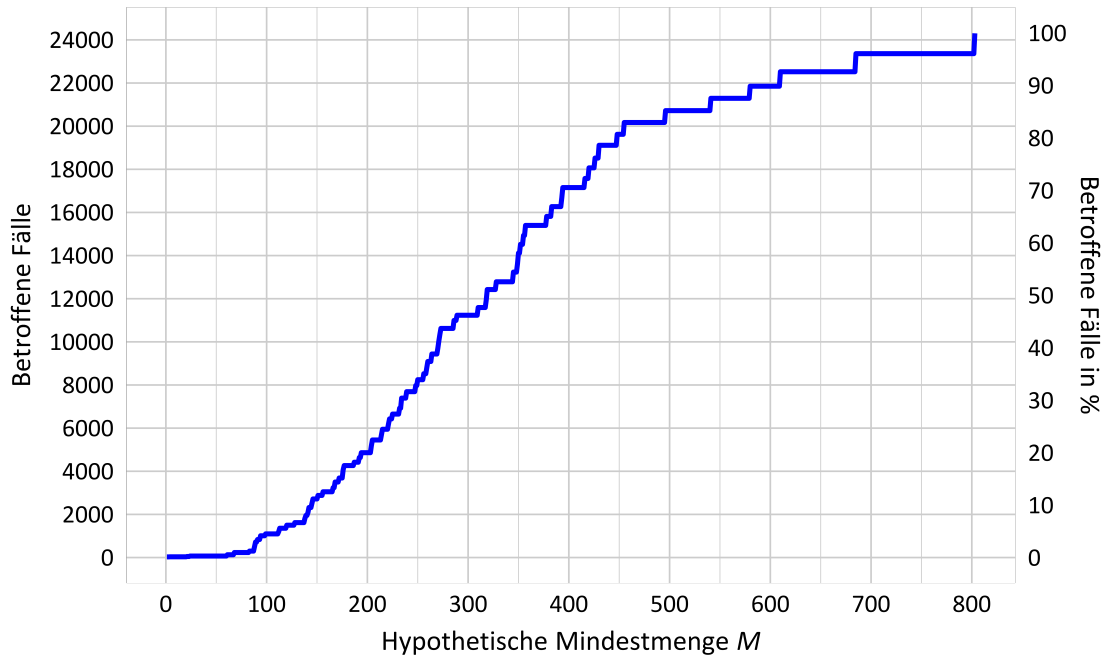


Abbildung 10: Von einer hypothetischen Mindestmenge betroffene Fälle 2019

Entsprechend der in Abschnitt 2.2 beschriebenen Prinzipien ergeben sich für die beiden Szenarien jeweils folgende Werte für die absolute (Formel (4)) und mittlere Übersterblichkeit, Letztere dabei einerseits in Bezug zur Gesamtfallzahl (Formel (5), Abbildung 11) und andererseits in Bezug zur jeweiligen Anzahl der betroffenen Fälle (Formel (6), Abbildung 12). Zur besseren Sichtbarkeit sind die Grafiken zunächst auf den Bereich bis zur Fallzahl 300 eingeschränkt.

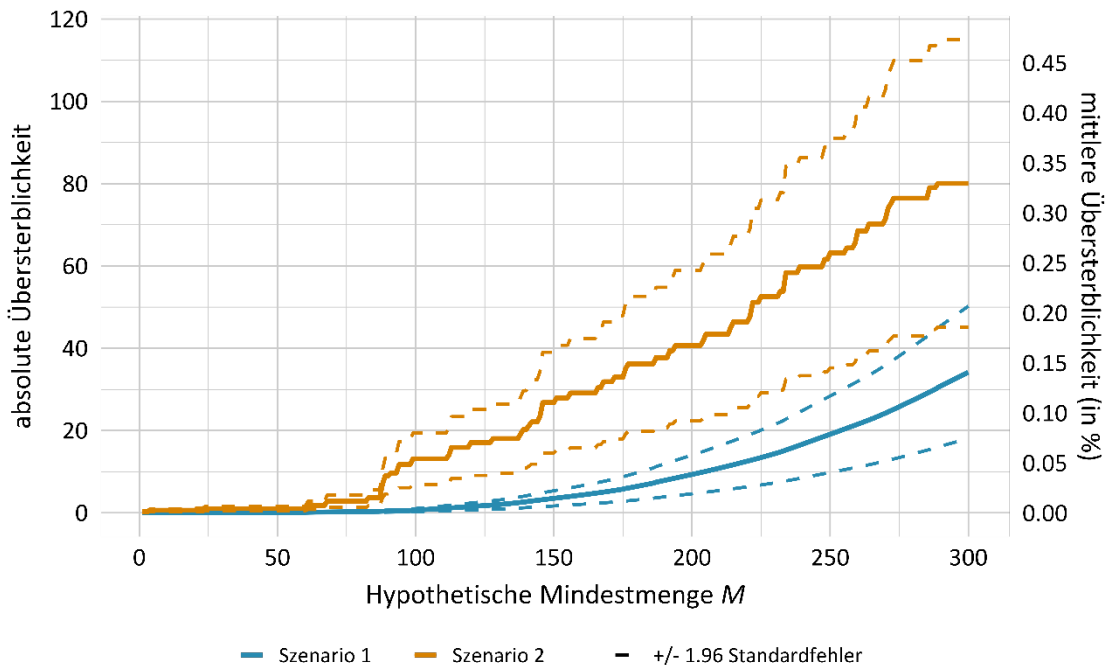


Abbildung 11: Absolute und mittlere Übersterblichkeit zu hypothetischen Mindestmengen

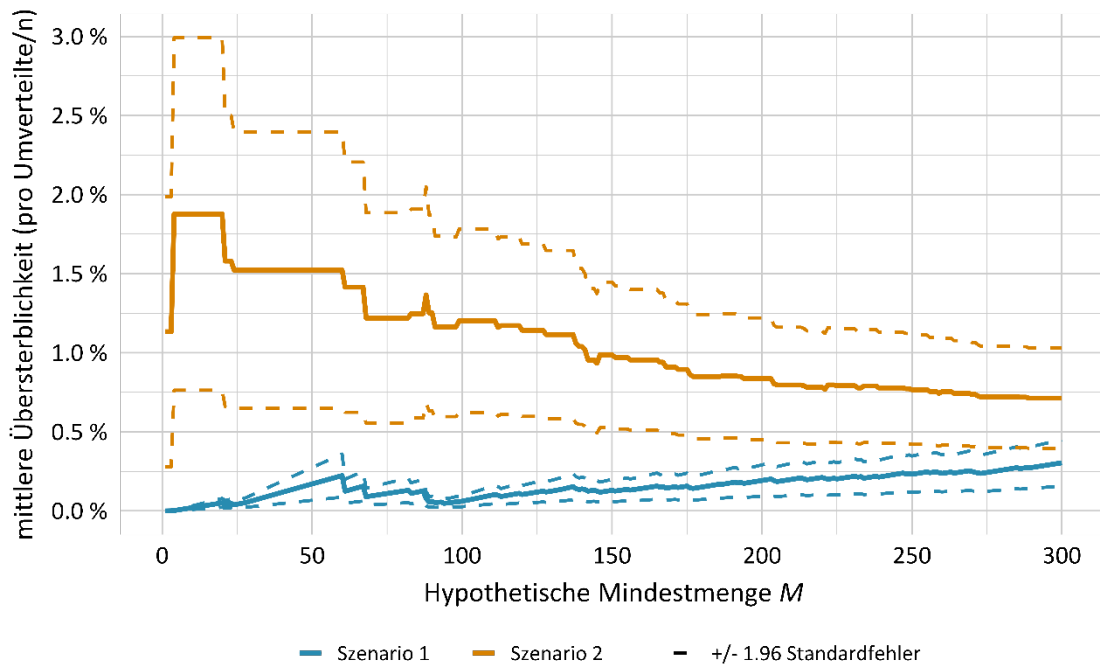


Abbildung 12: Mittlere Übersterblichkeit unter den umverteilten Fällen zu hypothetischen Mindestmengen

Zum besseren Verständnis dieser Grafiken und ihres Zusammenhangs zum Fallzahleffekt (Abbildung 5) soll die Betrachtung eines konkreten Beispiels dienen:

Angenommen, es wird eine Mindestmenge bei 150 angesetzt und Szenario 1 betrachtet. Dann geht zunächst aus Abbildung 10 hervor, dass diese Mindestmenge im Rückblick die Umverteilung von 2.716 Fällen bedeutet. Für die erwartete Reduktion der Sterblichkeit lassen sich nun zwei Grafiken heranziehen: Für einen durchschnittlichen Fall entspricht der Unterschied zwischen Fallzahl 1 und Fallzahl 150 gemäß Abbildung 5 einem Unterschied in der Wahrscheinlichkeit des Versterbens von 4 % zu 3,3 %, d. h. eine Verbesserung um 0,7 Prozentpunkte. Demgegenüber steht gemäß Abbildung 12 eine Verbesserung der Wahrscheinlichkeit des Versterbens (mittlere Übersterblichkeit unter Umverteilten) um nur 0,13 Prozentpunkte. Ein Aspekt, der diesen Unterschied mit erklärt, besteht darin, dass bei der Berechnung der Übersterblichkeit der betroffene Case-Mix berücksichtigt wird. Deutlich ausschlaggebender ist jedoch die Tatsache, dass die Fallzahlerhöhung in aller Regel deutlich geringer ist als von 1 auf 150. Im Median wurden die betroffenen (d. h. potenziell umverteilten) Fälle an Standorten mit der Fallzahl 98 behandelt. Insofern sollten für einen sinnvollen Vergleich in Abbildung 5 eher die Werte an den Stellen 98 und 150 verglichen werden, was natürlich dennoch sehr grob ist. Dabei ergibt sich eine Differenz von nur etwa 0,2 Prozentpunkten. Bezogen auf alle Fälle – nicht nur die umverteilten – fällt die Verbesserung natürlich noch einmal deutlich geringer aus. Die globale Rate von etwa 2,61 % (Tabelle 1) verringert sich nur um 0,015 Prozentpunkte, was etwa 3,5 potenziell vermeidbaren Todesfällen entspricht (Abbildung 11).

Für Szenario 2 kann analog vorgegangen werden. Hier bedeutet die Mindestmenge 150, dass sich die Fallzahl bei den umverteilten Fällen von im Median 98 auf im Median 280 erhöht. Die Wahrscheinlichkeit für einen durchschnittlichen Fall in Abbildung 5 fällt zwischen diesen beiden

Stellen um etwa 0,71 Prozentpunkte. Ohne die Medianbildung für die umverteilten Fälle und unter Berücksichtigung des Case-Mixes liegt die Verbesserung unter den umverteilten Fällen bei etwa einem Prozentpunkt (Abbildung 12) und bezogen auf alle Fälle bei etwa 0,11 Prozentpunkten. Dies entspricht etwa 26,8 potenziell vermeidbaren Todesfällen.

Diese Beispiele sind selbstverständlich nur die Punktschätzer, die von z. T. großer statistischer Unsicherheit behaftet sind. Die Unsicherheit drückt sich durch das Konfidenzband (Abbildung 5) bzw. die punktweise berechneten Konfidenzintervalle (Abbildung 11 und Abbildung 12) aus.

Die folgenden Tabellen (Tabelle 12 und Tabelle 13) enthalten noch weitere in Abbildung 11 und Abbildung 12 dargestellte Ergebnisse in Zahlen, wobei zusätzlich jeweils die NNT, Formel (7), angegeben ist, die in Abschnitt 2.3 definiert wurde.

Tabelle 12: Übersterblichkeit nach Szenario 1 in Zahlen mit 95%-Konfidenzintervallen

M	absolute Übersterblichkeit	mittlere Übersterblichkeit (gesamt)	mittlere Übersterblichkeit (Umverteilungen)	NNT
10	0,007 (± 0,004)	0,000 % (± 0,000 %)	0,020 % (± 0,011 %)	4893,358 (± 2695,313)
20	0,017 (± 0,009)	0,000 % (± 0,000 %)	0,051 % (± 0,028 %)	1974,552 (± 1080,104)
30	0,044 (± 0,025)	0,000 % (± 0,000 %)	0,064 % (± 0,037 %)	1561,012 (± 890,440)
40	0,081 (± 0,048)	0,000 % (± 0,000 %)	0,118 % (± 0,069 %)	850,383 (± 499,230)
50	0,118 (± 0,070)	0,000 % (± 0,000 %)	0,171 % (± 0,101 %)	586,084 (± 348,436)
60	0,154 (± 0,092)	0,001 % (± 0,000 %)	0,223 % (± 0,133 %)	448,119 (± 267,967)
70	0,222 (± 0,131)	0,001 % (± 0,001 %)	0,095 % (± 0,056 %)	1047,156 (± 617,194)
80	0,288 (± 0,169)	0,001 % (± 0,001 %)	0,124 % (± 0,073 %)	804,316 (± 472,352)
90	0,423 (± 0,244)	0,002 % (± 0,001 %)	0,059 % (± 0,034 %)	1699,185 (± 977,840)
100	0,674 (± 0,377)	0,003 % (± 0,002 %)	0,061 % (± 0,034 %)	1626,298 (± 908,861)
120	1,554 (± 0,838)	0,006 % (± 0,003 %)	0,104 % (± 0,056 %)	964,450 (± 520,080)
140	2,712 (± 1,433)	0,011 % (± 0,006 %)	0,139 % (± 0,073 %)	719,638 (± 380,283)

M	absolute Übersterblichkeit	mittlere Übersterblichkeit (gesamt)	mittlere Übersterblichkeit (Umverteilungen)	NNT
160	4,344 (± 2,254)	0,018 % (± 0,009 %)	0,142 % (± 0,074 %)	701,953 (± 364,313)
180	6,441 (± 3,290)	0,027 % (± 0,014 %)	0,151 % (± 0,077 %)	662,000 (± 338,150)
200	9,342 (± 4,698)	0,038 % (± 0,019 %)	0,192 % (± 0,097 %)	520,348 (± 261,710)
220	12,569 (± 6,239)	0,052 % (± 0,026 %)	0,211 % (± 0,105 %)	473,145 (± 234,845)
240	16,701 (± 8,176)	0,069 % (± 0,034 %)	0,217 % (± 0,106 %)	460,439 (± 225,392)
260	21,559 (± 10,415)	0,089 % (± 0,043 %)	0,237 % (± 0,115 %)	421,582 (± 203,662)
280	27,405 (± 13,070)	0,113 % (± 0,054 %)	0,258 % (± 0,123 %)	387,591 (± 184,845)
300	34,153 (± 16,087)	0,141 % (± 0,066 %)	0,304 % (± 0,143 %)	328,898 (± 154,921)

Tabelle 13: Übersterblichkeit nach Szenario 2 in Zahlen mit 95%-Konfidenzintervallen

M	absolute Übersterblichkeit	mittlere Übersterblichkeit (gesamt)	mittlere Übersterblichkeit (Umverteilungen)	NNT
10	0,638 (± 0,371)	0,003 % (± 0,002 %)	1,877 % (± 1,092 %)	53,274 (± 30,988)
20	0,638 (± 0,371)	0,003 % (± 0,002 %)	1,877 % (± 1,092 %)	53,274 (± 30,988)
30	1,050 (± 0,590)	0,004 % (± 0,002 %)	1,522 % (± 0,855 %)	65,693 (± 36,911)
40	1,050 (± 0,590)	0,004 % (± 0,002 %)	1,522 % (± 0,855 %)	65,693 (± 36,911)
50	1,050 (± 0,590)	0,004 % (± 0,002 %)	1,522 % (± 0,855 %)	65,693 (± 36,911)
60	1,050 (± 0,590)	0,004 % (± 0,002 %)	1,522 % (± 0,855 %)	65,693 (± 36,911)
70	2,829 (± 1,513)	0,012 % (± 0,006 %)	1,220 % (± 0,652 %)	81,995 (± 43,834)

M	absolute Übersterblichkeit	mittlere Übersterblichkeit (gesamt)	mittlere Übersterblichkeit (Umverteilungen)	NNT
80	2,829 (± 1,513)	0,012 % (± 0,006 %)	1,220 % (± 0,652 %)	81,995 (± 43,834)
90	9,003 (± 4,364)	0,037 % (± 0,018 %)	1,252 % (± 0,607 %)	79,862 (± 38,709)
100	13,160 (± 6,236)	0,054 % (± 0,026 %)	1,201 % (± 0,569 %)	83,285 (± 39,470)
120	17,125 (± 8,010)	0,070 % (± 0,033 %)	1,142 % (± 0,534 %)	87,532 (± 40,943)
140	20,288 (± 9,422)	0,083 % (± 0,039 %)	1,039 % (± 0,483 %)	96,214 (± 44,680)
160	29,116 (± 13,283)	0,120 % (± 0,055 %)	0,955 % (± 0,436 %)	104,721 (± 47,777)
180	36,229 (± 16,413)	0,149 % (± 0,068 %)	0,850 % (± 0,385 %)	117,695 (± 53,319)
200	40,638 (± 18,318)	0,167 % (± 0,075 %)	0,836 % (± 0,377 %)	119,617 (± 53,918)
220	46,422 (± 20,828)	0,191 % (± 0,086 %)	0,781 % (± 0,350 %)	128,107 (± 57,476)
240	59,818 (± 26,478)	0,246 % (± 0,109 %)	0,778 % (± 0,344 %)	128,557 (± 56,904)
260	68,518 (± 30,125)	0,282 % (± 0,124 %)	0,754 % (± 0,331 %)	132,652 (± 58,323)
280	76,459 (± 33,444)	0,315 % (± 0,138 %)	0,720 % (± 0,315 %)	138,924 (± 60,766)
300	80,065 (± 34,893)	0,329 % (± 0,144 %)	0,713 % (± 0,311 %)	140,299 (± 61,144)

3.3 Ergänzungen

Robustheit der Ergebnisse

Wie an verschiedenen Stellen in Kapitel 2 erwähnt, hängt die Brauchbarkeit der vorgestellten Ergebnisse auch davon ab, inwiefern sie gegenüber gewissen Operationalisierungen oder Wahlmöglichkeiten robust sind. Im Rahmen der Analyse wurden folgende Alternativen geprüft (sog. Sensitivitätsanalysen). Dabei konnte kein nennenswerter Einfluss auf die Ergebnisse festgestellt werden.

- **Spezifikation der Splines.** Neben den genutzten „thin plate regressen splines“ stellen auch kubische Splines, genauer „cardinal“ splines (Wood 2017), eine bewährte Wahl dar. Bei diesen ist ein anderer Parameter zu wählen als der Bestrafungsgrad, und zwar die Anzahl der sogenannten Knoten, zwischen denen jeweils ein Polynom dritten Grades den gesuchten Zusammenhang abbilden soll. Die Variation der Knotenzahl bietet eine gute Möglichkeit zur Überprüfung der Robustheit.
- **Operationalisierung der Fallzahl.** Die in diesem Bericht genutzten Fallzahleffekte beruhen auf der durchschnittlichen Fallzahl der Standorte 2018 bis 2019. Alternativ ist zum Beispiel denkbar, direkt die jeweilige Fallzahl des Jahres zu nutzen (dann würde auch n_i in Formel (1) den Index k erhalten) oder auch nur die Fallzahl des Jahres 2018, um besser die zeitliche Abfolge zu berücksichtigen.
- **Risikoadjustierung.** Hinsichtlich der beiden Komplikationsoutcomes (siehe Abschnitte 3.1.2 und 3.1.3) verbleibt eine gewisse Unsicherheit, ob die patientenseitigen Einflüsse ausreichend berücksichtigt wurden. Bei den Gefäßkomplikationen wurden die recherchierten Risikofaktoren in das Modell aufgenommen. Darüber hinaus erfolgten für beide Komplikationsoutcomes Analysen, bei denen die Risikofaktoren aus dem Modell für die Sterblichkeit (Tabelle 6) verwendet wurden.

Grafiken zur Übersterblichkeit über den gesamten Fallzahlbereich

Während bei Abbildung 11 und Abbildung 12 der Fallzahlbereich auf den mutmaßlich relevantesten Bereich eingeschränkt war, sind die folgenden Abbildungen (Abbildung 13 und Abbildung 14) zur Ergänzung nicht in diesem Sinne abgeschnitten.

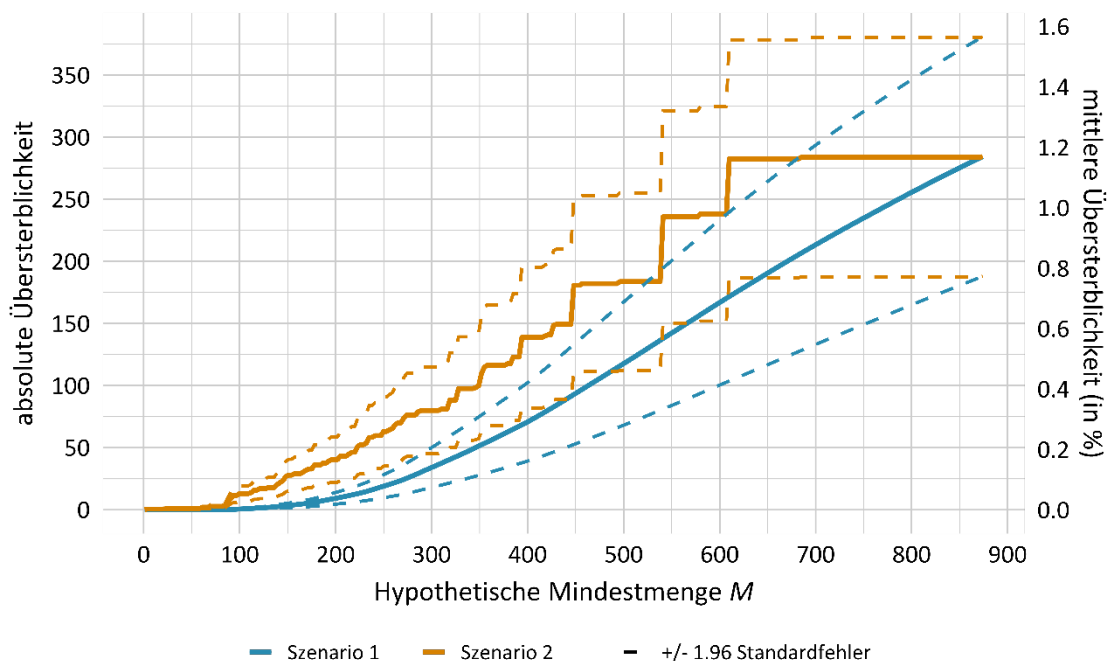


Abbildung 13: Absolute und mittlere Übersterblichkeit zu hypothetischen Mindestmengen, gesamter Bereich

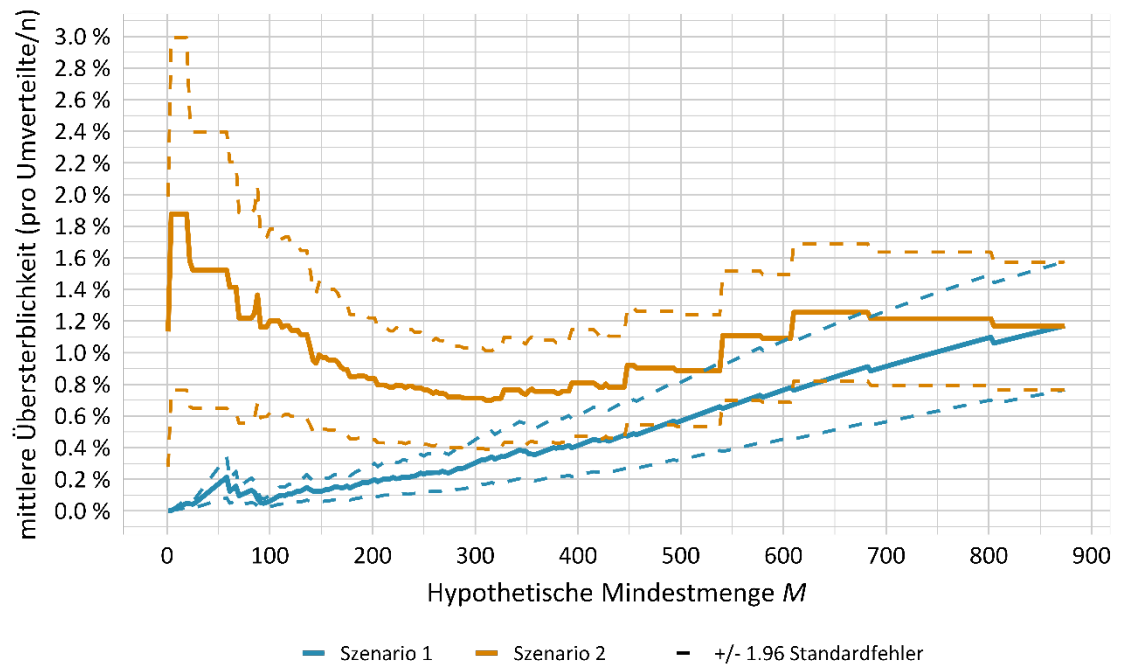


Abbildung 14: Mittlere Übersterblichkeit unter den umverteilten Fällen zu hypothetischen Mindestmengen, gesamter Bereich

4 Diskussion und Empfehlungen

Unter Verwendung der esQS-Daten der Jahre 2018 und 2019 zeigt sich für die risikoadjustierte Sterblichkeit ein statistisch signifikanter Volume-Outcome-Zusammenhang, der über den gesamten Fallzahlbereich streng monoton fallend ist. Somit sinkt die Sterbewahrscheinlichkeit für Patientinnen und Patienten mit einer TAVI bei einer Zunahme der Fälle bei sonst identischer Fallkonstellation. Da es sich bei der Auswertung um eine retrospektive Beobachtungsstudie handelt, sollte dieses Ergebnis allerdings als statistische Assoziation, nicht kausaler Zusammenhang verstanden werden.

Der beobachtete Zusammenhang sowie die darauf basierende Simulation von Mindestmengen legen keine obere Grenze für eine etwaige Mindestmenge nahe und zugleich tritt ein relevanter Effekt auf die Behandlungsqualität erst bei hohen Mindestmengen ein. Gerade im Bereich hoher Mindestmengen weist die Übersterblichkeit allerdings eine besonders große statistische Unsicherheit auf; hinzu kommt die nicht abgebildete Dynamik in Bezug auf die Versorgungslandschaft und Fallzahlverteilung.

Im Gegensatz zur Sterblichkeit zeigten sich für die untersuchten Komplikationen (Intraprozedurale Komplikation, Gefäßkomplikation) auf Basis der zur Verfügung stehenden Daten keine Volume-Outcome-Zusammenhänge. Ein Sachverhalt, der zunächst verwunderlich erscheint, als intuitiv vermutet werden könnte, dass eine erhöhte Sterblichkeit in Standorten mit geringerer Fallzahl auf Basis von (potenziell) vermeidbaren Komplikationen erklärbar ist. Dabei ist an dieser Stelle anzumerken, dass auch in der Literaturrecherche des IQWiG kein Zusammenhang zwischen Fallzahl und Komplikationen für TAVI bestätigt wurde. Zusätzlich lassen sich ähnliche Ergebnisse im Sinne von Volume-Outcome-Zusammenhängen für Mortalität, aber keine Volume-Outcome-Zusammenhänge für Komplikationen bzw. entstehende Morbiditäten für verschiedene andere medizinische Leistungsgruppen in der Literatur finden (z. B. bei Frühgeborenen mit einem Aufnahmegewicht unter 1.250 g, vgl. z. B. IQTIG (2020b)).

Hierfür können verschiedene Gründe bestehen:

1. Es könnte tatsächlich kein Zusammenhang existieren. Diese Annahme kann allerdings vor dem Hintergrund der deutlichen Volume-Outcome-Zusammenhänge mit dem Endpunkt Mortalität als unwahrscheinlich angesehen werden, zum einen, weil die Mortalität als eine ultimative Ausprägung von Komplikationen angesehen werden kann, zum anderen, weil die Datenvalidität der Mortalität als sehr hoch angenommen werden kann.
2. Zudem könnte auch eine fehlende Risikoadjustierung zu diesem Ergebnis beitragen. Die beiden betroffenen Qualitätsindikatoren sind grundsätzlich nicht risikoadjustiert und im Rahmen dieses Auftrags konnte nur relativ grob eine Risikoadjustierung geprüft werden. Ein entscheidender Einfluss auf den jeweiligen Volume-Outcome-Zusammenhang zeigte sich dabei allerdings nicht.

Generell ist zu beachten, dass eine nicht systematisch entwickelte bzw. unzureichende Risikoadjustierung eine relevante Limitation für Volume-Outcome-Analysen darstellen kann. Der

geschätzte Fallzahleffekt ist nur für diejenigen Einflüsse adjustiert, die explizit im Modell vorkommen (ob patienten- oder standortbezogen). Jeder relevante nicht berücksichtigte Faktor vermag den Fallzahleffekt daher grundsätzlich mit zu hoher Unsicherheit zu versehen oder sogar zu verfälschen.

3. Schließlich könnte sich eine einheitliche objektive Abbildung der Komplikationen erheblich schwieriger gestalten als eine Abbildung der Mortalität. So finden sich bei den 1.184 verstorbenen Patientinnen und Patienten in unserer Datenbasis lediglich bei 246 Patientinnen und Patienten intraprozedurale Komplikationen (Anteil 20,7 %) bzw. bei 117 Patientinnen und Patienten Gefäßkomplikationen (Anteil 9,9 %). Beispielsweise könnte vermutet werden, dass entsprechende Komplikationen, insbesondere bei akut durch Komplikationen verstorbenen Patientinnen und Patienten, weniger wahrscheinlich entdeckt und kodiert werden als bei überlebenden Patientinnen und Patienten. Eine solche selektive Kodierung könnte diese bei unterschiedlichen medizinischen Leistungsgruppen beschriebenen Phänomene (Volume-Outcome für Mortalität nachvollziehbar, für Morbidität bzw. für Komplikationen nicht) erklären.

Mit Blick auf die Modellspezifikation wäre es ggf. wünschenswert, mehr Strukturmerkmale der Krankenhausstandorte einbeziehen zu können. Andererseits würde dann ein Fallzahleffekt jenseits der Strukturmerkmale geschätzt. Es würde also ausgeglichen, dass Standorte mit definierten Strukturmerkmal(en) ggf. bessere (oder schlechtere) Ergebnisse aufweisen, und ein Fallzahleffekt bereinigt nach diesen Strukturmerkmalen ausgewiesen. Zusätzlich konnten sekundäre Effekte einer Mindestmengenvereinbarung in diesen Analysen nicht abgebildet werden.

5 Fazit und Ausblick

Der gefundene Volume-Outcome-Zusammenhang mit Blick auf den Endpunkt Sterblichkeit zeigte keinen Schwellenwert und spricht für eine möglichst hohe Mindestmenge, da hier die größten Effekte zu erwarten bzw. bei geringeren Mindestmengen auch vergleichsweise geringe Effekte zu erwarten sind. Mit Blick auf die Endpunkte Intraprozedurale Komplikationen und Gefäßkomplikationen konnten dagegen, wie in der uns bekannten Literatur zum Thema, keine relevanten Volume-Outcome-Effekte oder Schellenwerte beschrieben werden.

Der vorliegende Bericht soll in Kombination mit der Auswertung zu Folgenabschätzungen von Mindestmengen den G-BA in die Lage versetzen, geeignete Mindestmengen auszuwählen. Aus dieser Abwägung ist dann ein konkreter Schwellenwert als Kompromiss zwischen Zentrierung und wohnortnaher, regionaler Versorgung abzuleiten.

Literatur

- Agresti, A (2013): *Categorical Data Analysis*. Third Edition. (Wiley Series in Probability and Statistics). Hoboken, US-NJ: Wiley. ISBN: 978-0-470-46363-5.
- Fahrmeir, L; Kneib, T; Lang, S; Marx, B (2013): *Regression. Models, Methods and Applications*. Berlin [u. a.]: Springer. ISBN: 978-3-642-34332-2.
- G-BA [Gemeinsamer Bundesausschuss] (2020a): Beschluss des Gemeinsamen Bundesausschusses über die Einleitung eines Beratungsverfahrens zur Festlegung einer Mindestmenge für die Durchführung von kathetergestützten Aortenklappenimplantationen (Transcatheter aortic-valve implantation – TAVI) gemäß § 136b Absatz 1 Satz 1 Nummer 2 SGB V. [Stand:] 18.06.2020. Berlin: G-BA. URL: https://www.g-ba.de/downloads/39-261-4351/2020-06-18_Mm-R_Einleitung-Beratungsverfahren_Mindestmenge-TAVI.pdf (abgerufen am: 28.06.2022).
- G-BA [Gemeinsamer Bundesausschuss] (2020b): Vorschlag der Vorsitzenden des Gemeinsamen Bundesausschusses. Festsetzung einer Mindestmenge für die Durchführung von kathetergestützten Aortenklappenimplantationen (Transcatheter aortic-valve implantation / TAVI) von jährlich 150 Eingriffen pro Krankenhausstandort und von jährlich 65 Eingriffen pro Operateur [Anlage zum Beschluss des Gemeinsamen Bundesausschusses über die Einleitung eines Beratungsverfahrens zur Festlegung einer Mindestmenge für die Durchführung von kathetergestützten Aortenklappenimplantationen (Transcatheter aortic-valve implantation – TAVI) gemäß § 136b Absatz 1 Satz 1 Nummer 2 SGB V]. [Stand:] 08.06.2020. Berlin: G-BA. URL: https://www.g-ba.de/downloads/40-268-6675/2020-06-18_Mm-R_Einleitung-Beratungsverfahren_Mindestmenge-TAVI_Anlage.pdf (abgerufen am: 28.06.2022).
- G-BA [Gemeinsamer Bundesausschuss] (2022): Beschluss des Gemeinsamen Bundesausschusses über eine Beauftragung des IQTIG mit der Auswertung von esQS-Daten hinsichtlich Volume-Outcome-Beziehungen bei TAVI. [Stand:] 19.05.2022. Berlin: G-BA. URL: https://www.g-ba.de/downloads/39-261-5441/2022-05-19_IQTIG-Beauftragung_esQS-Daten-TAVI.pdf (abgerufen am: 28.06.2022).
- Gonska, B; Reuter, C; Mörike, J; Rottbauer, W; Buckert, D (2021): Vascular Access Site Complications Do Not Correlate With Large Sheath Diameter in TAVI Procedures With New Generation Devices. *Frontiers in Cardiovascular Medicine* 8:738854. DOI: 10.3389/fcvm.2021.738854.
- Grouven, U; Küchenhoff, H; Schröder, P; Bender, R (2008): Flexible regression models are useful tools to calculate and assess threshold values in the context of minimum provider volumes. *Journal of Clinical Epidemiology* 61(11): 1125-1131. DOI: 10.1016/j.jclinepi.2007.11.020.
- Heller, G (2018): Optionen der Regionalisierung der Frühgeborenenversorgung: Mindestmengen und Ergebnisqualität – Analysen zu Volume-Outcome-Zusammenhängen

und zur Auswirkung einer Einführung von Mindestmengen im Vergleich zu einer Planung über Ergebnisqualität. Kapitel I. Abschnitt 6. In: Dormann, F; Klauber, J; Kühlen, R: *Qualitätsmonitor 2018*. Berlin: MMW – Medizinisch Wissenschaftliche Verlagsgesellschaft, 85-103. ISBN: 978-3-95466-348-4. URL: https://www.gesundheitsstadt-berlin.de/fileadmin/user_upload/Downloads/Qualitätsmonitor_2018_E-Book_V02_anGesundheitsstadt.pdf (abgerufen am: 11.09.2019).

Heller, G; Gutzeit, M; Rauh, J; Cederbaum, J; Rossi, R; Thomas, T; et al. (2022): Reanalyse: Wie hoch ist die optimale Mindestmenge für die Behandlung Frühgeborener mit einem Geburtsgewicht unter 1250 g in Deutschland? *Zeitschrift für Geburtshilfe und Neonatologie* 226(1): 68-69. DOI: 10.1055/a-1701-6686.

IQTIG [Institut für Qualitätssicherung und Transparenz im Gesundheitswesen] (2019): Volumen-Outcome-Beziehungen bei Revisionseingriffen in der Knieendoprothetik. Sonderauswertung des vormaligen Leistungsbereichs Knie-Endoprothesenwechsel und -komponentenwechsel. Stand: 22.12.2019. Berlin: IQTIG. [unveröffentlicht].

IQTIG [Institut für Qualitätssicherung und Transparenz im Gesundheitswesen] (2020a): Aortenklappenchirurgie, isoliert (Kathetergestützt). Qualitätsindikatoren und Kennzahlen. Bundesauswertung zum Erfassungsjahr 2019. Stand: 14.07.2020. Berlin: IQTIG. URL: https://iqtig.org/downloads/auswertung/2019/hchaortkath/QSKH_HCH-AORT-KATH_2019_BUAW_V02_2020-07-14.pdf (abgerufen am: 12.09.2022).

IQTIG [Institut für Qualitätssicherung und Transparenz im Gesundheitswesen] (2020b): Datenauswertung zu Mindestmengen in der Versorgung von Frühgeborenen mit einem Aufnahmegegewicht unter 1.250 g. Abschlussbericht. Stand: 22.06.2020. Berlin: IQTIG. URL: https://iqtig.org/downloads/berichte/2020/IQTIG_Mindestmengen-bei-Fruehgeborenen_Abschlussbericht_2020-06-22_barrierefrei.pdf (abgerufen am: 28.06.2022).

IQTIG [Institut für Qualitätssicherung und Transparenz im Gesundheitswesen] (2022a): Evaluation zu Auswirkungen der Anforderungen der MHI-RL auf die Versorgungsqualität in Deutschland. Evaluationsbericht. Stand: 21.01.2022. Berlin: IQTIG. URL: https://iqtig.org/downloads/berichte/2022/IQTIG_Auswirkungen-der-MHI-RL_Evaluationsbericht_2022-01-21-barrierefrei.pdf (abgerufen am: 06.09.2022).

IQTIG [Institut für Qualitätssicherung und Transparenz im Gesundheitswesen] (2022b): Folgenabschätzungen zu Mindestmengen Kathetergestützte Aortenklappen-implantationen (*Transcatheter aortic-valve implantation* – TAVI). Datenanalysen im Rahmen der Beratungen zu Mindestmengen. Stand: 29.09.2022. Berlin: IQTIG. [unveröffentlicht].

IQWiG [Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen] (2021): Zusammenhang zwischen Leistungsmenge und Qualität des Behandlungsergebnisses bei der Durchführung von kathetergestützten Aortenklappenimplantationen (*Transcatheter aortic-valve implantation* – TAVI). Rapid Report. Stand: 30.06.2021. (IQWiG-Berichte, Nr. 1153). Köln: IQWiG. V20-04. URL: <https://www.iqwig.de/download/v20->

04 zusammenhang-zwischen-leistungsmenge-und-qualitaet-des-behandlungsergebnisses-bei-der-durchfuehrung-von-tavi_rapid-report_v1-0.pdf (abgerufen am: 10.05.2022).

- Larsen, K; Petersen, JH; Budtz-Jørgensen, E; Endahl, L (2000): Interpreting Parameters in the Logistic Regression Model with Random Effects. *Biometrics* 56(3): 909-914. DOI: 10.1111/j.0006-341x.2000.00909.x.
- Mach, M; Okutucu, S; Kerbel, T; Arjomand, A; Fatihoglu, SG; Werner, P; et al. (2021): Vascular Complications in TAVR: Incidence, Clinical Impact, and Management. *Journal of Clinical Medicine* 10(21): 5046. DOI: 10.3390/jcm10215046.
- Marra, G; Wood, SN (2012): Coverage Properties of Confidence Intervals for Generalized Additive Model Components. *Scandinavian Journal of Statistics* 39(1): 53-74. DOI: 10.1111/j.1467-9469.2011.00760.x.
- Potluri, SP; Hamandi, M; Basra, SS; Shinn, KV; Tabachnick, D; Vasudevan, A; et al. (2020): Comparison of Frequency of Vascular Complications With Ultrasound-Guided Versus Fluoroscopic Roadmap-Guided Femoral Arterial Access in Patients Who Underwent Transcatheter Aortic Valve Implantation. *The American Journal of Cardiology* 132: 93-99. DOI: 10.1016/j.amjcard.2020.07.013.
- R Core Team (2022): R: A Language and Environment for Statistical Computing [*Open Source Software*]. R version 4.2.1 (Funny-Looking Kid). Vienna, AT: R Foundation for Statistical Computing. URL: <https://cran.r-project.org/> [Source Code for all Platforms > Source code of older versions of R is available here >R-4/ > R-4.2.1.tar.gz] (abgerufen am: 02.09.2022).
- Rothman, KJ; Greenland, S; Lash, TL (2008): Modern Epidemiology. Third Edition. Philadelphia, US-PA [u. a.]: Lippincott Williams and Wilkins. ISBN: 978-1-4511-9005-2.
- Scarsini, R; De Maria, GL; Joseph, J; Fan, L; Cahill, TJ; Kotronias, RA; et al. (2019): Impact of Complications During Transfemoral Transcatheter Aortic Valve Replacement: How Can They Be Avoided and Managed? *Journal of the American Heart Association* 8(18): e013801. DOI: 10.1161/JAHA.119.013801.
- Ueshima, D; Barioli, A; Nai Fovino, L; D'Amico, G; Fabris, T; Brener, SJ; et al. (2020): The impact of pre-existing peripheral artery disease on transcatheter aortic valve implantation outcomes: A systematic review and meta-analysis. *Catheterization and Cardiovascular Interventions* 95(5): 993-1000. DOI: 10.1002/ccd.28335.
- Varaganam, M; Hutchings, A; Black, N (2015): Relationship between Patient-reported Outcomes of Elective Surgery and Hospital and Consultant Volume. *Medical Care* 53(4): 310-316. DOI: 10.1097/MLR.0000000000000318.
- Wetzel, H (2006): Mindestmengen zur Qualitätssicherung: Konzeptionelle und methodische Überlegungen zur Festlegung und Evaluation von Fallzahlgrenzwerten für die klinische Versorgung. *Zeitschrift für ärztliche Fortbildung und Qualität im Gesundheitswesen* 100(2): 99-106.

Wood, SN (2003): Thin plate regression splines. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)* 65(1): 95-114. DOI: 10.1111/1467-9868.00374.

Wood, SN (2013): On p -values for smooth components of an extended generalized additive model. *Biometrika* 100(1): 221-228. DOI: 10.1093/biomet/ass048.

Wood, SN (2017): Generalized Additive Models. An Introduction with R. Second Edition. (Texts in Statistical Science). Boca Raton, US-FL [u. a.]: Chapman & Hall/CRC. ISBN: 978-1-4987-2833-1.

Wood, SN (2022): mgcv: Mixed GAM Computation Vehicle with Automatic Smoothness Estimation [*Open Source Software*]. R package version 1.8-40. Published: 29.03.2022. Vienna, AT: R Foundation for Statistical Computing. URL: <https://CRAN.R-project.org/package=mgcv> [Package source > mgcv_1.8-40.tar.gz] (abgerufen am: 02.09.2022).