

Modellsteckbriefe für QS GYN-OP

Auswertungsjahr 2024

Stand: 29. November 2024

Dieses Dokument enthält Hintergrundinformationen zu den im QS-Verfahren *Gynäkologische Operationen (QS GYN-OP)* verwendeten Risikoadjustierungsmodellen. Für Erläuterungen zu den dargestellten Informationen wird auf die Leseanleitung zu den Modellsteckbriefen verwiesen.

Inhaltsverzeichnis

1	QI 51906: Verhältnis der beobachteten zur erwarteten Rate (O/E) an Organverletzungen bei laparoskopischer Operation	2
2	Leseanleitung zu den Modellsteckbriefen	7
	Impressum	10

1 QI 51906: Verhältnis der beobachteten zur erwarteten Rate (O/E) an Organverletzungen bei laparoskopischer Operation

Grundgesamtheit Alle Patientinnen mit ausschließlich laparoskopischem OP-Zugang oder mit ausschließlich laparoskopischem und abdominalem OP-Zugang bei der ersten Operation. Ausgeschlossen werden Patientinnen, bei denen eine Uterusexstirpation mit einer ausgedehnten retroperitonealen Präparation und einseitiger oder beidseitiger Salpingoovarektomie erfolgte (OPS 5-683.6* und 5-683.7*).

Zähler Patientinnen mit mindestens einer Organverletzung

1.1 Datenbasis und Modellentwicklung

Die Modellschätzung basiert auf der Grundgesamtheit der Erfassungsjahre 2021 und 2022.

Anzahl Fälle in der Modellschätzung	Davon mit Zählerereignis	Anteil
259.827	2.253	0,87 %

1.1.1 Leistungserbringereffekte

Das Modell wurde unter Berücksichtigung von Leistungserbringereffekten als *zufällige Effekte* geschätzt. Die geschätzte Standardabweichung der Leistungserbringereffekte beträgt $\hat{\tau} = 0,471$. Das genaue Vorgehen wird im Begleitdokument *Leistungserbringereffekte bei der Risikoadjustierung* beschrieben, welches u.a. mit den endgültigen Rechenregeln zum Auswertungsjahr 2022 ([hier](#) herunterladen) veröffentlicht wurde.

1.1.2 Normative Setzung von Koeffizienten

Das Modell enthält Koeffizienten, die normativ gesetzt wurden: Das Modell für die Risikofaktoren wurde zusammen mit zufälligen Effekten für Leistungserbringer geschätzt. Diese wurden für die Vorhersage der fallbasierten Risiken auf Null gesetzt und die Regressionskonstante entsprechend so angepasst, dass das Modell auf dem Entwicklungsdatensatz gesamt-kalibriert ist ($O/E = 1$).

1.1.3 Veränderungen zum Vorjahr

Das Vorjahresmodell wurde auf den Schätzdaten neu gefittet und leicht angepasst. Neue Einflussvariablen kamen nicht hinzu.

1.1.4 Erklärung zu „in sample“-Angaben

Mit „in sample“ bezeichnete Angaben in diesem Modellsteckbrief basieren auf dem zur Modellentwicklung genutzten Datensatz und auf dem geschätzten Modell vor normativer Setzung von Koeffizienten. Dieses kann daher vom für die Bundesauswertung eingesetzten Modell im Abschnitt „Risikomodell des Qualitätsindikators“ leicht abweichen. Die geschätzten

Leistungserbringereffekte werden außerdem für die Schätzung der „in sample“-Risiken berücksichtigt. Siehe dazu die anderen Informationen in diesem Abschnitt.

1.1.5 Weiterführende Informationen

Detaillierte Informationen über die Datenerhebung und die Berechnung der Qualitätsindikatoren entnehmen Sie bitte den über die [Verfahrensübersicht des IQTIG](#) zugänglichen Dokumenten zu Spezifikation und Rechenregeln. Die [Bundesauswertungen des IQTIG](#) liefern im Kapitel *Basisauswertung* zudem beschreibende Statistiken zur Grundgesamtheit des QS-Verfahrens.

1.2 Risikomodell des Qualitätsindikators

Risikoeffizienten aus der logistischen Regression. Die Referenzwahrscheinlichkeit beträgt 0,48 % (Odds: 0,0048).

Risikofaktor	Regressionskoeffizient	Std.-Fehler	Z-Wert	Odds-Ratio (mit 95 %-Vertrauensbereich)
Konstante	-5,304211	0,059976	-89,10	
Alter bei nicht vorliegender Endometriose (linear, in Jahren)	0,019376	0,001657	11,69	
ASA-Klassifikation				
▪ ASA-Klassifikation 2	0,147254	0,053985	2,73	1,159 (1,042 – 1,288)
▪ ASA-Klassifikation 3	0,309656	0,078678	3,94	1,363 (1,168 – 1,590)
▪ ASA-Klassifikation 4 oder 5	0,719994	0,315235	2,28	2,054 (1,108 – 3,811)
Adhäsiolyse	0,498221	0,073432	6,79	1,646 (1,425 – 1,901)
Endometriose	0,332214	0,066455	5,00	1,394 (1,224 – 1,588)
Exzision	-0,494220	0,066883	-7,39	0,610 (0,535 – 0,695)
Voroperation im OP-Gebiet	0,324449	0,059054	5,49	1,383 (1,232 – 1,553)
Gynäkologisches Malignom	0,358639	0,068975	5,20	1,431 (1,250 – 1,639)
Interaktion: Voroperation im OP-Gebiet mit Adhäsiolyse	0,225862	0,091152	2,48	1,253 (1,048 – 1,499)

1.2.1 Odds-Ratios

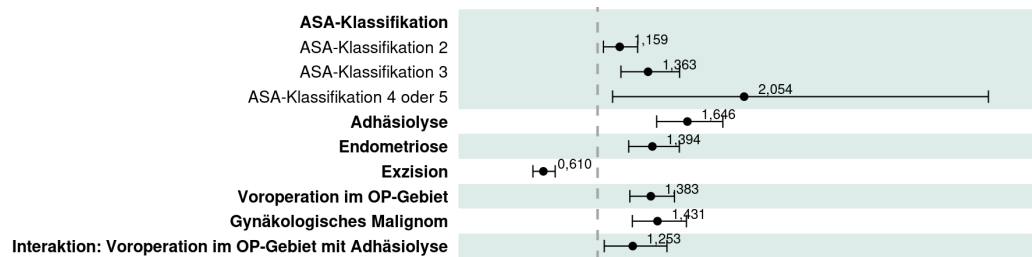


Abbildung 1: Odds-Ratios (grafische Darstellung).

1.2.2 Einfluss stetiger Variablen

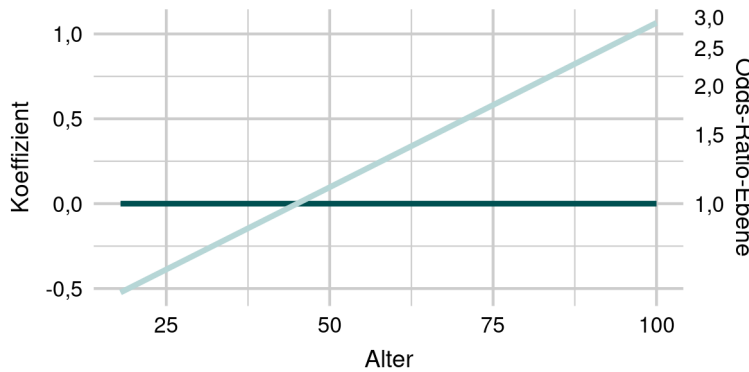


Abbildung 2: Einfluss der stetigen Variable Alter (dunkelgrün: „bei vorliegender Endometriose“; hellgrün: „bei nicht vorliegender Endometriose“).

1.2.3 Verteilung der Risiken (Ergebnisse der Bundesauswertung)

Die folgende Grafik und Tabelle zeigt die bundesweite Verteilung der geschätzten Risiken für das Auswertungsjahr 2024, basierend auf dem oben dargestellten Risikomodell.

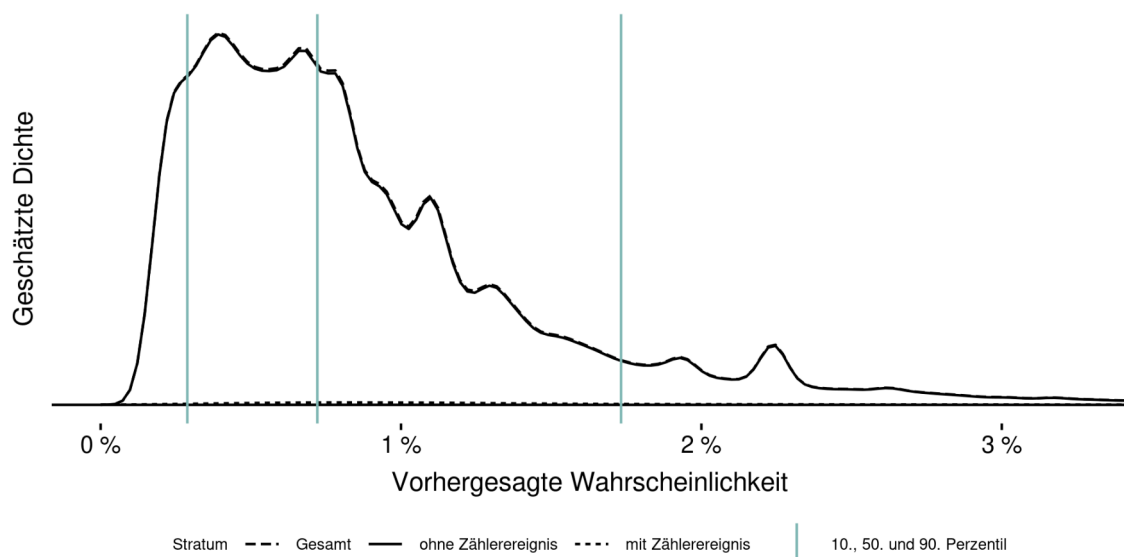


Abbildung 3: Dichtediagramm zur Verteilung der Risiken (Ergebnisse der Bundesauswertung).

Statistiken zur Verteilung der Risiken (Ergebnisse der Bundesauswertung).

Geschätzte Risiken			
Ereignis	Anzahl Fälle	Mittelwert	Median
mit Zählerereignis	1.257	1,30 %	1,07 %
ohne Zählerereignis	133.130	0,89 %	0,72 %
Gesamt	134.387	0,89 %	0,72 %

1.3 Eigenschaften des geschätzten Modells

1.3.1 Kennzahlen zur Vorhersagegüte

	AUC	Brier-Score	Nagelkerkes Pseudo-R ²
in sample	0,749	0,009	0,073

1.3.2 Kalibrierung (in sample)

Kalibrierungstabelle nach Risiko-Dezilen (in sample).

Dezil	Erwartet	Beobachtet	Teststatistik	Kalibrierungsdiagramm
1 [0,00042 ... 0,0024]	0,18 %	0,11 %	6,94	
2 (0,0024 ... 0,00334]	0,29 %	0,19 %	8,24	
3 (0,00334 ... 0,0043]	0,38 %	0,25 %	11,84	
4 (0,0043 ... 0,00532]	0,48 %	0,35 %	9,77	
5 (0,00532 ... 0,00646]	0,59 %	0,50 %	3,71	
6 (0,00646 ... 0,00784]	0,71 %	0,65 %	1,59	
7 (0,00784 ... 0,00963]	0,87 %	0,90 %	0,31	
8 (0,00963 ... 0,0123]	1,09 %	1,11 %	0,12	
9 (0,0123 ... 0,0175]	1,46 %	1,69 %	9,85	
10 (0,0175 ... 0,158]	2,63 %	2,93 %	9,18	

In den Spalten 'Erwartet' und 'Beobachtet' sind die jeweiligen Mittelwerte dargestellt. Ein Aufsummieren der Spalte 'Teststatistik' ergibt die Teststatistik nach Hosmer und Lemeshow zur Modellkalibrierung.

2 Leseanleitung zu den Modellsteckbriefen

Zu einigen Risikoadjustierungsmodellen veröffentlicht das IQTIG ergänzend zu den Angaben in der QIDB sogenannte Modellsteckbriefe. Sie enthalten detaillierte Informationen über die Entstehung der Modelle sowie ihre statistischen Eigenschaften. Für eine Einführung in die Risikoadjustierung von Qualitätsindikatoren siehe das entsprechende Kapitel in der Leseanleitung zur Bundesauswertung.

Die Modellsteckbriefe gliedern sich in drei Abschnitte: „Datenbasis und Modellentwicklung“, „Risikomodell des Qualitätsindikators“ und „Eigenschaften des geschätzten Modells“.

2.1 Datenbasis und Modellentwicklung

Dieser Abschnitt enthält allgemeine Informationen über das Zustandekommen des Modells, die zugrundeliegenden Daten und eventuelle Veränderungen zum Vorjahr.

2.2 Risikomodell des Qualitätsindikators

Dieser Abschnitt enthält eine Beschreibung des für die Berechnung des Qualitätsindikators / der Qualitätsindikatoren genutzten Modells. Neben der bereits in der QIDB enthaltenen Tabelle der Koeffizienten der Risikofaktoren werden die Einflüsse der diskreten und stetigen Risikofaktoren graphisch dargestellt.

Schließlich wird die empirische Verteilung der Risiken mittels Dichteschätzer dargestellt. Dies vermittelt einen Eindruck darüber, wie sich die Grundgesamtheit des Modells in Hochrisiko- und Niedrigrisikofälle verteilt. Die Verteilung wird stratifiziert für die Fälle mit und ohne Zählerereignisse dargestellt (grob- bzw. feingestrichelte Linie). Die Darstellung ist so gewählt, dass die Fläche unter der grobgestrichelten, feingestrichelten bzw. durchgezogenen Linie proportional ist zur Anzahl an Fällen mit bzw. ohne Zählerereignissen bzw. zur Gesamtanzahl an Fällen. Vertikale Linien markieren das 10., 50. (Median) und 90. Perzentil der Gesamtverteilung.

Zusätzlich zur grafischen Darstellung werden hier Maße für die Lage der geschätzten Risiken tabellarisch berichtet, ebenfalls stratifiziert für die Fälle mit und ohne Zählerereignisse. Dies vermittelt, wie stark sich die Modellvorhersagen im Mittel und im Median zwischen Fällen mit und ohne Zählerereignis unterscheiden. Die Größe dieses Unterschieds wird auch oft als Maß für die Diskriminationsfähigkeit eines Modells interpretiert. Dazu muss allerdings auch die Gesamtprävalenz des Zählerereignisses berücksichtigt werden.

2.3 Eigenschaften des geschätzten Modells

In diesem Abschnitt werden statistische Eigenschaften des Risikoadjustierungsmodells dargestellt. Die in diesem Abschnitt dargestellten Modelleigenschaften beschreiben das Verhältnis des Modells mit der empirischen Grundgesamtheit des Modells. Falls das Risikoadjustierungsmodell normativ gesetzte Koeffizienten enthält, so werden die entsprechenden Setzungen in diesem Abschnitt nicht berücksichtigt (siehe dazu auch den Unterabschnitt „Erklärung zu *in sample*-Angaben“ im Modellsteckbrief). Dies ist darin begründet, dass die normative Setzung datenunabhängig geschieht.

Die statistischen Eigenschaften lassen sich dabei grundsätzlich *in sample* und *out of sample* berechnen. *In sample* bedeutet, dass die jeweilige Eigenschaft auf Grundlage des Datensatzes berechnet wird, der auch für die Schätzung des Modells selbst genutzt wurde. *Out of sample* bedeutet, dass die jeweilige Modelleigenschaft auf einem anderen Datensatz (z. B. den Daten aus einem anderen Erfassungsjahr) berechnet wurde. Jeder Modellsteckbrief enthält mindestens Informationen zu den *in-sample*-Eigenschaften des Modells. Darüber hinaus sind die Modelleigenschaften unter Umständen auch für einen oder mehrere *out-of-sample*-Datensätze dargestellt. Unterschiede zwischen *in-sample*- und *out-of-sample*-Eigenschaften deuten auf Überanpassung hin.

2.3.1 Kennzahlen

Bei den drei dargestellten Kennzahlen AUC, Brier-Score und Nagelkerkes Pseudo- R^2 handelt es sich um weitverbreitete und wichtige Maßzahlen für klinische Vorhersagemodelle. Für die genaue Definition wird z.B. auf das Buch *Clinical Prediction Models* von Ewout Steyerberg (2. Auflage, Springer 2019) verwiesen.

Die **AUC** (*area under the ROC-curve*, auch *c-statistics* bzw. Konkordanzstatistik) ist ein Maß für die Diskriminationsfähigkeit des Modells. Der Wert liegt zwischen $\frac{1}{2}$ und 1. Der Wert $\frac{1}{2}$ tritt nur dann auf, wenn sämtlichen Fällen das gleiche Risiko zugewiesen wird.

Der **Brier-Score** beschreibt, wie gut das Modell das Eintreten von Zählerereignissen vorher-sagen kann. Der Wert liegt in der Regel zwischen 0 und $\frac{1}{4}$. Der Wert $\frac{1}{4}$ tritt nur dann auf, wenn sämtlichen Fällen das Risiko $\frac{1}{2}$ zugewiesen wird.

Nagelkerkes Pseudo- R^2 ist ein Maß für den Anteil an der Variabilität oder Unsicherheit über das Zählerereignis, den das Modell erklärt. Der Wert liegt zwischen 0 und 1. Der Wert 0 tritt nur dann auf, wenn sämtlichen Fällen das gleiche Risiko zugewiesen wird.

Es handelt sich bei diesen Kennzahlen nicht um Gütekriterien für Risikoadjustierungsmodelle: Perfekte Werte (also ein AUC von 1, ein Brier-Score von 0 und ein Pseudo- R^2 von 1) sind bei Risikoadjustierungsmodellen nicht möglich und auch nicht wünschenswert, da solche

perfekten Werte nur dann auftreten können, wenn sich aus den Risikofaktoren sicher vorher-sagen lässt, ob ein interessierende Zählerereignis auftritt oder nicht. In der externen Quali-tätssicherung werden jedoch Ereignisse betrachtet, bei denen die Leistungserbringer einen starken Einfluss haben. „Schlechte“ Werte der Kennzahlen deuten also auf einen schwachen Einfluss der berücksichtigten Risikofaktoren auf das Zählerereignis hin. Solche Werte treten daher unter anderem dann auf, wenn es, wie beispielsweise bei vielen Prozessindikatoren, nur wenige patientenseitige Risikofaktoren gibt, die berücksichtigt werden sollen.

2.3.2 Kalibrierung

Die Kalibrierung beschreibt, inwiefern die vorhergesagten Risiken mit beobachteten Häufig-keiten in den Daten zusammen passen: Unter allen Fällen mit einem Risiko von x % wird er-wartet, dass der Anteil von Fällen mit beobachtetem Zählerereignis in etwa x % beträgt.

Die Kalibrierung wird gelegentlich mit dem Hosmer-Lemeshow-Test überprüft. Dieser Test ist jedoch umstritten: Bei großen Fallzahlen kann der Test sehr kleine p -Werte liefern, auch wenn die Verletzung der Kalibrierung nicht relevant erscheint. Zudem wird in der wissen-schaftlichen Literatur diskutiert, welche Anzahl an Freiheitsgraden man für die Verteilung der Teststatistik annehmen sollte. Aussagekräftiger sind die Kalibrierungstabelle sowie das Ka-librierungsdiagramm, aus denen ersichtlich wird, in welchen Bereichen die Kalibrierung ver-letzt wird. In der Kalibrierungstabelle werden die Fälle nach ihrem vorhergesagten Risiko in zehn Gruppen eingeteilt und geprüft, wie weit innerhalb dieser Gruppen die vorhergesagte und beobachtete Anzahl an Zählerereignissen auseinanderliegen. Diese Gruppierung liegt auch dem Hosmer-Lemeshow-Test zugrunde. Für jede Gruppe ist daher der Beitrag zur Hos-mer-Lemeshow-Statistik dargestellt. Im Kalibrierungsdiagramm ist die Kalibrierung ohne Diskretisierung mit Hilfe eines Glätters dargestellt (siehe z.B. *Clinical Prediction Models* von Ewout Steyerberg, 2. Auflage, Springer 2019). Ergänzt wird die Darstellung durch die Werte für die Dezile aus der Kalibrierungstabelle, die als Punkte eingetragen sind.

Impressum

HERAUSGEBER

IQTIG – Institut für Qualitätssicherung
und Transparenz im Gesundheitswesen
Katharina-Heinroth-Ufer 1
10787 Berlin

Telefon: (030) 58 58 26-0

info@iqtig.org

iqtig.org