

# Umstellung der Methodik für die Ermittlung rechnerischer Auffälligkeiten und verteilungs- abhängiger Referenzbereiche in *QS CHE, QS GYN-OP* und *QS MC*

**Begleitdokument zu den endgültigen Rechenregeln  
Auswertungsjahr 2025**

## Inhaltsverzeichnis

Tabellenverzeichnis .....	2
1 Hintergrund .....	3
2 Ermittlung rechnerischer Auffälligkeit .....	4
3 Ermittlung verteilungsabhängiger Referenzbereiche .....	7
3.1 Ratenbasierte Qualitätsindikatoren .....	7
3.2 Risikoadjustierte O/E-Qualitätsindikatoren .....	8
Literatur .....	9
Impressum .....	10

## Tabellenverzeichnis

Tabelle 1: Ermittlung rechnerischer Auffälligkeit für ratenbasierte Qualitätsindikatoren .....	4
Tabelle 2: Ermittlung rechnerischer Auffälligkeit für risikoadjustierte O/E- Qualitätsindikatoren .....	4
Tabelle 3: Auszug aus der Ergebnistabelle für ein hypothetisches Leistungserbringerergebnis ...	6

# 1 Hintergrund

Mit dem Auswertungsjahr 2025 kommt in drei Qualitätssicherungsverfahren (QS-Verfahren) eine neue statistische Methodik zur Berechnung der rechnerischen Auffälligkeiten für Qualitätsindikatoren zum Einsatz.<sup>1</sup> Es handelt sich um die QS-Verfahren

- Cholezystektomie (QS CHE),
- Gynäkologische Operationen (QS GYN-OP) sowie
- Mammachirurgie (QS MC).

Bereits zum Auswertungsjahr 2022 war die neue Methodik für das QS-Verfahren *Perkutane Koronarintervention (PCI)* und *Koronarangiographie (QS PCI)* eingeführt worden. Dieses Dokument dient dazu, über die Umstellung in weiteren QS-Verfahren zu informieren und die statistischen Details der Berechnung darzulegen. Weitere, leicht verständliche Informations-Materialien zur neuen Berechnung der rechnerischen Auffälligkeiten sind auf der Webseite des IQTIG öffentlich abrufbar.<sup>2</sup>

Die Umstellung und die daraus resultierende Festlegung der statistischen Methodik für die rechnerische Auffälligkeit sind dadurch motiviert, dass (s. Hengelbrock et al. 2022, IQTIG 2024):

- die bisherige Methodik die statistische Unsicherheit über die zugrundeliegende Qualität von Leistungserbringern nicht angemessen berücksichtigt. Dies entspricht nicht den wissenschaftlichen und internationalen Standards. Zudem führt es dazu, dass alleine aus stochastischen Gründen bei Leistungserbringern mit geringer Fallzahl viele Auffälligkeiten auftreten.
- die bisherige Methodik für die Berechnung perzentilbasierter Referenzwerte nicht gewährleistet, dass der durch das Perzentil vorgegebene Anteil an Standortergebnissen rechnerisch auffällig wird (sondern in der Regel ein höherer Anteil).

---

<sup>1</sup> Für eine Beschreibung der bisherigen Methodik für die anderen QS-Verfahren siehe die Erläuterung zur *rechnerischen Auffälligkeit* im Glossar der Leseanleitung zur Bundesauswertung.

<sup>2</sup> <https://iqtig.org/das-iqtig/wie-wir-arbeiten/grundlagen/biometrische-grundlagen/biometrische-methodik-zur-auffaelligkeitseinstufung/>

## 2 Ermittlung rechnerischer Auffälligkeit

Für Qualitätsindikatoren in den auf die neue Methodik umgestellten QS-Verfahren wird die rechnerische Auffälligkeit von Leistungserbringerergebnissen gemäß der in Hengelbrock et al. (2022) publizierten statistischen Methodik im Rahmen einer analytischen Zielsetzung ermittelt (vgl. auch IQTIG 2024). Das bedeutet, dass die statistische Unsicherheit der beobachteten Indikatorergebnisse bei ihrer rechnerischen Einstufung berücksichtigt wird. Die rechnerische Auffälligkeitseinstufung erfolgt anhand des Vergleichs der A-posteriori-Wahrscheinlichkeit für die Einhaltung des Referenzbereichs mit einem vorab festgelegten Schwellenwert  $\alpha$  (ähnlich einem Signifikanzniveau). Die A-posteriori-Wahrscheinlichkeit wird für ratenbasierte Qualitätsindikatoren dabei anhand eines Beta-Binomial-Modells (siehe Tabelle 1) und für risikoadjustierte O/E-Qualitätsindikatoren<sup>3</sup> anhand eines Poisson-Gamma-Modells (siehe Tabelle 2) berechnet. Dabei bezeichnet  $\theta$  den zugrunde liegenden Parameter des Leistungserbringers für den Indikator,  $J$  seine Fallzahl,  $o$  seine Anzahl interessierender Ereignisse und  $e$  seine erwartete Anzahl interessierender Ereignisse.

Tabelle 1: Ermittlung rechnerischer Auffälligkeit für ratenbasierte Qualitätsindikatoren

Statistisches Modell	Beta-Binomial-Modell, mit A-posteriori-Verteilung $\theta \mid o, J \sim \text{Beta}\left(\frac{1}{2} + o, \frac{1}{2} + J - o\right)$
Kriterium für die rechnerische Auffälligkeit	$P(\theta \in \text{Referenzbereich} \mid o, J) \leq \alpha$
Schwellenwert	$\alpha$
Pseudocode <sup>4</sup>	<code>compute_rate_dc(o, J, R, alternative = "greater") &lt;= alpha</code>

Tabelle 2: Ermittlung rechnerischer Auffälligkeit für risikoadjustierte O/E-Qualitätsindikatoren

Statistisches Modell	Poisson-Gamma-Modell mit A-posteriori-Verteilung $\theta \mid o, e \sim \text{Gamma}\left(\frac{1}{2} + o, e\right)$
Kriterium für die rechnerische Auffälligkeit	$P(\theta \in \text{Referenzbereich} \mid o, J) \leq \alpha$
Schwellenwert	$\alpha$
Pseudocode <sup>4</sup>	<code>compute_oe_dc(o, e, R, alternative = "greater") &lt;= alpha</code>

<sup>3</sup> O/E bezeichnet das Verhältnis der beobachteten Rate (*observed*, O) eines interessierenden Ereignisses zur erwarteten Rate (*expected*, E).

<sup>4</sup> Durch das Argument `alternative` wird dabei die Richtung des Referenzbereichs spezifiziert ("`greater`" für Referenzbereich  $[0; R]$  und "`less`" für Referenzbereich  $[R; 1]$  bzw.  $[R; \infty)$ ).

Je nach Polung des Indikators ist der Referenzbereich bei Referenzwert  $R$  entweder  $[0; R]$  bzw.  $[R; 1]$  (ratenbasierter Indikator) oder  $[0; R]$  bzw.  $[R; \infty)$  (risikoadjustierter O/E-Indikator).

Das in Tabelle 1 und Tabelle 2 dargestellte Kriterium für die rechnerische Auffälligkeit zeigt: Das Ergebnis eines Leistungserbringers wird dann rechnerisch auffällig, wenn, nach Kenntnis des Indikatorergebnisses, die Wahrscheinlichkeit, den Referenzbereich einzuhalten, so klein ist, dass sie den Schwellenwert  $\alpha$  erreicht oder unterschreitet. Das bedeutet, es wird dann auffällig, wenn es ausreichend Evidenz dafür gibt, dass der Referenzbereich nicht eingehalten wird. Bei einem kleineren Schwellenwert  $\alpha$  fordert die Einstufungsmethodik somit gravierendere Evidenz für das Verletzen des Referenzbereichs (z.B. die Beobachtung stärkerer Abweichungen vom Referenzbereich oder von mehr Fällen). Der Schwellenwert  $\alpha$  wird vorab für alle Indikatoren festgelegt.

Aus Gründen der Nachvollziehbarkeit und methodischen Transparenz sind die Funktionen `compute_rate_dc` und `compute_oe_dc` im R-Paket *iqtigbdt* exemplarisch implementiert.<sup>5</sup> Shiny-Apps zur Erstellung von Funnelplots und der Analyse von theoretischer Sensitivität und Spezifität der Auffälligkeitseinstufung sind im R-Paket implementiert und auch verfügbar unter <https://iqtig.shinyapps.io/funnelplot/> bzw. [https://iqtig.shinyapps.io/sensitivitaet\\_spezifitaet/](https://iqtig.shinyapps.io/sensitivitaet_spezifitaet/).

Bei der Interpretation von Funnel-Plots ist zu beachten, dass die rechnerische Auffälligkeit in den auf die neue Methodik umgestellten QS-Verfahren nicht daraus abgeleitet werden kann, ob ein QI-Ergebnis (als Zähler / Nenner) außerhalb des Referenzbereichs liegt oder nicht. Stattdessen ist eine rechnerische Auffälligkeit daran erkennbar, dass ein QI-Ergebnis oberhalb (bzw. unterhalb, je nach Polung des QI) eines fallzahlabhängigen Grenzwerts liegt. Diese zeigt an, für welche QI-Ergebnisse unter Berücksichtigung statistischer Unsicherheit genügend datenbasierte Evidenz für eine Klassifikation als „rechnerisch auffällig“ vorliegt.

<sup>5</sup> <https://github.com/iqtigorg/iqtigbdt>. Für einen ratenbasierten Indikator mit Referenzbereich  $\leq 15\%$  und Schwellenwert  $\alpha = 2,5\%$  und ein Leistungserbringerergebnis von  $J = 10, o = 4$ , kann die Auffälligkeitseinstufung beispielsweise berechnet werden über den Aufruf der R-Funktion:

```
iqtigbdt::compute_rate_dc(4, 10, 0.15) <= 0.025
```

```
## [1] TRUE
```

(wobei TRUE bedeutet, dass das Ergebnis *rechnerisch auffällig* ist).

### Beispiel

Ob das Ergebnis eines Leistungserbringers rechnerisch auffällig nach dieser neuen Methodik ist, ist anhand der Ergebnisliste erkennbar. Für einen ratebasierten Indikator mit Referenzbereich  $\leq 15\%$  und Schwellenwert  $\alpha = 2,5\%$  ist ein Ergebnis von  $o = 4$ ,  $J = 10$  beispielsweise rechnerisch auffällig. Dies ist an dem Wert TRUE der Spalte BEWERTUNG\_RECH erkennbar. Dies folgt daraus, dass die A-posteriori-Wahrscheinlichkeit kleiner gleich dem Schwellenwert von  $2,5\%$  ist.

Tabelle 3: Auszug aus der Ergebnistabelle für ein hypothetisches Leistungserbringerergebnis.

QI-TYP	ZAehler	NENNER	ERGEBNIS	VERTRBER_UNTEN	VERTRBER_OBEN	BEWERTUNG_RECH	SIGNIFIKANZWERT	REFWERT1	REFOPE-RATOR
Rate	4	10	40	15,30671011	69,63205144	TRUE	0,02317645	15	$\leq$

### 3 Ermittlung verteilungsabhängiger Referenzbereiche

Verteilungsabhängige Referenzbereiche werden für die auf die neue Methodik umgestellten QS-Verfahren anhand aller Leistungserbringer ermittelt, welche im betrachteten Erfassungsjahr mindestens einen Fall in der Grundgesamtheit des jeweiligen Qualitätsindikators aufweisen. Anhand der A-posteriori-Wahrscheinlichkeiten der einzelnen Leistungserbringer, des vorgegebenen Schwellenwertes ( $\alpha$ ) sowie des vorgegebenen Perzentils wird der Referenzbereich so bestimmt, dass der Anteil  $q$  an Indikatorergebnissen als rechnerisch auffällig eingestuft wird, der durch das Perzentil definiert ist.<sup>6</sup> Dabei wird die in Kapitel 2 beschriebene Methodik sowie der Schwellenwert  $\alpha$  des Indikators verwendet. Es ist nicht immer möglich, dass genau die durch das Perzentil vorgegebene Anzahl an Leistungserbringern  $q \cdot I$  rechnerisch auffällig wird. In dem Fall wird die beste Näherung verwendet.

#### 3.1 Ratenbasierte Qualitätsindikatoren

Für ratenbasierte Qualitätsindikatoren seien  $P(\theta_i \leq R | o_i, J_i)$  die A-posteriori-Wahrscheinlichkeiten für die Einhaltung des Referenzbereichs  $[0; R]$  von Leistungserbringer  $i, i = 1, \dots, I$ , mit  $o_i$  als Anzahl interessierender Ereignisse und Fallzahl  $J_i$ . Der Referenzwert  $R$  wird so gewählt, dass die absolute Differenz zwischen der Anzahl an rechnerisch auffälligen Indikatorergebnissen und dem durch das Perzentil definierten Anzahl (im Folgenden  $q \cdot I$ ) minimiert wird:

$$R = \arg \min_R \left| \sum_{i=1}^I [P(\theta_i \leq R | o_i, J_i) \leq \alpha] - q \cdot I \right|,$$

bzw. mit  $P(\theta_i \geq R | o_i, J_i)$  für Indikatoren mit Referenzbereich  $[R; 1]$ .

#### Beispiel

Liegen bei einem Indikator mit  $q = 0.1$  (also einem vorab definierten Anteil an rechnerischen Auffälligkeiten von 10%) genau  $I = 11$  Leistungserbringerergebnisse mit  $J_i > 0$  vor, und sind die Ergebnisse als  $(o_i, J_i)$  Paare eindeutig, so wird der Referenzbereich nach oben beschriebener Methodik so gewählt, dass genau ein Leistungserbringerergebnis als rechnerisch auffällig klassifiziert wird. Dies entspricht einem Anteil von  $\frac{1}{11} = 9,1\%$ . Eine bessere Näherung an den geforderten Anteil von 10% ist nicht möglich.

<sup>6</sup> Auf diese Weise ermittelte Referenzbereiche können daher weder exakt noch annäherungsweise interpretiert werden als Perzentil der Verteilung der beobachteten Leistungserbringerergebnisse. Sie resultieren einzig aus der Anforderung, bei gleichzeitiger Berücksichtigung statistischer Unsicherheit den definierten Anteil an rechnerischen auffälligen Ergebnissen zu erzielen.

### 3.2 Risikoadjustierte O/E-Qualitätsindikatoren

Für risikoadjustierte O/E-Indikatoren erfolgt die Ermittlung analog anhand der Anzahl interessierender Ereignisse  $e_i$  (statt  $J_i$ ). Zusätzlich gilt bei diesen Indikatoren, dass für den Referenzwert ein unteres Limit von 1 festgelegt wird, welches der Referenzwert bei Indikatoren mit Referenzbereich  $[0; R]$  nicht unterschreitet:

$$R = \arg \min_{R \geq 1} \left| \sum_{i=1}^I [P(\theta_i \leq R | o_i, e_i) \leq \alpha] - q \cdot I \right|,$$

bzw. mit  $P(\theta_i \geq R | o_i, e_i)$  für Indikatoren mit Referenzbereich  $[R; \infty)$ .

Sowohl für ratenbasierte als auch für risikoadjustierte O/E-Indikatoren kann das Minimum über eine numerische Lösung des Optimierungsproblems bestimmt werden. Falls die obere Grenze eines Referenzbereichs der Form  $[0; R]$  nach der Optimierung unter dem Bundesergebnis liegt, wird sie auf das Bundesergebnis gesetzt. Analog dazu wird die untere Grenze eines Referenzbereichs der Form  $[R; 1]$  bzw.  $[R; \infty)$  auf das Bundesergebnis gesetzt, falls sie darüber liegt. Ein verteilungsbasierter Referenzwert kann dadurch nie nominell strenger werden als das Bundesergebnis im Indikator.



## Literatur

Hengelbrock, J; Rauh, J; Cederbaum, J; Kähler, M; Höhle, M (2022): Hospital profiling using Bayesian decision theory. *Biometrics*, Epub 19.11.2022. DOI: 10.1111/biom.13798.

IQTIG [Institut für Qualitätssicherung und Transparenz im Gesundheitswesen] (2024): Methodische Grundlagen. Version 2.1. [Stand:] 27.11.2024. Berlin: IQTIG.  
URL: [https://iqtig.org/downloads/berichte/2024/IQTIG\\_Methodische-Grundlagen\\_Version-2.1\\_2024-11-27-1.pdf](https://iqtig.org/downloads/berichte/2024/IQTIG_Methodische-Grundlagen_Version-2.1_2024-11-27-1.pdf) (abgerufen am: 11.12.2024).

# 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](mailto:info@iqtig.org)

[iqtig.org](http://iqtig.org)