

# LOST IN STATISTICS

## DIE NEUE AUFFÄLLIGKEITSEINSTUFUNG IM VERFAHREN QS PCI AUS PERSPEKTIVE DER LAG UND LEISTUNGSERBRINGER

Bund-Länder-Konferenz Kardiologie | Frankfurt a. Main, 29.11.2024

**Tanja Grath, MPH**

Stellvertretende Geschäftsführung LAG Bayern GbR, Projektleitung QS PCI

**Prof. Dr. med. Wolfgang von Scheidt**

Emer. Direktor I. Med. Klinik, Universitätsklinikum Augsburg

Ehemaliges Mitglied der Fachkommission PCI/Koronarangiographie Bayern



Landesarbeitsgemeinschaft zur datengestützten,  
einrichtungübergreifenden Qualitätssicherung in Bayern

# Gliederung

<b>TOP 1</b>	Ausgangssituation
<b>TOP 2</b>	Kritikpunkte an der neuen Auffälligkeitseinstufung
<b>TOP 3</b>	Verantwortung und Zielkonflikt
<b>TOP 4</b>	Fazit und Verbesserungsvorschläge

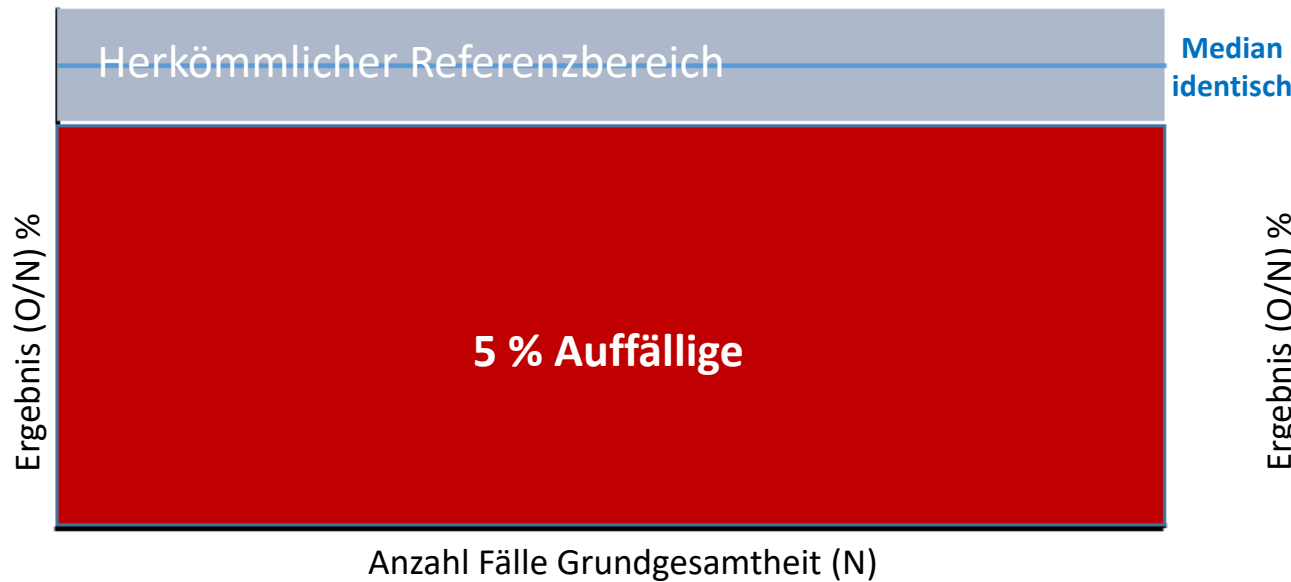
# Ausgangssituation

- DeQS-Richtlinie kennt ausschließlich „Auffälligkeiten“  
→ keine Unterscheidung in rechnerische vs. statistische Auffälligkeiten
- Bisher (für alle Verfahren gem. DeQS-RL)
  - Bestimmung Referenzwert (fest oder perzentilbasiert)
  - **Referenzwert bildet lineare Grenze** zur Auffälligkeit
  - Jedes Ergebnis ober- bzw. unterhalb des Referenzwerts (je nach Polung des QI) ist auffällig
- Seit dem Auswertungsjahr (AJ) 2022 neues Vorgehen zur Auffälligkeitseinstufung im Verfahren QS PCI
  - Minimierung der fallzahlabhängigen Irrtumswahrscheinlichkeit und Generierung von exakt 5 % Auffälligen (im Falle perzentilbasierter Ref.bereiche)
  - **Schwellenwert mit Hyperbelverlauf** als Grenze zur Auffälligkeit
- Ab dem AJ 2025 Ausweitung der geänderten Methodik auf 3 weitere QS-Verfahren

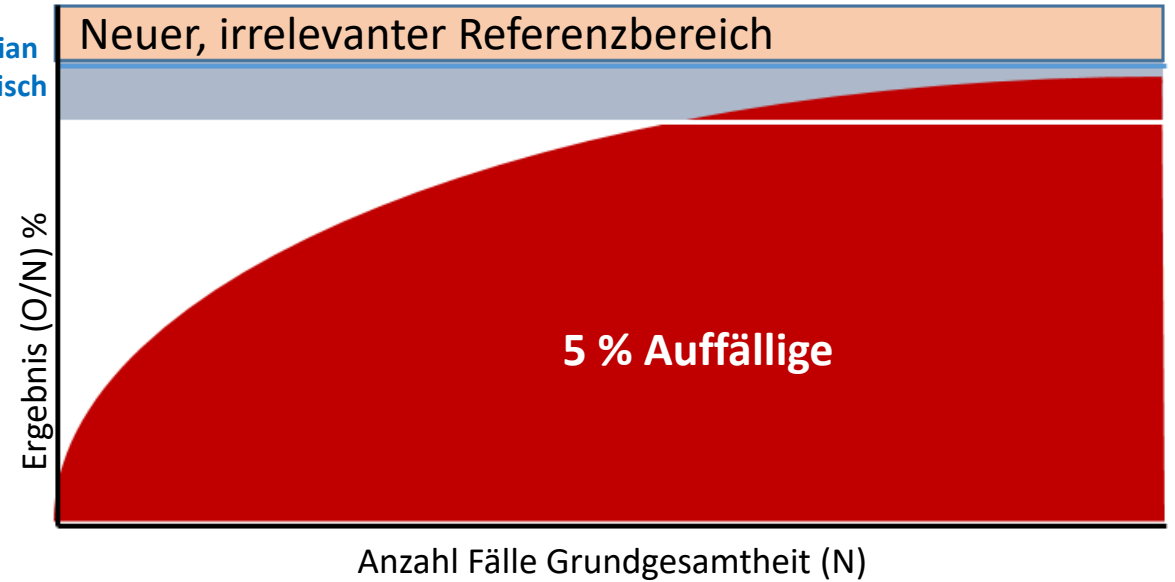
# Ausgangssituation

## Vergleich Auswertungsmethodik alt vs. neu

Alt (bis AJ 2021)



neu (seit AJ 2022)



# Kritikpunkt: Irreführende Darstellung eines Referenzwertes

---

Verringerte Verständlichkeit und Akzeptanz des Verfahrens

# „Leichte Uminterpretation von perzentilbasierten Referenzwerten“

(gem. Antwort des IQTIG vom 26.08.2022 auf eine Verfahrenssupport-Anfrage der LAG Bayern)

Erfassungsjahr	Referenzbereich QI ID 56014 (Interventionsziel bei STEMI-PCI)
<b>2021</b>	> <b>93.26 %</b> (5. Perzentil)
2020	> 85.36 % (5. Perzentil)
2019	> 84.21 % (5. Perzentil)
2018	> 83.33 % (5. Perzentil)
2017	> 82.42 % (5. Perzentil)
2016	> 80.71 % (5. Perzentil)

*„Die starke Verschiebung der Referenzbereiche ist in der Tat darauf zurückzuführen, dass mit der Einführung der neuen Auffälligkeitseinstufung auch die verteilungsbasierte Ermittlung der Referenzbereiche an die neue Methodik angepasst wurde.“*

(aus IQTIG-Antwort an LAG Bayern vom 01.02.2023)

# Perspektive IQTIG vs. Leistungserbringer

„Gemäß den methodischen Grundlagen, Version 2.0, ist die Eigenschaft, dass der Referenzwert gleichzeitig die Grenze zur Auffälligkeit bildet, keine definierende Eigenschaft des Referenzwertes, sondern eine Eigenschaft der alten Methodik der rechnerischen Auffälligkeit.“

(aus IQTIG-Antwort an LAG Bayern vom 01.02.2023)



Wikipedia

[https://de.wikipedia.org > wiki > Referenzbereich\\_\(Med...](https://de.wikipedia.org/wiki/Referenzbereich_(Medizin))

## Referenzbereich (Medizin)

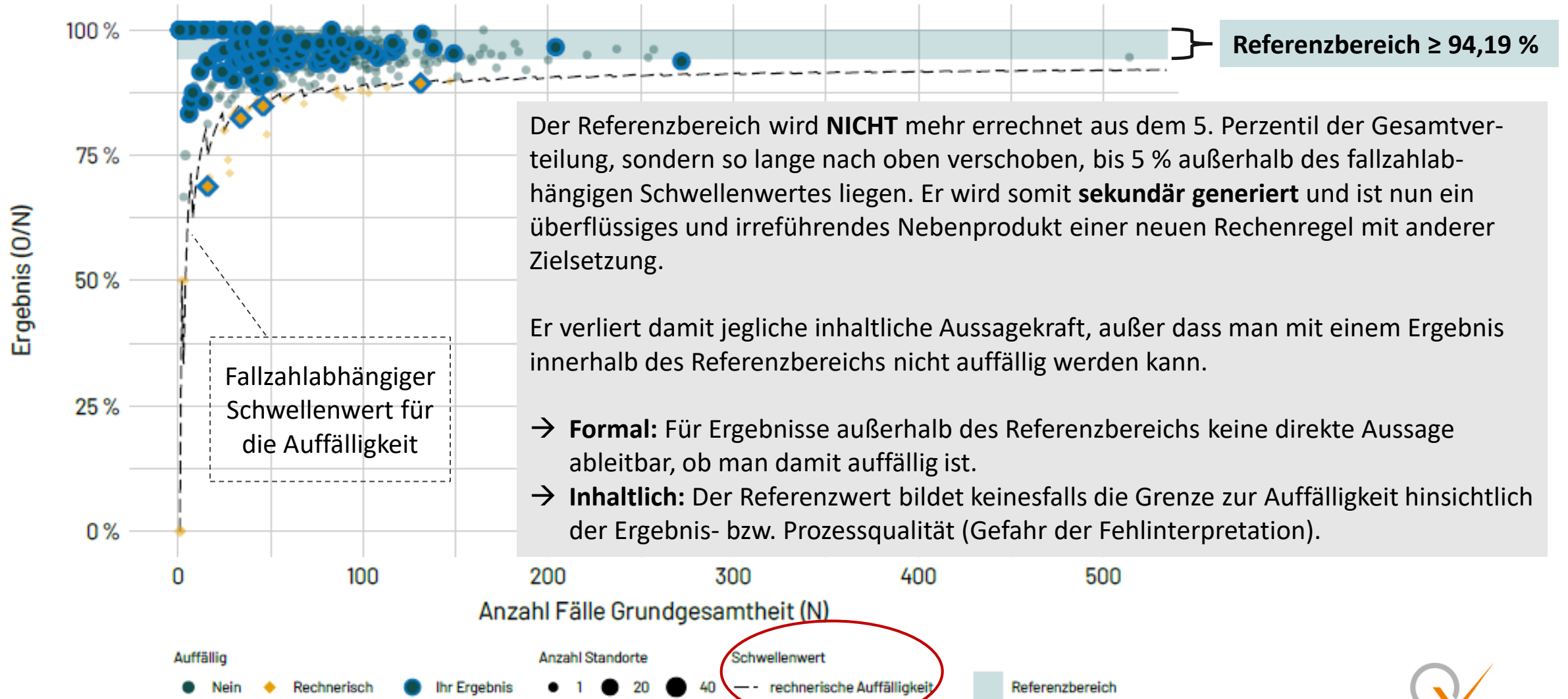
Umgangssprachlich werden Werte innerhalb des Referenzbereichs als normal, Werte außerhalb als pathologisch (krankhaft) bezeichnet. Dabei ist aber zu beachten, ...

Konträr zu gängigem Verständnis von Referenzbereichen/-werten

„Perzentilbasierte Referenzbereiche sind nicht für die Bewertung der bundesweiten Leistungsqualität geeignet.“

(aus IQTIG-Antwort an LAG Bayern vom 01.02.2023)

Nicht mehr der Referenzwert, sondern der Schwellenwert definiert die Grenze zur rechnerischen Auffälligkeit!





...dennoch weiterhin Verwendung in den Auswertungen:

Indikatoren und Kennzahlen		Ihr Ergebnis	Bund (gesamt)
ID	Bezeichnung des Indikators oder der Kennzahl	Ergebnis	Ergebnis
56014	Erreichen des wesentlichen Interventionsziels bei PCI mit der Indikation ST-Hebungsinfarkt	95,36 % O = 5.402 N = 5.665	95,01 % O = 36.610 N = 38.531



- Tatsächlich ist der Referenzbereich aber irrelevant für die Auffälligkeitseinstufung
- **Die Beibehaltung der Referenzbereich-Angabe mit der Suggestion einer Auffälligkeits-Bedeutung und methodischer Kontinuität ist irreführend und fördert weder Verständnis noch Akzeptanz**
- Die Angabe 5. Perzentil beim Referenzbereich ist u. E. falsch, da die 5. Perzentilgrenze durch den Schwellenwert definiert wird, nicht mehr durch den Referenzbereich

# In Wahrheit: Verteilungsabhängiger Schwellenwert

Indikatoren und Kennzahlen			Ihr Ergebnis	Bund (gesamt)
ID	Bezeichnung des Indikators oder der Kennzahl	Referenzber.	Ergebnis	Ergebnis
56014	Erreichen des wesentlichen Interventionsziels bei PCI mit der Indikation ST-Hebungsinfarkt	<del>                     ≥ 94,19 %                      (5. Perzentil)                 </del>	95,36 % O = 5.402 N = 5.665	95,01 % O = 36.610 N = 38.531

- Korrekt wäre die alleinige Benennung des verteilungsabhängigen Schwellenwertes als Grenze zur (rechnerischen) Auffälligkeit:

Indikatoren und Kennzahlen			Ihr Ergebnis	Bund (gesamt)
ID	Bezeichnung des Indikators oder der Kennzahl	Schwellenwert	Ergebnis	Ergebnis
56014	Erreichen des wesentlichen Interventionsziels bei PCI mit der Indikation ST-Hebungsinfarkt	Verteilungs- abhängiger Schwellenwert (5. Perzentil)	95,36 % O = 5.402 N = 5.665	95,01 % O = 36.610 N = 38.531

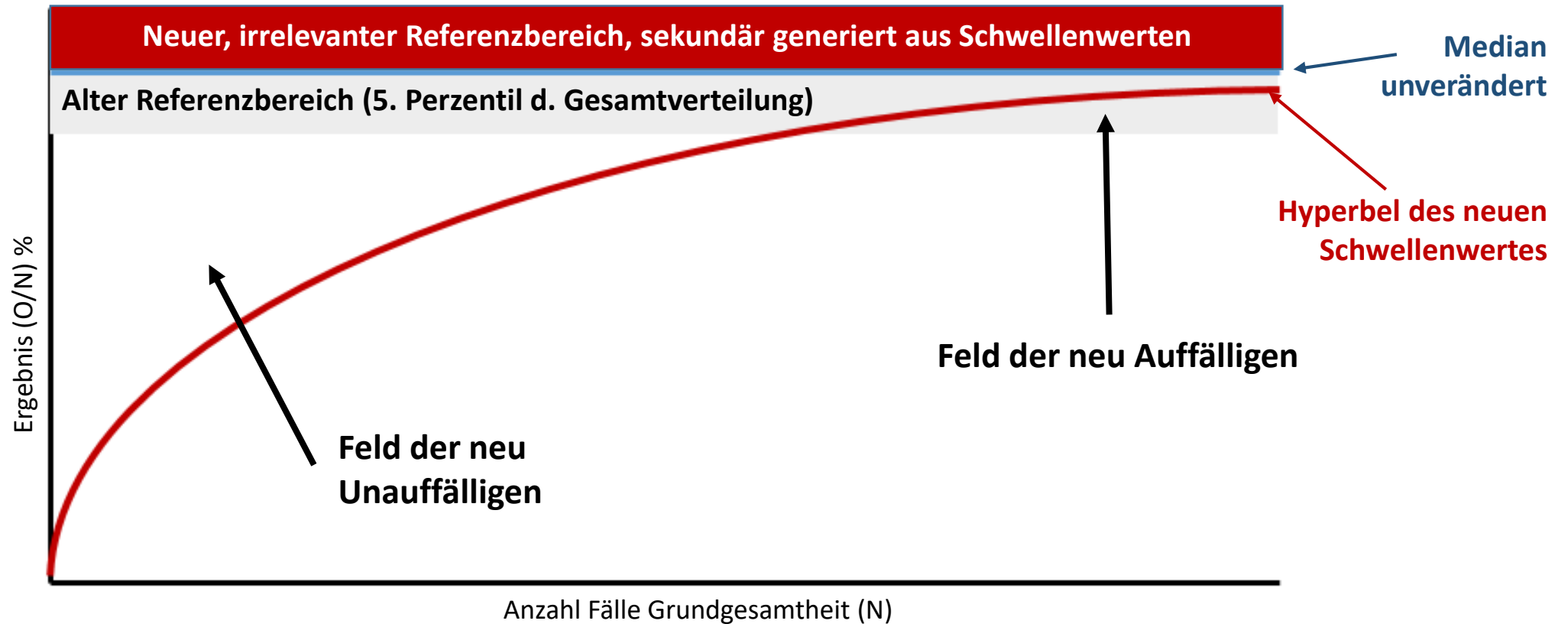


# Kritikpunkt: Inhaltliche Angemessenheit

---

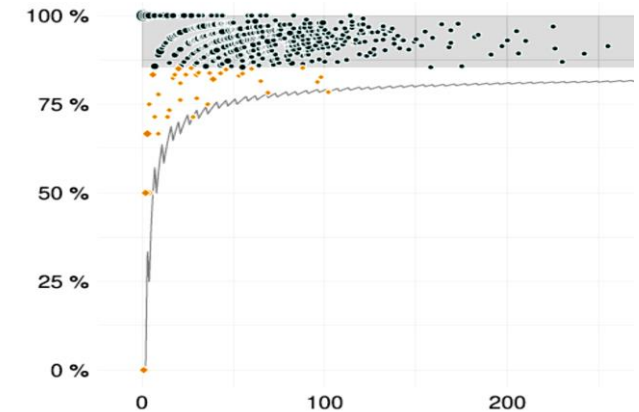
Werden die Richtigen auffällig?

Gnade bei kleinen Fallzahlen erfordert Strafe bei hohen Fallzahlen, um das Ziel zu erreichen, 5 % Auffällige zu generieren



# Inhaltliche Grundüberlegungen

- Mindestmengen-Diskussion
  - Leistungserbringer mit kleineren Fallzahlen zeigen schlechtere Resultate
- Komplexere Fälle werden bislang nicht risikoadjustiert
  - Leistungserbringer mit hohen Fallzahlen versorgen häufiger komplexe Fälle, die aber für eine Vielzahl der QI nicht risikoadjustiert werden
  - NB: Im Verfahren QS PCI soll dies im Rahmen der Weiterentwicklung des Verfahrens größtenteils behoben werden
- Gegenüber Kleineren Gnade zu üben und Größere anzuprangern, ist daher kontraproduktiv zur Erkennung von Qualität



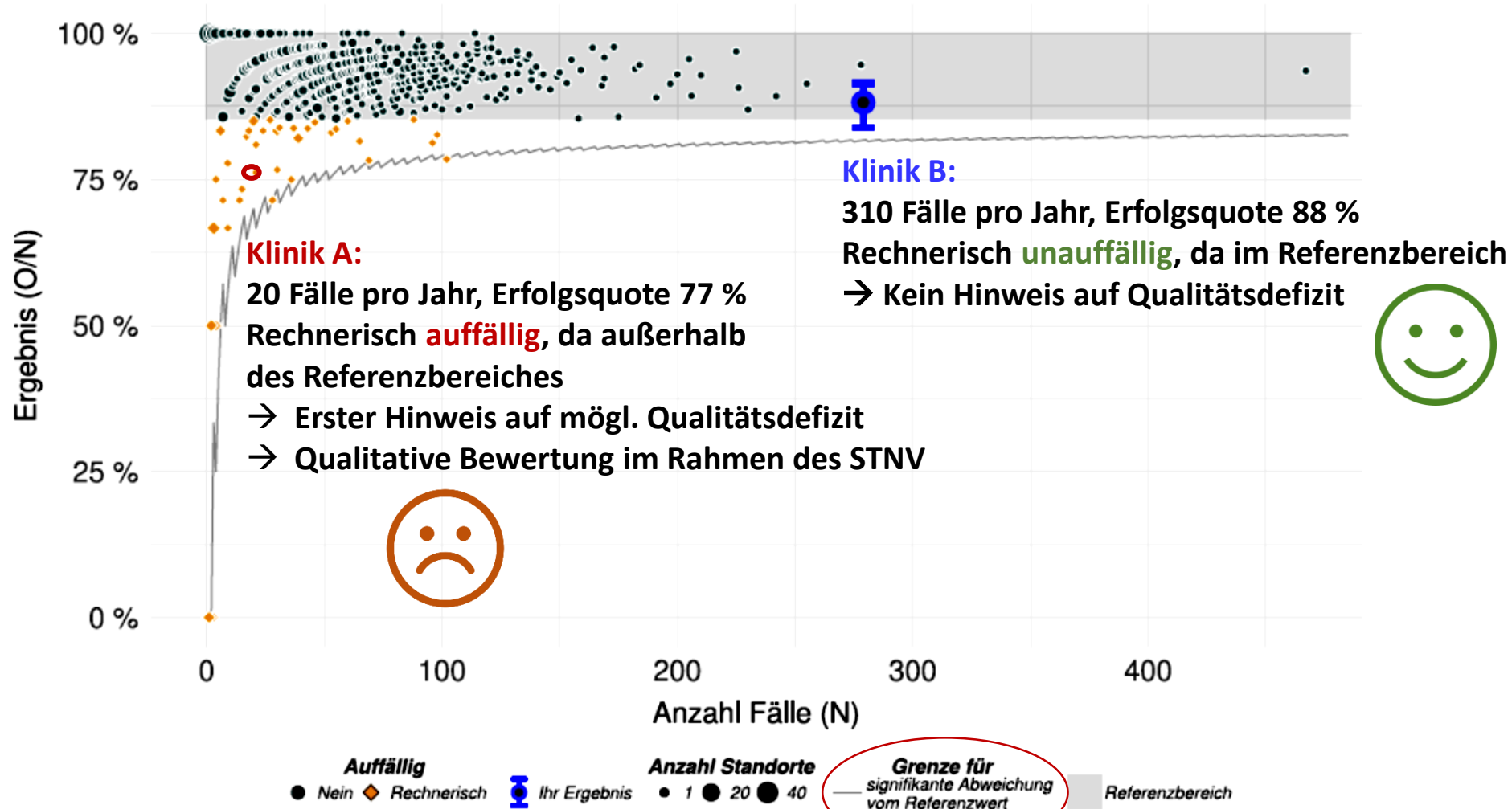
# Kritikpunkt: Außenwirkung

---

Wie werden Ergebnisse von der Öffentlichkeit ggf. genutzt und interpretiert?

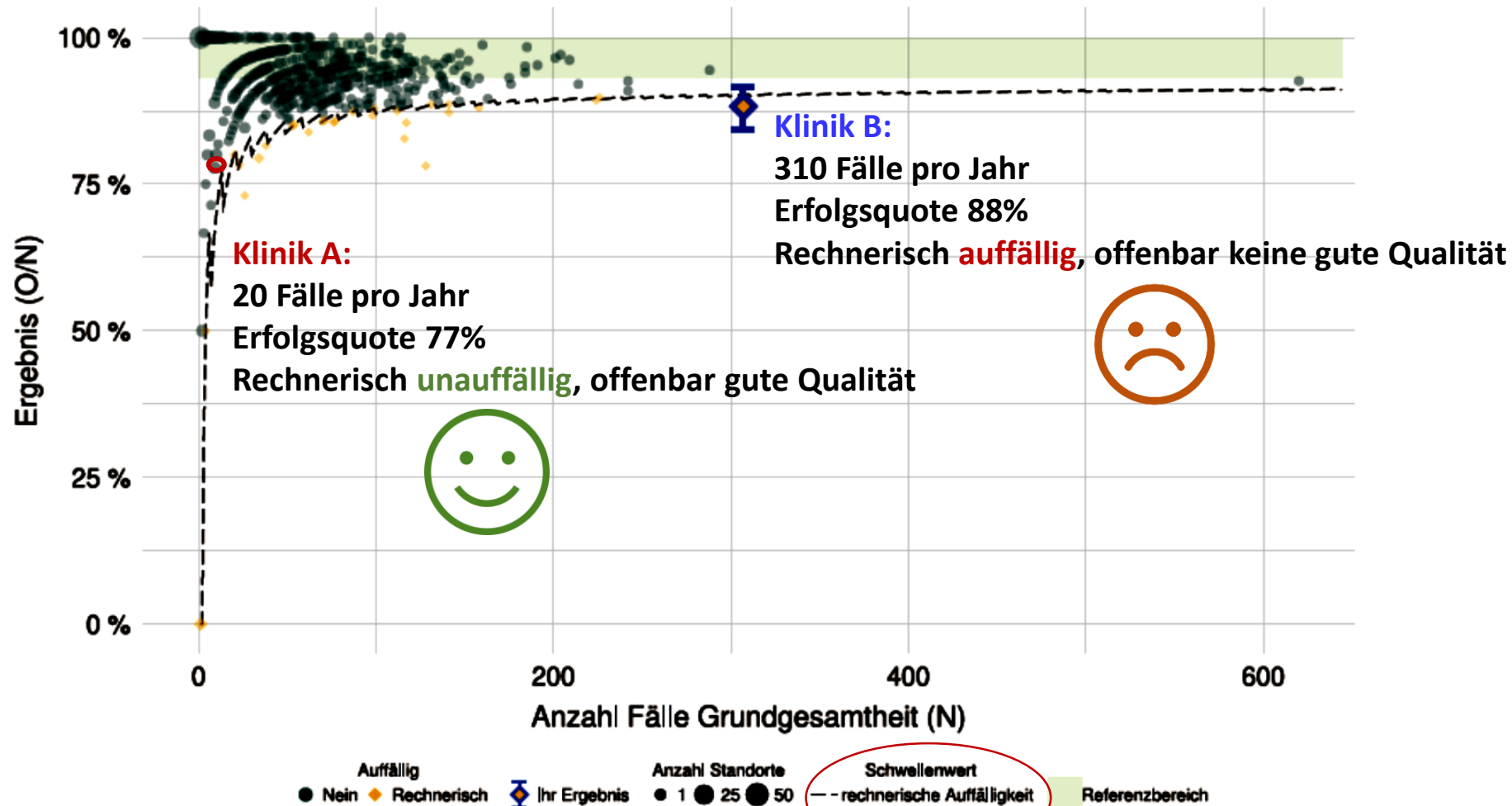
# Wo würden Sie als Patient hingehen? (alte Methodik)

QI ID 56014 Erreichen des wesentlichen Interventionsziels bei PCI bei STEMI (Bsp. aus dem AJ 2021)



# Wo würden Sie als Patient hingehen? (neue Methodik)

QI ID 56014 Erreichen des wesentlichen Interventionsziels bei PCI bei STEMI (Bsp. aus dem AJ 2022)





# Wo würden Sie als Patient hingehen?

- Gefahr der Verzerrung der inhaltlichen Qualitätswahrnehmung für Leistungserbringer und Patienten
- Irreale Erwartungshaltung:  
Welches Benchmark kann ich als Leistungserbringer und Patient:in überhaupt erwarten?

## Literatur-Benchmark TIMI- 3-Fluss bei STEMI-PCI ?!

Vonseiten einer LAG bzw. deren Fachkommission wird auf internationale, randomisierte Studien und Meta-Analysen hingewiesen, die belegten, dass die Rate des Erreichens des Interventionsziels bei PCI mit der Indikation ST-Hebungsinfarkt (definiert als TIMI-3-Fluss nach Intervention) selbst bei selektierten Studienpatientinnen und -patienten „nur“ um 88 – 89 % liegt, bei Schock-Patientinnen und -Patienten um 85 % (Van de Werf et al. 2006, Keeley et al. 2006, Guo und Yang 2019, Thiele et al. 2017). Das Bundesergebnis lag im Erfassungsjahr 2021 bei 93,25 % und somit deutlich über den literaturgestützten durchschnittlichen Raten. Die Interpretation der LAG hinsichtlich des Vorliegens eines möglichen Deckeneffekts hat sich bei der Prüfung durch das IQTIG nicht bestätigt.

Das IQTIG empfiehlt, Patientinnen und Patienten mit Nicht-ST-Hebungsinfarkt (NSTEMI) und Patientinnen und Patienten mit subakutem ST-Hebungsinfarkt (STEMI) zukünftig in den Nenner des Indikators aufzunehmen und den Zähler um das Kriterium „TIMI-2-Fluss“ zu erweitern. Die Einführung eines festen Referenzbereichs, die nicht Gegenstand der vorliegenden Beauftragung war, wäre unter Berücksichtigung der Auswirkungen der zuvor genannten Empfehlungen auf das Indikatorergebnis gesondert zu prüfen.

# Verantwortung und Zielkonflikt

---

Ziel sollte die Abbildung von Ergebnis- und Prozessqualität sein, nicht die Generierung von 5 % rechnerischen Auffälligkeiten mit geringstmöglicher Irrtumswahrscheinlichkeit.

# Weder gemeinsame Ebene, noch gleiches Ziel

*„Es ist die Rolle der Qualitätsindikatoren, und nicht der Einstufungsmethodik, die Qualität messbar zu machen.“*

(aus IQTIG-Antwort an LAG Bayern vom 01.02.2023)

## **Einspruch:**

Ein QI per se ist kein Maß für Qualität, es bedarf eines sachdienlichen Bewertungsmaßstabs (=Einstufungsmethodik), der die Auffälligkeitsgrenze definiert.

Erderwärmung ist ein QI für Klimawandel, aber zur Bewertung von Qualität von Klimaschutzmaßnahmen bedarf es z. B. des  $\leq 2\%$ -Ziels.

# Weder gemeinsame Ebene, noch gleiches Ziel

*„Das Perzentil der beobachteten Indikatorergebnisse ist kein geeignetes Maß für die Leistungsqualität. Perzentilbasierte Referenzbereiche haben das Ziel, die Anzahl an rechnerischen Auffälligkeiten und damit den Aufwand des Verfahrens zu steuern. Mit der neuen Methodik kann dieses Ziel erstmals zuverlässig erfüllt werden.“*

(aus IQTIG-Antwort an LAG Bayern vom 01.02.2023)

## Heißt übersetzt:

Qualitätserkennung ist nicht das primäre Ziel, sondern die Schaffung einer Rechenregel, die exakt 5 % Auffällige abliefert. Deren Leistungsqualität kann jedoch lt. IQTIG im Falle von QI mit perzentilbasiertem Referenzbereich nicht korrekt beurteilt werden.



Dies vernachlässigt jedoch, dass u. a. die qualitative Bewertung im STNV z. B. im **Bundesklinikatlas** nicht Teil der veröffentlichten Qualitätsdaten ist → verzerrte Außenwahrnehmung.

# Zielkonflikt

**Einer (vermeintlichen) statistischen Sicherheit wurde die (prioritäre) inhaltliche Korrektheit geopfert.**

Dadurch geht leider das selbst definierte **Ziel verloren**, nämlich die bestmögliche Bewertung und Förderung der Versorgungsqualität.

*Empfehlungen zur Weiterentwicklung der Verfahren QS PCI, QS HSMDEF und QS KEP*

*Abschlussbericht*

## Hintergrund

Das IQTIG setzt im Auftrag des Gemeinsamen Bundesausschusses (G-BA) zum Zeitpunkt der Berichtslegung 15 Qualitätssicherungsverfahren (QS-Verfahren) gemäß der Richtlinie zur datengestützten einrichtungsübergreifenden Qualitätssicherung (DeQS-RL) um. Das Ziel dieser QS-Verfahren ist gemäß der DeQS-RL insbesondere die **Förderung der Versorgungsqualität.**

# Fazit und Verbesserungsvorschläge

---

# Fazit

- Grundsätzlich ist eine neue Methodik zur Auffälligkeitseinstufung durchaus zu begrüßen
- Insbesondere eine Differenzierung zwischen „rechnerisch vs. statistisch auffällig“ ist aus Sicht der LAG Bayern grundsätzlich sinnvoll
- Aber...



# Fazit: Lost in Statistics

- Neue Auffälligkeitseinstufung ist aus Sicht der LAG Bayern und vieler Leistungserbringer v. a. für perzentilbasierte QI eine formal und inhaltlich **irreführende Rechenregel**, die im Verfahren QS PCI nicht zur besseren Erkennbarkeit der Versorgungsqualität geführt hat.
- Verteilungsbezogene (fallzahlabhängige) Berücksichtigung statistischer Irrtumswahrscheinlichkeiten ist **keine „One size fits all“ Lösung**.
  - Es gibt Leistungsbereiche und/oder Qualitätsindikatoren, bei denen insbesondere Leistungserbringer mit kleinen Gesamtfallzahlen potenziell eine schlechtere Qualität aufweisen.
  - Problem: keine Auffälligkeit → kein Stellungnahmeverfahren → keine Konsequenzen
- **Gefahr der Fehlinterpretation**
  - besondere Bedeutung v. a. im Hinblick auf Bundesklinikatlas, in dem QS-Ergebnisse ohne eine qualitative Bewertung der Fachkommission verwendet werden

# Verbesserungsvorschläge

- Adäquate **Risikoadjustierung** für alle Ergebnisindikatoren
- **Feste Referenzwerte** an Stelle perzentilbasierter als bevorzugtes Maß

*„Im Verfahren PCI ist auch geplant, die perzentilbasierten Referenzbereiche nach und nach durch feste Referenzbereiche zu ersetzen.“* (aus IQTIG-Antwort an LAG Bayern vom 01.02.2023)

- Bislang nicht erfolgt; auch künftig wird dies schwer umsetzbar sein
  - Dann aber bitte keine pauschale Unauffälligkeit für LE mit kleiner Fallzahl aufgrund der Irrtumswahrscheinlichkeit
  - Oder würden Sie den Referenzwert von 50 km/h innerorts bei 50 Ortsdurchfahrten pro Jahr auf einen Schwellenwert von 70 km/h erhöhen und bei 300 Ortsdurchfahrten pro Jahr auf einen Schwellenwert von 40 km/h verschärfen?
- Darstellung von Angaben zu **rechnerischer UND statistischer Auffälligkeit** ggü. LAG und LE
    - Möglichkeit eines Stellungnahmeverfahrens bleibt auch bei kleinen Fallzahlen erhalten

# Vielen Dank für Ihre Aufmerksamkeit

---