

## Elektronisches Supplement 2

### Beschreibung des ROC-Verfahrens und der Vorgehensweise

Für die Konstruktion der Cut-Off-Werte wurde die Methode der *Receiver-Operating-Characteristics*-Analyse (ROC-Analyse) verwendet, welche aus der Signalentdeckungstheorie von Green und Swets (1966) stammt. Bei der Methode wird für jeden möglichen Cut-Off-Wert die Sensitivität und Spezifität abgetragen (Goldhammer & Harting, 2012; Schäfer, 1989). Sensitivität und Spezifität stehen in gegenläufiger Beziehung zueinander, was bedeutet, dass eine steigende Sensitivität mit hoher Wahrscheinlichkeit mit einer sinkenden Spezifität einhergeht und umgekehrt. Angewandt auf diese Studie bedeutet Sensitivität, dass ein Test bei einem Probanden mit einem objektiv hohen bzw. gut gestalteten Tätigkeitsspielraum auch tatsächlich einen hohen Tätigkeitsspielraum anzeigt. Umgekehrt bedeutet Spezifität, dass ein Test bei einem Probanden mit objektiv niedrigem bzw. schlecht gestaltetem Tätigkeitsspielraum auch tatsächlich einen geringen Tätigkeitsspielraum anzeigt. Die ROC-Kurve gibt an, wie gut der Test geeignet ist, um zwischen zwei Gruppen zu unterscheiden. Dabei ist die *area under the curve* (AUC) ein Maß für die diagnostische Güte eines Tests. Nach Swets (1988) spricht man bei einer AUC zwischen 0.5 und 0.7 von einem kleinen, zwischen 0.7 und 0.9 von einem mittleren und zwischen 0.9 und 1 von einem hohen diagnostischen Nutzen des Tests. Zusätzlich wurden Konfidenzintervalle berechnet, deren Enge Aufschluss darüber geben, wie genau die geschätzten Cut-Off-Werte sind (Carpenter & Bithell, 2000; Hilgers, 1991). Bei einer ROC-Analyse werden somit viele verschiedene Cut-Off-Werte und deren Spezifitäten und Sensitivitäten berechnet. Je nach den Konsequenzen, muss abgewogen werden, auf welchen Cut-Off-Wert die Entscheidung fällt. Die Berechnung des Youden-Index ( $Y = \text{Sensitivität} + \text{Spezifität} - 1$ ) ist in der Literatur die favorisierte Methode zum Finden des optimalen Cut-Off-Wertes (Perkins & Schisterman, 2006). Dabei wird das optimale Verhältnis von Sensitivität und Spezifität gesucht. Nach Goldhammer und Harting (2012) kann der Cut-Off-Wert aber in eine Richtung verschoben werden, wenn die Konsequenzen eines falsch-positiven Testwertes höher sind als die Konsequenzen eines falsch-negativen Testwertes oder umgekehrt. Daher wurde bei der Konstruktion des Cut-Off-Wertes in dieser Untersuchung darauf geachtet, dass je nachdem, ob z. B. die Sensitivität für das Kriterium wichtiger ist, der Cut-Off-Wert, ausgehend von dem optimalen Cut-Off-Wert nach Youden (1950), solange zu Gunsten der Sensitivität verschoben wird, bis die Sensitivität die Spezifität übersteigt. Allerdings nur, wenn die Sensitivität bei dem optimalen Wert des Youden-Index kleiner ist als die Spezifität. Theoretisch wäre es möglich die Sensitivität oder Spezifität zu gewichten (Perkins & Schisterman, 2006). Da sich jedoch die Kosten einer Fehlklassifikation nicht genau beziffern lassen und nicht nur ökonomische, sondern auch ethische Aspekte zu berücksichtigen sind, wurde auf eine Gewichtung und die damit verbundene Manipulation mathematischer Kennziffern verzichtet.

### Literatur

Carpenter, J. & Bithell, J. (2000). Bootstrap confidence intervals: when, which, what? A practical guide for medical statisticians. *Statistics in Medicine*, 19 (9), 1141–1164. [https://doi.org/10.1002/\(SICI\)1097-0258\(20000515\)19:9<1141::AID-SIM479>3.0.CO;2-F](https://doi.org/10.1002/(SICI)1097-0258(20000515)19:9<1141::AID-SIM479>3.0.CO;2-F)

- Goldhammer, F. & Harting, F. (2012). Interpretation von Testresultaten und Testeichung. In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 173–201). Heidelberg: Springer-Verlag. [https://doi.org/10.1007/978-3-642-20072-4\\_2](https://doi.org/10.1007/978-3-642-20072-4_2)
- Green, D. M. & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: John Wiley and Sons.
- Hilgers, R. A. (1991). Distribution-free confidence bounds for ROC curves. *Methods of Information in Medicine*, 30, 96–101. <https://doi.org/10.1055/s-0038-1634822>
- Perkins, N. J. & Schisterman, E. F. (2006). The Inconsistency of “Optimal” Cutpoints Obtained using Two Criteria based on the Receiver Operating Characteristic Curve. *American Journal of Epidemiology*, 163 (7), 670–675. <https://doi.org/10.1093/aje/kwj063>
- Schäfer, H. (1989). Constructing a cut-off point for a quantitative diagnostic test. *Statistics in Medicine*, 8, 1381–1391. <https://doi.org/10.1002/sim.4780081110>
- Swets, J. A. (1988). Measuring the accuracy of diagnostic systems. *Science*, 240, 1285–1293. <https://doi.org/10.1126/science.3287615>
- Youden, W. J. (1950). Index for rating diagnostic tests. *Cancer*, 3, 32–35. [https://doi.org/10.1002/1097-0142\(1950\)3:1<32::AID-CNCR2820030106>3.0.CO;2-3](https://doi.org/10.1002/1097-0142(1950)3:1<32::AID-CNCR2820030106>3.0.CO;2-3)