

Was ist ein Konfidenzintervall?

R. Bender¹, St. Lange²

¹ Fakultät für Gesundheitswissenschaften, AG Epidemiologie und medizinische Statistik, Universität Bielefeld

² Abteilung für Medizinische Informatik, Biometrie und Epidemiologie der Ruhr-Universität Bochum

Ein Konfidenzintervall (1, 3, 4) ist ein Unsicherheitsbereich für die Schätzung eines interessierenden Parameters (zum Beispiel Mittelwert, Median, Differenz zweier Mittelwerte, Regressionskoeffizient, Differenz zweier Wahrscheinlichkeiten, Risk Ratio, Odds Ratio etc.) aus einer Stichprobe. Das Ergebnis einer solchen Schätzung ist abhängig von der gezogenen Stichprobe und weist damit eine Zufallsschwankung auf. Zur Berechnung eines Konfidenzintervalls muss man die gewünschte Überdeckungswahrscheinlichkeit spezifizieren (häufig 95%). Ein 95%-Konfidenzintervall zum Beispiel ist ein Bereich, der den theoretischen (unbekannten) Wert des interessierenden Parameters mit einer Wahrscheinlichkeit von 95% beinhaltet. Um einem verbreiteten Missverständnis vorzubeugen, sollte betont werden, dass der zu schätzende Parameter fest ist, also keine zufällige Größe darstellt.

Ein Konfidenzintervall liefert (im Gegensatz zum *p*-Wert [2]) Information in der gleichen Skala wie die untersuchte Variable. Wird zum Beispiel der systolische Blutdruck (in mm Hg) untersucht, so ist die Einheit der Werte des Konfidenzintervalls ebenfalls mm Hg. Damit lässt sich die Unsicherheit über die Stärke eines Effekts mit Hilfe eines Konfidenzintervalls zum Beispiel für die Differenzierung zweier Blutdruck-Mittelwerte klinisch interpretieren. In **Tab.1** findet man für sechs Gruppenvergleiche basierend auf hypothetischen Daten (systolischer Blutdruck in mm Hg) die 95%-Konfidenzintervalle für die Differenzierung der Mittelwerte. Die Breite der Konfidenzintervalle hängt ab von den Stichprobenumfängen (n_1 , n_2) und den Standardabweichungen (SD) (5) in den Gruppen.

Aus einem Konfidenzintervall lassen sich Rückschlüsse auf die statistische Signifikanz ableiten. Enthält ein 95%-Konfidenzintervall nicht den Wert des »Null-Effekts« (zum Beispiel 0 bei der Differenz zweier Mittelwerte oder 1 bei einem Odds Ratio), so bedeutet dies ein signifikantes Ergebnis zum Niveau $\alpha = 5\%$ für den Ausschluss eines »Null-Effekts«, das heißt der *p*-Wert ist kleiner als 5%. Um wie viel kleiner der *p*-Wert ist, kann man jedoch am Konfidenzintervall nicht ablesen. Dafür

muss man den *p*-Wert selbst über die entsprechende Teststatistik (2) berechnen.

Die Angabe eines Konfidenzintervalls macht damit die Angabe eines *p*-Werts nicht völlig überflüssig. Der *p*-Wert ist die genaue Angabe der statistischen Evidenz, das Konfidenzintervall liefert klinisch interpretierbare Informationen über die Unsicherheit des beobachteten Effekts (1, 3).

kurzgefasst: Das Konfidenzintervall ist der Unsicherheitsbereich für die Schätzung eines bestimmten, nicht bekannten Parameters. Ein 95%-Konfidenzintervall beispielsweise enthält den gesuchten Parameter mit einer Wahrscheinlichkeit von 95%. Aus dem Konfidenzintervall lassen sich Schlüsse bezüglich der statistischen Signifikanz ziehen.

Literatur

- 1 Altman DG. Use of confidence intervals to indicate uncertainty in research findings. *Evidence-Based Med* 1996; 1: 102–104
- 2 Bender R, Lange S. Was ist der P-Wert? *Dtsch med Wschr* 2001; 126: T39–T40
- 3 Gardner MJ, Altman DG. Confidence intervals rather than P values: Estimating rather than hypothesis testing. *Brit med J* 1986; 292: 746–750
- 4 Guyatt GH, Jaeschke R, Heddle N, Cook D, Shannon H, Walter S. Basic statistics for clinicians: 2. Interpreting study results: Confidence intervals. *Can Med Ass J* 1995; 152: 169–172
- 5 Lange S, Bender R. Variabilitätsmaße. *Dtsch med Wschr* 2001; 126: T29–T30

Korrespondenz

Dr. Ralf Bender
 Fakultät für Gesundheitswissenschaften
 AG3: Epidemiologie und medizinische Statistik
 Universität Bielefeld
 Postfach 100131
 33501 Bielefeld
 E-Mail: ralf.bender@uni-bielefeld.de

Tab.1 95%-Konfidenzintervalle für die Differenz von Mittelwerten zweier Gruppen bei hypothetischen Daten (systolischer Blutdruck in mm Hg) mit variierender Differenz, Stichprobengröße und Variabilität.

Nr.	Stichprobenumfänge	Medikament Mittelwert (SD)	Placebo Mittelwert (SD)	Differenz der Mittelwerte	95%-Konfidenzintervall	P-Wert
1	$n_1 = n_2 = 10$	160 (22)	180 (22)	20	[-0,7; 40,7]	0,057
2	$n_1 = n_2 = 10$	160 (15)	180 (15)	20	[5,9; 34,1]	0,008
3	$n_1 = n_2 = 20$	160 (22)	180 (22)	20	[5,9; 34,1]	0,007
4	$n_1 = n_2 = 50$	170 (22)	180 (22)	10	[1,3; 18,7]	0,025
5	$n_1 = n_2 = 5$	140 (50)	180 (50)	40	[-32,9; 112,9]	0,242
6	$n_1 = n_2 = 1000$	178 (12)	180 (12)	2	[0,9; 3,1]	<0,001

SD = Standardabweichung