Median oder Mittelwert?

St. Lange[1], R. Bender[2]

[1] Abteilung für Medizinische Informatik, Biometrie und Epidemiologie der Ruhr-Universität Bochum
[2] Fakultät für Gesundheitswissenschaften, AG Epidemiologie und medizinische Statistik, Universität Bielefeld

In nahezu jeder (klinischen) Studie ist es erforderlich, die gewonnenen Daten in geeigneter und einfacher Weise zusammenzufassen. Für kategoriale Merkmale, wie zum Beispiel Geschlecht oder Blutgruppe, liegt es nahe, die Anzahl von Beobachtungen innerhalb jeder Kategorie darzustellen, gewöhnlich als absolute Häufigkeit oder Prozentwert bezogen auf die Gesamtzahl beobachteter Patienten. Für die statistische Beschreibung von quantitativen, stetigen Merkmalen wird üblicherweise ein »Durchschnittswert« angegeben. Ein solcher Wert soll einen für die beobachtete Population typischen Wert repräsentieren. Unter »Durchschnitt« wird im allgemeinen Sprachgebrauch der arithmetische Mittelwert verstanden, der definiert ist als die Summe aller beobachteten Werte geteilt durch die Gesamtzahl der Beobachtungen.

Neben dem arithmetischen Mittelwert gibt es noch ein weiteres häufig verwendetes Lagemaß, den Median. Der Median ist derjenige Wert der sortierten Stichprobe, der genau in der Mitte liegt. Er teilt die Stichprobe also in zwei gleich große Hälften (bei geradem Stichprobenumfang liegen genau zwei Werte in der Mitte; der Median ist dann als arithmetischer Mittelwert dieser beiden Werte definiert). Die eine Hälfte weist Werte auf, die kleiner als der Median sind, während die Werte der anderen Hälfte größer als der Median sind. In **Tab.1** sind die Werte für den systolischen Blutdruck zum Zeitpunkt der Krankenhausaufnahme von 25 Patienten mit akutem Myokardinfarkt angegeben. Der Mittelwert beträgt 128 mmHg (3200/25), der Median 123 mmHg (der 13. Wert der sortierten Stichprobe).

Median und arithmetischer Mittelwert haben unterschiedliche Eigenschaften: Der Median wird von extremen Werten (Ausreißern) praktisch kaum beeinflusst. Dies ist jedoch kein Qualitätskriterium, sondern eine Eigenschaft. Sie bedeutet, dass der Median weniger von Ausreißern »gestört« wird, andererseits jedoch auch, dass auf Ausreißer weniger deutlich aufmerksam gemacht wird. Da bei jeder Anwendung ohnehin über Extremwerte gesondert nachgedacht werden muss, und ihre Auswirkungen berücksichtigt werden müssen, ist diese Eigenschaft des Medians meist nicht von erheblicher Relevanz.

Entweder stellt ein Extremwert einen plausiblen Wert der Stichprobe dar, dann ist der Mittelwert unter dessen Einbeziehung eine sinnvolle Beschreibung, oder es ist davon auszugehen, dass der Extremwert unplausibel ist, dann kann der Mittelwert auch ohne diesen Extremwert berechnet werden. Letzteres erfordert natürlich eine Begründung.

Tab.1 Systolische Blutdruckwerte von 25 Patienten mit akutem Myokardinfarkt zum Zeitpunkt der Krankenhausaufnahme (aufgelistet nach Patientennummer und aufsteigend sortiert nach Höhe des Blutdrucks)

Patientennummer	systolischer Blutdruck (mm Hg)	sortierte Blutdruckwerte (mm Hg)
1	81	81
2	170	99
3	120	106
4	127	108
5	190	110
6	118	113
7	140	118
8	132	120
9	152	120
10	123	120
11	106	123
12	120	123
13	130	123
14	99	126
15	123	127
16	108	130
17	123	131
18	110	132
19	131	138
20	126	140
21	160	140
22	113	152
23	120	160
24	140	170
25	138	190

Mittelwert: 128 mm Hg ([81 + 170 + 120 + ... + 138]/25 = 3200/25);
Median: 123 mm Hg (der 13. Wert der sortierten Stichprobe).

Bei schiefen, unsymmetrischen Verteilungen, wie sie für Laborwerte typisch sind, kann der Median besser interpretiert werden als der Mittelwert (1). Bei einer Erhebung in einer internistischen Notfallaufnahmestation werden beispielsweise die Werte der Kreatinkinase (CK) für die meisten Patienten zwischen 0 und 50 U/l liegen, allerdings werden einige Patienten auch Werte bis 1000 U/l und darüber aufweisen. Ohne dass hier Ausreißer im eigentlichen Sinne vorliegen, lässt ein Mittelwert von zum Beispiel 100 U/l überhaupt keine sinnvolle Interpretation zu, während die Interpretation des Medians unbeeinflusst bleibt: Die Hälfte der Messwerte ist niedriger

Dtsch. Med. Wschr. 2001; 126: T 25–T 26
© Georg Thieme Verlag Stuttgart · New York

Dieses Dokument wurde zum persönlichen Gebrauch heruntergeladen. Vervielfältigung nur mit Zustimmung des Verlages.

Tab. 2 Übersetzungen (deutsch – englisch)

(arithmetischer) Mittelwert	(arithmetic) mean
Median	Median
Ausreißer	Outlier
Stichprobenumfang	Sample size
schiefe Verteilung	Skewed distribution
zensierte Daten	Censored data

als der Median von zum Beispiel 25 U/l, die andere Hälfte größer. Häufig lassen sich aber schiefe Verteilungen durch eine geeignete Transformation in eine mehr oder weniger symmetrische Form umwandeln, wofür der Mittelwert dann wieder ein geeignetes Lagemaß darstellt (2, 3).

Bei der Betrachtung von Überlebenszeiten schließlich, das heißt bei Studien, in denen man sich für die Zeit bis zum Eintreten eines bestimmten Ereignisses (zum Beispiel Tod) interessiert, liegen typischerweise nicht für alle Patienten diese Zeiten vor, sondern es ist nur bekannt, dass innerhalb eines bestimmten Zeitraumes das interessierende Ereignis nicht eingetreten ist (zensierte Daten). In einer solchen Situation ist die Berechnung eines Mittelwertes nicht sinnvoll, während eine mediane Überlebenszeit spätestens dann angegeben werden kann, wenn die Hälfte der beobachteten Patienten gestorben ist (1).

kurzgefasst: **Der Median teilt eine Stichprobe in zwei gleiche Hälften. Er wird von extremen Werten (Ausreißern) praktisch kaum beeinflusst. Deshalb kann der Median zum Beispiel bei schiefen, unsymmetrischen Verteilungen (Laborwerte) oder bei der Betrachtung von Überlebenszeiten besser interpretiert werden.**

Für die Interpretation englischsprachiger Studien sind die Übersetzungen der wichtigsten in diesem Beitrag besprochenen Termini in **Tab. 2** aufgelistet.

Literatur

1 Altman DG, Bland JM. Quartiles, quintiles, centiles, and other quantiles. Brit med J 1994; 309: 996
2 Bland JM, Altman DG. Logarithms. Brit med J 1996; 312: 700
3 Bland JM, Altman DG. Transforming data. Brit med J 1996; 312: 770

Korrespondenz
Dr. Stefan Lange
Abteilung für Medizinische Informatik, Biometrie
und Epidemiologie Ruhr-Universität
Universitätsstraße 150
44780 Bochum
E-Mail: stefan.f.lange@ruhr-uni-bochum.de

Dieses Dokument wurde zum persönlichen Gebrauch heruntergeladen. Vervielfältigung nur mit Zustimmung des Verlages.