

Verlaufskurven

– Artikel Nr. 10 der Statistik-Serie in der DMW –

Serial data curves

Autoren

R. Bender¹ S. Lange¹

Institut

¹ Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen, Köln

Schlüsselwörter

- ▶ Verlaufskurven
- ▶ Messwertwiederholungen
- ▶ Kenngrößen
- ▶ Fläche unter der Kurve

Key words

- ▶ Serial data curves
- ▶ Repeated measurements
- ▶ Summary measures
- ▶ Area under the curve (AUC)

Bibliografie

DOI 10.1055/s-2007-959033
Dtsch Med Wochenschr 2007;
132: e22–e23 · © Georg Thieme
Verlag KG Stuttgart · New York ·
ISSN 0012-0472

Korrespondenz

Privatdozent Dr. rer. biol. hum.
Ralf Bender
Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG)
Dillenburger Straße 27
51105 Köln
eMail Ralf.Bender@iqwig.de

Ein häufiges Problem in der biomedizinischen Forschung ist die Analyse von Verlaufskurven. Eine Verlaufskurve ist eine zeitlich angeordnete Datenreihe, gemessen an einem Probanden, zum Beispiel die glomeruläre Filtrationsrate (GFR) über 2 Jahre im Abstand von 3 Monaten oder eine Blutzuckerkurve über 2 Stunden mit 5minütlichen Blutzuckermessungen. Kennzeichnend für solche Situationen ist also, dass es zu einer Vielzahl von Messwertwiederholungen kommt. Da zur Beurteilung und zum Testen auf Unterschiede von Verlaufskurven keine einfachen statistischen Standardverfahren zur Verfügung stehen, werden Verlaufskurven häufig inadäquat ausgewertet, zum Beispiel durch den Vergleich von Gruppen zu jedem einzelnen Messzeitpunkt.

Solche multiplen Vergleiche können zu inhaltlichen Problemen bei der Interpretation führen, wenn beispielsweise nur bei einem von 10 Messzeitpunkten ein Unterschied auffällt, oder wenn sich die Kurven überkreuzen. Bei den multiplen Vergleichen entstehen auch Probleme, wenn multiple Signifikanztests eingesetzt werden (zum Beispiel 10 verschiedene *t*-Tests), da bei einer solchen Vorgabe das Signifikanzniveau [2] nicht mehr eingehalten werden kann. Das heißt, die Irrtumswahrscheinlichkeit als Grundpfeiler der statistischen „Sicherheit“ ist nicht mehr quantifizierbar. Auf diese besondere Problematik des multiplen Testens und geeignete Lösungsmöglichkeiten werden wir in einem weiteren Beitrag dieser Reihe näher eingehen [4].

Je nach Fragestellung und Design kann eine angemessene Analyse von Verlaufskurven sehr komplex werden. Eine vollständige Übersicht über die vorhandenen statistischen Methoden zur Auswertung von Messwertwiederholungen kann hier nicht gegeben werden. Eine sehr gute Übersicht findet man zum Beispiel bei Cowder und Hand [5].

Neben komplizierten Verfahren, wie zum Beispiel Varianzanalyse für Messwertwiederholungen [1] oder Regressionsanalyse mit zufälligen Effekten, kann man oft auf einfachere Methoden zurückgreifen, um Verlaufskurven adäquat auszuwerten. Vor allem bei langen Verlaufskurven interessiert häufig nicht jeder einzelne Datenpunkt, sondern nur gewisse Charakteristiken der gesamten Kurve. Lässt sich der Informationsgehalt einer Kurve sinnvoll durch eine oder wenige Kenngrößen wiedergeben, so kann die Auswertung dadurch vereinfacht werden, dass nicht sämtliche Einzelpunkte, sondern lediglich diese Kenngrößen der Verlaufskurven analysiert werden [8]. Häufig verwendete Kenngrößen sind die Fläche unter der Kurve (AUC), bei monotonen Kurven die Steigung aus der einfachen linearen Regression [7] und bei eingipfligen Kurven das Kurvenmaximum und der Zeitpunkt des Kurvenmaximums.

Welcher Test zur weiteren Analyse der Kenngrößen geeignet ist, hängt von der Fragestellung und vom Studiendesign ab. Grundsätzlich kommen zur Analyse von Kurvenkenngrößen die üblichen parametrischen und nichtparametrischen Tests in Frage [3]. Man muss lediglich beachten, ob die zu vergleichenden Verlaufskurven in unabhängigen Gruppen (Verfahren für unabhängige Stichproben, zum Beispiel ungepaarter *t*-Test) oder ob mehrere Verlaufskurven beim gleichen Probanden (Verfahren für abhängige Stichproben, zum Beispiel gepaarter *t*-Test) gemessen wurden.

Als Beispiel betrachten wir die Serum-Insulin-Kurven von $n = 14$ Probanden, die in einer Studie zum Vergleich zweier Insulin-Präparate (Insulin-Analogon und Humaninsulin) gemessen wurden [6]. Nach 1 Stunde Vorlaufphase wurde zum Zeitpunkt 0 subkutan je ein Insulin-Präparat verabreicht und über 6 Stunden das Serum-Insulin gemessen. Insgesamt liegen pro Kurve 20 Mess-

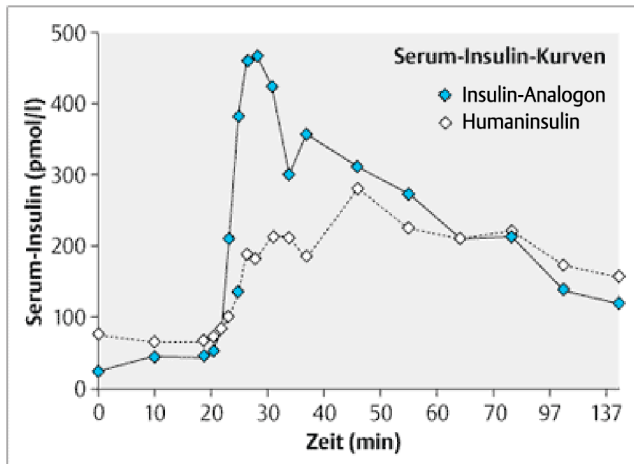


Abb. 1 Serum-Insulin-Kurven eines Probanden nach subkutaner Injektion von Humaninsulin und eines Insulin-Analogons.

werte vor. Jeder Proband erhielt an 2 verschiedenen Tagen beide Insulin-Präparate. In **Abb. 1** findet man beispielhaft die beiden Serum-Insulin-Kurven eines Probanden.

Eine sinnvolle Kenngröße zur Beschreibung der maximalen Wirkung der Insulin-Präparate ist das Kurvenmaximum (C_{\max}). In **Tab. 1** findet man die C_{\max} -Werte aller 28 Kurven. Es zeigt sich ein deutlicher Unterschied zwischen den beiden Präparaten bezüglich des Kurvenmaximums. Da es sich um abhängige Stichproben handelt und eine Normalverteilung der C_{\max} -Werte nicht angenommen werden kann, ist ein adäquater Test auf Unterschied zwischen den Kurven bezüglich C_{\max} , der Vorzeichenrangtest von Wilcoxon. Die Anwendung dieses Tests liefert einen signifikanten Unterschied zwischen den beiden Insulin-Präparaten bezüglich der maximalen Wirkung ($p < 0,001$).

kurzgefasst

Verlaufskurven sind zeitlich angeordnete Datenreihen gemessen an einem Probanden. Die richtige Auswertung kann anhand von komplexen statistischen Verfahren, zum Beispiel Varianzanalyse für Messwertwiederholungen oder Regressionsanalyse mit zufälligen Effekten aber auch anhand einfacherer Kenngrößen, zum Beispiel Fläche unter der Kurve, bei monotonen Kurven die Steigung aus der einfachen linearen Regression oder bei eingipfligen Kurven durch die Bestimmung des Kurvenmaximums bzw. den Zeitpunkt der Kurvenmaximums erfolgen.

Dieser Beitrag ist eine überarbeitete Fassung aus dem Supplement Statistik aus dem Jahr 2001.

Tab. 1 C_{\max} -Werte (in pmol/l) von Serum-Insulin-Kurven gemessen an $n = 14$ Probanden nach subkutaner Verabreichung eines Insulin-Analogons bzw. eines Human-Insulins [6].

| Proband | Insulin-Analogon | Humaninsulin | Differenz |
|-----------------|------------------|--------------|--------------|
| 1 | 510 | 345 | 156 |
| 2 | 428 | 212 | 216 |
| 3 | 595 | 329 | 266 |
| 4 | 456 | 300 | 156 |
| 5 | 444 | 258 | 186 |
| 6 | 538 | 344 | 194 |
| 7 | 472 | 281 | 191 |
| 8 | 516 | 293 | 223 |
| 9 | 476 | 201 | 275 |
| 10 | 785 | 286 | 499 |
| 11 | 585 | 321 | 264 |
| 12 | 490 | 294 | 196 |
| 13 | 442 | 261 | 181 |
| 14 | 515 | 355 | 160 |
| Mittelwert (SD) | 518.0 (92.0) | 291.4 (47.0) | 226.6 (87.6) |
| Median | 500 | 293.5 | 195 |

Literatur

- Bender R, Grouven U, Ziegler A. Varianzanalyse für Messwertwiederholungen. Dtsch Med Wochenschr 2007; 132: e61–e64
- Bender R, Lange S. Was ist der p -Wert? Dtsch Med Wochenschr 2007; 132: e15–e16
- Bender R, Lange S, Ziegler A. Wichtige Signifikanztests. Dtsch Med Wochenschr 2007; 132: e24–e25
- Bender R, Lange S, Ziegler A. Multiples Testen. Dtsch Med Wochenschr 2007; 132: e26–e29
- Crowder MJ, Hand DJ. Analysis of Repeated Measures. Chapman & Hall: London, 1990
- Heinemann L, Heise T, Jorgensen LN, Starke AAR. Action profile of the rapid acting insulin analogue. Human Insulin B28Asp. Diabetic Med 1993; 10: 535–539
- Lange S, Bender R. Lineare Regression und Korrelation. Dtsch Med Wochenschr 2007; 132: e9–e11
- Matthews JNS, Altman DG, Campbell MJ, Royston P. Analysis of serial measurements in medical research. BMJ 1990; 300: 230–235