

PRO EXPERIMENTIS

Zur Schätzung des kleinsten zulässigen Stichprobenumfanges für stereologische Messungen an histologischen Schnitten

Das in der Statistik gebräuchliche Verfahren zur Schätzung des erforderlichen Stichprobenumfanges beruht auf dem Begriff des Mutungintervales (confidence interval)¹⁻³. Dieses Verfahren wird auch für stereologische Messungen an histologischen Schnitten angewendet³, wenn es darum geht zu entscheiden, wie viele Bildfelder ausgemessen werden sollen. Der Umfang der Stichprobe wird dabei – vereinfachend gesagt – so gross gewählt, dass die Streuung der verschiedenen möglichen Stichprobenmittelwerte um den wirklichen Mittelwert der Grundgesamtheit eine festgesetzte Grenze nicht überschreitet. Soll diese Grenze höchstens 5% des Mittelwertes der Grundgesamtheit oder weniger betragen³, so ergibt sich häufig ein Stichprobenumfang von einigen Hunderten bis Tausenden von auszumessenden Bildfeldern. Kann diese Forderung nicht erfüllt werden, so wird der Stichprobenumfang gesenkt, indem man eine grössere Irrtumswahrscheinlichkeit in Kauf nimmt³.

Werden aus einem Organ mehrere gleich strukturierte Regionen untersucht, so bildet das Messergebnis der Schnitte aus einer einzelnen Region eine Teilstichprobe, das Gesamtergebnis aus den Messungen aller Regionen die Gesamtstichprobe. Den optimalen Umfang der Gesamtstichprobe wird man nach dem üblichen Verfahren mit Hilfe des Mutungintervales schätzen¹⁻³. Es wäre jedoch unzweckmässig, die Forderung nach dem optimalen Umfang der Stichprobe auch auf die einzelnen Teilstichproben anzuwenden, da der so errechnete Umfang der Teilstichproben zu einem Umfang der Gesamtstichprobe führen würde, der das Optimum weit überschreitet, und einen unnötigen Aufwand an Zeit und Arbeit erfordern würde. Ausserdem würde damit die Wahrscheinlichkeit erhöht, dass beim statistischen Vergleich der Teilstichproben untereinander auch praktisch irrelevante Unterschiede als statistisch signifikant erscheinen würden¹.

Soll daher der Umfang der einzelnen Teilstichproben verringert werden, so erhebt sich die Frage: Welches ist der kleinste Umfang der Teilstichproben, der auf keinen Fall unterschritten werden sollte? Im folgenden wird ein Verfahren vorgeschlagen, das erlaubt, den kleinsten zulässigen Umfang der Teilstichproben zu schätzen.

Als Ausgangspunkt dient eine Überlegung von SITTE⁴: Er schlägt vor, während der Messung fortlaufend den Mittelwert aus der Gesamtheit der bisherigen Messungen zu bilden und die Messungen so lange fortzuführen, bis die neu hinzukommenden Messwerte nur noch geringfügige Schwankungen verursachen.

Um einen Begriff zu gewinnen, zu welcher Grössenordnung des Stichprobenumfanges dieses Verfahren führt, wurde ein Kleincomputer (Olivetti Programma 101) so programmiert, dass er anhand von Zahlenmaterial aus der stereologischen Ausmessung histologischer Schnitte durch fortlaufendes Bilden der Mittelwerte denjenigen Stichprobenumfang bestimmte, von dem an die Schwankung der Mittelwerte unter einem geforderten Prozentsatz blieb. Dabei zeigte sich, dass der nach diesem Verfahren bestimmte Mindestumfang der Stichproben einer einfachen linearen Funktion in Abhängigkeit von der relativen Standardabweichung der Stichprobe folgen muss.

Um diese Funktion mathematisch zu erfassen, müssen wir strenger definieren, was wir von der kleinsten zulässigen Stichprobe fordern: Der Umfang der kleinsten zulässigen Stichprobe soll aus einer so grossen Anzahl von Messwerten bestehen, dass das zufällige Auftreten eines

Extremwertes in der letzten Messung den Mittelwert der bisherigen Messungen um nicht mehr als einen gewissen zugelassenen Prozentsatz verändern kann. Diese Forderung können wir so schreiben:

$$\frac{(n \cdot M) + \delta}{n} - M = \frac{M \cdot p}{100}, \quad (1)$$

wobei n = Anzahl der Messwerte

M = Mittelwert der Stichprobe

δ = Differenz zwischen Mittelwert und Extremwert

p = Anzahl Prozent der maximal zugelassenen Veränderung.

Nach n aufgelöst, ergibt diese Gleichung:

$$n = \frac{\delta \cdot 100}{M \cdot p}. \quad (2)$$

Für die Differenz δ könnte man den extremsten im Verlauf der Messungen gefundenen Wert verwenden, zweckmässiger jedoch wird man ein Mehrfaches f der (empirisch ermittelten) Standardabweichung σ der Stichprobe einsetzen. Für $f = 3$ wäre – entsprechend der t -Tabelle von Student – die Wahrscheinlichkeit, dass einmal ein grösserer Extremwert auftritt, höchstens 1%, sofern die Stichprobe wenigstens 14 Messungen umfasst, für $f = 2$ im ungünstigsten Fall ca. 5%. Für das Ausmass der in Prozent ausgedrückten Schwankung p der Mittelwerte ist es, wie sich weiter unten zeigen wird, am zweckmässigsten, eine Grenze von 1% (bis höchstens 1,5%) festzusetzen. Wir erhalten damit für den kleinsten zulässigen Umfang n der Stichprobe die Formel

$$n = \frac{f}{p} \cdot \frac{\sigma \cdot 100}{M}, \quad (3)$$

wobei die Faktoren f und p die gewünschten Sicherheitsschwellen für den Extremwert und die Schwankung der Mittelwerte ausdrücken. Wir können also sagen: Der kleinste zulässige Stichprobenumfang n ist gleich der in Prozent ausgedrückten relativen Standardabweichung der Stichprobe ($\sigma \cdot 100/M$), multipliziert mit dem Quotienten aus den gewünschten Sicherheitsschwellen.

Wählen wir für den extremsten im Verlauf der Messungen zu erwartenden Wert 3 σ (wobei wir höchstens in 1% der Fälle fehlgehen werden) und fordern wir, dass die durch diesen Extremwert verursachte Veränderung des Mittelwertes höchstens 1% betragen dürfe, so erhalten wir die Formel

$$n = 3 \frac{\sigma \cdot 100}{M}, \quad (4)$$

¹ J. PFANZAGL, *Allgemeine Methodenlehre der Statistik*, 3. Aufl. (W. de Gruyter, Berlin 1968).

² L. SACHS, *Statistische Auswertungsmethoden* (Springer, Berlin-Heidelberg-New York 1968). – O. W. HASELOFF und H. J. HOFFMANN, *Kleines Lehrbuch der Statistik*, 3. Aufl. (W. de Gruyter, Berlin 1968).

³ R. T. DE HOFF, Proc. 2nd Int. Congr. Stereology, Chicago 1967 (Springer, Berlin-Heidelberg-New York 1967), p. 119.

⁴ H. SITTE, in *Quantitative Methods in Morphology* (Ed. E. R. WEIBEL und H. ELIAS; Springer, Berlin-Heidelberg-New York 1967), p. 167.

d. h. der kleinste nach unseren Forderungen noch zulässige Umfang n einer Teilstichprobe ist gleich dem dreifachen Wert der in Prozent ausgedrückten relativen Standardabweichung dieser Stichprobe. Nach dieser Regel lässt sich der kleinste zulässige Umfang der Stichproben auf einfache Weise abschätzen. Ist beispielsweise der Mittelwert einer Stichprobe $M = 60$ und die Standardabweichung $\sigma = 10$, so beträgt der kleinste zulässige Umfang n der Stichprobe nach Formel (4) $n = 3(10 \cdot 100/60) = 50$, d. h. es müssen für jede Teilstichprobe (also für jede der untersuchten gleichstrukturierten Regionen eines Organs) mindestens 50 Bildfelder ausgemessen werden.

Sollte es aus technischen Gründen nötig sein, kann man den Umfang der Teilstichproben noch weiter verringern, indem man eine grösse Veränderung des Mittelwertes beim zufälligen Auftreten eines Extremwertes zulässt oder sich mit der geringeren Sicherheitsschwelle von 2σ für den eingesetzten Extremwert begnügt. Da nach Formel (3) f und p in reziproker Verhältnis zueinander stehen, hat ein Erhöhen der tolerierten Veränderung des Mittelwertes auf 1,5% die gleiche Wirkung wie ein Senken des eingesetzten Extremwertes auf 2σ . Diesen Wert sollte man in keinem Fall unterschreiten, da sonst, wie die t -Tabelle von Student zeigt, die Irrtumswahrscheinlichkeit zu stark anwächst. Für das oben erwähnte Beispiel würde sich dabei ein unterster für den Umfang der Stichprobe noch zulässiger Wert von $n = 33$ ergeben.

Anderseits wird das Überschreiten eines Wertes von 4 für den Faktor f/p im Interesse einer höheren Messgenauigkeit der Teilstichprobe wenig sinnvoll sein, da dann die Irrtumswahrscheinlichkeit bereits unter 1% gesunken ist und eine weitere Vermehrung der Messungen nur noch eine sehr geringe Steigerung der Sicherheit mit sich bringt.

Wie sich der optimale Umfang der Gesamtstichprobe mit Hilfe des Mutungintervall schätzen lässt, so ermöglicht das angegebene Verfahren, für die Teilstichproben den kleinsten Umfang zu bestimmen, der den gestellten Forderungen noch genügt.

Summary. If, for technical reasons, the sample size of stereological measurements has to be kept as small as possible, a method is given to determine the smallest number of measurements tolerable. It should never be less than twice the relative standard deviation of the sample, expressed as percentage.

R. BAUR

Anatomisches Institut der Technischen Hochschule Aachen (Deutschland), c/o Anatomisches Institut der Universität Basel (Schweiz), 11. Februar 1969.

A New Method for Determination of α -Amylase

There are various methods available for the determination of the enzyme, α -amylase, based on (a) decrease in viscosity of starch solution, (b) decrease in turbidity of starch suspension, (c) decrease in the starch-iodine colour intensity and (d) the increase of reducing groups.

In the past there were many attempts to simplify the native substrate for α -amylase, but so far this approach has met with only partial success. For example, α -*p*-nitrophenyl maltosides have been used as substrates for amylases¹⁻³. However, it was found that α -amylases exhibit lower affinity (by several orders of magnitude) for the low molecular weight substrates. It is therefore assumed that α -amylase requires many points of attachment to long chain substrate molecules to obtain full catalytic activity.

A new method for the determination of α -amylase activity was reported by RINDERKNECHT et al.⁴. Starch grains were labelled covalently with a coloured substance, Remazolbrilliant Blue R. A suspension of such material was then incubated with an α -amylase source for a given length of time. The reaction was terminated by addition of diluted acetic acid. The mixture was then filtered and the filtrate was assayed colorimetrically.

For some years, we have been working on the development of an α -amylase method based on the use of new types of substrates. The synthesis of these substrates involves essentially 2 major steps. First, soluble starch was coloured under alkaline conditions using Cibacron blau F 3 G-A. This procedure resulted in the formation of covalent bonds between starch molecules and dye molecules. Second, the coloured starch was cross-linked by the addition of 1,4-butandiolglycidether. This gave a three-dimensional, insoluble network which swells in water. The degree of swelling could be controlled by the

amount of cross-linking. The enzymatic hydrolysis of such insoluble starch derivatives yields soluble starch split products carrying the coloured marker.

The resulting blue starch polymer is homogenized in a Waring Blender to small particle size and washed several times in distilled water and several times in 0.02 M Phosphate Buffer (pH 7.0) containing 0.01 M NaCl. This blue cross-linked starch polymer in the form of small gel grains is then resuspended in the above mentioned Phosphate buffer.

For testing, 1 ml of blue starch polymer suspension (containing 0.143% of dry polymer) was used. To 1 ml samples, 20 μ l portions of crude porcine pancreas extract were added. The α -amylase extract was prepared from 10 mg of porcine pancreatin powder from ORGANON, BATCH 32 A/59, which was suspended in 10 ml of the above phosphate buffer. The mixture was stirred for 5 min at room temperature and centrifuged at 5000 rpm for 10 min. The clear supernatant was the source of crude α -amylase extract.

The enzymatic reaction was performed at 37°C for 0, 10, 20 and 30 min. The samples were shaken during the incubation time. The enzymatic hydrolysis was terminated by placing the tubes in a boiling water bath or

¹ S. AKABORI, K. UEHARA, Y. SHIMAZU and K. NAKANISHI, J. chem. Soc. Japan 64, 1046 (1943).

² S. MATSUBARA, T. IKENAKA and S. AKABORI, J. Biochem., Tokyo 46, 425 (1959).

³ A. P. JANSEN and P. G. A. B. WYDEVELD, Nature 182, 525 (1958).

⁴ H. RINDERKNECHT, P. WILDING and B. J. HAVERBACH, Experientia 23, 805 (1967).