

Medicina Generalis Helvetica

«Check-up des Check-ups» – Qualitätssicherung oder Schildbürgerstreich?

Ein Labor-Resultat bedeutet eine Wahrscheinlichkeitsausgabe; aus einer Vortest-Wahrscheinlichkeit (pre-test likelihood) gewinnt man mit einem Labor-Test eine Nachtest-Wahrscheinlichkeit (post-test likelihood), mit der eine vermutete Krankheit vorliegt. Der Wahrscheinlichkeits-Charakter von Labor-Resultaten ist in einer Reihe von unvermeidlichen «Fehlern» begründet. Fehler, die nicht mit einem fehlerhaften Gebrauch der Testmethode verwechselt werden dürfen.

Als Betreiber eines kleinen und sparsam eingesetzten Praxislabors habe ich bereits vor einigen Jahren an sogenannten Ringversuchen zur «Qualitätskontrolle» teil-

genommen – etwas unüberlegt dieser zunächst als sinnvoll erscheinenden Modetorheit folgend. Die jeweils zugeschickte graphische Auswertung der Ergebnisse eines solchen Ringversuches liess dann aber den biometrisch-statistischen Trugschluss hinter dieser angeblichen «Qualitätssicherung» erkennen. Diese Graphik zeigte nämlich sehr schön eine Resultatbeziehungswise Noten-Verteilung, wie sie durch ein Würfelspiel zustandekommt; wir kennen diese Verteilung aus der Mathematik als «Gaussche Normalverteilung» (wobei dieses «normal» nicht mit «physiologisch normal» verwechselt werden darf).

Die Auszählung von Blutzellen ist in der Tat ein Würfelspiel. Wenn wir kunstgerecht das Blut entnommen, genauestens verdünnt, gefärbt, geschüttelt und richtig in die Zählkammer eingefüllt haben, wird «gewürfelt»: Vier Felder werden nach dem Zufallsprinzip ausgezählt (und die Zählung wird hochgerechnet). Auch wenn die Laborantin alles fehlerfrei macht, kann sie den stochastischen «Fehler» des Zufalls (random error) nicht vermeiden. Mit der Zufalls-Wahrscheinlichkeit von 68 Prozent wird das Resultat innerhalb einer sogenannten Standardabweichung (standard error) vom richtigen Wert (Sollwert) liegen, mit 95 Prozent innerhalb von zwei Standardabweichungen; mit fünf Prozent wird der fehlerfrei durchgeführte Test ein Resultat ergeben, das über zwei Standardabweichungen vom richtigen Wert abweicht, wobei mit einem Prozent Wahrscheinlichkeit sogar

ein solcher stochastischer Fehler von mehr als 2,5 Standardabweichungen möglich ist.

In Ringversuchen wird der Sollwert einer hämatologischen Probe durch manuelles Zählen zwar durch «erfahrenes Personal» ermittelt (wie mir versichert wurde), aber auch dieses erfahrene Personal hat keine Möglichkeit, diesen Zufallsfehler zu verhindern; einzelne Blutkörperchen-Auszählungen müssen auch hier entsprechend der stochastischen Gesetzmässigkeiten stark vom richtigen Wert abweichen. Der Zufalls-Fehler kann natürlich dadurch verringert werden, dass eine grössere Zahl Auszählungen vorgenommen wird. Damit wird oder würde aber der Referenzwert im Ringversuch auf eine praxisfremde Art und Weise erzeugt, und jede teilnehmende Praxis müsste dann gleich verfahren, um einen fairen Qualitätsvergleich überhaupt möglich zu machen. Der Ringversuch-Betreiber ist im Prinzip gar in der Lage, einen «wahreren» Wert zu ermitteln als ein teilnehmendes Praxislabor. Der wahre Wert wird am genauesten über den Durchschnittswert der teilnehmenden Praxislabors ermittelt. Wie bereits gesagt, kommt ein mehrfaches (unabhängiges) Zählen der gleichen Probe im Referenzlabor näher an den wahren Wert (wie Ringversuch-Betreiber versichern, liegt der Mittelwert aus allen teilnehmenden Praxislabors jeweils «sehr nahe am wahren Wert»), es entspricht aber einem Überlegungsfehler, wenn das zentrale Referenzlabor in einem Ringversuch den

«Check-up des Check-up» – Qualitätssicherung oder Schildbürgerstreich?

wahren Wert gepachtet zu haben glaubt, weil – wie gesagt – auch dieser durch Zufalls-Fehler «gestört» wird.

Der Zufalls-Fehler ist aufgrund des zufallsbedingten Zählverfahrens bei hämatologischen Bestimmungen von einiger Bedeutung, die «Normal-Streuung» von Testresultaten bei kunstgerechter Durchführung hängt indes – je nach Testverfahren – von einer Reihe weiterer unvermeidbarer Fehler ab, auf die hier nicht weiter eingegangen werden soll (dazu kommen auch noch zum Teil erhebliche natürliche Schwankungen im Patientenblut). So wie auch ein aufwendigeres Zählen von viel mehr Gesichtsfeldern diese Streuung (gemessen am Variationskoeffizienten [= Standardabweichung : Mittelwert]) verschmälern und somit präziser machen kann, kann mit aufwendigeren Anlagen in Grosslabors bei Ringversuch-Betreibern die Präzision bei vielen Analysen möglicherweise etwas erhöht werden, was aber wiederum zu einer unfairen Kontrolle des Praxislabors führt. Andererseits führen Fehlmanipulationen oder «unsauberes» Arbeiten zu einer breiteren Streuung der Verteilungskurve von Laborresultaten.

Entscheidend ist erstens, wieweit solche «wirkliche» Fehler durch den Ringversuch von den vorher genannten unvermeidlichen Fehlern unterschieden werden können und zweitens, ob der mögliche Grenznutzen in der Verbesserung der Labor-Präzision den technischen und personellen Aufwand rechtfertigt. Wie wir gesehen haben, muss nur schon der Zufall in einzelnen Fällen zu grotesken Falsch-Resultaten führen, und viele eventuell verbesserungsbedürftige Manipulationsfehler führen andererseits oft nur zu

relativ kleinen Fehlern. Die Varianz durch unvermeidbare Fehler bei Ringversuchen ist bei hämatologischen Laboruntersuchungen sicher grösser als die Varianz durch vermeidbare Manipulationsfehler, und dies dürfte bei den meisten Analysen der Fall sein. Ein Ringversuch kann deshalb «wirkliche» Fehler eines Praxislabors nicht von unvermeidbaren «Fehlern» unterscheiden; anhand der Streubreite der Resultate kann lediglich vage festgestellt werden, ob mehr oder weniger Fehler irgendwo im untersuchten System aufgetreten sind, jedoch ohne diese Fehler orten zu können. Qualitätssicherung von individuellen Praxislabors durch Ringversuche ist deshalb eine Fiktion! Das Qualitäts-Attest in Form der üblichen Benotung der teilnehmenden Praxislabors ist nun ein wirklicher (Denk-) Fehler, der vermeidbar wäre und kaum eine gute Note verdient.

Interpretieren und richtiges Handeln – das ist Qualität

Ein Laborwert entspricht also schon rein messtechnisch praktisch nie dem wahren Patienten-Wert; beim Quick-Wert etwa spricht es sich langsam herum, dass wir eigentlich kaum wissen, was wir da messen. Es ist in meinem Praxislabor schon vorgekommen, dass ich mit Quick-Werten konfrontiert worden bin, die ich nicht glauben konnte, und auch einer zweiten Bestimmung konnte ich keinen Glauben schenken. Die erneute Blutentnahme und Weitergabe der Probe in ein benachbartes Praxislabor ergab schliesslich einen «vernünftigeren» Wert. Aus dem Wissen um die Fehlerhaftigkeit von Laborergebnissen das Richtige tun, das ist Qualität! Wir konnten nicht herausfinden, ob im eigenen Labor ein Fehler passiert war und welche Fehlmanipulation in Frage gekommen wäre. Hätten wir in diesem Moment auf das falsche Würfelspiel des Ringversuchs als «Qualitätsgarantie» gesetzt, wäre eine gefährliche Behandlung die Folge gewesen.

Ein bedeutendes Qualitätsproblem liegt in

der mangelnden Kunst der Interpretation von Laborresultaten. Dieser Mangel besteht im Übersehen des Umstandes, dass auch die biologische Aussagekraft eines Laborwertes weiteren Fehlermöglichkeiten unterworfen ist. Das Labor ist da, um einen klinischen Verdacht zu bestätigen; eine Bestätigung durch ein positives Resultat kann von einer noch zweifelhaften Vortest-Wahrscheinlichkeit zu einer relativ sicheren Nachttest-Wahrscheinlichkeit führen, aber nur bei einem von vornherein relativ grossen Krankheitsverdacht. Auch das technisch perfekte Labor taugt aber nicht zum «Durch-checken» (1) oder «Absichern» (2)! Hier machen falsch positive und falsch negative Ergebnisse, wiederum durch die Zufalls-Mathematik bedingt, eine Laborbestimmung wertlos. Im Check-up treten meist rein zufälligstochastisch bedingte «Pathologien» auf, die oft als krankhafte Befunde fehlinterpretiert werden. Nicht zuletzt darauf dürfte die Nutzlosigkeit des Check-ups (3) zurückzuführen sein. Die einfache Tatsache, dass falsch positive Fehler sich bei geringer Krankheits-Wahrscheinlichkeit multiplizieren und die Test-Aussage stark reduzieren, ist aber den meisten Ärzten nicht bekannt, denn wir hören in unserer Ausbildung kaum etwas davon (es sei auf unseren Einsiedler Kurs «Kritische Beurteilung medizinischen Wissens» vom 25. bis 27. August 1995 verwiesen, in welchem eine moderne, kunstgerechte Beurteilung von Testresultaten gelernt werden kann).

Eine Laborbestimmung muss in einem relevanten biologischen Informationsgewinn resultieren; das heisst es muss ein klinisch relevanter Unterschied zwischen Vortest- und Nachttest-Wahrscheinlichkeit erwartet werden können, damit ein Labortest gerechtfertigt ist. In der Praxis besteht heute indes eine weitverbreitete irrationale Laborgläubigkeit (das irreführende Credo «harter» Daten (4)); angesichts der naiven Verkennung der oft sehr geringen «biologischen Präzision» von Labordaten sind Anstrengungen zur Verbesserung der messtechnischen Präzision eben eine Geschichte aus Schilda. Wieso soll denn soviel Aufmerksamkeit

Erlernen einer anderen, kunstgerechten Beurteilung von Testresultaten

Einsiedler Kurs

«Kritische Beurteilung medizinischen Wissens» vom 25. bis 27. August 1995.

«Check-up des Check-up» – Qualitätssicherung oder Schildbürgerstreich?

auf die Präzision von Laborresultaten in der Praxis gerichtet werden, wenn deren biologisch-klinische Aussagekraft oft nicht klar ist und in der Praxis häufig verkannt wird? Neben der verbesserungsbedürftigen Interpretations-Qualität nimmt sich die angeblich verbesserungsbedürftige technische Labor-Qualität, der nun mit obligatorischen Massnahmen zu Leibe gerückt werden soll, als Bagatelle aus.

Fragwürdige Ringversuche

Obligatorische Ringversuche beruhen auf Fehlüberlegungen. Auch wenn Qualitätskontrollen nicht in jedem Fall abzulehnen sind, muss eine billige, modische Propagierung undurchdachter Pseudo-Aktivitäten vermieden werden, die von wichtigeren Fragen abhalten. Die für unseren Beruf viel wesentlichere Frage ist die einer rationalen Verwendung und Interpretation von Laboruntersuchungen. Es ist zu hoffen, dass interkollegiale Qualitätsanstrengungen («peer review») die heu-

tige Vorliebe für administrativ-bürokratische «Qualitätssicherung» ersetzen werden. Vielleicht gilt es, von Ärzteseite das neue KVG zu nutzen, um neue, durchdachtere und wahrhaft qualitätsfördernde Versicherungen und Honorierungs-Systeme auf die Beine zu stellen.

Auch wenn dieser Beitrag zuerst auf Verwunderung oder sogar Ablehnung bei den Befürwortern der Qualitätskontroll-Mode stossen könnte, so kann man doch davon ausgehen, dass die Berechtigung dieser Kritik mit der Zeit erkannt werden wird; ich hoffe, es dauert nicht allzu lange. Nicht nur die Ärzteschaft sollte wegen unwirksamer Therapien kritisiert werden, auch Versicherer sollten Trugschlüsse erkennen und Wirksamkeit und Unbedenklichkeit ihrer Massnahmen belegen müssen.

Literatur

1. *Bucher H, Gutzwiller F*: Checkliste Gesundheitsberatung und Prävention, Stuttgart: Georg Thieme Verlag, 1993.

2. *Schmidt J.G.*: Früherkennung und Umgang mit Risikofaktoren. In: Kochen MM, ed. Allgemeinmedizin. Stuttgart: Hippokrates Verlag, 1992: 133-149.
3. *Friedman G.D.* et al.: Multiphasic health check-up evaluation: A 16-year follow-up. *J Chronic Dis* 1986; 39: 453-463.
4. *Feinstein A.R.*: The hard data creed in current clinical practice: Its spurious validity and the challenge to define meaningful clinical variables. In: *Schmidt J.G., Steele R.E.* eds. Kritik der medizinischen Vernunft. Verlag Kirchheim, Mainz 1994: 210-218 (Deutsch: Das irreführende Credo harter Daten in der heutigen klinischen Praxis – das Problem sinnvoller klinischer Erfolgskriterien. *Forsch Komplementärmed* 1994; 1: 260-265).

Anschrift des Verfassers: Dr. med. Johannes G. Schmidt, Allgemeinpraxis und Institut für Klinische Epidemiologie, 8840 Einsiedeln.

8. Zürcher Kinder- und Jugendpsychiatrisches Symposium

Thema: Gewalt und Delinquenz bei Kindern und Jugendlichen
 Datum: Donnerstag, 28. September 1995
 Ort: Universitätsspital Zürich; Frauenklinik: grosser Hörsaal D Nord, Frauenklinikstrasse 22
 Zeit: 9.00 bis 12.00 Uhr und 14.00 bis 16.30 Uhr
 Teilnahmegebühr: Fr. 120.– pro Person inklusive Zwischenverpflegung in den Kaffeepausen

Kontaktadresse und Detailinformation:

Psychiatrische Universitäts-Poliklinik für Kinder und Jugendliche, Sekretariat Ärztliche Direktion, Postfach, 8028 Zürich, Tel. 01/251 96 94 (Direktwahl).

Impressum

Medicina Generalis Helvetica
 (SGAM-Bulletin/Bulletin de la Société Suisse de Médecine Générale/Bolletino della Società Svizzera di Medicina Generale)

Redaktor

Dr. med. M. Peltenburg, Hinwil

Redaktionskommission

Dr. med. R. Böni, Madiswil
 Dr. med. B. Croisier, Grône
 Dr. med. R. L. Meyer, Möhlin

Sekretariat SGAM/SSMG und Redaktionssekretariat

Luzia Schneider, Oberplattenstrasse 73, 9620 Lichtensteig,
 Tel: 074/7 66 40, Fax: 074/7 66 41