

Tumormarker-Profilauswertung – und aus Messdaten wird Information

Analysis of tumor marker profiles – turning measured data into information

Norman Bitterlich*

Medizin & Service GmbH, Chemnitz, Deutschland

Zusammenfassung

In die Tumormarker-Profilauswertung fließen in der praktischen Anwendung die Erfahrungen der Einzelmarker-Interpretation ein. Aus mathematischer Sicht eröffnen sich dabei verschiedene Möglichkeiten, die einzelnen Messwerte miteinander zu verknüpfen. Vergleicht man solche Lösungsansätze, sind wesentliche Unterschiede in der Leistungsfähigkeit zu beobachten. Im Allgemeinen geht ein Zuwachs an Sensitivität zulasten der Spezifität. Die Auswahl des mathematischen Verfahrens wird deshalb entscheidend von den subjektiven Erwartungen geprägt.

Ziel der Tumormarker-Profilauswertung ist eine Sensitivitätssteigerung bei Erhalt einer hohen Spezifität. Ausgehend von der Analyse häufig verwendeter Strategien der Datenauswertung werden Effekte beschrieben, aus denen sich ein allgemeingültiges Konzept der Datenverknüpfung ableiten lässt. Besondere Beachtung findet dabei die Unabhängigkeit der mathematischen Verfahren von konkreten Cut-offs und damit von diagnostischen Testmethoden. Anstelle quantitativer Messwerte werden qualitative Bewertungen der Entscheidungssicherheit verarbeitet. Auf diese Weise gelingt nicht nur eine Sensitivitätserhöhung, sondern die Ergebnisse lassen sich auch plausibel aus Anwendersicht interpretieren. Die Verarbeitung der Tumormarker-Profile erscheint in diesem Sinne (fuzzy) logisch und eröffnet neue Chancen, mit den Laboruntersuchungen zur Klärung klinischer Fragestellungen weitreichend beizutragen.

Schlüsselwörter: Entscheidungshilfesystem; fuzzybasierte Mustererkennung; Sensitivitätserhöhung; Tumormarker-Profilauswertung.

Abstract

In practical use, the experiences of single marker interpretation influence the analysis of tumor marker profiles. From a mathematical point of view, there are different possibilities to link individual measured data. When comparing such methods of resolution, significant differences in their efficiency are observed. An increase in sensitivity generally results in a loss of specificity. Therefore, the choice of the mathematical process is decisively influenced by subjective expectations.

The goal of the analysis of tumor marker profiles is raising the sensitivity while maintaining a high specificity. Based on the study of frequently used strategies of data evaluation, effects are described from which a universally valid concept of data evaluation can be derived. Special attention is given to the independence of the mathematical processes from specified cut-offs and thereby from diagnostic testing methods. Instead of quantitative measured data, qualitative valuations of decision security are handled. This not only enables a rise in sensitivity but also allows a plausible interpretation of results from the user's view. Thus the analysis of tumor marker profiles seems (fuzzy) logical and opens up new avenues to make far-reaching contributions to the clarification of clinical problems by laboratory analysis.

Keywords: analysis of tumor marker profiles; decision support system; fuzzy based pattern recognition; increase of sensitivity.

Einführung

Trotz aller kritischer Diskussion über den effektiven Einsatz von Tumormarkern bieten diese Laboruntersuchungen bei korrekter Anwendung eine wertvolle Ergänzung für die Tumordiagnostik oder die Verlaufskontrolle. Auch wenn Tumormarker den durch Bildgebung oder Endoskopie erhobenen Befund zur Diagnosestellung nicht ersetzen können, vermögen sie in nicht-invasiver Weise die Diagnose zu bestätigen und zusätzliche Information für die Therapieentscheidung zu geben. Eine Möglichkeit, der häufig nicht zufriedenstellenden Sensitivität von einzelnen Tumormarkern entgegenzuwirken, ist in der

*Korrespondenz: Dr. rer. nat. Norman Bitterlich, Medizin & Service GmbH, Zwickauer Str. 227, 09116 Chemnitz, Deutschland
Tel.: +49-371-4660751
Fax: +49-371-4660752
E-mail: norman.bitterlich@t-online.de

Auswertung von Marker-Kombinationen zu sehen. Der gezielte Einsatz von Markern erster Wahl und Ersatzmarkern steigert die Entscheidungssicherheit [1].

Grundlage für die Leistungssteigerung ist jedoch ein differenziertes Bewerten der Einzelinformationen. Die Übertragung der einfachen Entscheidungsfindung beim Einzelmarker durch Vergleich mit einem Schwellenwert (Cut-off) auf den multiparametrischen Einsatz eröffnet vielfältige Bewertungsstrategien. Um die teilweise noch bestehenden Probleme der Vergleichbarkeit von Tumormarker-Tests verschiedener Hersteller nicht durch die weitere Verarbeitung der Messdaten zu verschärfen, sollten die mathematischen Methoden solche Unterschiede möglichst eliminieren können. In diesem Beitrag werden deshalb einige Aspekte der multiparametrischen Tumormarker-Auswertung vorgestellt und es wird auf Lösungsansätze hingewiesen, um die Entwicklung von leistungsfähigen Testmethoden durch eine adäquat erfolgreiche Verarbeitungsstrategie zu fördern.

Datenmaterial

Anstelle abstrakter Formeln werden die mathematischen Aspekte anhand eines konkreten Datensatzes dargestellt, der am Institut für Klinische Chemie (Dr. Petra Stieber) in Kooperation mit der Chirurgischen Klinik (PD Dr. Rudolf Hatz) des Klinikums München-Großhadern der Ludwig-Maximilians-Universität München erhoben wurde. Er resultiert aus Untersuchungen, in denen die Leistungsfähigkeit des Tumormarkers Pro-Gastrin-Releasing Peptide (ProGRP; ALSI, Tokio) zur Differenzierung kleinzelliger (SCLC) und nichtkleinzelliger (NSCLC) Bronchialkarzinome analysiert und – neben der Behandlung anderer wissenschaftlicher Fragestellungen – mit dem etablierten Marker Neuronspezifische Enolase (NSE; ROCHE Diagnostics, Mannheim) verglichen wurde.

Insgesamt standen Daten von 263 Patienten mit primär diagnostizierten Bronchialkarzinomen zur Verfügung. Darunter waren 198 Fälle mit NSCLC (75,3%) und 65 Fälle mit SCLC (24,7%). Die klinische Fragestellung zielte auf die Erkennung kleinzelliger Tumore ab, ein Befund "SCLC" gilt in diesem Sinne im Folgenden als positiv, der Befund "NSCLC" entsprechend als negativ. Die Vergleichsgruppe zur Festlegung von Referenzwerten besteht in diesem Kollektiv somit aus den Patienten mit einem nachgewiesenen nichtkleinzelligen Tumor.

Im vorliegenden Beitrag wird nicht die klinische Beschreibung von Einzelfällen des Patientenkollektivs oder die prinzipielle Bedeutung der Tumormarker als diagnostisches Verfahren für diese Indikation diskutiert. Die konkreten Daten dienen lediglich der exemplarischen Demonstration und können auf ähnliche Fragestellungen und anderes Datenmaterial übertragen werden.

Tumormarker mit Diagnostikpotenzial

Die Diagnoseunterstützung eines Laborparameters ist besonders dann effektiv, wenn ein Vergleich des ermit-

telten Messwertes mit einem festgelegten Cut-off eine statistisch gesicherte Zuordnung zur Erkrankung ermöglicht. So gelten erhöhte Werte von Tumormarkern stets als weiter abzuklärende Hinweise auf malignes Geschehen. Im Allgemeinen wird die diagnostische Leistungsfähigkeit quantitativ mittels (diagnostischer) Sensitivität und Spezifität beschrieben. Die Wahl des Cut-offs bestimmt dabei wesentlich die Erkennungsraten der Testpositiven zu allen Kranken bzw. der Testnegativen zur Vergleichsgruppe. Da sich jedoch Sensitivität und Spezifität gegenläufig beeinflussen, ist die Festlegung des Cut-offs letztlich ein Kompromiss zwischen beiden statistischen Kenngrößen. Im Allgemeinen legt man eine Spezifität von 95% fest, um verschiedene Tumormarker oder Auswerteverfahren bei Tumormarker-Kombinationen anhand der zugehörigen Sensitivität zu vergleichen.

Unter den Indikationen für kleinzellige Bronchialkarzinome wird ProGRP empfohlen [2]. Der Hersteller gibt als obere Referenzgrenze einen Cut-off von 46 pg/mL an. Messwerte über diesem Wert sind als positiver Befund zu bewerten. Im verfügbaren Datensatz konnte die 95%-Spezifität bestätigt werden: Unter den 198 Fällen mit NSCLC zeigten 187 einen niedrigeren Wert (Spezifität: 94,4%). Die Leistungsfähigkeit des Tumormarkers ergab eine Sensitivität von 78,5%: Von den 65 Fällen mit SCLC wurde bei 51 ein höherer Messwert als der gegebene Cut-off nachgewiesen.

NSE gilt (neben CYFRA 21-1) als weiterer relevanter Marker für die Diagnostik kleinzelliger Bronchialkarzinome. Um im Datensatz hierbei ebenfalls eine Spezifität von ca. 95% zu realisieren, ist ein Cut-off von 41 ng/mL zu wählen. In diesem Fall zeigen tatsächlich 187 der 198 Patienten mit NSCLC einen niedrigeren Wert (Spezifität: 94,4%). Von den 65 Patienten mit SCLC findet man bei 31 Fällen einen gegenüber dem Cut-off erhöhten NSE-Wert (Sensitivität: 47,7%).

Multiparametrische Tumormarker-Auswertung

Um die Sicherheit der Abklärung eines Karzinomverdachts zu verbessern, können mehrere Tumormarker untersucht werden. Die Möglichkeit, dass trotz eines malignen Geschehens mittels Labormessungen kein erhöhter Einzelwert nachgewiesen werden kann, verringert sich durch die Untersuchung mehrerer Marker. Vereinfacht man die biochemischen Zusammenhänge in der Kombination und betrachtet (idealisierend und wirklichkeitsfremd!) die Messwerte von Tumormarkern als voneinander statistisch unabhängig, so kann man die Sensitivitätssteigerung abschätzen: Die ermittelten Werte von 78% und 48% bedeuten, dass bei der Einzelmessung 22% bzw. 52% der Testergebnisse falsch-negativ sind. Die Wahrscheinlichkeit, dass beide Marker gleichzeitig zu einem falsch-negativen Ergebnis führen, ergibt sich aus dem Produkt beider Werte, also $0,22 \times 0,52 \approx 0,11$. Folglich steigt die Sensitivität bei Auswertung zweier Marker (theoretisch) auf ca. 89%.

Tabelle 1 Differenzierte Messwert-Bewertung für ProGRP und NSE (Fallzahlen).

		“Unauffällig”	“Grenzwertig”	“Erhöht”
ProGRP		< 26 pg/mL	26...66 pg/mL	> 66 pg/mL
SCLC	65	8	9	48
NSCLC	198	156	38	4
Summe	263 (100%)	164 (62,3%)	47 (17,8%)	52 (19,8%)
NSE		< 23 ng/mL	23...59 ng/mL	> 59 ng/mL
SCLC	65	19	22	24
NSCLC	198	159	33	6
Summe	263 (100%)	178 (67,7%)	55 (20,9%)	11,4%

Diese Abschätzung wird im Datensatz nahezu bestätigt. Bei 57 Fällen mit SCLC liegt entweder der ProGRP- oder der NSE-Wert (oder beide gleichzeitig) über den oben festgelegten Cut-offs – lediglich acht SCLC-Fälle werden mit dieser Entscheidungsstrategie nicht gefunden (Sensitivität: 87,7%). Allerdings sinkt dabei, wie allgemein erwartet, die Spezifität auf 89,9%.

Um die Entscheidungssicherheit bei der Auswertung der Marker-Kombination genauer zu analysieren, unterteile man die Messbereiche beider Marker in drei Abschnitte, die in Bezug zum Cut-off interpretiert werden:

- “unauffällig”: Es gibt keine Hinweise auf malignes Geschehen (negativer Befund),
- “grenzwertig”: Der Messwert ist in der Nähe des Cut-offs (unsicherer Befund) und
- “erhöht”: Der Messwert liegt weit über dem Cut-off (positiver Befund).

Mit willkürlich festgelegten Bereichsgrenzen lassen sich die Einzelwerte wie in Tabelle 1 angegeben zahlenmäßig einteilen.

In der Auswertung der Marker-Kombination erscheinen in 131 Fällen beide Messwerte als “unauffällig” und in 19 Fällen beide als “erhöht”. Die gleichlautenden Bewertungen bei diesen Fällen (57,0%) unterstützen offensichtlich die Entscheidungssicherheit. Dagegen widersprechen sich in 17 Fällen (immerhin 6,5%) die Bewertungen der Einzelmarker, einer erscheint als “unauffällig” und der andere als “erhöht”. Die Zweifachmessung irritiert also in diesen Fällen mehr als sie hilft. Im verbleibenden Drittel aller Fälle ist mindestens ein Messwert “grenzwertig” – nur mit fachärztlicher Erfahrung kann hier über die Bewertung entschieden werden.

Tabelle 2 Anteil der SCLC-Fälle in Abhängigkeit von den Messbereichen.

/ProGRP NSE/	“Unauffällig”	“Grenzwertig”	“Erhöht”
“Unauffällig”	5%	6%	79%
“Grenzwertig”	3%	33%	100%
“Erhöht”	33%	62%	95%

Mit der folgenden Argumentation kann die Entscheidungsfindung in den unklaren und unsicheren Situationen unterstützt werden. Durch die Aufteilung der Messbereiche in jeweils drei Abschnitte sind insgesamt neun Varianten möglich. Ermittelt man in jeder dieser Varianten den Anteil der SCLC-Fälle an allen betreffenden Datensätzen, so erhält man eine Art Verteilungsdichte. Die Entscheidung für einen positiven Befund ist dann zu empfehlen, wenn im betreffenden Bereich der Anteil der SCLC-Fälle einen vorgegebenen Wert übersteigt. Zum Anteil der SCLC-Fälle in Abhängigkeit von den Messbereichen siehe Tabelle 2.

Entscheidet man sich für Bereiche mit mehr als 30% SCLC-Fällen für den positiven Befund, so findet man eine Sensitivität von 86,2% bei einer Spezifität von 93,4%. Man beachte: Ohne Verlust an Spezifität gelingt durch dieses differenzierte Bewerten eine Sensitivitätssteigerung gegenüber dem besten Einzelmarker!

Cut-off-orientierte Verfahren

Die willkürlich vorgenommene Messbereichs-Einteilung in “unauffällig”, “grenzwertig” oder “erhöht” führt dazu, dass geringe Variationen dieser Bereichsgrenzen zwar qualitativ vergleichbare, aber quantitativ unterschiedliche Aussagen ergeben. Zur Objektivierung der eigentlich plausibel erscheinenden Bewertung ist folglich ein Algorithmus notwendig, der die subjektive Beeinflussung der Ergebnisfindung einschränkt. Das mathematische Problem besteht hierbei in der Verknüpfung zweier verschiedener Messgrößen. Nicht nur die unterschiedlichen Maßeinheiten lassen eine direkte Verarbeitung fragwürdig erscheinen.

Ein pragmatischer und verbreiteter Lösungsansatz besteht in der “Normalisierung” der Werte. Durch Bezug auf den Cut-off wird anstelle des absoluten Messwertes eine relative Messgröße bestimmt. Dividiert man die ProGRP-Messwerte durch den Wert 46 pg/mL und die NSE-Messwerte durch den Wert 41 ng/mL, so liegen alle relativen Größen der Messwerte unterhalb ihres Cut-offs zwischen 0 und 1 und der Werte oberhalb des Cut-offs über 1. Die relativen Messgrößen sind dimensionslos. Zahlenmäßig gleiche Werte haben eine vergleichbare Interpretation und ermöglichen damit problemlos eine

Tabelle 3 Sensitivitätssteigerung durch multiparametrische Auswertung.

	Einzelmarker		Zweiparametrisches Markerprofil (ProGRP – NSE)		
	ProGRP	NSE	Maximum	Arithmetischer Mittelwert	Geometrischer Mittelwert
Spezifität	94,4%	94,4%	89,9%	94,9%	96,0%
Sensitivität	78,5%	47,7%	87,7%	84,6%	78,5%

mathematische Weiterverarbeitung. Allerdings gibt es dabei vielfältige Möglichkeiten der Verknüpfung, beispielsweise:

- Ein positiver Befund liege vor, wenn das Maximum beider relativer Messgrößen mindestens 1 beträgt. Diese Entscheidungsstrategie entspricht der eingangs geschilderten Auswertung: mindestens ein Tumormarker zeigt einen Wert gleich oder über dem entsprechenden Cut-off.
- Die in Tabelle 2 angedeutete Vorschrift kann durch den arithmetischen Mittelwert beider relativer Messgrößen beschrieben werden. Ist dieser mindestens 1, so entscheide man sich für einen positiven Befund.
- Anstelle des arithmetischen Mittelwertes kann analog der geometrische Mittelwert für die Entscheidungsfindung genutzt werden. Ist dieser mindestens 1, so liege ein positiver Befund vor.

Während die Maximumbildung im Allgemeinen die Spezifität senkt, erhöht die Verwendung des Mittelwertes die Spezifität – “grenzwertige” Fälle werden nur dann zum positiven Befund erhoben, wenn die Interpretation in Richtung “erhöht” deutlicher ausgeprägt ist. Tabelle 3 zeigt die Auswirkung der unterschiedlichen Verfahren auf die Leistungsfähigkeit der Bewertung.

Sensitivitäts-orientierte Verfahren zur multiparametrischen Tumormarker-Auswertung

Die Verwendung relativer Messgrößen birgt den Vorteil, den Einfluss methodenabhängiger Cut-offs weitgehend eliminieren zu können. Die mathematisch abstrakte “Normierung” erlaubt aber außer dem Abstand zum Cut-off

keine weitere Interpretation. Was bedeutet ein relativer ProGRP-Wert von 0.5, etwa im Sinne von “halb so wie der Cut-off”? Können zahlenmäßig gleiche relative ProGRP- und NSE-Werte auch klinisch gleich interpretiert werden?

Um eine Interpretation der Entscheidungssicherheit über den gesamten Messbereich zu gewährleisten, nutze man zur Beschreibung den Sensitivitätsverlauf. In Analogie zur bisherigen Betrachtungsweise sei die Entscheidungssicherheit für einen positiven Befund gleich 1, wenn der Messwert einem Cut-off mit 95%-Spezifität entspricht. Zudem wird für einen Messwert von 0 die Entscheidungssicherheit mit 0 und für einen deutlich erhöhten Messwert (hier der doppelte Cut-off) mit 2 festgelegt. Die Werte für die Entscheidungssicherheit zwischen diesen drei fixierten Punkten wird mittels linearer Transformation durch den Sensitivitätsverlauf geprägt (vergleiche Abbildung 1).

Alle bisherigen Eigenschaften der multiparametrischen Tumormarker-Auswertung bleiben bestehen, so auch die Potenziale der mathematischen Verknüpfungen. Zusätzlich erhalten die zu verrechnenden Kenngrößen eine ganz neue Interpretation: Es werden nicht mehr quantitative Laborergebnisse, sondern qualitative Bewertungen der Entscheidungssicherheit miteinander verarbeitet, unabhängig von den eingesetzten Testmethoden und dem gewählten Cut-off. Die Auswertung der Kombination entspricht nun der Fragestellung “Wie sicher ist mit der Marker-Kombination die Entscheidung für einen positiven Befund SCLC?”, wenn für jeden Einzelmarker die Entscheidungssicherheit bekannt ist.

Dieses Vorgehen entspricht genau dem subjektiven Abwägen von Einzelinformationen für eine Gesamtbewertung – ein typisches Einsatzgebiet von fuzzy-basierten Verfahren [3]: Anstelle der konkreten Zahlenwerte von Messungen, die durch vielfältige Aspekte beeinflusst sein können, werden zielorientierte qualitative Eigenschaften dieser Messergebnisse verarbeitet. Im Falle der Tumormarker-Profilauswertung sind darunter die Beiträge der Einzelmarker zur Befunderstellung zu verstehen. Man stützt sich hierbei auf Sensitivität und Spezifität als diagnostische Kenngrößen des jeweiligen Markers. Weil ROC-Kurven als Darstellungsform für Sensitivität und Spezifität bekanntermaßen unabhängig von Cut-offs sind, geraten labortechnische Abhängigkeiten für die Datenauswertung in den Hintergrund und unterschiedliche Patientenkollektive werden für die ausgewählte klinische Fragestellung objektiv vergleichbar. Mit den so transformierten Daten tragen fuzzy-basierte Verfahren

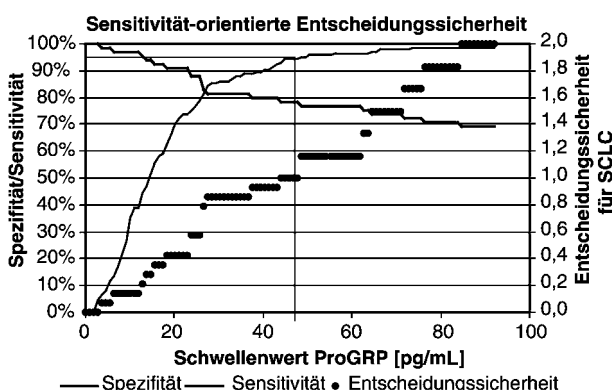


Abbildung 1 Sensitivitäts-orientierte Entscheidungssicherheit (Beispiel ProGRP).

der multiparametrischen Datenanalyse [4] zur Leistungsfähigkeit moderner Laboruntersuchungen bei.

Literatur

1. Lamerz R, Stieber P. Tumormarker. *Onkologe* 1998; 4:1067–76.
2. Stieber P, Dienermann H, Schmitt NM, Reinmiedl Y, Has-holzner N, Reiter W, et al. Pro-gastrin-releasing peptide (ProGRP) – a new marker in small-cell lung carcinomas. *J Lab Med* 1997;21:336–44.
3. Bitterlich N. Ist die Tumormarker-Profilauswertung wirklich (fuzzy) logisch? *J Lab Med* 2004;28:116–214.
4. Keller T, Bitterlich N, Hilfenhaus S, Bigl H, Löser T, Leonhardt P, et al. Tumour markers in the diagnosis of bronchial carcinoma: new options using fuzzy logic-based tumour marker profiles. *J Cancer Res Clin Oncol* 1998;124:565–74.