

# Von der zellbiologischen Messdatenflut zur medizinischen Erkenntnis

Nicht nur die Entwicklung der Nanotechnologie, sondern auch die im Rahmen der Genomforschung miniaturisierten und hochleistungsfähigen so genannten Mikroarray-Chips beschleunigten die Generierung von Messdaten in einer zuvor unvorstellbaren Weise. Damit wird aber heute ein Problem evident, für dessen Lösung im Prinzip noch keine guten und für die Routine einsetzbaren Verfahren entwickelt sind: Wie können wir diese unüberschaubare Fülle von Daten bezüglich ihrer Wertigkeit und vor allem ihrer Zuverlässigkeit und Reproduzierbarkeit interpretieren? Noch gibt es keine zuverlässigen Verfahren mit denen wir solche Analysenergebnisse auswerten und bewerten können.

Hoch relevant wurden diese Probleme bei der Auswertung von Genanalysen und beim Versuch, diese Techniken für die Diagnostik von meist durch Mutationen bedingten onkologischen Prozessen nutzbar zu machen. Der Ansatz war verführerisch einfach. Jedoch zeigte sich bald, dass diese vereinfachte Sicht von Genmutation und Ungleichgewicht zwischen Proliferation und Apoptose der multiplexen genetischen Situation nicht entsprach. Es genügt eben nicht, festzustellen, ob ein Gen mutiert, aktiviert oder ausgeschaltet ist, um so das Krankheitsgeschehen zu erkennen.

Daher sollen heute nicht mehr Einzeldaten in Bezug auf die Stratifizierung des Prozesses mit Referenzwertebereichen verglichen werden, sondern man versucht, Mustererkennungs-Verfahren als Grundlagen der Erkennung von Krankheitsprozessen zu entwickeln. Doch es war schon sehr schwer, für Einzelmessungen wie Enzymaktivitäten oder Substratkonzentrationen Referenzwertebereiche so genannter normaler Individuen festzulegen. Der Versuch, den Wertemustern von „gesunden“ biologischen Systemen Muster pathologischer Biosysteme gegenüberzustellen, ist noch ungleich komplizierter.

In diesem Zusammenhang wird das Kriterium der diagnostischen Sensitivität und Spezifität noch einmal um mehrere Stufen diffiziler, was sich auch nicht durch formale Standardisierungs- und Normierungsverfahren ausschalten lässt. Die moderne Bioinformatik wird sich dieser Fragestellung mithilfe von Informatik, Stochastik und weiteren mathematischen Disziplinen an der Schnittstelle zwischen analytischer Messtechnologie und biologischer Wertung widmen, um diese Musterbewertungen im Vergleich von physiologischen und pathologischen Zuständen zu ermöglichen.

So evident das Problem durch die genetischen Untersuchungen in unser Bewusstsein rückte, so ist es dennoch keineswegs auf diesen Bereich beschränkt. Es betrifft im Prinzip ein altes Problem der Mustererkennung: Denn wir wissen längst, dass Krankheitsprozessen fast immer mehr

oder weniger komplexe Muster von Messdaten zugrunde liegen. Immer hochleistungsfähigere Multiparameteranalysen und die dadurch mögliche Miniaturisierung der Messprozesse selbst sowie die Miniaturisierung des erforderlichen Probenmaterials, machen es dringend erforderlich, die erzeugte massive Datenflut solchen bioinformatischen Prozeduren zu unterziehen, um nicht mit falsch positiven bzw. falsch negativ bewerteten Ergebnissen falsche Zuordnungen von Krankheitsprozessen zu verursachen.

Wie sich diese Prozeduren der modernen Bioinformatik auch für die Bewertung von komplexen Datenmustern klassischer Art einsetzen lassen, bleibt spannend. Jedenfalls sind diese Verfahren zwingend erforderlich, um mit der bereits vorhandenen und keineswegs zuverlässig bewertbaren Datenflut etwas anfangen zu können. Unsere Messverfahren haben eine extreme Präzision erreicht, die aus Mangel einer verbindlichen Referenzebene die Kategorisierung zwischen kranken und gesunden Zuständen mit hinreichender Richtigkeit tatsächlich nicht feststellen lassen. Das Dilemma ist, dass die Leistungsfähigkeit der modernen instrumentellen, miniaturisierten, roboterbetriebenen, extrem schnellen und in Parallelprozessen „unbegrenzten“ Analytik keineswegs mit den Verfahren der biologisch erforderlichen Evaluierungen und Bewertungen Schritt gehalten haben.

Es ist ein spannender Vorgang, wie sich die Krankheitsbilder und ihre heute verwendeten Charakterisierungen letzten Endes mit diesen neuen Ansätzen einer genetisch bedingten Entwicklung von Befundmustern darstellen werden. Auch die im vorliegenden Heft abgehandelte Osteoporose, die sich ja ebenfalls Referenzwerten – sowohl zur Definition als auch zur Verlaufsbeurteilung – bedient, wird letzten Endes davon profitieren, durch eine Multiparametermustererkennung präziser einordenbar zu sein, um so genetisch bedingte von erworbenen, aber in ihrer Symptomatologie gleich erscheinenden Krankheitsformen zu unterscheiden.



Prof. Dr. Dr. A. Grünert, Ulm