

Gleason-Grading des Prostatakarzinoms

Konsens statt Evidenz

ALFRED BÖCKING

Die Malignitätsgradierung des Prostatakarzinoms nach Gleason bestimmt wesentlich die Eingangskriterien zur Active Surveillance. Laut S3-Leitlinien qualifizieren sich nur Patienten mit Gleason-Score 6 dafür – obwohl die interindividuelle Reproduzierbarkeit des modifizierten Scores nur 60 % beträgt, die an Stanz- und Resektaten ermittelten Scores in 30 % der Fälle voneinander abweichen und sich die Prognosen von Patienten mit Scores 6 und 7 nicht signifikant voneinander unterscheiden. Daher bedarf dieses subjektive Malignitätsmaß dringend einer Ergänzung durch objektivere, reproduzierbarere und validere Verfahren wie der prognostischen DNA-Bildzytometrie.

Tumoren werden nach dem Gewebe beziehungsweise dem Zelltyp benannt, von dem sie abstammen und dem sie daher mikroskopisch ähnlich sehen (histogenetische Tumor-Klassifikation, International Classification of Diseases, Oncology, Morphology, ICD-O-M). So stammt ein Adenokarzinom der Prostata von dessen Drüsen-bildendem Epithel ab. Der Malignitätsgrad eines einzelnen Tumors bezeichnet darüber hinaus seine individuelle maligne Potenz. Darunter versteht man ein Maß für die Wahrscheinlichkeit, dass ein Tumor innerhalb einer gegebenen Zeit lokal infiltrativ und destruktiv wächst sowie Metastasen setzt. Für das Prostatakarzinom ist dazu derzeit der modifizierte Gleason-Score (GS) geeignet [15]. Das Tumor-Stadium dagegen benennt organspezifisch den Ausbreitungsstand (TNM-System). So wie sich aus noch zurückzulegender Wegstrecke und Geschwindigkeit eine verbliebene Reisezeit berechnen lässt, sollte sich aus Tumorstadium und Malignitätsgrad eines Patienten die Zeit bis zum Eintritt eines tumorbedingten fatalen Ereignisses abschätzen lassen.

Während es sich bei der Bestimmung der histogenetischen Klassifikation von Tumoren um eine qualitative mikroskopische Beurteilung handelt, wird beim mikroskopischen Malignitäts-Grading eine skalare Größe quantitativ geschätzt. Da das menschliche Auge nicht messen

kann, unterliegt die subjektive Schätzung visueller Größen erheblichen interindividuellen Schwankungen. Handelt es sich zusätzlich, wie beim Gleason-Score, um die quantitative Beurteilung von Mustern, die durch mehrere Variablen definiert werden (Form, Größe, Konfluenz, Variabilität, Packungsdichte, Invasionsverhalten, papilläre oder kribri-forme Drüsen), ist eine reproduzierbare subjektive Einschätzung noch problematischer.

Biologie zunehmender Malignität

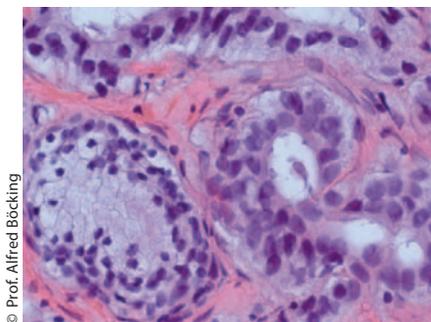
Der entscheidende formal-pathogenetische Schritt der malignen Transformation einer Zelle ist nach Duesberg et al. [9, 10] die Etablierung eines Klons chromosomal aneuploider Zellen. Die primären Chromosomen-Aberrationen verursachen mit weiteren Zellteilungen im Sinne einer Kettenreaktion sekundäre und tertiäre Aberrationen, was als zytogenetische Tumorprogression bezeichnet wird [5, 11, 12]. Diese beschleunigt wiederum die klinische Progression des Tumors. Das Ausmaß und die Variabilität dieser Chromosomen-Aberrationen korrelieren auch beim Prostatakarzinom mit der Prognose [18]. Daraus resultieren zyto- und histomorphologische Veränderungen, die bei der Gradierung nach Gleason quantitativ beurteilt werden. Aus der zytogenetischen Progression resultiert eine chromosomale, zytologische und histologische Hetero-

genität des Tumors. Später entstandene Zellklone höherer maligner Potenz verdrängen dabei früher entstandene niedrige Potenz [12, 39]. Daher ändert sich die Bösartigkeit maligner Tumoren spontan, wenn überhaupt, nur zu höheren Graden, nie zu niedrigeren.

Kein anderer Tumor des Menschen hat ein derart breites Spektrum einer Mehrzahl sehr wenig maligner, klinisch insignifikanter bis zu einer Minderzahl hoch maligner, aggressiver und schnell zum Tode führender Tumoren wie das Prostatakarzinom. Dies hat dazu geführt, dass manche Kliniker Pathologen dazu drängen, den niedrig malignen Karzinomen (z.B. des GS ≤ 6) eine andere, Patienten weniger beängstigende Bezeichnung zu geben [24].

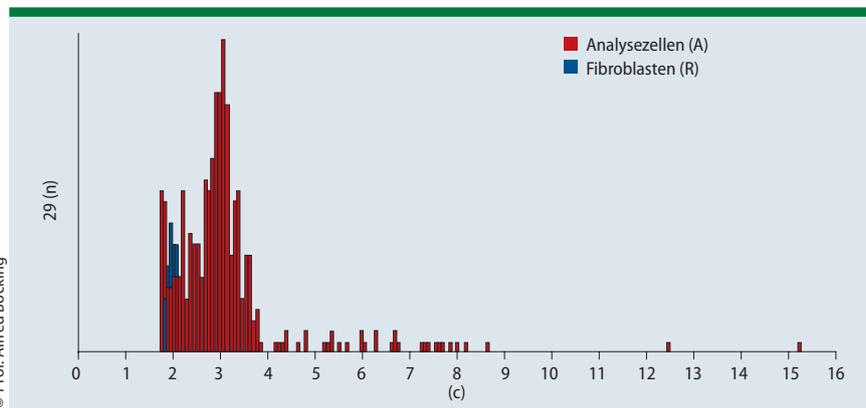
Anforderungen an Grading-Systeme

Um anhand einer Zell- oder Gewebsanalyse eines Tumors auf sein zukünftiges Verhalten zu schließen, sollte vor allem auf dessen bösartigsten Anteile geachtet werden, nicht auf die häufigsten, wie es der Gleason-Score tut. Um die maligne Potenz eines individuellen Tumors verlässlich zu bestimmen, sollte man möglichst das dafür ursächliche biologische Phänomen selbst messen (Ausmaß chromosomaler- bzw. DNA-Aneuploidie), statt die Ausprägung eines seiner Epiphänome (feingewebliche Wachstums-Muster, Histomorphologie) zu schätzen.



© Prof. Alfred Böcking

Abbildung 1a: Stanzbiopsie der Prostata. Links: normale Drüse, rechts: Karzinom, als Gleason-Score 3 + 3 = 6 befundet. 200 x. Untergradierung, da die hohe Anaplasie der Zellkerne keine Berücksichtigung findet



© Prof. Alfred Böcking

Abbildung 1b: zugehöriges DNA-Histogramm mit multiploider Verteilung, d.h. stark schwankenden und sehr hohen DNA-Gehalten, Typ D, DNA-MG 4; blaue Säulen: gesunde Fibroblasten als Referenzzellen, rote Säulen: Tumorzellen

Ein Tumor-Malignitäts-Grading ist umso besser, je reproduzierbarer und valider es das maligne Verhalten eines individuellen Tumors vorhersagt. Eine interindividuelle Reproduzierbarkeit von > 95 % und eine Vorhersage-Wahrscheinlichkeit eines Patienten-relevanten Ereignisses, wie dem Auftreten von Metastasen innerhalb von fünf Jahren, von > 95 % sind erstrebenswert.

Gleason-Score

Der modifizierte Gleason-Score [15] ergibt sich aus einem subjektiven Rating des häufigsten und des zweithäufigsten histologischen Wachstumsmusters in fünf Grade einer Gewebeprobe und deren Addition. In Stanzbiopsien wird zum Grad des häufigsten derjenige des malignesten Musters addiert („the most and the worst“). Sie berücksichtigen allerdings keine zellulären Aspekte unterschiedlicher Malignität, wie dies andere Grading-Systeme für das Prostatakarzinom tun [3, 20, 31]. Es resultieren Scores zwischen 2 und 10. Laut Konsensus der International Society of Urologic Pathology [15] werden an Stanzbiopsien Scores zwischen 2 und 5 nicht mehr vergeben, um den Anteil der späteren Höhergradierungen an Resektaten zu reduzieren.

Methodische Kritik

Vernachlässigung der zellulären Ebene: Der Gleason-Score lässt die prognostisch relevanten zellulären Aspekte unterschiedlicher maligner Potenz unberück-

sichtigt. Dies führt zu Untergradierungen, wenn sich in vermeintlich gut differenzierten Drüsen, beispielsweise des Gleason-Grades 3, hochgradig entdifferenzierte Zellen finden, wie dies Helpap und Köllerman beschrieben haben [20] und **Abbildung 1** dokumentiert. Solche Fälle sind mit verantwortlich für die unerwartet hohen Progressionsraten bei Patienten mit GS 6-Karzinomen unter Active Surveillance zwischen 30 % [23], 33 % [29] und 36 % [33].

Unangemessenes tumorbiologisches Konzept:

Die Betonung des häufigsten statt des malignesten Anteiles missachtet die Tatsache, dass es im Rahmen der Tumorprogression zu einer Abnahme weniger maligner Zellen und einer Zunahme höher maligner kommt. Dies ist bei zyto-genetisch unterschiedlichen Zellpopulationen im Tumor bedingt durch eine Selektion für höhere Malignität [12]. Die malignesten Anteile im Tumor bestimmen das Schicksal des Patienten. Daher sollte ein Tumor-Malignitätsgrad nach dem malignesten Anteil benannt werden [3] und nicht nach dem häufigsten.

Rating von Bildern: Ein subjektives visuelles Rating von Mustern, wie es der Gleason-Score verlangt, ist nachvollziehbar weniger reproduzierbar als eine objektive Messung prognostisch valider Variablen, wie zum Beispiel der chromosomalen Instabilität, die sich in der Verteilung der Zellkern-DNA-Gehalte manifestiert.

Tumor-Heterogenität: Aufgrund der durch genetische Instabilität verursachten erheblichen Heterogenität der meisten Prostatakarzinome, vor allem fortgeschrittener Stadien, ist das an wenigen Biopsien gewonnene Ergebnis oft nicht repräsentativ für den Tumor als Ganzes [19]. Sowohl die Erhöhung der Zahl von Biopsien als auch die Betrachtung eines weniger heterogenen Markers, wie etwa der DNA-Verteilung [42], können diesem Problem begegnen.

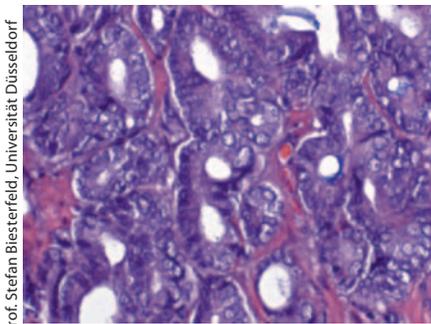
Systematische Höhergradierung:

Durch eine systematische Höhergradierung der Muster 1 und 2 an Stanzbiopsien als Gleason-Grad 3, kommen Scores < 6 darin nicht mehr vor. Zusätzlich werden durch Verschiebung eines Teiles des früheren Grad-3-Musters in Grad 4 mehr Fälle als GS 7 bewertet, als zuvor. Die betroffenen Patienten kommen damit nicht mehr für eine Strategie der aktiven Überwachung infrage [2].

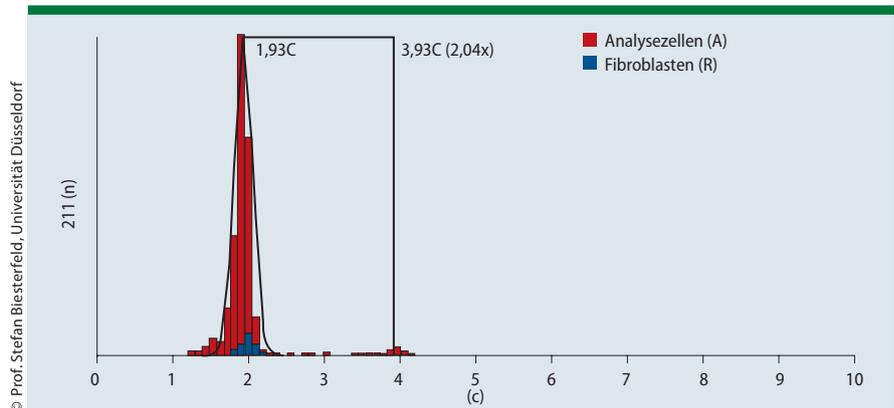
Performance des Gleason-Scores

Interindividuelle Reproduzierbarkeit:

Der Gleason-Score wurde auch revidiert [15], um die bekannt schlechte interindividuelle Reproduzierbarkeit der ursprünglichen Anleitung [16] zu verbessern. Die seitdem publizierten diesbezüglichen Studien belegen aber, dass dieses Ziel nicht erreicht wurde. Veloso et al. berichten über 58,6–69,6 % an 110 Biopsien [41], Burchard et al. über 45,7 % an 331 Fotos [7], Netto et al. über 70,0 % an 257 Prostataektomien [32] und Helpap



© Prof. Stefan Biesterfeld, Universität Düsseldorf
Abbildung 2a: Radikale Prostatektomie. Karzinom, wegen Fusion von Drüsen als Gleason-Score 4 + 3 = 7b befundet. 200 x. Präparat überfärbt. Übergradierung, da die geringe Anaplasie der Zellkerne unberücksichtigt bleibt.



© Prof. Stefan Biesterfeld, Universität Düsseldorf
Abbildung 2b: zugehöriges Histogramm mit peridiploider DNA-Verteilung (Zahl der Kerne vs. DNA-Gehalt in c), Typ A, DNA-Grad 1; blaue Säulen: gesunde Fibroblasten, rote Säulen: Tumorzellen

& Köllerman über 59,5% an 134 Biopsien [20]. 40% der zunächst als GS 6 eingestuften Biopsien gradierten sie im Rahmen ihres Zweitmeinungs-Service höher. Vergleicht man den Mittelwert dieser neueren Ergebnisse (60,7%) mit demjenigen aus zehn, die Reproduzierbarkeit des ursprünglichen GS (67,3%) untersuchenden Publikationen [6, 13, 17, 26–28, 30, 34, 37, 38], so fällt sogar eine Verschlechterung um 6,6% auf.

Zu wenig niedrig maligne Karzinome: Durch die Höherbewertung eines Teils des früheren Gleason-Grades 3 (kribri-formes Muster) als 4 nahm der Anteil an GS 6-Fällen an Stanzbiopsien um 19,2% (n=172) ab, die nun nur noch 49,4% ausmachen [2]. In weiteren Studien ergab sich ein Abfall der GS 6-Häufigkeit an Biopsien von 48% auf 22% [19] beziehungsweise 35,1% [20].

Prognostische Validität: Das therapeutisch relevanteste Ziel des modifizierten GS hätte sein sollen, die prognostisch besonders günstigen Patienten mit einem Score von ≤ 6 von denen mit einem Score 7 hinsichtlich des Eintretens einer Tumorprogression signifikant zu unterscheiden, um ersteren eine sichere Strategie der aktiven Überwachung anbieten zu können. Wie an 118 Patienten mit GS ≤ 6 nach radikaler Prostatektomie und 44 mit GS 7 gezeigt werden konnte, unterschieden sich diese innerhalb von sieben Jahren aber nicht hinsichtlich der progressionsfreien Überlebenszeit [2]. In

einer weiteren Studie zeigte sich zwischen Patienten mit GS 7a und 7b nach radikaler Prostatektomie kein Unterschied im rezidivfreien Überleben [35]. Lediglich die besonders bösartigen Tumoren, deren Träger in wenigen Jahren daran sterben, erkennt der modifizierte GS besser als der ursprüngliche [2].

Repräsentativität: Auch das Ziel, die Übereinstimmung zwischen GS an Stanzbiopsien und Resektaten wesentlich zu verbessern, wurde durch den modifizierten Score nicht erreicht: In verschiedenen Studien wurden Übereinstimmungen von 70–72% ermittelt [19, 40, 41].

Therapeutische Konsequenzen

Zu wenig niedrig gradierte Karzinome: Die durch die Modifikation des GS von 2005 bewirkte systematische Höhergradierung schließt ungerechtfertigterweise viele Patienten von einer leitliniengerechten Strategie der Active Surveillance aus, da ihnen biologisch ein zu hoher Malignitätsgrad zuerkannt wird. Die Letalität am Prostatakarzinom betrug in den USA im Jahre 2006 lediglich 8,5% [22], 91,5% der Patienten sind also mit und nicht an ihrem Karzinom gestorben. Daher müsste der Anteil niedrig maligner Karzinome eines GS ≤ 6 höher sein als nur 22% [19].

Übergradierungen: Übergradierungen können dazu führen, dass Patienten vermeidbaren, belastenden Operationen oder Bestrahlungen zugeführt werden.

Die derzeit gültigen S3-Leitlinien Prostatakarzinom schließen Patienten mit einem GS > 6 aber von einer aktiven Überwachung aus. Mit großer Wahrscheinlichkeit befindet sich daher unter den als GS 7 eingestuften Patienten ein hoher Anteil, der damit übergradiert ist und deswegen von einer Active Surveillance ausgeschlossen wird. **Abbildung 2** demonstriert das Gewebesbild eines mit GS 7 befundenen Prostatektomie-Präparates, an dem aber eine prognostisch sehr günstige peridiploide DNA-Verteilung (Typ A, DNA-Malignitätsgrad 1) gemessen wurde. Dieser Tumor wurde damit also übergradiert. Bei einer Fehlerrate von 40% muss man davon ausgehen, dass etwa 20% der als GS 7 eingestuften Patienten in Wirklichkeit einen niedrigeren Score haben und damit für eine aktive Überwachung infrage kämen. Burchard et al. berichten über 15,1% übergradierte GS7-Befunde [7].

Untergradierungen: Untergradierungen können dazu führen, dass eigentlich indizierte Interventionen verworfen werden und der Patient unter Active Surveillance einen Progress erleidet, der hätte vermieden werden können. Helpap & Oehler gradierten 40% der zuvor als GS 6 befundenen Biopsien als GS 7a nach [21]. Die oben genannten Raten zwischen 30% und 36% für Progress von Patienten unter Active Surveillance innerhalb von sechs Jahren lassen darauf schließen, dass ein relevanter Anteil der als GS- ≤ 6 -Tumoren damit untergra-

diert ist. Einen solchen Fall zeigt die **Abbildung 2**. Am Karzinomgewebe eines GS von 6 wurde eine DNA-multiploide Verteilung (Typ D) gemessen, die allgemein als hoch maligne gilt [39]. Bei der bekannten Fehlerrate von etwa 40 % für die Bestimmung des GS muss man davon ausgehen, dass bei mindestens 20 % der GS 6-Befunde in Wirklichkeit ein höherer Score vorliegt, diese Patienten also von einer aktiven Überwachung ausgeschlossen werden sollten.

Konsens statt Evidenz

Die durch die ISUP-Konferenz von 2005 vorgeschlagenen Modifikationen des GS sind nicht etwa evidenzbasiert, sondern lediglich unter den eingeladenen Pathologen konsentiert. Ihre prognostische Relevanz bei den niedrig bis mittelgradig malignen Karzinomen konnte bisher nicht hinreichend belegt werden. Die Empfehlungen der 2013 aktualisierten S3-Leitlinien Prostatakarzinom zu den Einschlusskriterien in eine aktive Überwachung sind ebenfalls ein nicht durch Studien hoher Evidenzgrade gestützter „Expertenkonsens“. Die fehlenden wissenschaftlichen Evidenzen, die eine valide prognostische Unterscheidung zwischen Patienten mit GS 6 und 7 belegen, können nicht durch einen Konsens unter „Experten“ ersetzt werden. Schwerwiegende therapeutische Entscheidungen, wie der Entschluss zu einer radikalen Prostatektomie, sollten daher nicht alleine vom GS abhängig gemacht werden.

Adjuvante DNA-Zytometrie

In Karzinomen nehmen, bedingt durch chromosomale Instabilität im Rahmen der zytogenetischen Tumorprogression, Ausmaß und Variabilität chromosomaler Aberrationen (Aneuploidie) zu. Damit ist eine zunehmende Malignität der Tumoren assoziiert. Indem man den DNA-Gehalt in hunderten von Tumorzellen aus Stanzbiopsien misst, lässt sich das Ausmaß der DNA-Aneuploidie als objektives Maß für die Bösartigkeit individueller Karzinome bestimmen. Niedrigrisiko-Prostatakarzinome zeigen mit den gängigen Messmethoden von gesunden Zellen nicht unterscheidbare, peridiploide DNA-Gehalte (Typ A, DNA-Malignitätsgrad 1). Die interindi-

viduelle Reproduzierbarkeit der DNA-Malignitäts-Gradierung des Prostatakarzinoms beträgt 93 % [14]. Vier Studien mit Evidenzlevel 1b belegen, dass die DNA-Zytometrie ein organüberschreitendes Wachstum besser vorhersagen kann als der Gleason-Score. Zwölf Studien des Evidenzlevels 2b konnten für Patienten unter verschiedenen Therapien zeigen, dass die prognostische DNA-Zytometrie den Gleason-Score um signifikante prognostische Information ergänzen kann. Eine Level-1b- und drei Level-2b-Studien beweisen, dass die DNA-Ploidie dem histologischen oder zytologischen Grading bei Patienten unter aktiver Überwachung hinsichtlich der tumorspezifischen Überlebenszeit und der Nicht-Progression prognostisch signifikante Information zufügen kann [4, 5]. Die Korrelation der Grading-ergebnisse zwischen Stanzbiopsien und Resektaten ist für die DNA-Zytometrie besser als für den GS [1, 25, 36].

Patienten mit solchen, prognostisch sehr günstigen DNA-Verteilungen, die aufgrund der aktuellen S3-Leitlinien Prostatakarzinom für eine aktive Überwachung infrage kommen, können sich dieser Strategie beruhigter anvertrauen als aufgrund eines niedrigen Gleason-Scores ≤ 6 alleine. Für Patienten mit Prostatakarzinomen höherer DNA-Malignitäts-Grade sollten jedoch unmittelbar kurative Ansätze gewählt werden, da ihr Risiko für einen Progress nachweislich erhöht ist.

Die DNA-Malignitäts-Gradierung bietet eine Art Zweitmeinung mit anderer Methodik an, welche den histopathologischen Erstbefund nicht kompromittiert.

Literatur beim Verfasser

Prof. Dr. Alfred Böcking
Em. Direktor des Instituts für
Cytopathologie
Universitätsklinikum Düsseldorf;
Institut für Pathologie am Krankenhaus
Düren
Roonstr. 30
52309 Düren
E-Mail: alfred.boecking@web.de