

# Numerus clausus: Finden wir mit dem «Eignungstest für das Medizinstudium» die Geeigneten?

K.-D. Hänsgen, B. Spicher

## Zusammenfassung

Für die Zulassung zum Studium der Medizin wird seit 1998 ein Numerus clausus an den Universitäten Basel, Bern, Freiburg und Zürich angewendet. Als Zulassungskriterium dient ein Eignungstest. Welche Erfahrungen liegen bezüglich dieser Zulassung vor – berücksichtigt sie wie gefordert die Eignung? Diese Frage konnte nun auch erstmals in der Schweiz anhand konkreter empirischer Daten untersucht werden, indem für zwei Jahrgänge (1998 und 1999) die Prüfungsergebnisse im Studium mit den Testwerten im EMS verglichen worden sind. Der Testwert erlaubt eine gute Vorhersage von Studiendauer und Prüfungsnoten. Es erhalten diejenigen bevorzugt einen Studienplatz, die das Studium in kürzerer Zeit und mit besseren Leistungen absolvieren. Gleiche Testwerte bedeuten auch in verschiedensten Teilgruppen (nach Sprache, Alter, Geschlecht, Disziplin, Universität) eine vergleichbare Bestehenswahrscheinlichkeit der Prüfung – die Chancengleichheit ist gegeben.

Der EMS ist nicht nur bezüglich Ökonomie, Rechtsgleichheit oder Willkürfreiheit gegenüber anderen Zulassungskriterien konkurrenzlos – er differenziert nachweislich Personen nach der Studieneignung.

## Einleitung

Die gesetzlichen Grundlagen für den Numerus clausus der Kantone regeln, dass bei der Zulassung zum Medizinstudium von der Eignung auszugehen ist. Die *Übereinstimmung von persönlichen Leistungsvoraussetzungen und den Anforderungen* ist das Gemeinsame aller Definitionen von Eignung. Bei der Bestimmung der Eignung ist daher zu beachten:

- Es muss das «Wofür geeignet» hinsichtlich der *Anforderungen* bekannt sein. Die Beschränkung auf Studieneignung ist sinnvoll, denn auch der Kapazitätsengpass, mit dem fair umgegangen werden muss, betrifft das Studium. Berufseignung wäre beispielsweise

bezüglich der Anforderungen nur unzureichend definierbar – es gibt verschiedene Disziplinen und Tätigkeiten innerhalb der Medizin, die sich stark unterscheiden. Dennoch sind Studien- und Berufseignung natürlich nicht unabhängig voneinander.

- Es müssen Zulassungskriterien vorhanden sein, welche den *Erfolg einer Person* vorhersehbar machen und eine Unterscheidung der Personen nach der Eignung zulassen. Das Zulassungskriterium ist nur dann rechtfähig, wenn wissenschaftlich nachgewiesen ist, dass die Geeigneten tatsächlich eine höhere Chance erhalten, zum Studium zugelassen zu werden.

Zwei *Erfolgskriterien* sind für die Studieneignung bestimmend: Studienplätze sollen bevorzugt an diejenigen vergeben werden, die

1. in möglichst kurzer bzw. *optimaler Zeit* das Studium absolvieren sowie
2. das Studium in *guter Qualität* bzw. mit auf hohem Niveau erworbenen Fähigkeiten und Fertigkeiten absolvieren.

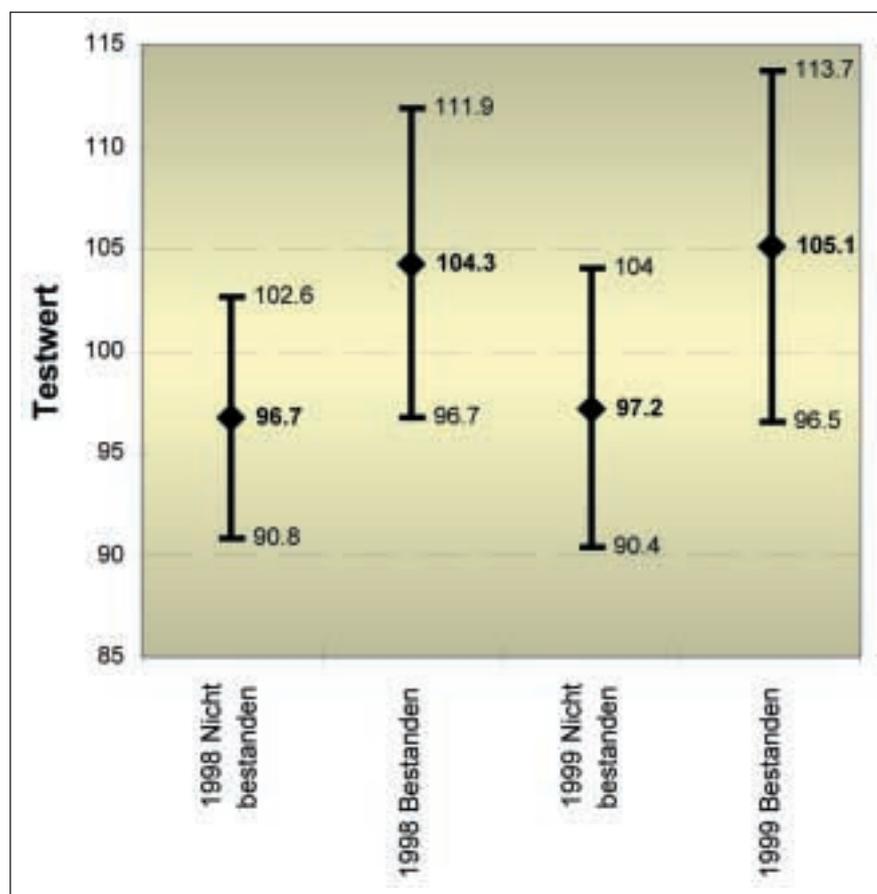
Durch Vermeidung einer verlängerten Studiendauer oder gar von Studienabbrüchen werden die begrenzt vorhandenen Ausbildungsressourcen nicht mehr als notwendig beansprucht und früher wieder für Nachfolgende frei. Im Medizinstudium ergeben sich Begrenzungen beispielsweise durch eine nicht beliebig erweiterbare Patientenzahl an den Universitätsspitalern und teuer ausgestattete Übungs- und Praktikumsplätze. Man muss aber nicht nur diese «volkswirtschaftliche» Perspektive sehen, denn auch für die Studierenden besteht der Nutzen darin, dass weniger verlorene Lebenszeit auf dem Weg zum Beruf entsteht, wenn weniger Abbrüche nach einem oder sogar zwei Jahren auftreten.

Eignung sollte dabei auch nicht mit *Neigung* verwechselt werden. Der Begriff der Neigung wird eher mit dem Interesse für den Medizinalberuf in Verbindung gebracht. Es wäre ideal, wenn die Informationen vor der Studienbewerbung so ausreichend gegeben würden, dass

Korrespondenz:  
Prof. Dr. Klaus-Dieter Hänsgen  
Universität Freiburg  
Departement für Psychologie  
Zentrum für Testentwicklung  
und Diagnostik  
Route Englisberg 9  
CH-1763 Granges-Paccot

Abbildung 1

Mittelwerte (◆) und Standardabweichungen des Testwertes für beide Kohorten nach dem Bestehen der ersten Vorprüfung.



realistische Erwartungen in bezug auf Studium und den späteren Beruf bestehen. Interesse ist hier aber wie in vielen anderen Bereichen eben nicht mit Fähigkeit bzw. Eignung gleichzusetzen.

In der Schweiz wird ein Eignungstest (EMS) als Zulassungskriterium verwendet. Über Erfahrungen in der Schweiz wurde von Hänsgen und Spicher (2000) berichtet. Der EMS basiert dann auf der Eignung, wenn 1. Beziehungen zwischen Testleistung im EMS und dem Bestehen der Prüfungen zur Vorhersage der Studiendauer und 2. zwischen Testleistung im EMS und den Prüfungsnoten zur Vorhersage der Studienleistungen nachgewiesen werden können. Der EMS ist eine Adaptation des deutschen Tests für Medizinische Studiengänge (TMS). Auch die deutschen Ergebnisse belegen eine gute Prognosegüte für Studienerfolg (dazu u.a. Trost et al. [1]).

Für die Analyse des Zusammenhangs von Testergebnis und Prüfungserfolg stehen bisher die Testteilnehmenden 1998 und 1999 zur Verfügung. 1998 war der Numerus clausus für

Humanmedizin, 1999 für Human- und Veterinärmedizin an den Universitäten Basel, Bern, Freiburg und Zürich notwendig. Die 1998 Immatrikulierten konnten bisher vier mögliche Prüfungstermine wahrnehmen, eine zweijährige Studiendauer bis zur zweiten Vorprüfung wäre damit das Optimum. Die 1999 Immatrikulierten konnten bisher im Sommer oder Herbst 2000 nur die erste Vorprüfung ablegen.

Die Prüfungsdaten wurden im Bundesamt für Gesundheit (BAG) erfasst.<sup>1</sup> Die Vorhersage des Prüfungserfolges aufgrund des Testwertes wird als wichtigstes Evaluationskriterium betrachtet. Bekanntlich ist der EMS ein Test, der auf die Vorhersage der Studierfähigkeit hin ausgerichtet wurde. Er ist kein Wissenstest – es wird geprüft, inwieweit die Erarbeitung von Wissen (z.B. aus Texten), das Erkennen und die Anwendung logischer Regeln oder die Konzentrations- und Merkfähigkeit gelingt. Bei den Vorprüfungen handelt es sich dagegen um den tatsächlich im Studium erworbenen Wissensstand (Fähigkeiten und später auch Fertigkeiten). Die hier vorgestellten Ergebnisse werden in einer ausführlichen Fassung bei Hänsgen und Spicher [2] dargestellt.

### Testwert und erste Vorprüfung

Die mittleren Testwerte der einzelnen Gruppen nach dem Prüfungsverlauf sind in Tabelle 1 dargestellt. Unterschiedliche Zeiträume (ein oder zwei Jahre bis zur Prüfung) haben dabei nur einen geringen Einfluss.

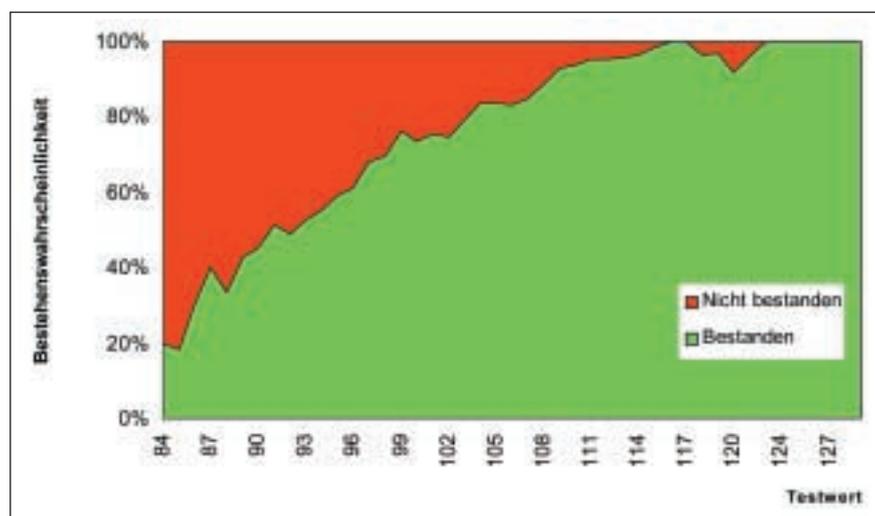
Die Personengruppe «Prüfung sofort bestanden» unterscheidet sich von allen anderen signifikant, ihr mittlerer Testwert liegt signifikant höher als der aller anderen Gruppen. Personen, die im zweiten Anlauf bestanden haben oder noch nicht angetreten sind, liegen bei einem mittleren Testwert um 100. Personen, welche die Prüfung ein- oder zweimal nicht bestanden haben, unterscheiden sich wiederum auch von diesen Gruppen – deren Mittelwert liegt deutlich unter 100.

Die Testwerte derjenigen Personen, die zur ersten Vorprüfung angetreten sind, wurden nach dem Bestehen dieser Prüfung unterschieden und in Abbildung 1 dargestellt. In beiden Jahren ergeben sich etwa gleiche signifikante Unterschiede zwischen den Gruppen nach dem Erfolg.

Zwischen Testwert und der Wahrscheinlichkeit des Prüfungsbestehens besteht ein monotoner linearer Zusammenhang (Abb. 2): je höher der Testwert, desto höher die Wahrscheinlichkeit des Bestehens. Dies bedeutet allerdings auch,

<sup>1</sup> Wir danken Herrn Megert, Frau Hirvi sowie Herrn Neuhaus vom BAG an dieser Stelle herzlich für die Zusammenarbeit.

**Abbildung 2**  
Wahrscheinlichkeit des Bestehens der 1. Vorprüfung pro Testwert für Kohorte 1998 und 1999 zusammengefasst.



dass unterhalb eines bestimmten Testwertes die Wahrscheinlichkeit, eine Prüfung zu bestehen, recht gering wird. Durch die erneute Zuweisung von durch Rückzüge freiwerdenden Studienplätzen an diesen Personenkreis wird die Gesamterfolgsrate des Jahrganges dann nachweislich verringert.

### Effekt der Zulassung von der Warteliste

Durch Rückzüge freiwerdende Plätze werden bekanntlich an Personen vergeben, deren Testwert unterhalb des kapazitätsentsprechenden Testwertes (nachfolgend als Kriterium bezeichnet) liegt. Dieses Kriterium wird jährlich nur nach der Kapazität festgelegt – es trennt keinesfalls Geeignete von Ungeeigneten. Stünden weniger Plätze zur Verfügung, würde das Kriterium ein höherer Testwert sein, bei mehr Plätzen wäre es entspre-

**Tabelle 1**  
Testwert Mittelwert (m) und Standardabweichung (s) für Gruppen nach der ersten Vorprüfung. Doppellinien: homogen im multiplen Mittelwertvergleich, n: Anzahl.

	Test 1998 (nach 2 Jahren)			Test 1999 (nach 1 Jahr)			Gesamt		
	n	m	s	n	m	s	n	m	s
zweimal nicht bestanden	46	97,2	6,4	11	95,9	7,1	57	96,9	6,5
einmal nicht bestanden	47	96,3	5,4	172	97,3	6,8	219	97,1	6,6
noch nicht angetreten	33	99,6	6,9	83	98,3	7,7	116	98,6	7,5
zweiter Anlauf bestanden	101	99,9	5,4	30	101,7	7,9	131	100,3	6,1
sofort bestanden	373	105,5	7,7	370	105,4	8,6	743	105,4	8,1

chend niedriger. Aus Tabelle 2 wird ersichtlich, dass 63 Personen zur ersten Vorprüfung angetreten sind, die das Kriterium eigentlich nicht erreicht hatten und über die Warteliste zugelassen worden sind. Deren Erfolgsaussicht zum Bestehen dieser Prüfung beträgt nur 40% – im Vergleich zu 78% für Personen mit einem Testwert über dem Kriterium. Die zweite Vorprüfung erreichten bisher sogar überhaupt nur zwei Personen mit einem Testwert unterhalb des Kriteriums. Es ist also damit zu rechnen, dass sich die Unterschiede im Studienverlauf noch verstärken.

Diese Befunde zeigen das Risiko auf, welches eine Kapazitätserhöhung bei gleichbleibendem Bewerbungsverhalten birgt: die Studiendauer wird länger und die Erfolgsaussichten werden geringer. Vermutlich wird auch die Abbrecherzahl steigen.

Es wäre darüber nachzudenken, wie sich mehr Geeignete für ein Medizinstudium bewerben – ob die Dissuasionsfaktoren möglicherweise die «Falschen» abschrecken und selektiv wirken. Mit dem EMS steht jetzt ein Instrument zur kapazitätsentsprechenden Zulassung nach der Eignung zur Verfügung, und die Gefahr der Überlastung der Universitäten ist gebannt. Die Zugelassenen haben gute Chancen, das Studium mit Erfolg abzuschliessen. Unterstellt man, dass steigende Maturanzahlen eigentlich auch ein steigendes Interesse für das Medizinstudium hervorbringen müssten, bleiben nur Dissuasionsfaktoren als eine Ursache für das hohe, aber gleichbleibende Interesse für das Fach Medizin. Ein Teil davon wird in beruflichen Faktoren begründet sein. Wenn mehr Personen den Test als Chance sehen würden, ihre Eignung festzustellen (und ihre Neigung zu diesem Studium prüfen), würden mehr Geeignete das Studium beginnen.

### Chancengleichheit für verschiedene Teilgruppen

Eine weitere wichtige Forderung an das Zulassungskriterium ist die Chancengleichheit – bei gleicher Eignung müssen unabhängig von Alter, Geschlecht, Herkunft usw. gleiche Erfolgschancen bestehen. Vergleicht man die Testwerte für verschiedene Teilgruppen nach dem Prüfungserfolg in der ersten Vorprüfung, so fällt auf, dass die Erfolgreichen sich von den nicht Erfolgreichen bezüglich des mittleren Testwertes auch in jeder einzelnen Teilgruppe signifikant unterscheiden und dass die mittleren Testwerte der Erfolgreichen sich untereinander wenig unterscheiden. Die Vorhersagekraft des Testwertes ist

**Tabelle 2**

Erfolg in beiden Vorprüfungen und kapazitätsentsprechender Testwert (Kriterium), Erläuterungen siehe Text.

<i>Erfolg 1. Vorprüfung</i>				
		nicht bestanden	bestanden	n
Testwert unter Kriterium	Anzahl	38	25	63
	Prozentverhältnis	60,3%	39,7%	
Testwert über Kriterium	Anzahl	238	849	1087
	Prozentverhältnis	21,9%	78,1%	
<i>Erfolg 2. Vorprüfung</i>				
		nicht bestanden	bestanden	n
Testwert unter Kriterium	Anzahl	1	1	2 (!)
	Prozentverhältnis	50,0%	50,0%	
Testwert über Kriterium	Anzahl	51	293	344
	Prozentverhältnis	14,8%	85,2%	

also universell. In der Tabelle 3 werden die Mittelwerte und Standardabweichungen des Testwertes gegenübergestellt. Zu beachten sind die unterschiedlichen Bestehenswahrscheinlichkeiten in den Teilgruppen – nicht jede Teilgruppe weist das gleiche Verhältnis von Geeigneteren und weniger Geeigneten auf.

Die Vorhersagekraft ist nicht spezifisch für einzelne Universitäten. In allen Universitäten beträgt der Unterschied mindestens 5 Testwertpunkte zwischen den Gruppen nach dem Prüfungsbestehen. Unterschiede finden sich allerdings in den Bestehensquoten – die liegen aber offenbar nicht (nur) in unterschiedlicher «Strenge» bei den Prüfungen begründet, sondern auch in Leistungsunterschieden der Personen (unterschiedliche «Bewerberprofile» der Universitäten und Folgen der Umleitungen sind zu beachten). Die besondere Situation von Freiburg wird deutlich: Bekanntlich ist die Rate der umgeleiteten Studierenden dort am höchsten. Würde man die Kapazitäten weiter erhöhen, führte dies vermutlich zur Zulassung von noch mehr Personen, die geringere Aussichten auf das Bestehen der Prüfung haben.

Auch bezüglich der beiden Geschlechter findet sich der gleiche deutliche Unterschied des Testwertes für die beiden Gruppen nach dem Bestehen der Prüfung. Bekanntlich waren 1998 und 1999 keine signifikanten Unterschiede der Testwerte für beide Geschlechter nachweisbar, wobei tendenziell die Testwerte der Männer immer etwa 1,5 Punkte besser waren. Für die beiden Gruppen, welche die Prüfung bestanden

haben, sind die Testwerte hier nun nahezu identisch – die Gruppen, welche die Prüfung nicht bestanden haben, unterscheiden sich nur um 0,4 Punkte. Dies ist ein Beleg, dass mögliche Testwertunterschiede zwischen den Geschlechtern keinesfalls als Chancenunterschiede zu interpretieren wären, sondern sehr wahrscheinlich die unterschiedliche Repräsentativität mit einem Überwiegen weiblicher Studierender dafür verantwortlich sein wird. Die testwertabhängigen «Chancen», die erste Vorprüfung zu bestehen, sind praktisch bei gleicher Eignung identisch. Dieser Befund entkräftet jegliche Befürchtungen, dass es zu Ungleichbehandlungen bezüglich der Geschlechter kommt.

In allen Testjahren ergab sich bisher das Ergebnis, dass die Testwerte von Personen, die älter sind und die Maturitätsprüfung spät abgelegt haben, geringer sind als von Gleichaltrigen, welche die Maturität zum frühestmöglichen Zeitpunkt abgelegt haben. Es wurden zwei Gruppen «älterer Personen» (für Geburtsjahre vor 1975) gebildet, die nach dem Ablegen der Maturität nach 1995 («späte Maturität») oder vor 1996 («frühe Maturität») unterschieden werden. Diese werden mit den «jüngeren Personen» verglichen. Die Testwerte unterscheiden in allen drei Gruppen Erfolgreiche von weniger Erfolgreichen deutlich (mindestens  $\frac{1}{2}$  Standardabweichung). In der älteren Gruppe mit später Maturität können mit anderen Gruppen vergleichbare Erfolge allerdings auch mit niedrigeren Testwerten erreicht werden. Dies könnte ein Beleg dafür sein, dass die grössere «Lebenserfahrung» dann für einen Bonus sorgt, wenn die «Lernerfahrung» durch eine unlängst abgelegte Maturitätsprüfung aufgefrischt worden ist.

Auch für die drei Sprachgruppen treten vergleichbare Unterschiede nach dem Prüfungserfolg auf, wobei er in der italienischen Sprachgruppe tendenziell etwas weniger deutlich ist (sie ist allerdings auch die zahlenmässig kleinste Gruppe). Dieser Befund ist bedeutsam, weil er auch zeigt, dass es keine unterschiedliche Eignung bei gleichem Testwert zwischen den Sprachgruppen gibt. Die Chancengleichheit zwischen den Sprachgruppen ist gegeben.

Die Vorhersagekraft des Eignungstests ist ebenfalls nicht disziplinspezifisch für Humanmedizin oder Veterinärmedizin. Die Erfolgreichen weisen jeweils in beiden Disziplinen nahezu identische Testwerte auf, gleiches gilt für die nicht Erfolgreichen. Die Bestehensquote ist allerdings unterschiedlich.

Tabelle 3

Mittelwerte (m) und Standardabweichungen (s) des Testwertes für verschiedene Gruppen nach dem Bestehen der ersten Vorprüfung.

Universität	nicht bestanden			bestanden		
	n	m	s	n	m	s
Bern	46 (16%)	98,0	6,2	235 (84%)	104,2	7,8
Basel	48 (21%)	97,1	5,4	186 (79%)	102,6	8,2
Freiburg	69 (40%)	93,8	6,2	104 (60%)	102,3	9,0
Zürich	113 (25%)	98,6	6,7	347 (75%)	106,8	7,4
Geschlecht	nicht bestanden			bestanden		
männlich	100 (20%)	96,8	7,0	407 (80%)	104,6	8,3
weiblich	176 (27%)	97,2	6,3	467 (73%)	104,8	7,9
Alter und Maturität	nicht bestanden			bestanden		
geb. vor 1975, Matur vor 1996	15 (44%)	96,5	4,7	19 (56%)	105,7	8,1
geb. vor 1975, Matur nach 1995	7 (16%)	94,9	5,1	37 (84%)	101,1	6,0
geb. nach 1974	254 (24%)	97,1	6,7	816 (76%)	104,8	8,1
Sprache	nicht bestanden			bestanden		
deutsch	216 (21%)	97,1	6,3	806 (79%)	104,8	8,1
französisch	41 (53%)	95,9	7,5	37 (47%)	104,1	7,3
italienisch	19 (38%)	98,3	6,7	31 (62%)	102,7	7,7
Disziplin	nicht bestanden			bestanden		
Humanmedizin	124 (27%)	96,9	7,2	336 (73%)	105,3	8,8
Veterinärmedizin	59 (48%)	97,9	6,0	64 (52%)	104,4	6,8

### Testwert und zweite Vorprüfung

Die zweite Vorprüfung kann bisher nur durch die Personen der Kohorte 1998 absolviert worden sein. Dabei sind nur Personen angetreten, welche die erste Vorprüfung bestanden haben; Personen mit geringeren Testwerten sind in dieser Gruppe noch unterrepräsentiert. Erst in den Folgejahren werden mehr Personen mit geringeren Testwerten über Prüfungsergebnisse für die 2. Vorprüfung verfügen. Dennoch unterscheiden sich die Testwerte für Erfolgreiche und nicht Erfolgreiche signifikant (Tab. 4).

Dass die Prognosekraft des Testwertes für die 2. Vorprüfung keinesfalls geringer sein wird als für die erste Vorprüfung belegt ein Vergleich der punktbiserialen Korrelationen (Assoziations-

koeffizient Eta als Ausdruck des Zusammenhangs) zwischen Testwert und dem Bestehenskriterium. Für die erste Vorprüfung beträgt dieser Wert 0,42 und für die zweite Vorprüfung 0,40. Rechnet man bei der zweiten Vorprüfung alle Personen, die schon zur Prüfung hätten antreten können als «Nicht bestanden» dazu, beträgt Eta sogar 0,44.

### Testwert und Prüfungsnoten (erste Vorprüfung)

Neben der Vorhersage der Studiendauer ist die Studienleistung das zweite wichtige Evaluationskriterium. Dessen Untersuchung wird dadurch erschwert, dass sich die Bewertungssysteme zwischen den Universitäten und zwischen den Jahren unterscheiden.

Die erste Vorprüfung wurde in Bern 1999 (teilweise) sowie in beiden Jahren in Basel, Freiburg und Zürich in vier Fächern abgelegt. Erfasst wurde jeweils die letzte Note, wenn eine Prüfung zweimal abgelegt wurde.

Die Zusammenhänge werden mit Korrelationskoeffizienten erfasst. Werte um 0,30 bis 0,40 sind im Vergleich zu anderen Tests als Vor-

Tabelle 4

Mittelwert (m) und Standardabweichungen (s) des Testwertes für 2. Vorprüfung nach Bestehen.

Prüfungserfolg	n	m	s
Nicht bestanden	52	102,2	6,6
Bestanden	294	106,0	7,6

Tabelle 5

Korrelation des Testwertes und der Punktwerte für Untertests mit den Noten 1. Vorprüfung (Bern «Reform 2» und zusammengefasst Basel, Zürich, Freiburg).

	Mündliche Prüfung gesamt (Bern)	Multiple-Choice-Prüfung gesamt (Bern)	Physik und Physiologie	Chemie und Biochemie	Allgemeine und Humanbiologie I	Allgemeine und Humanbiologie II	Durchschnitts note
Gesamt	,55	,59	,51	,49	,50	,48	,55
Humanmedizin			,50	,47	,48	,44	,53
Veterinärmedizin			,55	,57	,60	,66	,66

Tabelle 6

Korrelation der Noten für die 1. Vorprüfung mit dem Testwert für verschiedene Teilgruppen (alle Korrelationen auf 1%-Niveau signifikant).

	n	Physik und Physiologie	Chemie und Biochemie	Allgemeine und Humanbiologie I	Allgemeine und Humanbiologie II	Durchschnitts note
<i>Studienort</i>						
Bern (Teil 1999)	64	,64	,51	,48	,58	,59
Basel	233	,39	,35	,30	,38	,41
Freiburg	169	,50	,53	,38	,39	,53
Zürich	458	,52	,49	,51	,50	,58
<i>Alter und Maturität</i>						
geb. vor 1975, Matur vor 1996	30	,50	,58	,52	,43	,59
geb. vor 1975, Matur nach 1995	34	,41	,34	,45	,31	,46
geb. nach 1974	859	,47	,44	,45	,46	,52
<i>Geschlecht</i>						
männlich	411	,46	,40	,43	,41	,49
weiblich	512	,50	,48	,47	,50	,56
<i>Sprache</i>						
deutsch	829	,47	,43	,45	,46	,52
französisch	52	,57	,55	,46	,50	,60
italienisch	42	,41	,42	,41	,50	,47

hersagewerte dabei als hoch, über 0,40 bereits als sehr hoch zu bewerten. Die Korrelationen des Testwertes beider untersuchten Jahrgänge mit den Noten der ersten Vorprüfung sind für alle Fächer und den Notendurchschnitt hoch. Auch die neu vergebenen Kreditpunkte der Prüfung in Bern korrelieren sehr hoch mit dem Testwert und den einzelnen Testvariablen. Das ist auch auf die höhere Variation und damit auf die bessere Differenzierung der Bewertung zwischen den Personen zurückzuführen.

Festzuhalten ist eine hohe Korrelation der Noten mit den Testergebnissen für beide Disziplinen, wobei sie in der Veterinärmedizin tendenziell sogar etwas höher sind. Die Prognosegüte des Tests für die Studienleistungen ist in beiden Disziplinen gegeben.

Weitere Teilgruppen werden in der Tabelle 6 verglichen. Die Prognosegüte des Testwertes für die Studiennoten ist nicht spezifisch für Universitäten (wo es auch Variationen der Prüfungsarten «Schriftlich Essay», «Schriftlich Multiple-Choice» und «Mündlich» gibt, die keinen Einfluss auf die Prognosegüte haben – vergleiche dazu [2]). Auch bezüglich der Geschlechter, der Sprachgruppen und der Altersgruppen gibt es keine Hinweise auf Unterschiede hinsichtlich der Prognosegüte des Tests, was wieder für die Wahrung der Chancengleichheit spricht.

### Schlussfolgerungen zur Eignungsfeststellung

Die Ergebnisse der vorliegenden Evaluation lassen es zu, für den Zeitraum bis zur zweiten ärztlichen Vorprüfung gesicherte Erkenntnisse zum Zusammenhang zwischen Zulassung aufgrund des Eignungstests und der Eignung zu gewinnen:

- Der Eignungstest für das Medizinstudium in der Schweiz (EMS) erfasst nachweislich Studieneignung, da 1. Personen mit höheren Testwerten die Prüfungen mit höherer Wahrscheinlichkeit und zeitlich eher bestehen, als die mit niedrigeren und 2. Personen mit höheren Testwerten auch eine höhere Studienqualität erreichen, die sich in besseren Prüfungsnoten ausdrückt. Der Zusammenhang zwischen Testergebnis und Studieneignung kann auf einem statistisch bedeutsamen Niveau festgestellt werden, welches die Verwendung des Testwertes als Zulassungskriterium rechtfertigt.
- Auch die Bedingung der Chancengleichheit ist erfüllt. Bei gleicher Eignung (gleichem Testwert) bestehen grundsätzlich die gleichen Chancen auf ein Bestehen der Prüfung unabhängig von Geschlecht, Sprache, Disziplin, Alter und Studienort. Einzig die Gruppe der Älteren mit spätem Ablegen der Maturität weist eine Besonderheit auf (höhere Bestehenswahrscheinlichkeit der Prüfung auch bei niedrigeren Testwerten – wobei die Erfolgreichen sich von den Nichterfolgreichen wie in allen anderen Gruppen unterscheiden).
- Obwohl die Ergebnisse der zweiten Vorprüfung erst für wenige Personen vorliegen, gibt es ausreichend Hinweise, dass die Vorhersagekraft des Testwertes für beide Vorprüfungen mindestens gleichermassen hoch ist.
- Die Universitäten sind jetzt in der Lage, eine kapazitätsentsprechende Zulassung vorzunehmen und dadurch die Studienbedingungen zu verbessern. Das ist der wichtigste Vorteil des EMS, der seine Anwendung rechtfertigt. Ausserdem kann mit seiner Hilfe fair entschieden werden, wer an der «Wunschuniversität» studieren kann und wer eine Umleitung in Kauf nehmen muss.
- Was sich offenbar nicht ändert, ist die Zahl der Geeigneten, die sich für ein Medizinstudium anmelden. Der Numerus clausus ist mit rund 20% abgewiesener Bewerbungen sehr mild. Von den 80% Zugelassenen tritt ein

Teil das Studium nicht an – Umleitungen von der Wunsch- an eine andere Universität sind eine wichtige Ursache. Dies führt nun dazu, dass die freiwerdenden Studienplätze an weniger Geeignete aus den 20% der Personen mit den geringeren Testwerten vergeben werden und das Niveau der Eignung der Studierenden gegenüber der Situation ohne Numerus clausus praktisch unverändert bleibt. Die Zulassung von weniger Geeigneten (nur um Kapazitäten auszunutzen) hat nachweislich wenig Erfolgsaussicht: diese Personen haben um die Hälfte verringerte Erfolgchancen bei der ersten Vorprüfung, bei der zweiten Vorprüfung sind bisher sogar nur zwei Personen aus der «Warteliste» angelangt.

- Das jetzt praktizierte Verfahren könnte ohne jegliche Systemveränderung im schlimmsten Falle dazu führen, dass die Studiendauern sich nur langsam verkürzen, die Abbrecherzahlen und Notenverteilungen sich nur unwesentlich gegenüber dem Zeitraum vor Einführung des Numerus clausus verändern. Vor allem die bisher immer auch gewollten Dissuasionseffekte durch Test und Information über Studien- und Berufsbedingungen wären hinsichtlich möglicher unerwünschter selektiver Wirkungen zu hinterfragen.

Welche Möglichkeiten zur Systemoptimierung gäbe es, um den relativen Anteil der Geeigneten im Medizinstudium zu erhöhen und dadurch auch den Einsatz der Ausbildungsressourcen zu optimieren?

- *Einflüsse auf das Bewerbungsverhalten*, indem vermehrt Geeignete angesprochen werden. Zu analysieren ist dabei, ob durch die «gewollten» Dissuasionsfaktoren Numerus clausus, Eignungstest und geschmälerter Berufsperspektive («Ärztenschwemme», Kostenregulierung u. a.) selektiv die Geeigneteren stärker abgeschreckt werden, sich zu bewerben. Dazu gehört auch die Klärung der Frage, warum der Männeranteil bei den Bewerbungen so viel geringer ist und von Jahr zu Jahr weiter abnimmt. Diese Analyse ist ohne Auswirkung auf die Chancengleichheit zu verstehen. Ein wichtiger und ermutigender Befund dieser Evaluation ist, dass Männer und Frauen bei gleichem Testwert und damit gleicher Eignung nahezu identische Chancen im Studium haben, Prüfungen in optimalen Zeiträumen zu bestehen und gute Noten zu erreichen.

- *Verzicht, Personen zum Studium zuzulassen, die nachweislich ein «Mindestniveau» der Eignung nicht erreichen* – wo die Wahrscheinlichkeit, Prüfungen zu bestehen, von vornherein als gering eingeschätzt werden muss. Dabei würde man in Kauf nehmen müssen, die festgelegten Kapazitäten zu unterschreiten (die aktuellen Kapazitäten der ersten beiden Jahre beinhalten ihrerseits noch Schätzungen von Rückzügen, Abbrüchen und Verzögerungen, die auf den Zahlen der Situation vor dem Numerus clausus beruhen).
- *Breite Einführung der Reformstudiengänge, um den Geeigneten dann solche Studienbedingungen zu verschaffen, die eine Optimierung von Studiendauer und Studienqualität erlauben, wodurch ein Studienabbruch dann einmal die Ausnahme würde.*

## Literatur

- 1 Trost G, Blum F, Fay E, Klieme E, Maichle U, Meyer M, Nauels H-U. Evaluation des Tests für Medizinische Studiengänge (TMS): Synopse der Ergebnisse. Bonn: ITB; 1998.
  - 2 Hänsgen K-D, Spicher B. EMS Eignungstest für das Medizinstudium in der Schweiz. Bericht über die Durchführung und Ergebnisse, Evaluation der Vorhersage von Prüfungserfolg – avec un résumé en français. Berichte des Zentrums für Testentwicklung und Diagnostik Band 7. Fribourg; 2001. Auch im Internet abrufbar unter [www.unifr.ch/ztd/EMS](http://www.unifr.ch/ztd/EMS).
- Hänsgen K-D, Spicher B. Zwei Jahre Numerus clausus und Eignungstest für das Medizinstudium in der Schweiz (EMS). Teil 1: Erfahrungen mit dem EMS als Zulassungskriterium. Schweiz Ärztezeitung Heft 2000;81(12):666-72.
  - Hänsgen K-D, Spicher B. Zwei Jahre Numerus clausus und Eignungstest für das Medizinstudium in der Schweiz (EMS). Teil 2: EMS und Chancengleichheit. Schweiz Ärztezeitung Heft 2000; 81(13):723-30.