



Medizinische Fakultäten

Der Ausbildungserfolg im Vergleich (I)

Die Erfolgsraten im schriftlichen Teil der Ärztlichen Vorprüfung können ein irreführendes Bild von der Qualität eines Standortes geben.

Thomas Zimmermann¹, Karl Wegscheider², Hendrik van den Bussche¹

Ranking und Benchmarking gehören heute zu den Schlüsselbegriffen der Hochschulpolitik. Auch die Medizinischen Fakultäten sollen sich nach den Vorstellungen von Wissenschaftsrat und Bundesregierung ein eigenes Profil geben und miteinander um Qualität und Produktivität der ärztlichen Ausbildung konkurrieren. Als möglicher Parameter für den Leistungsvergleich in der Ausbildung gelten die Resultate der vom Institut für medizinische und pharmazeutische Prüfungsfragen (IMPP) durchgeführten Prüfungen, darunter die zum Abschluss der vorklinischen Ausbildung durchgeführte Ärztliche Vorprüfung (ÄVP), kurz „Physikum“ genannt.

Auch in den Fakultäten werden die Resultate im Physikum als Indikatoren für die Effektivität der eigenen vorklinischen Ausbildung angesehen. Schlechte Platzierungen führen regelmäßig zu Spekulationen über die Ursachen und mögliche Auswege aus der vorgeblichen Misere; gute Resultate hingegen werden den eigenen Lehrbemühungen gutgeschrieben. Im Gegensatz zum angeblich subjektiven Charakter von Evaluationsergebnissen mittels Befragungen gelten die Ergebnisse der schriftlichen externen Prüfungen als objektiv, was implizit mit valide gleichgesetzt wird.

Vor diesem Hintergrund stellte sich die Frage, inwiefern sich die 36 Medizinischen Fakultäten der Bundesrepublik be-

züglich der Leistungen ihrer Studierenden in der ÄVP über den zehnjährigen Zeitraum 1994 bis 2004 unterscheiden. Dieser Zeitraum wurde gewählt, weil die Medizinischen Fakultäten in den fünf neuen Bundesländern erst seit 1994 an der bundesweiten schriftlichen Prüfung teilnehmen. Der quantitative Ausbildungserfolg wird im Folgenden als prozentuale Erfolgsrate im schriftlichen Teil der ÄVP definiert. Um die Ausbildungsqualität zu vergleichen, sollen später in einem weiteren Beitrag im Deutschen Ärzteblatt die Notendurchschnitte der erfolgreichen Prüflinge untersucht werden.



Die Resultate im Physikum gelten allgemein als Indikatoren für die Qualität und die Effektivität der vorklinischen Ausbildung.

Der schriftliche Teil der ÄVP bestand in den Jahren von 1970 bis 2004 aus einem Multiple-Choice-Test, der 320 Fragen aus vier Stoffgebieten (Chemie und Biochemie, Biologie und Anatomie, Physik und Physiologie, medizinische Psychologie und medizinische Soziologie) enthält. Ein Teilnehmer, der

mindestens 60 Prozent der Fragen richtig beantwortet, hat die Prüfung bestanden. Für diese Untersuchung wurden für jede Fakultät die Zahlen über Prüflinge und Prüfungserfolge aus den jährlichen Berichten des IMPP über den gesamten Untersuchungszeitraum addiert.

Indikatoren und Auswertungsmethoden

Als Bezugsgröße für den Prüfungserfolg konnten die Zulassungskohorten aus den Jahren 1992 bis 2002, die die Zentralstelle für die Vergabe von Studienplätzen (ZVS) in Dortmund zur Verfügung gestellt hat, genutzt werden. Die Zahl der Prüfungsteilnehmer, die vom IMPP ausgewiesen wird, ist als Bezugsgröße für die Einschätzung des Prüfungserfolgs wenig geeignet, denn sie ist bereits selbst erfolgsabhängig (1).

Folgende quantitative Erfolgsindikatoren wurden untersucht:

- **4-Semester-Erfolgsrate:** Quotient aus der Zahl

der erfolgreichen Prüfungsteilnehmer innerhalb der Mindeststudienzeit von vier Semestern und der Zahl der durch die ZVS vier Semester zuvor zugelassenen Studierenden. Die 4-Semester-Erfolgsrate sagt aus, wie viele Studierende einer Kohorte schnell und erfolgreich studieren, und ist somit ein Indikator

für den Ausbildungserfolg einer Fakultät in der Mindeststudienzeit.

- **Gesamterfolgsrate:** Quotient aus der Zahl der erfolgreichen Prüfungsteilnehmer und der Zahl der durch die ZVS vier Semester zuvor zugelassenen Studierenden. Die Gesamterfolgsrate entspricht der Zahl der Studierenden einer Fakultät, die die ÄVP erfolgreich abschließen, unabhängig von Semesterzahl und Zahl der Prüfungswiederholungen. Die Gesamterfolgsrate misst somit den Wirkungsgrad beziehungsweise die Ausschöpfungsrate einer Fakultät in Bezug auf die Gesamtpopulation ihrer vorklinischen Studierenden.

Auf der Basis dieser beiden Erfolgsraten wurden Fakultätsrankings erstellt. Dies sind nicht-adjustierte, „naive“ Rankings, die sich nur auf die vom IMPP berichteten Erfolgsraten beziehen. Weil an den Fakultäten unterschiedliche Rahmenbedingungen vorzufinden sind, unterliegen die beobachteten Erfolgsraten möglichen Verzerrungen. Ein fairer Vergleich der Fakultäten in Bezug auf die Ausbildungserfolge kann somit nur vorgenommen werden, wenn im Rankingmodell für verschiedene Einflussfaktoren korrigiert wird (2). Welche Messgrößen für die Adjustierung verwendet werden, muss auch pragmatisch entschieden werden, weil nur jene Einflussfaktoren in das Rankingmodell einbezogen werden können, für die den Untersuchern vergleichbare Daten zur Verfügung stehen. In dieser Untersuchung wurden folgende Einflussgrößen in das Modell einbezogen:

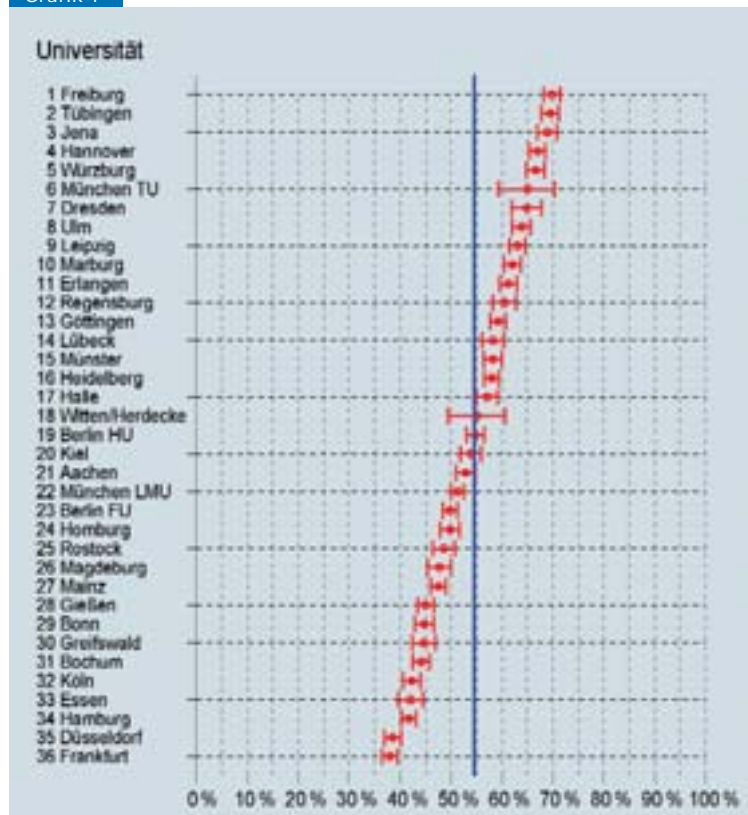
- **Summe der ZVS-Zulassungen:** Sie kann als Maß für die Größe der Fakultät und damit als ein potenzieller Einflussfaktor auf die Studienbedingungen angesehen werden.

- **Einwohnerzahl der Stadt, in der die Universität liegt (logarithmiert):** Die Größe des Fakultätsstandortes wird als Surrogatparameter für das Lernmilieu beziehungsweise das „Ablenkungspotenzial“ eines Studienortes verstanden.

- **Anteil weiblicher Studenten an den Prüfungsteilnehmern:** Ein Einfluss des Geschlechts auf den Studiererfolg wurde a priori nicht postuliert, aber auch nicht ausgeschlossen.

- **Anteil der Studierenden mit nicht-deutscher Staatsangehörigkeit bei den Prüfungsteilnehmern:** Es sei darauf hin-

Grafik 1



Ligatabelle der Medizinischen Fakultäten auf der Basis der 4-Semester-Erfolgsrate der schriftlichen Prüfungen (ohne Adjustierung)

gewiesen, dass eine nicht-deutsche Staatsangehörigkeit im Einzelfall nicht mit geringeren Kenntnissen der deutschen Sprache einhergehen muss.

- **Bundeslandspezifische NC-Werte:** Jedes Bundesland erhält für seine Abiturienten ein bestimmtes Studienplatzkontingent. Der bundeslandspezifische NC-Wert ist die Abiturnote des Bewerbers, der den letzten Platz der Bewerberliste für das jeweilige Bundeslandkontingent einnimmt. Da laut Kultusministerkonferenz (KMK 2005) durchschnittlich zwei Drittel der Studienplatzbewerber im selben Bundesland studieren, in dem sie ihr Abitur abgelegt haben, kann der NC-Wert als Indikator für das Schulleistungsniveau der Studierenden an einer Fakultät verstanden werden.

- **Personalausstattung:** Einer jüngeren Stellungnahme des Wissenschaftsrates sind die Zahlen der Mitarbeiter im wissenschaftlichen Dienst im Jahr 2001 für jede Fakultät zu entnehmen (3). Diese Zahlen stehen für die personellen Ressourcen, die prinzipiell für die Ausbil-

dung zur Verfügung stehen. Diese Zahlen liegen nur für 34 der 36 Fakultäten vor (nicht für Bochum und Witten/Herdecke, da diese Fakultäten in großem Umfang mit externen Lehrbeauftragten arbeiten) und erlauben somit nur ein Ranking für 34 anstatt 36 Fakultäten.

Die Adjustierung wurde in zwei Schritten durchgeführt, und die daraus resultierenden Rankings werden gesondert dargestellt: Zunächst wurden für die Adjustierung die populationsbezogenen Einflussfaktoren (die Größenvariablen „Größe der Stadt“ und „Größe der Fakultät“) sowie die Studierendenvariablen „Anteil Frauen“, „Anteil Nicht-Deutscher“ und „bundeslandspezifischer NC-Wert“ berücksichtigt. Beiden Variablengruppen ist gemeinsam, dass sie von den Fakultäten nicht direkt beeinflussbar sind. Daran schloss sich die zusätzliche Adjustierung für die Personalausstattung der Fakultäten an. Die Personalausstattung steht für Ausbildungsressourcen, die durch die Fakultät oder die Landesregierung verändert werden können. ▷

Die Auswertung basiert auf Meta-regressionsanalysen. Sämtliche Modellanpassungen wurden mit der im Internet verfügbaren aktuellen (Januar 2006) Version der STATA-8.0-Prozedur „metareg“ (4) durchgeführt. Die statistische Methodik wurde von Thompson und Higgins (5) beschrieben. Die Übereinstimmung unterschiedlicher Rankings wurde mit einem nicht-parametrischen Korrelationskoeffizienten, Kendall's τ , gemessen.

Im Durchschnitt über den Zehnjahreszeitraum variiert der Anteil ausländischer Studierender beträchtlich (Bundesdurchschnitt: neun Prozent, Dresden: drei Prozent, Frankfurt: 17 Prozent). Der Frauenanteil variiert zwischen 48 Prozent in Freiburg und 60 Prozent in Witten/Herdecke (Bundesdurchschnitt: 53 Prozent).

Auch der Anteil derjenigen Prüfungsteilnehmer, die im vierten Semester erstmals zur Prüfung antreten, variiert stark (Bundesdurchschnitt: 55 Prozent, Düsseldorf: 40 Prozent, Jena: 74 Prozent). Der Anteil der Prüfungswiederholer beträgt im Bundesdurchschnitt 21 Prozent und variiert ebenfalls deutlich (Düsseldorf: 28 Prozent, Witten/Herdecke: sie-

ben Prozent). Die entsprechende Tabelle findet sich in einer ausführlicheren Textfassung unter www.aerzteblatt.de/plus2506.

Die Erfolgsraten in den schriftlichen Prüfungen

Von den ZVS-zugelassenen Studierenden traten nach vier Semestern im Durchschnitt 59,9 Prozent erstmals zur Prüfung an. 53,8 Prozent der zugelassenen Studierenden bestanden die ÄVP beim ersten Versuch (4-Semester-Erfolgsrate), 6,1 Prozent fielen beim ersten Versuch durch.

Die 4-Semester-Erfolgsrate

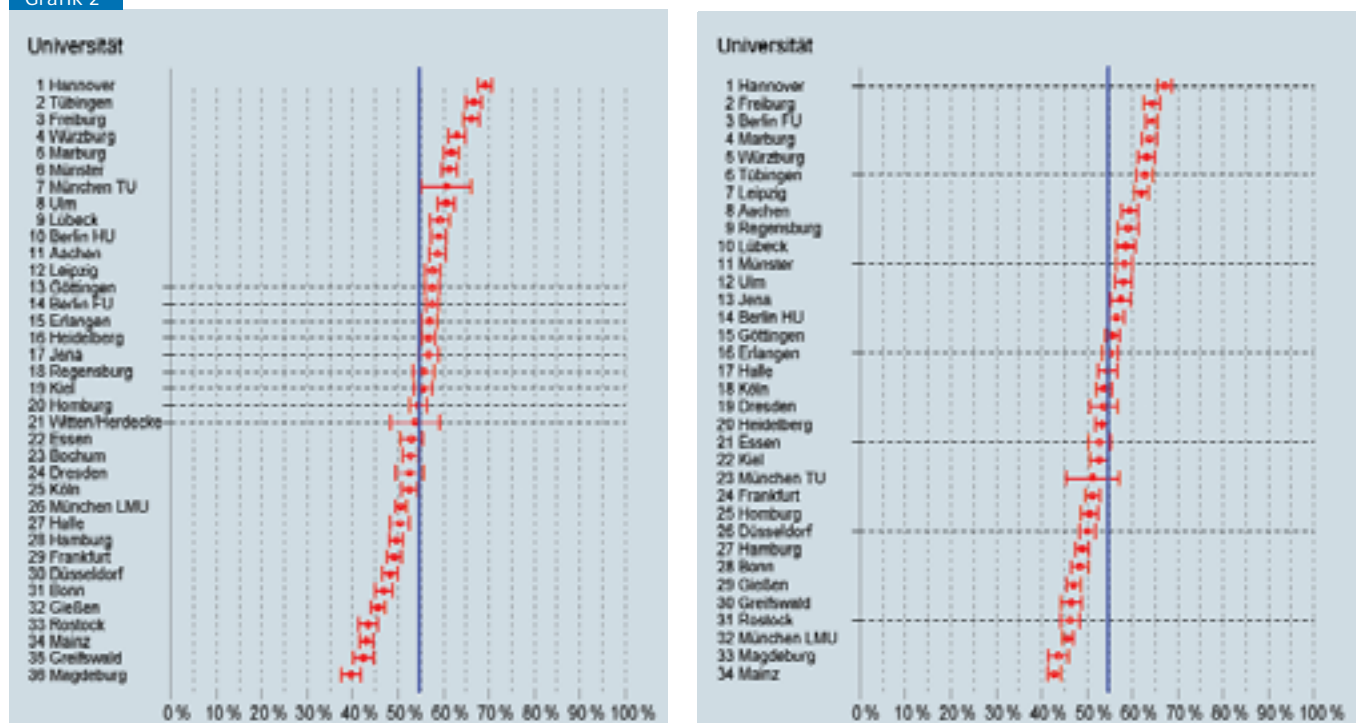
Grafik 1 zeigt die Ligatabelle auf der Basis der vom IMPP veröffentlichten, nicht-adjustierten 4-Semester-Erfolgsraten der einzelnen Fakultäten mit den 95-Prozent-Konfidenzbereichen. Die durchgehende Linie stellt das Gesamtmittel dar.

Die 4-Semester-Erfolgsraten hängen in univariaten Metaregressionen mit

mehreren Adjustierungsvariablen zusammen. Die 4-Semester-Erfolgsrate sinkt mit steigendem Ausländeranteil ($p = 0,0003$, erklärt 29,7 Prozent der Fakultätsunterschiede) und höherem bundeslandspezifischen NC-Wert ($p = 0,0133$, Erklärungswert: 14,6 Prozent). Ebenfalls gibt es einen Trend zu niedrigeren Erfolgsraten mit wachsender Größe der Stadt ($p = 0,093$). Frauenanteil und Größe der Fakultät hingegen weisen keine signifikante Assoziation mit der 4-Semester-Erfolgsrate auf.

Die linke Grafik in *Grafik 2* zeigt das in Bezug auf die Größen- und Studierendenvariablen adjustierte Ranking. Die Raten können als diejenigen 4-Semester-Erfolgsraten interpretiert werden, die die einzelnen Fakultäten aufweisen würden, wenn alle Fakultäten und Städte gleich groß wären und die Studentenschaft in allen Fakultäten die gleiche Zusammensetzung in Bezug auf Frauenanteil, Ausländeranteil und Abiturnote aufweisen würde. Die Adjustierungsvariablen erklären zusammen 33,4 Prozent der Varianz zwischen den Fakultäten. Im Modell dominiert der Ausländeranteil, dessen eigenständiger Beitrag zum Modell auch nach Absicherung

Grafik 2



Ligatabelle der Medizinischen Fakultäten auf der Basis der 4-Semester-Erfolgsrate der schriftlichen Prüfungen (linke Grafik: Adjustierung für Größen- und Studierendenvariablen; rechte Grafik: zusätzliche Adjustierung für Personalausstattung)

für mehrfaches Testen statistisch gesichert ist (multiples $p = 0,0366$). Nimmt man zusätzlich die Personalausstattung ins Modell (multiples $p = 0,0134$), so erhöht sich der Erklärungswert auf 41,9 Prozent. Im erweiterten Modell leisten auch die Stadtgröße ($p = 0,010$) und die Fakultätsgröße ($p = 0,011$) einen signifikanten Erklärungsbeitrag: Die 4-Semester-Erfolgsrate steigt bei gleicher Personalausstattung und gleichem Ausländeranteil mit der Größe der Fakultät und nimmt mit zunehmender Stadtgröße ab. Aus der erweiterten Adjustierung resultiert das in der rechts in *Grafik 2* dargestellte Ranking.

Die beiden adjustierten Rankings weisen gegenüber dem naiven Ranking in *Grafik 1* erhebliche Unterschiede auf (Kendall- τ lediglich 0,575 beziehungsweise 0,476). In *Grafik 2* sehen wir im adjustierten Ranking ohne Berücksichtigung der Personalausstattung (linke Grafik) Hannover, Tübingen und Freiburg mit 4-Semester-Erfolgsraten von mehr als 65 Prozent an der Spitze. Im unteren Tabellenbereich finden wir drei Universitäten aus den neuen Bundesländern (Rostock, Magdeburg und Greifswald) sowie Mainz.

Unter Einbeziehung der Personalausstattung ergibt sich erneut ein deutlich verändertes Ranking (Kendall- τ zwischen den beiden adjustierten Rankings = 0,701). Halle und Regensburg, insbesondere aber die FU Berlin steigen um ungefähr zehn Rangplätze nach oben, während die beiden Münchner Fakultäten deutlich sinken. Hannover und Freiburg behaupten sich an der Spitze. Rostock, Mainz und Magdeburg verbleiben auch nach dieser Korrektur am Tabellenende.

Die Gesamterfolgsrate

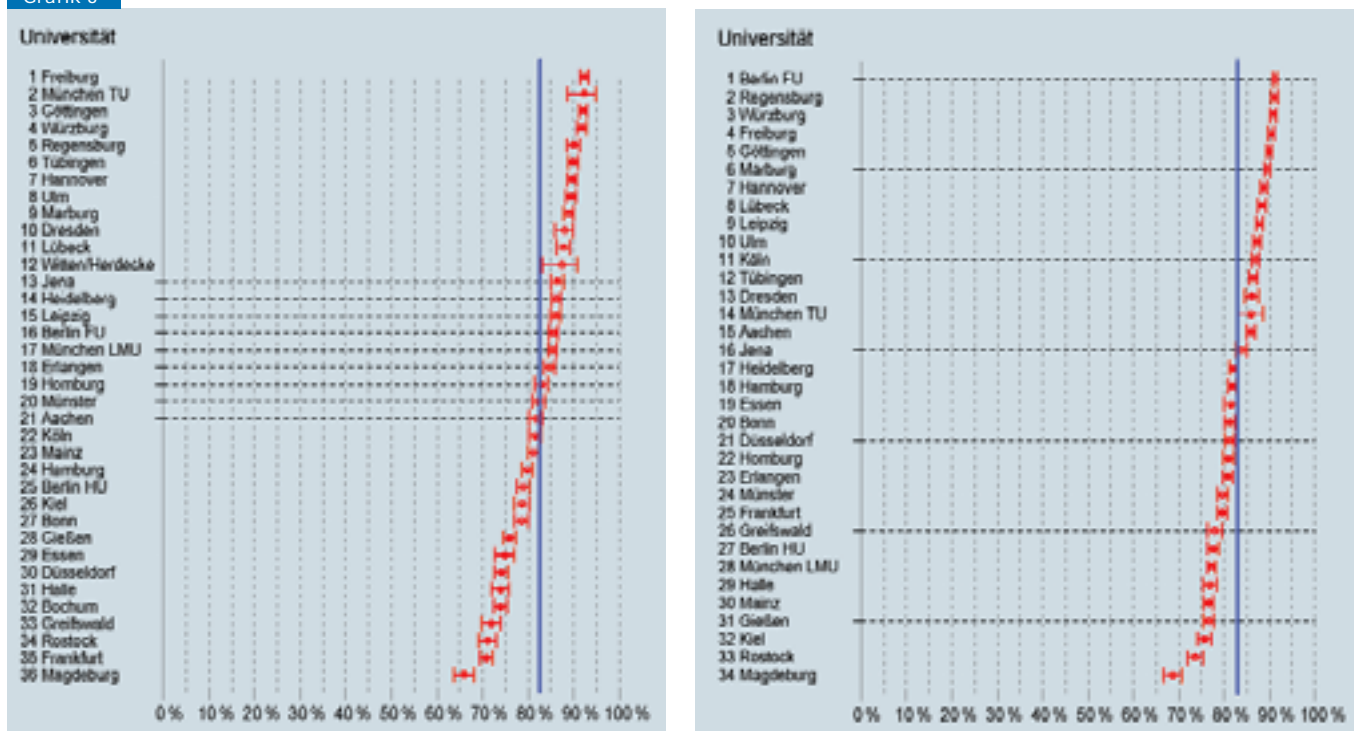
Das nicht-adjustierte Ranking auf Basis der Gesamterfolgsrate („Ausschöpfungsrate“) ist in *Grafik 3* links dargestellt. Dieses Ranking weist keine besonders enge Übereinstimmung mit der 4-Semester-Erfolgsrate auf (Kendall- $\tau = 0,635$), die Gesamterfolgsrate ist also ein eigenständiges Bewertungskriterium. Anders als bei der 4-Semester-Erfolgsrate ist keine der Größen- oder Studierendenvariablen signifikant mit der Gesamterfolgsrate assoziiert. Dementsprechend ändert sich das Ranking nach der Adjustierung für Größen- und Stu-

dierendenvariablen nur geringfügig, weswegen auf eine gesonderte Darstellung dieses Rankings verzichtet wird.

Assoziationen treten jedoch zutage, wenn man die Personalausstattung als erklärende Variable hinzunimmt. Bei simultaner Betrachtung aller sechs Einflussgrößen leisten die Personalausstattung ($p = 0,0007$), die Fakultätsgröße ($p = 0,013$) und in der Tendenz die Stadtgröße ($p = 0,061$) einen eigenständigen Erklärungsbeitrag: Wie die 4-Semester-Erfolgsrate steigt auch die Gesamterfolgsrate mit der Personalausstattung sowie mit der Größe der Fakultät und sinkt mit zunehmender Größe der Stadt. Das adjustierte Ranking ist in *Grafik 3* rechts dargestellt.

Durch die Adjustierung ändert sich erneut die Platzierung einzelner Universitäten erheblich (Kendall- τ für beide Rankings = 0,604). Während ohne Adjustierung Freiburg, TU München, Göttingen und Würzburg das Feld anführen, treten nach Adjustierung die Freie Universität Berlin und Regensburg an die Spitze. Auch verbessern die Fakultäten Köln, Frankfurt/Main und Essen ihren Rangplatz um circa zehn Plätze. Die Münchner Fakultäten sinken erneut dra-

Grafik 3



Ligatabelle der Medizinischen Fakultäten auf der Basis der Gesamterfolgsrate der schriftlichen Prüfungen (linke Grafik: ohne Adjustierung; rechte Grafik: Adjustierung für Größen- und Studierendenvariablen sowie Personalausstattung)

stisch. Magdeburg und Rostock sind in beiden Rankings am Ende der Ligatabelle zu finden.

Vergleich der Erfolgsraten

In der 4-Semester-Erfolgsrate und in der Gesamterfolgsrate beträgt der Unterschied zwischen der best- und der letztplatzierten Fakultät jeweils circa 25 Prozent.

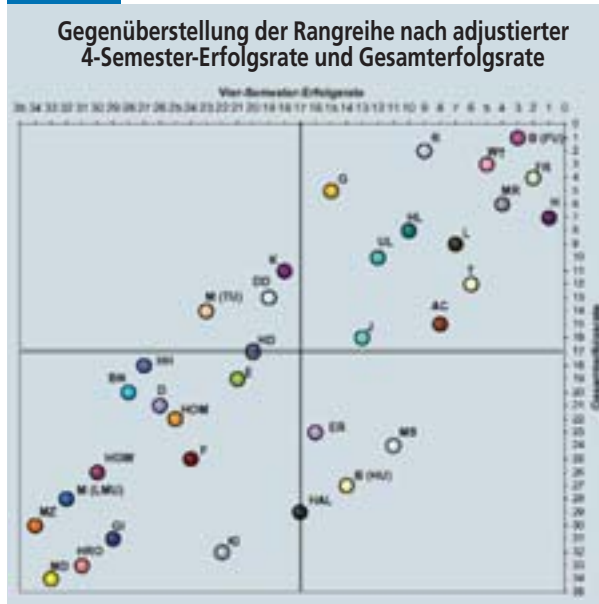
In *Grafik 4* werden die jeweiligen Rangplätze in Bezug auf die 4-Semester- und die Gesamterfolgsrate für 34 Fakultäten (ohne Bochum und Witten/Herdecke) einander gegenübergestellt. Die Verteilung zeigt eine ähnliche Grundausrichtung, im Einzelfall aber auch einzelne erhebliche Abweichungen. In einer Portfolio-Analyse lassen sich vier Prototypen unterscheiden, die jeweils durch mehrere Fakultäten repräsentiert werden:

- Fakultäten mit hohen 4-Semester- und hohen Gesamterfolgsraten (rechter oberer Quadrant: FU Berlin, Freiburg, Würzburg, Marburg sowie – mit Einschränkungen – Hannover und Regensburg): An diesen Fakultäten wird im Sinne einer geringeren Drop-out-Quote effektiv und bezogen auf einen Abschluss in der Mindeststudienzeit zugleich effizient studiert.

- Fakultäten mit niedrigen 4-Semester- und niedrigen Gesamterfolgsraten (linker unterer Quadrant: Mainz, Magdeburg, Rostock, Greifswald, Gießen und LMU München): An diesen Fakultäten wird in der Vorklinik im Vergleich zu den anderen Fakultäten sowohl relativ ineffektiv wie ineffizient studiert.

- Fakultäten mit einer hohen 4-Semester-Erfolgsrate und einer niedrigen Gesamterfolgsrate (rechter unterer Quadrant: HU Berlin, Erlangen und Münster): An diesen Fakultäten findet zwar ein relativ effizientes, jedoch kein besonders effektives Studieren in der Vorklinik statt. Das Studium wirkt relativ stark selektierend, die Studierendenpopulation teilt sich in eine schnell erfolgreiche und eine weniger erfolgreiche Gruppe.

Grafik 4



- Fakultäten mit einer niedrigen 4-Semester-Erfolgsrate und einer hohen Gesamterfolgsrate (linker oberer Quadrant). Diese Fakultäten sind gekennzeichnet durch eine hohe Effektivität und zugleich eine relativ geringe Effizienz, das Studium ist wenig selektierend. Abgesehen von der TU München entsprechen nur wenige Fakultäten diesem Typ.

Ungerechtigkeiten vermeiden

Mit der zunehmenden Verfügbarkeit zentraler Daten zur Leistungserbringung in Forschung und Lehre wächst die Möglichkeit, gängige Einschätzungen des Leistungsvermögens einzelner Universitäten empirisch zu überprüfen. Die Erwartung, eine Analyse der vor-klinischen Ausbildungserfolge an den Medizinischen Fakultäten ergäbe eine eindeutige Rangfolge nach Art der Fußball-Bundesligatabelle mit Meister, Vizemeister und Abstiegs-kandidaten muss jedoch enttäuscht werden: Zu vielfältig sind die möglichen Erfolgskriterien, zu groß die Unterschiede in den Lehr- und Lernbedingungen, zu gewichtig die Begrenzungen in der Datenstruktur und zu groß die methodischen Unsicherheiten, als dass sich die Landschaft der Fakultäten auf eine einfache Rangordnung reduzieren ließe.

Zwar liegt mit den IMPP-Daten eine zentral, unter identischen Bedingungen

erhobene Datensammlung vor, die, weil mit beruflichen Konsequenzen für die Prüflinge behaftet, ein hohes Zuverlässigkeitsniveau aufweisen dürfte. Dennoch lassen sich aus diesem Datensatz die jeweiligen Bezugsgrundgesamtheiten, das heißt die tatsächlichen Studierendenzahlen, nur ungenau rekonstruieren. Unterschiede bezüglich der Zahl der eingeklagten Studienplätze, der Studienort- oder Fachwechsler und Stichprobenverluste am Anfang und Ende des Analysezeitraums können zu Unschärfen, möglicherweise sogar zu systematischen

Verzerrungen in den Analysen führen. Zu Recht stellt der Wissenschaftsrat 2005 fest, dass „die Datensituation in nahezu allen Bereichen der Hochschulmedizinistik vollkommen unbefriedigend“ ist.

Die Analysen zeigen auch, dass bei der Untersuchung der Frage der Erfolgsraten völlig unterschiedliche Rankings resultieren – je nachdem, ob man sich auf die Mindeststudienzeit („4-Semester-Erfolgsrate“) oder auf die reale Studiendauer („Gesamterfolgsrate“) bezieht. Das erste Kriterium bildet das Ausmaß des schnellen und erfolgreichen Studierens ab, das zweite den Studienerfolg der Gesamtgruppe unabhängig von der Studiendauer und der Anzahl der Prüfungswiederholungen. Schon deshalb kann es nicht die *eine* allgemeingültige Rangliste geben. Die noch folgende Darstellung der Qualifikationsunterschiede, gemessen an den schriftlichen Notendurchschnitten, wird dies ebenfalls deutlich zeigen.

Als „Spitzengruppe“ schälen sich in dieser Untersuchung die Medizinischen Fakultäten der FU Berlin, Freiburg, Würzburg, Marburg sowie – eingeschränkt – Hannover und Regensburg heraus. Auffällig ist es, dass keine regionale Clusterung vorliegt. Auch gibt es keine Übereinstimmung mit der am 20. Januar 2006 von Wissenschaftsrat und DFG veröffentlichten Liste der Spitzenuniversitäten. ▷

Bei beiden Erfolgsindikatoren wurden ein negativer Zusammenhang mit der Größe der Stadt und positive Zusammenhänge mit dem Schulleistungsniveau der Studierenden, der Personalausstattung der Fakultät und der Größe der Fakultät gefunden. Letzteres Ergebnis stützt die These des Wissenschaftsrates, dass Medizinische Fakultäten nicht unter eine Ressourcen-Mindestgröße sinken sollten. Darüber hin-



Foto: Barbara Krbath

Personalausstattung und Größe einer Fakultät sind wichtige Faktoren, die bei einer Bewertung berücksichtigt werden müssen.

aus erklärte bei der 4-Semester-Erfolgsrate der Ausländeranteil die Unterschiede zwischen den Fakultäten in weit höherem Maße als alle anderen Faktoren, nicht jedoch bei der Gesamterfolgsrate. Vermutlich brauchen die Prüflingskohorten in Fakultäten mit einem hohen Anteil nicht-deutscher Studierender etwas länger, um die ÄVP zu bestehen, sind aber insgesamt nicht weniger erfolgreich.

Die Ergebnisse dieser Untersuchung lassen Zweifel daran aufkommen, ob die bundeseinheitlichen Prüfungen tatsächlich „Vergleiche der Prüfungsleistungen von Studierenden an den

verschiedenen Standorten“ erlauben, wie der Wissenschaftsrat 2005 behauptet. Dies ist nur haltbar, wenn man tatsächliche Unterschiede der Fakultäten (Populationsfaktoren, Betreuungsrelationen et cetera) unberücksichtigt lässt und die Ergebnisse nicht entsprechend adjustiert. Erst die Kontrolle verschiedener Störgrößen stellt aber eine wirkliche Vergleichbarkeit zwischen den Standorten her, nicht der Prüfungsmodus. Diese Untersuchung bekräftigt die Forderung anderer Autoren (1, 6), die vom IMPP veröffentlichten Daten nicht ohne weiteres für den Leistungsvergleich zwischen den Fakultäten heranzuziehen. Dem IMPP wird empfohlen, in seiner Berichterstattung auf die gesonderte Ausweisung der Ergebnisse der IMPP-Referenzgruppe zu verzichten. Sinnvoller erscheint es, die Zulassungszahl-vor-vier-Semestern als Bezugsgröße zu verwenden, weil die Größe und Zusammensetzung dieser Gruppe nicht abhängig von den jeweiligen Ausbildungsbedingungen und erfolgen ist.

Ohne Adjustierung und darüber hinaus auf der Grundlage der schmalen Basis von drei Prüfungsterminen ordnet der Wissenschaftsrat unter anderem Magdeburg, Rostock und die LMU München den „erfolgreichen Standorten nach IMPP-Ergebnissen“ zu. Diese Standorte finden sich in unserem Ranking, nach Adjustierung und im Zehnjahreszeitraum betrachtet, jedoch im unteren Tabellenbereich. Umgekehrt werden Köln und insbesondere Göttingen in nicht gerechtfertigter Weise vom Wissenschaftsrat zu den weniger erfolgreichen Standorten gezählt. Diese Beispiele mögen zeigen, wie schwierig es ist, ein wissenschaftlich haltbares Ranking zu erstellen. Auch die Stellungnahme des Wissenschaftsrates zur Qualität der Ausbildung an der Fakultät in Witten/Herdecke berührt diesen Aspekt (7).

Problematisch könnte auch die Nutzung nichtadjustierter IMPP-Daten für eine leistungsorientierte Mittelvergabe in

der Lehre auf interfakultärer Ebene sein. Dies wäre am Beispiel des Landes Baden-Württemberg näher zu untersuchen (8). Hier sind die Fragen nach der Gültigkeit der Indikatoren und Kontrollvariablen alles andere als nur akademische.

Rankings sind aus unserer Wettbewerbskultur nicht mehr wegzudenken. Die Untersuchung hat gezeigt, dass die Frage der Adjustierung für unterschiedliche Rahmenbedingungen für das Ergebnis entscheidend ist. Hier stellt sich allgemein das Problem der Verfügbarkeit und Qualität der entsprechenden Daten, vor allem aber die Frage nach der Auswahl der Adjustierungsvariablen und der Begründung für deren Gültigkeit. Wenn diese konfundierenden Einflüsse erfassbar sind, ist es nicht gerechtfertigt, nichtadjustierte Ligatabellen als Vergleichsmaßstab heranzuziehen (2). Es ist ein Gebot der Fairness, auf leichtthin erstellte einfache Zahlenvergleiche zu verzichten, um Ungerechtigkeiten zu vermeiden.

■ Zitierweise dieses Beitrags:

Dtsch Arztebl 2006; 103(25): A 1732–8

Literatur

1. Brähler E, Wittig U, Beckert C: Der Studienerfolg an Medizinischen Fakultäten – Wie viel Studienanfänger bestehen nach vier Semestern die ärztliche Vorprüfung? *Gesundheitswesen* 1998; 60: 317–21.
2. Wegscheider K: Methodische Anforderungen an Einrichtungsvergleiche („Profiling“) im Gesundheitswesen. *Z. ärztl. Fortbild. Qual. Gesund.wes.* 2004; 98: 647–54.
3. Wissenschaftsrat: Stellungnahme zu Leistungsfähigkeit, Ressourcen und Größe universitätsmedizinischer Einrichtungen. Drs. 6913-05. (2005a).
4. Sharp SJ: Meta-analysis regression. *STATA Technical Bulletin* 1998; 42: 16–22.
5. Thompson SG, Higgins JPT: How should meta-regression analyses be undertaken and interpreted? *Statist Med* 2002; 21:1559–73.
6. Robra BP, Schmitt H: Ergebnisse der IMPP-Prüfungen sind keine geeignete Grundlage für eine jährliche leistungsbezogene Mittelzuweisung an deutschen medizinischen Fakultäten. *Med Ausbildung* 2001; 18: 159–64.
7. Bussche H van den, Zimmermann T: Witten/Herdecke besser als ihr Ruf. *Dtsch Arztebl* 2005; 102(40): 2678.
8. Ministerium für Wissenschaft, Forschung und Kunst Baden-Württemberg: Bericht zum Staatshaushaltsplan für 2005/2006. Stuttgart, November 2004.

Anschrift für die Verfasser:

Dr. Thomas Zimmermann, Dipl. Psych.
 Institut für Allgemeinmedizin
 Universitätsklinikum Hamburg-Eppendorf
 Martinistraße 52
 20246 Hamburg
 E-Mail: t.zimmermann@uke.uni-hamburg.de

Für detaillierte Informationen über die Beschreibungen des Prüfungsmodus, die Berichterstattung des IMPP, die verwendeten Erfolgsindikatoren und die Auswertungsmethodik wird auf das gesonderte Dokument zur Datenbasis und Auswertungsmethodik unter www.aerzteblatt.de/plus2506 verwiesen.