

Prüfung zum Arzt für Allgemeinmedizin - PAA Jahresbericht **2007**

Monika Himmelbauer
Martin Lischka

November 2008

Medizinische Universität Wien
Besondere Einrichtung für medizinische Aus- und
Weiterbildung (BEMAW)

Inhaltsverzeichnis

1	Einleitung, Übersicht und Zusammenfassung	2
1.1	Einleitung, Übersicht	2
1.2	Zusammenfassung der Ergebnisse	2
2	Teilnehmerinnen und Teilnehmer	3
3	Zusammensetzung der Prüfungen.....	4
3.1	Leistungskriterien.....	4
3.2	Patientenalter und -geschlecht	5
3.3	Interventionsbereiche	7
3.4	Zeitliche Dynamik der Fälle	8
3.5	Leistungskategorien nach Bundesländern	9
4	Fälle.....	13
4.1	Quantitative Aspekte	13
4.2	Qualitative Problembeschreibung.....	13
4.3	Wiederverwendete Prüfungsfälle.....	16
5	Kandidaten	18
5.1	Auswertung nach Bundesländern	18
5.2	Nach Geschlecht und zeitlichem Abstand zur Promotion	21
6	Spezielle Aspekte:	22
6.1	Prüfungswiederholung	22
6.2	Qualität der Korrekturen.....	23
6.3	Ausblick, Empfehlungen.....	23

Tabellenverzeichnis

Tabelle 1:	Terminübersicht: Zahl der Kandidaten.....	3
Tabelle 2:	Leistungskriterien: Anteil an den Prüfungsfragen.....	5
Tabelle 3:	Leistungskriterien: Schwierigkeit	5
Tabelle 4:	Patientenalter: Anteil an den Prüfungsfragen	6
Tabelle 5:	Patientenalter: Schwierigkeit.....	6
Tabelle 6:	Patientengeschlecht: Anteil an den Prüfungsfragen.....	6
Tabelle 7:	Patientengeschlecht: Schwierigkeit	6
Tabelle 8:	Interventionsbereiche: Anteil an den Prüfungsfragen.....	7
Tabelle 9:	Interventionsbereiche: Schwierigkeit	8
Tabelle 10:	Dynamik: Anteil an den Prüfungsfragen	8
Tabelle 11:	Dynamik: Schwierigkeit.....	9
Tabelle 12:	Durchschnittliche Schwierigkeit p der Leistungskriterien je Bundesland.....	10
Tabelle 13:	Durchschnittliche Schwierigkeit der Leistungskriterien je Bundesland (Anova; unter Ausschluss der Kategorie ÖÄK)	10
Tabelle 14:	Schwierigkeit p der Interventionsbereiche je Bundesland	11
Tabelle 15:	Schwierigkeit der Interventionsbereiche je Bundesland (Anova; unter Ausschluss der Kategorie ÖÄK).....	11
Tabelle 16:	Wiederverwendete Fragen 200702	16
Tabelle 17:	Wiederverwendete Fragen 200705	17
Tabelle 18:	Wiederverwendete Fragen 200708	17
Tabelle 19:	Wiederverwendete Fragen 200711	18
Tabelle 20:	Mittlere Prüfungsleistung je Bundesland.....	19
Tabelle 21:	Prüfungsleistung: Multipler Paarvergleich der Bundesländer	19
Tabelle 22:	KandidatInnen: Geschlechterverteilung	21
Tabelle 23:	Prüfungsleistung je Geschlecht.....	21
Tabelle 24:	Prüfungsleistung je Geschlecht, Gruppenvergleich.....	21
Tabelle 25:	Wiederholungsantritte / davon positiv	22

1 Einleitung, Übersicht und Zusammenfassung

1.1 Einleitung, Übersicht

Im Berichtsjahr 2007 wurde die Prüfung zum Arzt für Allgemeinmedizin (PAA) wiederum an vier Terminen durchgeführt. Die Anmeldezahlen lagen von minimal 180 bis maximal 294 Anmeldungen. Damit ist die Notwendigkeit, einen weiteren Termin einzuführen, neuerlich bestätigt worden. Die angedachte Gleichverteilung konnte nicht ganz erreicht werden; ≈ 250 pro Termin wurde dreimal über- und einmal deutlich unterschritten.

Dieser Bericht orientiert sich am Bericht für das Jahr 2006. Die vier Termine wurden wiederum einzeln ausgewertet. Die Darstellung im Bericht wurde aber der besseren Übersichtlichkeit halber auf Tabellen reduziert, die den Jahresüberblick ermöglichen. Weiters wurden in den Tabellen Rohwerte eliminiert, um Vergleiche mit anderen Terminen oder dem Vorjahr übersichtlicher zu machen. Einzelergebnisse zu den vier Terminen sind dem Bericht als (elektronische) Tabellenbeilage angeschlossen, müssen aber nicht zusammen mit dem Bericht gelesen werden. Neuerlich werden aus den einzelnen Berichten je Prüfungstermin auch die qualitativen Kommentare zusammengefasst und die Qualität der Datenerfassung beschrieben.

Erstmals wird versucht, die im Bericht zu jedem Einzeltermin angeführten, besonders auffälligen Items, im Überblick des gesamten Jahres inhaltlich zu analysieren. Damit sollten Hinweise auf besonders schwierige, ausbildungsrelevante Themen gefunden und Möglichkeiten der Rückmeldung an die Auszubildenden diskutiert werden.

1.2 Zusammenfassung der Ergebnisse

Insgesamt sind 967 Personen zu den Prüfungen angetreten; der Anteil der negativen Prüfungen lag bei 6% (min. 5 . max. 7,6). Das bedeutet eine geringgradige Erhöhung der Misserfolgsquote im Vergleich zu 2006.

Die Detailergebnisse zur Zusammensetzung der Prüfungen gemäß Blueprint entsprechen in etwa denen des Jahres 2006; bei den auf Fallebene festgelegten Kriterien fällt allerdings die deutliche Überrepräsentation von Fällen mit $\approx 60\%$ Berufsalter auf. Die sich auf der Ebene einzelner Items auswirkenden $\approx 10\%$ Leistungskriterien sind von Termin zu Termin sehr unterschiedlich stark vertreten; auch ihr Verhältnis zueinander erscheint unausgewogen.

Schwierigkeitsgrade bei den einzelnen Kriterien nicht stark verändert. Bei den Leistungsbereichen fällt auf, dass die in den ersten Jahren meist sehr schwierigen & Unspezifische Allgemeinreaktionen%& nun zu den leichtesten Kategorien gehören. Neuerlich gibt es sowohl bei den Leistungs- als auch bei den Interventionsbereichen signifikante Unterschiede.

Die erstmals im Jahresüberblick untersuchten und versuchsweise inhaltlich interpretierten & Ausreißer%& (d.s. auffallend schwierige Items, die von den anderen Items innerhalb des betreffenden Falles deutlich abweichen) entfielen in höherem Ausmaß als erwartet auf fachspezifische Probleme als auf & allgemeine%& Aspekte.

Der Vergleich der Mittelwerte der erzielten Prüfungsergebnisse in der Auswertung nach & Anmeldung bei Landesärztekammer%& ergibt für Wien im Berichtsjahr den niedrigsten Wert (und die größte Streuung).

Neuerlich sind die Prüfungsergebnisse von Ärztinnen deutlich besser als von Ärzten. Der negative Zusammenhang zwischen erzieltm Prüfungsergebnis und zunehmendem Abstand der Prüfung von der Graduierung wird auch 2007 bestätigt. Bei Prüfungswiederholungen sinkt die Erfolgsrate ab dem 3. Antritt deutlich; es werden wieder einzelne, sehr hohe Antrittsraten registriert.

2 Teilnehmerinnen und Teilnehmer

Die Anmeldezahlen pro Prüfungstermin lagen bei drei Terminen über der angestrebten Größenordnung von 250 pro Termin und einmal deutlich darunter (Tab. 1). Die Gesamtzahl der KandidatInnen war im Vergleich zum Vorjahr deutlich höher.

Die für 2007 registrierten Misserfolgsquoten und die sonstigen Kennwerte der einzelnen Termine können als einigermaßen stabilisierte Werte im Bezug auf die zu prüfende Population und die verwendete Prüfungs- und Bewertungsmethodik angesehen werden; die Prüfung kann als & etabliert%& bezeichnet werden.

Für das Jahr 2007 liegen keine bedeutsamen Informationen vor, dass Prüfungsfälle samt Lösungsschlüssel bekannt geworden wären.

Tabelle 1: Terminübersicht: Zahl der Kandidaten

Termine	200702	200705	200708	200711	Gesamt
Angemeldet	294	276	180	263	1.013
Angetreten	279	265	173 + 1	249	966 + 1
Durchgefallen	14	20	9	15	58
Durchgefallen %	5%	7,6%	5,2%	6%	6%

Quelle: *Österreichische Akademie der Ärzte, BEMAW 2007*
+1: überlange Ausbildungszeit im Ausland

3 Zusammensetzung der Prüfungen

3.1 Leistungskriterien

Es gibt für die Verteilung der Leistungskriterien, die zwar auf Ebene einer Unterfrage definiert sind - sich aber auf Item-Ebene auswirken, keine anteilmäßigen Vorgaben im Sinn des Blueprints. Dieser führt Gewichtungen nur für Kriterien an, die auf Fallebene definiert (und bei der Zusammenstellung einer Prüfung handhabbar) sind. Es erscheint aber sinnvoll, auch die Verteilung der inhaltlich bedeutsamen Leistungskriterien zu monitieren. Dabei wurde diesmal das relative Gewicht der Unterfragen so berücksichtigt, dass als Grundgesamtheit die Summe der Items genommen wurde und damit die Zahl der Items je Unterfrage in die Berechnung eingegangen ist.

Die Gegenüberstellung der Anteile der auf diese Weise berechneten Leistungskriterien bei den 4 Terminen zeigt von Termin zu Termin beträchtliche Unterschiede (Tab. 2). Anamnese hat beispielsweise eine Bandbreite von 9,2% bis 23,1%, Therapiekonzept von 2,7% bis 25% - Beratung ist dagegen vergleichsweise ausgeglichen. Ethik war diesmal in 3 Terminen vertreten (2006: nur einmal).

Es muss im Rahmen dieses Berichts unklar bleiben,

- ob die diesmal festgestellten Schwankungen als Ergebnis einer Überbewertung durch Berücksichtigung der Zahl der Items pro Unterfrage anzusehen sind,
 - oder ob sie die Verteilung der Leistungskriterien im Pool spiegeln,
 - oder ob sie sich aus Effekten bei der Zusammenstellung eines konkreten Prüfungstermins ergeben.
 - Im ersten Fall (Überbewertung) könnte man einwenden, dass die Unterfrage mit einer durchschnittlichen Schwierigkeit unterschiedlich vieler Items letztlich die inhaltlich bedeutsame Einheit ist (eine Unterfrage mit zwei Items hat bei Betrachtung auf Ebene der Unterfrage das gleiche Gewicht wie eine mit vier Items, bei Betrachtung auf Item-Ebene aber nur halb so viel).
 - Im zweiten Fall (Charakteristik des Pools) müsste überlegt werden, ob es bei den FragenautorInnen eine Tendenz gibt, in bevorzugter Weise auf bestimmte Leistungskriterien abzustellen und andere Leistungsarten weniger zu berücksichtigen.
 - Im dritten Fall wäre zu fragen, ob bei der Fragenauswahl neben den Blueprintkategorien auf Fallebene auch Aspekte der Verteilung der Leistungskriterien zu beachten wären . bzw. ob das überhaupt leistbar erscheint.
- Eine eingehende Analyse sollte möglich werden, wenn ein neuer PVA eingesetzt wird.

Das Schwierigkeitsprofil der Leistungskriterien entspricht auch bei Berechnung auf Item-Ebene etwa dem des Vorjahres (Anamnese am schwierigsten, Untersuchung am leichtesten) . Tabelle 3; Ethik fällt heraus und wird wegen der insgesamt weiterhin geringen Fallzahl nicht speziell bewertet.

eil an den Prüfungsfragen

	200702	200705	200708	200711
Anamnese	13,5%	12,2%	9,2%	23,1%
Untersuchung	26,9%	33,6%	19,7%	36,2%
Diagnose	29,1%	14,1%	23,2%	21,4%
Dringliches Handeln	8,5%	5,7%	6,6%	7%
Therapiekonzept	11,7%	18,3%	25%	2,7%
Beratung	9,9%	15,7%	14,5%	9,6%
Ethik	0,4%	0,4%	1,8%	0%

Quelle: BEMAW 2007

Tabelle 3: Leistungskriterien: Schwierigkeit

		Gesamt	
		p	SD
Anamnese	2006	0,59	0,12
	2007	0,61	0,13
Untersuchung	2006	0,74	0,09
	2007	0,72	0,09
Diagnose	2006	0,68	0,09
	2007	0,70	0,10
Dringliches Handeln	2006	0,70	0,12
	2007	0,66	0,13
Therapiekonzept	2006	0,66	0,08
	2007	0,67	0,12
Beratung	2006	0,68	0,10
	2007	0,68	0,10
Ethik	2006	0,74	0,14
	2007	0,53	0,44

Quelle: BEMAW 2007

p: Schwierigkeitsindex

3.2 Patientenalter und -geschlecht

Während Säuglings- und Jugendalter gemäß Blueprint vertreten sind, erscheinen Klein- und Schulkinder unterrepräsentiert. sBerufsalter%ist neuerlich in hohem Maß überrepräsentiert (Tab. 4).

Die in Tab. 5 gegenüber gestellten Schwierigkeitswerte schwanken bei niedrigen Blueprintvorgaben im Jahresvergleich; in den stark besetzten Bereichen sind die Schwierigkeiten stabil und gut vergleichbar.

Die Ausgewogenheit der Repräsentierung männlicher und weiblicher PatientInnen ist etwa im selben Ausmaß gegeben wie 2006. Auch die Schwierigkeit von Fällen mit männlichen oder weiblichen PatientInnen ist im Mittel wiederum ausgeglichen (Tab. 6 und 7).

Anteil an den Prüfungsfragen

	200702	200705	200708	200711	Soll
Säuglinge	4%	4%	4%	4%	5%
Kleinkinder	4%	8%	8%	4%	10%
Schulkinder	4%	4%	8%	4%	10%
Jugendalter	12%	4%	4%	4%	5%
Berufsalter	52%	45%	44%	59%	35%
Pensionsalter	20%	31%	28%	21%	30%
Hochbetagte	4%	4%	4%	4%	5%

Quelle: BEMAW 2007

Tabelle 5: Patientenalter: Schwierigkeit

		Gesamt	
		p	SD
Säuglinge	2006	0,55	0,34
	2007	0,77	0,22
Kleinkinder	2006	0,73	0,13
	2007	0,64	0,22
Schulkinder	2006	0,61	0,18
	2007	0,58	0,22
Jugendalter	2006	0,70	0,17
	2007	0,66	0,15
Berufsalter	2006	0,69	0,08
	2007	0,68	0,08
Pensionsalter	2006	0,70	0,08
	2007	0,67	0,10
Hochbetagte	2006	0,70	0,18
	2007	0,83	0,15

Quelle: BEMAW 2007

p: Schwierigkeitsindex

Tabelle 6: Patientengeschlecht: Anteil an den Prüfungsfragen

	200702	200705	200708	200711	Soll
Weiblich	56%	54%	48%	54%	50%
Männlich	44%	46%	52%	46%	50%

Quelle: BEMAW 2007

Tabelle 7: Patientengeschlecht: Schwierigkeit

		Gesamt	
		p	SD
Weiblich	2006	0,68	0,08
	2007	0,66	0,08
Männlich	2006	0,68	0,08
	2007	0,70	0,10

Quelle: BEMAW 2007

p: Schwierigkeitsindex

3.3 Interventionsbereiche

Wie in früheren Jahren konnte auch diesmal die Verteilung der Fälle auf die verschiedenen Interventionsbereiche gut an die Vorgaben des Blueprint angepasst werden (geringfügige Abweichungen von Termin zu Termin sind auf Grund der beschränkten Fallzahl unvermeidbar Tab. 8).

Die drei schwierigsten Interventionsbereiche sind diesmal Augenerkrankungen (0,58) und Haut- und Geschlechtskrankheiten (0,55). Die drei leichtesten Interventionsbereiche sind HNO (0,74) und Stoffwechsel, Endokriniem, Leber (0,74) . Tabelle 9. Der früher besonders schwierige Bereich unspezifische Allgemeinreaktion lag diesmal mit 0,73 im leichteren Bereich.

Tabelle 8: Interventionsbereiche: Anteil an den Prüfungsfragen

	200702	200705	200708	200711	Soll
Unspezifische Allgemeinreaktion	12%	15%	12%	13%	15%
Brustraum, Herz-Kreislaufsystem	8%	11%	8%	13%	10%
Bauchraum	12%	8%	8%	8%	10%
Muskel-, Skelett-, Gelenksystem	8%	8%	8%	13%	10%
Stoffwechsel, Endokriniem, Leber	8%	11%	8%	8%	10%
Periphere Gefäße	4%	4%	4%	4%	3%
Blut-, Lymphsystem	4%	4%	4%	4%	3%
Hals-, Nasen-, und Ohrenkrankheiten	4%	4%	4%	4%	5%
Augenerkrankungen	4%	4%	8%	4%	3%
Gehirn und Nervensystem	4%	4%	8%	4%	5%
Psyche und Psychosoziale Problematik	8%	7%	12%	8%	8%
Haut- und Geschlechtskrankheiten	8%	4%	4%	4%	5%
Urogenitaler Bereich	4%	4%	4%	4%	5%
Gynäkologischer Bereich	4%	4%	4%	4%	5%
Unfälle, häufige Verletzungen	4%	4%	4%	4%	3%

Quelle: BEMAW 2007

Schwierigkeit

		Gesamt	
		p	SD
Unspezifische Allgemeinreaktion	2006	0,68	0,10
	2007	0,73	0,11
Brustraum, Herz-Kreislaufsystem	2006	0,74	0,10
	2007	0,66	0,13
Bauchraum	2006	0,76	0,10
	2007	0,70	0,14
Muskel-, Skelett-, Gelenksystem	2006	0,68	0,13
	2007	0,69	0,13
Stoffwechsel, Endokriniem, Leber	2006	0,69	0,17
	2007	0,74	0,14
Periphere Gefäße	2006	0,70	0,20
	2007	0,64	0,24
Blut-, Lymphsystem	2006	0,59	0,16
	2007	0,66	0,17
Hals-, Nasen-, und Ohrenkrankheiten	2006	0,71	0,17
	2007	0,74	0,18
Augenerkrankungen	2006	0,61	0,23
	2007	0,58	0,19
Gehirn und Nervensystem	2006	0,69	0,16
	2007	0,65	0,17
Psyche und Psychosoziale Problematik	2006	0,64	0,14
	2007	0,62	0,12
Haut- und Geschlechtskrankheiten	2006	0,69	0,16
	2007	0,55	0,21
Urogenitaler Bereich	2006	0,69	0,18
	2007	0,71	0,14
Gynäkologischer Bereich	2006	0,65	0,20
	2007	0,66	0,18
Unfälle, häufige Verletzungen	2006	0,68	0,16
	2007	0,71	0,16

Quelle: BEMAW 2007

p: Schwierigkeitsindex

3.4 Zeitliche Dynamik der Fälle

Aus Tabelle 10 ist zu entnehmen, dass akute Erkrankungen tendenziell überrepräsentiert sind; insgesamt ist aber eine Verbesserung der Anpassung an die Vorgaben des Blueprints festzustellen.

Bei diesen Kriterien hat sich das Schwierigkeitsprofil im Berichtsjahr verändert. Notfälle repräsentieren nicht mehr die leichtesten Fragen; die Schwierigkeiten der einzelnen Bereiche haben sich einander angenähert.

Tabelle 10: Dynamik: Anteil an den Prüfungsfragen

	200702	200705	200708	200711	Soll
Notfall	16%	15%	12%	17%	20%
akute Erkrankung	44%	46%	48%	33%	35%
chronische Erkrankung	18%	31%	32%	38%	35%
Prävention, Prophylaxe	12%	8%	8%	12%	10%

Quelle: BEMAW 2007

		Gesamt	
		p	SD
Notfall	2006	0,72	0,11
	2007	0,63	0,13
akute Erkrankung	2006	0,70	0,07
	2007	0,68	0,08
chronische Erkrankung	2006	0,67	0,08
	2007	0,70	0,08
Prävention, Prophylaxe	2006	0,63	0,10
	2007	0,68	0,11

Quelle: BEMAW 2007

p: Schwierigkeitsindex

3.5 Leistungskategorien nach Bundesländern

Die Aufgliederung nach Leistungskriterien und Bundesländern (exakte Bezeichnung des Merkmals „Anmeldung bei der Landesärztekammer“) ergibt zwar ein ähnliches Bild wie 2006; die Unterschiede erscheinen für den Berichtszeitraum aber größer ausgefallen zu sein.

Tabelle 12 gibt die durchschnittliche Schwierigkeit der Leistungskriterien je Bundesland an. Die Schwierigkeit von „Anamnese“ liegt (gerundet) zwischen 0,58 und 0,63 (die Schwierigkeit wird hier als Wert p zwischen 0 und 1 angegeben; 0,63 heißt, dass durchschnittlich 63% richtig geantwortet haben. je höher der Wert, desto leichter die Frage). Die Abstände sind unterschiedlich groß (Dringliches Handeln 0,63 . 0,72; Untersuchung 0,69 . 0,74; Diagnose 0,67 . 0,73; Therapiekonzept 0,65 . 0,69; Beratung 0,65 . 0,72; Ethik . min. 0,44 . max. 0,76 . ist für eine Bewertung mit zu wenig Fragen vertreten).

Bei den Leistungskriterien sind die Unterschiede zwischen den Bundesländern mit Ausnahme von „Therapiekonzept“ signifikant. zumindest zwischen dem je Kriterium besten und schlechtesten Bundesland (Tab. 13). Wie im Vorjahr wird hier neuerlich darauf hingewiesen, dass es bei der Bewertung dieser Ergebnisse vor allem darauf ankommt, welche Unterschiede man als inhaltlich bedeutsam einstuft.

Bei einer Aufgliederung nach Interventionsbereichen kann man feststellen, dass die im Vorjahr bei „Unspezifische Allgemeinreaktionen“ noch stärker schwankenden Ergebnisse diesmal ziemlich ausgeglichen sind. Tabelle 14 gibt eine Übersicht über die Bundesländerergebnisse für alle fünfzehn Interventionsbereiche. (Interventionsbereich ist nicht fachspezifisch definiert: Interventionsbereich „Brustraum“ oder „Bauchraum“ kann in der Allgemeinmedizin sowohl für ein internistisches wie ein chirurgisches Problem stehen.)

Signifikante Bundesländerunterschiede findet man diesmal bei vielen Interventionsbereichen (Kautelen für die Interpretation. siehe oben!). Hervorzuheben sind Brustraum, Bauchraum, Muskel-, Skelet- und Gelenksystem sowie Blut und Lymphsystem (Tab. 15).

Wichtigkeit p der Leistungskriterien je Bundesland

B.-Länder	Anamnese	Dringliches Handeln	Untersuchung	Diagnose	Therapiekonzept	Beratung	Ethik
B	,5952	,6636	,7336	,7312	,6825	,6976	,4737
K	,6248	,6722	,7380	,7090	,6639	,6765	,6000
NÖ	,6257	,6672	,7395	,7170	,6862	,6965	,4393
OÖ	,6284	,6636	,7285	,7086	,6806	,6953	,6150
NÖ	,6154	,6563	,7228	,7039	,6778	,6898	,5652
St	,6121	,6566	,7195	,7025	,6768	,6853	,4894
T	,5989	,6465	,6894	,6934	,6802	,6791	,5375
V	,6669	,7189	,7466	,7232	,6646	,7195	,7600
W	,5766	,6266	,6854	,6673	,6450	,6462	,4451

Quelle: BEMAW 2007

Tabelle 13: Durchschnittliche Schwierigkeit der Leistungskriterien je Bundesland (Anova; unter Ausschluss der Kategorie ÖÄK)

			Quadratsumme	df	Mittel der Quadrate	F	Signifikanz
Anamnese * Ärztammer	Zwischen den Gruppen	(Kombiniert)	,443	8	,055	3,330	,001
	Innerhalb der Gruppen		15,904	957	,017		
	Insgesamt		16,347	965			
Dringliches Handeln * Ärztammer	Zwischen den Gruppen	(Kombiniert)	,345	8	,043	2,788	,005
	Innerhalb der Gruppen		14,794	957	,015		
	Insgesamt		15,139	965			
Untersuchung * Ärztammer	Zwischen den Gruppen	(Kombiniert)	,422	8	,053	6,719	,000
	Innerhalb der Gruppen		7,520	957	,008		
	Insgesamt		7,942	965			
Diagnose * Ärztammer	Zwischen den Gruppen	(Kombiniert)	,308	8	,038	4,361	,000
	Innerhalb der Gruppen		8,447	957	,009		
	Insgesamt		8,755	965			
Therapiekonzept * Ärztammer	Zwischen den Gruppen	(Kombiniert)	,200	8	,025	1,894	,058
	Innerhalb der Gruppen		12,639	957	,013		
	Insgesamt		12,839	965			
Ethik * Ärztammer	Zwischen den Gruppen	(Kombiniert)	4,999	8	,625	3,198	,001
	Innerhalb der Gruppen		138,324	708	,195		
	Insgesamt		143,323	716			
Beratung * Ärztammer	Zwischen den Gruppen	(Kombiniert)	,361	8	,045	5,027	,000
	Innerhalb der Gruppen		8,594	957	,009		
	Insgesamt		8,955	965			

Quelle: BEMAW 2007

Interventionsbereiche je Bundesland

	Bundesländer								
	B	K	NÖ	OÖ	S	St	T	V	Wien
Unspezifische Allgemeinreaktion	,7267	,7343	,7374	,7393	,7409	,7357	,7211	,7471	,7070
Brustraum, Herz-Kreislaufsystem	,6630	,6978	,6766	,6779	,6734	,6595	,6410	,7132	,6329
Bauchraum	,7126	,7176	,7076	,7131	,7109	,7096	,6999	,7458	,6590
Muskel-, Skelett- und Gelenksystem	,6957	,6982	,7181	,6989	,6784	,6823	,6814	,6888	,6417
Stoffwechsel, Endokrinium, Leber	,7789	,7590	,7607	,7446	,7358	,7470	,7128	,7599	,7115
Periphere Gefäße	,7110	,6336	,6497	,6651	,6309	,6321	,6098	,6356	,5957
Blut- und Lymphsystem	,6656	,6755	,6760	,6761	,6662	,6862	,6563	,7072	,6034
Hals-, Nasen- und Ohrenkrankheiten	,7622	,7542	,7341	,7358	,7605	,7306	,7440	,7195	,7167
Augenerkrankungen	,5716	,6026	,6015	,6055	,5995	,5665	,5461	,6470	,5421
Gehirn und Nervensystem	,6740	,6575	,6862	,6408	,6601	,6313	,6302	,7000	,6152
Psyche und Psychosoziale Problematik	,5949	,5971	,6309	,6362	,6164	,6208	,6087	,6495	,5928
Haut- und Geschlechtskrankheiten	,5806	,5543	,5752	,5435	,5329	,5624	,5233	,5711	,5191
Urogenitaler Bereich	,7557	,7136	,7111	,7260	,7194	,7115	,7183	,7684	,6797
Gynäkologischer Bereich	,6926	,6616	,6863	,6849	,6754	,6587	,6230	,7171	,6278
Unfälle und häufige Verletzungen	,6915	,6707	,7186	,7214	,6937	,7058	,7344	,7652	,6709

Quelle: BEMAW 2007

Tabelle 15: Schwierigkeit der Interventionsbereiche je Bundesland (Anova; unter Ausschluss der Kategorie ÖÄK)

			Quadrat summe	df	Mittel der Quadrate	F	Signifikanz
Unspezifische Allgemeinreaktion * Ärztekammer	Zwischen den Gruppen	(Kombiniert)	,150	8	,019	1,665	,103
	Innerhalb der Gruppen		10,767	957	,011		
	Insgesamt		10,916	965			
Brustraum, Herz-Kreislaufsystem * Ärztekammer	Zwischen den Gruppen	(Kombiniert)	,452	8	,057	3,498	,001
	Innerhalb der Gruppen		15,464	957	,016		
	Insgesamt		15,916	965			
Bauchraum * Ärztekammer	Zwischen den Gruppen	(Kombiniert)	,456	8	,057	3,032	,002
	Innerhalb der Gruppen		17,989	957	,019		
	Insgesamt		18,445	965			
Muskel-, Skelett- und Gelenksystem * Ärztekammer	Zwischen den Gruppen	(Kombiniert)	,589	8	,074	4,314	,000
	Innerhalb der Gruppen		16,325	957	,017		
	Insgesamt		16,914	965			
Stoffwechsel, Endokrinium, Leber * Ärztekammer	Zwischen den Gruppen	(Kombiniert)	,367	8	,046	2,586	,008
	Innerhalb der Gruppen		16,969	957	,018		
	Insgesamt		17,336	965			

		(Kombiniert)	,688	8	,086	1,477	,161
		n	55,755	957	,058		
	Insgesamt		56,443	965			
Blut- und Lymphsystem * Ärztchamber	Zwischen den Gruppen	(Kombiniert)	,845	8	,106	3,839	,000
	Innerhalb der Gruppen		26,347	957	,028		
	Insgesamt		27,192	965			
Hals-, Nasen- und Ohrenkrankheiten * Ärztchamber	Zwischen den Gruppen	(Kombiniert)	,162	8	,020	,602	,777
	Innerhalb der Gruppen		32,186	957	,034		
	Insgesamt		32,348	965			
Augenerkrankungen * Ärztchamber	Zwischen den Gruppen	(Kombiniert)	,789	8	,099	2,767	,005
	Innerhalb der Gruppen		34,133	957	,036		
	Insgesamt		34,922	965			
Gehirn und Nervensystem * Ärztchamber	Zwischen den Gruppen	(Kombiniert)	,647	8	,081	2,822	,004
	Innerhalb der Gruppen		27,444	957	,029		
	Insgesamt		28,091	965			
Psyche und Psychosoziale Problematik * Ärztchamber	Zwischen den Gruppen	(Kombiniert)	,297	8	,037	2,635	,007
	Innerhalb der Gruppen		13,493	957	,014		
	Insgesamt		13,790	965			
Haut- und Geschlechtskrankheiten * Ärztchamber	Zwischen den Gruppen	(Kombiniert)	,421	8	,053	1,172	,313
	Innerhalb der Gruppen		43,007	957	,045		
	Insgesamt		43,428	965			
Urogenitaler Bereich * Ärztchamber	Zwischen den Gruppen	(Kombiniert)	,382	8	,048	2,418	,014
	Innerhalb der Gruppen		18,887	957	,020		
	Insgesamt		19,269	965			
Gynäkologischer Bereich * Ärztchamber	Zwischen den Gruppen	(Kombiniert)	,669	8	,084	2,596	,008
	Innerhalb der Gruppen		30,811	957	,032		
	Insgesamt		31,480	965			
Unfälle und häufige Verletzungen * Ärztchamber	Zwischen den Gruppen	(Kombiniert)	,584	8	,073	2,870	,004
	Innerhalb der Gruppen		24,342	957	,025		
	Insgesamt		24,926	965			

Quelle: BEMAW 2007

4 Fälle

4.1 Quantitative Aspekte

Wie im Vorjahr werden auch heuer die Auflistungen der Kennwerte zu den einzelnen Prüfungsfällen im Jahresbericht nicht wiederholt; diese Kennwerte (insbesondere Mittelwert und Trennschärfe) sind in den zu den einzelnen Terminen erstellten Berichten enthalten und werden in der jeweiligen sGrenzfallsitzung%von der Prüfungskommission diskutiert.

4.2 Qualitative Problembeschreibung

Wie im Vorjahr kamen auch in diesem Berichtszeitraum qualitative Probleme in Form von sAusreißern%in zwei Hauptformen vor. Als Ausreißer werden Items bezeichnet, deren Schwierigkeit deutlich höher als die der anderen zugehörigen Items einer Unterfrage bzw. eines Falls.

Einige wenige sogenannte Ausreißer waren auf methodische Probleme zurückzuführen. Mitunter ist es vorgekommen, dass im Antwortschlüssel eine Antwort verlangt wird, die sich durch Angaben im Fall eigentlich erübrigt. Dann schreiben nur einige wenige Kandidaten diese Antwort an, sodass dieses Item als überproportional schwierig erscheint.

Pro Termin gab es im Durchschnitt ein methodisches Problem, der Rest wurde als inhaltlich i.S. eines sAusbildungsproblems%eingestuft. Um die immer wieder geforderte Rückmeldung an die Auszubildenden qualitativ formulieren zu können, wurde versucht, diese als inhaltlich interpretierten Ausreißer näher zu charakterisieren.

Bei den vier Terminen wurden Ausreißer in jeweils vergleichbarer Größenordnung gefunden; das waren beim Termin im Februar 13, im Mai 18, im August 19 und im November 20 Items. Häufig erscheinen die Ausreißer auf inhaltliche Probleme rückführbar zu sein.

Zur Methodik: Insgesamt wurden 70 sausreißende%Items bearbeitet. Alle verfügbaren Informationen (Fallname, Problemstellung, Leistungskriterium des Ausreißers und Metadaten zum übergeordneten Fall) wurden in einer excel-Tabelle zusammengefasst. Zusätzlich wurde pro Item eine Einschätzung des Berichtstatters festgehalten: Allgemeines oder fachspezifisches Problem, inhaltliche oder sonstige Auffälligkeiten, Mehrfachzuordnungen waren möglich. Diese Einschätzungen wurden erstmals vorgenommen, die Ergebnisse sind jedenfalls als vorläufig zu betrachten. Sie gehen auf nur eine Person zurück und dienen zunächst der Suche nach einer Arbeitshypothese für weitere Auswertungen.

Die ergab sich bei den meisten Kriterien kein wesentlicher Unterschied zur angestrebten Verteilung gemäß Blueprint. Größere Abweichungen gab es lediglich im Bereich Augenerkrankungen (7,1% gegenüber SOLL 3%), Psyche und soziale Problematik (15,7% gegenüber SOLL 8%), Haut- und Geschlechtskrankheiten (11,4% gegenüber SOLL 5%). (Abweichungen nach unten sind häufig, werden hier . weil positiv%o naturgemäß aber nicht im Detail dargestellt.)

Wenn man die Zuordnung der Ausreißer zur Altersstufe überprüft, findet man für 2007 Ausreißer etwas überproportional bei sBerufsalter%(42,9% gegenüber SOLL 35 . siehe aber auch die Abweichung dieses Kriteriums vom Sollwert), die Verteilung entspricht sonst aber im Wesentlichen dem Blueprint.

Auch bezüglich des Geschlechts der Patienten und der zeitlichen Dynamik der Fälle gibt es keine auffallenden Abweichungen. Die Verteilung der Ausreißer entspricht auch hier in etwa der Gewichtung im Blueprint.

Bei den Leistungsbereichen liegen die sAusreißer%o meist in der gleichen Verteilung vor, wie diese in der Prüfung vertreten sind (Tab. 2) . mit 14,3 % Ausreißern für sDiagnose%ist dieses Kriterium bei den Problemfällen allerdings UNTERrepräsentiert, während 18,6% bei sBeratung% tendenziell höher liegen als der Ausgangsverteilung entspricht. Dieser Befund könnte in die Richtung weisen, dass es . wie schon in früheren Berichten festgestellt . eine Tendenz zur sDiagnoselastigkeit%bei Wissensvermittlung und . überprüfung gibt, während andere Aspekte der ärztlichen Leistung weniger stark betont werden.

Ein erster (noch nicht durch weitere Beurteiler überprüfter) Versuch einer inhaltlichen Zuordnung der bei den sAusreißern%vorliegenden Problematik ergibt folgendes Bild (Mehrfachzuordnungen waren möglich):

Wenn man die inhaltlichen Probleme in sbanale%o mehr auf Hausverstand basierende und allgemeine Inhalte einerseits und in fachspezifische Inhalte andererseits gliedert, dann sind die fachspezifischen Probleme (rund 55) etwa vier bis fünf mal so häufig wie die sbanalen%o Management- und Beratungsthemen sind in beiden Gruppen vertreten. Bei den oft unerwartet schwierigen, allgemeinen, mitunter sbanalen%o Fragen konnte man immer wieder den Kandidaten zugute halten, dass diese im Spital und in der Spitalsorganisation eben kaum auftreten und den Kandidaten daher bei der Prüfung oft nicht in den Sinn kommen%o

Die hohe Zahl der vom Referenten als fachspezifisch eingestuft Probleme signalisiert allerdings eine Ausbildungsproblematik: Es kommt immer wieder vor, dass von den in die differentialdiagnostischen Überlegungen einzubeziehenden Diagnosen sfast alle%o angegeben werden, aber eben eine wichtige von sehr vielen Kandidaten nicht genannt wird. Wenn man lediglich Wissen messen möchte, könnte man das akzeptieren. Wenn es um (handlungsorientierte) Kompetenz geht, müssen vom Arzt aber alle wichtigen Punkte in Betracht gezogen . und daher auch in der Prüfung verlangt . werden. Diese Überlegung gilt nicht nur für das obige Beispiel (Differential)Diagnose sondern sinngemäß für alle Aspekte des Leistungsspektrums, wie sie in den Leistungskriterien abgebildet sind.

me in der Prüfung, die man wohl darauf zurückführen kann, dass sie nur in der All gemeinmedizin erfahren und praxisgerecht erlernt werden können. Deshalb wurde seit Einführung der PAA auf die Notwendigkeit der verpflichtenden Lehrpraxis am Ende der Ausbildungszeit hingewiesen. Es scheint aber auch eine beträchtliche Zahl von Problemen zu geben, die im Rahmen der Spitalsausbildung beeinflussbar sein dürften (aktuelles Beispiel: Die Anwendung von Kautel en im Umgang mit antikoagulierten Patienten ist im Spital gleichermaßen wie extramural gefordert; schlechte Ergebnisse bei einschlägigen Fragen sind nicht mit dem Fehlen allgemeinmedizinischer Ausbildungselemente erklärbar). Diese sintramuralen Probleme sollten in der nächsten Zeit weiter analysiert und Wege zu deren Behebung gesucht werden.

Unter den inhaltlichen Problemen auffallend häufig vertreten sind die Themen sDepression und seelisch-körperliche Zusammenhänge, sie sind fast ebenso häufig wie die allgemeinen Probleme; sie stellen inhaltlich die größte Gruppe dar und kommen unter den 70 Ausreißern 10 mal vor).

Aufgrund dieser ersten Durchsicht der sAusreißer kann anhand der bis jetzt vorliegenden Ergebnisse vorläufig zusammengefasst werden:

- Die Zahl der methodisch bedingten Ausreißer ist gegenüber früheren Jahren deutlich geringer geworden.
- Die inhaltlichen Ausreißer zeigen im Wesentlichen die Verteilung der Blueprintkategorien. Die Interventionsbereiche sHaut, sPsyche und sAugen sind allerdings überrepräsentiert. Damit werden die früheren Analysen der unterschiedlichen Schwierigkeiten verschiedener Interventionsbereiche auf anderem Weg bestätigt.
- Es gibt eine Häufung fachspezifischer Probleme, die man im weitesten Sinn als Ausbildungsprobleme bezeichnen muss. Die inhaltliche Aufschlüsselung bedarf weiterer Analysen, wobei fraglich ist, ob sich diese Probleme auf wenige, gemeinsame Faktoren (sNenner) zurückführen lassen werden, oder ob sie viele, fachspezifische Einzelfragen mit fachspezifischem Ausbildungs- und Supervisionsbedarf darstellen.
- Es gibt also sowohl sSpitalsprobleme wie sallgemeinmedizinische Probleme, erstere möglicherweise sogar in höherem Ausmaß als bisher diskutiert wurde.
- Insgesamt scheint es jedenfalls deutlichen Ausbildungsbedarf zu geben, weil man vom lizenzierten Arzt ja verlangen muss, dass er ALLE wichtigen Aspekte berücksichtigen kann.
- Es sollte daher überlegt werden, auf welchem Weg die sCommunity der Auszubildenden über Probleme informiert werden soll, die man aus den Prüfungsergebnissen erkennen kann. Hier scheinen sowohl persönliche Informationen im Sinn von Workshops, Vorträgen, Beratungen etc. überlegenswert zu sein wie auch verschiedenartige Publikationen; solche könnten als Fachartikel konzipiert sein, in speziell geeigneten Medien abgehandelt werden, sollten aber auch immer wieder im Sinn einer sKolumne in den standeseigenen traditionellen Medien, vielleicht aber auch mittels neuer Medien transportiert werden.

4.3 Wiederverwendete Prüfungsfälle

Die Unterschiede in der Schwierigkeit bei mehrmaligem Einsatz sind in der Regel gering und liegen oft im Rahmen des \pm Messfehlers; häufig ist der Unterschied von erster zu zweiter Verwendung größer als von der zweiten zu einer dritten (siehe Tab. 16 . 19).

Die Unterschiede könnten einerseits durch Weitergabe der Problemstellungen unter den Kandidaten bedingt sein; in einzelnen Fällen mag auch die Überarbeitung von Fragen zu einem fallweise eingetretenen, größeren Unterschied beigetragen haben.

Auch für dieses Berichtsjahr gibt es zwar Gerüchte um die Weitergabe von Fällen durch Prüfungskandidaten, es gibt aber keinen Hinweis, dass ANTWORT-SCHLÜSSEL bekannt geworden wären. Die Unterschiede, die in den Tab. 16 . 19 dargestellt sind, können als der üblichen Entwicklung folgend eingestuft werden.

Tabelle 16: Wiederverwendete Fragen 200702

Titel	Nr. 0702	Nr. damals		p 0702	p damals	p _{diff}
Arteriitis temporalis	02	PAfA 05 / 2001	Nr. 24	0,69	0,66	+0,03
Fieber	03	PAfA 09 / 2003 PAfA 11 / 1999	Nr. 03 Nr. 10	0,85	0,65 0,67	+0,20 +0,18
Antikoagulantien- therapie	04	PAfA 05 / 2004 PAfA 01 / 2003 PAfA 05 / 2001	Nr. 04 Nr. 24 Nr. 09	0,71	0,70 0,63 0,59	+0,01 +0,08 +0,12
Parästhesien	06	PAfA 09 / 2003 PAfA 11 / 1999	Nr. 12 Nr. 04	0,70	0,62 0,61	+0,08 +0,09
Hypoglykämie	07	PAfA 09 / 2003	Nr. 23	0,76	0,59	+0,17
Präsuizidales Syndrom	10	PAfA 09 / 2003	Nr. 17	0,68	0,59	+0,09
Hüftschmerz	11	PAfA 09 / 2003 PAfA 11 / 1999	Nr. 24 Nr. 17	0,82	0,80 0,75	+0,02 +0,07
Hypercholesterinämie	12	PAfA 09 / 2002	Nr. 03	0,90	0,75	+0,15
Pneumonie	13	PAfA 09 / 2001	Nr. 11	0,74	0,71	+0,03
Lendenschmerz	14	PAfA 03 / 2000	Nr. 06	0,66	0,68	-0,02
Diarrhoe	15	PAfA 06 / 2005 PAfA 09 / 2003 PAfA 11 / 1999	Nr. 11 Nr. 15 Nr. 18	0,79	0,67 0,71	+0,12 +0,08
Varikophlebitis	16	PAfA 05 / 2003 PAfA 09 / 2001	Nr. 08 Nr. 08	0,59	0,55 0,49	+0,04 +0,10
DD Magen-Ca	20	PAfA 03 / 2000	Nr. 14	0,83	0,78	+0,05
Empfängnisverhütung	21	PAfA 05 / 2003 PAfA 09 / 2001	Nr. 05 Nr. 07	0,68	0,66 0,68	+0,02 +/-0,00
Thoraxverletzung	23	PAfA 05 / 2004 PAfA 09 / 2001	Nr. 06 Nr. 10	0,73	0,68 0,75	+0,05 -0,02
Verruca plana	25	PAfA 09 / 2003	Nr. 25	0,75	0,72	+0,03

p: Schwierigkeitsindex

igen 200705

Titel	Nr. 0705	Nr. damals		p 0705	p damals	p _{diff}
Husten	01	PAfA 05 / 2001	Nr. 10	0,99	0,73	+0,026
		PAfA 01 / 2004	Nr. 25		0,95	+0,04
Spontanhämatom	02	PAfA 09 / 2000	Nr. 07	0,63	0,59	+0,04
		PAfA 01 / 2004	Nr. 11		0,62	+0,01
Ca-Angst	11	PAfA 11 / 1999	Nr. 06	0,64	0,66	-0,02
		PAfA 09 / 2003	Nr. 10		0,60	+0,04
Verhaltensstörung im Wochenbett	14	PAfA 09 / 2000	Nr. 10	0,53	0,45	+0,08
art. Hypertonie	16	PAfA 01 / 2004	Nr. 18	0,77	0,69	+0,08
kolorektales Karzinom	17	PAfA 01 / 2004	Nr. 20	0,83	0,72	+0,11
Osteoporose	19	PAfA 01 / 2004	Nr. 07	0,74	0,64	+0,10
Nikotinabusus	20	PAfA 01 / 2004	Nr. 04	0,74	0,73	+0,01
Miktionsbeschwerden	21	PAfA 05 / 2002	Nr. 10	0,76	0,72	+0,04
retrosternaler Schmerz	23	PAfA 09 / 2000	Nr. 09	0,78	0,51	+0,27
		PAfA 01 / 2004	Nr. 06		0,64	+0,14
Wolkengehen	25	PAfA 01 / 2003	Nr. 08	0,69	0,37	+0,32
TIA	26	PAfA 05 / 2001	Nr. 21	0,64	0,67	-0,03

p: Schwierigkeitsindex

Tabelle 18: Wiederverwendete Fragen 200708

Titel	Nr. 0708	Nr. damals		p 0708	p damals	p _{diff}
Erbrechen	03	PAfA 05/2002	Nr. 09	0,64	0,64	+/-0,0
		PAfA 05/2004	Nr. 14		0,62	+0,02
Demenz	06	PAfA 09/2000	Nr. 14	0,46	0,39	+0,07
		PAfA 01/2003	Nr. 18		0,37	+0,09
Schlaganfall	07	PAfA 05/2004	Nr. 03	0,74	0,59	+0,15
Windeldermatitis	11	PAfA 01/2004	Nr. 02	0,45	0,42	+0,03
Diabetische Retinopathie	14	PAfA 05/2004	Nr. 19	0,79	0,68	+0,09
Epistaxis	15	PAfA 03/2000	Nr. 01	0,62	0,62	+/-0,0
urethraler Ausfluss	16	PAfA 11/1999	Nr. 07	0,66	0,60	+0,06
Lymphangitis	17	PAfA 06/2000	Nr. 06	0,70	0,70	+/-0,0
		PAfA 09/2003	Nr. 05		0,78	-0,08
kolorektales Karzinom	20	PAfA 05/2004	Nr. 23	0,70	0,70	+/-0,0
Bauchschmerz	21	PAfA 03/2000	Nr. 11	0,59	0,57	+0,02
		PAfA 05/2004	Nr. 21		0,49	+0,1
Coxarthrose rechts	22	PAfA 03/2000	Nr. 19	0,59	0,53	+0,06
		PAfA 05/2004	Nr. 16		0,57	+0,02
Hyperthyreose	23	PAfA 05/2004	Nr. 07	0,90	0,72	+0,18
uncharakt. Fieber	24	PAfA 03/2000	Nr. 09	0,73	0,58	+0,15
		PAfA 05/2004	Nr. 09		0,71	+0,02

p: Schwierigkeitsindex

Titel	Nr. 0711	Nr. damals		p 0711	p damals	p _{diff}
Skabies	02	PAfA 09/2002	Nr. 12	0,62	0,55	+0,07
schreiendes Kind	03	PAfA 09/2004	Nr. 11	0,77	0,67	+0,10
Pneumonie	04	PAfA 09/2004	Nr. 08	0,68	0,68	+/-0,0
Hypoparathyreoidismus	05	PAfA 09/2000	Nr. 19	0,63	0,53	+0,10
		PAfA 01/2004	Nr. 05		0,51	+0,12
Oberarmschmerzen	07	PAfA 06/2000	Nr. 18	0,58	0,50	+0,08
		PAfA 09/2004	Nr. 10		0,66	-0,08
Leberzirrhose	08	PAfA 09/2002	Nr. 14	0,78	0,58	+0,20
PAVK	09	PAfA 09/2001	Nr. 19	0,78	0,44	+0,34
		PAfA 05/2003	Nr. 04		0,51	+0,27
Anämie	10	PAfA 08/1999	Nr. 23	0,62	- - -	+0,04
		PAfA 09/2002	Nr. 15		0,58	+0,06
		PAfA 06/2005	Nr. 14		0,56	
Asthma bronchiale	11	PAfA 06/2000	Nr. 14	0,76	0,60	-0,16
		PAfA 09/2004	Nr. 13		0,62	+0,14
Otitis media	12	PAfA 09/2003	Nr. 02	0,65	0,65	+/-0,0
M. Parkinson	14	PAfA 09/2000	Nr. 16	0,52	0,48	+0,04
		PAfA 09/2004	Nr. 16		0,55	-0,03
Angehörige Ca-Patient	15	PAfA 06/2000	Nr. 13	0,63	0,43	+0,20
		PAfA 09/2004	Nr. 22		0,38	+0,25
Hydrocele testis	17	PAfA 05/2003	Nr. 15	0,70	0,68	-0,02
unerfüllter Kinderwunsch	18	PAfA 06/2000	Nr. 12	0,68	0,66	+0,02
		PAfA 09/2004	Nr. 03		0,52	+0,16
Arthrose	20	PAfA 09/2003	Nr. 16	0,77	0,57	+0,30
Thoraxschmerz	22	PAfA 01/2002	Nr. 4	0,81	0,57	+0,24

p: Schwierigkeitsindex

5 Kandidaten

5.1 Auswertung nach Bundesländern

Die Auswertung nach Bundesländern in Tabelle 20 ergibt neuerlich eine Veränderung der Rangreihenfolge. Diesmal liegt Vorarlberg auf dem ersten, Wien auf dem letzten Rang. Auf dem von vier Bundesländern geteilten 2. Rang liegen solche mit kleinen ebenso wie solche mit sehr großen Kandidatengruppen.

Ein diesmal ausführlicher als sonst dargestellter multipler Paarvergleich der Bundesländer zeigt die jeweils signifikanten Unterschiede auf (Tab. 21).

ind in den (elektronischen) Beilagen enthalten.

Tabelle 20: Mittlere Prüfungsleistung je Bundesland

	Mittel	SD	Rang*	n
Burgenland	69,7	7,1	2*	24
Kärnten	69,3	6,7	2	76
Niederösterreich	69,9	6,5	2	181
ÖÄK	67,7	---	7	1
Oberösterreich	69,5	7,0	2	205
Salzburg	68,7	6,2	5	61
Steiermark	68,3	7,4	5	130
Tirol	67,0	7,2	7	79
Vorarlberg	71,7	6,0	1	33
Wien	65,0	9,1	10	177

Quelle: BEMAW 2007

* bezogen auf Wert vor dem Komma

Tabelle 21: Prüfungsleistung: Multipler Paarvergleich der Bundesländer

Abhängige Variable: Gesamtpunkte in Prozent

	(I) B.-Länder	(J) B.-Länder	Mittlere Differenz (I-J)	Standardfehler	Signifikanz	95%-Konfidenzintervall	
						Obergrenze	Untergrenze
Bonferroni	B	K	,38629	1,71836	1,000	-5,1234	5,8959
		NÖ	-,25742	1,59426	1,000	-5,3692	4,8543
		OÖ	,19810	1,58330	1,000	-4,8785	5,2747
		S	,96626	1,76834	1,000	-4,7036	6,6362
		St	1,34138	1,63046	1,000	-3,8864	6,5692
		T	2,61657	1,71051	1,000	-2,8679	8,1011
		V	-2,03040	1,96880	1,000	-8,3431	4,2823
		Wien	4,63921	1,59637	,135	-,4793	9,7577
	K	B	-,38629	1,71836	1,000	-5,8959	5,1234
		NÖ	-,64371	1,00311	1,000	-3,8600	2,5726
		OÖ	-,18818	,98559	1,000	-3,3483	2,9720
		S	,57998	1,26158	1,000	-3,4651	4,6250
		St	,95509	1,05970	1,000	-2,4427	4,3528
		T	2,23028	1,17916	1,000	-1,5505	6,0111
		V	-2,41668	1,52995	1,000	-7,3222	2,4889
		Wien	4,25292(*)	1,00645	,001	1,0259	7,4800
	NÖ	B	,25742	1,59426	1,000	-4,8543	5,3692
		K	,64371	1,00311	1,000	-2,5726	3,8600
		OÖ	,45552	,74852	1,000	-1,9445	2,8555
		S	1,22368	1,08650	1,000	-2,2600	4,7074
		St	1,59880	,84372	1,000	-1,1064	4,3040
		T	2,87399	,98960	,136	-,2990	6,0470
		V	-1,77298	1,38911	1,000	-6,2270	2,6810
		Wien	4,89663(*)	,77579	,000	2,4092	7,3841

			-,19810	1,58330	1,000	-5,2747	4,8785
			,18818	,98559	1,000	-2,9720	3,3483
		NÖ	-,45552	,74852	1,000	-2,8555	1,9445
		S	,76816	1,07035	1,000	-2,6638	4,2001
		St	1,14327	,82281	1,000	-1,4949	3,7815
		T	2,41846	,97184	,468	-,6976	5,5345
		V	-2,22850	1,37652	1,000	-6,6421	2,1851
		Wien	4,44111(*)	,75300	,000	2,0267	6,8555
	S	B	-,96626	1,76834	1,000	-6,6362	4,7036
		K	-,57998	1,26158	1,000	-4,6250	3,4651
		NÖ	-1,22368	1,08650	1,000	-4,7074	2,2600
		OÖ	-,76816	1,07035	1,000	-4,2001	2,6638
		St	,37511	1,13896	1,000	-3,2768	4,0270
		T	1,65030	1,25087	1,000	-2,3604	5,6610
		V	-2,99666	1,58588	1,000	-8,0815	2,0882
		Wien	3,67295(*)	1,08959	,028	,1793	7,1666
	St	B	-1,34138	1,63046	1,000	-6,5692	3,8864
		K	-,95509	1,05970	1,000	-4,3528	2,4427
		NÖ	-1,59880	,84372	1,000	-4,3040	1,1064
		OÖ	-1,14327	,82281	1,000	-3,7815	1,4949
		S	-,37511	1,13896	1,000	-4,0270	3,2768
		T	1,27519	1,04692	1,000	-2,0816	4,6320
		V	-3,37177	1,43051	,670	-7,9585	1,2149
		Wien	3,29783(*)	,84769	,004	,5798	6,0158
	T	B	-2,61657	1,71051	1,000	-8,1011	2,8679
		K	-2,23028	1,17916	1,000	-6,0111	1,5505
		NÖ	-2,87399	,98960	,136	-6,0470	,2990
		OÖ	-2,41846	,97184	,468	-5,5345	,6976
		S	-1,65030	1,25087	1,000	-5,6610	2,3604
		St	-1,27519	1,04692	1,000	-4,6320	2,0816
		V	-4,64696	1,52113	,083	-9,5242	,2303
		Wien	2,02264	,99299	1,000	-1,1612	5,2065
	V	B	2,03040	1,96880	1,000	-4,2823	8,3431
		K	2,41668	1,52995	1,000	-2,4889	7,3222
		NÖ	1,77298	1,38911	1,000	-2,6810	6,2270
		OÖ	2,22850	1,37652	1,000	-2,1851	6,6421
		S	2,99666	1,58588	1,000	-2,0882	8,0815
		St	3,37177	1,43051	,670	-1,2149	7,9585
		T	4,64696	1,52113	,083	-,2303	9,5242
		Wien	6,66961(*)	1,39153	,000	2,2079	11,1313
	Wien	B	-4,63921	1,59637	,135	-9,7577	,4793
		K	-4,25292(*)	1,00645	,001	-7,4800	-1,0259
		NÖ	-4,89663(*)	,77579	,000	-7,3841	-2,4092
		OÖ	-4,44111(*)	,75300	,000	-6,8555	-2,0267
		S	-3,67295(*)	1,08959	,028	-7,1666	-,1793
		St	-3,29783(*)	,84769	,004	-6,0158	-,5798
		T	-2,02264	,99299	1,000	-5,2065	1,1612
		V	-6,66961(*)	1,39153	,000	-11,1313	-2,2079

Quelle: BEMAW 2007

* Die Differenz der Mittelwerte ist auf dem Niveau .05 signifikant.

5.2 Nach Geschlecht und zeitlichem Abstand zur Promotion

Der Anteil männlicher Kandidaten liegt im Berichtsjahr bei insgesamt 37,3 % - Tab. 22 (2006: 38,3%).

Neuerlich sind die Prüfungsleistungen der Kandidatinnen deutlich (und signifikant) besser als die der männlichen Prüflinge (Tab. 23 und 24).

Tabelle 22: KandidatInnen: Geschlechterverteilung

		Häufigkeit n	Prozent
2007	männlich	361	37,3
	weiblich	606	62,7
	Gesamt	967	

Quelle: BEMAW 2007

Tabelle 23: Prüfungsleistung je Geschlecht

Geschlecht	p	n	SD
männlich	66,9	361	8,5
weiblich	69,3	606	6,7
Insgesamt	68,4	967	7,5

Quelle: BEMAW 2007

p: Schwierigkeitsindex

Tabelle 24: Prüfungsleistung je Geschlecht, Gruppenvergleich

	Quadrat summe	df	Mittel der Quadrate	F	Signifikanz
Zwischen den Gruppen (Kombiniert)	1302,568	1	1302,568	23,489	,000
Innerhalb der Gruppen	53514,448	965	55,455		
Insgesamt	54817,017	966			

Quelle: BEMAW 2007

Wie bereits wiederholt festgestellt wurde, sind auch diesmal die Ergebnisse von Kandidaten deutlich schlechter, wenn sie diese Prüfung lange nach der Graduierung abgelegt haben. (Die bisher zur Illustration dieses Befundes verwendeten Daten und Abbildungen werden nicht wiedergegeben, weil sie keine wesentlichen neuen Informationen enthalten.)

6 Spezielle Aspekte:

6.1 Prüfungswiederholung

Im Berichtszeitraum wurden insgesamt 42 Antritte zu Wiederholungen registriert. Das sind bezogen auf die Teilnehmerzahl etwa so viele Antritte zu Wiederholungen wie 2006. Von diesen 42 Antritten waren 25 erfolgreich. Die Mehrzahl der Antritte waren Zweitansätze (22 von 32 positiv). Weitere Einzelheiten siehe Tab. 25.

Unter den Personen, die zu einer Wiederholung angetreten sind, ist wiederum . soweit aus den Angaben zur Person zu schließen - ein guter Teil aus Österreich oder einem deutschsprachigen Land.

Tabelle 25: Wiederholungsansätze / davon positiv

Antritt	2007 / 02	2007 / 05	2007 / 08	2007 / 11	Gesamt
2	8 / 6	11 / 8	9 / 5	4 / 3	32 / 22
3		1 / 0	1 / 1	5 / 0	7 / 1
4			2 / 1		2 / 1
5					
6	1 / 1				1 / 1
Gesamt	9 / 7	12 / 8	12 / 7	9 / 3	42 / 25

Quelle: BEMAW 2007

Aus diesen Ergebnissen ergeben sich nach Meinung der Berichtersteller zwei Schlussfolgerungen:

Personen, die mehr als drei Mal zu dieser Prüfung antreten müssen (und dann unter Umständen immer noch nicht erfolgreich sind), dürften auch in den Augen der Öffentlichkeit letzten Endes nicht geeignet sein, als eigenberechtigte Allgemeinärzte tätig zu sein. Ab dem dritten Antritt ist die Erfolgsquote sehr niedrig. Es sollte jedenfalls mit Nachdruck überlegt werden, wie die Zahl der Antritte limitiert werden kann. (Dabei muss natürlich den Aspekten des Rechtsschutzes ebenso Rechnung getragen werden, wie den Erfordernissen der Prüfungsmethodik; es erscheint aber dringend geboten zu sein, eine Form zu finden, die auch in der Öffentlichkeit vertreten werden kann.)

gestellte hohe Zahl an Personen mit (aller
wünschenswerten nach) deutscher Muttersprache unter den Wiederholungs-
kandidaten könnte wiederum als Indikator gewertet werden, dass Sprachlichkeit an
sich nicht den manchmal ins Treffen geführten hohen Anteil an den
Versagensursachen hat. (Das wäre auch insoferne verwunderlich, als die
KandidatInnen ja in praktisch allen Fällen jahrelang in Österreich mit PatientInnen
gearbeitet haben und daher schon aus diesem Grund in der Lage sein müssten, sich
entsprechend zu verständigen.)

6.2 Qualität der Korrekturen

Im Berichtszeitraum 2007 lag die Zahl der Datenerfassungsfehler bei den vier
Terminen zwischen 2,5 und 4,2 Fehlern pro Fall.

Diese Kennzahlen markieren einen aus der Sicht der Auswertung durchaus
zufriedenstellenden Qualitätsstandard. Auf Grund dieser Qualitätsbemühungen ist
die Zahl der technischen Auswertungsprobleme nach wie vor als niedrig zu
bezeichnen.

6.3 Ausblick, Empfehlungen

Verteilung der Leistungskriterien: Die Variabilität von Termin zu Termin bei einem
Kriterium, aber auch die vorgefundene Gewichtung der Kriterien zueinander
erfordern eine Überprüfung der Betrachtungsweise (Ebene der Unterfrage oder
Ebene der Items einer Unterfrage) und des Ergebnisses der Zusammenstellung der
jeweiligen Prüfung bei späteren Terminen.

Es sei auch nochmals auf das Problem der *Key features* hingewiesen: Wenn man
die Prüfungsfälle nach der Erfassung *essentieller* für die ärztliche Kompetenz
kritischer Elemente bewertet, findet man keineswegs in jedem Fall ein solches,
besonders *kritisches* Element (siehe vorangegangene Berichte). Es sollte überlegt
werden, ob nicht eine Ausrichtung der Fälle forciert werden sollte, bei der jeder Fall
mindestens ein derartiges *Key feature* enthält, das als essentielle Antwort
besonders gewichtig erscheint. Zumindest sollte geprüft werden, ob eine . und
konkret welche - wichtige Problemstellung, die sich in mehr als einem Item
niederschlagen kann, hier geprüft wird. Dabei sollte ein systematisches Abarbeiten
von Anamnese . Diagnose . Therapie vermieden werden. Die jeweilige
Problemstellung sollte in Kürzestform explizit formuliert werden.

Weiters wird an den Vorschlag aus früheren Berichten erinnert, die Verteilungs-
kriterien um bestimmte allgemeinmedizinisch relevante Themen zu ergänzen (z. B.
Palliation, Ethik und Rechtsmedizin u.a.) und Fälle gezielt zu bestimmten
Aufgabenstellungen zu rekrutieren.

Analyse besonders schwieriger Themen hat ergeben, dass es neben den bekannten allgemeinmedizinischen Problembereichen viele, jedenfalls nicht spitalsfremde Ausbildungsdefizite geben dürfte. Die Bemühungen um Charakterisierung und Interpretation der aus den quantitativen und vor allem auch qualitativen Prüfungsergebnissen ableitbaren Hinweise auf Ausbildungsdefizite sollten fortgesetzt werden. Insbesondere sollte überlegt werden, in welcher Weise die Betroffenen davon informiert und wie Verbesserungsmaßnahmen angestoßen werden können.

Die schon in den früheren Berichten jeweils angeführten signifikant besseren Prüfungsergebnisse der Ärztinnen sollten ebenso eingehend diskutiert werden wie die unterschiedlichen Ergebnisse der einzelnen Bundesländer.

Das Problem der unlimitierten, oftmaligen Prüfungswiederholung besteht nach wie vor; die vereinzelt vorkommenden hohen Wiederholungsraten und die ab dem dritten Antritt zu registrierenden (Miss)Erfolgsquoten erfordern eine möglichst baldige Lösung.