

AMBOSS-Nutzung ist mit höherer Erfolgsquote und besseren Punktzahlen im schriftlichen ersten Abschnitts der Ärztlichen Prüfung (M1) assoziiert – eine retrospektive Analyse der 4 Kohorten Frühjahr 2023 bis Herbst 2024

AMBOSS Research Team, Email: research@amboss.com

Abstract

Der schriftliche Teil des ersten Abschnitts der Ärztlichen Prüfung (M1) stellt eine entscheidende Hürde im Studium der Humanmedizin in Deutschland dar, und die Wahl geeigneter Lernmittel kann maßgeblich zum Prüfungserfolg beitragen. AMBOSS ist über das gesamte Studium der Humanmedizin hinweg die am weitesten verbreitete digitale Lernressource. Im zweiten Abschnitt der Ärztlichen Prüfung (M2) nutzen seit Jahren über 90 % der Studierenden die Plattform zur Prüfungsvorbereitung.

Die vorliegende Studie untersuchte in insgesamt vier unabhängigen Examenskohorten den Zusammenhang zwischen AMBOSS-Nutzung und Prüfungsergebnissen im M1. Grundlage hierfür waren die Antworten aller Studierenden, die im Examensservice angaben, sich mit AMBOSS auf das M1 vorbereitet zu haben ($n = 8.803$). Ihre Ergebnisse wurden mit den Gesamtergebnissen der jeweiligen Kohorten verglichen ($n = 19.518$). Zur statistischen Analyse wurden vier Einstichproben- t -Tests zum Vergleich der durchschnittlich erreichten Punktzahlen sowie vier Chi-Quadrat-Anpassungstests zum Vergleich der Erfolgsquoten herangezogen. Die Ergebnisse zeigten in allen untersuchten Kohorten höhere Punktzahlen und Erfolgsquoten unter den AMBOSS-Nutzer:innen. In drei der vier Kohorten (Frühjahr 2023, Herbst 2023, Herbst 2024) waren diese Unterschiede statistisch signifikant ($p \leq 0,001$). Die durchschnittliche Punktzahl der AMBOSS-Nutzer:innen lag dabei über der jeweiligen Kohortenleistung, und auch die Erfolgsquote war signifikant höher.

Die vorliegenden Ergebnisse deuten darauf hin, dass die Nutzung von AMBOSS mit einer höheren Punktzahl sowie einer höheren Erfolgsquote im M1 assoziiert ist.

Hintergrund

Die Vorbereitung auf den schriftlichen Teil des ersten Abschnitts der Ärztlichen Prüfung (M1) ist für Studierende der Humanmedizin eine entscheidende Hürde. Erst nach erfolgreicher Absolvierung ist eine Fortführung des Studiums im klinischen Abschnitt möglich. Die Wahl geeigneter Lernmittel stellt Studierende oftmals vor Herausforderungen und kann sich maßgeblich auf den Prüfungserfolg auswirken.

Bislang liegen keine aktuellen Studien zum tatsächlichen Effekt von AMBOSS in der Vorbereitung auf das M1 vor. Im zweiten Abschnitt der Ärztlichen Prüfung (M2) nutzen seit über

10 Jahren mehr als 90 % der Studierenden AMBOSS als primäre Lernressource^[1]. Berechnet wird dieser Wert durch den Abgleich der vom Institut für medizinische und pharmazeutische Prüfungsfragen (IMPP) veröffentlichten Gesamtteilnehmerzahlen mit internen Nutzungsdaten. Zur Identifikation der AMBOSS-Nutzer:innen wurde ein *Detection Tag* definiert, der auf der Erreichung spezifischer Nutzungsschwellen basiert (Anzahl gekreuzter Fragen: M1: ≥ 2.500 vorklinische Fragen innerhalb von 90 Tagen; M2: ≥ 4.000 klinische Fragen innerhalb eines festgelegten Zeitraums vor dem Examen). Studierende, die diese Kriterien erfüllen, werden der jeweiligen Nutzerkohorte zugeordnet. Der daraus resultierende Anteil wird in Relation zur Gesamtzahl der Prüfungsteilnehmenden gesetzt und als Prozentwert ausgewiesen.

Durch den hohen Anteil der AMBOSS-Nutzer:innen ist ein Vergleich mit einem Gesamtkollektiv kaum möglich. Im US-amerikanischen Raum konnte hingegen ein positiver Effekt der AMBOSS-Nutzung auch für ein klinisches Examen nachgewiesen werden^[2].

Im M1 ist der Anteil der AMBOSS-Nutzer:innen in den vergangenen Jahren kontinuierlich gestiegen und liegt inzwischen bei ca. 60 %^[1], was weiterhin einen direkten Vergleich mit der Gesamtkohorte ermöglicht. Dadurch kann die Assoziation der AMBOSS-Nutzung mit dem Prüfungserfolg im ersten Abschnitt der Ärztlichen Prüfung quantifiziert werden.

Bereits im Frühjahr 2017 wurde eine retrospektive Kohortenanalyse durchgeführt, um den Zusammenhang zwischen Vorbereitungsmethode und Prüfungsleistung im M1 zu evaluieren. Die Kohorte der Studierenden, die sich hauptsächlich mit AMBOSS auf das Examen vorbereitet hatten ($n = 355$), erzielte eine höhere durchschnittliche Punktzahl. Der ermittelte Leistungsvorsprung betrug dabei 11,8 Punkte im Vergleich zur Gesamtkohorte^[3]. Die Datenbasis für diesen Vergleich stammte aus dem Examensservice, der die Auswertung der Antwortmuster der AMBOSS-Nutzer:innen im Verhältnis zu den Gesamtergebnissen aller M1-Teilnehmenden ermöglichte.

Es ist zu berücksichtigen, dass das Frühjahr 2017 das erste Examensjahr darstellte, in dem AMBOSS für die M1-Vorbereitung verfügbar war, wodurch der Anteil der AMBOSS-Nutzer:innen an der Gesamtkohorte damals noch verhältnismäßig niedrig war (zwischen 10 und 15%)^[1]. Aktuelle Daten des Examensservices belegen jedoch eine veränderte Verbreitung: Inzwischen bereiten sich mehr als 60 % aller Studierenden mithilfe von AMBOSS auf ihre M1-Prüfung vor.

Im Folgenden soll untersucht werden, ob der zuvor beobachtete positive Unterschied fortbesteht und über mehrere Examen hinweg nachweisbar ist. Während in der früheren Analyse lediglich die Differenz zwischen der durchschnittlichen Punktzahl der AMBOSS-Nutzer:innen und derjenigen der Teilnehmergesamtheit genannt wurde, wird in der vorliegenden Studie zusätzlich die statistische Signifikanz dieses Unterschieds geprüft.

Methode & Design

In der vorliegenden Studie werden Daten genutzt, die im Rahmen des Examensservice erhoben wurden. Dieser ermöglicht es Studierenden, nach Abschluss eines oder beider Prüfungstage im M1 vorläufige Informationen zum Abschneiden in der Prüfung zu erhalten. Dabei wird zudem das jeweils primär eingesetzte Lernmittel erfasst.

Für unsere Analyse stehen die Antworten aller derer zur Verfügung, die angegeben haben, hauptsächlich mit AMBOSS gelernt zu haben. Für jede einzelne Person werden die einzelnen

Antworten der jeweils 160 Fragen an Tag 1 und 2 des M1 übermittelt. Im Anschluss können die von den Studierenden gegebenen Antworten mit den vom IMPP veröffentlichten offiziellen Lösungen abgeglichen werden. Hierbei werden diejenigen Fragen ausgeschlossen, die in der Auswertung keine Berücksichtigung finden. Auf dieser Grundlage lässt sich für jede teilnehmende Person eine individuelle Punktzahl berechnen.

Die so erfassten Daten werden daraufhin mit den vom IMPP veröffentlichten Gesamtergebnissen des jeweiligen M1 verglichen^[4,5,6,7]. Zur Analyse wird für jedes Examen ein einseitiger Einstichproben-*t*-Test angewandt, um eine Stichprobe (hier: Ergebnisse der AMBOSS-Nutzer:innen) mit einem bekannten Bevölkerungsdurchschnitt (hier: die Gesamtergebnisse des M1) zu vergleichen.

Ebenso wird die Anzahl derer, die die Prüfung erfolgreich bestanden haben, mit der Anzahl derer verglichen, die eine Punktzahl unterhalb der Bestehensgrenze erreicht haben, um die Erfolgsquote zu vergleichen. Hierzu wird ein Chi-Quadrat-Anpassungstest angewandt, mit dem die tatsächliche Frequenz der Erfolge und Misserfolge (= Nichtbestehen des Exams) mit den erwarteten Ergebnissen verglichen werden können.

Ergebnisse & Analyse

Im Rahmen des Examensservice wurden die anonymisierten Daten der folgenden M1-Kohorten übermittelt: Frühjahr 2023 (n = 1.363), Herbst 2023 (n = 3.568), Frühjahr 2024 (n = 1.107), Herbst 2024 (n = 2.890). Von der Analyse ausgeschlossen wurden die Datenpunkte derer, die nur für einen der beiden Prüfungstage ihre Antworten gegeben hatten (Frühjahr 2023, n = 19; Herbst 2023, n = 44; Frühjahr 2024, n = 17; Herbst 2024, n = 45).

Im Folgenden werden die deskriptiven Ergebnisse der Gesamtkohorte, veröffentlicht durch das IMPP, sowie die Resultate der AMBOSS-Nutzer:innen dargestellt, deren Daten im Rahmen des Examensservice übermittelt wurden. Es werden die durchschnittliche Punktzahl, die Standardabweichung sowie der Anteil von Erfolgen und Misserfolgen aufgeführt. Für die Gruppe der AMBOSS-Nutzer:innen wird außerdem aufgeführt, welchen Anteil der Gesamtkohorte unsere Stichprobe repräsentiert.

Tabelle 1: Deskriptive Statistik für die Gesamtkohorten von F23 bis H24

	<i>F23</i>	<i>H23</i>	<i>F24</i>	<i>H24</i>
N	2.927	7.050	2.914	7.094
Durchschnittliche Punktzahl (SD)	215,87 (37,64)	230,10 (33,80)	214,28 (33,91)	229,97 (33,43)
% Erfolge (N)	81,96 (2.399)	90,20 (6.359)	83,49 (2.433)	90,48 (6.419)
% Misserfolge (N)	18,04 (528)	9,80 (691)	16,51 (481)	9,52 (675)

Tabelle 2: Deskriptive Statistik für die AMBOSS-Nutzer:innen von F23 bis H24

	<i>F23</i>	<i>H23</i>	<i>F24</i>	<i>H24</i>
N	1.344	3.524	1.090	2.845
Anteil an der Gesamtkohorte (%)	45,92	49,99	37,41	40,10
Durchschnittliche Punktzahl (SD)	221,13 (34,87)	232,10 (31,41)	215,91 (32,77)	231,76 (31,09)
% Erfolge (N)	85,79 (1.153)	92,28 (3.252)	84,86 (925)	92,55 (2.633)
% Misserfolge (N)	14,21 (191)	7,72 (272)	15,14 (165)	7,45 (212)

In allen vier untersuchten Prüfungen war die Nutzung von AMBOSS zur M1-Vorbereitung im Vergleich zur Gesamtkohorte sowohl mit höheren durchschnittlichen Punktzahlen als auch mit höheren Erfolgsquoten assoziiert. Die Ergebnisse unterstützen somit die geplante Durchführung eines einseitigen *t*-Tests.

Im Folgenden werden die Ergebnisse der Einstichproben-*t*-Tests präsentiert. Aufgeführt werden die Differenz zwischen der durchschnittlichen Punktzahl der Gesamtkohorte und der AMBOSS-Nutzer:innen sowie der *t*- und *p*-Wert.

Tabelle 3: Vergleich der Punktzahlen von AMBOSS-Nutzer:innen und Gesamtteilnehmer:innen mithilfe von Einstichproben-*t*-Tests

	<i>Gesamtmittel</i>	<i>AMBOSS-Mittel</i>	<i>df (AMBOSS)</i>	<i>t-Wert</i>	<i>p-Wert</i>
F23	215,87	221,13	1.343	5,5301	<i>p</i> < 0,001
H23	230,10	232,10	3.523	3,7799	<i>p</i> < 0,001
F24	214,28	215,91	1.089	1,6422	<i>p</i> = 0,050
H23	229,97	231,76	2.844	3,071	<i>p</i> = 0,001

In allen vier der untersuchten Examen schnitten diejenigen, die angegeben hatten, hauptsächlich mit AMBOSS zu lernen, besser ab als der Durchschnitt der Gesamtteilnehmer:innen. In drei der vier Examen (F23, H23, H24) unterschritt der *p*-Wert den Signifikanzgrenzwert von 0,05, sodass die Unterschiede als statistisch signifikant einzustufen sind (*p* ≤ 0,001).

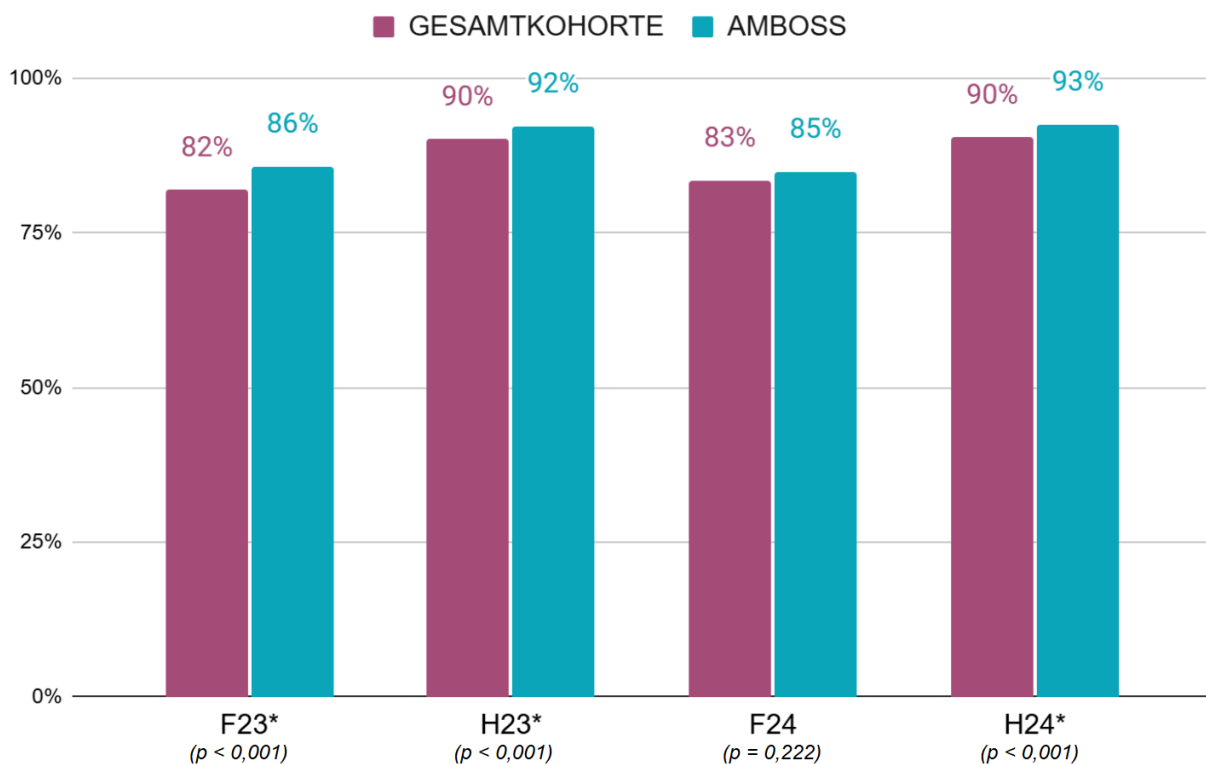
Es wurden außerdem die Erfolge und Misserfolge der AMBOSS-Kohorte mit denen der Gesamtteilnehmer:innen verglichen. Hierzu wurden Chi-Quadrat-Anpassungstests

durchgeführt, deren zusammengefassten Ergebnisse in Tabelle 4 dargestellt sind (vgl. Anhang, Tabellen A1–4).

Tabelle 4: Vergleich der Erfolge/Misserfolge von AMBOSS-Nutzer:innen und der Gesamtkohorten mithilfe von Chi-Quadrat-Anpassungstests

	<i>Chi-Quadrat-Wert</i>	<i>p-Wert</i>
F23	13,32	$p < 0,001$
H23	17,27	$p < 0,001$
F24	1,49	$p = 0,222$
H24	14,13	$p < 0,001$

Abbildung 1: Vergleich der Erfolgsquoten von AMBOSS-Nutzer:innen und der Gesamtkohorten (vgl. Anhang, Tabellen A1–4)



In allen vier Examen gab es anteilig mehr Erfolge unter AMBOSS-Nutzer:innen, und in drei von vier Examen (F23, H23, H24) war der Unterschied statistisch signifikant ($p \leq 0,001$). Es war wahrscheinlicher, dass AMBOSS-Nutzer:innen im Examen erfolgreich waren, als aufgrund der Gesamtergebnisse erwartet werden konnte.

Diskussion

In allen untersuchten Examen erzielten die Studierenden, die sich mit AMBOSS vorbereitet hatten, im Durchschnitt eine höhere Punktzahl als die Grundgesamtheit. In drei von vier Examen war dieser Unterschied statistisch signifikant. Zudem konnte unter AMBOSS-Nutzer:innen eine niedrigere Misserfolgsquote beobachtet werden.

Die vorliegenden Ergebnisse bestätigen den positiven Unterschied, der bereits in der 2017 durchgeführten Analyse nachgewiesen wurde. Zusätzlich zeigte sich erstmals, dass dieser Unterschied über mehrere Examen hinweg konsistent nachweisbar ist.

Wie in Tabelle 2 ersichtlich, lagen durch die Verwendung der Daten von MEDI-LEARN die Informationen eines großen Anteils der Gesamtkohorte zur Analyse vor. Die absolute Fallzahl und die hohe Abdeckungsquote der relevanten Referenzkohorte steigern das Vertrauen in die Robustheit der festgestellten Leistungsunterschiede.

Die Nutzung der Daten eines Drittanbieters bedeutet außerdem, dass eine Verzerrung der Daten zugunsten von AMBOSS weniger wahrscheinlich ist, da die Daten nicht von AMBOSS selbst erhoben wurden.

Der Unterschied zwischen den Gruppen fiel in den vorliegenden Examen weniger ausgeprägt aus als in der Analyse von 2017. Möglicherweise ist dieser frühere, stärkere Effekt auf die damalige geringe Verbreitung von AMBOSS zurückzuführen. Studierende, die sich im Jahr 2017 für die Nutzung von AMBOSS entschieden, waren „Early Adopters“ dieses neuen Lernmittels. Deshalb ist die Möglichkeit einer positiven Selbstselektion in Betracht zu ziehen, bei der diese Studierenden bereits intrinsisch höhere Leistungsstärken aufwiesen^[8,9].

Angesichts der Etablierung von AMBOSS als meistgenutztes Lernmittel auch im M1 – mit einer aktuellen Nutzungsprävalenz von über 60 % – ist es nicht mehr plausibel, dass die Nutzergruppe weiterhin eine positive Selbstselektion bzgl. der intrinsischen Leistungsstärke darstellt. Die Nutzerbasis ist heute deutlich heterogener und repräsentiert einen breiteren Querschnitt der Studierendenpopulation. Die Persistenz eines statistisch signifikanten Leistungsvorsprungs bei AMBOSS-Nutzer:innen in den nachfolgenden Examen, ungeachtet der signifikanten Steigerung und Heterogenisierung der Nutzerbasis, legt somit nahe, dass die beobachtete Assoziation nicht mehr primär durch diese mögliche Selbstselektion erklärt werden kann. Vielmehr deutet dieser Befund auf einen konsistenten positiven Zusammenhang zwischen Lernmittel und Examensleistung hin, der über die potenzielle Anfangsselektion hinausgeht.

Bei der Interpretation der Ergebnisse sind mehrere wesentliche Limitationen zu berücksichtigen. Die berichteten Punktzahlen basieren auf Selbstauskünften der Studierenden zu den im Examen gegebenen Antworten und nicht auf den offiziellen Ergebnissen. Dies setzt eine hinreichend verlässliche Dokumentation der eigenen Antworten voraus. Es ist jedoch davon auszugehen, dass ein hohes Interesse der Studierenden an einer wahrheitsgetreuen Angabe besteht, da nur so eine realistische Einschätzung der eigenen Leistung möglich ist. Vor diesem Hintergrund wurden die erhobenen Daten als ausreichend valide eingestuft.

Eine wesentliche methodische Einschränkung der Analyse ergibt sich aus der verwendeten Datengrundlage, da ausschließlich Leistungsdaten von AMBOSS-Nutzerkohorten übermittelt

wurden. Die Durchführung eines direkten Vergleichs zwischen strikt unabhängigen Gruppen (AMBOSS-Nutzer:innen vs. reine Nicht-Nutzer:innen) war folglich nicht möglich. Stattdessen wurde die AMBOSS-Nutzer:innen-Performance mit der Performance der Gesamtkohorte der jeweiligen M1-Prüfung verglichen. Die Gesamtkohorte schließt die AMBOSS-Nutzer:innen notwendigerweise mit ein. Der Anstieg der Nutzungsprävalenz von AMBOSS (von einem geringen Anteil in 2017 auf aktuell über 60 %) hat dabei zur Folge, dass die Leistung der im Durchschnitt besser abschneidenden AMBOSS-Nutzer:innen die Referenzpunktzahl der Gesamtkohorte im Zeitverlauf angehoben hat. Dies führt zu einer methodisch bedingten Verringerung des beobachteten Leistungsunterschieds im Vergleich zur Analyse von 2017.

Der Vergleich zweier vollständig unabhängiger Gruppen würde voraussichtlich andere Ergebnisse hervorbringen. Im Laufe der Analyse wurde in Erwägung gezogen, die Ergebnisse der „Nicht-AMBOSS“-Gruppe rechnerisch zu ermitteln und darauf basierend zwei unabhängige Gruppen zum Vergleich zu erstellen. Allerdings ist die Zusammensetzung dieser errechneten Gruppe für eine verlässliche Analyse zu uneindeutig. Wie in Tabelle 2 dargestellt, wurde ein sehr großer Anteil der Gesamtkohorte als AMBOSS-Nutzer:innen identifiziert und in die Analyse einbezogen. Interne Nutzungsdaten von AMBOSS weisen jedoch darauf hin, dass die tatsächliche Zahl der Studierenden, die AMBOSS zur Vorbereitung auf die jeweiligen Examen verwendet haben, noch höher lag. Ein Vergleich mit der rechnerisch ermittelten Restgruppe hätte mit hoher Wahrscheinlichkeit dazu geführt, dass in beiden vermeintlich unabhängigen Gruppen AMBOSS-Nutzer:innen enthalten gewesen wären. Daher wurde der Vergleich mit der Gesamtkohorte gewählt, da hierfür die vollständige Unabhängigkeit der Gruppen keine Voraussetzung darstellt.

Darüber hinaus war es aufgrund der limitierten zur Verfügung stehenden Informationen nicht möglich, weitere Lernmittel in die Analyse einzubeziehen. Ein häufig genutztes ergänzendes Lernmittel ist Anki, das von Studierenden vielfach zusätzlich zur primären Lernplattform eingesetzt wird. Der Einfluss solcher zusätzlicher Lernmittel ist daher nicht zu vernachlässigen und sollte, sofern möglich, in zukünftigen Analysen berücksichtigt werden.

Fazit

Das Ziel dieser Studie war es, den Zusammenhang zwischen der Nutzung von AMBOSS für die Vorbereitung auf den ersten Abschnitt der Ärztlichen Prüfung und dem tatsächlichen Prüfungserfolg zu untersuchen. Die Ergebnisse können Studierende bei der Wahl ihrer Lernmittel unterstützen und sie somit entlasten. In der vorliegenden Auswertung war festzustellen, dass die Nutzung von AMBOSS als Hauptlernmittel mit einer höheren Punktzahl sowie einer höheren Erfolgsquote im M1 assoziiert war. Für die Praxis bedeutet dies, dass Studierende, die sich mit AMBOSS vorbereiten, in der Wahl ihres Lernmittels bestärkt werden. Zugleich unterstreichen die Ergebnisse die Relevanz, künftige Analysen stärker vergleichend auszurichten – etwa im direkten Kontrast zu anderen Lernressourcen oder unter Berücksichtigung der parallelen Nutzung mehrerer Lernmittel.

Quellen

1. Interne Daten von AMBOSS SE Zeitraum 2018–2023. Verfügbar auf Anfrage beim Hauptautor.
2. Lee, E. Cherkaoui, O. Tolman, Z. (2022). *The Impact of Digital Medical Resources on USMLE Step 2 CK Scores—A Retrospective Study of 1,985 US Medical Students*. [White Paper]
<https://drive.google.com/file/d/11qCBTEKGS7C1WghG33IFATWbIJDMjbUz/view>
3. AMBOSS-Nutzer erreichten höhere Punktzahl im Physikum | AMBOSS. (2017). Amboss.
<https://web.archive.org/web/20210621184344/https://www.amboss.com/de/physikum/hoehere-punktzahl>
4. Herbst 2024 Ergebnisse Medizin M1 | IMPP. (2024).
<https://www.impp.de/pruefungen/medizin/archiv-medizin.html>
5. Frühjahr 2024 Ergebnisse Medizin M1 | IMPP. (2024).
<https://www.impp.de/pruefungen/medizin/archiv-medizin.html>
6. Herbst 2023 Ergebnisse Medizin M1 | IMPP. (2023).
<https://www.impp.de/pruefungen/medizin/archiv-medizin.html>
7. Frühjahr 2023 Ergebnisse Medizin M1 | IMPP. (2023).
<https://www.impp.de/pruefungen/medizin/archiv-medizin.html>
8. Abercrombie, S., Bang, H., & Vaughan, A. (2022). Motivational and disciplinary differences in academic risk taking in higher education. *Educational Psychology*, 42(7), 895–912.
9. Leinonen, J., & Denny, P. (2024, December). "Sometimes You Just Gotta Risk It for the Biscuit": A Portrait of Student Risk-Taking. In *Proceedings of the 2024 on ACM Virtual Global Computing Education Conference V. 1* (pp. 102-108).

Anhang:

Fig. 1.1: Ergebnisinformation Frühjahr 2023 (inkl. Misserfolge, Bestehensgrenze)

Frühjahr 2023

Ergebnisinformation

gewertete Aufgaben: 318
 Bestehensgrenze: 183
 Stand: 04.04.2023

Hochschule	Gesamteilnehmer					Referenzteilnehmer						
	Teilnehmer	Misserfolge abs.	Misserfolge in %	Mittelwert abs.	Mittelwert in %	Std. abw.	Teilnehmer	Misserfolge abs.	Misserfolge in %	Mittelwert abs.	Mittelwert in %	Std. abw.
Berlin ¹	1						0					
Berlin MS	61	11	18,0	209,5	65,9	35,2	55	10	18,2	211,9	66,6	35,0
Bochum	78	30	38,5	194,7	61,2	34,7	0					
Bonn ¹	106	26	24,5	216,4	68,1	40,1	1					
Dresden	48	9	18,8	217,5	68,4	38,3	2	0	0,0	240,5	75,6	56,5
Erfurt HMU	109	46	42,2	188,8	59,4	41,5	67	20	29,9	201,8	63,5	41,4
Erlangen	224	18	8,0	231,9	72,9	32,8	101	2	2,0	244,4	76,9	28,4
Essen	49	12	24,5	203,2	63,9	29,0	0					
Frankfurt	94	19	20,2	203,4	64,0	29,8	2	0	0,0	257,0	80,8	12,0
Freiburg	115	16	13,9	224,3	70,5	32,0	3	0	0,0	272,3	85,6	6,1
Gießen	180	27	15,0	213,8	67,2	34,5	133	16	12,0	217,8	68,5	33,1
Göttingen	167	13	7,8	230,6	72,5	33,5	125	3	2,4	237,1	74,6	27,7
Greifswald	36	5	13,9	202,1	63,5	24,9	0					
Halle	50	12	24,0	197,1	62,0	25,9	0					
Hamburg MS	118	37	31,4	199,6	62,8	38,4	52	6	11,5	225,2	70,8	31,3
Heidelberg	35	9	25,7	212,9	67,0	45,2	3	0	0,0	230,0	72,3	11,0
Homburg	59	21	35,6	192,1	60,4	32,6	3	0	0,0	226,3	71,2	28,2
Jena	41	14	34,1	198,8	62,5	40,8	0					
Kiel	28	8	28,6	195,2	61,4	35,5	0					
Leipzig	52	13	25,0	200,9	63,2	28,5	10	2	20,0	203,1	63,9	23,4
Lübeck	27	5	18,5	213,3	67,1	32,5	0					
Magdeburg	25	4	16,0	212,9	66,9	35,8	0					
Mainz	212	15	7,1	231,5	72,8	33,1	114	2	1,8	244,5	76,9	25,8
Mannheim	36	9	25,0	201,6	63,4	39,4	5	0	0,0	260,0	81,8	20,7
Marburg	140	47	33,6	196,4	61,8	31,6	0					
München LMU	220	42	19,1	212,2	66,7	35,6	5	0	0,0	237,2	74,6	22,3
Münster	157	16	10,2	229,3	72,1	37,3	108	3	2,8	239,0	75,2	32,8
Regensburg	47	7	14,9	217,7	68,5	34,5	3	0	0,0	241,0	75,8	23,3
Rostock	42	3	7,1	221,6	69,7	25,3	0					
Tübingen	177	8	4,5	238,5	75,0	32,1	139	1	0,7	244,8	77,0	27,8
Ulm	62	11	17,7	211,4	66,5	33,6	2	0	0,0	213,5	67,1	10,5
Würzburg	131	14	10,7	229,4	72,1	36,9	92	1	1,1	243,4	76,5	25,3
Gesamt	2927	528	18,0	215,9	67,9	37,6	1025	67	6,5	233,5	73,4	33,4

¹ Aus Datenschutzgründen werden Ergebnisse einzelner Teilnehmer nicht angegeben, bei den Berechnungen aber berücksichtigt.

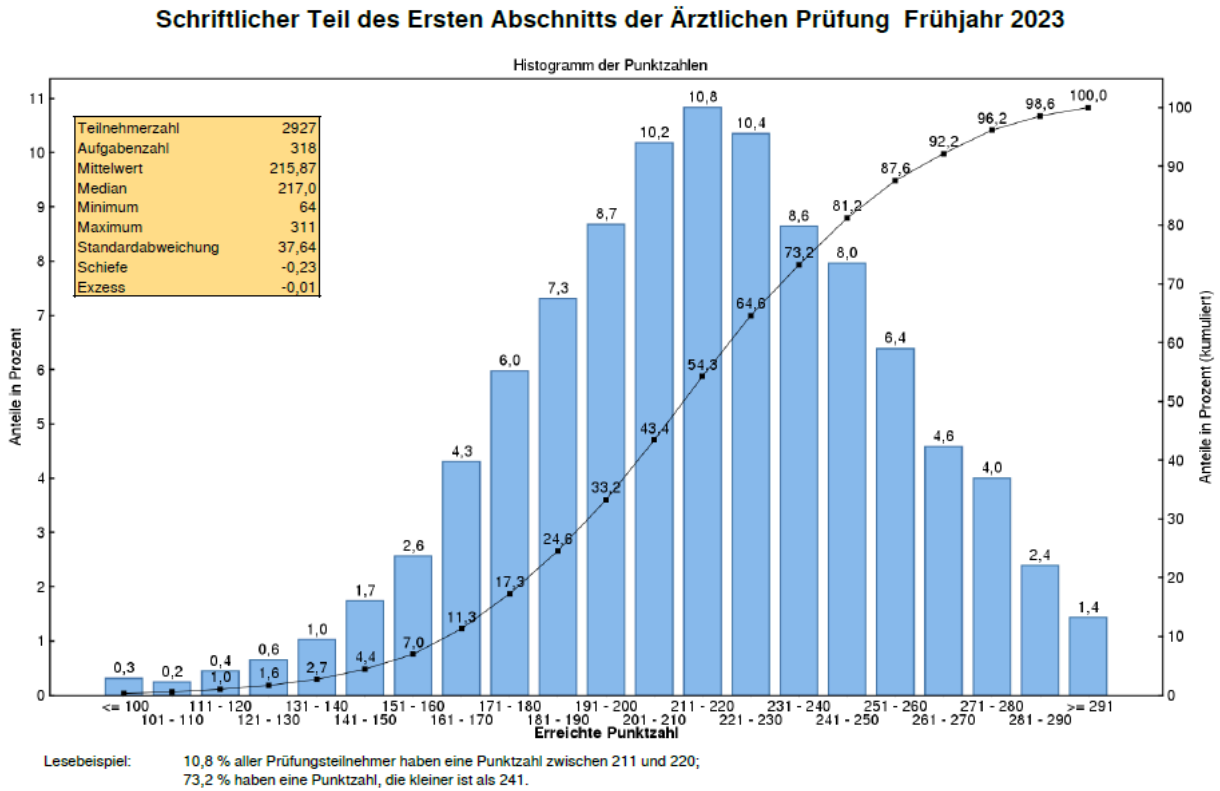
Referenzgruppe = Erstteilnehmer mit Mindeststudienzeit

Die Misserfolge sind unter Berücksichtigung des Nachteilsverbots (§ 14 Abs. 4 Satz 5 AAppO) ermittelt.

Quelle: Frühjahr 2023 Ergebnisse Medizin M1 | IMPP. (2023).

<https://www.impp.de/pruefungen/medizin/archiv-medizin.html>

Fig. 1.2: Ergebnisinformation Frühjahr 2023 (inkl. Durchschnitt, SD)



Quelle: Frühjahr 2023 Ergebnisse Medizin M1 | IMPP. (2023).
<https://www.impp.de/pruefungen/medizin/archiv-medizin.html>

Fig. 2.1: Ergebnisinformation Herbst 2023 (inkl. Misserfolge, Bestehensgrenze)

Herbst 2023

Ergebnisinformation

gewertete Aufgaben: 313 Stand: 14.09.2023
 Bestehensgrenze: 187

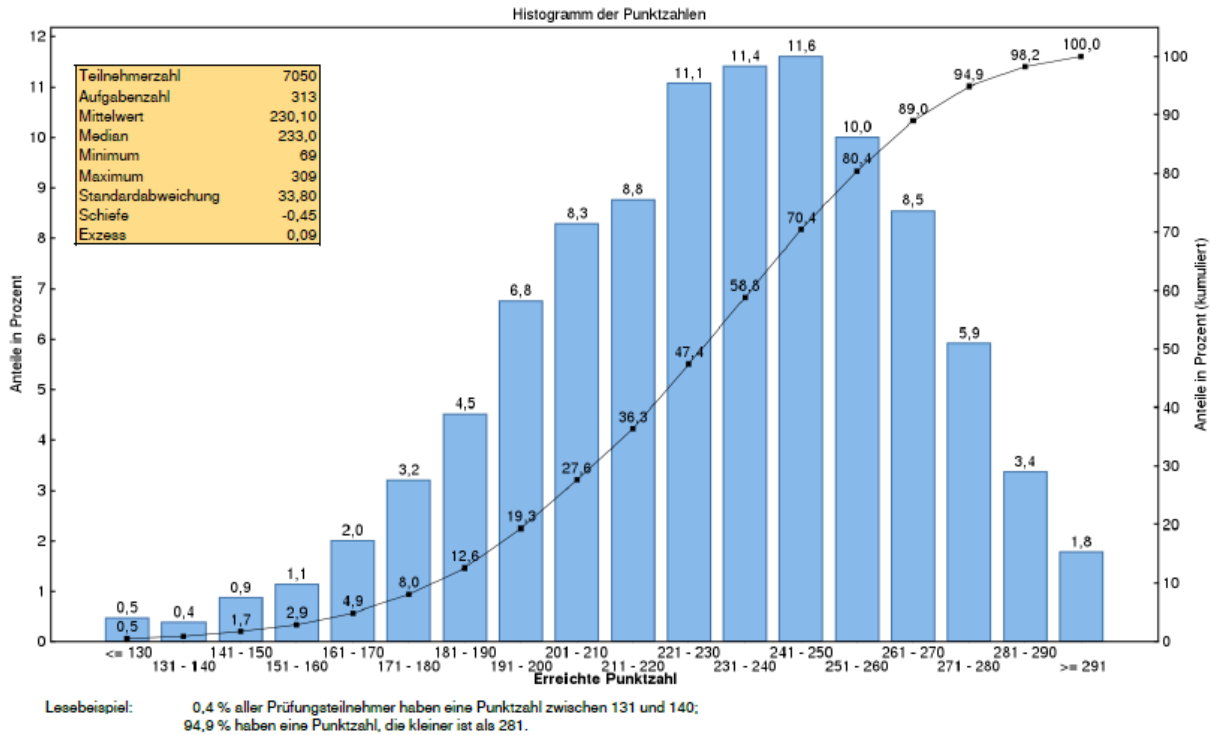
Hochschule	Gesamteilnehmer					Referenzteilnehmer						
	Teilnehmer	Misserfolge		Mittelwert		Std. abw.	Teilnehmer	Misserfolge		Mittelwert		Std. abw.
		abs.	in %	abs.	in %			abs.	in %	abs.	in %	
Berlin MS	142	50	35,2	198,7	63,5	36,3	98	25	25,5	205,9	65,8	33,6
Bochum	286	21	7,3	229,1	73,2	29,7	233	2	0,9	237,0	75,7	23,6
Bonn	251	26	10,4	230,4	73,6	34,5	198	9	4,5	239,7	76,6	27,6
Dresden	157	12	7,6	233,6	74,6	30,2	134	3	2,2	240,4	76,8	25,6
Erlangen	234	31	13,2	224,0	71,6	31,8	122	7	5,7	239,0	76,3	28,0
Essen	175	21	12,0	225,3	72,0	30,9	118	3	2,5	236,3	75,5	25,5
Frankfurt	333	27	8,1	229,7	73,4	30,3	282	9	3,2	236,2	75,5	25,1
Freiburg	265	14	5,3	237,5	75,9	30,4	218	3	1,4	244,1	78,0	25,7
Gießen	162	28	17,3	215,0	68,7	31,7	108	12	11,1	222,4	71,1	28,4
Göttingen	177	14	7,9	230,8	73,7	34,4	133	6	4,5	239,9	76,7	28,2
Greifswald	162	14	8,6	222,5	71,1	29,8	117	2	1,7	230,8	73,7	26,1
Halle	180	12	6,7	230,8	73,7	30,6	144	0	0,0	240,1	76,7	23,3
Hamburg MS	191	48	25,1	208,9	66,8	37,4	126	15	11,9	223,5	71,4	31,2
Heidelberg	329	11	3,3	255,5	81,6	31,6	297	2	0,7	260,0	83,1	27,0
Homburg	194	25	12,9	220,4	70,4	31,1	142	4	2,8	232,1	74,1	23,0
Jena	256	22	8,6	229,4	73,3	31,1	216	4	1,9	235,0	75,1	26,6
Kiel	148	3	2,0	231,1	73,8	26,8	134	1	0,7	233,6	74,6	25,3
Leipzig	310	10	3,2	241,5	77,2	26,6	287	4	1,4	244,7	78,2	24,1
Lübeck	162	12	7,4	239,8	76,6	35,1	143	1	0,7	248,5	79,4	24,7
Magdeburg	171	18	10,5	226,6	72,4	31,3	140	5	3,6	233,9	74,7	26,2
Mainz	204	15	7,4	236,3	75,5	30,9	105	0	0,0	253,1	80,8	20,0
Mannheim	244	14	5,7	240,5	76,8	32,4	219	4	1,8	246,0	78,6	26,5
Marburg	332	55	16,6	215,4	68,8	35,4	250	18	7,2	226,5	72,4	29,9
München LMU	716	62	8,7	230,7	73,7	33,2	603	25	4,1	236,4	75,5	29,6
Münster	175	21	12,0	236,4	75,5	40,3	126	6	4,8	250,9	80,2	30,5
Potsdam HMU	139	39	28,1	204,3	65,3	28,0	99	14	14,1	213,9	68,3	25,5
Regensburg	180	12	6,7	239,0	76,3	31,6	157	5	3,2	244,6	78,1	26,5
Rostock	140	1	0,7	248,7	79,5	24,1	121	0	0,0	253,7	81,1	19,1
Tübingen	203	17	8,4	233,5	74,6	33,6	154	4	2,6	243,9	77,9	28,5
Ulm	262	16	6,1	231,4	73,9	27,0	227	11	4,8	234,8	75,0	26,0
Würzburg	170	20	11,8	224,3	71,6	35,3	110	3	2,7	237,8	76,0	26,7
Gesamt	7050	691	9,8	230,1	73,5	33,8	5561	207	3,7	238,6	76,2	28,5

Referenzgruppe = Erstteilnehmer mit Mindeststudienzeit
 Die Misserfolge sind unter Berücksichtigung des Nachteilverbots (§ 14 Abs. 4 Satz 5 ÄAppO) ermittelt.

Quelle: Herbst 2023 Ergebnisse Medizin M1 | IMPP. (2023).
<https://www.impp.de/pruefungen/medizin/archiv-medizin.html>

Fig. 2.2: Ergebnisinformation Herbst 2023 (inkl. Durchschnitt, SD)

Schriftlicher Teil des Ersten Abschnitts der Ärztlichen Prüfung Herbst 2023



Quelle: Herbst 2023 Ergebnisse Medizin M1 | IMPP. (2023).
<https://www.impp.de/pruefungen/medizin/archiv-medizin.html>

Fig. 3.1: Ergebnisinformation Frühjahr 2024 (inkl. Misserfolge, Bestehensgrenze)

Frühjahr 2024

Ergebnisinformation

gewertete Aufgaben: 315
Bestehensgrenze: 183

Stand: 10.04.2024

Hochschule	Gesamtteilnehmer						Referenzteilnehmer					
	Teilnehmer	Misserfolge abs.	Misserfolge in %	Mittelwert abs.	Mittelwert in %	Std. abw.	Teilnehmer	Misserfolge abs.	Misserfolge in %	Mittelwert abs.	Mittelwert in %	Std. abw.
Berlin MS	138	52	37,7	191,4	60,8	36,9	79	12	15,2	208,7	66,2	31,5
Bochum	57	11	19,3	206,8	65,6	27,6	0					
Bonn	101	22	21,8	204,9	65,1	31,6	2	1	50,0	215,5	68,4	36,5
Dresden	49	7	14,3	205,9	65,4	29,9	0					
Erlangen	233	22	9,4	222,0	70,5	31,0	119	3	2,5	233,1	74,0	24,9
Essen	45	8	17,8	203,8	64,7	26,9	0					
Frankfurt	48	8	16,7	203,2	64,5	30,0	0					
Freiburg ¹	100	15	15,0	214,7	68,1	34,5	1					
Gießen	190	15	7,9	222,3	70,6	29,3	120	2	1,7	235,9	74,9	24,1
Göttingen	166	9	5,4	227,2	72,1	29,7	114	3	2,6	235,2	74,7	26,2
Greifswald	23	2	8,7	207,5	65,9	19,1	0					
Halle	50	13	26,0	203,6	64,6	33,6	0					
Hamburg MS	138	37	26,8	204,1	64,8	31,6	77	9	11,7	219,5	69,7	27,8
Heidelberg	24	6	25,0	208,2	66,1	33,5	2	0	0,0	214,5	68,1	16,5
Homburg	69	17	24,6	201,6	64,0	24,7	6	0	0,0	220,8	70,1	27,9
Jena	45	12	26,7	202,7	64,4	30,4	0					
Kiel ¹	36	3	8,3	206,3	65,5	30,6	1					
Leipzig	33	1	3,0	218,3	69,3	22,5	5	0	0,0	222,2	70,5	18,3
Lübeck	31	5	16,1	213,6	67,8	33,4	0					
Magdeburg	37	9	24,3	203,0	64,4	28,0	0					
Mainz	246	15	6,1	235,0	74,6	32,1	121	1	0,8	248,3	78,8	24,7
Mannheim	27	4	14,8	205,6	65,3	25,1	0					
Marburg	135	65	48,1	186,6	59,2	28,5	2	0	0,0	244,0	77,5	2,0
München LMU	186	48	25,8	201,9	64,1	30,7	0					
Münster	147	16	10,9	228,2	72,4	32,9	108	3	2,8	239,0	75,9	27,5
Potsdam HMU	105	15	14,3	209,1	66,4	30,7	52	1	1,9	227,9	72,4	24,7
Regensburg	42	5	11,9	206,8	65,7	27,1	0					
Rostock ¹	40	4	10,0	226,2	71,8	36,3	1					
Tübingen	159	11	6,9	237,4	75,4	33,0	124	1	0,8	246,6	78,3	26,5
Ulm	63	7	11,1	205,6	65,3	24,5	2	0	0,0	248,5	78,9	11,5
Würzburg	151	17	11,3	221,2	70,2	30,0	92	3	3,3	231,9	73,6	25,6
Gesamt	2914	481	16,5	214,3	68,0	33,9	1028	39	3,8	234,2	74,4	28,4

¹ Aus Datenschutzgründen werden Ergebnisse einzelner Teilnehmer nicht angegeben, bei den Berechnungen aber berücksichtigt.

Referenzgruppe = Erstteilnehmer mit Mindeststudienzeit

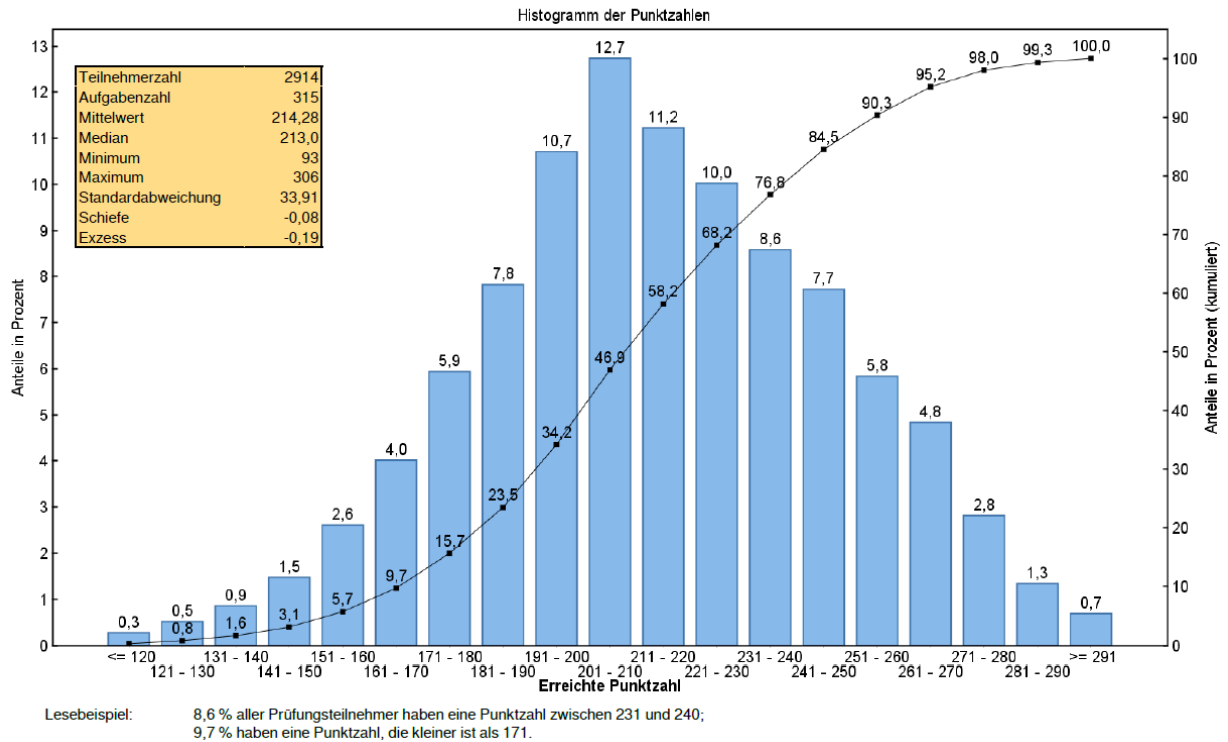
Die Misserfolge sind unter Berücksichtigung des Nachteilsverbots (§ 14 Abs. 4 Satz 5 ÄAppO) ermittelt.

Quelle: Frühjahr 2024 Ergebnisse Medizin M1 | IMPP. (2024).

<https://www.impp.de/pruefungen/medizin/archiv-medizin.html>

Fig. 3.2: Ergebnisinformation Frühjahr 2024 (inkl. Durchschnitt, SD)

Schriftlicher Teil des Ersten Abschnitts der Ärztlichen Prüfung Frühjahr 2024



Quelle: Frühjahr 2024 Ergebnisse Medizin M1 | IMPP. (2024).

<https://www.impp.de/pruefungen/medizin/archiv-medizin.html>

Fig. 4.1: Ergebnisinformation Herbst 2024 (inkl. Misserfolge, Bestehensgrenze)

Herbst 2024

Ergebnisinformation

gewertete Aufgaben: 315
Bestehensgrenze: 186

Stand: 12.09.2024

Hochschule	Gesamtteilnehmer					Referenzteilnehmer						
	Teilnehmer	Misserfolge abs.	Misserfolge in %	Mittelwert abs.	Mittelwert in %	Std. abw.	Teilnehmer	Misserfolge abs.	Misserfolge in %	Mittelwert abs.	Mittelwert in %	Std. abw.
Berlin ¹	1						0					
Berlin MS	146	53	36,3	198,0	62,8	35,0	90	25	27,8	207,9	66,0	35,4
Bochum	264	28	10,6	227,8	72,3	32,8	218	13	6,0	234,2	74,3	29,4
Bonn	255	25	9,8	226,9	72,0	34,4	206	8	3,9	234,0	74,3	28,7
Dresden	161	15	9,3	230,4	73,1	33,9	130	3	2,3	238,5	75,7	29,8
Erlangen	234	28	12,0	226,9	72,0	32,1	131	6	4,6	241,1	76,5	27,8
Essen	147	11	7,5	228,9	72,7	29,7	96	2	2,1	236,1	74,9	28,2
Frankfurt	338	19	5,6	228,8	72,6	29,0	280	10	3,6	233,2	74,0	27,6
Freiburg	271	22	8,1	237,3	75,3	34,1	225	1	0,4	247,3	78,5	23,5
Gießen	185	17	9,2	223,6	71,0	31,2	133	4	3,0	233,6	74,2	27,1
Göttingen	197	19	9,6	231,4	73,5	33,6	136	6	4,4	240,8	76,4	30,3
Greifswald	168	15	8,9	230,3	73,1	31,7	134	6	4,5	236,9	75,2	29,2
Halle	190	15	7,9	230,6	73,2	30,7	153	5	3,3	237,5	75,4	26,6
Hamburg MS	198	49	24,7	209,2	66,4	35,5	128	12	9,4	223,5	70,9	31,1
Heidelberg	329	5	1,5	255,8	81,2	30,0	299	5	1,7	259,7	82,4	27,5
Homburg	191	17	8,9	224,8	71,4	28,9	143	4	2,8	233,1	74,0	23,7
Jena	242	17	7,0	230,9	73,3	29,1	213	8	3,8	234,7	74,5	26,9
Kiel	158	15	9,5	226,6	71,9	29,2	146	10	6,8	229,0	72,7	27,7
Leipzig	295	10	3,4	237,7	75,5	27,6	281	5	1,8	239,4	76,0	25,7
Lübeck	167	11	6,6	239,5	76,0	32,7	148	5	3,4	245,3	77,9	27,7
Magdeburg	186	24	12,9	222,7	70,7	32,4	144	11	7,6	228,4	72,5	29,3
Mainz	199	12	6,0	235,5	74,8	30,0	105	0	0,0	249,6	79,2	24,3
Mannheim	251	16	6,4	239,4	76,0	31,4	232	9	3,9	241,6	76,7	29,0
Marburg	374	58	15,5	219,3	69,6	32,1	274	16	5,8	230,4	73,1	27,3
München LMU	696	67	9,6	230,5	73,2	32,9	600	29	4,8	236,1	74,9	29,4
Münster	143	13	9,1	243,8	77,4	38,4	108	3	2,8	256,9	81,5	29,0
Potsdam HMU	143	23	16,1	217,3	69,0	35,2	90	4	4,4	232,5	73,8	29,0
Regensburg	189	11	5,8	235,7	74,8	32,8	165	5	3,0	240,3	76,3	30,3
Rostock	145	3	2,1	241,9	76,8	27,2	109	0	0,0	249,8	79,3	20,7
Tübingen	194	9	4,6	231,3	73,4	30,4	140	3	2,1	239,2	75,9	28,0
Ulm	278	28	10,1	224,4	71,2	31,5	214	8	3,7	230,9	73,3	27,3
Würzburg	159	19	11,9	223,8	71,1	32,9	98	5	5,1	235,1	74,6	29,2
Gesamt	7094	675	9,5	230,0	73,0	33,4	5569	231	4,1	237,7	75,5	29,3

¹ Aus Datenschutzgründen werden Ergebnisse einzelner Teilnehmer nicht angegeben, bei den Berechnungen aber berücksichtigt.

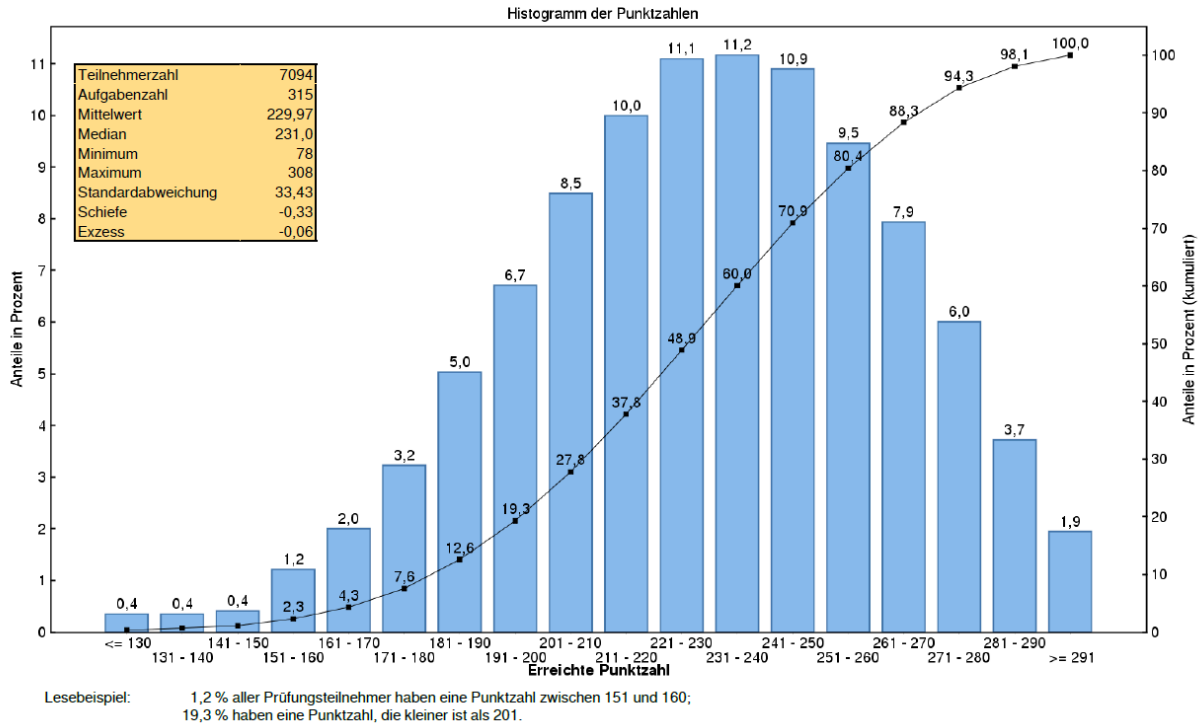
Referenzgruppe = Erstteilnehmer mit Mindeststudienzeit

Die Misserfolge sind unter Berücksichtigung des Nachteilsverbots (§ 14 Abs. 4 Satz 5 ÄAppO) ermittelt.

Quelle: Herbst 2024 Ergebnisse Medizin M1 | IMPP. (2024).
<https://www.impp.de/pruefungen/medizin/archiv-medizin.html>

Fig. 4.2: Ergebnisinformation Herbst 2024 (inkl. Durchschnitt, SD)

Schriftlicher Teil des Ersten Abschnitts der Ärztlichen Prüfung Herbst 2024



Quelle: Herbst 2024 Ergebnisse Medizin M1 | IMPP. (2024).
<https://www.impp.de/pruefungen/medizin/archiv-medizin.html>

Tabelle A1: Vollständige Ergebnisse der Chi-Quadrat-Anpassungstests F23

F23	Group	Observed Counts (Proportions)	95% CI1	Expected Counts (Proportions)	Z Value	p Value
	Erfolge	1153 (0.86)	0.84, 0.88	1102 (0.82)	3.65	<0.001
	Misserfolge	191 (0.14)	0.12, 0.16	242 (0.18)	-3.65	<0.001

Chi-Quadrat: 13.32, p -Wert: < .001, CI = confidence interval

Tabelle A2: Vollständige Ergebnisse der Chi-Quadrat-Anpassungstests H23

H23	Group	Observed Counts (Proportions)	95% CI1	Expected Counts (Proportions)	Z Value	p Value
	Erfolge	3252 (0.92)	0.91, 0.93	3179 (0.90)	4.16	<0.001
	Misserfolge	272 (0.08)	0.07, 0.09	345 (0.10)	-4.16	<0.001

Chi-Quadrat: 17.27, p -Wert: < .001, CI = confidence interval

Tabelle A3: Vollständige Ergebnisse der Chi-Quadrat-Anpassungstests F24

F24	Group	Observed Counts (Proportions)	95% CI1	Expected Counts (Proportions)	Z Value	p Value
	Erfolge	925 (0.85)	0.83, 0.87	910 (0.83)	-	-
	Misserfolge	165 (0.15)	0.13, 0.17	180 (0.17)	-	-

Chi-Quadrat: 1.49, p -Wert: .222, CI = confidence interval

Tabelle A4: Vollständige Ergebnisse der Chi-Quadrat-Anpassungstests H24

H24	Group	Observed Counts (Proportions)	95% CI1	Expected Counts (Proportions)	Z Value	p Value
	Erfolge	2633 (0.93)	0.92, 0.94	2574 (0.90)	3.76	<0.001
	Misserfolge	212 (0.07)	0.06, 0.08	271 (0.10)	-3.76	<0.001

Chi-Quadrat: 14.13, p -Wert: < .001, CI = confidence interval