

Spontaneously retrievable knowledge of German general practitioners depending on time since graduation, measured with the progress test medicine

Abstract

Background: General practitioners and general internists occupy a key position in German and Austrian healthcare systems. They provide primary care and act as gatekeepers between medical disciplines and sectors of care. Their explicit medical knowledge levels, however, can be quite disparate.

Objective: This study analyses whether general practitioners' performances on a standardised knowledge test changes with four relevant socio-demographic variables.

Design: The survey was based on the Progress Test Medicine (PTM), a standardised 200 item knowledge test on graduate level. After formal blueprinting and item analysis, 60 items of PTM were selected ("PTM-GP") for our study.

Participants: PTM-GP was presented ad hoc to general practitioners and internists from Germany and Austria at a number of professional meetings in 2011. 161 volunteers completed the survey.

Main measures: For evaluation, correlation analysis (Spearman), Kruskal Wallis-tests for non-parametric data and an analysis of covariance (ANCOVA) were calculated.

Results: Overall, four indicators turned out to be slightly significant for the performance on the PTM-GP, namely:

1. time passed since graduation,
2. the grade received in the licensing examination,
3. the type of institution for postgraduate training, and
4. the medical specialisation.

Conclusions: Recent graduates performed better in the PTM-GP; a doctor's licensing examination grade as well as training at a university hospital correlated positively with PTM-GP performance.

A general doctor's knowledge level is moderately influenced by exam grades, time since graduation, the institutional affiliation of postgraduate training and medical specialisation. Individual changes in knowledge over time have to be deliberately considered in lifelong learning. In consequence, the on-going teaching of medical knowledge should be integrated mandatory and verifiable into general doctors' everyday practices, e.g. through repetitive knowledge tests with individual feedback and recommendations for further continuing medical education (CME).

Keywords: general practitioners, replication study, progress test medicine, knowledge development, life-long learning

Introduction

General practitioners (GPs) are one of the central contact points for patients in the German and Austrian health care systems. They should have broad knowledge [1] in all medical disciplines to serve their role as a "gatekeeper" [2] in patient care. Driven by new technologies and procedures as well as a better exchange of information,

available medical knowledge is increasing rapidly and doctors are required to continue their education during their practice in order to keep up with new developments. It is estimated that the half-life of medical knowledge is currently five to ten years [3], [4] and that medical knowledge as a whole approximately doubles every six to eight years [5]. Up-to-date medical knowledge is of high importance to general practitioners. Since the as-

Michaela Zupanic¹

Jelena Kreuer²

Daniel Bauer³

Zineb M. Nouns⁴

Jan P. Ehlers⁵

Martin R. Fischer⁶

1 Witten/Herdecke University, Faculty of Health, Interprofessional and Collaborative didactics, Witten, Germany

2 Clenia Schlössli AG, Klinik für Psychiatrie und Psychotherapie, Oetwil am See, Switzerland

3 University of Bern, Faculty of Medicine, Institute for Medical Education (IML), Bern, Switzerland

4 Helios Kliniken GmbH, Academy at Fresenius Helios, Berlin, Germany

5 Witten/Herdecke University, Faculty of Health, Didactics and Educational Research in Health Science, Witten, Germany

6 LMU Munich, University Hospital, Institute for Medical Education, Munich, Germany

assessment of the current status of knowledge of primary care physicians has so far received little attention in Germany, our study aims to shed some light on the question as to what determines the medical knowledge of general practitioners and how their knowledge base changes over time. The publication "Changes over time in the knowledge base of practicing internists" by Ramsey [6] served as an inspiration and methodological model for the research question of our study. Ramsey identified several predictors of a good track record in knowledge-based exams. The procedure allowed a comparison of their results with the performance level of current graduates of that time. The results showed, among others findings, a significant negative correlation between the time elapsed since the medical licensing exam and the test performance; that the better the result of the original Internal Medicine Specialization Exam, the better the knowledge test performance; that internists trained at an academic hospital achieved significantly better results than colleagues who had their residency at non-academic hospitals; that, interestingly, there was no difference in the knowledge test result with regard to participation in further education courses; and that general internists achieved better test results than internists of technical sub-disciplines such as cardiologists or gastroenterologists [6].

Ramsey's results inspired the research questions underlying this work. Thus, we made the following assumptions for general practitioners in Germany and Austria:

1. The more years since graduation from medical school have passed, the less medical knowledge can be spontaneously called up in a knowledge test.
2. The better the grade in the final medical examination, the better the performance of GPs on a knowledge test.
3. GPs who started their postgraduate clinical training at university and teaching hospitals perform better on a knowledge test than those who were trained in hospitals without university affiliation.
4. The performance on a knowledge test concerning internal medicine depends on the specific specialist training of GPs.

With these assumptions, the study aims to address the research gap concerning GPs' knowledge changes over time in Germany and Austria and at the same time looks for starting points for further development of continuing education and training for general practitioners.

Subjects and methods

Subjects

For this cross-sectional study, general practitioners were contacted during professional meetings between October 2010 and September 2011 at nine German quality circles as well as in Salzburg, Austria, during the annual conference of the German College of General Practitioners and

Family Physicians (DEGAM). In total 161 participants, including 25 from Austria, completed the survey for evaluation. The statistically necessary minimum sample size was met for Kruskal-Wallis-tests (expected differences=7, power=0.8, n=112) and Spearman correlation analyses (correlation=0.3, power=0.8, n=84), but failed for analysis of covariance (ANCOVA) (covariates=2, power=0.8, n=269) (G*Power 3) [7]. The survey with overall 58 questions is divided into four sections:

1. socio-demographic information of study participants,
2. specialist training and additional training,
3. professional progression with career milestones and
4. CME (Continuing Medical Education) habits (see attachment 1).

After collecting the data, each participant received her/his personal results in comparison to the entire group as a feedback concerning the individual knowledge status by mail.

Study instruments

Medical knowledge was assessed using a modified and shortened Progress Test Medicine (PTM). The original PTM of Charité Berlin consists of 200 Multiple Choice (MC)-questions from all medical specialties at German graduate level [8]. The three resulting PTM-GP-scores were "correct", "false" and "total" (correct minus false). In contrast to common multiple choice tests, the PTM includes a "don't know"-option, which is supposed to reduce guessing when uncertain [9], [10]. The PTM no. 21 (Autumn 2009) served as basis for the "PTM for General Practitioners" (PTM-GP) with 60 MC-questions used in the present study. Using item analysis (based on data from graduate level test takers), the best items according to formal criteria with an item difficulty between $P=.40-.80$ [11] and an item total correlation $r>.30$ [12] were selected. Despite the reduction of items, there was no marked decrease in meaningfulness, since the items were selected respecting the original blueprint based on both organ systems and clinical subjects, and were examined for internal and external consistency. The number of items stating the subject distribution is shown in table 1. With 13 items the largest share was attributed to internal medicine topics.

Statistical analyses

The hypotheses that there are relationships between subjects' socio-demographic variables (time since graduation, examination grade, residency, academic teaching hospital) and performance on PTM-GP Scores (correct, false and total) was investigated using correlation analyses (Spearman), with the correlation coefficient r as effect size. Kruskal-Wallis tests for non-parametric data were performed to prove that the middle ranks in the groups are the same. The test variable can be converted analogously into the effect size d [13], [14]. The influence of biographical variables on performance on PTM-

Table 1: Number of items stating the subject distribution of PTM-GP(n=60)

Preclinical subjects:	2
Anatomy, biology	1
Physiology, physics	1
Cross-sectional subjects:	10
Pathology	3
Pharmacology, toxicology	6
Radiology, nuclear medicine	1
Clinical subjects	48
General medicine	4
Occupational and social medicine	1
Ophthalmology	1
Surgery	3
Gynecology and obstetrics	3
Otorhinolaryngology	2
Internal medicine	13
Pediatrics	9
Naturopathic medicine, physical medicine	1
Neurology	5
Orthopedics	1
Psychiatry, Psychosomatics	4
Forensic medicine	1

GP total score was determined with analysis of covariance (ANCOVA) [15]. All statistical analyses were performed using SPSS 21.0 and R. The α -level was set at 5%.

Results

Demographic characteristics

The sample consisted of 161 participants (109 male, 67.7%) with an average age of 51 years (SD 8.4 years; range 31 to 67 years). The mean year of graduation was 1988, so that an average of 23 years had passed since graduation (SD 8.6 years; range 3 to 42 years). 125 participants (77.6%) reported their grade in the final licensing examination (missing $n=36$, 22.4%).

89.4% ($n=144$) of the participants had completed their residency training. Completion was 16 years ago on average (SD 7.9 years; range 2 to 37 years). Training qualifications were as GP (Facharzt für Allgemeinmedizin in German; $n=97$, 60.2%), GP with internal medicine certification ($n=21$, 13.0%), internist ($n=16$, 9.9%) or family doctor (Praktischer Arzt in German; $n=10$, 6.2%).

At the time of the survey, most participants ($n=103$, 64%) worked in a private practice, 37 (23%) in a non-academic teaching hospital, 14 (8.7%) in an academic teaching hospital and only 7 (4.3%) in a university hospital.¹ The first employment was in a non-academic teaching hospital for most participants ($n=69$, 42.9%), 57 (35.4%) in an academic teaching hospital and 23 ($n=14.3%$) in a university hospital. 5 participants started their career in a private practice (3.1%; missing $n=7$, 4.3%).

PTM-GP performance

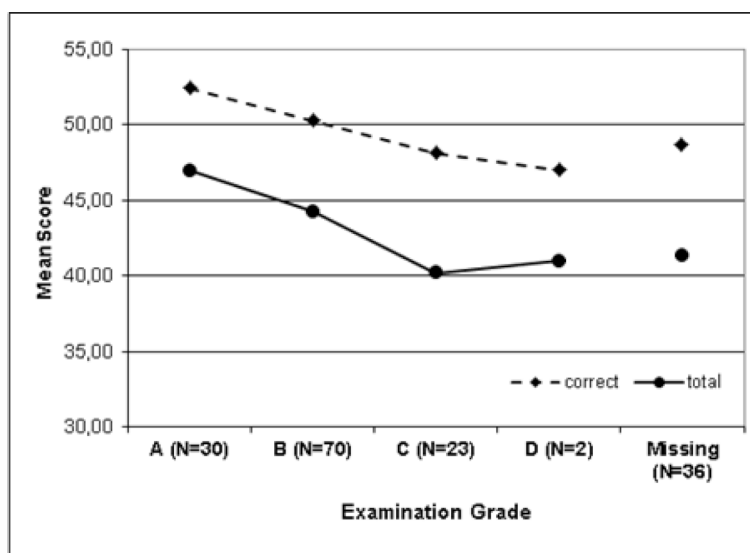
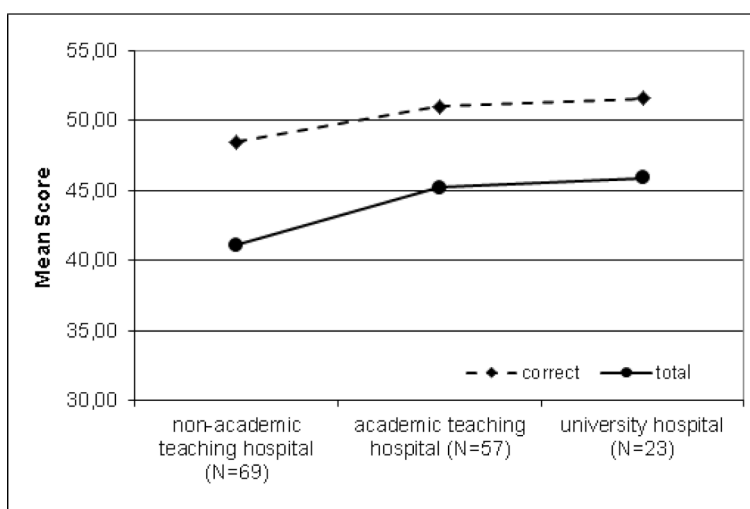
The obtained PTM-GP-scores of the 161 participants were 50.0 ± 5.1 correct responses (range 33-60), and 6.5 ± 3.7 false responses (range 0-24) with the resulting mean test score of 43.5 ± 8.0 (range 11-60). The more time passed since graduation, resp. year of graduation, the lower the score "correct" with a weak significant inverse correlation ($r=-.166$, $p=.035$) and the higher the score "false" with a weak significant positive correlation ($r=.196$, $p=.013$). Table 2 indicates the descriptive statistics (mean and standard deviation) in four categories, showing minor differences in time since graduation.

The indicated examination grade of 125 participants was 2.0 ± 0.7 (minimum 1, equivalent to A; maximum 4, equivalent to D). With a better score at graduation, correlation analyses yield fewer false answers ($r=.188$, $p=.036$) and significantly more correct answers ($r=-.289$, $p=.001$). The differences in PTM-GP scores depending on the examination grade is moderate significant for correct answers (Kruskal-Wallis test: $\chi^2=9.5$, $p=.023$, $d=.382$). Results for the five groups of participants, grade A ($n=30$, 18.6%), grade B ($n=70$, 43.5%), grade C ($n=23$, 14.3%), grade D ($n=2$, 1.2%), and missing ($n=36$, 22.4%) are shown in figure 1.

In most participants ($n=69$, 42.9%), the first employment was in a non-academic teaching hospital, 57 (35.4%) in an academic teaching hospital and 23 ($n=14.3%$) in a university hospital (see figure 2). These three groups were compared with a non-parametric Kruskal-Wallis test. The PTM-GP scores "correct" ($\chi^2=8.8$, $p=.012$, $d=.424$) and "total" are moderate significantly higher ($\chi^2=7.9$, $p=.019$, $d=.394$), when the first employment was at a university hospital (45.9 ± 6.6 , 95% CI 43.0-48.8) or academic teaching hospital (45.2 ± 6.2 , 95% CI 43.6-46.9), com-

Table 2: Categories regarding time since graduation and results of PTM-GP (n=161)

year of graduation	n	„correct“ M ± SD	95 % CI	„total“ M ± SD	95 % CI
before 1981	41	48.7 ± 5.7	46.9 – 50.5	41.2 ± 8.9	38.4 – 44.0
1981-1987	38	49.4 ± 4.8	47.8 – 51.0	42.8 ± 7.4	40.3 – 45.2
1988-1994	43	50.8 ± 4.9	49.3 – 52.3	44.5 ± 8.0	42.0 – 47.0
1995-2008	39	50.9 ± 4.6	49.4 – 52.4	45.4 ± 7.4	43.0 – 47.8

**Figure 1: Examination grade and PTM-GP scores of participants (n=161)****Figure 2: First employment and PTM-GP scores of participants (n=161; Missing 12 participants with n=5 in private practice and n=7 Missing)**

pared to first employment at non-academic teaching hospitals (41.1±9.3, 95% CI 38.9–43.3). However, there were no significant differences in results by taking the participants' workplaces at the time of the survey (see above) into account (see figure 2).

The PTM-GP scores of the four specialist groups general practitioner (n=97, 60.2%), general practitioner with internal medicine certification (n=21, 13.0%) and internists (n=16, 9.9%) or family doctor (n=10, 6.2%) differ moderately significant in non-parametric examination with Kruskal-Wallis test. This is true for "correct" answers ($\chi^2=16.4$, $p=.001$, $d=.611$), "false" answers ($\chi^2=15.1$,

$p=.002$, $d=.578$) and the PTM-GP score "total" ($\chi^2=16.9$, $p=.001$, $d=.623$). Figure 3 shows that the best performance on the PTM-GP was achieved by the internists (49.4±8.4, 95% CI 44.9-53.8) and the lowest performance by the family doctors (36.3±7.8, 95% CI 30.7) (see figure 3).

Comparable results were found in the 13 questions on internal medicine. The PTM-GP scores of the four specialist groups were examined by using the non-parametric Kruskal-Wallis test. Moderate significant differences in the PTM-GP scores of the specialist groups for "correct" ($\chi^2=11.3$, $p=.010$, $d=.473$), "false" ($\chi^2=13.4$, $p=.004$,

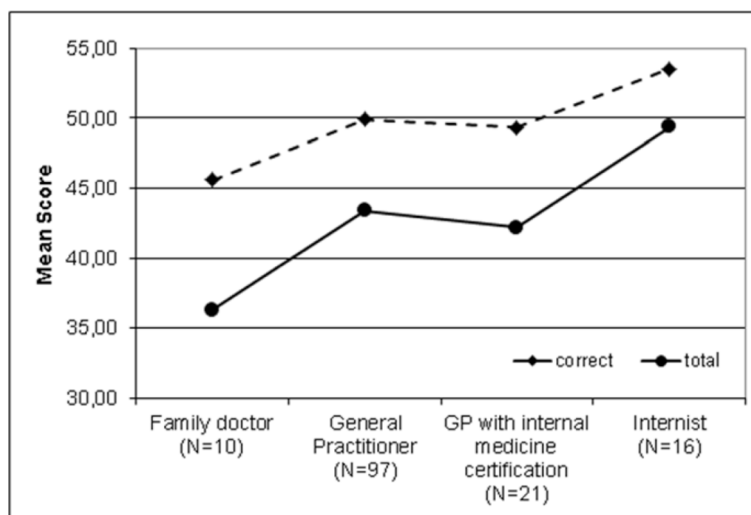


Figure 3: Specialist groups and PTM-GP scores of participants (n=161; Missing 17 participants not yet having completed residency training)

Table 3: Specialist groups and results of PTM-GP concerning internal medicine (n=161)

	n	„correct“ M ± SD	95 % CI	„total“ M ± SD	95 % CI
family doctor	10	8.9 ± 2.4	7.2 – 10.6	5.8 ± 4.0	2.9 – 8.7
general practitioner	97	10.5 ± 1.7	10.2 – 10.9	8.9 ± 2.9	6.8 – 9.9
GP with internal medicine certification	21	10.3 ± 2.0	9.4 – 11.2	8.3 ± 3.3	6.8 – 9.9
internist	16	11.6 ± 1.7	10.6 – 12.5	10.6 ± 2.6	9.2 – 12.0

d=.533) and “total” ($\chi^2=12.1$, $p=.007$, $d=.496$) became clear and are shown in table 3.

To investigate the influence of the independent socio-demographic variables on the PTM-GP total score, an analysis of covariance (ANCOVA) was carried out with the specialist group and type of institution for postgraduate training, adjusted for time passed since graduation and the grade received in the licensing examination. When interpreting the results, it must be taken into account that the statistically necessary minimum sample for ANCOVA has not been reached (see above) and that 36 participants did not report their examination grade. Resulting sample for ANCOVA consist therefore of 105 participants (65% of the whole sample). Table 4 shows the results for the main effects of the within-groups factors and the covariates on the PTM-GP total score. The corrected model of ANCOVA explained moderate 27% of the variability (partial eta square=0.267). The examination grade had the greatest influence on performance (highly significant 10.8% clarification of variance) and time since graduation had a significant but lesser influence on PTM-GP total score (significant 6.8% clarification of variance). There is no effect of the two factors specialist group and residency, nor their interaction.

A further model of ANCOVA without the covariate examination grade and n=137 participants resulted in 19% explained variability (partial eta square=0.190). The greatest influence on performance had the type of residency (7.7%

clarification of variance; significant). There were no other significant effects.

Discussion

The assumptions of this replication study based on the results of the study by Ramsey can mostly be confirmed [6]. As shown in the results, there is a weak to moderate significant correlation between the spontaneously available general practitioner's knowledge and the

1. time elapsed since graduation,
2. the score in the licensing examination,
3. first career station and
4. specialist status.

Time that has passed since the graduation

The slight decrease of medical knowledge in the course of working life could be due to the fact that older GPs are less confronted with the patient problems of university hospital level in their daily work. In addition, it is likely they are less familiar with multiple-choice questions as their younger colleagues. This would mean that the exam questions do not represent a GP's everyday routine well [16].

The “don't know” option could also have caused a distortion. This option is included in the progress test medicine

Table 4: Socio-demographic variables and results of PTM-GP total Score in ANCOVA (n=105)

Source	Type III sum of squares	df	Mean squares	F	p-value	Partial eta square
Specialist group	74.14	3	24.71	1.06	.370	.034
Residency	176.11	3	58.70	2.52	.063	.077
Time since graduation	153.00	1	153.00	6.57	.012	.068
Examination grade	253.62	1	253.62	10.89	.001	.108
Specialist groups*residency	61.43	6	10.24	0.44	.851	.028

to discourage guessing and to simulate authentic clinical practice [17]. In the present study, 36 participants (22%) never chose the “don’t know” option. It may be that the “don’t know” option was ruled out, as it is an audit without real consequence (that is, without real decision-making situations on the patient with potentially threatening consequences), prompting bolder decisions from participants than they would have shown with real patients. Carline [18], in contrast to Ramsey [6] and Norcini [19], found no significant correlation between the times elapsed since certification as an internist and the completion of a recertification test for internists [18]. However, they only examined physicians who had passed their first certification exam 5-10 years earlier. Due to the limitations mentioned, the results of Ramsey in 1991 and Norcini in 1985 may be considered more meaningful [6], [16]. Also, a multistage study from Canada comparing different methods of assessing primary care physicians found that medical literacy was inversely and significantly related to age and time since the exam [20].

Final grade in the licensing examination

In the literature, the importance of exam grades for further degrees occupies a wide space. Ramsey showed that physicians with a higher examination grade performed better at the specialist examination, too [6]. In a German study it was found that the performance on the internist part of the progress test medicine in 2010 correlated significantly with the school leaving examination grade and the first part of the medical examination grade [21]. A similar correlation between the final grade and the examinations in the course of studies was demonstrated in numerous other studies [22], [23], [24]. The relationship between the level of knowledge and the completion of the specialist recertification examination and further examinations before physicians entered medical practice is described in several studies. In addition to school grades, other studies found influences of the country of origin, the field of study, the time between school and study achievement and the study section as relevant predictors [25], [26]. The obtained exam grade seems a good predictor for later performance on standard knowledge-based tests such as the knowledge test used in this study. In the model of ANCOVA, the examination grade had the greatest influence on PTM-GP performance and explained 10.8% of variance.

First career station

The results showed that physicians who chose to continue education at a university hospital at the beginning of their careers achieved significantly better results in the knowledge test than physicians in non-academic hospitals. The nature of the first training facility, in turn, correlated with the score in the test score, as well as within the individual score groups. The extent of university affiliation of the first employment explains 7.7% of variance on PTM-GP performance, when the best predictor “exam grade” is not taken into account in the ANCOVA because of the many missing data points. It can be assumed that physicians at university hospitals are more likely to be involved in research and teaching and that their knowledge will be updated more reliably. This also has an impact on answering knowledge questions in tests. Due to the large teams of university hospitals, younger physicians are more challenged to justify themselves to their peers and to be more frequently audited in hierarchical structures. The more frequent “exam situation” trains quick-to-learn knowledge and may lead to better results in tests.

It remains unclear whether the different training and further education on offer at training institutions has an impact on the level of knowledge. A study by Lehmann and Schultz found no significant difference in training behavior between physicians working in university hospitals and physicians working in peripheral care [27]. At university hospitals it takes longer to obtain the status of a specialist than that of houses with standard care, which is attributed, among other things, to the higher time spent on research and teaching [28]. Haffner and Schmidt postulate the creation of a training regulation that is based on the realistic competencies of future general practitioners and promotes their acquisition of medical knowledge. This is in their view the only way to strengthen the outpatient sector [29].

Specialist status

Specialists in internal medicine on average score more points in the test score of the knowledge test than general practitioners. One reason for this might be that practitioners are versed in a broad routine, but less well-prepared for specialized theoretical questions in the knowledge test. Another reason may be a sample matching age-based confounding effect, as more than 60% of the family doctors are among the older ones in the sample

(median 52 years), whereas more than 60% of the internists belong to the younger participants. In the model of ANCOVA, the time since graduation resp. inverse age of participants explained significant 6.8% of the variance. Ramsey showed a rather opposite result than that found in this study [6]. According to their research, the well-educated generalists performed better than the specialized cardiologists and gastroenterologists. The contradiction is certainly also related to the different specialist training in Germany and the US.

Limitations of the current study

While the comparison with data from the Federal Register of Physicians [30] shows a similar distribution pattern nationwide of general practitioners in terms of gender, age and specialist training, the recruitment of study participants to quality circles and congresses could have led to a selection bias, since it can be assumed that especially motivated, inquisitive general practitioners attend congresses and quality circles. Although all primary care physicians are required to undergo training and to collect CME points, a distortion of the study results cannot be ruled out. A randomized random sample was not possible in this study design. Possible effects due to the selection of participants are therefore not excluded. The motivation factor gains additional weight because the participants had to invest a relatively long time in answering the questionnaire with a scope of seven pages and the 60 test questions. Participation in the study was voluntary, reward or sanction was not expected in any case. An incentive may have been the offer of feedback concerning one's individual knowledge status. But this is not comparable to a summative assessment for an academic and/or professional qualification, so that the lower motivation of the participants could have affected the results of the study. Ramsey found a similar limitation when comparing the results of voluntary study participants with the results of those physicians who had to undergo recertification and had a marked different degree of success [6]. A study focusing on low-stake tests, i. e. non-impact testing, may support the assumption that the external motivation or demotivation factors of the subjects could have played a role in the present study [31].

Conclusion and outlook

Study results from the US [6] could be replicated in Germany and Austria although general practitioner's education and career paths in the US and Germany are not directly comparable. The decline of medical knowledge over time seems to be relatively independent of the respective health care system.

This study was able to assess the performance of GPs in a standardized progress test and presents different predictors for their performance. In total, four indicators proved to be weak to moderate significant for performance:

1. The time elapsed since the licensing exam. There was a weak negative correlation between the time elapsed since graduation and the results of the knowledge test.
2. The grade achieved in the licensing exam. The results of the study show a moderate correlation between a good final grade and a good result in the knowledge test. The exam grade showed the greatest effect of the used socio-demographic variables on performance in ANCOVA.
3. The type of hospital in which the specialist training took place. Doctors who opted for continuing education in a university hospital at the beginning of their careers achieved moderate significantly better results in the knowledge test than doctors in non-academic hospitals did.
4. Whether the participant is a general practitioner or an internist. Internists achieved moderate better results on the knowledge test than general practitioners did.

While factual knowledge is a very important resource in the working life of a general practitioner, the experience-based and practical competence of a physician or the attention paid to the patient can be equally important. In order to keep it at the highest possible level over the entire period of active practice, the teaching of medical knowledge should be integrated into the non-academic everyday life of the general practitioner in a contemporary way, e.g. through repetitive progress tests with feedback on the individual knowledge profiles for planning their own further continuing medical education.

Note

We chose not to exclude the 58 participants from the study who reported having their current primary workplace in inpatient care, because their attendance at quality groups aimed at GPs or their participation at the annual conference of the German College of General Practitioners and Family Physicians (DEGAM) respectively, suggested a mutual professional interest in primary care medicine, suggesting that albeit having different workplace settings they shared mutual patient cohorts.

Acknowledgements

The authors thank all general practitioners who participated in this study and we owe sincere gratitude to the President of the DEGAM conference in Salzburg, Mr. Univ.-Prof. Dr. Andreas Sönnichsen, for supporting the implementation of this research. We would also like to thank Prof. Dr. med. Jörg Schelling and Mr. Matthias Holzer for supporting our study.

Competing interests

The authors declare that they have no competing interests.

Attachments

Available from

<https://www.egms.de/en/journals/zma/2020-37/zma001342.shtml>

- Attachment_1.pdf (647 KB)
Additional questionnaire for the Progress Test Medicine

References

- Woods N, Brooks L, Norman G. The role of biomedical knowledge in diagnosis of difficult clinical cases. *Adv Health Sci Educ Theory Pract.* 2007;12(4):417-426. DOI: 10.1007/s10459-006-9054-y
- Linden M, Gothe H, Ormel J. Der Hausarzt als Gatekeeper. *Dtsch Arztebl.* 2004;101(39):2103-2110.
- Hall J, Platell C. Half-life of truth in surgical literature. *Lancet.* 1997;350(9093):1752. DOI: 10.1016/S0140-6736(05)63577-5
- Masic I, Novo A, Delkovic S, Onerhodzic I, Piralic A. How to assess and improve quality of medical education: lessons learned from Faculty of Medicine in Sarajevo. *Bosn J Basic Med Sci.* 2007;7(1):74-78. DOI: 10.17305/bjbm.2007.3097
- Mantovani F, Castelnuovo G, Gaggioli A, Riva G. Virtual reality training for health-care professionals. *Cyberpsychol Behav.* 2003;6(4):389-395. DOI: 10.1089/109493103322278772
- Ramsey PG, Carline JD, Inui TS, Larson EB, LoGerfo JP, Norcini JJ, Wenrich MD. Changes over time in the knowledge base of practicing internists. *JAMA.* 1991;266(8):1103-1107. DOI: 10.1001/jama.1991.03470080073032
- Faul F, Erdfelder E, Lang AG, Buchner A. G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods.* 2007;39(2):175-191. DOI: 10.3758/BF03193146
- Nouns ZM, Georg W. Progress testing in German speaking countries. *Med Teach.* 2010;32(6):467-470. DOI: 10.3109/0142159X.2010.485656
- Osterberg K, Kölbel S, Brauns K. Der Progress Test Medizin: Erfahrungen an der Charité Berlin. *GMS Z Med Ausbild.* 2006;23(3):Doc46. Zugänglich unter/available from: <https://www.egms.de/static/de/journals/zma/2006-23/zma000265.shtml>
- Schauber SK, Hecht M, Nouns ZM, Kuhlmei A, Dettmer S. The role of environmental and individual characteristics in the development of student achievement: a comparison between a traditional and a problem-based-learning curriculum. *Adv Health Sci Educ Theory Pract.* 2015;20:1033-1052. DOI: 10.1007/s10459-015-9584-2
- Möltner A, Schellberg D, Jünger J. Grundlegende quantitative Analysen medizinischer Prüfungen. *GMS Z Med Ausbild.* 2006;23(3):Doc53. Zugänglich unter/available from: <https://www.egms.de/static/de/journals/zma/2006-23/zma000272.shtml>
- Lienert GA, Raatz U. Testaufbau und Testanalyse. 6. Auflage. Weinheim: Psychologie Verlags Union; 1998.
- Cohen J. *Statistical power analysis for the behavioral sciences.* Hillsdale: Erlbaum; 1988.
- Fritz CO, Morris PE, Richler JJ. Effect size estimates: Current use, calculations, and interpretation. *J Experim Psychol Gen.* 2012;141(1):30. DOI: 10.1037/a0026092
- Lenhard W, Lenhard A. Berechnung von Effektstärken. Dettelbach: Psychometrica; 2016. Zugänglich unter/available from: <https://www.psychometrica.de/effektstaerke.html>
- Fleming AW. Examination scores fall with time: but so what? *JAMA.* 1992;267(3):366. DOI: 10.1001/jama.1992.03480030044031
- Ravesloot CJ, van der Schaaf MF, Muijtjens AM, Haaring C, Kruitwagen CJ, Beek FJ, Bakker J, van Schaik JP, Ten Cate TJ. The don't know option in progress testing. *Adv Health Sci Educ Theory Pract.* 2015;20(5):1325-1338. DOI: 10.1007/s10459-015-9604-2
- Carline J, Inui T, Larson E, Logerfo J, Ramsey P. The knowledge base of certified internists: relationships to training, practice type, and other physician characteristics. *Arch Intern Med.* 1989;149(10):2311-2313. DOI: 10.1001/archinte.1989.00390100115025
- Norcini J, Lipner R, Benson J, Webster G. An analysis of the knowledge base of practicing internists as measured by the 1980 recertification examination. *Ann Intern Med.* 1985;102(3):385-389. DOI: 10.7326/0003-4819-102-3-385
- Norman G, Davis D, Lamb S, Hanna E, Caulford P, Kaigas T. Competency assessment of primary care physicians as part of a peer review program. *JAMA.* 1993;270(9):1046-1051. DOI: 10.1001/jama.1993.03510090030007
- Schmidmaier R, Holzer M, Angstwurm M, Nouns Z, Reincke M, Fischer MR. Using the Progress-Test Medizin (PTM) for evaluation of the Medical Curriculum Munich (MeCuM). *GMS Z Med Ausbild.* 2010;27(5):Doc70. DOI: 10.3205/zma000707
- Dlugosch S. Prognose von Studienerfolg: dargestellt am Beispiel des Auswahlverfahrens der Bucerius Law School. Aachen: Shaker; 2005.
- Gentsch S. Richtig ausgewählt? Eine Evaluation neuer Verfahren der Studierendenauswahl in den Fächern Medizin und Pharmazie an der Goethe-Universität. Berlin: Logos-Verlag; 2009.
- Schiefele U, Streblov L, Ermgassen U, Moschner B. Lernmotivation und Lernstrategien als Bedingungen der Studienleistung. Ergebnisse einer Längsschnittstudie. *Z Päd Psychol.* 2003;17:185-198. DOI: 10.1024//1010-0652.17.34.185
- Tamblyn R, Abrahamowicz M, Brailovsky C, Grand Maison P, Lescop J, Norcini J, Girard N, Haggerty J. Association between licensing examination scores and resource use and quality of care in primary care practice. *JAMA.* 1998;280(11):989-996. DOI: 10.1001/jama.280.11.989
- Trapmann S, Hell B, Weigand S, Schuler H. Die Validität von Schulnoten zur Vorhersage des Studienerfolgs-eine Metaanalyse. *Z Päd Psychol.* 2007;21:11-27. DOI: 10.1024/1010-0652.21.1.11
- Lehmann K, Schultz J. Zur Lage der anästhesiologischen Weiter- und Fortbildung in Deutschland Ergebnisse einer Repräsentativumfrage. *Anaesthesist.* 2001;50:248-261. DOI: 10.1007/s001010170028
- Ansorg J, Hassan I, Fendrich V, Polonius M, Rothmund M, Langer P. Qualität der chirurgischen Weiterbildung in Deutschland. *Dtsch Med Wochenschr.* 2005;130:508-513. DOI: 10.1055/s-2005-863084
- Haffner C, Schmidt M. Weiterbildung Allgemeinmedizin: Sind Hausärzte wirklich noch erwünscht? *Dtsch Arztebl.* 2008;105(30):A-1635/B-1411/C-1379.

30. Bundesärztekammer. Statistics at a glance. Berlin: Bundesärztekammer; 2011. Zugänglich unter/available at: <https://www.bundesaerztekammer.de/ueber-uns/aerztestatistik/>
31. Silm G, Must O, Taeht K. Test-taking effort as a predictor of performance in low-stakes tests. J Human Soc Sci. 2013;17(4):433-448. DOI: 10.3176/tr.2013.4.08

Please cite as

Zupanic M, Kreuer J, Bauer D, Nouns ZM, Ehlers JP, Fischer MR. Spontaneously retrievable knowledge of German general practitioners depending on time since graduation, measured with the progress test medicine. GMS J Med Educ. 2020;37(5):Doc49. DOI: 10.3205/zma001342, URN: urn:nbn:de:0183-zma0013421

This article is freely available from

<https://www.egms.de/en/journals/zma/2020-37/zma001342.shtml>

Corresponding author:

Prof. Dr. phil. Michaela Zupanic
Witten/Herdecke University, Faculty of Health,
Interprofessional and Collaborative didactics,
Alfred-Herrhausen-Str. 50, D-58448 Witten, Germany
Michaela.Zupanic@uni-wh.de

Received: 2020-02-04

Revised: 2020-05-23

Accepted: 2020-06-30

Published: 2020-09-15

Copyright

©2020 Zupanic et al. This is an Open Access article distributed under the terms of the Creative Commons Attribution 4.0 License. See license information at <http://creativecommons.org/licenses/by/4.0/>.

Spontan abrufbares Wissen deutscher Hausärzt*innen in Abhängigkeit von der Zeit seit Studienabschluss, gemessen mit dem Progress Test Medizin

Zusammenfassung

Hintergrund: Hausärzt*innen und Internist*innen nehmen eine Schlüsselposition im deutschen und österreichischen Gesundheitswesen ein. Sie sind für die Primärversorgung zuständig und fungieren als Gatekeeper zwischen den unterschiedlichen medizinischen Fach- und Versorgungsbereichen. Hinsichtlich ihres medizinischen Wissensstandes können sie sich jedoch deutlich unterscheiden.

Zielsetzung: Ziel dieser Studie ist die Beantwortung der Fragestellung, ob die Prüfungsleistungen von Hausärzt*innen in einem standardisierten Wissenstest von vier relevanten soziodemografischen Merkmalen beeinflusst werden.

Material: Der für die Studie verwendete Fragebogen basierte auf dem Progress Test Medizin (PTM), einem standardisierten Wissenstest auf Absolvierendenniveau bestehend aus 200 Fragen. Durch Berücksichtigung einer genauen Vorgabe zur Verteilung von klinischen Fächern und Organsystemen, dem sogenannten „Blueprint“, sowie nach einer Bewertung der Fragen, wurden 60 Fragen des PTM für diese Studie ausgewählt: PTM – General Practitioners („PTM-GP“).

Teilnehmende: Der PTM-GP wurde Hausärzt*innen und Internist*innen aus Deutschland und Österreich ad hoc im Rahmen von mehreren Fachtagungen im Jahr 2011 vorgelegt. Es nahmen 161 Freiwillige am Wissenstest teil.

Methodik: Zur Bewertung der Ergebnisse wurden Korrelationsanalysen (Spearman), Kruskal-Wallis-Tests für nicht-parametrische Daten sowie eine Kovarianzanalyse (ANCOVA) durchgeführt.

Ergebnisse: Insgesamt erwiesen sich vier Indikatoren als leicht signifikant für das Abschneiden beim PTM-GP:

1. die Zeit seit Studienabschluss,
2. die Note im 3. Abschnitt der Ärztlichen Prüfung,
3. die Art der Einrichtung, an der die Fachärzt*innenausbildung absolviert wurde, sowie
4. das medizinische Fachgebiet.

Schlussfolgerung: Kürzlich Graduierte schnitten beim PTM-GP besser ab; die Note im 3. Abschnitt der ärztlichen Prüfung sowie die Fachärzt*innenausbildung an einem Universitätsklinikum hatten einen positiven Einfluss auf die Prüfungsleistung im PTM-GP.

Der Wissensstand von Hausärzt*innen wird moderat beeinflusst von Examensnoten, der Zeit seit Studienabschluss, der institutionellen Zugehörigkeit in der postgradualen Fachärzt*innenausbildung und dem Fachgebiet. Individuelle Veränderungen des Wissensstandes im zeitlichen Verlauf müssen bei der Planung von Fortbildungsmaßnahmen (lebenslanges Lernen) sorgfältig berücksichtigt werden. Daher sollte eine kontinuierliche medizinische Fortbildung verpflichtend und nachprüfbar in den hausärztlichen Berufsalltag integriert werden, beispielsweise durch wiederholte Wissenstests mit individuellem Feedback und Empfehlungen für die ärztliche Fortbildung (CME).

Schlüsselwörter: Hausärzt*innen, Replikationsstudie, Progress Test Medizin, Wissensentwicklung, lebenslanges Lernen

Michaela Zupanic¹

Jelena Kreuer²

Daniel Bauer³

Zineb M. Nouns⁴

Jan P. Ehlers⁵

Martin R. Fischer⁶

1 Universität Witten/Herdecke, Fakultät für Gesundheit, Interprofessionelle und Kollaborative Didaktik in Medizin- und Gesundheitsberufen, Witten, Deutschland

2 Clenia Schlössli AG, Klinik für Psychiatrie und Psychotherapie, Oetwil am See, Schweiz

3 Universität Bern, Medizinische Fakultät, Institut für Medizinische Lehre (IML), Bern, Schweiz

4 Helios Kliniken GmbH, Hochschule Fresenius Helios, Berlin, Deutschland

5 Universität Witten/Herdecke, Fakultät für Gesundheit, Didaktik und Bildungsforschung im Gesundheitswesen, Witten, Deutschland

6 LMU München, Klinikum der Universität München, Institut für Didaktik und Ausbildungsforschung in der Medizin, München, Deutschland

Einleitung

Hausärzt*innen sind eine der zentralen Anlaufstellen für Patient*innen im deutschen und österreichischen Gesundheitswesen. Sie sollten über ein breitgefächertes Wissen [1] in allen medizinischen Fachrichtungen verfügen, um ihrer Rolle als Gatekeeper [2] in der Patientenversorgung gerecht zu werden. Durch neue Technologien und Verfahren sowie einer Verbesserung des Informationsaustausches, nimmt der Umfang verfügbaren medizinischen Wissens schnell zu, und so müssen Ärzt*innen sich auch während der Ausübung ihres Berufes fortbilden, um mit neuen Entwicklungen Schritt halten zu können. Es wird angenommen, dass die aktuelle Halbwertszeit medizinischen Wissens fünf bis zehn Jahre beträgt [3], [4] und dass sich das gesamte medizinische Wissen alle sechs bis acht Jahre verdoppelt [5]. Aktuelles medizinisches Wissen ist für Hausärzt*innen von großer Bedeutung. Da eine Einschätzung des aktuellen Wissensstandes von Ärzt*innen in der Primärversorgung in Deutschland bisher wenig Beachtung fand, ist es das Ziel unserer Studie, den medizinischen Wissensstand von Hausärzt*innen und dessen Veränderung im Laufe ihres Berufslebens zu untersuchen. Die Studie "Changes over time in the knowledge base of practicing internists" von Ramsey [6] diente als Inspiration und methodische Vorlage für die Forschungsfrage unserer Untersuchung. Ramsey ermittelte mehrere Prädiktoren für gute Leistungen in wissensbasierten Prüfungen. Das Verfahren ermöglichte einen Vergleich der Ergebnisse mit dem Leistungsstand damaliger Absolvent*innen. Die Studie ergab unter anderem einen signifikant negativen Zusammenhang zwischen der Zeit seit der medizinischen Abschlussprüfung und der Prüfungsleistung; dass die Leistung im Wissenstest umso besser war, je besser das Ergebnis der ursprünglichen Fachärzt*innenprüfung für Innere Medizin war; dass Internist*innen, die ihre Fachärzt*innenausbildung an einem Universitätsklinikum antraten, signifikant bessere Ergebnisse erzielten als Kolleg*innen, die ihre Fachärzt*innenausbildung an nicht-akademischen Krankenhäusern absolvierten; dass die Teilnahme oder Nicht-Teilnahme an Fortbildungsmaßnahmen interessanterweise keinen Unterschied bei den Ergebnissen des Wissenstests ergab, und dass allgemeinärztlich tätige Internist*innen bessere Ergebnisse erzielten als Internist*innen, die in Spezialgebieten tätig waren, wie z. B. Kardiologen oder Gastroenterologen [6].

Die Ergebnisse der Studie von Ramsey waren die Grundlage für die Formulierung der Forschungsfragen dieser Untersuchung. Folgende Annahmen für Hausärzt*innen in Deutschland und Österreich wurden getroffen:

1. Je mehr Jahre seit Abschluss des Medizinstudiums vergangen sind, desto weniger medizinisches Wissen kann spontan in einem Wissenstest abgerufen werden.

2. Je besser die Note in der medizinischen Abschlussprüfung war, umso besser ist das Abschneiden von Hausärzt*innen bei einem Wissenstest.
3. Hausärzt*innen, die ihre Fachärzt*innenausbildung an Universitätskliniken und akademischen Lehrkrankenhäusern begannen, schneiden im Wissenstest besser ab als Kolleg*innen, die ihre Weiterbildung an Krankenhäusern ohne Anbindung an eine Universität absolvierten.
4. Das Abschneiden bei einem Wissenstest zu Innerer Medizin hängt von der spezifischen Fachärzt*innenausbildung der Hausärzt*innen ab.

Basierend auf diesen Annahmen ist das Ziel dieser Studie, die Veränderung des medizinischen Wissensstandes von Hausärzt*innen im zeitlichen Verlauf in Deutschland und Österreich zu untersuchen und gleichzeitig Ansatzpunkte für eine Weiterentwicklung der Fort- und Weiterbildungsmaßnahmen für Hausärzt*innen zu präsentieren, um die bestehende Forschungslücke zu schließen.

Studienteilnehmende und Methodik

Studienteilnehmende

Für diese Querschnittsstudie wurden Hausärzt*innen im Rahmen von neun deutschen Qualitätszirkeln sowie in Salzburg, Österreich, während der Jahrestagung der Deutschen Gesellschaft für Allgemeinmedizin und Familienmedizin (DEGAM) im Zeitraum zwischen Oktober 2010 und September 2011 rekrutiert. Insgesamt 161 Ärzt*innen, davon 25 aus Österreich, füllten den Fragebogen aus. Die statistisch erforderliche Mindeststichprobe wurde für die Durchführung von Kruskal-Wallis-Tests (erwartete Unterschiede=7, Power=0.8, N=112) sowie von Korrelationsanalysen nach Spearman (Korrelation=0.3, Power=0.8, N=84) erreicht, jedoch nicht für die Durchführung einer Kovarianzanalyse (ANCOVA) (Kovariate=2, Power=0.8, N=269) (G*Power 3) [7]. Der Fragebogen mit insgesamt 60 Fragen ist in vier Abschnitte unterteilt:

1. soziodemografische Daten der Teilnehmenden,
2. Fachärzt*innenausbildung und Zusatzausbildung,
3. beruflicher Werdegang inklusive beruflicher Stationen,
4. ärztliche Fortbildungsgewohnheiten (siehe Anhang 1).

Nach der Datenerhebung erhielten alle Teilnehmende ihre persönlichen Ergebnisse im Vergleich zu allen anderen Teilnehmenden als Rückmeldung zum persönlichen Wissensstand per E-Mail.

Erhebungsmethoden

Der medizinische Wissensstand wurde anhand einer modifizierten und gekürzten Fassung des Progress Tests Medizin (PTM) ermittelt. Der ursprüngliche PTM der Charité Berlin besteht aus 200 Fragen im Multiple-Choice-

Tabelle 1: Anzahl der Fragen je Fachgebiet im PTM-GP (N=60)

Vorklinische Fächer:	2
Anatomie, Biologie	1
Physiologie, Physik	1
Querschnittsfächer:	10
Pathologie	3
Pharmakologie, Toxikologie	6
Radiologie, Nuklearmedizin	1
Klinische Fächer	48
Allgemeinmedizin	4
Arbeits- und Sozialmedizin	1
Augenheilkunde	1
Chirurgie	3
Frauenheilkunde und Geburtshilfe	3
Hals-Nasen-Ohrenheilkunde	2
Innere Medizin	13
Kinderheilkunde	9
Naturheilverfahren, Physikalische Medizin	1
Neurologie	5
Orthopädie	1
Psychiatrie, Psychosomatik	4
Rechtsmedizin	1

(MC)-Format zu allen medizinischen Fachbereichen auf dem Niveau deutscher Absolvent*innen [8]. Die drei resultierenden PTM-Werte waren „richtig“, „falsch“ und „gesamt“ (Anzahl der richtig beantworteten Fragen minus Anzahl der falsch beantworteten Fragen). Im Unterschied zu herkömmlichen Multiple-Choice-Tests beinhaltet der PTM auch eine „Weiß nicht“-Antwortoption, um zu vermeiden, dass Antworten geraten werden [9], [10]. Der PTM Nr. 21 (Herbst 2009) diente als Basis für den PTM für Hausärzt*innen/PTM for General Practitioners („PTM-GP“) mit 60 MC-Fragen für die vorliegende Studie. Nach einer Bewertung der Fragen (basierend auf den Daten von Teilnehmenden auf Absolvierendenniveau) wurden die für Hausärzt*innen am besten geeigneten Fragen anhand formaler Kriterien mit einem Schwierigkeitsgrad zwischen $P=.40-.80$ [11] und einer Trennschärfe der Fragen von $r>.30$ [12] ausgewählt. Obwohl der für diese Studie verwendete Test weniger Fragen enthielt, wurde seine Aussagekraft nicht wesentlich geschmälert, da die Auswahl der Fragen auf Grundlage einer festgelegten Verteilung nach klinischen Fächern und Organsystemen der Originalversion („Blueprint“) erfolgte und die Fragen auf interne und externe Konsistenz überprüft wurden. Tabelle 1 gibt die Anzahl der Fragen je Fachgebiet an. Die meisten Fragen bezogen sich auf internistische Themen (13 Fragen).

Statistische Auswertungen

Die Annahmen über eine Korrelation zwischen soziodemografischen Daten der Teilnehmenden (Zeit seit Studienabschluss, Examensnote, Fachärzt*innenausbildung, akademisches Lehrkrankenhaus) und der Werte im PTM-GP (richtig, falsch und gesamt) wurden mit Hilfe von Korrelationsanalysen (Spearman) untersucht, mit dem

Korrelationskoeffizienten r als Effektstärke. Ferner wurden Kruskal-Wallis-Tests für nicht-parametrische Daten durchgeführt, um zu belegen, dass die mittleren Rangplätze bei allen Gruppen die gleichen sind. Die Testgröße kann analog in die Effektstärke d umgewandelt werden [13], [14]. Der Einfluss biografischer Daten auf den PTM-GP-Gesamtwert wurde mit einer Kovarianzanalyse (ANCOVA) berechnet [15]. Alle statistischen Auswertungen wurden in SPSS 21.0 und R vorgenommen, bei einem Signifikanzniveau α von 5%.

Ergebnisse

Demografische Merkmale

Die Stichprobe umfasste 161 Teilnehmende (davon waren 109 männlich, 67.7%) mit einem Durchschnittsalter von 51 Jahren (Standardabweichung (SD) 8.4 Jahre; Range 31 bis 67 Jahre). Das mittlere Abschlussjahr war 1988, sodass im Mittel 23 Jahre seit Studienabschluss vergangen waren (SD 8.6 Jahre; Range 3 bis 42 Jahre). Von den Teilnehmenden gaben 125 (77.6%) die Note ihrer Abschlussprüfung an (Missing: $N=36$, 22.4%).

Von den Teilnehmenden hatten 89.4% ($N=144$) ihre Fachärzt*innenausbildung abgeschlossen, im Schnitt vor 16 Jahren (SD 7.9 Jahre; Range 2 bis 37 Jahre). Sie erzielten ihre Qualifikation als Fachärzt*innen für Allgemeinmedizin ($N=97$, 60.2%), als Fachärzt*innen für Allgemeinmedizin mit Schwerpunkt in Innerer Medizin ($N=21$, 13.0%), Internist*innen ($N=16$, 9.9%) oder als Praktische Ärzt*innen ($N=10$, 6.2%).

Zur Zeit der Teilnahme am Wissenstest arbeiteten die meisten Teilnehmenden ($N=103$, 64%) als niedergelassene Ärzt*innen, 37 (23%) in einem nicht-akademischen

Lehrkrankenhaus, 14 (8.7%) in einem akademischen Lehrkrankenhaus und lediglich 7 (4.3%) in einem Universitätsklinikum.¹ Die erste Anstellung der meisten Teilnehmenden erfolgte in einem nicht-akademischen Lehrkrankenhaus (N=69, 42.9%), während 57 Teilnehmende (35.4%) zunächst in einem akademischen Lehrkrankenhaus arbeiteten und 23 (N=14.3%) in einem Universitätsklinikum. Fünf der Teilnehmenden begannen ihre Berufslaufbahn als niedergelassene Ärzt*innen (3.1%; Missing: N=7, 4.3%).

PTM-GP-Ergebnisse

Die 161 Teilnehmenden erzielten die folgenden PTM-GP-Werte: 50.0±5.1 richtige Antworten (Range 33-60) sowie 6.5±3.7 falsche Antworten (Range 0-24) mit einem daraus resultierenden mittleren Testwert von 43.5±8.0 (Range 11-60). Je mehr Zeit seit dem Studienabschluss bzw. seit dem Jahr des Studienabschlusses vergangen war, umso geringer war die Anzahl der richtigen Antworten mit einer schwach signifikanten negativen Korrelation ($r=-.166$, $p=.035$) und umso höher die Anzahl der falschen Antworten mit einer schwach signifikanten positiven Korrelation ($r=.196$, $p=.013$). Tabelle 2 zeigt die statistischen Werte (Mittelwert und Standardabweichung) in vier Kategorien mit geringfügigen Unterschieden in der Zeit seit Studienabschluss.

Die von 125 Teilnehmenden angegebene Examensnote war 2.0±0.7 (beste Note 1, vergleichbar mit „A“ im anglo-amerikanischen Raum; schlechteste Note 4, vergleichbar mit „D“). Die Korrelationsanalysen ergaben weniger falsche Antworten ($r=.188$, $p=.036$) und signifikant mehr richtige Antworten ($r=-.289$, $p=.001$) bei einer besseren Examensnote. Die Unterschiede in den PTM-GP-Werten sind in Abhängigkeit von der Examensnote moderat signifikant in Bezug auf richtige Antworten (Kruskal-Wallis-Test: $\chi^2=9.5$, $p=.023$, $d=.382$). Abbildung 1 veranschaulicht die Ergebnisse für die fünf Gruppen der Teilnehmenden: Note 1 (N=30, 18.6%), Note 2 (N=70, 43.5%), Note 3 (N=23, 14.3%), Note 4 (N=2, 1.2%) und Missing (N=36, 22.4%).

Die meisten Teilnehmenden (N=69, 42.9%) waren im Anschluss an den Studienabschluss zunächst an einem nicht-akademischen Lehrkrankenhaus, 57 (35.4%) an einem akademischen Lehrkrankenhaus und 23 (N=14.3%) an einem Universitätsklinikum angestellt (siehe Abbildung 2). Diese drei Gruppen wurden mit Hilfe eines nicht-parametrischen Kruskal-Wallis-Tests verglichen. Die Anzahl der richtigen Antworten im PTM-GP ($\chi^2=8.8$, $p=.012$, $d=.424$) und der Gesamtwert waren moderat signifikant höher ($\chi^2=7.9$, $p=.019$, $d=.394$), wenn die Teilnehmenden zuerst an einem Universitätsklinikum (45.9±6.6, 95% CI 43.0-48.8) oder einem akademischen Lehrkrankenhaus (45.2±6.2, 95% CI 43.6-46.9) angestellt waren, verglichen mit einer Erstanstellung an einem nicht-akademischen Lehrkrankenhaus (41.1±9.3, 95% CI 38.9-43.3). Eine Berücksichtigung des Arbeitsplatzes der Teilnehmenden zum Befragungszeitpunkt (siehe

oben) führte jedoch nicht zu signifikanten Ergebnisunterschieden (siehe Abbildung 2).

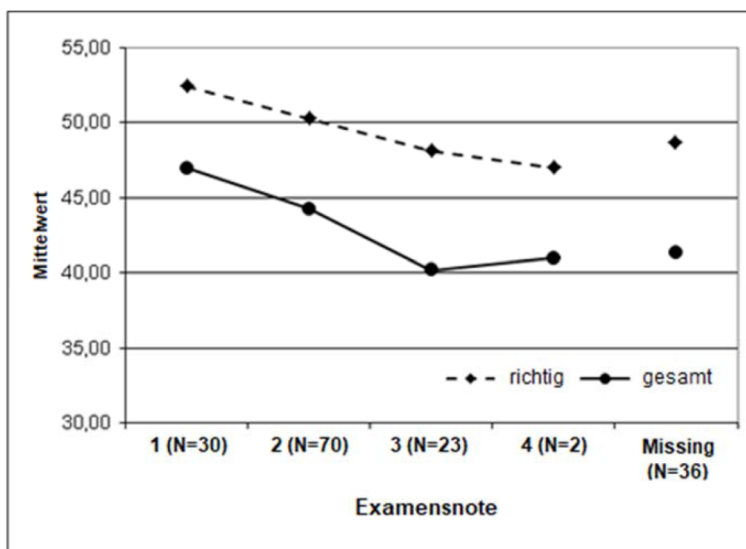
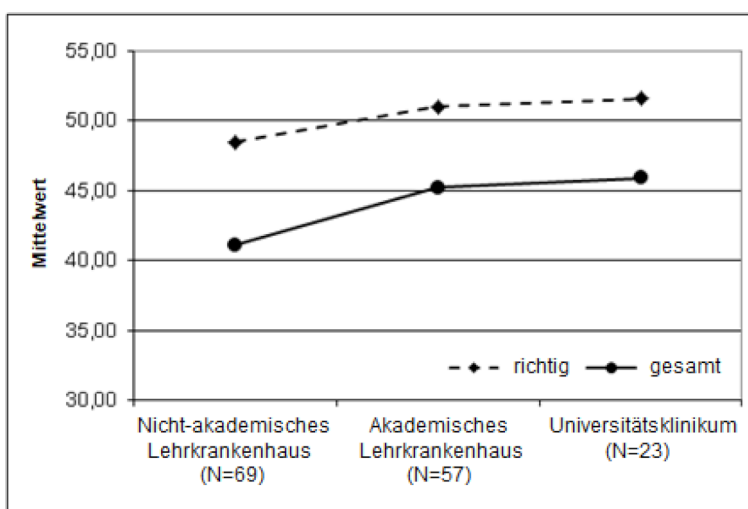
Die nicht-parametrische Auswertung durch den Kruskal-Wallis-Test ergab, dass sich die PTM-GP-Werte der vier Gruppen Fachärzt*innen für Allgemeinmedizin (N=97, 60.2%), Fachärzt*innen für Allgemeinmedizin mit Schwerpunkt in Innerer Medizin (N=21, 13.0%), Internist*innen (N=16, 9.9%) und Praktische Ärzt*innen (N=10, 6.2%) moderat signifikant unterschieden. Dies trifft auf die Anzahl der richtigen Antworten ($\chi^2=16.4$, $p=.001$, $d=.611$), die Anzahl der falschen Antworten ($\chi^2=15.1$, $p=.002$, $d=.578$) und auf den PTM-GP-Gesamtwert ($\chi^2=16.9$, $p=.001$, $d=.623$) zu. Abbildung 3 verdeutlicht, dass die besten PTM-GP-Werte von Internist*innen erzielt wurden (49.4±8.4, 95% CI 44.9-53.8) und die schlechtesten von Praktischen Ärzt*innen (36.3±7.8, 95% CI 30.7).

Vergleichbare Ergebnisse zeigten sich bei den 13 Fragen auf dem Gebiet der Inneren Medizin. Die PTM-GP-Werte der vier Fachärzt*innengruppen wurden mit Hilfe des nicht-parametrischen Kruskal-Wallis-Tests berechnet. Es bestanden moderat signifikante Unterschiede zwischen den PTM-GP-Werten der Fachärzt*innengruppen hinsichtlich der Anzahl der richtigen Antworten ($\chi^2=11.3$, $p=.010$, $d=.473$), der falschen Antworten ($\chi^2=13.4$, $p=.004$, $d=.533$) und dem Gesamtwert ($\chi^2=12.1$, $p=.007$, $d=.496$), dargestellt in Tabelle 3.

Um den Einfluss der unabhängigen soziodemografischen Daten auf den Gesamtwert des PTM-GP zu untersuchen, wurde eine Kovarianzanalyse (ANCOVA) mit der Fachärzt*innengruppe und der Art der postgradualen Ausbildungseinrichtung durchgeführt, bereinigt um die Zeit seit Studienabschluss und die Examensnote. Bei der Interpretation der Ergebnisse muss berücksichtigt werden, dass die für eine ANCOVA statistisch erforderliche Mindeststichprobe nicht erreicht wurde (siehe oben) und dass 36 Teilnehmende ihre Examensnote nicht angaben. Die Stichprobe für die Durchführung der ANCOVA bestand daher aus 105 Teilnehmenden (65% der gesamten Stichprobe). Tabelle 4 stellt die Ergebnisse in Bezug auf die wesentlichen Auswirkungen der Faktoren innerhalb der jeweiligen Gruppen sowie der Kovariaten auf den PTM-GP-Gesamtwert dar. Das berichtete ANCOVA-Modell ergab eine mittlere Variabilität von 27% (partielles Eta-Quadrat=0.267). Die Examensnote hatte den größten Einfluss auf das Testergebnis (hoch signifikante 10.8% Varianzaufklärung). Die Zeit seit Studienabschluss hatte einen signifikanten, aber geringeren Einfluss auf den PTM-GP-Gesamtwert (signifikante 6.8% Varianzaufklärung). Die beiden Faktoren Fachärzt*innengruppe und Fachärzt*innenausbildung hatten weder einzeln noch in ihrer Wechselwirkung einen Einfluss auf die Ergebnisse. Ein weiteres ANCOVA-Modell ohne die Kovariate Examensnote und N=137 Teilnehmende ergab eine erklärte Variabilität von 19% (partielles Eta-Quadrat=0.190). Die Art der Einrichtung der Fachärzt*innenausbildung hatte den größten Einfluss auf das Abschneiden im Test (7.7% Varianzaufklärung; signifikant). Es gab keine weiteren signifikanten Einflüsse.

Tabelle 2: Kategorien für die Zeit seit Studienabschluss und PTM-GP-Ergebnisse (N=161)

Jahr des Studienabschlusses	N	„richtig“ M ± SD	95 % CI	„gesamt“ M ± SD	95 % CI
vor 1981	41	48.7 ± 5.7	46.9 – 50.5	41.2 ± 8.9	38.4 – 44.0
1981-1987	38	49.4 ± 4.8	47.8 – 51.0	42.8 ± 7.4	40.3 – 45.2
1988-1994	43	50.8 ± 4.9	49.3 – 52.3	44.5 ± 8.0	42.0 – 47.0
1995-2008	39	50.9 ± 4.6	49.4 – 52.4	45.4 ± 7.4	43.0 – 47.8

**Abbildung 1: Examensnote und PTM-GP-Werte der Teilnehmenden (N=161)****Abbildung 2: Erstanstellung und PTM-GP-Werte der Teilnehmenden (N=161; Missing: 12 Teilnehmende; N=5 niedergelassene Ärzt*innen und N=7 Missing)**

Diskussion

Die Annahmen dieser Replikationsstudie basierend auf den Ergebnissen einer Studie von Ramsey können weitgehend bestätigt werden [6]. Wie die Ergebnisse der vorliegenden Studie zeigen, existiert eine schwache bis moderate Korrelation zwischen dem spontan abrufbaren Wissen von Hausärzt*innen und

1. der Zeit seit Studienabschluss,
2. der Examensnote,
3. der ersten beruflichen Station und

4. dem Status als Fachärzt*innen.

Zeit seit Studienabschluss

Die geringfügige Abnahme medizinischen Wissens im Laufe des Berufslebens könnte auf die Tatsache zurückzuführen sein, dass ältere Hausärzt*innen weniger mit solchen Gesundheitsproblemen konfrontiert sind, wie sie im Berufsalltag an einem Universitätsklinikum behandelt werden. Darüber hinaus haben sie vermutlich weniger Erfahrung mit Multiple-Choice-Fragen als ihre jüngeren Kolleg*innen. Dies würde bedeuten, dass die gestellten

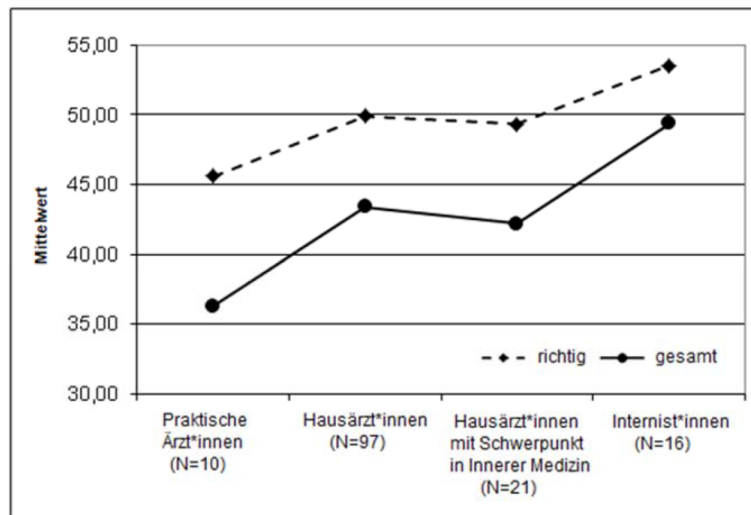


Abbildung 3: Fachärzt*innengruppen und PTM-GP-Werte der Teilnehmenden (N=161; Missing: 17 Teilnehmende ohne abgeschlossene Fachärzt*innenausbildung)

Tabelle 3: Fachärzt*innengruppen und PTM-GP-Ergebnisse bei den Fragen zur Inneren Medizin (N=161)

	N	„richtig“ M ± SD	95 % CI	„gesamt“ M ± SD	95 % CI
Praktische Ärzt*innen	10	8.9 ± 2.4	7.2 – 10.6	5.8 ± 4.0	2.9 – 8.7
Hausärzt*innen	97	10.5 ± 1.7	10.2 – 10.9	8.9 ± 2.9	6.8 – 9.9
Hausärzt*innen mit Schwerpunkt in Innerer Medizin	21	10.3 ± 2.0	9.4 – 11.2	8.3 ± 3.3	6.8 – 9.9
Internist*innen	16	11.6 ± 1.7	10.6 – 12.5	10.6 ± 2.6	9.2 – 12.0

Tabelle 4: Soziodemografische Merkmale und ANCOVA-Ergebnisse des PTM-GP-Gesamtwertes (N=105)

Quelle	Quadratsumme vom Typ III	df	Mittel der Quadrate	F	p-Wert	Partielles Eta-Quadrat
Fachärzt*innengruppe	74.14	3	24.71	1.06	.370	.034
Fachärzt*innenausbildung	176.11	3	58.70	2.52	.063	.077
Zeit seit Studienabschluss	153.00	1	153.00	6.57	.012	.068
Examensnote	253.62	1	253.62	10.89	.001	.108
Fachärzt*innengruppen * Fachärzt*innenausbildung	61.43	6	10.24	0.44	.851	.028

Fragen den hausärztlichen Berufsalltag nicht gut widerspiegeln [16].

Die „Weiß nicht“-Option könnte ebenso zu einer Verzerrung der Ergebnisse geführt haben. Diese Antwortoption ist im Progress Text Medizin enthalten, um ein Raten von Antworten zu verhindern und echte klinische Praxis zu simulieren [17]. Im Rahmen der vorliegenden Studie machten 36 Teilnehmende (22%) nie von der „Weiß nicht“-Option Gebrauch. Eventuell kam die „Weiß nicht“-Option für sie nicht in Betracht, da es sich um eine formative Prüfung handelte (d. h., ohne echte Entscheidungssituationen mit möglicherweise bedrohlichen Konsequenzen für die Patient*innen) und damit zu mutigeren Entscheidungen auf Seiten der Teilnehmenden führte als dies in Situationen mit echten Patienten der Fall gewesen wäre.

Im Gegensatz zu Ramsey [6] und Norcini [19], stellte die Studie von Carline [18] keinen signifikanten Zusammenhang zwischen der Zeit seit der Zertifizierung als Internist*innen und der Zeit seit Absolvierung einer Rezertifizierungsprüfung für Internist*innen fest [18]. Allerdings hatten lediglich Ärzt*innen, die ihre erste Zertifizierungsprüfung vor 5 bis 10 Jahren absolvierten, an der Studie teilgenommen. Aufgrund der genannten Limitationen könnte den Studienergebnissen von Ramsey (1991) und Norcini (1985) eine größere Bedeutung zukommen [6], [16]. Ferner zeigen die Ergebnisse einer mehrstufigen Studie aus Kanada, die unterschiedliche Methoden zur Bewertung von Hausärzt*innen verglich, dass eine negative, signifikante Korrelation zwischen medizinischer Kompetenz sowie dem Alter und der Zeit seit Studienabschluss besteht [20].

Examensnote im 3. Abschnitt der Ärztlichen Prüfung

In der Literatur nimmt die Bedeutung von Examensnoten für weiterführende Abschlüsse einen großen Raum ein. Ramsey zeigte, dass Ärzt*innen mit einer besseren Examensnote auch bei der Fachärzt*innenprüfung besser abschnitten [6]. Die Ergebnisse einer deutschen Studie deuten darauf hin, dass das Abschneiden beim internistischen Teil des Progress Tests Medizin im Jahr 2010 signifikant mit der Abiturnote und der Note im 1. Abschnitt der Ärztlichen Prüfung zusammenhing [21]. Eine ähnliche Korrelation zwischen der Abiturnote und den Prüfungen im Verlauf des Studiums wurde durch zahlreiche weitere Studien belegt [22], [23], [24].

Der Zusammenhang zwischen dem Wissensstand und dem Bestehen einer Rezertifizierungsprüfung für Fachärzt*innen sowie weiteren Prüfungen vor Aufnahme einer praktischen medizinischen Tätigkeit wird von mehreren Studien beschrieben. Zusätzlich zu den Schulnoten entdeckten weitere Studien Einflüsse durch das Herkunftsland, das Studienfach, die Zeitspanne zwischen Schul- und Studienleistung sowie den Studienabschnitt der Studienleistung als zuverlässige Indikatoren [25], [26]. Die erzielte Examensnote ist offenbar ein guter Prädiktor für die spätere Leistung bei standardisierten, wissensbasierten Tests wie dem für diese Studie verwendeten Test. Im ANCOVA-Modell hatte die Examensnote den größten Einfluss auf das Abschneiden beim PTM-GP und erklärte 10.8% der Varianz.

Erste berufliche Station

Die Ergebnisse zeigten, dass Ärzt*innen, die zu Beginn ihrer beruflichen Laufbahn ihre Fachärzt*innenausbildung an einem Universitätsklinikum absolvierten, signifikant bessere Ergebnisse im Wissenstest erzielten als Ärzt*innen in nicht-akademischen Krankenhäusern. Die Art der ersten Weiterbildungseinrichtung wiederum korrelierte mit dem Wert des Testergebnisses sowie mit dem Wert innerhalb der einzelnen Gruppen. Der Umfang der Angliederung an eine Universität im Rahmen der Erstanstellung erklärt 7.7% der Varianz beim Abschneiden im PTM-GP, wenn der beste Prädiktor „Examensnote“ im ANCOVA-Modell aufgrund der vielen fehlenden Datenpunkte nicht berücksichtigt wird. Es kann davon ausgegangen werden, dass Ärzt*innen an Universitätskliniken eher in Forschung und Lehre eingebunden sind und dass ihr Wissen daher verlässlicher auf den neuesten Stand gebracht wird, was sich wiederum auch auf die Beantwortung von Wissensfragen in entsprechenden Tests auswirkt. Aufgrund der großen Teams an Universitätskliniken sind jüngere Ärzt*innen dort mehr gefordert, sich gegenüber ihren Kolleg*innen zu behaupten, und werden in hierarchischen Gefügen öfter geprüft. Die häufigeren „Prüfungssituationen“ schulen rasch erlernbares Wissen und führen so eventuell zu besseren Testergebnissen. In diesem Zusammenhang bleibt offen, ob die unterschiedlichen Fort- und

Weiterbildungsmaßnahmen, die von den Einrichtungen angeboten werden, einen Einfluss auf den Wissensstand haben. Eine Studie von Lehmann und Schultz ergab keinen signifikanten Unterschied im Weiterbildungsverhalten zwischen Ärzt*innen, die an Universitätskliniken arbeiten, und solchen, die in der dezentralen medizinischen Versorgung tätig waren [27]. An Universitätskliniken dauert es länger, bis man den Status einer Fachärztin/eines Facharztes erhält als an Krankenhäusern der Regelversorgung, was unter anderem auf den höheren Zeitaufwand für Forschung und Lehre zurückgeführt wird [28]. Haffner und Schmidt fordern die Einführung einer Weiterbildungsordnung, die auf den realitätsnahen Kompetenzen zukünftiger Hausärzt*innen basiert und den Erwerb medizinischen Wissens fördert. Nur so kann ihrer Meinung nach der ambulante Bereich gestärkt werden [29].

Fachärzt*innenstatus

Fachärzt*innen für Innere Medizin erzielten durchschnittlich mehr Punkte beim Wissenstest als Hausärzt*innen. Ein Grund dafür könnte darin liegen, dass Hausärzt*innen zwar breitgefächerte Routineuntersuchungen durchführen, jedoch weniger gut auf die speziellen theoretischen Fragen des Wissenstests vorbereitet sind. Ein weiterer Grund kann ein Effekt durch altersbedingte Störgrößen sein, der auf die Zusammensetzung der Stichprobe zurückzuführen ist, da mehr als 60% der Praktischen Ärzt*innen zu den älteren Teilnehmenden der Stichprobe zählen (Median: 52 Jahre), während mehr als 60% der Internist*innen zu den jüngeren Teilnehmenden gehören. Im ANCOVA-Modell erklärte die Zeit seit Studienabschluss bzw. das umgekehrt proportionale Alter der Teilnehmenden signifikante 6.8% der Varianz. Das Ergebnis der Studie von Ramsey war eher gegenläufig zu den Ergebnissen der vorliegenden Studie [6]. Den Ergebnissen von Ramsey zufolge schnitten gut ausgebildete Generalisten besser ab als spezialisierte Kardiologen und Gastroenterologen. Dieser Widerspruch ist nicht zuletzt in der unterschiedlichen Fachärzt*innenausbildung in Deutschland und den USA begründet.

Limitationen der vorliegenden Untersuchung

Während ein Vergleich mit Daten der Bundesärztekammer [30] deutschlandweit eine ähnliche Verteilung für Hausärzt*innen hinsichtlich Geschlecht, Alter und Fachärzt*innenausbildung ergibt, könnte die Rekrutierung der Studienteilnehmenden im Rahmen von Qualitätszirkeln und Tagungen zu einer Auswahlverzerrung geführt haben, da davon ausgegangen werden kann, dass besonders motivierte und interessierte Hausärzt*innen an solchen Veranstaltungen teilnehmen. Obwohl alle Hausärzt*innen zur Fortbildung und zum Sammeln von CME-Punkten verpflichtet sind, kann eine Verzerrung der Studienergebnisse nicht ausgeschlossen werden. Das gewählte Studiendesign ließ eine randomisierte Zufallsstichprobe nicht zu. Mögliche Effekte durch die Auswahl der Teilnehmenden können daher nicht ausgeschlossen

werden. Der Motivationsfaktor gewinnt zusätzlich dadurch an Gewicht, dass die Teilnehmenden relativ viel Zeit für die Beantwortung des Fragebogens mit einem Umfang von sieben Seiten und 60 Fragen aufwenden mussten. Die Teilnahme an der Studie war freiwillig. Gegenleistungen oder Sanktionen waren nicht zu erwarten. Ein Anreiz bestand möglicherweise in dem Angebot, eine Rückmeldung zum eigenen Wissenstand zu erhalten. Dieses Feedback ist jedoch nicht vergleichbar mit der summativen Evaluation einer wissenschaftlichen und/oder beruflichen Qualifikation, sodass eine geringere Motivation der Teilnehmenden die Studienergebnisse beeinflusst haben könnte. Die Studie von Ramsey stellte bei einem Vergleich der Ergebnisse von freiwilligen Studienteilnehmenden mit den Ergebnissen von Ärzt*innen, die sich einer Rezertifizierung unterziehen mussten und deutlich erfolgreicher abschnitten, eine ähnliche Limitation fest [6]. Eine Studie, die sich mit nicht bestehensrelevanten formativen Tests beschäftigte, unterstützt die These, dass externe motivierende oder demotivierende Faktoren möglicherweise eine Rolle beim Abschneiden der Teilnehmenden in der vorliegenden Studie spielten [31].

Schlussfolgerung und Ausblick

Studienergebnisse aus den USA [6] konnten in Deutschland und Österreich wiederholt werden, auch wenn die Ausbildung und die berufliche Laufbahn von Hausärzt*innen in den USA und in Deutschland nicht direkt vergleichbar sind. Die Abnahme medizinischen Fachwissens im zeitlichen Verlauf scheint relativ unabhängig vom jeweiligen Gesundheitssystem zu sein.

Die vorliegende Untersuchung konnte das Abschneiden von Hausärzt*innen bei einem standardisierten Progress Test bewerten und unterschiedliche Prädiktoren für deren Prüfungsleistung identifizieren. Insgesamt vier Indikatoren waren schwach bis moderat signifikant für das Abschneiden beim PTM-GP:

1. Die Zeit seit Studienabschluss: Hier bestand eine schwach negative Korrelation zwischen der Zeit seit Studienabschluss und dem Abschneiden beim Wissenstest.
2. Die Note im 3. Abschnitt der Ärztlichen Prüfung: Die Studienergebnisse zeigen eine moderate Korrelation zwischen einer guten Endnote und einem guten Ergebnis beim Wissenstest. Die ANCOVA ergab, dass von den betrachteten soziodemografischen Merkmalen die Examensnote den größten Einfluss auf das Abschneiden beim Test hatte.
3. Die Art des Krankenhauses, an dem die Fachärzt*innenausbildung absolviert wurde: Ärzt*innen, die sich zu Beginn ihrer beruflichen Laufbahn für eine Fachärzt*innenausbildung an einem Universitätsklinikum entschieden, erzielten moderat signifikant bessere Ergebnisse beim Wissenstest als Ärzt*innen an nicht-akademischen Krankenhäusern.
4. Die Tatsache, ob die Teilnehmenden Hausärzt*innen oder Internist*innen sind: Internist*innen erzielten moderat bessere Ergebnisse im Wissenstest als Hausärzt*innen.

Während Faktenwissen eine äußerst wichtige Ressource im Berufsleben einer Hausärztin/eines Hausarztes ist, können erfahrungsbasierte und praktische Kompetenzen sowie die Aufmerksamkeit gegenüber Patient*innen gleichermaßen wichtig sein. Um sicherzustellen, dass medizinisches Fachwissen während der gesamten Berufslaufbahn auf höchstmöglichem Niveau verfügbar ist, sollten entsprechende Fortbildungsmaßnahmen in den nicht-akademischen Berufsalltag von Hausärzt*innen zeitgemäß integriert werden, beispielsweise durch wiederholte Progress Tests mit einer Rückmeldung zum persönlichen Wissensprofil für eine bessere Planbarkeit der eigenen medizinischen Fortbildung.

Anmerkung

Wir beschlossen, die 58 Teilnehmenden, die angaben, aktuell hauptberuflich in der stationären Versorgung tätig zu sein, nicht von der Studie auszuschließen, da ihre Teilnahme an Qualitätszirkeln für Hausärzt*innen bzw. an der Jahrestagung der Deutschen Gesellschaft für Allgemeinmedizin und Familienmedizin (DEGAM) ein allgemeines berufliches Interesse an der medizinischen Grundversorgung nahelegt und damit vermuten lässt, dass sie dieselben Patient*innenkohorten behandeln, auch wenn das Arbeitsumfeld ein anderes ist.

Danksagung

Die Autor*innen bedanken sich bei allen Hausärzt*innen, die an dieser Untersuchung teilnahmen, und wir danken dem Präsidenten der DEGAM-Tagung in Salzburg, Herrn Univ.-Prof. Dr. Andreas Sönnichsen, herzlich für seine Unterstützung bei der Umsetzung dieses Forschungsvorhabens. Ferner möchten wir uns bei Herrn Prof. Dr. med. Jörg Schelling und Herrn Matthias Holzer für die Unterstützung unserer Studie bedanken.

Interessenkonflikt

Die Autor*innen erklären, dass sie keinen Interessenkonflikt im Zusammenhang mit diesem Artikel haben.

Anhänge

Verfügbar unter

<https://www.egms.de/de/journals/zma/2020-37/zma001342.shtml>

1. Anhang_1.pdf (644 KB)

Ergänzender Fragebogen zum Progress Test Medizin

Literatur

1. Woods N, Brooks L, Norman G. The role of biomedical knowledge in diagnosis of difficult clinical cases. *Adv Health Sci Educ Theory Pract.* 2007;12(4):417-426. DOI: 10.1007/s10459-006-9054-y
2. Linden M, Gothe H, Ormel J. Der Hausarzt als Gatekeeper. *Dtsch Arztebl.* 2004;101(39):2103-2110.
3. Hall J, Platell C. Half-life of truth in surgical literature. *Lancet.* 1997;350(9093):1752. DOI: 10.1016/S0140-6736(05)63577-5
4. Masic I, Novo A, Delkovic S, Onerhodzic I, Piralic A. How to assess and improve quality of medical education: lessons learned from Faculty of Medicine in Sarajevo. *Bosn J Basic Med Sci.* 2007;7(1):74-78. DOI: 10.17305/bjbm.2007.3097
5. Mantovani F, Castelnuovo G, Gaggioli A, Riva G. Virtual reality training for health-care professionals. *Cyberpsychol Behav.* 2003;6(4):389-395. DOI: 10.1089/109493103322278772
6. Ramsey PG, Carline JD, Inui TS, Larson EB, LoGerfo JP, Norcini JJ, Wenrich MD. Changes over time in the knowledge base of practicing internists. *JAMA.* 1991;266(8):1103-1107. DOI: 10.1001/jama.1991.03470080073032
7. Faul F, Erdfelder E, Lang AG, Buchner A. G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods.* 2007;39(2):175-191. DOI: 10.3758/BF03193146
8. Nouns ZM, Georg W. Progress testing in German speaking countries. *Med Teach.* 2010;32(6):467-470. DOI: 10.3109/0142159X.2010.485656
9. Osterberg K, Kölbel S, Brauns K. Der Progress Test Medizin: Erfahrungen an der Charité Berlin. *GMS Z Med Ausbild.* 2006;23(3):Doc46. Zugänglich unter/available from: <https://www.egms.de/static/de/journals/zma/2006-23/zma000265.shtml>
10. Schaubert SK, Hecht M, Nouns ZM, Kuhlmeier A, Dettmer S. The role of environmental and individual characteristics in the development of student achievement: a comparison between a traditional and a problem-based-learning curriculum. *Adv Health Sci Educ Theory Pract.* 2015;20:1033-1052. DOI: 10.1007/s10459-015-9584-2
11. Möltner A, Schellberg D, Jünger J. Grundlegende quantitative Analysen medizinischer Prüfungen. *GMS Z Med Ausbild.* 2006;23(3):Doc53. Zugänglich unter/available from: <https://www.egms.de/static/de/journals/zma/2006-23/zma000272.shtml>
12. Lienert GA, Raatz U. Testaufbau und Testanalyse. 6. Auflage. Weinheim: Psychologie Verlags Union; 1998.
13. Cohen J. Statistical power analysis for the behavioral sciences. Hillsdale: Erlbaum; 1988.
14. Fritz CO, Morris PE, Richler JJ. Effect size estimates: Current use, calculations, and interpretation. *J Experim Psychol Gen.* 2012;141(1):30. DOI: 10.1037/a0026092
15. Lenhard W, Lenhard A. Berechnung von Effektstärken. Dettelbach: Psychometrica; 2016. Zugänglich unter/available from: <https://www.psychometrica.de/effektstaerke.html>
16. Fleming AW. Examination scores fall with time: but so what? *JAMA.* 1992;267(3):366. DOI: 10.1001/jama.1992.03480030044031
17. Ravestloot CJ, van der Schaaf MF, Muijtjens AM, Haaring C, Kruitwagen CJ, Beek FJ, Bakker J, van Schaik JP, Ten Cate TJ. The don't know option in progress testing. *Adv Health Sci Educ Theory Pract.* 2015;20(5):1325-1338. DOI: 10.1007/s10459-015-9604-2
18. Carline J, Inui T, Larson E, Logerfo J, Ramsey P. The knowledge base of certified internists: relationships to training, practice type, and other physician characteristics. *Arch Intern Med.* 1989;149(10):2311-2313. DOI: 10.1001/archinte.1989.00390100115025
19. Norcini J, Lipner R, Benson J, Webster G. An analysis of the knowledge base of practicing internists as measured by the 1980 recertification examination. *Ann Intern Med.* 1985;102(3):385-389. DOI: 10.7326/0003-4819-102-3-385
20. Norman G, Davis D, Lamb S, Hanna E, Caulford P, Kaigas T. Competency assessment of primary care physicians as part of a peer review program. *JAMA.* 1993;270(9):1046-1051. DOI: 10.1001/jama.1993.03510090030007
21. Schmidmaier R, Holzer M, Angstwurm M, Nouns Z, Reincke M, Fischer MR. Using the Progress-Test Medizin (PTM) for evaluation of the Medical Curriculum Munich (MeCuM). *GMS Z Med Ausbild.* 2010;27(5):Doc70. DOI: 10.3205/zma000707
22. Dlugosch S. Prognose von Studienerfolg: dargestellt am Beispiel des Auswahlverfahrens der Bucerius Law School. Aachen: Shaker; 2005.
23. Gentsch S. Richtig ausgewählt? Eine Evaluation neuer Verfahren der Studierendenauswahl in den Fächern Medizin und Pharmazie an der Goethe-Universität. Berlin: Logos-Verlag; 2009.
24. Schiefele U, Streblow L, Ermgassen U, Moschner B. Lernmotivation und Lernstrategien als Bedingungen der Studienleistung. Ergebnisse einer Längsschnittstudie. *Z Päd Psychol.* 2003;17:185-198. DOI: 10.1024//1010-0652.17.34.185
25. Tamblyn R, Abrahamowicz M, Brailovsky C, Grand Maison P, Lesco J, Norcini J, Girard N, Haggerty J. Association between licensing examination scores and resource use and quality of care in primary care practice. *JAMA.* 1998;280(11):989-996. DOI: 10.1001/jama.280.11.989
26. Trapmann S, Hell B, Weigand S, Schuler H. Die Validität von Schulnoten zur Vorhersage des Studienerfolgs-eine Metaanalyse. *Z Päd Psychol.* 2007;21:11-27. DOI: 10.1024/1010-0652.21.1.11
27. Lehmann K, Schultz J. Zur Lage der anästhesiologischen Weiter- und Fortbildung in Deutschland Ergebnisse einer Repräsentativumfrage. *Anaesthesist.* 2001;50:248-261. DOI: 10.1007/s001010170028
28. Ansorg J, Hassan I, Fendrich V, Polonius M, Rothmund M, Langer P. Qualität der chirurgischen Weiterbildung in Deutschland. *Dtsch Med Wochenschr.* 2005;130:508-513. DOI: 10.1055/s-2005-863084
29. Haffner C, Schmidt M. Weiterbildung Allgemeinmedizin: Sind Hausärzte wirklich noch erwünscht? *Dtsch Arztebl.* 2008;105(30):A-1635/B-1411/C-1379.
30. Bundesärztekammer. Statistics at a glance. Berlin: Bundesärztekammer; 2011. Zugänglich unter/available at: <https://www.bundesaerztekammer.de/ueber-uns/aerzttestatistik/>
31. Silm G, Must O, Taeht K. Test-taking effort as a predictor of performance in low-stakes tests. *J Human Soc Sci.* 2013;17(4):433-448. DOI: 10.3176/tr.2013.4.08

Korrespondenzadresse:

Prof. Dr. phil. Michaela Zupanic
 Universität Witten/Herdecke, Fakultät für Gesundheit,
 Interprofessionelle und Kollaborative Didaktik in Medizin-
 und Gesundheitsberufen, Alfred-Herrhausen-Str. 50,
 58448 Witten, Deutschland
 Michaela.Zupanic@uni-wh.de

Bitte zitieren als

Zupanic M, Kreuer J, Bauer D, Nouns ZM, Ehlers JP, Fischer MR.
Spontaneously retrievable knowledge of German general practitioners
depending on time since graduation, measured with the progress test
medicine. *GMS J Med Educ.* 2020;37(5):Doc49.
DOI: 10.3205/zma001342, URN: urn:nbn:de:0183-zma0013421

Artikel online frei zugänglich unter

<https://www.e-gms.de/en/journals/zma/2020-37/zma001342.shtml>

Eingereicht: 04.02.2020

Überarbeitet: 23.05.2020

Angenommen: 30.06.2020

Veröffentlicht: 15.09.2020

Copyright

©2020 Zupanic et al. Dieser Artikel ist ein Open-Access-Artikel und steht unter den Lizenzbedingungen der Creative Commons Attribution 4.0 License (Namensnennung). Lizenz-Angaben siehe <http://creativecommons.org/licenses/by/4.0/>.