

Poor Performance Among Trainees in a Dutch Postgraduate GP Training Program

Margit I. Vermeulen, MD, PhD; Marijke M. Kuyvenhoven, PhD; Esther de Groot, PhD; Nicolaas PA Zuithoff, PhD; Honore M. Pieters, MD, PhD; Yolanda van der Graaf, MD, PhD; Roger AMJ Damoiseaux, MD, PhD

BACKGROUND AND OBJECTIVES: Poor performance among trainees is an important issue, for patient safety and economic reasons. While early identification might enhance remediation measures, we explored the frequency, nature, and risk factors of poor performance in a Dutch postgraduate general practitioner (GP) training program.

METHODS: All trainees who started the GP training between 2005 and 2007 were included. Multivariate logistic regression analysis was applied to examine associations between individual characteristics; early assessments of competencies and knowledge, training process characteristics (eg, illness, maternal leave), and the outcome poor performance; sub-analyses were performed for each year.

RESULTS: A total of 215 trainees started the 3-year GP program, and 49 (22.8%) exhibited poor performance (in one or more years). In the first and second years, problem areas among poor performers were equally distributed across the roles of “medical expert,” “communicator,” and “professional.” In the third year, shortcomings in “professionalism” were the most common problem. Increasing age was a risk factor for poor performance as were insufficient scores in communication and knowledge. Poor performance in the previous year was a risk factor for poor performance in the second and third years; OR=4.20 (CI=1.31–13.47) and OR=5.40 (CI=1.58–18.47), respectively.

CONCLUSIONS: Poor performance is prevalent but primarily occurring within a single training year. This finding suggests that trainees are capable of solving trainee problems. Increasing age, insufficient assessment scores early in the training, and poor performance in a previous year constitute risk factors for poor performance.

(Fam Med 2016;48(6):430-8.)

Poor performance might be associated with poor professional performance after the training program, with a potential effect on patient safety.^{1,6} Postgraduate medical training should ensure that medical doctors have satisfactory competencies. Persistent poor performance inevitably has to lead to dismissal. This involuntary attrition, together with voluntary, is a common problem in postgraduate medical training programs leading to economic loss and potential shortage of specialists. Rates are varying depending on specialty, eg, in the United States, attrition rates range from 3.3% in psychiatry to 8.6% in family medicine and almost 30% in some surgery programs.⁷⁻¹⁴ In The Netherlands, approximately 10% of internal medicine and surgery trainees withdraw from their programs prematurely.¹⁵

Poor performance is defined as “having significant problems with knowledge, attitudes, or skills that require monitoring and remediation.”¹⁶⁻¹⁹ Poor performance often creates a significant burden for the trainee, trainers, and staff members.^{1,4,20} Trainers and staff members are known to be reluctant to

From the University Medical Center Utrecht, Julius Center for Health Science and Primary Care, The Netherlands.

confront trainees with their shortcomings and, if necessary, impose consequences.¹⁶ Therefore, involuntary attrition rate might even underreport the number of trainees exhibiting poor performance. Early detection of poor performance could be fruitful, as it would provide remediation efforts with a greater chance to succeed.^{10,13,18,19,21-23} From the perspective of patient safety, societal, and personal cost-benefit balances, we explored the frequency and nature of poor performance in a Dutch postgraduate GP training program in order to identify possible risk factors of “trainees at risk” early in training.^{1-6,24}

The following research questions were posed:

- How many trainees exhibit poor performance during the 3-year program, and in which competency roles do shortcomings appear?²⁵
- To what degree are individual characteristics, early competency/knowledge assessments, and training process characteristics associated with poor performance in each year and in the total program?

Methods

The Utrecht Trainee Competence Study in General Practice (GP-UTCS) is a retrospective observational cohort study of all trainees who started the GP training in Utrecht between March 1, 2005, and September 1, 2007. At the time of the study, ethical approval was not mandatory for routinely gathered data. We executed the study according to “the code of conduct” for the use of personal data in scientific research.²⁶ Prior to data processing, all data were anonymized.

Postgraduate GP Training

The Dutch postgraduate GP training takes 3 years. The first and third years provide practical training in general practice under the supervision of an experienced GP (the “trainer”). The second year is dedicated to hospital rotations, primarily in the emergency department (6 months), nursing homes (3 months),

and mental health institutions (3 months). Dispensation is offered for one or more rotations during the second year for those trainees who have relevant previous experience in authorized institutions, so the duration of the second training year may be 0, 3, 6, 9, or 12 months. Each week, the trainee attends a 1-day tutorial in a small group of 12 trainees, with two staff members (a GP and a psychologist) providing training in theoretical, practical, and reflective skills.

Every 3 months, each trainee’s performance is evaluated according to the National Assessment Protocol by his or her trainer or hospital supervisor and staff members in an evaluation session; beforehand the trainer and staff members independently assess and score the competency roles on the Compass, a validated Competence Scoring list with the seven CanMEDS roles.²⁷ In addition, twice a year, the trainees’ knowledge progress is assessed via a National GP Knowledge Test.²⁸

At the end of each year, relying on all assessments, the head of the training department decides whether a trainee can continue with the program (“go”) or whether certain conditions must be met before he/she is allowed to continue (“go, unless”). These conditions are individualized requirements designed to help trainees improve in particular problem areas, eg, additional time in the program, further competency or knowledge training and assessment, or an extra rotation in another clinical setting. If a trainee does not meet these conditions, he/she is dismissed (“no go”).

Data Collection

All data were derived from trainee dossiers.

Individual Characteristics. Gender, age (in years) at the time of entry into the program, the region where medical school was completed (NW Europe, elsewhere), past performance—clinical experience as a medical doctor (< 1 year or ≥ 1 year), the number of times the trainee applied

to the training program (once, more than once), the duration of dispensation (0, 3, 6, or ≥ 9 months) and the application score on a 3-point scale—below standard (1), standard (2), and above standard (3), as evaluated by a selection committee with respect to motivation, orientation on the job, learning needs, and personal attributes.²⁹

Early Assessments. Evaluation scores of the competency roles (“medical expert,” “communicator,” and “professional,” the main focus in this period) on a 4-point scale (1=poor, 2=insufficient, 3=sufficient, and 4=good) after the first 3 months. Result of the National GP Knowledge Test (insufficient or sufficient) in the second month.²⁸

Training Process Characteristics. These time-dependent variables may vary per year: the percentage of time the trainee spent on weekly employment (< 85%, 85%–95%, > 95%), the duration of any recorded illnesses (longer than 2 weeks), the frequency of maternity leave, and the quality of the trainee’s performance during the previous year (“go,” “go, unless”).

Outcome Measure. Poor performance (“go, unless” or “no go”) assessed by the head of the department at the end of each year, based on the competency and knowledge assessments of the trainers and staff members. In the evaluation report the nature of the shortcomings of the poor performer was described, classified in the competency roles “medical expert” (including medical knowledge), “communicator,” “collaborator,” “manager,” “health advocate,” “scholar,” and “professional.”^{25,27}

Analysis

First, trainees’ individual characteristics and early competency/knowledge assessment scores were collected. Next, the course of the cohort and training process characteristics were determined. The associations between individual

characteristics, early assessments, training process characteristics, and outcome were estimated as follows: per category multivariate logistic regression analysis (odds ratios, 95% confidence interval) was performed with all the determinants of this category and poor performance as outcome. Per step the determinant with the highest *P* value was removed until determinants with a *P* value <0.10 remained (stepwise backward analysis). The final analysis was done similarly with the “remaining-determinants per category” and with gender and age. Using a comparable approach, analyses of the data of the total program were conducted. We applied multiple imputation for missing data (1.3%).³⁰ Analyses were performed using SPSS version 20.

Results

Individual Characteristics

In total, 215 trainees started GP training between March 1, 2005, and September 1, 2007 (Table 1). One third of these were male, and the mean age was 29.5 years (SD

4.2). Most trainees (93%) had completed medical school in northwestern Europe. Almost half of them had worked less than 1 year as a medical doctor at the time of application, and one fourth had applied for the GP training program more than once. The mean application score was 2.4 (SD=0.3) on a 3-point scale. More than one fourth had no dispensation, 13% had 3 months, 43% had 6 months, and the remaining trainees had 9 or 12 months.

Early Assessments. After the first 3 months of training, approximately one fourth of the trainees had received a score of “poor” or “insufficient” with respect to the roles as “medical expert,” “communicator,” or “professional.” Fifteen per cent received the score “insufficient” on the GP knowledge test.

Training Process Characteristics. The percentage of trainees who maintained full-time employment (> 95%) decreased from about 80% in the first year to 70% in the second

year, and 60% in the third year (Table 2). Sixteen per cent had been ill for more than 2 consecutive weeks. Among these individuals, the median duration of illness was 7.5 weeks. During the course of the program, of the 148 female trainees, 47 went on maternity leave once and 19 twice. Most maternity leaves occurred in the third year (42), 18 in the first year, and 25 in the second year.

Poor Performance. At the end of the first year, 21 of the 215 trainees were assessed as exhibiting poor performance, and three of them were eventually dismissed (Figure 1). Four trainees voluntarily left the program after year 1; three of them went to another GP program, and one changed specialization. One trainee received dispensation for all rotations; thus, 207 trainees entered year 2. In that year, 27 trainees exhibited poor performance. Of these, 19 had not experienced problems in the first year. Eight of the 18 trainees who exhibited poor performance in year 1 did so in year 2 as well,

Table 1: Individual Characteristics, Early Competency, and Knowledge Assessments of the Trainees*

Individual Characteristics	
Gender, % male	31.2
Age in years, mean (SD)	29.5 (4.2)
Region medical school, % NW Europe	93.1
Past performance, % < 1 year as MD	48.3
Times of application, % first	76.7
Mean selection score: mean (SD)	2.4 (0.3)
Dispensation, %	
0 month	27.9
3 months	12.6
6 months	43.3
9 months	13.5
12 months	2.8 (n=6**)
Early Competency and Knowledge Assessments, % “Sufficient, Good”	
Medical expert	73.5
Communicator	71.3
Professional	77.5
Knowledge	84.6

* n=215

** n=6: one trainee received 12-month dispensation and could directly pass to year 3

Five trainees (military doctors): 6 months dispensation for year 2; 3 months in year 1; 3 months in year 3

Table 2: Training Process Characteristics and Poor Performance

Training Process Characteristics	Total n=215	Year 1 n=215	Year 2 n=207	Year 3 n=203
Employment rate % < 85%		3.3	3.9	5.4
85%–95%		14.4	23.2	36.5
>95%		82.3	72.9	58.1
Illness periods % none	84.2	95.8	94.2	93.1
Once	14.4	4.2	5.3	6.4
Twice	1.4	0	0.5	0.5
Median duration of illness in weeks	7.5 n=34	10.0 n=9	6.5 n=12	6.0 n=14
Women	n=148	n=148	n=142	n=140
Maternal leave: once	n=47	n=18	n=23	n=40
Twice	n=19	n=0	n=1	n=1
Outcome poor performance %	22.7 n=49	9.7 n=21	13.0 n=27	7.4 n=15

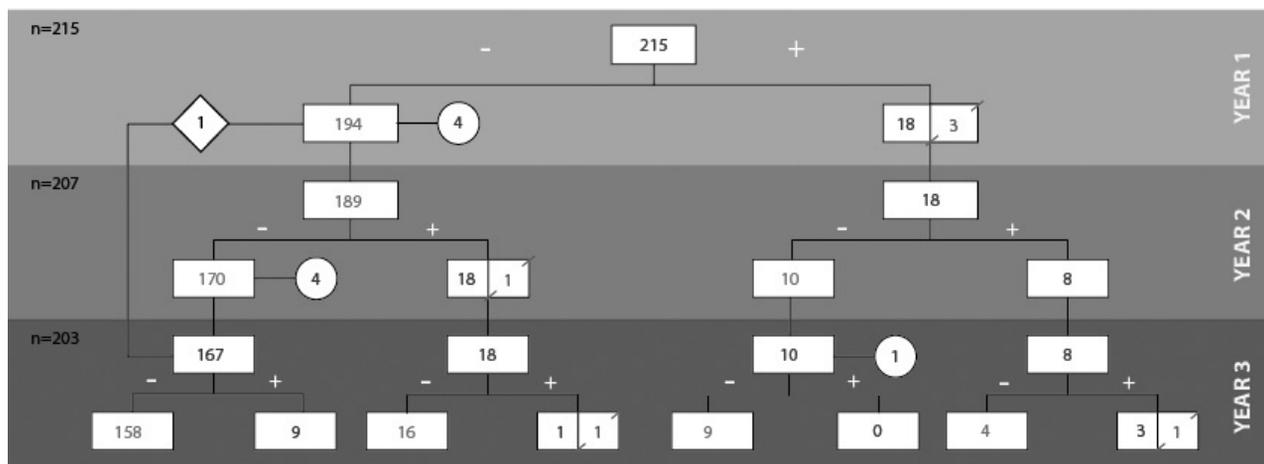
whereas 10 did not. At the end of year 2, one trainee was dismissed, and four trainees left for another GP department. Eventually, 203 trainees entered year 3. Fifteen trainees exhibited poor performance in that year. Nine trainees exhibited problems for the first time, two trainees for the second time, and four for the third time. Two trainees were dismissed in the last year of the program, and one trainee left the program for another GP department.

The involuntary attrition rate of this cohort was 2.8%; among trainees who exhibited poor performance, the rate was 12.2%. One additional trainee left voluntarily to another postgraduate program; thus, the total attrition rate was 3.3%.

Overall, 49 trainees exhibited poor performance: 39 in 1 of the 3 years, six in two different years, and four in every year of the program. Trainees exhibiting poor performance primarily demonstrated shortcomings

in the following three roles: “medical expert,” “communicator,” and “professional” (Table 3). In years 1 and 2, trainee shortcomings were split equally among these three roles; in year 3, shortcomings were primarily related to the role of “professional.” The number of shortcomings decreased the longer the poor performer was in training. In year 1, more than half of these trainees had shortcomings in two or more competencies; in year 2, one-third and, in

Figure 1: Flowchart of the Trainees During Their Postgraduate Training (n=215), Poor Performance (+)



At the right side of the figure poor performance is indicated (+). Further explanation in the text.

- ◊ : 12 months dispensation
- : another GP-department
- : attrition



Table 3: Distribution of Shortcomings Per Competency Roles of Poor Performers Per Year

Trainee#	Medical Expert			Communicator			Professional			Remaining Roles			Years
	Year 1	2	3	Year 1	2	3	Year 1	2	3	Year 1	2	3	
001											X		1
003									X				1
006		X			X								1
010	X			X									1
016				X			X						1
021							X						1
023		X	X		X	X		X	X				2
029								X					1
048	X												1
051			X						X				1
056	X	X		X	X		X	X	X				3
059									X				1
062	X												1
063	X			X			X	X	X				3
071		X						X					1
073					X			X					1
076	X	X			X			X		X			2
085	X	X		X			X						2
092								X					1
101		X											1
102		X			X								1
104		X											1
105	X			X			X						1
106					X	X		X	X				2
111									X				1
114									X				1
121	X			X									1
130							X						1
139				X				X	X				3
142				X									1
143								X					1
146				X									1
148								X					1
167								X					1
171								X					1
178					X			X		X			1
185	X												1
189				X			X	X	X				3
190									X				1
195		X											1
200									X				1
205									X				1
210	X			X			X						1
214				X				X					2
215	X	X		X			X						2
216								X					1
222							X			X			1
226									X				1
227								X					1

year 3, one-fourth of the trainees had shortcomings in two or more competencies.

Risk Factors for Poor Performance

Increasing age was a risk factor for poor performance in the first and second years (Table 4); OR=1.14 (CI=1.02–1.26) and OR=1.12 (CI=1.02–1.23), respectively; insufficient scores in “medical expertise” and knowledge in the first year; OR=0.49 (CI=0.29–0.84) and OR=0.12 (CI=0.04–0.35), respectively or “communication” in

the second year; OR=0.47 (CI=0.30–0.75). Poor performance in the first and second years was independently associated with poor performance in years 2 and 3; OR=4.20 (CI=1.31–13.47) and OR=5.40; (CI=1.58–18.47), respectively. In the results of the total program, older trainees ran a greater risk of exhibiting poor performance, OR=1,16 (CI=1,06–1,27). Trainees with sufficient scores in “communication” and knowledge were at lower risk of exhibiting poor performance; OR=0.50

(CI=0.33–0.77) and OR=0.16 (CI=0.07–0.40), respectively.

Discussion

Main Findings

Forty-nine trainees exhibited poor performance, most of them in a single year. Six of the 49 trainees were eventually dismissed. Most shortcomings fell into the roles of “medical expert,” “communicator,” and “professional.” In years 1 and 2, trainee weaknesses were distributed equally across these three competency roles, whereas in year 3,

Table 4: Association Between Individual, Early Assessments and Training Process Characteristics and Poor Performance in Respectively First, Second, Third Year and in Total Program (Multivariate Logistic Regression)

	First Year n=215	Second Year n=207	Third Year n=203	Program n=215
	ORs (95% CI)	ORs (95% CI)	ORs (95% CI)	ORs (95% CI)
Individual Characteristics				
Gender ¹	1.27 (0.39–4.18)	0.82 (0.31–2.21)	3.89 (0.83–18.11)	1.29 (0.56–2.97)
Age (years)	1.14 (1.02–1.26)	1.12 (1.02–1.23)	1.10 (0.98–1.23)	1.16 (1.06–1.27)
Region medical school ²				
Past performance ³				
Times of application ⁴				
Dispensation				
Application score				
Early Assessments				
Medical expert	0.49 (0.29–0.84)			
Communicator		0.47 (0.30–0.75)		0.50 (0.33–0.77)
Professional				
Knowledge ⁵	0.12 (0.04–0.35)			0.16 (0.07–0.40)
Training Process Characteristics				
Employment rate ⁶ year 1				
Illness year 1 (weeks)				
Maternity leave ⁷ year 1				
Poor performance year 1		4.20 (1.31–13.47)		
Employment rate ⁶ year 2				
Illness year 2 (weeks)				
Maternity leave ⁷ year 2				
Poor performance ⁸ year 2			5.40 (1.58–18.47)	
Employment rate ⁶ year 3				

¹ Gender (0: male; 1: female)

² Region medical school (0: NW Europe; 1: elsewhere)

³ Past performance (0: <1 year; 1: ≥ 1 year)

⁴ Times of application (0: 1st time; 1: 2nd/3rd)

⁵ Knowledge (0: insufficient; 1: sufficient)

⁶ Employment rate (0: <85%; 1: 85%–95%; 2: > 95%)

⁷ Maternity leave (0: no; 1: yes)

⁸ Poor performance (0: no; 1: yes)

OR—odds ratio, CI—confidence interval

shortcomings primarily pertained to “professionalism.” In the first and second years, age was a risk factor, just as insufficient scores (as “medical expert” and in knowledge in the first year and as “communicator” in the second year). Poor performance in the previous year was a risk factor for poor performance in years 2 and 3. In the results of the total program increasing age and insufficient scores in “communication” and knowledge were risk factors for poor performance.

Discussion of Main Findings

Twenty-three per cent of all trainees exhibited poor performance in one or more years during their GP training. A comprehensive range of poor performance rates across medical disciplines have been mentioned: low rates (<6%) for psychiatry and geriatric medicine, median rates for internal and family medicine (8%–15%), and higher rates (20%–30%) for surgical training programs.^{1,10,11,31–34} In Canada, the mean rate of “residents in difficulty” across several postgraduate training programs was 3%.²⁵ However, it is difficult to compare these percentages because they are influenced by documentation, identification criteria, and the assessment culture of the department. The 23% in our study seems fairly high and may reflect our specific program with three-monthly performance evaluations. Using the competency framework for evaluation does flag trainees as problematic in an early stage but seems useful as in most cases trainees exhibited poor performance only in 1 year. This could mean that poor performance was self-limiting (trainees just needed more time to come to a satisfactory competency level), the intervention of giving the negative feedback during the evaluation was sufficient for trainees to self-remediate their shortcomings, or a tailored remediation plan had helped to overcome the weaknesses.

From the literature it is known that some performance deficiencies could be resistant to remediation and

remain “chronic.”^{1,17,31} Almost 5% of trainees exhibited poor performance in 2 or 3 years, with shortcomings mostly in the same competency roles. This finding may indicate that remediation efforts were not effective enough, trainee problems were resistant to these efforts, or decisions regarding dismissal were postponed.¹⁶ Although the frequency is low, persistent poor performance is problematic if these trainees are at risk of exhibiting poor performance as future GPs.^{1–5}

In years 1 and 2, approximately 10% of trainees exhibited poor performance, whereas in year 3, this percentage was lower. Reamy (2006) also found that most trainee problems were identified during the first 2 years of a family medicine program.³¹ The changes in setting that occur during training may contribute to this difference as the maturing of the trainees themselves.

Among the trainees in this study, poor performance primarily pertained to shortcomings as a “medical expert,” “communicator,” and “professional.” This finding aligns with previous research that cited insufficient knowledge, poor clinical judgment, poor communication, poor interpersonal skills, and attitudinal problems as reasons for poor performance.^{1,4,10,17,19,31,32} Almost half of the poor performers had shortcomings in two or more competency roles, findings that echo other studies.^{1,4,10,17} The distribution pattern of trainee shortcomings changed over time. During the first 2 years, trainee shortcomings were distributed equally across the roles as “medical expert,” “communicator,” and “professional,” in the third year, however, most shortcomings fell into the role as “professional.” If a trainee exhibited poor performance in 2 or 3 years, the role “professionalism” was always involved. This finding suggests that shortcomings as a professional could be more difficult to remediate.^{1–6} More insight into this problem should be gathered by means of qualitative research.

Risk Factors

Increasing age is a risk factor for poor performance in years 1 and 2 and in the total program. Some of the previous literature on this topic mentioned increasing age as a risk factor for poor performance and attrition, while other studies did not support this relationship.^{4,9,10,13,35,36} Conceivable reasons may be that older trainees have more family responsibilities, possess outdated knowledge, or they may be less resilient in adapting to new clinical situations or academic settings. Unfortunately, we did not have data on performance in possible earlier training programs or work situations as trainees might have exhibited poor performance before.

Early insufficient assessments in the roles as “medical expert” or “communicator” and knowledge were associated with poor performance among trainees. The relation between earlier and later performance within the same training or school mirrors findings of others.^{10,20,21,37–39} This finding suggests the utility of performing early competency and knowledge assessments in a programmatic assessment approach.^{32,40–45} Poor performance in a previous year was a strong risk factor for poor performance in the next one, which corroborates earlier findings.^{9,10,21}

In our study, the application score was not associated with poor performance. This finding aligns with earlier studies that indicate it is difficult to predict performance problems from application data.^{20,46} Likewise, illness was not associated with poor performance, even though health problems, substance abuse, and psychiatric illness are often mentioned in the literature as possible reasons for voluntary and involuntary attrition.^{4,10,25,32} Part-time employment and maternity leave were not associated with poor performance; only programmatic schedulers appear to be disadvantaged by the special circumstances that they require.

Strengths and Limitations

As far as we know, this is the first study that has tracked a cohort of 215 GP trainees over consecutive years. We are aware that many determinants regarding the number of outcomes have been used in the model, but our results were stable; the overall cohort results were corroborated by results for each training year. Using administrative collected data resulted in a limitation. However, the study design did not allow us to collect data regarding additional possible risk factors, such as personality traits, medical school records, or the nature of any recorded illnesses.^{47,48}

We assume that the study's internal validity is satisfactory, due to imputation of missing values.³⁰ As most postgraduate GP programs, national as well as international, are similar in duration (3–4 years), in curriculum design (combination of training in general practice and rotations), and in assessment programs (knowledge and competency assessments), we assume a certain generalizability of this study. By exploring evaluation reports of the poor performers, we have attempted to choose the most valid source of information that offers the least risk of bias with respect to trainee shortcomings.

Conclusions

Poor performance among trainees in a postgraduate GP training program is quite common. Most trainees exhibited poor performance in only 1 year of the program, suggesting that trainees are capable of solving problems. However, increasing age, insufficient early assessments, and poor performance in a previous year represent risk factors for poor performance during postgraduate GP training. Because poor performers are at risk of repeated or ongoing problems, they need extra monitoring and supervision by the trainers and staff.

CORRESPONDING AUTHOR: Address correspondence to Dr Vermeulen, Broederplein 43, Zeist, 3703 CD, The Netherlands. 0031-88-756975. m.i.vermeulen@umcutrecht.nl.

References

- Williams RG, Dunnington GL, Klamen DL. Forecasting residents' performance—partly clouded. *Acad Med* 2005;80(5):415-22.
- Papadakis MA, Hodgson CS, Teherani A, Kohatsu ND. Unprofessional behaviour in medical school is associated with subsequent disciplinary action by a state medical board. *Acad Med* 2004;79(3):244-9.
- Resnick AS, Mullen JL, Kaiser LR, Morris JB. Patterns and predictions of resident misbehaviour—a 10-year retrospective look. *Curr Surg* 2006;63(6):418-25.
- Yao DC, Wright SM. National survey of internal medicine residency program directors regarding problem residents. *JAMA* 2000;284(9):1099-104.
- Tamblyn R, Abrahamowicz M, Dauphinee D, et al. Physician scores on national clinical skills examination as predictors of complaints to medical regulatory authorities. *JAMA* 2007;298(9):993-1001.
- Adams KE, Emmons S, Romm J. How resident unprofessional behavior is identified and managed: a program director survey. *Am J Obstet Gynecol* 2008;198(6):692.e1-692.e5.
- Roback HB, Crowder MK. Psychiatry resident dismissal. A national survey of training programs. *Am J Psychiatry* 1989;146(1):96-8.
- Laufenburg HF, Turkal NW, Baumgardner DJ. Resident attrition from family practice residencies: United States versus international medical graduates. *Fam Med* 1994;26(10):614-7.
- Yeo H, Bucholz E, Ann Sosa J, et al. A national study of attrition in general surgery training: which residents leave and where do they go? *Ann Surg* 2010;252(3):529-34.
- Bergen PC, Littlefield JH, O'Keefe GE, Rege RV, Antony TA, Kim LT. Identification of high risk residents. *J Surg Res* 2000;92(2):239-44.
- Yaghoobian A, Galante J, Kaji A, et al. General surgery resident remediation and attrition. *Arch Surg* 2012;147(9):829-33.
- Longo WE, Seashore J, Duffy A, Udelsman R. Attrition of categorical general surgery residents: results of a 20-year audit. *Am J Surg* 2009;197(6):774-8.
- Sullivan MC, Yeo H, Roman SA, et al. Surgical residency and attrition: defining the individual and programmatic factors predictive of trainee losses. *J Am Coll Surg* 2013;216(3):461-71.
- Dodson TF, Webb ALB. Why do residents leave general surgery? The hidden problem in today's program. *Curr Surg* 2005;62(1):128-31.
- Katzenbauer M. That culture did not suit me. (In Dutch: Die cultuur paste niet bij mij). *Dutch Cont* 2009;64(14):580-3.
- Dudek NL, Marks MB, Regehr G. Failure to fail: the perspectives of clinical supervisors. *Acad Med* 2005;80(10 suppl):S84-S87.
- Williams RG, Roberts NK, Schwind CJ, Dunnington GL. The nature of general surgery resident performance problems. *Surgery* 2009;145(6):651-8.
- Evans DE, Alstead EM, Brown J. Applying your clinical skills to students and trainees in academic difficulty. *Clin Teach* 2010;7(4):230-5.
- Steinert Y. The "problem" learner: whose problem is it? AMEE Guide No. 76. *Med Teach* 2013;35(4):e1035-e1045.
- Brenner AM, Mathai S, Jain S, Mohl PC. Can we predict "problem residents"? *Acad Med* 2010;85(7):1147-51.
- Winston KA, van der Vleuten CPM, Scherpbier AJJA. Prediction and prevention of failure: an early intervention to assist at-risk medical students. *Med Teach* 2014;36(1):25-31.
- Guerrasio J, Garrity M, Aagaard EM. Learner deficits and academic outcomes of medical students, residents, fellows and attending physicians referred to a remediation program, 2006-2012. *Acad Med* 2014;89(2):352-8.
- Yao DC, Wright SM. The challenge of problem residents. *J Gen Intern Med* 2001;16(7):486-92.
- O'Neill LD, Wallstedt B, Eika B, Hartvigsen J. Factors associated with dropout in medical education: a literature review. *Med Educ* 2011;45(5):440-54.
- Zbieranowski I, Takahashi SG, Verma S, Spadafora SM. Remediation of residents in difficulty: a retrospective 10-year review of the experience of a postgraduate board of examiners. *Acad Med* 2013;88(1):1-6.
- Association of Universities The Netherlands. Code of conduct for use of personal data in scientific research (<http://www.vsnun.nl/code-pers-gegevens.html>, in Dutch) 2005. Accessed May 1, 2015.
- Tromp F, Vernooij-Dassen M, Grol R, Kramer A, Bottema B. Assessment of CanMEDS roles in postgraduate training: the validation of the compass. *Patient Educ Couns* 2012;89(1):199-204.
- Van Leeuwen YD, Pollemans MC, Mol SSL, Eekhof JAH, Grol R, Drop MJ. Dutch knowledge test for general practice: issues of validity. *Eur J Gen Pract* 1995;1(3):113-7.
- Vermeulen MI, Kuyvenhoven MM, Zuithoff NPA, Graaf van der Y, Damoiseaux RAMJ. Dutch postgraduate GP selection procedure; reliability of interview assessments. *BMC Fam Pract* 2013;14:43. doi:10.1186/1471-2296-14-43.
- Donders AR, van der Heijden GJ, Stijnen T, Moons KG. Review: a gentle introduction to imputation of missing values. *J Clin Epidemiol* 2006;59(10):1087-91.
- Reamy BV, Harman JH. Residents in trouble: an in-depth assessment of the 25-year experience of a single family medicine residency. *Fam Med* 2006;38(4):252-7.
- Mitchell C, Bhat S, Herbert A, Baker P. Workplace based assessments of junior doctors: do scores predict training difficulties? *Med Educ* 2011;45(12):1190-8.
- Frank JR. ABIM: "The problem resident." The CanMEDS 2005 Physician Competency Framework. Better standards. Better physicians. Better care. Ottawa: The Royal College of Physicians and Surgeons of Canada, 2005.

34. Kuyvenhoven MM, Vermeulen MI, van Campen SM, Schmidt JET. Many trainees drop out. (In Dutch: veel aiossen haken af.) *Med Cont* 2010;65(42):2206-7.
35. Naylor RA, Reisch JS, Valentine RJ. Factors related to attrition in surgery residency based on application data. *Arch Surg* 2008;143(7):647-51.
36. McAlister RP, Andriole DA, Brotherton SE, Jeffe DB. Attrition in residents entering US obstetrics and gynecology residencies: analysis of national GME Census data. *Am J Obstet Gynecol* 2008;199(5):574.e1-6.
37. Van Leeuwen YD, Mol SS, Polleman MC, van der Vleuten CP, Grol R, Drop MJ. Selection for postgraduate training for general practice: role of knowledge tests. *Br J Gen Pract* 1997;47(419):359-62.
38. Yates J. Development of a "toolkit" to identify medical students at risk of failure to thrive on the course: an exploratory retrospective case study. *BMC Med Educ* 2011;11:95. doi:10.1186/1472-6920-11-95.
39. Stegers-Jager KM, Cohen-Schotanus J, Themmen APN. The effect of a short integrated study skills programme for first-year medical students at risk of failure: a randomised controlled trial. *Med Teach* 2013;35(2):120-6.
40. Van der Vleuten CPM, Schuwirth LW. Assessing professional competence: from methods to programmes. *Med Educ* 2005;39(3):309-17.
41. Thundiyil JG, Modica RF, Silvestri S, Papa L. Do United States medical licensing examination (USMLE) scores predict in-training test performance for emergency medicine residents? *J Emerg Med* 2010;38(1):65-9.
42. Shellito JL, Osland JS, Helmer SD, Chang FC. American Board of Surgery: can we identify surgery residency applicants and residents who will pass the examinations on the first attempt? *Am J Surg* 2010;199(2):216-22.
43. Spitzer AB, Gage MJ, Looze CA, Walsh M, Zuckerman JD, Egol KA. Factors associated with successful performance in an orthopaedic surgery residency. *J Bone Joint Surg Am* 2009; 91(11): 2750-5.
44. Perez JA Jr, Greer S. Correlation of United States Medical Licensing Examination and Internal In-Training Examination performance. *Adv Health Sci Educ* 2009;14(5):753-8.
45. Park YS, Riddle J, Tekian A. Validity evidence of resident competency ratings and the identification of problem residents. *Med Educ* 2014;48(6):614-22.
46. Dubovsky SL, Gendel M, Dubovsky AN, Rosse J, Levin R, House R. Do data obtained from admission interview and resident evaluations predict later personal and practice problem. *Acad Psych* 2005;29(5):443-7.
47. Lievens F, Coetsier P, De Fruyt F, De Maeseeneer J. Medical student's personality characteristics and academic performance: a five factor model perspective. *Med Educ* 2002;36(11): 1050-6.
48. Hamdy H, Prasad K, Anderson MB, et al. BEME systematic review: predictive values of measurements obtained in medical schools and future performance in medical practice. *Med Teach* 2006;28(2):103-16.