Promoting high quality research on the Royal College of General Practitioner assessments: a summary of key peer reviewed research on the MRCGP examinations

Over the past two decades, the RCGP has commissioned, supported and encouraged publication of research on the assessments which form part of the MRCGP examination.

The research mission developed in relation to the exam is to support systematic research and evaluation relating to the MRCGP, conducted to the highest standards of quality, ethics and governance, to ensure the validity, reliability and fairness of the exam and to increase our understanding of the determinants of these.

According to a recent systematic review of research on medical specialty certification exams studied according to the Ottawa Quality Criteria, the MRCGP exam was "the most extensively studied specialty certification exam regarding the Ottawa [Quality] Criteria".[1] In addition two papers from those listed below were awarded the RCGP Research Paper of the Year in Medical Education 2019 and 2022.

The MRCGP exam is committed to working with academic teams to conduct high quality research on the assessment. In the first instance interested individuals and team should contact the MRCGP Research and Development Lead via exams@rcgp.org.uk in order to obtain appropriate advice.

Professor Niroshan Siriwardena
Research and Development
Lead for Assessment, 2017 –

Updated 5 October 2023

1. Staudenmann D, Waldner N, Lorwald A, Huwendiek S. Medical specialty certification exams studied according to the Ottawa Quality Criteria: a systematic review. *BMC Med Educ.* 2023;23(1):619.

Papers by category

Differential attainment by ethnicity

- Siriwardena AN, Botan V, Williams N, Emerson K, Kameen F, Pope L, Freeman A, Law GR. Academic performance of ethnic minority versus White doctors in the MRCGP assessment 2016-2021: cross sectional study. BJGP 2023; 73 (729): e284-e293. DOI: 10.3399/BJGP.2022.0474.

- Pattinson J, Blow C, Sinha B, Siriwardena AN. Exploring reasons for differences in performance between UK and international medical graduates in the Membership of the Royal College of General Practitioners Applied Knowledge Test: a cognitive interview study. BMJ Open 2019; 9:e030341. DOI:10.1136/bmjopen-2019-030341.

- Hawthorne K, Roberts C, Atkins S. Sociolinguistic factors affecting performance in the Clinical Skills Assessment of the MRCGP: a mixed-methods approach. BJGP Open. 2017;1(1):bjgpopen17X100713. DOI: 10.3399/bjgpopen17X100713

- Low S, Sheppard ZA, Tomkins S. Identifying factors associated with performance in a mock CSA: perceptions from trainees and trainers from one deanery. Educ Prim Care. 2013;24(6):452-460. DOI: 10.1080/14739879.2013.11494216

- Esmail A, Roberts C. Academic performance of ethnic minority candidates and discrimination in the MRCGP examinations between 2010 and 2012: analysis of data. BMJ. 2013;347:f5662. DOI: 10.1136/bmj.f5662

- Denney ML, Freeman A, Wakeford R. MRCGP CSA: are the examiners biased, favouring their own by sex, ethnicity, and degree source? Br J Gen Pract. 2013;63(616):e718-725.

Differential attainment by disability or neurodiversity

- Botan V, Williams N, Siriwardena AN, Law GR. The effect of specific learning difficulties on GP written and clinical assessments. Medical Education 2023; 57(6): 548-555. DOI: 10.1111/medu.15008.

- Asghar Z, Williams N, Denney, M L, Siriwardena A N. Performance in candidates declaring vs not declaring dyslexia in a licensing clinical exam. Medical Education: 53: . 2019;53:1243–1252. DOI: 10.1111/medu.13953

- Asghar ZB, Siriwardena AN, Elfes C, et al. Performance of candidates disclosing dyslexia with other candidates in a UK medical licensing examination: cross-sectional study. Postgrad Med J. 2018;94(1110):198-203. DOI: 10.1136/postgradmedj-2017-135326

Differential attainment by sex

- Pope L, Hawkridge A, Simpson R. Performance in the MRCGP CSA by candidates' gender: differences according to curriculum area. Educ Prim Care. 2014;25(4):186-193. DOI: 10.1080/14739879.2014.11494276

- Siriwardena AN, Irish B, Asghar ZB, et al. Comparing performance among male and female candidates in sex-specific clinical knowledge in the MRCGP. Br J Gen Pract. 2012;62(599):e446-450. DOI: 10.3399/bjgp12X649142

Validity of assessments

- Botan V, Laparidou D, Phung VH, Cheung P, Freeman A, Wakeford R, Denney ML, Law GR, Siriwardena AN. Examiner perceptions of the MRCGP recorded consultation assessment for general practice licensing during COVID-19: cross-sectional study. BMC Medical Education 2023; 23: 65 DOI: 10.1186/s12909-023-04027-4.

- Botan V, Laparidou D, Phung VH, Cheung P, Freeman A, Wakeford R, Denney M, Law GR, Siriwardena AN. Candidate perceptions of the UK Recorded Consultation Assessment: cross-sectional data linkage study. Educ Prim Care 2022, 33(1):32-40. DOI: 10.1080/14739879.2021.1970630.

- Botan V, Williams N, Law GR, Siriwardena AN. How is performance at selection to general practice related to performance at the endpoint of GP training? Report to Health Education England. University of Lincoln, Lincoln, 2022.

- Wakeford R, Ludka K, Woolf K, McManus IC. Fitness to practise sanctions in UK doctors are predicted by poor performance at MRCGP and MRCP(UK) assessments: data linkage study. BMC Med. 2018;16(1):230. DOI: 10.1186/s12916-018-1214-4

- McLoughlin K, Pope L, Walsh E, Jennings A, Foley T. The MRCGP Clinical Skills Assessment: an integrative review of evidence. Educ Prim Care. 2018;29(3):132-137. DOI: 10.1080/14739879.2018.1427510

- Neden CA, Parkin C, Blow C, Siriwardena AN. Has there been a change in the knowledge of GP registrars between 2011 and 2016 as measured by performance on common items in the Applied Knowledge Test? Educ Prim Care. 2018:1-7. DOI: 10.1080/14739879.2018.1467737

- Bodgener S, Denney M, Howard J. Consistency and reliability of judgements by assessors of case based discussions in general practice specialty training programmes in the United Kingdom. Educ Prim Care. 2017;28(1):45-49

- Naidoo S, Lopes S, Patterson F, Mead HM, MacLeod S. Can colleagues', patients' and supervisors' assessments predict successful completion of postgraduate medical training? Med Educ. 2017;51(4):423-431. DOI: 10.1111/medu.13128

- Wakeford R, Denney M, Ludka-Stempien K, Dacre J, McManus IC. Cross-comparison of MRCGP & MRCP(UK) in a database linkage study of 2,284 candidates taking both examinations: assessment of validity and differential performance by ethnicity. BMC Med Educ. 2015; 15:1.

- Dixon H, Blow C, Milne P, Siriwardena N, Milne H, Elfes C. Quality assurance of the Applied Knowledge Test (AKT) of the MRCGP examination - an immediate post-test questionnaire evaluation of the candidates' views. Educ Prim Care. 2015;26(4):223-232. DOI: 10.1080/14739879.2015.11494346

- McManus IC, Wakeford R. PLAB and UK graduates' performance on MRCP(UK) and MRCGP examinations: data linkage study. BMJ. 2014;348:g2621. DOI:10.1136/bmj.g2621

- Dixon H, Blow C, Milne P, Siriwardena N. Reliability of non-pretested versus pretested questions in the applied knowledge test (AKT) of the MRCGP: evidence of quality assurance. Educ Prim Care. 2014;25(3):149-154. DOI: 10.1080/14739879.2014.11494265

- Ahmed H, Rhydderch M, Matthews P. Do general practice selection scores predict success at MRCGP? An exploratory study. Educ Prim Care. 2012;23(2):95-100. DOI: 10.1080/14739879.2012.11494083

- Siriwardena AN, Dixon H, Blow C, Irish B, Milne P. Performance and views of examiners in the Applied Knowledge Test for the nMRCGP licensing examination. Br J Gen Pract. 2009;59(559):e38-43. DOI:10.3399/bjgp09X395111

- McManus IC. Does performance improve when candidates resit a postgraduate examination? Med Educ. 1992;26(2):157-162. DOI: 10.1111/j.1365-2923.1992.tb00142.x

Workplace based assessment, failure to progress and remediation

- Winter R, Norman RI, Patel R. A qualitative exploration of the lived experience of GP trainees failing to progress in training. Educ Prim Care 2021, 32(1):10-18. DOI:10.1080/14739879.2020.1831970

- Shaw B, Fox J, Brown J, Hart A, Mamelok J. An investigation of factors affecting the outcome of the Clinical Skills Assessment(CSA) in general practice specialty training. Education for Primary Care; 2014, 25; 91-5. DOI: 10.1080/14739879.2014.11494253

- Patterson F, Tavabie A, Denney M, et al. A new competency model for general practice: implications for selection, training, and careers. Br J Gen Pract. 2013;63(610):e331-338.

- Sabey A, Harris M. 'It's the conversation they'll learn from': improving assessments for GP Specialist Trainees in hospital posts. Educ Prim Care. 2012;23(4):263-269. DOI: 10.1080/14739879.2012.11494119

- Sabey A, Harris M. Training in hospitals: what do GP specialist trainees think of workplace-based assessments? Educ Prim Care. 2011;22(2):90-99. DOI:10.1080/14739879.2011.11493974

- Bodgener S & Tavabie A. Is there a value to case-based discussion? Education for Primary Care,2011; 22: 223-228.

- Miller A, Archer J. Impact of workplace based assessment on doctors' education and performance: a systematic review. BMJ. 2010;341:c5064. DOI: 10.1136/bmj.c5064

Role players

- Bell EA, Cleland J, Gambhir N. 'It clarified a lot': GP trainees as peer role players in a formative Clinical Skills Assessment (CSA). Educ Prim Care 2021, 32(2):85-90. DOI:10.1080/14739879.2020.1836521

- Denney M, Wakeford R. Do role-players affect the outcome of a high-stakes postgraduate OSCE, in terms of candidate sex or ethnicity? Results from an analysis of the 52,702 anonymised case scores from one year of the MRCGP CSA. Educ Prim Care. 2016;27(1):39-43. DOI: 10.1080/14739879.2015.1113724

- Russell D, Simpson R, Eden M, Foreman P. The use of child actors to simulate paediatric patients in the MRCGP CSA: developing an effective model. Educ Prim Care. 2015;26(4):242-247. DOI: 10.1080/14739879.2015.11494349

- Russell D, Etherington C, Hawthorne K. How can simulated patients' experiences suggest ways to improve candidate performance in the MRCGP clinical assessments? Educ Prim Care. 2012;23(6):391-398.

Papers by date (most recent first)

Siriwardena AN, Botan V, Williams N, Emerson K, Kameen F, Pope L, Freeman A, Law GR. Academic performance of ethnic minority versus White doctors in the MRCGP assessment 2016-2021: cross sectional study. BJGP 2023; 73 (729): e284-e293. DOI: 10.3399/BJGP.2022.0474.

What this study tells us

- This study examined differential attainment in all components of GP licensing assessments, including the Workplace-Based Assessment, considering scores at selection to GP specialty training.
- Multi-Specialty Recruitment Assessment [MSRA] scores were the strongest predictor of success or failure in all assessments.
- Ethnic background did not reduce the chance of passing GP licensing tests once sex, country of primary medical qualification, declared disability, and selection (MSRA) scores were considered.

What this means

- It has been suggested that subjective bias due to racial discrimination in clinical skills assessments may be a cause of examination failure for UK-trained ethnic minority candidates and international medical graduates, but this study showed that this was unlikely to be the case.
- Doctors admitted to GP specialty training, who are in the lowest MSRA score bands, may need additional support during training to maximise their chances of achieving licensing, regardless of their ethnic group or other demographic characteristics.

Botan V, Williams N, Siriwardena AN, Law GR. The effect of specific learning difficulties on GP written and clinical assessments. Medical Education 2023; 57(6): 548-555. DOI: 10.1111/medu.15008.

What this study tells us:

- Candidates declaring a specific learning difficulty (SpLD) were significantly less likely to pass the Clinical Skills Assessment and Workplace Based Assessment but not the Applied Knowledge Test or Recorded Consultation Assessment.

What this means:

- Candidates with SpLDs encounter difficulties in multiple domains of the licensing and in-training assessments, suggesting that adjustments tailored to their needs should be put in place for the applied clinical skills tests and during training.

Botan V, Laparidou D, Phung VH, Cheung P, Freeman A, Wakeford R, Denney ML, Law GR, Siriwardena AN. Examiner perceptions of the MRCGP recorded consultation assessment for general practice licensing during COVID-19: cross-sectional study. BMC Medical Education 2023; 23: 65 DOI: 10.1186/s12909-023-04027-4.

What this study tells us:

- The Recorded Consultation Assessment (RCA) was introduced rapidly during the COVID-19 pandemic to enable doctors to be undertake a clinical licensing assessment.
- The RCA was considered by examiners to be feasible and broadly acceptable, although they experienced challenges from candidate case selection, case content and judgments leading to suggested areas for improvement.

What this means:

- The RCA was considered by examiners to be feasible and broadly acceptable with some challenges and suggestions for improvement.

Botan V, Laparidou D, Phung VH, Cheung P, Freeman A, Wakeford R, Denney M, Law GR, Siriwardena AN. *Candidate perceptions of the UK Recorded Consultation Assessment: cross-sectional data linkage study*. Educ Prim Care 2022, 33(1):32-40. DOI: 10.1080/14739879.2021.1970630.

What this study tells us:

- The RCA, was broadly acceptable and a feasible alternative to the Clinical Skills Assessment (CSA).
- Candidates were positive about the resources provided and the online platform, but less positive about the time they had to prepare the evidence needed and to record the consultations.
- Candidates' performance in the RCA expressed as pass or fail was not influenced by their perceptions on the assessment, but ethnicity, training, and English as first language were all significant predictors of exam pass rates.
- Recommendations were made for improvement by trainees responding including providing more guidance on case selection, more feedback, providing greater feedback, increasing consultation length, and offering further support or time to candidates based in practices with a higher number of patients coming from more deprived socio-economic backgrounds or with language barriers.

What this means:

- The RCA was broadly acceptable but some candidates experienced challenges and suggested areas for improvement, many of which have been implemented because of feedback from candidates.

Botan V, Williams N, Law GR, Siriwardena AN. *How is performance at selection to general practice related to performance at the endpoint of GP training*? Report to Health Education England. University of Lincoln, Lincoln, 2022.

What this study tells us:

- This is the first study to link performance at selection with all outcomes at licensing for doctors undertaking specialty training for general practice.
- The Multi-Specialty Recruitment Assessment (MSRA) scores for doctors at selection into training predicted general practice licensing outcomes for the MRCGP Applied Knowledge Test, Clinical Skills Assessment, Recorded Consultation Assessment, and Workplace Based Assessment – Annual Review of Competence Progression, within five years of starting training.
- The optimal MSRA threshold score for predicting an uncomplicated training pathway to licensing was around 500 in this large cohort.
- The Selection Centre added little to the predictive validity of the MSRA, so this analysis supports the decision made during the pandemic to discontinue the Selection Centre.
- Doctors' ethnicity did not reduce the chance of passing GP licensing tests once sex, place of primary medical qualification, declared disability and MSRA scores were taken into account.
- Doctors scoring below the MRSA threshold of 500 may need additional support during training to maximise their chances of achieving licensing.

What this means:

Ethnicity did not reduce the chance of passing GP licensing tests once gender, place of primary medical qualification, declared disability and MRSA scores were considered. Comparing candidate scores by ethnicity creates a false impression of differential attainment which should be addressed by routinely taking these factors into account.

Bell EA, Cleland J, Gambhir N. 'It clarified a lot': GP trainees as peer role players in a formative Clinical Skills Assessment (CSA). Educ Prim Care 2021, 32(2):85-90. DOI:10.1080/14739879.2020.1836521

What this study tells us:

- This qualitative semi-structured interview study explored 15 GP trainee role players' perceptions of peer role play in a formative CSA and found that role play provided an insight into the exam.
- It helped trainees know what to expect and how to approach the exam, emphasised the importance of communication skills, acknowledging the patient's perspective and aided professional development, particularly in observing good feedback.

What this means:

- GP trainees who role played patients in a formative CSA reported increased confidence and educational gains from their experience.

Winter R, Norman RI, Patel R. A qualitative exploration of the lived experience of GP trainees failing to progress in training. Educ Prim Care. 2021;32(1):10-18. DOI:10.1080/14739879.2020.1831970

What this study tells us:

- Professional factors, personal factors, and social factors affecting progression included difficulties with managing work-load, poor motivation, lack of family time and psychological ill-health.

What this means:

- Failure to fully understand trainees' journeys and associated challenges reduces opportunities to provide bespoke packages of care and remediation that fully address their needs.

Pattinson J, Blow C, Sinha B, Siriwardena AN. Exploring reasons for differences in performance between UK and international medical graduates in the Membership of the Royal College of General Practitioners Applied Knowledge Test: a cognitive interview study. BMJ Open 2019; 9:e030341. DOI:10.1136/bmjopen-2019-030341.

What this study tells us:

- This was the first study worldwide exploring reasons for differences in performance between UK graduates (UKGs) and international medical graduates (IMGs) in a licensing (applied knowledge test) examination using in-depth cognitive (think aloud) interviews.
- There are common causes of poor performance in the AKT whatever the ethnic background of the doctor, which are related to training and educational experience, knowledge skills and insight into these.
- IMG participants experienced additional difficulties because of differences (gaps) in their previous educational experience or lack of familiarity with the UK NHS.

What this means:

- Performance could be improved for all doctors in training by emphasising gaining clinical experience, increasing familiarly with the curriculum and receiving feedback to enhance personal insight into their knowledge and deficiencies.
- For IMGs a longer period of induction during UK training, addressing specific areas of difficulty, plugging gaps in undergraduate experience, and increasing understanding of NHS systems is also likely to aid performance.

Asghar Z, Williams N, Denney, M L, Siriwardena A N. . Performance in candidates declaring vs not declaring dyslexia in a licensing clinical exam. Medical Education: 53: . 2019;53:1243–1252. DOI: 10.1111/medu.13953

What this study tells us:

- Candidates who declared dyslexia late (after one or more failed attempts) were significantly more likely to fail compared with those candidates who declared dyslexia early (40.6% versus 9.2%; $p < 0.001$).
- Candidates who declared dyslexia late were more likely to have a non-UK medical qualification (79.3% versus 15.6%; $p < 0.001$) or come from a minority ethnic group (84.9% versus 39.2%; $p < 0.001$).
- The chance of passing was lower for candidates declaring dyslexia compared to those who never declared dyslexia and lower in those declaring late (incident rate ratio [IRR], 0.82; 95% confidence interval [CI], 0.70-0.96) compared with those declaring early (IRR, 0.95; 95% CI, 0.93-0.97).

What this means:

- A small proportion of candidates declaring dyslexia were less likely to pass the CSA, particularly if dyslexia was declared late. Further investigation of potential causes and solutions is needed.

Wakeford R, Ludka K, Woolf K, McManus IC. Fitness to practise sanctions in UK doctors are predicted by poor performance at MRCGP and MRCP(UK) assessments: data linkage study. BMC Med. 2018;16(1):230. DOI: 10.1186/s12916-018-1214-4

What this study tells us:

- Doctors sanctioned by the GMC performed substantially less well on MRCGP and MRCP(UK),
- Doctors on the 2.5th percentile of exam performance were about 12 times more likely to have to practice problems than those on the 97.5th percentile.
- Knowledge assessments and clinical assessments were independent predictors of future sanctions, with clinical assessments predicting GMC sanctions significantly better.
- MRCGP and MRCP(UK) performance were predictors of GMC fitness to practice categories conduct and trust.

What this means:

- The psychological processes involved in successfully studying, understanding and practising medicine at a high level may share similar mechanisms to those underlying conduct and trust.

McLoughlin K, Pope L, Walsh E, Jennings A, Foley T. The MRCGP Clinical Skills Assessment: an integrative review of evidence. Educ Prim Care. 2018;29(3):132-137. DOI: 10.1080/14739879.2018.1427510

What this study tells us:

- 11 of the 31 studies included were varied looking at the use of simulated patients, child actors and mock CSAs.
- Variables contributing to CSA performance, particularly in International Medical Graduates, were identified.
- A small pool of published evidence regarding the CSA exists, focused on factors influencing CSA performance.

What this means:

- Future research should explore effectiveness of interventions to improve CSA performance and candidates' experience of the CSA.

Neden CA, Parkin C, Blow C, Siriwardena AN. Has there been a change in the knowledge of GP registrars between 2011 and 2016 as measured by performance on common items in the Applied Knowledge Test? Educ Prim Care. 2018:1-7. DOI: 10.1080/14739879.2018.1467737

What this study tells us:

- There was no evidence of a change in performance of the question set as a whole.
- Candidates were more likely to get items on administration wrong compared with clinical medicine or research.

What this means:

- Candidate performance was stable over this 5-year period supporting the level of passing scores.

Asghar ZB, Siriwardena AN, Elfes C, et al. Performance of candidates disclosing dyslexia with other candidates in a UK medical licensing examination: cross-sectional study. Postgrad Med J. 2018;94(1110):198-203. DOI: 10.1136/postgradmedj-2017-135326

What this study tells us:

- The pass rate for candidates who declared dyslexia was 83.6% compared with 95.0% for other candidates.
- After adjusting for covariates linked to examination success including age, sex, ethnicity, country of primary medical qualification, stage of training, number of attempts and time spent completing the test dyslexia was not significantly associated with pass rates in the AKT.
- Candidates declaring dyslexia after initially failing the AKT were more likely to have a primary medical qualification outside the UK.

What this means:

- Performance was similar in AKT candidates disclosing dyslexia with other candidates once covariates associated with examination success were adjusted for.
- Earlier detection and support for candidates with dyslexia, particularly ethnic minority or International Medical Graduates may reduce higher failure rates in these groups of trainees.

McLoughlin K, Pope L, Walsh E, Jennings A, Foley T. The MRCGP Clinical Skills Assessment: an integrative review of evidence. Educ Prim Care. 2018;29(3):132-137. DOI: 10.1080/14739879.2018.1427510

What this study tells us:

- This review identified, critically appraised and synthesised published empirical research on the MRCGP and identified a small pool (n = 11) of heterogeneous studies mainly focusing on factors contributing to CSA performance, particularly International Medical Graduate (IMG) status and ethnicity.
- Methodological weaknesses were detected in five studies including insufficient or unclear information on study population, objectives, research questions, mixed-method approaches, and lack of consideration of researcher influence.

What this means:

- The study advocated future research examining reasons why these factors affect performance, exploring effectiveness of interventions to improve CSA performance, and considering candidates' experience of the CSA and the use of rigorous methods.

Hawthorne K, Roberts C, Atkins S. Sociolinguistic factors affecting performance in the Clinical Skills Assessment of the MRCGP: a mixed-methods approach. BJGP Open. 2017;1(1):bjgpopen17X100713. DOI: 10.3399/bjgpopen17X100713

What this study tells us:

- This study aimed to understand candidates' conversational contexts and behaviour in simulated CSA consultations to determine sociolinguistic factors for high- and low-performing candidates.
- There was more 'talk' in simulated consultations than in real life.
- On macroanalysis, there was little difference between poor- and well-performing candidates.
- Microanalysis found subtle differences in structuring consultations, metacommunication, picking up cues, and misunderstanding with and giving explanations to patients.
- Formulaic talk, contrary to examiners' perceptions was more common in successful candidates, but it was personalised and sited appropriately in the consultation.

What this means:

- The CSA is an interactionally demanding form of clinical assessment, that requires giving support to candidates and a more analytic approach to the development of interpersonal skills.

Bodgener S, Denney M, Howard J. Consistency and reliability of judgements by assessors of case based discussions in general practice specialty training programmes in the United Kingdom. Educ Prim Care. 2017;28(1):45-49

What this study tells us:

- Inconsistent judgements were made by different groups of assessors, varying with seniority.

What this means:

- More needs to be done to improve the reliability of case-based discussion (CbD), particularly in the hospital setting.

Naidoo S, Lopes S, Patterson F, Mead HM, MacLeod S. Can colleagues', patients' and supervisors' assessments predict successful completion of postgraduate medical training? Med Educ. 2017;51(4):423-431. DOI: 10.1111/medu.13128

What this study tells us:

- Multisource Feedback, Patient Satisfaction Questionnaire and Educational Supervisor's Review ratings were significantly correlated with subsequent performance in licensing exams ($r = 0.14$ to 0.47, $p < 0.01$),
- They also explained significant variance in the likelihood of requiring additional training time (15.6% to 25.6%, $p < 0.01$).

What this means:

- Review ratings have validity for predicting performance during training and are a useful tool for identifying trainees who are more likely to experience difficulty and could benefit from early additional support.

Denney M, Wakeford R. Do role-players affect the outcome of a high-stakes postgraduate OSCE, in terms of candidate sex or ethnicity? Results from an analysis of the 52,702 anonymised case scores from one year of the MRCGP clinical skills assessment. Educ Prim Care. 2016;27(1):39-43. DOI: 10.1080/14739879.2015.1113724

What this study tells us:

- There is a possibility that role-player subgroups (e.g., by ethnicity or sex) alter difficulty in the CSA for different candidate groups.
- Candidates were dichotomised by sex, by ethnicity and country of primary medical qualification (PMQ); role-players were dichotomised by sex and binary ethnicity; and the transaction of candidate/role-player encounters were classified as 'same' or 'different' in terms of the two parties' sex and of their ethnicity.
- Neither examinee nor role-player characteristics were found to predict any statistically significant portion of case score variance, where the significant ($p < .001$) predictors were source of PMQ (UK or elsewhere: 11% of case score variance), candidates' ethnicity (1%), candidates' sex (0.6%) and 'transactional' sex (0.1%).

What this means:

- Role-player subgroups did not systematically influence candidate subgroups' scores.

Wakeford R, Denney M, Ludka-Stempien K, Dacre J, McManus IC. Cross-comparison of MRCGP & MRCP(UK) in a database linkage study of 2,284 candidates taking both examinations: assessment of validity and differential performance by ethnicity. BMC Med Educ. 2015; 15:1.

What this study tells us:

- Understanding ethnic differences in performance in licensing assessment can be helped by comparing the performance of doctors who take both MRCGP and MRCP(UK).
- Performance for 2,284 candidates who had taken one or more parts of both assessments, MRCP(UK) typically being taken 3.7 years before MRCGP, was analysed on MCQs (MRCP(UK) Parts 1 and 2 and MRCGP Applied Knowledge Test (AKT)) and clinical exams (MRCGP Clinical Skills Assessment (CSA) and MRCP(UK) Practical Assessment of Clinical Skills (PACES)).
- Correlations between MRCGP and MRCP(UK) were high, disattenuated correlations for MRCGP AKT with MRCP(UK) Parts 1 and 2 being 0.748 and 0.698, and for CSA and PACES being 0.636.
- BME candidates performed less well on all five assessments ($P < .001$). Correlations disaggregated by ethnicity were complex, MRCGP AKT showing similar correlations with Part1/Part2/PACES in White and BME candidates, but CSA showing stronger correlations with Part1/Part2/PACES in BME candidates than in White candidates. CSA changed its scoring method during the study; multiple regression showed the newer CSA was better predicted by PACES than the previous CSA.
- High correlations between MRCGP and MRCP(UK) support the validity of each, suggesting they assess knowledge shared across both assessments.
- Although White candidates out-perform BME candidates, the differences are largely mirrored across the two examinations.

What this means:

- While the study does not give reasons for differential performance, similarity of effects in independent knowledge and clinical examinations suggests it is unlikely to result from specific features of either assessment and most likely to represent true differences in ability.

Russell D, Simpson R, Eden M, Foreman P. The use of child actors to simulate paediatric patients in the MRCGP Clinical Skills Assessment (CSA): developing an effective model. Educ Prim Care. 2015;26(4):242-247. DOI: 10.1080/14739879.2015.11494349

What this study tells us:

- An action research-based cycle was used to develop training procedures for child actors to ensure consistency for the CSA.
- Piloting allowed the use of natural pairs of adult and child role players in two teams for morning and afternoon circuits.
- A new process for training and calibrating children the day before and on the day of performance appeared to improve consistency.

What this means:

- This study shows how process for calibration and organisation led to greater consistency of performance in child role players.

Dixon H, Blow C, Milne P, Siriwardena N, Milne H, Elfes C. Quality assurance of the Applied Knowledge Test (AKT) of the MRCGP examination - an immediate post-test questionnaire evaluation of the candidates' views. Educ Prim Care. 2015;26(4):223-232. DOI: 10.1080/14739879.2015.11494346

What this study tells us:

- Most respondents believed that the test assessed their knowledge of problems relevant to general practice.
- Candidates identified training and knowledge needs particularly around research and practice administration.

What this means:

- This study supports the validity of the AKT content.

Pope L, Hawkridge A, Simpson R. Performance in the MRCGP CSA by candidates' gender: differences according to curriculum area. Educ Prim Care. 2014;25(4):186-193. DOI: 10.1080/14739879.2014.11494276

What this study tells us:

- Female GP trainees outperformed male peers in the CSA overall, in each assessment domain and in every curriculum area.
- Differences in performance were most marked in the areas of women's health and sexual health and least marked in cardiovascular problems and rheumatology and musculoskeletal.

What this means:

- GP trainees and trainers when planning educational activities and opportunities should prioritise development of consultation skills, and case mix presented to trainees.

Shaw B, Fox J, Brown J, Hart A, Mamelok J. An investigation of factors affecting the outcome of the Clinical Skills Assessment(CSA) in general practice specialty training.Education for Primary Care; 2014, 25; 91-5. DOI: 10.1080/14739879.2014.11494253

What this study tells us:

- Activity recorded for several different WBAs in a general practice trainee's portfolio did not correlate with passing the CSA examination.
- Years since qualification, and attending a non-European university were the strongest predictors of CSA fail, while being male and previous exam failure were also factors.

What this means:

- Portfolio activity was not a predictor of CSA outcome.

Dixon H, Blow C, Milne P, Siriwardena N. Reliability of non-pretested versus pretested questions in the applied knowledge test (AKT) of the MRCGP: evidence of quality assurance. Educ Prim Care. 2014;25(3):149-154. DOI: 10.1080/14739879.2014.11494265

What this study tells us:

- The AKT, which includes pretested and non-pretested questions, consistently shows high reliability.
- This study showed that inclusion of non-pretested questions did not reduce overall reliability of the test.

What this means:

- This study supports continued inclusion of new non-pretested questions to test new clinical knowledge and guidelines.

McManus IC, Wakeford R. PLAB and UK graduates' performance on MRCP(UK) and MRCGP examinations: data linkage study. *BMJ.* 2014;348:g2621. DOI:10.1136/bmj.g2621

What this study tells us:

- This study aimed to assess whether international medical graduates passing the two examinations set by the Professional and Linguistic Assessments Board (PLAB1 and PLAB2) of the General Medical Council (GMC) are equivalent to UK graduates at the end of the first foundation year of medical training (F1), as the GMC requires, and if not, to assess what changes in the PLAB pass marks might produce equivalence.
- Data linkage of GMC PLAB performance data with data from the Royal Colleges of Physicians and the Royal College of General Practitioners on performance of PLAB graduates and UK graduates at the MRCP(UK) and MRCGP examinations.
- PLAB1 marks were a valid predictor of MRCP(UK) Part 1, MRCP(UK) Part 2, and MRCGP AKT (r=0.521, 0.390, and 0.490; all P<0.001).
- PLAB2 marks correlated with MRCP(UK) PACES and MRCGP CSA (r=0.274, 0.321; both P<0.001).
- PLAB graduates had significantly lower MRCP(UK) and MRCGP assessments (Glass's Delta=0.94, 0.91, 1.40, 1.01, and 1.82 for MRCP(UK) Part 1, Part 2, and PACES and MRCGP AKT and CSA), and were more likely to fail assessments and to progress more slowly than UK medical graduates.
- IELTS scores correlated significantly with later performance but the effect of PLAB1 (beta=0.496) was much stronger than the effect of IELTS (beta=0.086).
- To produce equivalent performance on the MRCP and MRGP examinations, the pass mark for PLAB1 would require raising by about 27 marks (13%) and for PLAB2 by about 15-16 marks (20%) above the present standard.

What this means:

- PLAB is a valid assessment of medical knowledge and clinical skills, correlating well with performance at MRCP(UK) and MRCGP.
- PLAB graduates' knowledge and skills at MRCP(UK) and MRCGP are over one standard deviation below those of UK graduates, although differences in training quality cannot be taken into account.
- Equivalent performance in MRCGP(UK) and MRCGP would occur if the pass marks of PLAB1 and PLAB2 were raised considerably, but that would also reduce the pass rate, with implications for medical workforce planning.
- Increasing IELTS requirements would have less impact on equivalence than raising PLAB pass marks.

Patterson F, Tavabie A, Denney M, et al. A new competency model for general practice: implications for selection, training, and careers. Br J Gen Pract. 2013;63(610):e331-338.

What this study tells us:

- Eleven competency domains were identified including a new domain called Leading for Continuing Improvement.
- Empathy and Perspective Taking, Communication Skills, Clinical Knowledge and Expertise, and Professional Integrity were rated the most important domains.
- There was a significant increase in ratings of importance for each domain in future, except for Communication Skills and Empathy and Perspective Taking, which consistently remain high.

What this means:

- The breadth of competencies required for GPs has increased.
- GPs are now required to resolve competing tensions to be effective in their role, such as maintaining a patient focus while overseeing commissioning, with potential ethical conflicts between these aspects.
- Training provision arrangements should reflect the greater breadth of competencies now required.

Low S, Sheppard ZA, Tomkins S. Identifying factors associated with performance in a mock CSA: perceptions from trainees and trainers from one deanery. Educ Prim Care. 2013;24(6):452-460. DOI: 10.1080/14739879.2013.11494216

What this study tells us:

- Candidates who said that they had communication problems as a result of their accent, obtained significantly lower scores in the mock exam, suggesting that communication and language may be important factors for exam performance.
- The mock exam identified problems with performance, and stimulated expressions of interest for earlier preparation, working with peers and repeating the exercise.

What this means:

- Mock assessments are helpful for identifying performance problems and stimulating specific strategies to address these.

Esmail A, Roberts C. Academic performance of ethnic minority candidates and discrimination in the MRCGP examinations between 2010 and 2012: analysis of data. BMJ. 2013;347:f5662. DOI: 10.1136/bmj.f5662

What this study tells us:

- After controlling for age, sex, and performance in the AKT, significant differences persisted between white UK graduates and other candidate groups. Black and minority ethnic graduates trained in the UK or abroad were more likely to fail the clinical skills assessment at their first attempt than their white UK colleagues.

What this means:

- The authors stated that subjective bias due to racial discrimination in the clinical skills assessment may be a cause of failure for UK trained black and minority ethnic candidates and international medical graduates (but presented no evidence for this assertion of racial discrimination and did not take into account important confounders - see Siriwardena et al. DOI:10.3399/BJGP.2022.0474 above).
- The difference in training experience and other cultural factors between candidates trained in the UK and abroad could affect outcomes.
- Consideration should be given to strengthening postgraduate training for international medical graduates.

Denney ML, Freeman A, Wakeford R. MRCGP CSA: are the examiners biased, favouring their own by sex, ethnicity, and degree source? Br J Gen Pract. 2013;63(616):e718-725.

What this study tells us:

- Examiners show no general tendency to 'favour their own kind'.
- With confounding between variables, as far as the impact on candidates' case scores, substantial effects relate to candidate and not examiner characteristics.

What this means:

- Candidate-examiner interaction effects were inconsistent in their direction and slight in their calculated impact.

Russell D, Etherington C, Hawthorne K. How can simulated patients' experiences suggest ways to improve candidate performance in the MRCGP clinical assessments? Educ Prim Care. 2012;23(6):391-398.

What this study tells us:

- Simulated patients were confident they could portray a range of different cases.
- They offered clear views as to what made 'good' or 'poor' candidates.
- Positive characteristics included listening, interacting well, being empathetic, relaxed , non-judgmental showing positive non-verbal communication.
- While they did not feel able to mark or give feedback on candidates' clinical performance they made general observations about consulting behaviour in the examination, for example being approachable, showing interest, following logic rather than a checklist, avoiding making assumptions and being prepared to change ones mind.

What this means:

- Feedback from simulated patients could help guide learning and preparation prior to clinical examinations.

Siriwardena AN, Irish B, Asghar ZB, et al. Comparing performance among male and female candidates in sex-specific clinical knowledge in the MRCGP. Br J Gen Pract. 2012;62(599):e446-450. DOI: 10.3399/bjgp12X649142

What this study tells us:
Males performed significantly worse than females on female-specific questions, although the differences were small relative to the size of the test, but did not perform significantly better than females on male-specific questions.

What this means:

- Male doctors may need more exposure to clinical problems specific to women and children to address possible deficits in these areas.

Sabey A, Harris M. 'It's the conversation they'll learn from': improving assessments for GP Specialist Trainees in hospital posts. Educ Prim Care. 2012;23(4):263-269. DOI: 10.1080/14739879.2012.11494119

What this study tells us:

- Hospital assessors value WPBA because it enables feedback that GPSTs need, but they are not familiar with the GP curriculum or the standards required for GP training and struggle with giving feedback on poor performance.
- Concern that WBPA evidence may be used medico-legally inhibits honesty of feedback.
- WBPA was perceived to motivate doctors to pass assessments rather than to learn, focusing on a minimum standard of competence rather than excellence.

What this means:

- It should be determined how best, given limited consultant time and GP educator resources, to train hospital supervisors in relevant aspects of the GP curriculum and in giving feedback, particularly where GPSTs perform poorly.

Ahmed H, Rhydderch M, Matthews P. Do general practice selection scores predict success at MRCGP? An exploratory study. Educ Prim Care. 2012;23(2):95-100. DOI: 10.1080/14739879.2012.11494083

What this study tells us:

- The score for overall performance at selection achieved statistically significant correlation with examination performance ($r = 0.491$ for the AKT and $r = 0.526$ for the CSA, $P < 0.01$).

What this means:

- This relatively small study in Wales confirms that selection scores are important predictors of MRCGP success.

Sabey A, Harris M. Training in hospitals: what do GP specialist trainees think of workplace-based assessments? Educ Prim Care. 2011;22(2):90-99. DOI:10.1080/14739879.2011.11493974

What this study tells us:

- Trainees find face-to-face contact and verbal discussions useful.
- They place a low value on rating scale scores, and they perceive a lack of honesty in assessments, as well as bias and a ’box ticking’ attitude that undermines the credibility of WPBA.

What this means:

- WPBA needs changes to ensure trust in the assessment system, emphasising qualitative elements and ways to optimise the formative use of WPBA.

Bodgener S & Tavabie A. Is there a value to case-based discussion? Education for PrimaryCare,2011; 22: 223-228.

What this study tells us:

- Trainee performance in isolated case-based discussions was felt to be difficult to assess.
- A change in the trainee's ability to consult and perform could be demonstrated in subsequent assessments.
- Feedback received by trainees following assessment increased learning and learning outcomes irrespective of whether the assessment was felt to assess a performance.
- Educator skills in giving feedback was felt to be a limitation in developing trainees' performance.

What this means:

- Improving trainer feedback skills may improve the value and effectiveness of case-based discussion.

Miller A, Archer J. Impact of workplace based assessment on doctors' education and performance: a systematic review. BMJ. 2010;341:c5064. DOI: 10.1136/bmj.c5064

What this study tells us:

- Of 16 included studies 15 were non-comparative descriptive or observational and 1 was a randomised controlled trial.
- Studies reported positive results for the educational impact of workplace-based assessment tools but no objective evidence of improved performance.
- Multisource feedback can lead to performance improvement, although individual factors, feedback context, and facilitation have a profound effect on response.
- No evidence that alternative WPBA tools (mini-clinical evaluation exercise, direct observation of procedural skills, and case based discussion) lead to improvement in performance, although subjective reports on educational impact were positive.

What this means:

- Further research on how WPBA instruments can be used together, the role of facilitation and effects on performance are needed.

Siriwardena AN, Dixon H, Blow C, Irish B, Milne P. Performance and views of examiners in the Applied Knowledge Test for the nMRCGP licensing examination. Br J Gen Pract. 2009;59(559):e38-43. DOI:10.3399/bjgp09X395111

What this study tells us:

- This study evaluated the acceptability, feasibility, and validity of the new computer-based Applied Knowledge Test (AKT) delivered at local test centres.
- Most participants were satisfied with the new computer-based test content but also highlighted potential problems.
- Examiners had significantly higher scores than 'real' candidates who took an identical test.

What this means:

- Positive attitudes for most examiners and their higher success rate supports test validity.

McManus IC. Does performance improve when candidates resit a postgraduate examination? Med Educ. 1992;26(2):157-162. DOI: 10.1111/j.1365-2923.1992.tb00142.x

What this study tells us:

- A simple model based on pass rates at resits, and reliability of the old MRCGP examination was applied to the old MRCGP examination and found candidates increased their true ability before second and third attempts, after which ability declined.
- The model assumes alpha 0.8, that latent abilities of candidates are normally distributed, that candidates resitting examinations are a random subset of those who have failed; and that a linear logistic regression equation describes the relationship between latent ability and likelihood of success.

What this means:

- The model assumptions may not hold true for the current licensing assessment.