

Review of Recruitment to UK Foundation Programme

International Expert Panel

Chair, Professor Fiona Patterson

Summary Report

International Expert Panel Summary Report on Recruitment to UK Foundation Programme

1. Introduction

1.1. An independent International Expert Panel was convened by the Recruitment to Foundation Programme Steering Group to provide expert advice on a review of current and alternative methods for selection into the Foundation Programme. The role of the Expert Panel was to develop a consensus view regarding appropriate selection options for the Foundation Programme, based on their knowledge and expertise and informed by three commissioned literature reviews. The aim of this report is to summarise the views of the Panel and provide the Steering Group with a set of weighted selection options, along with suggested actions.

1.2. The members of the Expert Panel were as follows:

- Professor Fiona Patterson (Chair, City University, London)
- Dr Henk van Berkel (Maastricht University)
- Professor John Bligh (Peninsula Medical School)
- Professor Alison Bullock (Cardiff University)
- Dr Kelly Dore (McMaster University)
- Professor Eamonn Ferguson (University of Nottingham)
- Professor Richard Hays (Keele University)
- Professor Neil Johnson (University of Warwick)
- Professor Chris McManus (University College London)
- Professor Harold Reiter (McMaster University)

- Ms Victoria Carr (Supporting Officer, City University, London)
- Professor Paul O'Neill (Steering Group chair, University of Manchester)

1.3. The terms of reference for the International Expert Panel were that the Panel would:

- Consider the findings of literature reviews which have been commissioned to scope the empirical evidence from around the world for each of the different selection methods under consideration: *interviews, national examination, white space questions, structured record of achievement, any other method.*
- Suggest the range of options from the literature review evidence and their own expert knowledge.
- Provide a qualitative assessment of the options based on their expertise in medical education and recruitment.
- Provide an account as to how the information assessed was evaluated and conclusions drawn.
- Suggest mechanisms by which the option(s) viewed as most appropriate might be piloted and further developed to provide a workable mechanism for national adoption.
- Engage with stakeholders to explain the evidence and rationale behind the recommendations and to reflect upon the reactions and consider any necessary adjustments.
- Share information proactively and work closely with Cost Benefit Analysts to ensure all evidence presented to the Steering Group is consistent, and as well informed as possible.

1.4. An initial meeting of the Expert Panel was held on 27th March 2009. Following a discussion of the terms of reference, the Panel agreed to:

- Review the Foundation recruitment system, its strengths and weaknesses.
- Evaluate the options for selection methods using all available evidence.
- Assess the evidence regarding proposed selection methods, both from the literature reviews and from personal experience/expertise, in the context of several evaluation criteria (suggested by the Expert Panel, following guidance from the Steering Group), and placing the evidence in the appropriate context.
- Identify where there are gaps in the evidence regarding the current or suggested methods (*e.g. information on those who the system has failed, relative performance of Foundation doctors*).

- 1.5. At the initial meeting of the Expert Panel, the task requirements were reviewed. Following this, the Chair invited each Panel member to prepare and submit individual comments informed by the literature reviews. A preliminary analysis of existing Foundation Programme data was also conducted.
- 1.6. The Expert Panel was asked to provide a qualitative assessment of options based on their expertise in medical education and recruitment, and an account of how the information was evaluated and conclusions drawn. The initial evaluation criteria below were used to review each selection method. The Panel discussed whether certain criteria (e.g. reliability and validity) should be distinguished as more important or should constitute a minimum standard for evaluating potential selection methods.
- *Evidence of reliability*
 - *Evidence of validity*
 - *Capacity for discrimination / ranking*
 - *Fairness & legality*
 - *Candidate reactions*
 - *Stakeholder acceptance*
 - *Utility (time & costs to develop/implement)*
 - *Administrative practicality*
 - *National and regional considerations*
 - *Risks*
 - *Where further evidence is required*
- 1.7. A second meeting of the Expert Panel was held on 6th May, where the Panel discussed individual comments and began working towards a consensus view. After this meeting, an initial draft of the report was prepared on which Panel members commented. A final meeting of the Expert Panel was held on 29th May, following which Panel members presented their conclusions and implications for the Steering Group. A final version of this report was submitted in June 2009.

2. Purpose & Context of Foundation Programme Selection

- 2.1. The context of recruitment to the UK Foundation Programme contains a range of complex issues that need to be considered in evaluating potential selection methods. It was essential to clarify the nature and purpose of the selection process at the outset. In determining the scope of the project, the Panel identified three stages to the current Foundation selection process:
- **Stage 1:** National selection onto the Foundation Programme through (a) exclusion/ **de-selection** of those not qualified to proceed and (b) **allocation** to a Foundation School through a process of **matching to ranking/preference**.
 - **Stage 2:** **Allocation within Foundation Schools** to a specific Trust/scheme/post (*with variation in tools employed at this stage*).
 - **Stage 3:** **Pre-employment checks** by Trusts for jobs/employment (*with variation between Trusts in assessment tools employed at this stage*).
- It was agreed that only Stage 1 (a & b) would be considered by the Panel. Stages 2 & 3 were outside the scope of the review, although it was recognised that rankings from Stage 1 were likely to be used at Stage 2. It was also noted that issues of employability and patient safety remain crucial to the context and need to be considered.
- 2.2. Historically there have been Foundation Programme places available for all (or almost all) candidates. The purpose of the present system is therefore primarily *allocation* of posts, through deselection of those considered unsuitable and ranking of suitable candidates in order to facilitate matching. However, as the competition ratio may increase in future, the recommended selection option also needs to consider the potential for *selection* in a competitive environment. It is important to note that any increase in competition ratio beyond 1:1 (small or great) would have a major impact on the process. It was agreed that recommendations regarding selection methods would usefully take the form of a plan based primarily on allocation that could also account for the event of over-subscription to the Foundation Programme. It was suggested that the process could be seen as selection in either case, as candidates must meet specified criteria from the Person Specification.
- 2.3. Whilst the objective is to identify a single process for all applicants to the Foundation Programme, it was agreed that the selection process identified may potentially include assessments which certain groups of candidates would be exempt from, or referred to, on the basis of meeting/failing to meet certain pre-defined criteria. However, it must be ensured that

any such process was legally defensible and that any selection criteria must be made explicit in the Person Specification.

- 2.4. The Panel suggested that it is important to clarify and consider the different overall visions of how medical education and clinical care fit together in the UK and how the selection process accounts for these:
- A *unitary* system is based on assumptions that the quality of education and training provided by all medical schools (UK and EU) and UK Foundation schools are comparable and that the aim of the process is to provide graduates to the whole of the UK.
 - A *devolved* system accepts that there is variation in the attributes of graduates and the requirements of employers and the profile of training offered by NHS Trusts and therefore there will be competition to secure the best candidates/posts. In this scenario there would be added value to the NHS in being able to match the right candidates to the right Foundation schools and posts.

The assumptions made about the nature of the system will have implications for the structure and content of the selection process. The prevailing current assumption is that all Foundation schools are comparable (*unitary system*).

- 2.5. The Panel agreed that further work is urgently required regarding the criteria to be targeted in the Foundation selection process. Views on the critical attributes of Foundation doctors are likely to vary between patients and employers, candidates and educators. The issue of what attributes should be targeted for selection is complicated by the nature of the Foundation Programme as both the final stage of basic medical education and the first stage of employment in the NHS. The current UKFPO Person Specification is based on broad criteria derived from *Tomorrow's Doctors*, *The New Doctor* and *Good Medical Practice*. However, it was noted that there are no detailed interpretations of the standards required by *Tomorrow's Doctors* for candidates entering the Foundation programme and further research and development work on the UKFPO Person Specification is therefore recommended. Conducting a thorough job analysis is strongly recommended in order to develop robust, defensible selection criteria. Outputs from a specific job analysis (using established methodologies) will provide a more precise Person Specification for the Foundation Programme (NB It is assumed that there will be a single Person Specification to cover all Foundation Schools). This will be critical in ensuring that selection criteria are operationalised appropriately, in guiding choice of appropriate selection methods and enhancing the validity of the selection process in the long term.
- 2.6. In addition to the technical properties of proposed selection methods, it is essential to consider their appropriateness within the specific context and needs of selection into the Foundation Programme. In particular, it is vital that the preferred selection option should be acceptable to the various key stakeholder groups including candidates, employers, medical schools, Departments of Health, UKFPO, amongst others. Some methods that are legally defensible and thus 'fair' would not be acceptable (e.g. random allocation). Decisions about the most appropriate process for selection will be driven by various political factors as much as by scientific evidence, although it is essential that the selection process meets minimum evidence for reliability, validity and fairness. Candidates' perceived fairness of the recruitment process is an important factor regarding the acceptability of the system.

3. Evaluation of Current Methods

- 3.1. The Panel was asked to consider the strengths and weaknesses of the current process for Foundation selection. It was agreed that it was important to note how well the current selection system appears to have worked to date. Perceived advantages of the current system include the centralised process, the high proportion of candidates receiving their first choice, the fact that the system is well-understood, allows geographic flexibility and takes some account of individual needs.
- 3.2. It is important to understand the motivation, and evaluate the need, to adapt or change the current system. Some of the concerns about the current system relate to the specific methods used for ranking candidates, i.e. *academic quartiles* and *white space questions*, and the potential for these to be subject to legal challenge. Having reviewed these issues (informed by empirical analysis of recent data provided by the UKFPO; see 3.4), the Panel suggested that the rationale for using academic quartiles (as opposed to some other form of division) and for the points assigned to these, was unclear. Further, the Panel identified issues regarding standardisation and checking of international quartiles. Problems with the current white space questions approach include the potential for plagiarism/faking/coaching, limited shelf life and reliability issues. The potential for academic achievement questions to reward institutional structure rather than individual behaviour (e.g. schools where all students intercalate) was also noted.
- 3.3. The Panel identified other issues with the current system, including concerns about geographic flexibility and oversubscription in some areas (e.g. London), and depersonalisation (lack of a 'human element' to the process). A particular concern with the current system is that there is no available evidence regarding effectiveness, as predictive

validity data has not been analysed. This makes it difficult to gauge whether and how the current system requires changing. The Panel was unanimous in identifying a need to develop a *tracking methodology* to evaluate the current (or any new) system. Subject to appropriate approvals, data relating to the performance of selection of doctors for Foundation should be linked to assessment data during Foundation training, to enable evaluation of the predictive validity of the current or future selection process. In the long term, this evaluation would help inform speciality selection processes.

- 3.4. Technical reports on analyses of UKFPO data conducted on behalf of the Panel¹ highlighted the impact and potential concerns regarding the current allocation algorithm. One of the perceived strengths of the current system is that 91% of graduates are allocated to their first choice school, however it was noted that this is inherent in the use of the current 'First Choice First' algorithm and would occur even if ranking were to be allocated via a lottery. A concern of the current algorithm is the low choices allocated to the 9% of graduates who do not receive their first choice school; the median choice for those not placed in their first choice was 5th (2009) or 6th choice (2008). It was noted that the DCSF² has excluded the use of the First Choice First algorithm on the basis that it creates tactical behaviour¹. If this is the case, it must be acknowledged that all Foundation Programme data on choices made to date may have been influenced by tactical behaviour. The Panel and technical reports (Appendices 1 & 2) reviewed the properties of various alternative algorithms (e.g. *Serial Dictatorship*), although it was noted that special circumstances such as paired applications would need to be taken into account. If considering alternative algorithms, there needs to be agreement on the most important outcomes that the algorithm must achieve and identification of an algorithm that meets those outcomes with the highest degree of defensibility (e.g. *stability*³).

4. Evaluation of Literature Reviews & Selection Methods

- 4.1. The Expert Panel was asked to evaluate the evidence for proposed alternative selection methods (*interviews, national examination, white space questions, structured record of achievement, any other method*) using their own expertise and the commissioned literature reviews. The Panel extended its thanks to the research teams at the Universities of Durham, Newcastle and Warwick for their work with the literature reviews and suggested it would be helpful to make the three reports available publicly so that all parties can review the existing research literature.
- 4.2. The key points of evidence from the commissioned literature reviews can be summarised as follows:
- Traditional/unstructured *interviews* have low reliability and validity, whereas structured interviews have moderate reliability and predictive validity and positive user reactions. Multiple interview approaches offer stronger reliability and predictive validity than other interviewing techniques.
 - Performance on national licensing *exams* is a moderate predictor of performance in later clinical practice. The assumption that all graduates from UK medical schools can be ranked nationally on the basis of local results is not robustly supported.
 - Statements made in response to *white space questions* under tightly controlled conditions can have moderate predictive validity for subsequent performance. However, they are subject to plagiarism and can be expensive to score.
 - The predictive validity of portfolios/*structured records of achievement* is uncertain but they may show moderate reliability if designed and used appropriately. The cost of administering a portfolio assessment system comparable across medical schools is likely to be high and candidate acceptability appears to be low.
 - *Assessment centres* have value in allowing for a range of knowledge, skills and professional attributes to be evaluated in a time-efficient way, although they can be resource-intensive.
 - Behaviour which causes concern at undergraduate level has predictive validity for later clinical practice and a narrative summary of information relating to undergraduate behaviours may be of value for Foundation selection decisions.
 - Evidence suggests that selection tools focusing on professional attributes should feature in selection processes and will add a different insight into the candidate.
- 4.3. It was clearly apparent from the literature reviews and from the expertise of the Panel that there is *very little direct evidence* to support any individual or combination of selection tools, and even less evidence in this specific context. The

¹ Analyses conducted by Professor Chris McManus and by Drs Rob Irving & David Manlove, University of Glasgow.

² DCSF: Department for Children, Schools and Families.

³ If the matching algorithm is stable and a candidate asks why she was not allocated to a more preferred school, the answer is "because all of your preferred employers filled all of their posts with candidates having scores at least as high as you". If matching is not stable, this answer cannot be given in at least some cases (see Appendix 2)

current evidence base is severely limited, with a lack of experimental work, and focuses primarily on undergraduate selection to medical school, which may not generalise to this context and can at best be considered a rough guide. There is a small but emerging literature base for entry into UK specialty/postgraduate training but no direct research on Foundation selection in the UK, or indeed, directly comparable systems internationally. Validation research here is urgently required (see above). The Panel noted further issues relating to the literature include the *criterion problem* (i.e. identifying an appropriate outcome measure to evaluate the effectiveness of selection tools), use of existing tools not designed for selection/ matching, little focus on prediction of negative behaviours, little focus on incremental validity and lack of meta-analytic studies⁴ in this area. In light of these limitations, the Panel agreed that any proposed changes would need rigorous piloting and evaluation.

- 4.4. There was consensus within the Panel and from the literature reviews that multiple selection methods are needed to cover the broad range of relevant attributes required by the current Person Specification and to achieve sufficient reliability and validity of assessment. The Panel recognised that although multiple selection tools are optimal to enhance reliability and validity, the cost of operationalising a multi-method approach, such as assessment centres, may be prohibitive in this context if the focus is currently matching/allocation rather than high competition selection.
- 4.5. The Panel agreed that, given the nature of the job role, suitable selection options should cover assessment of both clinical knowledge/skills and professional attributes. In the literature, this distinction is often referred to in terms of cognitive and non-cognitive attributes. There is support in the current research literature for testing of non-cognitive attributes (e.g. on the grounds that graduating doctors are relatively homogeneous academically) and consensus that these professional attributes are important for employability. One approach identified was the use of situational judgement tests to assess professional attributes (see note 4.10 below).
- 4.6. **Interviews.** As reported in the literature reviews (and noted in 4.2), there is a considerable body of evidence to indicate that traditional unstructured interviews have low reliability and validity but structured interviews have moderate reliability and some evidence of moderate/good predictive validity. Approaches involving multiple interviews (e.g. Multiple Mini Interview (MMI) or Assessment Centre (AC)) could add further reliability and validity to the process. ACs offer advantages over other types of interview as they require candidates to demonstrate target attributes using a multivariate approach rather than relying on self-report. Interviews or other forms of face-to-face testing may be the most appropriate way to measure certain qualities and could also introduce the 'human element' perceived by candidates to be desirable. Considerable training and national calibration of interviewers would be required to ensure reliability and fairness. The Panel agreed that interviews are costly and complex to introduce on such a large scale (as with MMIs or ACs), making this approach difficult to justify unless the added value in this context could be clearly demonstrated.
- 4.7. **National examination.** There is evidence from other countries that performance on national licensing examinations (taken some time before qualifying), is a moderate predictor of performance in later clinical practice. A national examination in this context refers to a single national assessment used for ranking purposes (not licensing), although this could take a wide range of forms and requires further clarification (e.g. a written clinical MCQ⁵, OSCE-type⁶ skills stations, assessment of non-clinical attributes, administered on single or multiple days). Some key stakeholders are strongly opposed to a national examination and current acceptability is likely to be low. Concerns noted by the Panel included that a national examination could potentially lead to league tables of medical schools, and 'curriculum paralysis' as schools might 'teach to the examination'. Although there is positive evidence for national examinations in terms of reliability and predictive validity, a preferred approach might be some form of common content of assessment across medical schools (see 4.8 below).
- 4.8. **Common content of assessment.** The Panel agreed that some form of common content of assessment (or possibly test equating⁷) across medical schools (potentially involving banks of content for both written and clinical tests) might be a useful option to explore in order to assure equivalence of standards between medical schools. The Panel noted two key drivers (but potential challenges) for adopting a common content approach were (i) quality control of assessment and (ii) assessment of all candidates (UK and EU graduates). It was suggested that common content of assessment could incorporate elements testing professional attributes, and that non-UK graduates could be asked to take a composite examination which covered the common content elements which would be assessed in all UK schools. In particular, the Panel noted this would be a helpful approach to ensuring preparedness for employment for all candidates

⁴ Meta-analytic studies statistically combine the results of several independent studies addressing related research hypotheses in order to establish average effect sizes and provide a more accurate data analysis.

⁵ MCQ: Multiple Choice Questionnaire

⁶ OSCE: Objective Structured Clinical Examination

⁷ Test equating: use of a certain proportion of common questions in assessments across different locations in order to standard set assessments and ensure equivalent difficulty.

(UK and non-UK). The Panel noted some of the difficult issues regarding common content of assessment are timing, compensating for progression rates within different medical schools and accounting for/maintaining curriculum diversity. Test equating is a more complex process with potentially many unknown practical issues to resolve.

- 4.9. **White space questions.** There is little research literature on white space questions per se and evidence must be extrapolated from methods such as personal statements. White space questions can be seen as a pragmatic way of accessing professional attributes but can be expensive to score and there are concerns about longevity and plagiarism. Completing white space questions under invigilated and/or time-constrained conditions could help overcome some (not all) concerns about plagiarism/coaching but may work against the main purpose of white space questions in capturing reflection.
- 4.10. **Structured record of achievement.** Structured records of achievement are broadly defined as documents that capture and summarise the skills and achievements of a candidate, potentially similar to a portfolio. There is little evidence regarding the use of portfolios in medical selection and costs of introducing a standardised system across medical schools are likely to be high, although it is possible that some form of structured record of achievement could be used as a method for summarising relevant evidence from medical schools. The Panel noted emerging evidence that poor professional performance as a student is associated with poor professional performance as a clinician and suggested that this information could potentially be used in the ranking process, if captured appropriately. Such information could be provided to future employers through the 'transfer of information' process that is currently available but this process takes place after matching.
- 4.11. **Situational Judgement Tests (SJT).** The Panel discussed SJTs⁸ as an alternative to white space questions in some depth. Evidence on the reliability and acceptability of SJTs is available from the US NBME and selection into UK GP Specialty Training. If the financial resources for selection are limited, the SJT offers an effective and efficient option, although the development phase is resource intensive in the short term. SJTs can assess cumulative knowledge, skills and experience and aspects of preparedness for practice depending on the precise test specification. The Panel discussed whether SJTs should be MCQ or free-text questions, and raised some concerns about the potential for gamesmanship and about the security of SJT items, although this may be less of a concern than for white space questions. It was proposed that further research and possible piloting is necessary. Similar methods such as CMSENS⁹ could also be considered for 'at-a-distance' screening of professional attributes, with potential for increased efficiency through the use of computer-based approaches. The Panel noted that existing research for CMSENS is limited as this is a new area of development, although the approach is similar to SJTs.
- 4.12. **Personality Questionnaires.** Some of the literature reviews included consideration of personality measures as possible selection methods, in particular assessments of Conscientiousness, which has been found to relate to performance across a broad range of jobs. From the literature reviews, there is evidence that personality dimensions such as Conscientiousness is positively associated with undergraduate performance during medical school but this does not necessarily generalise to postgraduate performance. Given the nature and purpose of the recruitment focus here, the Panel agreed that it is unlikely that ranking individuals on Conscientiousness is a viable proposition. The selection literature on personality assessment suggests that personality questionnaires are best used to inform focused questioning during an interview process. Although there is evidence to suggest the predictive value of Conscientiousness in selection generally, issues of plagiarism and coaching may be heightened in the medical context and this may impact upon stakeholder acceptance.
- 4.13. The Panel was asked to consider international best practice of selection systems in its assessment of evidence, informed by both Panel member expertise and the literature reviews. The Panel discussed the evidence presented in the literature reviews from the US, Canada, Australia and Northern Europe. In the US and Canada, both students' and schools' preferences are taken into consideration in matching individuals with posts and there is a separate application process for international candidates, whereas in Australia, some states run lottery allocation systems. The UK selection context differs from some of these systems in that selection is also based on an assessment of candidate attributes. It was agreed that it would be desirable to have more evidence from other countries within Europe (where available), as these are also subject to EU employment law which is likely to preclude the separate processes for international candidates that are seen outside Europe.

⁸ SJT: a measurement method designed to assess candidates' judgement regarding situations encountered in the workplace. Candidates are presented with (usually) written depictions of scenarios they may encounter at work and are asked to identify an appropriate response from a list of alternatives.

⁹ CMSENS: Computer-Based Multiple Sampling Evaluation of Non-Cognitive Skills.

- 4.14. In addition to selection methods, the Panel considered possible changes to the structure of the recruitment process and the implications of these. One proposal was a system of competitive entry only for those who wish to *move locality*, either within the UK or non-UK graduates applying to the UKFPO. This suggestion was discussed at some length, as the Panel noted that the current selection process was constrained by the overproduction of graduates primarily in London and Scotland. One solution could be that the selection procedure need only be applied to those graduates declaring that they wish to move locality, although the legality of this approach is unclear (especially considering the principle of freedom of movement within the EU). Further evidence and advice on legality could be sought, as this approach could significantly increase the efficiency of the recruitment system.
- 4.15. The possibility was discussed of asking candidates to register an intention to apply in advance, as a way of estimating likely candidate numbers in time to implement a plan based on selection rather than allocation (see 2.2). It was envisaged that this pre-registration would only be used to give an estimate of numbers intending to apply to the Foundation Programme in total, rather than numbers intending to apply to each school. Potentially, UK medical schools could pre-register all final year students and international candidates could pre-register as individuals.
- 4.16. The possibility of charging candidates per Foundation School applied to, as in the Canadian and Australian systems, was discussed but was rejected owing to cultural differences in the UK regarding the acceptability of a charging system. The possibility of charging non-UK candidates who may be subject to a more extensive testing and screening process was also discussed.
- 4.17. It is essential to consider the timing of assessments and the implications of this for the feasibility and fairness of the selection process. The impact of any changes in the selection system on particular groups of candidates also needs to be considered, e.g. candidates applying for academic programmes, armed forces candidates, candidates with special circumstances and linked applications. These aspects were not reviewed in detail by the Panel.

5. International Expert Panel: Suggested Next Steps

- 5.1. The Panel provided various proposed actions in the course of its discussions and individual contributions. Some of these are actions relate to *selection methodologies* for the Foundation Programme, whereas others relate more broadly to *design and evaluation* of the selection system. Some of these actions have already been noted above.
- 5.2. Given the lack of clear criteria on which selection into the Foundation Programme is currently based, the Panel recommends an urgent task is to conduct a multi-method *job analysis* to define a robust set of selection criteria. This analysis should involve a range of key stakeholders and established methods following best practice (e.g. CIT interviews¹⁰ etc.) The outputs of this analysis will inform all subsequent recommendations regarding choice of selection methods. The Panel agreed that the outputs of a job analysis will augment, rather than replace, current policy documentation, the purpose being to inform the design and specific development of new selection tools.
- 5.3. Clarity regarding the purpose of the selection process is essential to inform appropriate choices of selection methodology. The present system, with places for all candidates, is concerned with *allocation*, but increased competition in future may require a system for *selection* in a competitive environment. Selection options therefore need to anticipate and be defensible for both allocation and selection in the event of over-subscription. It is acknowledged that any process will be at risk in the case of severe over-subscription.
- 5.4. The commissioned literature reviews were effective in documenting previous literature on selection and all researchers noted a lack of direct research evidence available in this specific area. In particular, the Panel proposes that meta-analytic summaries of the available data should be commissioned so that the research literature can be further synthesised for future reference. It is proposed that the wide range of literature reviewed and citations found are used to conduct a thorough, *structural meta-analysis* of selection methods in this context. An output from this work would be a table showing the relative accuracy of different selection methods based on all available research in this area.
- 5.5. Initial explorations of existing UKFPO data and the workings of the matching algorithm have been helpful in understanding the current selection system. Results from the Panel's analysis suggest that there is potential for improving the matching algorithm. This requires agreement on the most important outcomes that the algorithm must achieve and identification of an algorithm that meets those outcomes with the highest degree of defensibility.

¹⁰ CIT interviews: structured interviews using Critical Incidents Technique to collect observations of attributes and behaviours that are of critical significance to the job role.

- 5.6. The Panel discussed the lack of predictive validity data regarding the current system. It is recommended that a *tracking mechanism* is established to follow Foundation Doctors from ranking into the Programme through to selection into specialty training (and ideally, beyond). The two year time frame of the Foundation Programme would allow evidence to be generated relatively quickly for validation purposes. Data already held by the UKFPO and other stakeholders along the training pathway could also be combined to provide evidence on the relationships between various assessments.
- 5.7. As mentioned above, there was consensus amongst the International Expert Panel that *multiple methods* should be employed in the selection process which should include assessment of both clinical knowledge/skills and professional attributes, as defined by a job analysis.
- 5.8. The Panel agreed that *common content of assessment (or test equating)* where possible might be a positive move forward, although it is acknowledged that it may prove difficult to ensure acceptable levels of reliability, validity and fairness in practice.
- 5.9. The Panel generally supported the idea of making greater use of information collected during medical school. One proposal is that the nature of the information to be collected should be agreed at a local level, and incorporated into an overall reflection of performance at medical school, to replace current quartile ranking and allow a more fine-grained analysis of student performance. However, currently, it was noted that this information was not specifically designed for selection purposes. Potentially the information could include a common content of assessment (cf 5.8). Members suggested considering *narratives of performance* within the selection option, describing evidence of good and poor performance, as these may be relevant predictors of future performance and therefore of particular relevance to employers. The current transfer of information (TOI) post-matching does not allow any account to be taken of such predictors. However it was recognised that students may have concerns about such an approach. Collecting these measures of school performance could be seen by schools as empowering, or as an additional burden.
- 5.10. The Panel felt that the current system provides a useful platform for further development of the process in that it offers testing of both clinical knowledge/skills and professional attributes. One option discussed by the Panel, subject to the outcomes of the formal job analysis process, is to extend and improve the two elements (and the relative weighting of each) on which ranking is currently based, such that:
- Academic quartiles (or other divisions) become a composite measure of data existent within schools to reflect performance over the four/five year course, including narratives of poor performance (local decision). This would potentially provide more fine-grained information beyond quartiles for ranking purposes where appropriate.
 - White space questions are revised into a Situational Judgement Test under invigilated conditions covering domains relevant to preparedness for practice.
- 5.11. The Panel agreed that significant investment in selection tools, such as Assessment Centres or face-to-face interviews, may be more appropriate in a high stakes, competitive selection environment (e.g. selection into specialty training). Investment in developing assessment centres is recommended for specialty selection in the UK where, arguably, the stakes are higher.