

# Analysis of selection tools for appointment of specialty trainees in histopathology in Scotland

*This study explores possible benefits of structured selection methods in the employment of specialty trainees in histopathology in Scotland. It examines methods used in 2007 demonstrating novel approaches and discusses the positive reaction of consultants involved in the process.*

The medical training application system (MTAS) was introduced in 2007 in the UK, using new methods of selection for postgraduate doctors to specialty training grades. The administrative elements of MTAS have received wide, general and largely negative publicity (for example, Nachev, 2007) and are discussed in the recommendations of the Tooke report (2007). Within this process, however, were embedded potentially useful tools for shortlisting and assessing candidates for appointment at subsequent interview. This study was designed to investigate the reliability of the assessment tools used in selection and to evaluate the outcome of the use of these tools.

A generic person specification was applied to all first year (specialty training grade level 1; ST1) entry posts in all specialties. The application form was also generic for all ST1 candidates as were the shortlisting criteria and scoring protocols. These were also the subject of controversy which is outwith the scope of this study which considers the situation as it was presented to the specialty of histopathology (anatomic pathology) and how the histopathology community in Scotland responded to the challenges of the unfamiliar system.

A small but growing body of literature exists in the UK, investigating the development of tools for appointing doctors that can be standardized, fair, defensible, reliable, validated, cost effective and

achievable (Patterson et al, 2000, 2001, 2005; Randall et al, 2006; Patterson and Ferguson, 2007). Much of Patterson's work relates to general practice recruitment but more recently reports on public health (Pashayan et al, 2007) and surgery (Rowley, 2007) have appeared. These studies provide evidence supporting the use of structured selection methods targeted to personal attributes of candidates such as has been reviewed previously (McDaniel et al, 1994).

In considering the arguments and evidence summarized above, it is important that new methods of selection are not simply 'written off' because part of the administration was felt to be less than satisfactory. After all, the US National Residency Matching Program has used an electronic residency application system for many years (National Residency Matching Program Report, 2007), matching doctors to programmes in much the same way as MTAS was intended to do. The present analysis leaves national appointments administration processes aside and instead considers the reliability of the selection tools used, using appointment to programmes as an outcome measure. It addresses the issue of systematic bias, practicality and acceptability to selectors, drawing conclusions from that experience.

## Methods

### General

This was a prospective analysis of the selection process for specialty training in Scotland in histopathology in 2007. As this process was new to all concerned, it was decided to gather data from and examine the components of the Scottish histopathology process to find out how it worked in practice and to learn from it for future rounds. The 17 Scottish histopathology ST1 posts were advertised nationally through MTAS. The study dealt only with what occurred in Scotland after the

applications were lodged in the devolved national system. In Scotland in 2007, each of the four deaneries had responsibility for selection of a cohort of specialties. Histopathology, as one of the 'diagnostic' specialties, was handled by the North of Scotland Deanery in Aberdeen, with human resources support from NHS Highland in Inverness.

The selection process involved a 1-day shortlisting session and two sets of 2 days of interviews. The same core team was used for the entire process, augmented for the second interview set by other trained selectors.

### Selectors

From the pool of Scottish histopathologists (approximately 126), 12 were involved in the appointments process. Many had educational remits as college advisers, training programme directors and educational supervisors, while others held managerial positions in their respective departments. They were experienced in selection procedures.

These doctors had to provide evidence that they had undergone training in equality and diversity to demonstrate the need to eliminate discrimination on grounds of sex, marital status, race, disability, age, sexual orientation, language, social origin, or other personal attributes such as religious beliefs or political opinions and in accord with UK and Scottish legislation. NHS Education for Scotland North of Scotland Deanery ran two seminars to achieve this although other approved courses were also acceptable. The seminars were used as a vehicle for starting to devise and consider the interview tool in relation to the person specification and competencies relevant to histopathology.

### Shortlisting

Candidates who did not satisfy UK immigration and employment regulations, or

**Dr Peter W Johnston** is Consultant and Honorary Clinical Senior Lecturer in Pathology, Aberdeen Royal Infirmary, Aberdeen AB25 2ZD and Associate Postgraduate Dean, NHS Education for Scotland, North Deanery and **Ms Fiona H French** is Research and Development Officer, NHS Education for Scotland, North Deanery

Correspondence to: Dr PW Johnston

requirements regarding registration with the General Medical Council, were excluded. Shortlisting was to generic scoring guidelines over which those involved had little input. The selection was undertaken in one room on a single day with seven of the team present. The application form included fields to which shortlisters were not party and 14 scored areas:

1. Personal statement
2. Learning and personal development
3. Relevant activities/achievements
4. Clinical knowledge and expertise
5. Research skills
6. Additional qualifications
7. Prizes and other academic distinctions
8. Publications/presentations/posters
9. Experience of delivering teaching
10. Experience of audit
11. Communication skills
12. Coping with pressure
13. Managing others and team involvement
14. Professional integrity.

Shortlisting was started by a short discussion about scoring criteria, followed by independent scoring, although instances of concern or about which individuals felt the need for clarification were discussed to establish consistency and consensus reached. In accord with the protocol, each application was scored independently by two reviewers. The cut-off criterion for invitation to interview was based on providing at least four candidates for each post at interview as, at the time of shortlisting the selectors calculated this ratio would be likely to produce sufficient candidates to fill all 17 posts. The score sheets were later copied for the project and the details entered anonymously to a database for study.

### Interview

A question bank was constructed with reference to the person specification and the interview tool finalized by the interview panellists to ensure variety of format, novelty and appropriateness to the person specification, specialty and geographical location of the posts. Columns one and two of *Table 1* summarize the interview tool. Scoring was to a preset scale of one (poor response with many areas of concern and demonstration of several negative indicators) to four (good or excellent with no areas of concern and at most allusion to

negative indicators). Before the interview the appointable standard was set (see below). Interviews consisted of two meetings both with a panel of two interviewers. These interviewers alternately questioned or recorded notes. Both panels asked three of the six questions and assessed either language or communication skills. At the end of the interview, the panellists conferred and reached an agreed score for each area of assessment. These were recorded as the score for each candidate. The portfolio was assessed by the chairman or assistant chairman who were not interviewing. The roles of individuals and the pairing of interviewers were changed on a systematic basis over the days of interviewing to minimize systematic bias.

Interviews yielded numerical scores with a maximum of 36 from the nine assessment areas. A score of 34 was decided as the minimum 'appointable' score before the interviews on the basis that this level would exclude any candidate who scored zero or one in one or more fields or whose performance included the demonstration

of at least a few negative indicators. These scores were used to determine whether or not the panel judged the candidate appointable. Scores that were used in the application process were subsequently re-entered in an independent database in MS Excel for the purposes of this study. The scores for each interviewed candidate were correlated in SPSS. Combined scores were tested for reliability using Cronbach's alpha in SPSS.

Scores from the interview were used to provide NHS Scotland with a list of candidates recommended as being suitable for appointment. The appointments process was dealt with centrally by NHS Education for Scotland.

### Interviewer data

The time spent by the panellists in preparation and delivery of the system was also calculated. Views of interviewers about the system were noted at the final 'wash up' at the end of the interview process and on subsequent open e-mail discussions. Themes were identified.

**Table 1. Interview assessment tool**

Subject of element of assessment	Question, scenario, item assessed	Cronbach's alpha if item deleted
Histopathology report	The candidate will be presented with a histopathology report and asked to read it. The question will follow: 'Describe the clinical relevance of the findings reported to the management of the patient's case; how would you deal with this at a multidisciplinary team meeting?'	0.819
Audit	'Tell me about an audit in which you have been involved and explain its value to clinical care'	0.836
Incorrect diagnosis reported	'You are a specialist training registrar in histopathology. You have discovered an incorrect diagnosis in a report of yours of a fine needle aspirate from a neck lump. A metastatic cancer has been missed. Tell me how you would deal with this, please'	0.833
The picture	The candidate is given a picture, say of a country scene, and is asked to look at it. The candidate is asked to: 'Please describe the picture to someone at the end of a telephone.' And then: 'Can you tell me what feelings or messages the image creates?'	0.835
Presentation preparation	'Imagine you are now a specialist training registrar in histopathology. I am asking you to prepare a presentation to a group of junior doctors on the functions of a histopathology department. Tell me how you would go about this task.'	0.817
Commitment	Please explain to me why it is you want to be a histopathologist and why in Scotland?	0.800
Communication skills	Assessment of interview panel – not a specific question	0.829
Language skills	Assessment of interview panel – not a specific question	0.837
Portfolio	Assessment of submitted folder of achievements	0.851

**Results Analysis**

In Scotland in 2007, the four histopathology training programmes had collectively 17 available posts. There were 277 applicants for these posts, a ratio of 16.3:1. This ratio was greater than the Scottish average of 4.6 applicants to every available post. From these, 215 met UK regulations. A further 44 were excluded on person specification details. The remaining 171 applications were scored as described. Following this, 79 were called to interview and 76 attended. As a result of Scottish government intervention, the remainder (92 candidates) were invited to arrange interviews but only 61 made appointments for interviews and 41 attended. There are no data on why some candidates did not attend although the rates of non-attendance at the first and second interviews (three from 79, 20 from 61) appear different. Analysis was carried out on data from candidates who had undergone the whole process. Full data sets were available for review on 115 (98%) of the 117 interviewed.

**Shortlisting**

Analysis of the reliability of the shortlisting tool indicate a Cronbach's alpha for the whole tool of 0.857, reliability being defined as an alpha statistic of >0.700. Examination of the data shows all the components of the tool were discriminatory. The sections relating to academic qualifications, teaching and publications were slightly less so than the others. The correlation between the two shortlisters was 0.656 ( $P < 0.01$ ).

**Interview**

The reliability of the interview tool is shown in column three of *Table 1*. The overall Cronbach's alpha for this tool was 0.846. Again, all the component parts were discriminatory. The question on commitment to histopathology in Scotland was perhaps the most valuable while the examination of the portfolio was the least useful. Removal of these components from the Cronbach's alpha one by one changed the overall alpha to 0.800 and 0.851 respectively.

**Correlation of shortlisting and interview scores**

Comparisons of candidates' scores for shortlisting and interview are shown in

*Figure 1*. This figure demonstrates that there is a significant correlation between the two scores and that the shortlisting score predicted the interview score. It should be noted that from the first set of interviews, 57% of candidates interviewed (43/76) achieved scores defined as appointable. From the second set, only 7% (3/41) reached this score.

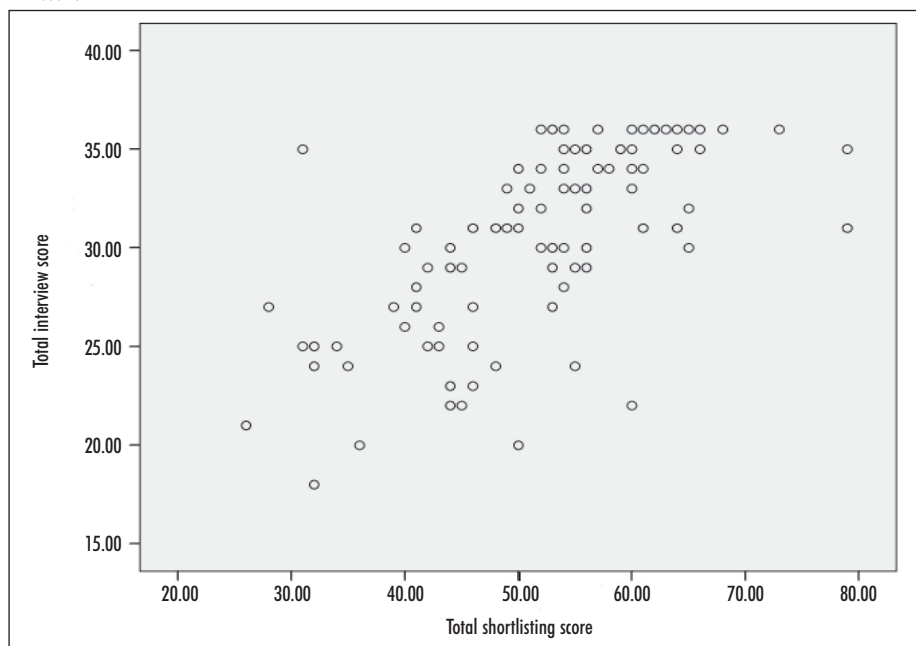
Overall, there was no evidence of systematic bias as regards the nationality,

years since graduation, post or specialty at the time of application (*Table 2*).

**Interviewer time and comment**

The panellists each spent half a day or a whole day on training depending on proximity to the venue for seminars or if they undertook the on-line version. The shortlisting took seven consultants one whole day excluding time taken in preparation by human resources staff. The first set of

**Figure 1. Comparison of shortlisting and interview score in histopathology. Pearson correlation = 0.656, P=0.01.**



**Table 2. Demographic details of candidates interviewed for histopathology training posts**

Demographic criterion		Applicants n (%)	Appointed n (%)	Not appointed n (%)	Chi squared
Nationality	UK	38 (35)	8 (50)	30 (32)	Not significant
	EU, not UK	9 (8)	2 (13)	7 (8)	
	non-EU	62 (57)	6 (38)	56 (60)	
		109 (100)	16 (100)	93 (100)	
Years since graduation	5 or less	65 (60)	9 (56)	56 (60)	Not significant
	6–10	28 (26)	5 (31)	23 (25)	
	11 or more	16 (15)	2 (13)	14 (15)	
		109 (100)	16 (100)	93 (100)	
Post at application	Foundation year 2	45 (42)	3 (19)	42 (46)	Not significant
	Senior house officer	40 (37)	9 (56)	31 (34)	
	Specialist registrar	4 (4)	1 (6)	3 (3)	
	Other	19 (17)	3 (19)	16 (17)	
	Total	108 (100)	16 (100)	92 (100)	

interviews used nine times 2.5 days and the second expended ten times 2.5 days. This amounted collectively to between 55 and 60 days in the preparation for and delivery of the selection process, bearing in mind that the second interview round took place as a result of political intervention. Overall this process selected 17 appointees from 277 candidates.

Questioning the panellists produced a number of themes. The group felt confidence in the system and expressed an opinion that it was fair, applied generic and specialty-specific selection criteria and did so on a standardized basis. The process of interviewing was less interesting than past experience of interviews had tended to be because the questioning was necessarily repetitive.

The group felt the outcome was satisfactory and that the size of the team involved and the coordinated approach were positive influences on this. The involvement in question design and blueprinting was felt to have contributed to a sense of ownership of the process and this too had been a positive experience. All those involved were happy to recommend that this or a similar set of tools should be used for subsequent selection of trainees.

At shortlisting, there had been scepticism about the usefulness of the shortlisting tool but experience about the perceived 'quality' of candidates interviewed and the scores achieved in the two sets of interviews led to comments about how well the shortlisting tool appeared to correlate with the interviewers' judgment of the ability of candidates to deal with the specialty-specific questions posed to them. These views were recorded before the completion of analysis.

### Discussion

This study was an evaluation of new selection tools used in the appointment of doctors to specialist training posts in histopathology in Scotland in 2007. The tools used were found to be reliable and discriminatory and the generic shortlisting methodology predicted performance in specialty-specific interview assessments. The process demonstrated that a group of senior histopathologists could work together effectively with relevant human resources staff across conventional geographic boundaries to maintain local cred-

ibility while ensuring a coordinated standardized approach across the Scottish nation.

The intervention of the Scottish government requiring that all eligible candidates be given the chance of an interview presented an opportunity to assess the shortlisting methods and extended the scope of the study. The results of the analysis were initially counter-intuitive but none the less reassuring as to the reliability of the shortlisting tool. It was reliable and its inter-observer consistency was high. It provided a good guide to who would be appointable and who would not. Similar correlation is noted by Pashayan et al (2007).

Some of the components of the selection tools used were novel but the principles conformed to established methods for structured selection (McDaniel et al, 1994) and to those used in other medical disciplines (Patterson et al, 2000; Rowley, 2007). The correlation of scores between shortlisters, although significant, might well have been improved by more focused training and discussion about the scoring of each field. This would be recommended practice in future.

While the shortlisting tool was generic, the interview was designed to complement attributes that are generic to all doctors with characteristics of specific significance to histopathological practice in mind. The generic shortlisting tool provided a baseline from which specialty selection could take place. The selectors' combined experience of training histopathologists led to the view that certain qualities in trainees contribute to a successful outcome to training while anecdotal evidence indicates trainees with, for example, difficulty in translating visual images into words fare less well.

From *Figure 1*, it is apparent that some candidates scored well on shortlisting but performed much less well at interview. If the assumption that the interview looked more specifically to sorting candidates into those who demonstrated attributes valuable to training in histopathology is accepted, it may be assumed that the outliers did not demonstrate these attributes. This may add confidence in the system as it noted those with generic qualities but was able to exclude those who lacked specialty-specific attributes. There was no way of testing this hypothesis in the current work, however.

In 2007, the selection process was complicated by the fact that many applicants had no or little experience of the discipline and so it was unrealistic to look for skills, like the interpretation of microscopic images, which require training in the specialty. A picture of a scene or still-life was used to test ability to observe and extract meaning from observations which was thought to be a fair achievable model of histopathological interpretation. Candidates were not expected to write a histopathology report but were expected to have seen and used the service and so be able to extract clinically relevant information from a condensed report. The scenario based on the issuing of an incorrect report produced interesting responses including disbelief and at times evidence of problems in dealing with errors of judgment.

The preparation of a presentation challenged some who were unable to untangle the generic presentational component that is essential for histopathology from the subject matter. The audit, commitment, language and communication skills components were reliable but less innovative as interview devices.

The portfolio proved more difficult, perhaps because these were very variable in length and content. Trainees from different backgrounds appeared to have varying perceptions as to the meaning of 'portfolio' and the opportunity to gather 'evidence' differed. In this way, scoring was a problem so as not to discriminate against those who would not have had the chance to collect much relevant material. The influence of each item in the Cronbach's alpha overall is not great but removal of the portfolio score shifted the overall alpha upwards more than other components' removal, suggesting it may be a less reliable element than, for example, commitment to the specialty.

The themes identified from discussions among the panellists point to a number of issues that are central to the perception of the process as successful. Central was the ability to ensure inclusion of representatives from the geographic regions of Scotland, the programmes and departments located there. Not all departments involved in training were represented but a spread was achieved. The size of the specialty and the numbers involved enabled

involvement of relevant people in a manageable and appropriate way.

The use of the training programme to start panellists thinking about blueprinting the interview against the person specification in a specialty-specific manner helped engagement and at the same time lent credibility to the process because people could see how questions set to generic standards across generic topics were relevant to their training expectations and understanding of how their own specialty worked. The ownership of the process that resulted from this was probably important in the willingness of panellists to participate especially in the second round of interviews and to their positive experience against a tide of negativity. Of course, the eventual validity of this as a selection process will emerge as the 17 appointed doctors progress through the training programme; although experience would tell us that they will do this despite us rather than because of us. **BJHM**

*The authors are grateful to Professor Gillian Needham, Dr Kim Walker and NHS Education for Scotland for encouragement and funding and to Findlay Stark for entering the data. Dr Maggie Cruickshank helped with the interpretation of the reliability data. Thanks to the histopathologists who spent time with the process and for their enthusiasm and support in the collection of data: Drs Stephen Dahill, Paul Fineron, Allan Howatson, Neil Kernohan, Mary McKean, Kathy Morton, Rosslyn*

*Rankin, Jenny Tolhurst, William Wallace, Shaun Walsh and Professor Stewart Fleming. These panellists would like to acknowledge the help, support and camaraderie of human resources and NHS Education for Scotland staff in Inverness: Jackie Sutherland, Kathy Black, Christine Holmes and Stephanie Grant.*  
*Conflict of interest: none.*

McDaniel MA, Whetzel DL, Schmidt FL, Steven D (1994) The validity of employment interviews: a comprehensive review and meta-analysis. *J Appl Psychol* **79**: 599–616

Nachev P (2007) MTAS or a tale of evidence heedless medicine. *BMJ* **335**: 615

National Resident Matching Program Report (2007) *Charting outcomes in the match. Characteristics of applicants who matched to their preferred specialty in the 2007*. NRMP Main Residency Match. 2nd edn. National Resident Matching Program and Association of American Medical Colleges, Washington, DC ([www.nrmp.org/data/chartingoutcomes2007.pdf](http://www.nrmp.org/data/chartingoutcomes2007.pdf) accessed 28 January 2008)

Pashayan N, Duff C, Mason BW (2007) Selection into specialty training in public health: performance of the Medical Training Application Service shortlisting. *J Pub Health* **29**: 331–7

Patterson F, Ferguson E (2007) *Selection for Medical Education and Training*. Association for the Study of Medical Education, Edinburgh

Patterson F, Ferguson E, Lane P, Farrell K, Martlew J, Wells A (2000) A competency model for general practice: implications for selection, training, and development. *Br J Gen Pract* **50**: 188–93

Patterson F, Lane P, Ferguson E, Norfolk T (2001) Competency based selection system for GP registrars. *BMJ* **323**: 2–3; Career Focus ([careerfocus.bmj.com/content/vol323/issue7311/](http://careerfocus.bmj.com/content/vol323/issue7311/) accessed 28 January 2008)

Patterson F, Ferguson E, Norfolk T, Lane P (2005) A new selection system to recruit general practice registrars: preliminary findings from a validation study. *BMJ* **330**: 711–14

Randall R, Davies H, Patterson F, Farrell K (2006) Selecting doctors for postgraduate training in paediatrics using a competency based assessment centre. *Arch Dis Child* **91**: 444–8

Rowley DI (2007) The new training landscape. *Surgeons' News Oct* ([www.surgeonsnews.info/Content/content.aspx?ID=137](http://www.surgeonsnews.info/Content/content.aspx?ID=137) accessed 28 January 2008)

Tooke J (2007) *Aspiring to excellence. Findings and recommendations of the independent inquiry into modernising medical careers*. MMC Inquiry, Aldridge Press, London

## KEY POINTS

- Structured criterion-based selection methods for appointing doctors to specialty training can be reliable.
- The process demonstrated no evidence of systematic bias.
- The recruitment team was engaged in the process by active participation in its design, content and delivery.
- Selectors felt ownership and enthusiasm for the future use of such a process.