Evaluation of the RCGP GP Training Curriculum

FINAL REPORT

May 2011

Dr Julie Bedward
Dr Ian Davison
Dr Sarah Burke
Professor Hywel Thomas

Centre for Research in Medical and Dental Education (CRMDE),
University of Birmingham

Professor Neil Johnson

Warwick Medical School
ACKNOWLEDGEMENTS

We would like to thank the five Deaneries: East Scotland; Kent Surrey and Sussex; Northern Ireland; North Western; and Wales for agreeing to be our research sites throughout the 3-year evaluation and for their valued help and support in arranging the focus groups each year. We would also like to thank all the Deanery Leads, GP Trainers, Programme Directors and GP STs who have given up their time to participate in the focus groups, interviews and survey.

Thank you to Dr Natasha Macnab for her valued contribution in conducting the case study interviews with trainees in one of the participating deaneries each year. Thanks also to Dr Vickie Firmstone and Dr Natasha Macnab for conducting one of the focus groups during the final year of the study. We would like to acknowledge the contribution made by Dr Mark Slater and thank him for the support he has provided with his IT expertise. Finally, thank you to Dr Magdalena Skrybant for the valuable administrative support provided and her painstaking work in transcribing many hours of audio recorded focus groups and interviews.
# Table of Contents

**ACKNOWLEDGEMENTS** ............................................................................................................................ iii  

**EXECUTIVE SUMMARY** .......................................................................................................................... 1  

- Introduction and Background and Context ........................................................................................................ 1  
- Methods .................................................................................................................................................... 1  
- The Qualitative Study ................................................................................................................................. 4  
  - Findings ............................................................................................................................................... 1  
  - Conclusions .......................................................................................................................................... 4  
  - Recommendations ................................................................................................................................. 4  

**Analyses of National Assessment Data** ........................................................................................................ 5  

- Findings ............................................................................................................................................... 5  
- Recommendations ................................................................................................................................. 6  

**1. INTRODUCTION** ................................................................................................................................. 7  

**2. BACKGROUND AND CONTEXT** .......................................................................................................... 8  

**3. METHODS** ........................................................................................................................................... 11  

- 3.1 Preliminary stage .................................................................................................................................. 11  
  - 3.1.1 Literature Review ......................................................................................................................... 11  
  - 3.1.2 Pilot Study .................................................................................................................................... 11  
  - 3.1.3 Selection of a Purposive Sample of Deaneries ........................................................................... 11  
  - 3.1.4 Preliminary interviews with Deanery Leads ................................................................................. 11  
  - 3.1.5 Findings from the preliminary phase of the study ......................................................................... 11  
- 3.2 Data Collection during the main phases of the Evaluation ................................................................. 12  
  - 3.2.1 Case studies .................................................................................................................................. 12  
  - 3.2.2 Focus Groups ................................................................................................................................. 13  
- Deanery ..................................................................................................................................................... 13  
  - Number of trainees ................................................................................................................................. 13  
  - Number of trainers ................................................................................................................................. 13  
- 3.2.3 Analysis of National Assessment Data ............................................................................................. 14  
- 3.2.4 National Survey ............................................................................................................................... 15
3.2.5 Interviews with Deanery Leads ................................................................. 16
3.3 The Qualitative Approach to Research used in the Evaluation .................. 16
  3.3.1 The nature of the research problem .......................................................... 16
  3.3.2 Philosophical influence ............................................................................. 16
  3.3.3 Sampling procedures ............................................................................... 17
  3.3.4 Approach to interviews and focus groups ............................................... 17
  3.3.5 Approach to data analysis ....................................................................... 18
3.4 Ethics ............................................................................................................. 19
3.5 Presentation of Findings .............................................................................. 19

4. FINDINGS: IMPLEMENTATION .......................................................... 21
  4.1 Initial Implementation .................................................................................... 21
  4.2 Findings from Year 3 ................................................................................... 22
    4.2.1 Conclusion ............................................................................................... 25
    4.2.2 Recommendations .................................................................................. 26

5. FINDINGS: HOW THE CURRICULUM IS EXPERIENCED ................. 27
  5.1 The written curriculum .................................................................................. 27
    5.1.1 Findings from Year 3 ............................................................................... 27
    5.1.2 Conclusions ............................................................................................. 35
    5.1.3 Recommendations .................................................................................. 35
  5.2 The ePortfolio ............................................................................................... 35
    5.2.1 Findings from Year 3 ............................................................................... 36
    5.2.2 Conclusion ............................................................................................... 44
    5.2.3 Recommendations .................................................................................. 45
  5.3 Assessment ..................................................................................................... 45
    5.3.1 Findings from Year 3 ............................................................................... 45
    5.3.2 Conclusion ............................................................................................... 62
    5.3.3 Recommendations .................................................................................. 63

6. FINDINGS: FITNESS FOR PURPOSE .............................................. 64
  6.1 Length and Structure of Training ................................................................. 64
    6.1.1 Findings from Year 3 ............................................................................... 64
    6.1.2 Conclusion ............................................................................................... 71
    6.1.3 Recommendations .................................................................................. 71
6.2 Preparedness for Practice of Newly Qualified GPs ................................................................. 71
6.2.1 Findings from Year 3 ........................................................................................................... 72
6.2.2 Conclusion ......................................................................................................................... 78
6.2.3 Recommendations .......................................................................................................... 79

7. ANALYSES OF NATIONAL ASSESSMENT DATA ................................................. 80

7.1 Introduction .......................................................................................................................... 80
7.2 Data manipulation and cleaning .......................................................................................... 81
7.3 Aims ....................................................................................................................................... 91
7.4 Results .................................................................................................................................. 92
  7.4.1 CBD ................................................................................................................................. 99
  7.4.2 COTs .............................................................................................................................. 102
  7.4.3 Mini-CEX ....................................................................................................................... 104
  7.4.4 CSR ............................................................................................................................... 107
  7.4.5 Correlations between the Workplace Based Assessments ............................................ 109
  7.4.6 ARCP ............................................................................................................................ 110
  7.4.7 AKT ................................................................................................................................ 113
  7.4.8 CSA ................................................................................................................................ 116
7.5 Discussion ............................................................................................................................ 118
7.6 Recommendations, limitations and potential further analyses ......................................... 124
  7.6.1 WPBAs and CSR ............................................................................................................ 124
  7.6.2 ‘Prediction’ of ARCP and summative assessments ......................................................... 125
  7.6.3 Data ............................................................................................................................... 126
  7.6.4 Descriptives .................................................................................................................. 126

8. DISCUSSION ......................................................................................................................... 128

8.1 Curriculum Implementation: Phase 1 .................................................................................. 128
  8.1.1 What are trainer and trainee attitudes towards the new curriculum? .............................. 128
  8.1.2 What support and resources would trainers and trainees value? ..................................... 130
  8.1.3 What operational issues have arisen from the implementation of the new curriculum? .... 131
8.2 How the Curriculum is Experienced – Phase 2 .................................................................. 132
  8.2.1 How has the new GP training curriculum impacted on trainees’ learning experience? ..... 132
  8.2.2 What do trainees perceive as the strengths and weaknesses of the curriculum? .............. 133
  8.2.3 How has the new GP training curriculum impacted on teaching practice and the role of the trainer? .................................................................................................................. 134
  8.2.4 How are WPBAs used by trainers and trainees in practice? .......................................... 135
8.2.5 How has the new GP curriculum impacted on teaching in the hospital based setting?.....136

8.3 Fitness for Purpose: Phase 3..................................................................................................136

8.3.1 To what extent is the training curriculum ‘fit for purpose’ in preparing GP specialty
registrars for independent practice? .......................................................................................136
8.3.2 What is the impact of the new curriculum on patient safety?..........................................137

9. CONCLUSION AND RECOMMENDATIONS ........................................ 138

9.1 Experience of implementation ..............................................................................................138
9.1.1 Recommendations..........................................................................................................138

9.2 Experience of the curriculum ................................................................................................138
9.2.1 Recommendations..........................................................................................................139

9.3 Fitness for purpose ...............................................................................................................140
9.3.1 Recommendations..........................................................................................................140

9.4 An Analysis of National Assessment Data ..............................................................................141
9.4.1 WPBAs and CSR ..............................................................................................................141
9.4.2 ‘Prediction’ of ARCP and summative assessments ...........................................................142
9.4.3 Data ...............................................................................................................................142
9.4.4 Descriptives ....................................................................................................................143

REFERENCES ........................................................................................ 144

Appendices .......................................................................................... 147

APPENDIX 1: Analysis of Case Study Eportfolios ........................................................................148
Characteristics ........................................................................................................................148
Learning log ...............................................................................................................................148
Personal Development Plan (PDP) ...........................................................................................149
Workplace-based assessment ....................................................................................................154
Case Based Discussion (CBD) .................................................................................................154
Consultation Observation Tool (COT) ......................................................................................156
Mini-CEX .................................................................................................................................158

Appendix 2: Focus Group Schedules Year 3 ................................................................................161
Evaluation of the RCGP Training Curriculum: Focus Groups 2008: Trainees.......................161
Evaluation of the RCGP Training Curriculum: Focus Groups 2010: Trainers and Programme
Directors ...................................................................................................................................165

APPENDIX 3: Questionnaire for GP Trainers ..............................................................................169
EXECUTIVE SUMMARY

Introduction and Background and Context

This Final Report presents findings from the three-year evaluation of the new training curriculum for general practice.

Methods

The research methods used in this evaluation include:

- A pilot study, preliminary interviews and a literature review which informed the design of research instruments.
- Focus groups were held with trainers and trainees in the 5 sample deaneries each year.
- The training experience of 10 case study trainees was documented in annual interviews.
- A content analysis examined case study trainees’ ePortfolios.
- A national survey of GP trainers and GP STs was conducted during the second year of the study.
- Semi-structured interviews were held with leads from each of the participating deaneries in the final year of the study.
- Analyses of national assessment data carried out during the first and final years of the evaluation.

The Qualitative Study

Findings

Implementation

- Initial implementation was a stressful process particularly for experienced trainers and ST3s who had to meet the new requirements during the first year. Trainers and trainees experienced difficulties learning how to use the ePortfolio and deaneries felt ill-prepared to support its use.
- Trainers have developed greater acceptance of the new curriculum and their changed roles.
- The work of the deaneries in providing training and developing resources to support the use of the training curriculum continues with particular focus now on hospital based clinical supervisors.

Experience of the written curriculum

- Initially trainers and trainees were overwhelmed by its size and had difficulty using it effectively to support teaching and learning. Training and the development of resources
have facilitated improved understanding and use of the written curriculum, particularly among GP trainers.

- The curriculum is used to identify trainees’ learning needs and gaps in their learning. However, trainees continue to find the written curriculum overwhelming and make limited use of it to support their learning.
- Resources are being developed to ensure that the GP curriculum is followed during the hospital based component of training.

**Experience of the ePortfolio**

- Initial difficulties have now eased and there is greater acceptance of the ePortfolio as an effective recording and assessment tool and its use in facilitating teaching and learning. Resources to support its use have proved useful and trainers and trainees have become more confident and competent in its use. Improvements have been welcomed but there is scope for making it easier to navigate, particularly to avoid user error.
- The ePortfolio can facilitate reflective learning and enable trainees to develop as mature self learners. Learning log entries linked to the PDP provide good evidence of trainees’ progress.
- Completing and commenting on the learning log is time consuming and a prescribed number of minimum entries is generally unpopular and viewed as detracting from their quality.
- The ePortfolio is used by educational supervisors to support their trainees in hospital posts.

**Experience of Assessment**

- There is continuing assimilation of the assessment system by trainers and trainees but the effectiveness of the WPBAs is assessor dependent.
- COTs and CBDs are used formatively in the general practice setting and this is generally valued by trainees. Trainers demonstrate increased confidence in ensuring that COTs and CBDs are used effectively and are based on challenging cases.
- Assessment carried out in hospitals is generally viewed as less effective and meaningful. Whilst training is provided for hospital clinical supervisors, many assessments are conducted by other staff that are not trained. Reluctance among some hospital based clinical supervisors to record concerns in the clinical supervisor’s report undermines its reliability in identifying weakness in trainees. The work of deaneries in providing training for hospital based assessors is continuing. It has to be recognised that the role of the clinical supervisor in hospital is quite different from the role of the GP educational supervisor because of the requirement to provide trainee supervision across a wide range of specialties, not just their own. The new training curriculum for general practice is one of many new specialty curricula that clinical supervisors are required to work with.
- The learning log, MSF and PSQ were highlighted as valuable in identifying weaknesses in trainees.
• The CSA and AKT are generally viewed positively.
• Community orientation and fitness to practice are considered the most difficult areas of competence to achieve.
• Clinical leadership, clinical governance, academic and research skills and the business skills in running a practice were identified as essential skills not sufficiently addressed by assessment.

**Length and Structure of Training**

• The additional post in general practice is viewed positively in facilitating learning and development.
• There is continued concern and dissatisfaction with the variable quality of the training provided in the hospital based component of training.
• Concerns were raised about gaps in experience in key specialties
• Study days are valued throughout the period of training and contribute to addressing gaps in learning.
• The educational supervisor’s role can contribute to increased general practice focus during hospital training.
• There is general acceptance that the overall period of general practice training needs to be extended to four or five years.

**Preparedness for Practice of Newly Qualified GPs**

• The new curriculum prepares trainees well for clinical practice and the ePortfolio prepares trainees for independent learning.
• ST3s are generally confident on embarking on their careers in general practice and recognise the need for continuous learning throughout their careers.
• The additional post in general practice allows trainees to maximise learning during their ST3 year.
• The written curriculum and assessments ensure broader curriculum coverage during training.
• There is improved coordination between the hospital based and general practice components of training facilitated by the educational supervisor’s role, the ePortfolio, the written curriculum and continuous assessment.
• The assessment system can enable identification of weakness in trainees.
• ST3s consider their main gap in learning is knowledge of the business side and administrative side of running a practice. Gaps in learning were identified in the following skills: clinical governance, academic and research, experience in the continuity of care, living with uncertainty and experience in key specialties.
• Most ST3s commence their careers in salaried or locum posts; this reflects the changing nature of employment in general practice, but most trainees consider these posts preferable to the commitment of a partnership at this initial stage.
Conclusions

- The implementation of the new training curriculum is generally perceived to have been successful. Its assimilation into practice has been a gradual process over the first three years. This process of assimilation is continuing and requires ongoing support.
- Taken together the three constituents of the new training curriculum, the written curriculum, the ePortfolio and the new assessment system have made a significant impact on the learning experience of trainees by providing clear aims and objectives; ensuring broader curriculum coverage; ensuring relevant learning occurs throughout training; allowing weaknesses to be identified; and empowering trainees as self directed learners.
- The new curriculum has increased the workload of the trainer, but new trainers generally welcome the structure it provides.
- The new training curriculum prepares trainees well for clinical practice, but does not prepare them fully for independent practice.
- The overall period of general practice training should be extended to allow greater consolidation of learning and to address the gaps in learning identified in the study.

Recommendations

- The process of implementation should be viewed as ongoing and support and training is necessary to ensure the intended meaning of the new training curriculum is assimilated into practice both within the general practice and hospital settings.
- Work on making the written curriculum more accessible for all GPs and those involved in general practice training should continue. This includes work currently being undertaken by the Curriculum Development Group and the Curriculum Guardians in refining and standardising the curriculum statements; and the work by deaneries in developing resources and training to support its effective use.
- Clearer guidance needs to be provided to trainees and educational supervisors on how log entries should be linked to curriculum headings.
- Further development is required to make the ePortfolio easier to navigate. The problems associated with user error with the ePortfolio need to be addressed.
- Training in assessment for trainers needs to continue along with workshops where best practice can be shared. Training in assessment for hospital based assessors needs to continue and should be extended to include all those who have involvement in assessment of GP STs.
- Work needs to continue to ensure greater consistency in the quality of hospital posts.
Training for educational supervisors needs to focus on developing their effectiveness in providing support to their ST1s and ST2s to ensure relevant general practice focus is achieved during their hospital based training.

Ways need to be explored of ensuring training includes coverage of skills in running a business, skills in clinical governance, academic and research skills; and also that experience is gained in continuity of care and living with uncertainty. English deaneries also need to consider how training can prepare future GPs for their proposed role in the commissioning of services. However, it is possible that this might only be achievable if the overall period of general practice training is extended.

Analyses of National Assessment Data

Findings

The ePortfolio database provided by the RCGP from the start until 31st July 2010 contained 19,929 Specialty Training (ST) years from 11,812 trainees. An average of 762 tagged events was recorded for each trainee i.e. about 21 per month. Most entries (73%) were to curriculum areas, tagged by the trainee, with the competency areas (27%) tagged by their supervisors. The total number of tagged events was a modest but significant predictor of passing ARCP first time at ST1, 2 and 3, perhaps because they indicate higher motivation or organisation. With AKT and CSA, tagged events were not significant.

The maximum correlation between years for the same WPBA (CBD, COTs, mini-CEX or CSR) was 0.24. The maximum correlation between different WPBAs in the same year was 0.19 with different assessors (COTs and mini-CEX), 0.35 with probably different assessors (other ST1 and 2 WPBAs) and 0.53 with probably the same assessors (COTs and CBDs in ST3). Differences between assessors may be of comparable importance to differences between trainees’ in determining assessment scores. As mini-CEX and CBDs are intended to feed into the CSR ratings, it is perhaps surprising that the correlations are so low: 0.32 max. This analysis suggests WPBAs were not sufficiently reliable to be used to evaluate trainee progress. Analysis of later cohorts would determine if there has been an improvement. Continued improvement of the reliability is recommended whilst maintaining acceptability and educational impact (we understand that work in this area is being done by the RCGP and deaneries).

In ST1, 93% of trainees were deemed satisfactory at ARCP first time, 92% in ST2 and 86% in ST3; correlations between these ARCPs were low (0.16 max). For all three years, the percentage was higher for younger, female trainees with more tagged events. A higher number of WPBAs, higher average score and taking them earlier in the training year were all associated with successful ARCP outcome. There is some evidence that undertaking sufficient WPBAs earlier in the training year may be more important for passing ARCP than the scores awarded. The importance of completing sufficient WPBAs in good time should be stressed to
trainees. Educational supervisors’ reports, qualitative interpretation of learning log entries and other naturally occurring evidence that is used to form ARCP judgements were not included in our analysis.

- For AKT; 85% of first attempts were passes; this percentage was slightly lower in ST1 and 2. For CSA, 98% of first attempts were in ST3 and 89% were passes.

- White trainees were more likely to be younger and UK trained; overseas trained and Asian trainees more likely to be male. Given this confounding, the combined effect of these demographics is more robust than differences between them. Younger, female, UK-trained, White trainees tended to do better at ARCP, AKT and CSA. These descriptives accounted for 20% of the variance for AKT and 39% for CSA, in line with findings in the MRCGP Annual reports. These large effects are not evidence of discrimination but are prima facie evidence that overseas, non-White, older, male trainees have greater difficulty with CSA than with AKT (we understand that socio-linguistic analysis of CSA cases by Kings College London and Cardiff University for the RCGP may aid understanding of an aspect of this issue). The explained variance was increase by a maximum of only 6% by WPBAs, CSR and ST1 and 2 ARCPs. The outcomes of ARCP and summative assessment depend far more on place of training, age, gender and ethnicity than on WBPAs.

- There were numerous coding errors in the ePortfolio data. The College should consider requiring legitimate entries for WPBAs to be recorded on the ePortfolio and should devise a system so that ST years and ARCP results are unambiguously recorded.

Recommendations

- Analyse WPBAs from later cohorts to see if reliability has been improved. Consider ways to increase the reliability of WPBAs so that they can be used to evaluate trainee progress. Undertake research to directly estimate the reliability of WPBAs.

- Undertake a reanalysis focussing on weaker trainees.

- (Continue to) stress to trainees the importance of completing sufficient WPBAs in good time.

- Explore qualitatively the ways in which ARCP judgments are made; also investigate the robustness of ARCP judgments.

- Consider requiring legitimate entries for WPBAs to be recorded on the ePortfolio and devise a system so that ST years and ARCP results are unambiguously recorded.

- Explore the reasons for the large differences in CSA performance between demographic groups.

- Incorporate estimates of criterion unreliability in future analyses of predictive validity for WPBAs and ARCP.
1. INTRODUCTION

This final report is based on a three-year evaluation of the new training curriculum for general practice in the UK that was introduced in August 2007. Prior to its implementation, the Royal College of General Practitioners (RCGP) commissioned a team of researchers from the University of Birmingham’s Centre for Research in Medical and Dental Education (CRMDE) and the University of Warwick’s Medical School, to conduct the evaluation.
2. BACKGROUND AND CONTEXT

The new RCGP Training Curriculum for general practice was introduced in August 2007 after approval from the Postgraduate Medical Education and Training Board (PMETB) that it met their required standards for new specialty curricula. There are three main constituents of the new curriculum: a written curriculum, an electronic portfolio and a standardised system of competence-based assessments. These changes were aimed at providing a standard of training that allows opportunity for valuable teaching and learning experiences, including assessments that are suitable, valid and reliable, so that newly qualified GPs are equipped with the knowledge and skills required to deliver high quality patient care.

Prior to the development and introduction of the new training curriculum, concerns had been expressed for more than a decade about the previous system of general practice training. In particular, a need was identified to provide a more coherent and relevant training experience that was more able to reconcile tensions between service provision and meeting the learning needs of trainees, historically most evident in the hospital based component of training (Reeve and Bowman, 1989; COPMeD, 1987; Dillner, 1993; Torry, 1996; Hand, 2000; Hand and McKee, 2001). There was also evidence that some newly qualified GPs did not consider their training had sufficiently met their training needs (Dixon and van Zwanenberg, 2001) and often did not feel sufficiently prepared to pursue partnerships at the beginning of their careers (Grant, 1998; Baker et al, 1995).

Over the three years of this study, the evaluation has explored stakeholder views and experience of how the new training programme has impacted on these issues. It is important to recognise that the implementation of major change often leads to tensions between old and new practices, resistance to change and uncertainty (Schon, 1971; Marris, 1975; Fullan, 2001; 2007). For change to be successful it has to occur at an individual level and, often, this involves a continuous process of feedback and adjustment (Fullan, 2001; 2007).

Consideration of these transitional issues was integrated into the design of the evaluation. Its phased approach has enabled regular feedback that has illuminated and facilitated these processes of adjustment. Importantly, the three-year duration of the study has provided opportunity to distinguish between initial reaction to change and the more long-term impact of the changed curriculum on general practice training. Over the three years of the study the evaluation has adopted a phased and layered approach that has explored stakeholders’ views and experience of the process of implementation; how the curriculum has been experienced by educators and learners; and the curriculum’s fitness for purpose. The evaluation was structured according to four overlapping phases of work:
• **Phase 1** (August 2007 to February 2010) focused on issues relating to curriculum implementation. The competence based curriculum and new modes of assessment required cultural and conceptual changes among trainers and trainees in their approach to education. Designed to encourage reflection and deep rather than surface learning, a period of adjustment was required to allow trainers and trainees to engage with these new approaches to teaching, learning and assessment and new operational systems. The evaluation explored these issues by investigating: trainee and trainer attitudes towards the new curriculum; the use of assessment tools in practice; and operational issues arising from implementation. The following research questions were addressed in this phase of the study:

- What are trainer and trainee attitudes towards the new curriculum?
- What support and resources would trainers and trainees value?
- What operational issues have arisen from the implementation of the new curriculum?

• **Phase 2** (June 2008 to November 2010) explored how the curriculum was experienced in practice by trainers, learners and other stakeholders, including their perceptions of the appropriateness of the curriculum structure and of teaching, learning and assessment tools and how these are used in practice to facilitate teaching and learning. The following research questions were addressed during this phase of the study:

- How has the new GP training curriculum impacted on trainees’ learning experience?
- What do trainees perceive as the strengths and weaknesses of the new curriculum?
- How has the new GP training curriculum impacted on teaching practice and the role of the trainer?
- How are workplace based assessment tools used by trainers and trainees in practice?
- How has the new GP training curriculum impacted on teaching practice in the hospital based setting?

• **Phase 3** (June 2009 to November 2010) explored the overall suitability of the curriculum in training competent doctors, its fitness for purpose and effect on how GPs practice. The following research questions were addressed in this phase:
To what extent is the training curriculum ‘fit for purpose’ in preparing GP specialty registrars for independent practice?

What is the impact of the new curriculum on patient safety?

- **Phase 4** (December 2010 to April 2011) this final report draws together evidence from all phases of the project and includes recommendations for development.

Feedback from the evaluation has been a continuous process which has contributed to informing the ongoing development of the curriculum. Channels for feedback have included Steering Group Meetings where regular updates on progress and preliminary findings have been reported; and membership and participation in the College’s Curriculum Development Group. In addition, Interim Reports have provided formal feedback at different stages of the evaluation. These include a report on initial deanery visits and pilot studies (Thomas et al, 2008); an Interim Report on findings from the first year of data collection, which included focus groups with trainers and trainees in the five research deaneries and a preliminary analysis of national assessment data on GP STs (Bedward et al, 2008); a report based on the second year of data collection, which included a national survey of trainers and GP STs and focus groups in the five research deaneries (Bedward et al, 2009).
3. METHODS

3.1 Preliminary stage

The research design and the design of research instruments was informed by the following methods: a literature review; a pilot study; interviews with deanery leads.

3.1.1 Literature Review

The design and first phase of the study was informed by a review of published literature on general practice training, educational change and change in postgraduate medical education. This included drawing on an earlier review of literature on general practice training that was conducted in 2004 by the Lead Researcher of the evaluation to inform the development of the new curriculum for general practice training. Documents produced by the RCGP and at deanery level were also included in the review.

3.1.2 Pilot Study

A pilot study was conducted between November 2007 and February 2008 to inform the development of research instruments. Separate focus groups were conducted in the West Midlands region with each of the following groups: GP STs, Programme Directors; Educational Supervisors. The focus groups aimed to explore participants’ perceptions and experience of the new training curriculum and to identify issues that they considered should be addressed by the evaluation.

3.1.3 Selection of a Purposive Sample of Deaneries

Five deaneries were selected as case study deaneries. To ensure UK wide representation, the Northern Ireland and Wales deaneries were invited to participate and both agreed to this. All Scottish and English deaneries were invited to express an interest in participating as a case study deanery. From the responses, one Scottish deanery, East Scotland, and two English deaneries, North Western and Kent, Surrey and Sussex (KSS) were selected.

3.1.4 Preliminary Interviews with Deanery Leads

Between December 2007 and February 2008 at least one visit was carried out in each of the five deaneries and semi-structured interviews were held with staff who held a key role in GP training including GP directors, associate deans, GP training managers, deanery administrators and programme directors.

3.1.5 Findings from the preliminary phase of the study

Findings from the pilot study were presented in a separate report (Thomas et al., 2008).
3.2 Data Collection during the main phases of the Evaluation

3.2.1 Case studies

During the first year of the study, ST1s in all the sample deaneries were invited to participate in the study over the first three years of their training. Ten trainees from three of the participating deaneries, Wales, North Western and KSS agreed to participate in this aspect of the study by taking part in semi-structured interviews each year. Their consent was also obtained to allow research access to their ePortfolios towards the end of their ST3 year. GP Educational supervisors and hospital based clinical supervisors who worked with these trainees were also invited to participate in interviews to discuss their experience of training with the new curriculum.

Separate interview schedules were designed for use with Trainees, GP trainers and hospital based clinical supervisors. These schedules were revised each year to ensure emergent and new topics were addressed. Interviews with the GP STs lasted approximately one hour. Interviews with GP trainers and hospital based clinical supervisors lasted approximately 30 minutes. All interviews relating to these case studies were conducted by telephone and were audio recorded with the interviewees’ consent. Recorded interviews were transcribed and thematically analysed.

During the final year of the evaluation semi-structured interviews were held with nine of the ten case study trainees between June and July 2010 as they were approaching the completion of their ST3 year and, in most cases, the completion of their GP specialty training. The trainees’ educational supervisors were also invited to participate in interviews again this year. Semi-structured interviews were held with two of these supervisors. Nine of the trainees consented to allow research access to their ePortfolios. Details from the ePortfolios were recorded and a content analysis was conducted between August and September 2010.

The case studies document the training experience of ten GP trainees over a three year period, which for nine of them constitutes their whole period of specialty training, and the first three years of a four year academic post for the other trainee. The findings from the case studies are not presented in detail in this report. Instead triangulation with other data highlighted the similarity between the experience of the case study trainees and that of trainees who participated in the focus groups each year. The more in depth data provided by these semi-structured interviews were drawn on to provide more detailed examples of the training experience. This allowed issues to be identified to explore in subsequent interviews and focus groups. The content analysis provides an example of how these trainees used the ePortfolio during their three years of training. Details of this are provided in Appendix 1. Interview schedules used in the final year of the study are presented in Appendix 2.
3.2.2 Focus Groups

During each of the three years of the study, focus groups were held with GP educational supervisors, GP clinical supervisors and programme directors in each of the five participating deaneries. Separate focus groups were also held on an annual basis with GP STs in each of these deaneries. Details of the number of participants from each deanery in each of the three years of the study are shown in the tables below.

Table 3.1: Focus group participants Year 1

<table>
<thead>
<tr>
<th>Number of focus groups held</th>
<th>Number of trainees</th>
<th>Number of trainers</th>
<th>Number of programme directors</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>114</td>
<td>57</td>
<td>12</td>
</tr>
</tbody>
</table>

Table 3.2: Focus group participants Year 2

<table>
<thead>
<tr>
<th>Deanery</th>
<th>Number of trainees</th>
<th>Number of trainers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deanery 1</td>
<td>11</td>
<td>12</td>
</tr>
<tr>
<td>Deanery 2</td>
<td>22</td>
<td>8</td>
</tr>
<tr>
<td>Deanery 3</td>
<td>8</td>
<td>5</td>
</tr>
<tr>
<td>Deanery 4</td>
<td>5</td>
<td>11</td>
</tr>
<tr>
<td>Deanery 5</td>
<td>21</td>
<td>8</td>
</tr>
<tr>
<td>Total no. of participants</td>
<td>67</td>
<td>44</td>
</tr>
</tbody>
</table>

Table 3.3: Focus group participants Year 3

<table>
<thead>
<tr>
<th>Deanery</th>
<th>Number of trainees</th>
<th>Number of trainers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deanery 1</td>
<td>9</td>
<td>22</td>
</tr>
<tr>
<td>Deanery 2</td>
<td>21</td>
<td>10*</td>
</tr>
<tr>
<td>Deanery 3</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td>Deanery 4</td>
<td>6</td>
<td>9</td>
</tr>
<tr>
<td>Deanery 5</td>
<td>4</td>
<td>11</td>
</tr>
<tr>
<td>Total no. of participants</td>
<td>50</td>
<td>61</td>
</tr>
</tbody>
</table>

\* 10 trainers in Deanery 2 completed questionnaires as it was not possible to arrange a focus group.

The format of the focus groups was the same throughout the study. The number of participants in individual focus groups ranged from 4-22. Participants’ views and experience of the new training curriculum were explored in the focus groups each year. The initial design of focus group schedules was informed by the preliminary phase of the study. The schedules were revised each year to ensure that emergent issues and topics were addressed in subsequent phases of the evaluation. The schedules used in the final year of the evaluation are contained in Appendix 3. The focus groups lasted between 60–90 minutes. All focus groups were audio recorded with the consent of the participants. Audio recordings were transcribed in full and thematically analysed.

During the final year of the evaluation focus groups were held in each of the five research deaneries between May and September 2010 to explore the views and experience of GP STs, GP trainers and
programme directors. Separate meetings were held for trainers and trainees. A total of 10 focus
groups were held. In Deanery 1 it was not possible to arrange a focus group with trainers, so trainers
from one district were invited to complete a questionnaire (see Appendix 4). Ten trainers returned
completed questionnaires.

3.2.3 Analysis of National Assessment Data

During the first and final years of the evaluation, analyses of the national ePortfolio assessment
dataset were carried out. In the final year of the study, analyses of these data were conducted
between September and December 2010. Data from the ePortfolio were received from the RCGP.
These data were from the start of the ePortfolio until 31st July 2010. The data were in several large
tables in Access with each line representing an event such as a workplace based assessment (WBPA).
Access, Excel and SPSS were used to transform this ‘event-level’ data in several tables to ‘trainee-
level’ data in one table.

To undertake these analyses, the assessment data needed to be coded in particular ways and these
differ from their use by the College. Whereas the College uses the clinical skills assessment (CSA) to
assess trainees, our analysis is designed to assess the WPBAs. For this reason, we developed a single
score to represent performance at CSA, whereas trainees were assessed on the number of stations
that they passed. The following coding scales were used: for case based discussion (CBD),
consultation observation tool (COT) and clinical supervisor’s report (CSR): 1= insufficient evidence; 2=
needs further development; 3= competent; 4= excellent. For mini clinical evaluation exercise (Mini-
CEX): 1= insufficient evidence; 2= below expectations; 3= borderline for completion; and 4= meets
expectations for completion; and 5= above expectation. For the CSA, candidates were assessed on 12
cases and their performance on each case was graded Clear Pass, Marginal Pass, Marginal Fail or Clear
Fail. Candidates generally needed to pass 8 of the 12 cases (9 in the first sitting). However, for our
purposes, the 12\(^2\) consultations were coded as follows: Clear Pass = 4, Marginal Pass = 3, Marginal Fail
= 2 and Clear Fail = 1. Therefore, the maximum score was 4 x 12 = 48. The pass mark for the applied
knowledge test (AKT) varied with each sitting to reflect different levels of difficulty. To account for
this, we considered a candidate’s performance in terms of how many marks they were awarded
above or below the pass mark. Those with negative ‘mark minus pass mark’ are ‘fails’.

Five outcomes for the annual review of competence progression (ARCP) were coded: 5= Satisfactory
Final – clinical or Satisfactory Progress – clinical; 4= Unsatisfactory - Insufficient evidence; 3=
Unsatisfactory - No additional training; 2= Unsatisfactory - additional training; and 1= Unsatisfactory -
Released.

\(^1\) The pilot station data were not included and the change to 13 scored consultations happened after the end of
data collection.
Start and end dates of each Specialty Training (ST) year for each trainee were taken from the ARCP data. From these, the length of each training year was calculated for each trainee. There has been concern that some (less good) trainees may undertake most of their WPBAs near the end of the year. Therefore, a ‘year decimal’ was calculated for each WBPA, using the start and end dates of the training year along with the date of the assessment. For example, a ‘year decimal’ of 0.25 would mean the assessment had been taken one quarter of the way through the ST year, whatever the duration of that ‘year’.

CBD, COT, mini-CEX and CSR were all linked with ST years. To make this manageable, only the date and overall assessment score were considered. These merged data allowed calculation of the number of entries per ARCP year along with total and average score and the average ‘year decimal’, which is a crude measure of whether the trainee leaves most of their assessments to the end of the year.

Many of the variables are not normally distributed. To ease comparison between them, all correlations used Kendall tau-b.

Multiple regressions where performed used stepwise entry in blocks and listwise deletion. Blocks of variables were used when some are logically prior to others: gender, age, place of training and ethnicity were in block 1; block 2 contained the WPBAs. Stepwise means that the variable that is the best predictor is entered first in the block, followed by the variable that best predicts the variance that is not accounted for by the first variable, etc. Variables were included if significant at $p=0.05$, and removed if $p>0.1$. Non-significant variables are not included, but they may be significant if it was not for the presence of the already entered variables. Listwise deletion was used, which means that trainees were only included in the regression if they had all the relevant data. Multicollinearity would have been considered problematic if the tolerance was less than 0.2. Plots of *ZRESID against *ZPRED were used to visually look for problems of heteroscedasticity and non-linearity. Lack of autocorrelation was assumed as the Durbin-Watson statistic was always between 1 and 3. Adjusted $R^2$ square was used as the measure of variance accounted for by the models.

### 3.2.4 National Survey

During the second year of the evaluation, a national survey which explored the views of trainers and trainees was conducted between June and September 2009. The survey was in electronic format and used Bristol Online Surveys. Separate questionnaires were designed for trainers and trainees with findings from Year 1 of the study informing their design. The survey was completed by 1471 trainees and 280 trainers. Findings from the survey were analysed using SPSS. These findings were presented in the Interim Report based on the second year of the evaluation (Bedward et al, 2009).
3.2.5 Interviews with Deanery Leads

During the final year of the study, between August and October 2010, semi-structured interviews were conducted by telephone with deanery leads from each of the five participating deaneries. See Appendix 5 for this interview schedule. Interviews lasted approximately one hour and were audio recorded with the interviewees’ consent. These interviews were transcribed and the data were thematically analysed.

3.3 The Qualitative Approach to Research used in the Evaluation

3.3.1 The nature of the research problem

The research team were commissioned to carry out an evaluation of the new RCGP GP Training Curriculum in practice during the first three years of its implementation. A major aim of the evaluation was to gain an improved understanding “of how GP Specialty Registrars learn through a competence based curriculum and the strengths and weaknesses of the curriculum” (RCGP, 2007). To fulfil this, and the other aims previously discussed, a predominantly qualitative approach to research was adopted. The introduction of the new training curriculum constitutes significant transitional educational change and this had to be reflected in the design. We subscribe to Fullan’s (2001) view that educational change is a process not an event. We also acknowledge the complexity of significant change and the importance of considering the interaction between the content, context and process of change (Pettigrew et al, 1992). The qualitative approach adopted facilitated exploration of the views and experience of trainers and trainees training with the new curriculum over the first three years of its implementation. It was successful in capturing the experience of educational change as a process and in highlighting the relationship between the content, context and process of change.

3.3.2 Philosophical influence

The qualitative approach adopted in the evaluation is underpinned by the ontological and epistemological assumptions of phenomenology, in particular the work of Schutz (1962; 1964; 1972). Schutz (1962) made important distinctions between natural and social sciences:

The world of nature, as explored by the natural scientist, does not ‘mean’ anything to molecules, atoms and electrons. But the observational field of the social scientist – social reality – has a specific meaning and relevance structure for the human beings living, acting, thinking within it. By a series of common-sense constructs they have pre-selected and pre-interpreted this world which they experience as the reality of their daily lives. It is these thought objectives of theirs which determine their behaviour by motivating it (Schutz, 1962: 59).
From this perspective, understanding of social phenomena can only be achieved by considering the meaning of social behaviour in relation to particular modes of social life: meaning is actively negotiated and constructed during interaction. This particular approach to qualitative research, phenomenological interactionism, aims to gain an understanding of phenomena from the viewpoint of participants. It seeks to preserve the integrity of the explored phenomena by aiming to faithfully reproduce how these are perceived and understood (Benson and Hughes, 1983). Its focus is on context and process and in providing “‘thick’ description”, not explanation or measurement (Bryman, 1988). It involves interpretation “based on systematically carried out inquiry” (Strauss and Corbin, 1998:8).

3.3.3 Sampling procedures

Quantitative research often aims to focus on a representative sample so that generalisations can be made. It is more usual for qualitative research populations to be selected to ensure diversity so that multiple perspectives are reflected (Patton, 1990). This was achieved in the qualitative aspects of the study through purposive sampling. The research deaneries were selected to ensure different regions of the UK were included. Within these deaneries, focus groups were held in urban and rural areas to capture different training experiences. The study was designed to ensure that the training experience in different roles was reflected. Research participants included GP STs, deanery leads, programme directors, educational and clinical supervisors in general practice and hospital based clinical supervisors. Whilst the voluntary nature of participation could result in focus primarily on particular types of trainee and practitioner, it could be argued that all invited participants had a vested interest in contributing to the study because of their involvement with general practice training. This proved to be the case with focus group participants reflecting diversity in views, experience, age, gender and ethnic group. The focus groups with trainers included new and experienced trainers. Participants in the trainees’ focus groups included representation from all the training years; although in the final year of the study we were particularly keen to focus on ST3 trainees who had trained entirely with the new curriculum.

3.3.4 Approach to interviews and focus groups

Pilot interviews and focus groups informed the design of interview schedules. This ensured focus on relevant key issues and areas during the first year of study. In subsequent years, findings from the previous year were used to adapt and refine these tools. A focused interview approach was adopted in the interviews and focus groups (Merton et al, 1956). This approach aims to explore participants’ views and experience as fully as possible and requires skilled and experienced researchers. Meaning was explored and clarified during the interactive process. The degree of consensus in the views and experience of focus group participants was examined across the range of topics explored. All focus
groups and interviews were audio recorded to provide reliable records and these recordings were fully transcribed. Whilst the researchers carried out some transcriptions, the timeframe of the study meant secretarial support was necessary to complete this work. Guides for transcription were used to ensure consistency in approach and researchers checked the accuracy of these transcripts against the original sound files.

3.3.5 Approach to data analysis

Data analysis was carried out manually using a thematic approach. During the first year of the study, two members of the research team were involved in this formal process of analysis. In the second and third years, analysis was conducted by the research lead working with the framework previously agreed. Expertise in qualitative research and the manual approach to analysis adopted ensured data did not become fragmented during this process and that meaning could be interpreted within its context.

Initial categories were generated by detailed line-by-line analysis of transcripts. There was congruence between many of these initial themes and those generated by the situational analysis which preceded the formal study, and also with the themes that emerged from the pilot study. Generation of themes was facilitated further by the informal processes of analysis which began during the first stages of the evaluation. These informal processes correspond with the documentary method of interpretation, which can be viewed as an integral part of the research process and a constituent of everyday social interaction (Garfinkel, 1967). The documentary method of interpretation relates to the process where the meaning of concepts is achieved through a continuous process of learning in which understanding is constantly reassessed and restructured in the light of further experience.

Further formal analysis focused on identifying finer categories within these themes, or relationships among categories; a process referred to by Strauss and Corbin (1998) as ‘axial coding’. Whilst the aim of this analysis was not to quantify findings, this process makes visible patterns within and between data in relation to shared meanings, allowing identification of views and experience that are held in common and also counter views and experience (Strauss and Corbin, 1998).

Data were first analysed in sets relating to the status of the participants. This allowed subsequent comparative analysis across themes in different data sets. This process allowed triangulation of findings from different participant groups and also triangulation of data generated from different methods. In the second year of the study findings from the qualitative aspects of the study where triangulated with findings from the national survey. The longitudinal approach to study allowed triangulation between findings from different years. The process of triangulation between data sets and methods is recognised as increasing the validity of findings (Gliner, 1994).
3.3.6 Presenting the findings

Findings are presented in separate chapters relating to the overarching themes of the evaluation. Findings in relation to how these categories are experienced and understood are organised around key themes and presented from the perspectives of the different participant groups. The interpretation of participants’ views and experience is illustrated with direct quotes from interviews and focus groups. These quotes are tagged to indicate the status of the participant making the quote and in most cases, their deanery. To protect confidentiality, the different deaneries have been assigned codes. However, deanery codes were not tagged to the quotes from deanery leads to ensure anonymity and confidentiality.

The volume of qualitative data generated by the study meant it was necessary to be selective in deciding which quotes to use. Some quotes have been selected to depict a commonly held view or experience. In these instances this is indicated in the text and, for example, the quote could be identified as illustrating a view that ‘most’ or ‘all’ participants subscribe to. Other selected quotes are used to illustrate differences between participants’ views and experience; an appropriate indicator precedes these, such as, ‘mixed views were held’ or ‘some considered’. Some selected quotes illustrate a particular view held by a single participant or an issue that emerged within a single deanery. In these cases this is specified in the preceding text.

The aim in presenting the findings is to faithfully reproduce how the new training curriculum is experienced and understood. The quoted material allows the reader to see how interpretations have been reached and to judge their validity (Green and Thorogood, 2009). Evaluation reports have been sent to participants providing opportunity for their comment on the reliability and validity of the data and findings. The feedback received has confirmed these constitute recognisable accounts of their training experience. The academic rigour of the design and conduct of the qualitative aspect of the study contributes to its value in informing the continuing development and implementation of the curriculum.

3.4 Ethics

This research was conducted in accordance with the ethical guidelines of the British Educational Research Association (www.bera.ac.uk/publications/guides.php). Ethical approval was obtained from the University of Birmingham Ethics Committee. National Research Ethics Service (NRES) approval was not required.

3.5 Presentation of Findings

Findings from each year of the Evaluation have been presented in separate Interim Reports (Bedward et al, 2008; Bedward et al, 2009). Whilst key findings from the earlier years of the study are referred
to in this report, its main focus is to present findings from the final year of data collection. Chapters 4-6 present findings from the qualitative aspect of the study from the focus groups and interviews. These focus on the key themes explored during the study. Chapter 4 discusses findings based on deanery leads’ views and experience of the implementation process. In Chapter 5, findings on how the three main constituents of the curriculum; the written curriculum, the ePortfolio and the assessment system; are experienced are presented and discussed. Findings that discuss the experience of key stakeholders and their views on the curriculum’s fitness for purpose are presented in Chapter 6. In Chapter 7 findings from the analysis of national assessment data collated between 2007 and 2010 are presented. In Chapter 8, the findings are discussed in relation to the research questions. Conclusions and recommendations are made in Chapter 9.
4. FINDINGS: IMPLEMENTATION

This chapter presents findings on stakeholder views and experience of the process of implementation of the new training curriculum for general practice.

4.1 Initial Implementation

The findings from the first year of the evaluation document the process of implementation and how this was experienced by deanery leads, programme directors, trainers and trainees. The new curriculum was viewed as being very much centrally constructed and centrally imposed. Trainers, in particular, felt alienated and many did not have any sense of ownership of the written curriculum, the ePortfolio or the new assessment system. Most stakeholders considered the ePortfolio when first launched was not fit for purpose and this was exacerbated by a lack of training, support, information and resources to support its use. Another major concern during the first year of implementation was the requirement for ST3s to complete their final year of training in accordance with the requirements of the new training programme. This meant that the process of implementation was particularly stressful for ST3s and their trainers who had an extremely limited timeframe in which to assimilate the meaning of the changed curriculum and to translate this into practice and fulfil the requirements of a new system of assessments.

During the second year of the study, findings suggest an easing of some of these initial anxieties. The focus groups demonstrated increased confidence among trainers with the requirements of the new programme and an improved understanding of the expectations of their training role. However, the learning process was clearly continuing and many still reported feeling unclear about certain aspects of the programme and their role as trainers. There was an interesting contrast between how new and established trainers related to the new curriculum. New trainers tended to be positive and welcomed the increased structure that it provided. Experienced trainers found the process far more challenging and tended to be critical of the changes and how these impacted on their previously established roles and increased their workload.
4.2 Findings from Year 3

The interviews with deanery leads held during the final year of the study provided an opportunity to reflect on the implementation process and to consider the deaneries’ roles in this process. Their views and experience of the challenges posed by the implementation of the new training curriculum and their efforts to address these are also considered.

One of the main challenges faced by the deaneries was dealing with the scale of change involved with the implementation of the new training curriculum. Previous experience of implementing and managing change in general practice training; for example, the changes that occurred when summative assessment was introduced during the 1990s, was felt to have eased the process of implementing the new system of assessments. However, whilst the introduction of the ePortfolio and a written curriculum were generally viewed as positive developments, “getting trainers on board” and ensuring they understood how to deliver the new curriculum was a major challenge. During the final year of the study, there was a general view that trainers had now developed greater acceptance of their changed roles:

Getting the trainers on board has been a huge challenge. New trainers are fine, but experienced trainers hugely resisted the changes and the fact they felt the expectation of how much work they were going to have to do was going to increase with all the ePortfolio checking they were going to have to do. And the fact they didn’t like doing the assessments very much. It took away the role they saw for themselves as trainers. But it’s got better. Year 1 was dreadful but now in Year 3 you have only a few disgruntled people (Deanery Lead).

Everybody is a little bit resistant to imposed change. By human nature if you feel you have ownership of change, you’re much more liable to embrace it rather than if it’s been imposed. This has pretty well been imposed. It’s a little harder to get people on board, but I think we’re moving forward. People have become used to their roles and people who have been in post for a long time find it difficult to change. But we are seeing the majority of our colleagues moving in the right direction; moving forward (Deanery Lead).

Resistance to change among trainers was considered to have been exacerbated by what the leads considered had been the “rushed launch” of the ePortfolio which had led to its introduction “before it was fit for purpose”. The leads described the anger and hostility that was expressed towards the ePortfolio and some highlighted how the deanery was often blamed for the problems that were being experienced by trainers and their trainees. In addition, most felt that reluctance by some trainers and trainees to embrace the ePortfolio, even after significant improvements had occurred, could be attributed to the negative experience of implementation:
A lot of negative flack was thrown our way because we were seen as being the bad guys when in actual fact it wasn’t our fault. We were just trying to do what needed to be done to move forward with this new curriculum and the nMRCGP (Deanery Lead).

The problems with the ePortfolio caused uproar with our trainers and at the deanery the number of calls we had to deal with, and trainers still go on about how badly the ePortfolio does this, that and the other, when actually it’s improved significantly (Deanery Lead).

Because of the degree of change involved in the new curriculum, ensuring that trainers understood the language and the meaning of concepts was also a significant challenge for the deaneries. One example of this was dealing with the meaning of the different competences associated with different aspects of the programme:

I think people have been slightly confused with the sudden change and all of the different terminologies that are going around of what are the competences expected within the MRCGP exam structure and the core competences of general practice that are in the curriculum. Sometimes it’s been the nomenclature that’s been the difficult bit to get our head round as well as just what we’re actually doing (Deanery Lead).

The deaneries addressed these challenges by providing training for trainers, including trainers’ conferences and trainers’ away days and supporting trainers’ groups. Programme directors often had a main responsibility for delivering this training, but first they needed to assimilate the meaning of the changes for themselves. Some leads were critical that communication and information from the College “wasn’t as good as it could have been”. This meant that at times deaneries were uncertain of what they should be doing:

We had to do a bit of mind reading as to what we were supposed to be doing and there wasn’t an awful lot of guidance. There was guidance, but it wasn’t easy to follow (Deanery Lead).

The start was a bit of a struggle, mainly I think, because the information we got, certainly initially with the ePortfolio, was not specific enough. We weren’t able to tell our trainers what it is that they wanted – what makes a good response in a log entry? What makes a good educational supervisor’s report? What are they looking for? And as we started to get a little bit more clarity ourselves within the deanery and get a little more clarity from the College, it has become easier for us to deliver better reports (Deanery Lead).

As greater clarity of meaning was achieved and the deaneries gained a better understanding of how the new curriculum should be translated into everyday practice, more material to support implementation was produced by the College and the deaneries were able to develop their own
materials to provide guidance to their trainers and trainees. A further tier of this diffusion was considered necessary by the deaneries to ensure GPs who were not trainers were made aware of the changes.

Another major challenge for the deaneries was “getting the hospital supervisors to deliver the GP curriculum”. This involved encouraging consultants “to deliver teaching that actually looks at the GP curriculum, not their specialty curriculum”. In all of the sample deaneries during the 2010–11 training year this process was ongoing and viewed as a continuing challenge. The written curriculum was considered helpful in supporting this and the deaneries had commenced mapping job descriptions in key specialties to the relevant statements. This process was described by one of the leads:

> We’ve produced some template job descriptions within the deanery that actually are mapped to the GP curriculum. So for example, for medical posts someone has gone through the curriculum and outlined the sorts of things in the curriculum you would expect a medical post to address. (Deanery Lead).

Whilst all of the deaneries have provided training in assessment for hospital based clinical supervisors, in some of the deaneries this training was now being extended to other aspects of the curriculum, including teaching them how to log on to the ePortfolio and comment on learning log entries. This training was still at a very early stage and by the third year of the study had not been rolled out to all of the hospital supervisors in the deaneries.

The decision during the first year of implementation that ST3s were required to complete their final year of training meeting the requirements of the new curriculum also posed a major challenge. Leads considered this resulted in the curriculum being “diluted”, because allowances had to be made for the newness of the system. This led to an expectation among trainers in subsequent years that “we’ll let them pass even though they haven’t actually done exactly everything they were supposed to have done”:

> The laxness of how we had to treat the ST3s in the beginning has followed through and it’s made it more difficult to implement all the changes as rigorously. For instance, in the first year we had to be very flexible about things like had they recorded 12 out of hours?; had they learning log entries in every single curriculum area?; and because we were flexible then, the trainers went on to expect us to be flexible every year. And even this year, I’ve had arguments with trainers who’ve said: ‘Well, you let them go through two years ago’: ‘Yes, but two years ago, they only had a year. This lot has had three years’. So it actually set an expectation that shouldn’t have been set. (Deanery Lead).
A further challenge for the deaneries was managing other changes that were occurring at the same time the curriculum was first introduced. In some of the deaneries there was an increase in the number of doctors undertaking GP specialty training which coincided with implementation of the new training programme. This meant that deaneries needed to find more training practices and more hospital posts, which in some cases proved difficult. Another deanery had to manage staff changes at a senior level during the early stages of implementation.

4.2.1 Conclusion

Three years on from the introduction of the new curriculum, deanery leads reflected on the overall success of the implementation process. The new training curriculum was introduced using a diffusion based model of implementation that was dependent on deaneries to support this process by providing guidance and training. Whilst this experience had been “turbulent at first”, particularly because of the pace of change and a lack of clarity in some of the requirements, they considered they were now “entering a period of calmer water”. The leads felt mistakes had been made during the initial stages of implementation, particularly in relation to the launch of the ePortfolio. However, on balance, they considered that the overall process had been successful. Most acknowledged that it had required a huge amount of hard work and had taken a full three years to get the curriculum in place and for assessments to be used in general practice “largely as they were designed to be used”. The leads considered that the new curriculum was now beginning to bed in and that trainees and trainers had adapted “remarkably well” to the changed system.

Whilst most trainers were viewed as having gained acceptance and understanding of the new curriculum, the deaneries’ ongoing provision of training for trainers to support the delivery of the new curriculum and extending this to their non training partners was considered necessary. The deaneries’ work with hospital based clinical supervisors was still at an early stage and the process of ensuring that the curriculum is used effectively within the hospital based component of training was viewed by leads as a continuing element of implementation.

The fact that implementation occurred on a national basis and that “all the trainees in the country felt they were being treated in the same way” was recognised as a significant achievement. However, the scale of change and the ambitious timescale for implementation had posed many challenges for the deaneries. Whilst the leads identified areas where they considered further improvements could be made, there was a strong view that these should be achieved by “evolution rather than revolution” to allow more time for the process of assimilation.
4.2.2 Recommendations

- The process of implementation should be viewed as ongoing and support and training is still necessary to ensure the intended meaning of the new training curriculum is assimilated and translated into practice.
5. FINDINGS: HOW THE CURRICULUM IS EXPERIENCED

This chapter presents findings on the experience of deanery leads, trainers and trainees of working with the new training curriculum and its three main constituents: the written curriculum, the ePortfolio and the system of assessment.

5.1 The written curriculum

Findings from the first two years of the study demonstrated that many trainers and trainees had not found the written curriculum accessible. In particular, during the first year of implementation, a general view expressed by both trainers and trainees was its content was overwhelming. Most had not identified how to use it effectively to support teaching and learning. During Year 2 of the study, there was evidence from both the focus groups and the national survey that greater use was being made of the written curriculum. Increased understanding of the structure of the core curriculum and supporting statements was evident among some GP trainers, especially those who had received training in its use. However, many still felt it was “inaccessible” and were uncertain of how to use it to support training. Trainees generally, continued to find the content of the written curriculum overwhelming and most reported using it only to link their log entries to the curriculum headings.

5.1.1 Findings from Year 3

Deanery Leads’ Views

Findings from the interviews with deanery leads showed that they welcomed the development and introduction of the written curriculum for general practice. They considered it was “helpful” in providing a “clear indication of the breadth of general practice” and “a good description of the competences required to be a good GP”. They also felt that it ensured that the coverage of the new training programme was more comprehensive as it prevented the gaps in learning that often occurred with the previous system:

I think it’s been very helpful in ensuring that there aren’t gaps. I think that was one of the problems with the old system; there could well have been gaps in areas of knowledge, particularly in clinical areas and things like that that were never addressed. I think the new curriculum certainly ensures that all those areas are covered (Deanery Lead).

Most recognised its potential beyond GP specialty training and considered, in particular, it was a valuable resource to support “significant learning post CCT”, especially for continuing professional development (CPD) and revalidation.

The leads acknowledged that the process of facilitating better understanding of the written curriculum among trainers had been gradual and that many had been initially daunted by its size.
However, they considered that most trainers had now embraced it and had achieved a good understanding of its potential and how it could be used to support training:

I think they've embraced it. There are very few that haven’t. I think over the past three years there’s been an increasing understanding of what the potential is and what the curriculum means. Year on year, I see a deeper understanding and more effect certainly in terms of GP trainers (Deanery Lead).

Whilst the written curriculum was viewed as having impacted on training by ensuring that it was GP focused from the start and not just in the final year, as it previously had been, most of the leads considered that trainees were only just starting to relate to it. Whilst trainees were using the curriculum headings on a regular basis, it had taken time for the deaneries to develop the content of the teaching programmes for the trainees so that it included guidance on how to use the written curriculum:

We guide all our trainees at the introductory course on how to use the curriculum domains and essential features and we have a session on the new curriculum and we guide them towards the ‘Condensed Curriculum Guide’ for that. I think it has been positive. It gives you a definite framework to work to and check to see if you’re on the right track. We try not to get them to use it as a complete tick list, but just to direct the sort of things and the sort of way they should be learning. So it’s more a method of how they’re learning rather than just a tick list of what they have to learn (Deanery Lead).

All the leads recognised the potential of the curriculum in providing a framework for the hospital component of training to ensure it provided relevant focus on general practice. However, the leads acknowledged that this was still very much “work in progress” and their development of job descriptions mapped to the curriculum was one aspect of this process.

**Trainers’ Views**

Findings from the focus groups held with trainers and programme directors during the final year of the evaluation showed increased use of the written curriculum to support teaching and learning. Some trainers pointed out that the written curriculum was helpful in outlining what needed to be covered during training, and many discussed how they used it as a reference document for planning teaching days and tutorials. However, most considered their trainees were still not making regular use of it. These views are captured in the following extracts from a focus group discussion:

I think from a trainer point of view, and as a (programme director), the curriculum document is fantastic. I mean for planning sessions that bible is brilliant. But the trainees haven’t got that and they don’t read it, but yet there is a huge amount of information and resources in there that they
could help themselves with, and they don’t do it. I’ve been a (programme director) for as long as the curriculum has been in, so it’s been wonderful to have the sort of tick box that you think, ‘Yeah, I’ve done that, done that’. And there are the resources that you can feed into, but I’m not sure how much the trainees get out of it (Trainer and Programme Director, Deanery 5).

I think from a programme director’s point of view, what it has helped is to actually start to say, ‘Well, where does someone learn this bit and where does someone learn that bit?’, and then start looking at the bits they won’t get in a hospital job, that they won’t get in general practice and then say, ‘Well I can focus my teaching on those areas of the curriculum’... I think you need a framework – we all had our own to some degree beforehand – but this is a universal framework. ... I think the great relief will be when learners start to look at it that way and say, ‘Well, so how am I going to make sure that my learning is wide enough? I know I can get this bit of the curriculum here, but where am I going to get this, that and the other?’ (Trainer and Programme Director, Deanery 4).

Trainers also described using the curriculum statements to develop job descriptions for hospital posts that provided guidance for hospital based clinical supervisors and for trainees in these posts:

I’ve a method of using it where I also try to draw up job descriptions and that’s really helpful because I can literally cut and paste and also I can show to the consultants this is what you need to do. So that’s been very helpful. Until I got into it I was a bit, what’s all this about? Now I can see the benefit of why it’s there (Trainer, Deanery 4).

Actually, in the past, we thought we knew what we needed to get out of the hospital posts, but we didn’t really know. And it’s actually focused us on looking at that and saying, ‘This is what we need to get out of a hospital post’ and ‘This is how you are going to do it’. And it’s putting the onus on the consultants to say, ‘This is what we’re going to deliver’ - well hopefully. We’re still in the early stages, but I think it’s going to be really useful (Trainer, Deanery 5).

Trainers identified other ways they were using the written curriculum to support teaching and learning. These include:

- Learning outcomes are cut and pasted and sent to speakers to ensure these points are covered when they make their presentations on certain topics at study days.

- Providing guidance for cases to be used in CBDs that focus on curriculum areas not covered.

- The learning objectives of the curriculum can be used as a self assessment tool to highlight what is known and where there are gaps in learning.
Whilst some trainers felt that the process of linking log entries to curriculum was helpful in drawing attention to areas where trainees had not had a great deal of exposure, many were critical this was of limited value as many trainees were not linking their log entries to the correct curriculum areas. This is illustrated in the following quotes:

I think one of the things you need to be careful about is trainees when they link their log entries to curriculum areas and you can very quickly tell that this trainee hasn’t read anything about the curriculum because for every single consultation they’ll put ‘management’ when it’s nothing to do with management (Trainer, Deanery 5).

They tick hundreds of bits of the curriculum that they’ve covered by seeing a lady with a cough (Trainer, Deanery 5).

In one focus group discussion, trainers with roles on the ARCP panels highlighted the random nature of how links to curriculum headings were often made and how this process was often not fully understood by either trainers or trainees:

When you do the panels, people haven’t a clue what they’re ticking and why they’re ticking it (Trainer 1, Deanery 5).

That’s why it’s worth doing a panel, because I didn’t really know what I was doing and you sit three hours doing a panel and you very soon come up to speed on the whole thing (Trainer 2, Deanery 5).

Some trainers suggested that limiting the number of links that could be made from an individual entry could ensure this process was carried out more appropriately:

You are supposed to pare them down to just one or two topics, but they should limit the linking to one main heading – that would make it more sensible – certainly no more than two (Trainer, Deanery 3).

In another focus group, trainers described how the links made to curriculum headings developed a tick box approach to training in both trainers and trainees, with trainers checking the number of ticks in different areas and then directing their trainees to focus on cases that had not frequently been ticked.

Whilst trainers held mixed views on the value of the written curriculum, the focus groups demonstrate that many have now developed a greater understanding and have assimilated its
meaning into their understanding of general practice. Most now have greater confidence in using it to inform teaching and learning. This process is captured in a trainer’s description of his initial reaction to the curriculum and how he felt alienated by the process of its top-down implementation. He felt it did not correspond with his meaning of general practice but rather imposed a European model that he felt no ownership of. However, over time, he has now gained a greater understanding and acceptance of it:

I was initially quite annoyed about the curriculum because I thought it had been imposed on us by Europe and it was basically WONCA, and we had to fit in with WONCA. And I was actually really extremely annoyed that we weren’t actually consulted. It just seemed to be sort of parachuted in on top of us. But having worked with it over the last three years, I can now actually see that ... it is actually really very good and a very practical and pragmatic framework, which is actually looking at the areas that are very important in day-to-day practice. So I’m a sort of late, reluctant convert really, and I was sort of dragged into it (Trainer, Deanery 4).

As in the earlier years of the study, trainers’ main criticism of the written curriculum was its size. Some felt this, combined with its language made it less accessible to trainees. Most trainers considered trainees were not engaging with it in a meaningful way, because it did not provide ‘a quick fix’; that is, immediate access to what they needed to do to get through their training:

It’s a very broad curriculum but it’s totally irrelevant to most of the trainees because you cannot and will not cover everything that is in that curriculum. There’s definitely stuff I’ve never heard of. It’s a useful document in that it gives you resources and places to look and guidance. It’s a useful document to help a learner prepare, to dip into and dip out of, but otherwise, it’s a wonderful creation ... but what drives the learner is the assessment process, so what they’re looking at is how they’re assessed at CSA and how they’re assessed at the AKT, the CBDs and COT. And that’s what they’re learning to pass and the rest of curriculum goes by the outside (GP Trainer, Deanery 3).

I can only think of two trainees we’ve had in the last three years who’ve used the ‘Condensed Curriculum Guide’ and they used it in a very systematic way, but they didn’t actually need it, they would have passed it on day one. It’s too big a document. It’s too scary a document, but if you sit down and actually start following the links, you could be there for about 12 hours reading about just one particular aspect of ENT or whatever. They can’t absorb so much information – they’re looking for quick fixes (Trainer and Programme Director, Deanery 3).

Other trainers considered trainees were keener to focus on clinical concepts rather than philosophical ones:
I’m not quite sure if this is what you’re getting at, but anything that’s strictly clinical and objective in that way then they’re absolutely fine with, but anything that’s what they call ‘woolly’ they dismiss or find it difficult or don’t want to engage with. So I think they’d say that quite a lot of it is wooly actually (Trainer, Deanery 5).

And I think that’s it, the term ‘woolly’ and ‘airy fairy’; they can relate to the clinical stuff much better than this concept of ‘Being a GP’ and... I spent time on the course trying to explore the core statement, but they just wanted to focus on the clinical stuff and the guidelines! (Trainer, Deanery 5).

When asked how the written curriculum could be improved, as in previous years, most trainers felt that a more concise user-friendly version was needed and that the language used should be clearer and less complicated. Most considered the ‘Condensed Curriculum Guide’ was not concise enough. Some trainers were also critical of confusion between “domains and outcomes, and trees with bits at the top and roots” finding this was “too complicated”.

**Trainees’ Views**

As in earlier years of the study, the findings show that for most of the trainees their familiarity with the written curriculum was limited to the curriculum headings that they linked to their log entries and its relationship to other aspects of the assessment system.

The assessment makes us, you know, touch on all parts of the curriculum, which we have to do, whether through CBDs or reflection (ST3, Deanery 1).

Several trainees found this was useful for identifying gaps in their learning:

I suppose having it on the ePortfolio you know you’re covering it. By filling it in, you can see where you’ve got zeros or low numbers, so I suppose you are exposed to it more than you think you are, but as we’re not quite sure about it, it just shows that perhaps we don’t know (ST2, Deanery 5).

Some ST3s had received guidance from their trainers to use the curriculum in this way to check for any gaps in their learning. Whilst some trainees were using it in this way, not all considered this was helpful:

It was suggested by my trainer to have it as a reference for ideas like practice management. If you have got gaps in your curriculum headings, like learning disability or something, and you are not really sure how you could beef that up on your ePortfolio. It is quite good as a reference manager because you can look it up and it has suggestions (ST3, Deanery 2).
I think, you’re meant to have entries in each of the curriculum headings but then it’s hard to see quite how many. Are you meant to have a certain number? What are you trying to achieve from that? Because you don’t see much genetics in primary care and it’s just, I’m not quite sure where it fits and what the purpose of having such a vast curriculum is but that’s the only thing I’m slightly confused about really (ST3, Deanery 4).

The focus groups during the final year of the evaluation were held mainly with ST3 trainees, who had trained throughout with the new curriculum. Yet many of these trainees had not accessed the written curriculum at all during their training, and among those who had, most felt it had not met their expectations of highlighting areas they needed to learn:

I think the curriculum really it’s quite vague. It just says about communication skills and stuff like that and doesn’t really specify any specific things that you should learn with certain conditions, or about treatment and management around certain conditions (ST3, Deanery 1).

I think that there are some things that are totally essential, and I think that it would be no bad thing if there were a list of things: ‘You must know how to manage this’, ‘You must know how to’ - like certain emergencies or something like that (ST3, Deanery 1).

In this year’s focus group discussion, the ST3s’ descriptions of feeling overwhelmed and deterred by the size of the written curriculum were virtually identical to those of trainees at the start of the study:

I agree with pretty much all that has been said before, the content of the curriculum is so vast that it is just unmanageable. At the start of the three years it is just too much of a task to look at it, because you scare yourself. Now at the end of the three years I haven’t needed to really specifically look at any part of it (ST3, Deanery 2).

However, others were able to reflect that its use might have been beneficial in their hospital posts:

I’ve never read it. Thinking back, I probably ought to have used it in my hospital jobs, it would be quite helpful to focus my hospital job work towards learning certain things but I think you can get quite caught up in hospital, just filling a service role and you get the teaching you’re given and that’s fine (ST3, Deanery 4).

A few of the trainees had used the curriculum statements in this way to inform their learning in hospital posts. Sometimes this was supervisor led:

I went to a hospice, … I’ve got a supervisor there …and she asked me if we had a curriculum for the end of life care, because she wanted to do a teaching session every week, to cover the
curriculum. So she’s gone through each bit and ticked off each area and that’s the only time I’ve
ever done that ...I didn’t do that that for the others (hospital posts) (ST2, Deanery 5).

The only time I found it useful was in the hospital posts to give me some direction or guidance as
to what I needed to know in hospitals when there was no one telling me what I needed to know.
That was the only time I found that it was useful, particularly when I was working in ENT (ST3,
Deanery 2).

Some trainees had used the written curriculum as a guide for their AKT revision:

I think if you’re revising for the AKT, that’s the only time that I’ve looked at the ‘Condensed
Curriculum’ to try and go through and see if there’s any pointers and that kind of thing (ST3,
Deanery 2).

Several suggestions were made by the trainees of how the written curriculum could be made more
meaningful in their training. Some suggested that more direction was needed at the start of training
to make them more aware of the curriculum and how it could be used to support their learning:

Probably just make us more aware of it. I know there’s a lot to take on when you start your ST1
and even the ePortfolio itself is giving you tasks to start with, but perhaps if we were just made
aware, ‘This is what you should be looking at’, perhaps ‘Print this out’ or ‘Use this’ (ST2,
Deanery 5).

Others suggested that the development of “a quick reference guide like they have in NICE
documents” would be helpful:

I think if you scored it at the beginning and it’s what you needed to learn that might be more
helpful, because we wouldn’t have needed to learn much medicine, or probably we were quite au
fait with paediatrics by the end of the hospital post, but what you wanted to learn was
dermatology, ophthalmology, those sorts of things and the other types of things that you can only
learn when you’re in the practice (ST3, Deanery 4).

One trainee described how the value of the core curriculum, ‘Being a GP’ was highlighted at a
conference she attended, however, she still admitted to never having read it:

I went to a talk, one of these annual conference things about GPs, and someone stood up there
and said “who has read ‘Being a GP’?” And my hand wasn’t up at that point. Then he said
“Everyone who hasn’t read it should just get up and leave because you don’t deserve to do
general practice”. That is how important it is apparently to read that (ST3, Deanery 2).
5.1.2 Conclusions

The findings from the interviews and focus groups demonstrate that learning the meaning of the written curriculum and translating how it can best be used to support general practice training has been a gradual process over the three years of implementation, and this process is continuing. This year, whilst some trainers still described finding the content of the written curriculum overwhelming and too complex, there was evidence of their increasing recognition of its potential and its more widespread use. More resources had been developed to support its use, and these generally were perceived as helpful in making the curriculum more accessible. Some deaneries had produced guides for trainees and hospital based clinical supervisors that identified areas of the curriculum to focus on whilst in particular hospital posts. By summarising key headings, aims and learning objectives, this had made the curriculum more accessible and manageable, and those using it to support training in hospital posts had generally found it helpful.

Many trainers described how they were using the curriculum to support teaching and learning and were increasingly finding it a valuable resource. Several trainers described using the written curriculum to identify learning needs in their trainees, particularly those experiencing difficulties. However, whilst there was some evidence of increased use of the written curriculum statements by trainees, for many its use was limited to linking log entries to curriculum headings. Deanery leads and trainers considered it was under-utilised by trainees. The increased confidence in their use of the written curriculum that was apparent among trainers and programme directors this year and the examples they gave of how they were now using it to support teaching, suggests that more time will be required before most GP STs assimilate its meaning and learn how best to use it to support their learning. The findings suggest that this year’s ST1s received more guidance and support on the written curriculum than ST1s during the first two years of implementation. Hopefully this will ensure that like the trainers, the trainees will in time develop more confidence in using the written curriculum effectively to support their training and their continuing development as GPs.

5.1.3 Recommendations

- Work on making the written curriculum more accessible for all GPs and those involved in general practice training should continue.

- Clearer guidance needs to be provided to trainees and educational supervisors on how log entries should be linked to curriculum headings.

5.2 The ePortfolio

Findings from the first year of the evaluation demonstrated that difficulties experienced by trainers and trainees in using the ePortfolio were a major concern. Most considered it was not user friendly
and many technical problems were reported. These difficulties were compounded by its late introduction, insufficient training and the requirement it had to be used in day-to-day training. By Year 2, many of these difficulties had eased and both trainers and trainees were generally more comfortable and competent in its use. However, many still felt that further improvements could be made to make it more user friendly. There was more evidence of its use in supporting teaching and learning, and it was recognised as an effective tool for assessment and providing a record of learning.

5.2.1 Findings from Year 3

Deanery Leads’ Views

The interviews with the deanery leads during the final year of the evaluation highlighted how they valued the ePortfolio as “an essential part of the system of general practice training”. The leads recognised that because of the expansion in the number of trainees it would have been impossible to have delivered the new training curriculum without this electronic tool. Whilst all of the leads were critical of how the ePortfolio was introduced and considered that initially it was not fit for purpose, in general, they considered that significant improvements had been made so that it was now “more user friendly”, and did “the job reasonably well”. The difficulties associated with its initial implementation contributed to its “fairly negative impact” on training and a hostile reaction from both trainers and trainees. The leads considered this had now “settled down” and both trainers and trainees were developing a greater acceptance of it:

There were always going to be a lot of problems with its implementation. It’s a change and people hate change and it was very, very badly released and it was late coming out. Our trainees started in the August and it came out in the October. There were loads of problems with it. It needed more of a testing before it was released on the world and I think it got a bad press because of that. But the new trainees aren’t aware of those problems, so I think it has become more accepted (Deanery Lead).

The leads considered that the ePortfolio had now made a positive impact on teaching and learning and that the PDP and learning log provided a good way of trainees demonstrating “learning”, “progression with learning” and “competence progression”. The leads acknowledged that making and commenting on entries was a time consuming process. Some considered that although the process was meant to be trainee led it was in fact deanery led, certainly at first when the leads had to ensure that both trainees and trainers were meeting the requirements of its intended use. The ePortfolio was viewed as meeting one of the requirements of the new curriculum to encourage reflective learning:

So the fact that the trainees need to do reflective log entries, the fact that supervisors should be reading these log entries and giving comments has raised the whole discussion of what is
reflective learning and how do you demonstrate competence through reflection. I think the ePortfolio has been the tool by which that aspect of the curriculum has been effectively delivered (Deanery Lead).

Another function of the ePortfolio that was welcomed by deanery leads was its facilitation of contact between educational supervisors and their trainees during the hospital based component of training. This enabled trainers to gain a better understanding of what their “remote” trainees were learning in their hospital posts and a lack of entries could alert them to potential problems that could be addressed. It was viewed as providing trainers with “a much better overview of the curriculum and the learning needs to meet”.

Most of the leads considered that the ePortfolio could become a valuable tool to support lifelong learning for GPs. In particular, its impact on the development of reflective learning and the concept of demonstrating competence through reflective learning was viewed as being helpful for revalidation.

**Trainers’ Views**

During the final year of the study, most trainers now seemed conversant with the use of the ePortfolio. Many described the initial difficulties they experienced learning how to use the ePortfolio when it was first introduced and there had been insufficient support, training and guidance available:

I think that the ePortfolio has been incredibly difficult to learn in spite of good attempts by the deanery to teach us that. And I think one of the reasons is that you couldn’t actually get in and play with it and you can’t learn until you actually start using it; it’s the sort of thing you have to learn by doing (GP Trainer, Deanery 4).

One Programme Director described the difficulties in learning how to use the ePortfolio, and also the difficulty of training new trainers in its use. Once these difficulties had been overcome, she could recognise its potential:

The ePortfolio has probably been the best thing. But I think ... if you’re not a trainer, and if you’re not looking at it every day – it probably took me twelve months easily for me to get my head around it – and that was looking at it nearly every day. God help ... (the Programme Directors) that aren’t trainers who haven’t got ST1s, 2s, 3s in their practice to really get a feel for it.

...Currently, I’m teaching on the basic trainers’ course and to try and explain to trainers who aren’t actually training now, who might not have access to an ePortfolio, it’s nigh on impossible to explain to them how it works. But I’ve been able to show them one of my trainees’ ePortfolios, and now we have got a dummy ePortfolio as well ... once they see it and the penny drops, you realise how fantastic it is, and how it links into the competences and things, and how evidence is
collected. And once you see the bigger picture, I think it works very well (Programme Director and Trainer, Deanery 5).

Whilst generally, there was increased confidence among trainers in using the ePortfolio, new trainers described difficulties learning how to use it. However, three years into implementation, trainers acknowledged that this learning process can be better facilitated by the knowledge and experience gained over time by other local trainers and programme directors, the GP STs and deanery staff, and the increased availability of resources, including the RCGP Helpline.

Most trainers recognised there had been significant improvements since the ePortfolio was first introduced. Educational supervisors generally found the new educator’s notes helpful, and some reported using this to highlight positive and negative attitudes and behaviour in their ST3s. Some considered this was the only opportunity to record these observations as workplace based assessments (WPBAs) did not cover this aspect of performance:

We’ve had a learner who we had issues with and yeah, we used the educator’s notes having done that with him, taken him into a room and discussed it and then you log it that there was an attitudinal problem here (Trainer, Deanery 5).

Whilst many felt there was further scope for improvement particularly in making it user-friendlier, technical difficulties now were rare. There were also few complaints this year about the slowness of the system when completing reviews. However, user error still caused problems. Among the examples highlighted in the focus groups, a common problem was assessments not being recorded correctly on the ePortfolio. In some cases this had led to trainees being told they would not be signed off. Often these errors were made by inexperienced trainers, but sometimes they occurred when trainees failed to share their entries:

I had an example recently when I was doing a final assessment, where the trainee had put her audit in at the last minute, but she hadn’t shared it, so it didn’t then show on the ePortfolio when the Deanery came to look at it. And so that is something that I will remember next time, to remind them to make sure that they have shared everything that needs to show up (Trainer, Deanery 4).

Other errors could be attributed to the lack of ease in navigating through the system and many trainers considered that further changes were necessary to address this problem. Some of these difficulties are highlighted in the following extract from a focus group discussion:
Yes, you should be able to do what you want to do which is to look at the COTs and CBDs that you’ve done and the competences that you have hit on each. I can’t get it to do that at the moment.

And navigating round it is a bit tricky too. The last two reviews I did we ended up with three versions of it open in parallel and trying to desperately remember which one we were supposed to be working on. (GP Trainers, Deanery 4).

Some trainers considered the ePortfolio provided a valuable link with their trainees whilst they were working in hospital posts which enabled them to provide support that enabled the trainee to benefit more from this experience:

The ePortfolio itself is a very useful form of dialogue in this area anyway, in terms of the educational supervisor’s comments. If they are used well, I think that it’s a really good method, you know, sort of assisting people to get the most out of the hospital jobs (Trainer, Deanery 2).

There was increasing recognition among trainers of the potential of the ePortfolio as a learning tool. Some trainers considered the ePortfolio empowered the trainee as a self-directed learner and considered that it had “given ownership very much back to the learner”:

Compared to previous training when it was very much the trainer’s job to say that things were OK, now it’s theirs; and if they fail it’s their failure (GP Trainer, Deanery 5).

As in previous years of the study, much of the discussion about the ePortfolio was based on the learning log. Although different views were expressed on the value of completing the learning log, many considered this aided reflective learning:

I think the new programme, using the ePortfolio to do the learning logs is preparing them to do more reflective work, much more than in the past and then they get better and better as time goes on. My current ST3 was very reluctant to do the whole process at the beginning, but he has started to develop and get into it (Trainer, Deanery 5).

I like the way that during learning log entries they’re asked questions to help them reflect. So you’ve got to be a very poor trainee, I think, not to be able to know how to reflect (Trainer, Deanery 5).

Another trainer highlighted that the function of the learning log in facilitating reflection was of value to qualified GPs in the appraisal process:
I think that the ePortfolio is good; I think it’s very good to have a Learning Log, and to keep a record of what they’re doing, and actually let them think of the value of and to reflect on aspects of any learning that they’ve done. And in fact, I’m going to use that myself for my own appraisal process, and I use my appraisal very much like an ePortfolio, and I’ve found that very useful (Trainer, Deanery 4).

However, not all trainers considered reflection could be taught, and many reported difficulties experienced by their trainees in providing evidence of reflection in their learning logs. Several deaneries had produced their own guides for trainees aimed at facilitating a better understanding of reflection, and how to complete the learning log. The implementation of a minimum number of log entries was generally viewed as detracting from the quality of entries. Trainers also pointed out that reading through their trainees’ log entries and providing feedback was a very time consuming process, with several having to do this in their own time. Many considered that “too much time is spent trying to get encourage trainees to complete the learning log”. There was some evidence to show that learning log entries were now being used by educational supervisors to identify difficulties in trainees including during their training in hospital posts.

I think that just the lack of entries in itself, is an identifier, because often you see that they don’t understand the purpose of it, and they don’t want to engage. So, the particular trainee that I’m thinking of didn’t put many entries on, and then when sort of forced to do so, they put some on, but they weren’t reflective, they were just descriptive. And you could sort of see that they weren’t engaging with the process. And that, in itself, is a sort of red flag sign, if they’re not doing their log, and they are not getting on with it. In a couple of instances, you know, these are the trainees that we had picked up in ST1, because they didn’t do the log, they weren’t doing their CBDS, they weren’t videoing, and they had turned out to be trainees with difficulties (Trainer, Deanery 4).

Other trainers highlighted a problem in assessing trainees based just on the number of learning log entries. They viewed it was necessary also for the content to be considered.

(T)he trainee that I’m thinking of had actually got letters of commendation, because it had been a sort of counting the ticks in the boxes, and she had put a lot of stuff on. And it was only when I went through and I read them all, and I discovered that some of them were where she had had a conversation with someone senior, about not knowing a really basic medical fact, you know (Trainer, Deanery 5).

There’s a danger of just counting the entries, never mind the quality, feel the width (Trainer, Deanery 4).
Trainers considered that the ePortfolio was underused as a communication tool. The messaging system was considered “very basic” and it was felt that there was scope for improvement. Some trainers experienced difficulties finding contact details for their trainees when attempting to send new messages and thought this was too complex. No difficulties were identified in sending replies to messages.

**Trainees’ Views**

The majority of trainees who participated in focus groups were in their ST3 year. Most now felt conversant using the ePortfolio, but described earlier difficulties they had experienced learning how to use it and learning what was required in relation to their entries. Most ST3s felt things would be better for the next cohort of trainees:

> I think now for ST1s starting they will have a little summary of what the requirements are in terms of log entries, because we didn’t have a clue. That has caused problems over the years. So if it was just more clear that would help. I still think it is too burdensome the way it is, but at least if you knew what you were supposed to be doing from the start there would be a chance of getting it right (ST3, Deanery 1).

Trainees considered that developing an understanding of the requirements of the ePortfolio had been compounded by conflicting and changing information received. Over their three years of training, trainees reported differences between the guidance they had received from their educational supervisors, the deanery and the College. Trainees from two of the participating deaneries highlighted how problems had occurred when new guidance was issued on a minimum number of PDP entries per week, which in some case only came to light during the review process in their ST3 year. No technical problems were reported, but many considered further improvements were required to make the ePortfolio more user friendly and easier to navigate:

> It’s an interesting tool designed by the NHS. If Apple had designed it, I suspect we’d all be finding it a lot easier. I think it’s got all the things you need in there. Once you get the hang of it, fine. I’m sure this is partly because we’re semi-piloting things, but the trainers, like the trainees, don’t really know where to put things always or how to link things and I think it could be made a lot better. I think it could be improved. There’s lots of things they could do (ST3, Deanery 4).

Some trainees described errors they or their trainers had made:

> I got an e-mail on Friday, having submitted all of my stuff a fortnight before, to say that the dates that my trainer had put on for my review was incorrect, and this was going to be a showstopper; I wouldn’t get my CCT. Also, my trainer hadn’t realised that he had to click a ‘validate evidence’ button, and then tick a few boxes at the end of each. He had written a comment for every single
entry that I had done, but because he hadn’t done this, again, I wasn’t applicable to be put forward for CCT. (My trainer) has never been shown that this is what he was expected to do, and he felt very bad, because he wrote to the deanery on my behalf to say “Please let this not affect my trainee”. But I mean, it wasn’t his fault, it was completely the inefficiency in his training; as a trainer, he didn’t know that he had to do this (ST3, Deanery 1).

One trainee described her frustration at repeatedly pressing the delete button instead of the save button and losing her work. She considered these buttons were positioned too closely together and also that there should be an ‘undo delete’ button. Others highlighted areas where they felt the navigation process could be simplified:

It’s just the home page - I think it could be much clearer to navigate around rather than traipsing through all the various headings (ST3, Deanery 2).

The trainees welcomed some changes made to the ePortfolio, particularly now being aware of when their educational supervisors had commented on their entries. Most of the trainees were positive about the function of the ePortfolio as a recording tool:

Overall it’s great actually because it means you can put everything you need in one place and its evidence of your training (ST3, Deanery 4).

It’s a record, that’s how I view it, as evidence and that’s why I fill it in, because you can go back and say actually yeah, I’ve been taught joint injections and I’ve been taught child protection and all those things that you’ve got to be able to demonstrate; it’s a useful, secure, portable way of having evidence that you’ve been trained (ST3, Deanery 4).

Many trainees considered the ePortfolio had contributed to their learning. Some trainees valued being able to work independently on the learning log and that, unlike other assessments, this was not dependent on a supervisor being present or observing practice. The trainees held mixed views on the value of completing the learning log. Whilst some considered that the process aided reflection, others felt that they were simply documenting reflection that had occurred previously:

I’m not sure you learn a lot through it. It encourages you, but most of the time I do my reflective during work. I’m seeing a patient, I’m reading about it there and then and about the case and what I need to learn and I just write it down. But then I have to go and put this on to the ePortfolio. I’ve already done the thinking, but I have to write it down in a way that would look nice to show that them I’ve reflected. But I do tons of reflection; more than just the ones I put together in the log (ST2, Deanery 5).
I think it’s taught me how to write a reflective entry [Laughs]. It’s taught me how to write more creatively. Now that’s not to say I haven’t had some of those thoughts, but the need to write it down and put it into words doesn’t mean that that’s me learning reflection. I think I’ve learnt reflection, like you guys, through being taught and discussing things (ST3, Deanery 4).

Many felt the format was too structured and was not appropriate for all of the entries they made. Several suggested that a more flexible approach with open boxes would work better for them:

I think it would be better if they just gave a blank box which said: ‘log entry title’ then another box that said: ‘reflect here’ (ST 3, Deanery 2).

But just in terms of the way you have got to put your log entries in, having to choose was it a clinical encounter? was it something you read? was it a professional discussion? and then log it into all these separate boxes. Actually what I perceive to be much easier would be to have just a box I can put an entry in to say: ‘I saw patient X with these symptoms; it made me wonder if I knew enough about the condition and why’; end it with, ‘I had a chat with my trainer and I read a little bit about it, and these are kind of what I learned’ (ST3, Deanery 2).

There was a general consensus that completing the learning log was time consuming. During their hospital posts when there was limited access to the internet, many had found it difficult to complete entries on a regular basis. Some trainees described “marathon sessions” on the ePortfolio where they made multiple entries, usually in their own time, which they felt also distracted from its value. In most of the sample deaneries a minimum number of entries per week had been prescribed and this tended to be unpopular. Trainees, in general, agreed with trainers that it was the quality not the quantity of entries that was important, and felt a prescribed minimum number could distract from the quality:

The problem with things that are reflective is that to give them a quantity it actually sometimes goes against what you’re trying to achieve. When it’s done properly and you reflect on a patient you’ve seen, then it’s actually a very valuable thing and I think we all do that anyway, because if you’re a fairly average doctor, that’s part of how you work anyway. And so you want to put those ones down. That’s fine, the big cases, the ones you reflect on, you’re happy to type in and look back on. But … you need to put two in a week. Well actually some weeks you don’t have that. I think some weeks, thankfully, you do just have coughs and colds and earaches and you go to a course and you don’t really learn anything new because you’ve done it before. Well then, there’s not much point in putting that in, so that’s the only point I’d add to it really, that it’s hard to quantify it actually (ST3, Deanery 4).

However, it was the time required to complete the entries that was often considered the greatest problem with many trainees arguing that protected time should be available for this:
You’re expected to go home and do it in your own time, and frankly, you just don’t want to (ST3, Deanery 1).

I do have a suggestion about this ePortfolio; it works very well actually and especially for the log entries. I think if the three entries are mandatory and it reflects your practice, we should be given a time during the week to fill up our log entries as an administrative time or something like that (ST3, Deanery 3).

5.2.2 Conclusion

Findings from the interviews and focus groups demonstrate that the introduction of the ePortfolio was one of the most stressful aspects of implementation. Initial teething problems with the technology and the demands placed on trainers and trainees to learn how to use this tool as an essential component of general practice training, and the expectation that the deanery would provide support and guidance from a limited knowledge base, contributed to an initial hostile response to its introduction. This was reflected in the findings from the first year of the evaluation. Now trainees and most trainers have become more confident and competent in using this tool, the findings from this year’s focus groups suggest it is beginning to bed down and become more established in the culture of general practice training. Although there is further scope for its improvement, particularly in making it more intuitive and user friendly to prevent errors, there is increasing recognition of its potential beyond its function as a recording and assessment tool.

This year, more emphasis was placed on its value as a learning tool. Apart from facilitating reflective learning, deanery leads and trainers considered the ePortfolio supported trainees’ development as mature self-directed learners. It was felt this prepared them well for revalidation and appraisal. There was increasing evidence of its use by educational supervisors to support learning, not just in their feedback on learning log entries, but also in providing guidance on how to maximise learning in hospital posts. This also demonstrates that when used effectively, the ePortfolio can ensure greater cohesion and relevance between hospital-based training and general practice.

Making learning log entries was considered time consuming by trainees and having to make a minimum number of entries each week was often resented. Educational supervisors also pointed out the heavy time commitment required to comment on these entries. In general, both trainers and trainees agreed that focus on quantity can detract from quality. However, learning log entries were highlighted as often providing a better indication of poor performance than other assessments because of the more in-depth qualitative data they provide.
5.2.3 Recommendations

- Further development is required to make the ePortfolio easier to navigate.
- The problems associated with user error need to be addressed.

5.3 Assessment

Findings from the first two years of the evaluation showed that the new system of assessment provided increased structure to training, but was considered time consuming and was often viewed as dominating tutorials. Initial findings suggested that the number of assessments should be reduced, but there was greater acceptance that the number was “about right” in the second year. The effectiveness of workplace based assessments (WPBAs) was seen to be assessor dependent, and generally those carried out in general practice were more highly valued for the formative feedback provided. During Year 2, findings suggested that a culture of minimum effort was developing among some trainees in relation to assessment with a preference to focus on ‘safe’ low challenge cases. The clinical skills assessment (CSA) and applied knowledge test (AKT) were generally popular and considered fair by both trainees and trainers. Most trainers considered these were preferable to the previous system of summative assessment.

5.3.1 Findings from Year 3

Deanery Leads’ Views

Interviews with deanery leads held during the final year of the study explored their views and experience of the new assessment system. All of the leads spoke positively about the summative exams, the AKT and the CSA. The AKT was described as “fit for purpose”, and “very well run”. The CSA was considered to be a “robust assessment” with one of the leads describing it as “the best College exam there is”. It was considered to be a good assessment of consultation skills and “much less intrusive” than the previous video assessment. The leads also considered it provided a “good objective measure” and was “less artificial than the previous membership exam.

The leads were more critical of the workplace based assessments with one pointing out they did not offer any increased reliability or validity than the previous system:

I think the assessments themselves were thought up by a committee. They’re certainly no more valid or reliable, or have any published evidence to suggest they are any better than what went before. I think that’s a shame, particularly basing it on the Dreyfus model of novice to expert, but then cutting it down to ‘needs further development’, ‘competent’ and ‘excellent’, when the end point is ‘competent’ was a fundamental mistake. I think that has caused problems and made the assessments far less useful than they could have been (Deanery Lead).
The leads generally felt that GP trainers were now quite comfortable conducting assessments and they were using the consultation observation tool (COT) well. The COT was generally popular and was considered by one lead as “a fantastic way for trainees to actually think about what they’ve done”. The deaneries had provided training in assessment for trainers over the three years since the new assessments were first introduced. Calibration of assessment was one aspect of this training. Whilst the leads generally considered that they were achieving consistency in how trainers were using the assessments, the calibration of case based discussion (CBD) was identified as causing some problems:

From my experience and my discussions with trainer colleagues, the CBD is artificial. We have always done case analysis and discussing of both random and problem cases, but the scoring seems to be more artificial, more arbitrary. And when I do calibration at our trainers’ workshops, that’s always the one that produces much more discussion and disagreement as it’s much more difficult to calibrate. In addition, doing it in the way that’s recommended by the College, I would say that the majority of people have now diverged from doing it that way; picking cases in the way they suggest – most people have moved away to doing it in a hybrid way (Deanery Lead).

Another lead highlighted the failure of CBDs to demonstrate a trainee’s progress:

I struggle to look at a series of CBDs and be able to judge real progress. Whereas I can look at learning log and judge progress much better, particularly if the trainee is following the learning cycle of a learning log with a PDP followed by a further learning log. And you can follow the cycle and you can actually watch the progress of the trainee. The CBDs tend to be standpoint assessments – this is where the trainee is at this point – but actually to monitor progress from CBDs is much harder (Deanery Lead).

Whilst the deanery leads acknowledged there were some problems with assessments carried out in hospital posts, particularly the high proportion that were not carried out by trained clinical supervisors, they generally felt that the hospital based assessments had contributed to ensuring there was a general practice focus in this element of training:

I think it’s been much more positive on the hospital component because it has pushed people to be focused. Previously hospital assessments were very much based on whether the assessor liked the person or not; whether the trainee turned up and did the work. They weren’t focused specifically on the competences required. I think the hospital assessments have focused hospital assessors on that and the competences that are required (Deanery Lead).

The leads held mixed views on the value of the direct observation of procedural skills (DOPS). One lead considered that now that the DOPS achieved during Foundation training could be used for the general practice assessment this had removed the duplication that many trainees had disliked.
However, one lead considered that the DOPS should “either be scrapped or changed” because the skills were too simple and not all of the competences were clearly defined.

The multisource feedback (MSF) and patient satisfaction survey (PSQ) were considered by most of the leads as useful in identifying weaknesses in trainees:

The MSF can be very powerful and I’ve seen it used actually within our own practice very powerfully for a trainee who was slightly underperforming. The first MSF in the ST3 year enabled the trainer to then discuss the comments made ... the constructive criticism enabled the trainer to point out the deficiencies of the particular trainee, and that trainee was then able to remedy those and by the time of the second MSF there was an appreciable difference in the comments. It’s more useful in the GP setting than in the hospital. Again there’s a risk that they’re using peers and people who’re junior to them who’re only going to say nice things. And I think actually it’s the administrative staff – the non clinical staff who actually have got quite an important role in that MSF, because in many ways they are the lay people, more like the patient, and therefore have an opportunity to show how the trainee is impacting on them. It’s a much more powerful tool than I would have expected, especially for the trainee who has a few difficulties (Deanery Lead).

For trainees with attitudinal issues in the way they consult especially when they think they’re doing fine, the PSQ will show that the patient felt they weren’t listened to; I think that can be a hugely valuable tool. It gives the trainer a bit more weight as well (Deanery Lead).

Whilst other WPBAs were considered useful in identifying general weaknesses they were viewed as less satisfactory in identifying and addressing specific weaknesses. One lead discussed how he found it useful to use the Cambridge Calgary Consultation Model to identify and address specific weaknesses in consultation skills highlighted by the COT:

COT is a global consultation tool with fifteen criteria, which is fine. There’s nothing wrong with the criteria, they’re all basic criteria we would agree that are part of a consultation. But if I had a trainee that has any specific problem in those, I will certainly revert to the Cambridge Calgary Model and I will analyse according to that and do my skills teaching on Cambridge Calgary for a focused area of training need for a trainee. ... The Cambridge Calgary Model is better particularly for things like explanation skills or structuring skills. If you’re talking about holistic care, the term ‘holistic care’ is a very global term within a CBD. There are more specific tools that have looked at holistic care and all the attributes of holistic care, and it needs that level of detail to break things down into the small skill areas (Deanery Lead).

The deanery leads held different views on the formative value of the WPBAs. Some considered they provided opportunities for teaching and learning “if they were done properly” and that the
opportunity they provided for formative feedback was the most useful aspect of these assessments. One lead considered that there were problems with using CBDs formatively because this turned the assessment into a tutorial and impaired its value as a summative assessment. Another lead pointed out that the WPBAs had focused trainees on the competences they needed to achieve and whilst there were some merits in this, it had led to the development of a tick box approach to assessment:

For instance, trainees talk about ‘passing’ workplace based assessments rather than treating it as a formative development process. I think it’s perhaps pushed them into that area more than I would have wanted, though I think that was probably inevitable and I don’t necessarily think it could be done another way (Deanery Lead).

The deanery leads identified the following gaps in skills they considered were not sufficiently included in the new system of assessments:

- Clinical leadership
- Clinical governance
- Critical reading
- Practice management.

The deanery leads identified ‘community orientation’ and ‘fitness to practice’ as the two areas of the competency framework that were most difficult to achieve:

There are probably some areas that they don’t understand; ‘community orientation’ is tricky. Even ‘fitness to practice’ I think there is still a lacking of understanding of what those areas mean and how they’re documented. One of the areas that’s difficult for trainers to comment on is how to improve in ‘fitness to practice’. But ‘community orientation is probably the hardest one – I think most GPs probably find it hard (Deanery Lead).

**Trainers’ Views**

During the final year of the study, the focus groups and interviews with trainers demonstrated an increased understanding of the assessment system and that many had found ways of combining assessment with teaching. Trainers generally, considered the new system of assessments was preferable to the previous summative assessment. Its multidimensional nature meant that, taken together, the assessments provided a holistic account of a trainee’s performance and development where “the whole is greater than the sum of its parts”. Trainers’ main criticism, as in previous years, was that assessment now tended to dominate training. The requirement to complete WPBAs and regular reviews was considerably time consuming, and many trainers reported having to complete reviews in their own time as there was insufficient time allocated to training for these to be
completed during the working day. Trainers described how most tutorials were spent entirely on assessment and many considered that their trainees’ primary focus was on completing assessments.

Findings from the second year of the study suggested that a culture of ‘minimum effort’ in relation to assessment was developing among some trainees. This culture did not seem evident among trainees during the final year. However, trainers considered that the competence based approach to assessment had resulted in many trainees adopting a totally summative view of assessment and their aim was to gain the required minimum number of assessments, or to see the ePortfolio “turn green”:

The negative from my point of view is the perception, “Oh I’ve just got to make it go green on the portfolio”... So when they’re doing their COTs, their case based, whatever, when it goes to six ... they see it as an exit; “That’s the bar I’ve got to jump over” and when it goes green ‘I’ve jumped over the bar’. They haven’t all of them made the link that actually that wasn’t really competent. That’s where some of them struggle (Trainer, Deanery 5).

Well I’d like to move away from numbers, you know. ‘Have your competencies been demonstrated?’ - the archetypal one is out of hours. The first question we get on the 2nd of August is, “Twelve sessions, how many hours is it I’ve got to do? Is four hours a session? Is six hours?” “No, it’s when your trainer thinks you’ve demonstrated competences.” And that analogy can be drawn across all of them. I honestly think that the notion of seeing it go green is actually a little bit unfair on them because that’s what they work on. That’s one thing I’d like to see change ... they’re very blinkered; six is the hurdle and jump over it and I’m there (Trainer, Deanery 5).

However, some trainers considered that by providing clear objectives for GP STs from the start of their training, the new system of assessment provided additional motivation to learn, especially during the first two years of training:

Now, I teach the ST1s and 2s on a Thursday afternoon and some of them are coming in with, ‘We want to learn about this because we need to know it for our AKT. They’re not just all sleeping through it and using it as an afternoon off which one had the feeling before. But some of them are more demanding as learners because they’ve got those objectives to meet (GP Trainer and Programme Director, Deanery 5).

I think having competences there has focused the mind on what they actually need to do and what they need to achieve at the end of it. Whereas in the past, I think it was more woolly. I think we thought we knew what they needed to achieve, but by ticking the boxes you actually, hopefully, ensure that they are achieving those competences. ... I think having the competences there to focus on has probably been invaluable (GP Trainer, Deanery 5).
Findings from the second year highlighted that trainees were adopting a minimum effort approach to COTs and CBDs by selecting ‘safe’ low challenge cases for assessments. This year’s findings demonstrated that trainers had developed approaches to ensure that more challenging cases were used in these assessments. The provision of ongoing training in assessment was valued in addressing this problem and trainers highlighted that preparation was critical to maximising the challenge and the learning value of assessments:

I think as an experienced teacher you can always make a challenge out of any case. ... The way I do it is as they bring their cases I choose the one that’s interesting to me. And if they come along with four respiratory tract infections we’ll make it challenging. But the key to that is preparation, doing it properly, not doing it on the hoof, as I suspect some people do, but having it before, preparing your questions before, knowing where you want to go, looking at the curriculum to match up against what you’re testing their knowledge of. So if you’re well prepared, it’s fine (GP Trainer, Deanery 5).

I have told them exactly the framework that they should be doing, that they should be providing me with three cases about a week before, and to just choose one of them, or two of them. They don’t, they give me a name, just as it is. They haven’t given it any thought, and it’s not that interesting a case. There’s not that much meat on it, and it’s very difficult to score them. And having done a few of the training workshops, I’ve changed the way that I do it. And I now very much go to them “Well OK, I would like you to find some cases and when we start, I want you to tell me which competences you hope to demonstrate in those” to make them think about what they’re doing (GP Trainer, Deanery 4).

In general, trainers considered that the WPBAs provided value opportunities for teaching and learning. However, there were differences in how trainers used assessments formatively. Some trainers considered it was very difficult to follow the guidance for conducting CBDs and adopt a summative approach to the assessment with the only formative element being in the feedback at the end:

I should be doing it as an assessment tool and it turns into a teaching tool and I don’t know if that’s my problem or whether it’s unrealistic to expect us to do an assessment without teaching on the back of it. But I’m very confused about my role and probably the registrar is as well. It’s an awful waste of time if we’re not going to use it as a teaching tool (Trainer, Deanery 3).

Yeah, from the CBD, if you stick to the letter on how it’s supposed to be done, they know you’re going to have a section at the end where you do all that. You’ve got to remember all the things you want to talk about in giving feedback and I think that’s hard because some people want to feedback as they go along (GP Trainer, Deanery 5).
Some trainers applied formative and summative approaches simultaneously when conducting COTs and CBDs. Others made a distinction between ‘formal’ assessments that were recorded on the ePortfolio and ‘informal’ assessments that were not recorded. In these cases, a summative approach was adopted in formal assessments and a formative approach used in informal ones:

> It’s meant to be learning rather than assessment and it may be that it’s misguided not to put them on, but if it goes on the ePortfolio it’s a permanent indelible part of the registrar’s record. And I suppose, since there is a finite number they require, perhaps in a sense you want to make sure that you represent the best of their ability on what will be available to the ARCP and you don’t want it cluttered up with things (GP Trainer, Deanery 3).

The focus groups demonstrated that GP trainers were developing an improved understanding of the WPBAs and adopting approaches to maximise challenge and learning potential. Training in assessment and support from trainer’s groups had contributed to this increased confidence:

> We’ve had some calibration, and sometimes the schools vary but what I’ve been impressed about is even the new trainers are very on board with it, so I’m seeing a lot of work going into training them to be trainers on the basic trainers course (GP Trainer and Programme Director, Deanery 5).

> I think overall in our deanery there’s a fair degree of consistency and I think certainly more trainers are getting to understand it. I think things have changed over three years; people are much more familiar with it and are able to work it, and I think within the trainer’s group, certainly in my patch, there is a lot of discussion about calibration, and how we actually ensure we were all doing the same kind of thing (GP Trainer and Programme Director, Deanery 5).

However, concerns were still expressed about the value of assessments in the hospital component of training. As in earlier years of the study, trainers described how hospital based assessments were often graded ‘excellent’ and this made it very difficult for educational supervisors to get any real understanding of how their ST1s and ST2s were performing during their hospital posts. Whilst more consultants were now trained in assessment for GP training, and this had led to some improvements, many hospital based assessors had not received any formal training. A further difficulty highlighted in the focus groups was the high rate of staff turn over in hospitals, which meant it was difficult in some cases to establish and maintain relationships with hospital clinical supervisors. Also, trainers pointed out that many hospital based assessments were not conducted by the clinical supervisor but by junior staff. The general view was these were not reliable:
So if the standard of recording in hospital improved this would be more useful. You usually don’t get a true impression of what the consultant thinks. A lot of the WPBAs in hospitals are done by staff grades - basically their friends are filling them in for them - and then the consultant often fills it in at the end. And I’ve often had ‘excellent, excellent, excellent’, and in the comment box, nothing (GP Trainer and Programme Director, Deanery 3).

I actually think that the more useful hospital assessments are the ones that are done by the consultants. Some that are done by the more junior staff are a complete waste of time (GP Trainer, Deanery 4).

Generally trainers considered the new system of assessment was valuable in identifying gaps in learning and weaknesses in trainees and allowing these to be addressed at an earlier stage:

The poor registrars it gives you a yardstick by which to measure them and you’re forced to measure them against that yardstick, which I think is quite useful when you’re trying to decide whether they should progress or not. Previously that wasn’t there and we know that, on the whole, trainers didn’t fail registrars because it was very difficult to actually do that. It still not done greatly through workplace based assessment, but it’s easier to do if they are not up to where they should be. (Trainer and Programme Director, Deanery 3).

I think the assessment tools and the way the evidence is gained from them, I think that structure is quite good. And I think it does allow you to highlight clearly where there are big gaps in their learning and performance. I think that is one of the useful elements of it. If you speak to somebody repeatedly about a clinical case they may never take on board the social aspects – you can mention it to them, but you actually have a tool where you can say, “Look, you are not filling this box”. It’s a bit like that with COTs; you can say that you have not demonstrated this to me. For those trainees who don’t see it, it’s a really useful tool (GP Trainer, Deanery 3).

However, this was dependent on the assessment tools being used correctly and on assessors being willing to document their concerns. Most trainers considered that if the clinical supervisor’s report was completed well and with depth, this was valuable in identifying weaknesses at an early stage:

The evidence provided by hospital-based supervisors is often not particularly deep evidence, but when it is deep it’s very helpful. I’ve had one or two registrars who’ve had certain consultants who’ve written very pertinent comments, which have been very helpful (GP Trainer, Deanery 3).

However, trainers generally considered this was not the norm. In one of the focus groups trainers highlighted reluctance among some consultants to document their concerns about trainees in the clinical supervisor’s review:
(I was speaking) with a consultant yesterday ... he had picked someone up in ST1 and he was saying to me “I think that there are going to be real problems with this trainee. He’s been with me for 3 months, and I can see the signs; he’s not engaging. And I’m asking him to do things, and he isn’t doing it. Whenever I set up to do a CBD or a mini-CEX, he always thinks that he is too busy to come. He is not engaging”. So he has isolated that this is a trainee with difficulties. And I said to him, “So will you write that in your Clinical Supervisor’s Report?” “Oh, well I’ve written it out on a piece of paper”. And I said “It’s no good on a piece of paper, you have to write it in that report, and come to the Faculty and say those things” (GP Trainer and Programme Director, Deanery 4).

That’s interesting, because we had a similar thing, between two hospital disciplines with a GP trainee who had gone from Paediatrics to Geriatrics. And Geriatrics were suddenly presented with three trainees all at once, who all had major problems, and they had not been flagged up. And then when they went back to the Paeds job, where they had been before, they said, “Oh yes, we knew there were problems”, but they hadn’t communicated them. And when we asked: “Well, why haven’t you?” “Oh, we didn’t want to blot their copy books, you know, we wanted them to have a fair assessment at the next job” (GP Trainer and Programme Director, Deanery 4).

Because of this reluctance to formally record negative views on the ePortfolio, some trainers considered it would be useful for educational supervisors to have dialogue with hospital based clinical supervisors prior to completing their end of year reviews on trainees:

I have got a trainee that we’ve been quite worried about, and actually, very little came out on the assessments. It came out to some extent, on reading through the lines of the Learning Log entries. But it also came out, most powerfully, when I actually rang up the current clinical supervisor, just before I did the ST2 review. And I just gave him open questions. I didn’t say, “I’m worried about this trainee”. I just said, “What do you think about him?” And this guy came up, in a ten minute telephone conversation, with all of the concerns that I’d had. And I think if we could find ways of actually building that much more into the system (GP trainer, Deanery 4).

However, weaknesses in trainees could also be overlooked by inexperienced GP trainers:

I’ve had one trainee with difficulty, and I found them (WPBAs) very good at actually pinning down what the problems were. And he was one actually, who finally had to leave the scheme, because he was never going to get through. It wasn’t picked up in his final ST3 year, because unfortunately this chap went to a brand new trainer, who was a bit soft on him, and it was only flagged up right at the end that there was a problem. And so then he came into a couple of us, and at that point it was very useful, so it depends on who is doing the assessments, doesn’t it? (GP Trainer, Deanery 4).
This year, trainers and programme directors spoke very positively in the focus groups about the summative exams, the CSA and AKT that were introduced with the new training curriculum. Most felt the AKT was a necessary exam that tested knowledge effectively. However, in one focus group, trainers raised concerns that it does not test basic knowledge and this allowed some weak candidates to pass:

Yes, my faith was slightly shaken with the AKT, which I thought in fact, would be an easier thing to get a good discrimination on the knowledge, in that my trainee with difficulty passed the AKT the second time. And yet he came to me with some real worrying gaps in his knowledge. He was someone who was very patchy (GP Trainer, Deanery 4).

I just wonder whether the AKT question base, which of course I have not seen, has a flaw that it is too high. It sort of assumes that there are some things that you must have learned in order to get on to the medical register, that aren’t tested, and therefore, you can really learn the NICE guidelines and all of the up-to-date stuff that you should be learning at this level, and get through the AKT, and if your foundations are weak, it doesn’t look at the foundations. I don’t know, I just wonder whether that is an issue (GP Trainer, Deanery 4).

Most of the focus group discussion on the exams was based on the CSA. Whilst, as in previous years, concerns were raised about the cost of taking the exam, and the additional costs incurred by trainees training in deaneries that were distant from the capital, generally, it was viewed by trainers as a fair exam and a great improvement on the previous video assessment:

I think it intuitively seems the right kind of exam because it is consulting which is what we do day in and day out and I think although it would be a very difficult surgery, probably each of those scenarios are reasonably real or realistic, so it seems a fair exam for testing the registrars (GP Trainer, Deanery 5).

I think the balance of the ST3 year is much better, I’m sure the trainers will agree. When we were doing the old summative process the video took over their lives. To an extent they now get the CSA taking over their lives, but I don’t anymore spend week after week watching videos, I watch some obviously but not in a coaching, perfecting way that I used to. I think that's been very positive because it gives me chance to actually teach what the registrar wants to learn about rather than the passing the hoops, so I thought it was a very good change (GP Trainer, Deanery 5).

During the second year, some trainers considered that trainees who had completed their earlier medical training outside of the UK could be at a disadvantage in the CSA because of their different
cultural approach to consultation. This year, this issue was not raised by trainers, but several discussed their experience of being ‘surprised’ by their trainees either passing or failing the CSA:

I’ve just been surprised with one person that I mentioned before, that she had passed, and I had felt that she was deficient in a number of areas. But equally I had a very good candidate - but it wasn’t my registrar, but he was in our practice, he was a very good consulter, and he scraped through it, and I would have thought that he would have done really brilliantly, and he didn’t seem to. I’m not quite sure why that happened (GP Trainer, Deanery 4).

I met a couple of trainees in my patch, who have been able to fly through the CSA, with 10 or 11 passes, but we felt that they were not competent, they’re not ready to be independent practitioners. But they have been able to pass the AKT, because they have got a good knowledge, and they can do consultations on a good day (GP Trainer, Deanery 4).

However, some trainers pointed out that there seemed to be a good correlation between good CSA passes and making good reflective log entries throughout training:

But what does intrigue me about the whole process, and certainly looking at ePortfolios from the panels, is that the people who fail just don’t do very well on all areas of it. And you can tell, it’s not just the final results, it’s what goes on there reflects the people putting it on. Overall, at the end of the year, the amount of work and the reflection of the whole quality tends to reflect them. And it’s a real good test you know, the people who fail panels and have bad comments, have failed the CSA. It’s a theme throughout the whole time. It is, from what I’ve seen, and I say to them at the start now, I say, “If you’re not putting it on, it’s telling me that you’re the kind of person who is not going to pass the CSA and get through, for whatever reason”. And those who do put it on, seem to do well. There is a very good link somewhere (GP Trainer, Deanery 5).

In one of the focus groups some of the participants who had observed the CSA earlier in the year discussed their surprise at its lack of focus on clinical skills:

A few of us have been down to London to look at the exam for the CSA and there’s three examples we saw and none of them were clinical. There were no clinical issues whatsoever. That would be one quarter of the examination which is not in the curriculum. They’re skills of diffusing situations; skills of dealing with anger. What boxes do you tick for them? You’ve got to sit down with them and talk them through and it’s mostly through consulting with them. There’s no amount of reading that can prepare you for diffusing an angry patient (GP Trainer, Deanery 3).
Trainees’ Views

Most of the ST3s in this year’s focus groups had experienced the new assessments throughout their three years of general practice specialty training, so had no experience of the previous summative assessment. As most of this cohort of trainees had completed the Foundation Programme prior to their general practice training, they had considerable experience of competence based WPBAs, and for them this was the norm. Some of the trainees described how this experience had often meant they were able to provide instruction to their supervisors in how these should be carried out:

We were the first lot to do foundation, the first lot to do the speciality training programme, and therefore we have always had to educate the consultants who we are giving out the forms to about the new process. So I wonder if the years coming behind us might find that a bit easier because we have laid the path for them you know (ST3, Deanery 2).

One aspect of trainees experiencing similar assessments in Foundation and GP specialty training was many considered that the DOPS in both programmes focused on the same skills. Trainees felt frustrated by having to repeat this exercise and have all the basic competences rechecked:

The DOPS in particular, for those of us who went through Foundation, and had to go through the whole process of getting somebody to watch you do all the basic skills; put in the venae flow, take blood, do an ECG, put in a catheter; all of those skills. And they’re actually called ‘foundation skills’ in the GP training programme. We had to do all those for our Foundation, and it was such a joke to make us do it again. It totally defeated the purpose of my Foundation Portfolio, and all of the assessments that we did (ST3, Deanery1).

Generally, the summative approach to assessment that trainers assigned to most trainees was evident in the trainees who participated in the focus groups. Assessments were viewed as hurdles that needed to be completed and many considered they were required to complete too many assessments. However, most of the trainees considered that, if done properly, assessments could provide good learning opportunities and some could see how an overview of a range of assessments could provide valuable insight into areas where development was needed:

I think the useful thing is with the ones we do in the registrar year is that we do quite a few of them, and by doing that, you start to see common threads in the way you practice which, ... can then change what you are doing, because you can see it. Sometimes it’s just an isolated thing and you might think, “OK, that’s something I won’t do again”, or, “That’s something that might have been a one off”; but also you will see characteristics of yourself cropping up again and again. So yes it can guide that and make you address those (ST3, Deanery 4).
The trainees considered that the WPBAs were “assessor dependent”. As in earlier years of the study, the trainees this year again highlighted the difference between assessments carried out in hospitals and in general practice. Difficulties in getting the assessments completed in hospital posts were highlighted and many felt this contributed to the process being reduced to a tick box exercise:

In hospital, they were really an inconvenience, and it was very difficult as a junior doctor to get them done. It depended on what rota you were on, and what your supervising doctors were like, but in a lot of cases, it was very difficult to get somebody to sit down with you and for example, discuss a case, or to go through a Mini-CEX and to come and watch you examine a patient, or watch you take an x-ray, and they really just weren’t interested. It was “How quickly can we get this done?” And “When is this due?”; “Tomorrow”; “OK”. So then we were doing it right before the deadline, because they don’t have the interest (ST3 Deanery 1).

This year, some trainees confirmed views expressed by trainers of “cherry picking” assessors for their hospital assessments to ensure they got good grades:

And also you want to do well of course, so people will very easily cherry pick someone that they are matey with, often a registrar that you’ve had a good week of nights with, you might ask to do a CBD or something, and of course they’re going to say nice things which doesn’t really help you (ST3, Deanery 4).

When ST3s reflected on their experience of assessment during their hospital posts, some felt they were perhaps not best prepared to get the most from these. Increased understanding and experience of the assessments had enabled them to benefit more from the process in their final year of training:

In the hospital it’s a bit like the blind leading the blind because you don’t really understand; maybe we should, but perhaps we didn’t really get enough teaching on it originally; what it is that we need to be getting out of them, what the domains are that we need to be focusing on; and you get so much more of that when you’re in practice (ST3, Deanery 4).

Very few of the trainees could identify any assessment that had been carried out well during their hospital training and for the few that had had this experience the key points were having time set aside for the assessment; having a supervisor that had an interest in general practice training and who attempted to ensure that areas of their specialty relevant to general practice were learnt; and being provided with constructive feedback. One trainee described positive WPBAs whilst in a psychiatry post:

I was with a psychiatrist, who was a locum psychiatrist, and he’d been in psychiatry for a bit more time. And I found it useful to do some of these assessments because it identified learning needs,
and then we went back. But it was to do totally with the motivation of the person filling it in. And
this doctor, he was a foreign trained doctor, and he thought this was very interesting, and was
able to say, “Right, this is what you’re going to do, and this is what you can learn”. And I found it
useful in that case, but it took a bit of extra time, and in other departments, other specialties
where the doctors, have got no motivation and are too busy, then they’ll just quickly tick - tick -
tick boxes, and you’ll not get anything out of it (ST3, Deanery 1).

Whilst most trainees considered there was little value to the hospital based WPBAs, several felt there
was generally more awareness among hospital based clinical supervisors of the learning needs of GP
trainees. One trainee highlighted that at least the requirement to get the assessments completed
ensured that some attention was given to addressing training needs in this context where service
provision tended to dominate the experience:

   I think in hospital jobs they are actually quite useful, because it means that the hospital
consultants will have to get them done and I need to get these done to be signed off. And in those
jobs it’s sometimes the only way you would get any contact one on one (ST3, Deanery 2).

Although many of the trainees had a summative approach to ensuring they completed the required
minimum number of assessments, this year, in particular, it was clear that many of the ST3s valued
the formative approach to assessment adopted in general practice. There was no evidence that
trainees were selecting low challenge cases for CBD and COTs, and often their views on attempting to
maximise the learning value of these assessments coincided with those expressed by trainers:

   The CBDs, I think, are only useful if you prepare them before you go and discuss them and if you
prepare them with a view of the domains in front of you. So you’re spending the week thinking,
“Who have I seen that makes sense of community orientation, or probity, or the wider issues that
are more important?” If you do that then they actually become really relevant and then I think
they change your practice, because if you look at one I did and a girl came in asking for the pill
and I think she was fifteen and she had had it before. But her dad was with her the second time
she came in, about something irrelevant, and I’d brought the screen up, as I would normally do
with people coming up and it had on the front this information. And if you look back on that it
changes what you would do. So now I don’t have the screen up for when people come in. I didn’t
know her dad was coming in with her. So if you discuss it like that, it changes things. But I don’t
see the value in hospital because no one has time, no one can be bothered and it’s such a clinical
discussion, you know: “I saw someone with pneumonia, I gave them amoxycillin”. It’s that sort of
discussion which doesn’t change anything (ST3, Deanery 4).

I always chose challenging ones because I thought that would help me move forward, but also you
can’t tick off all the boxes if you keep choosing the easy cases. I mean some people say you don’t
have to try and be marked in all of the different areas in one CBD. In which case it’s actually easy to choose simple ones, but the more complex ones actually lend themselves to be marked in all the areas (ST2, Deanery 5).

This year, the trainees seemed to understand better the approach towards assessment used in general practice and the different marking criteria adopted. Most accepted there was more value in being assessed as ‘needing further development’ than in gaining a tick box ‘excellent’, when gaps in learning and steps required to address these were identified during the assessment process. The cultural difference between hospital and general practice assessments was discussed by trainees, and most felt their training had prepared them for this transition:

I suppose from when we used to do it in the hospital I don’t think anybody would have put a ‘fail’ or ‘needs improvement’, because in this culture actually if they like you and they think you’re good enough, of course they’re going to pass you. And actually when you’re going into your registrar year, actually when you start off and you get a ‘needs further development’, and actually that’s the norm, and that was quite well explained. As you’re starting off, you’re being measured as an exit standard and actually the areas will all need development, and it wasn’t as threatening and that was useful then. So you can look back and see progress - that you’ve attained the standard of a GP without it being overly threatening (ST3, Deanery 4).

What I think is really important when we start GP VTS and the specialty training is that in our careers and professions we have been the top achievers at school and go to university and everyone’s trying to get top marks and everything’s ‘brilliant’, ‘brilliant’. And once you start work, obviously it’s not just about our academic achievements, but even through F1 and F2, basically nobody wants to put down anything bad because if there’s anything bad they will just think you’ll not get through, and you will need your registration. So there’s always this idea that if anything negative at all is put on your record it’s really bad, it’s complete failure and it’s something we’ve never come across before. I think a lot of medics are perfectionists and the problem is I think we need to be introduced to the concept that actually in a good assessment you’re being marked against what they expect from a fully qualified GP. So from year one, they’re saying, “No, you’re not good enough compared to what you should be like in three years time”. And it should be like this, you shouldn’t be upset if it’s like this and the aim is to improve and if there’s something which needs to improve that doesn’t mean you’re going to fail the whole lot, but it’s actually ok, and I think that’s quite important really. I was talking to other people that feel the same, not necessarily in medicine actually, but in other professions where they think everything is brilliant and perfect and good and then all of a sudden they just get slated, or have all these things where they need to improve and they don’t know how to handle it because they’ve not been taught. We’re getting it through the ePortfolio; we’re getting to understand the system. But I just think that it’s important that that is understood from the beginning (ST2, Deanery 5).
Generally, the trainees were positive about the AKT. Most trainees had taken the AKT in their ST2 year and although some thought there was too much focus on statistics, most passed on their first attempt and felt it was a fair exam. The arrangements for the exam were described as “well organised” in relation to the booking, travel and information. However, some trainees considered the fee for this computerised exam was too high at £400 and also felt they had to wait too long for the results.

Trainees’ views on the CSA were in general more mixed. Whilst many accepted that the OSCE approach with clinical scenarios was preferable to a written exam, several trainees considered these focused too much on social rather than clinical skills and this made it an unfair examination for a doctor, with some trainees suggesting that CSA should refer to “communication skills assessment” rather than clinical skills; or that, “an actor would do better than a doctor in that exam”:

And yet in my experience, certainly, it didn’t matter what the condition was, and I know there is a big social aspect in GP, but it seemed like the whole point of the entire thing was the social problem, in every single case, not in some of the cases, but in every single case. And I mean, it wasn’t medical things at all, and yet you can’t not study the medical things, so it just seems, it didn’t feel like a fair exam to judge you as a doctor. I personally felt that you could have gone in as somebody with very little medical experience at all, but it’s entirely about how you interact with the patient, and entirely about, you know, how good you cope if people start crying, and whether you ask the right question about their family background, and even if you, you know, you had got around to some of that, but not exactly the question that they wanted, you didn’t necessarily get the marks, and I think that was where I felt certainly, that it wasn’t necessarily a fair exam (ST3, Deanery 1).

In one of the focus groups the trainees pointed out there was a high proportion of overseas doctors training in their deanery. They considered that because of the social focus in the exam on communication skills these doctors were at a disadvantage in the CSA and were more likely to fail:

In any area ... a lot of GPs are overseas trained, and they are trained in a different way. They were brought up in a different cultural environment, so you shouldn’t deny the fact that there will be a little bit of difference in the culture, and their communication skills. When it is home grown graduates they are much more orientated to the new communication skills and cultural things and they are actually taught in the medical school about communication skills. In that case, I think overseas doctors find it difficult in the CSA exam to pass and that is why I think in general, the pass rate for the overseas doctors are less (ST3, Deanery 3).
The high cost of the CSA was a concern for most of the trainees, especially those training in the deaneries that were distant from the test centre in London, so had additional expenses for travel and an overnight stay. Several of the ST3s who participated in this year's focus groups had failed the CSA on their first attempt and this had added to the expense. One of the ST3s who failed the CSA described his shock when he received the results, and this would seem to reflect the trainers' experience, previously discussed of surprise failures in some of their trainees. However, this trainee considered his second experience of taking the exam was totally different:

The thing was I didn’t think it was a negative experience. We went in for the exam very positively, came out of the exam very positively. The only negative about it was when I got the results. ... But interestingly, comparing the two exams, if I didn’t know that I was sitting the same exam and if I wasn’t sitting it in the same building I wouldn’t have known they were the same exam. The correlation between cases, there was no correlation. The first exam was very medical, very surgical, very speciality specific; which I think isn’t something I had mentally prepared myself for, because I was going into a GP exam where I had been practicing how to be holistic and all sorts: whereas the second exam was very GP focused, lots of depression, lots of people crying, lots of anxiety so I think the focus of the cases was very different (ST3, Deanery 4).

This experience of a ‘surprise’ failure at the CSA was very similar to an ST3 from a different deanery, who had previously “never in my life failed an undergraduate or postgraduate exam”, however, in this example the exam that the trainee failed was described as focusing predominantly on social skills:

I had the opportunity to do the exam twice, and the first time, I was very disappointed, because I failed by a very short margin - by one mark, and I just felt really hard done by initially, you know. Having said that, when I got my feedback and reflected, I definitely knew that I could have done better, and to be fair, I worked hard, and did extra revision, and totally clearly passed the exam the next time. And I did notice a huge difference in the difficulty of the cases between the two episodes. And people said “Oh, that’s because you just worked harder, and you put in better, and you were more prepared”. ... Some of the cases in the first exam, there wasn’t a lot of hard clinical medicine, whereas in the second time that I did the exam, most of the cases were clinical, and I was actually worried that I had missed things, because the first exam was so social (ST3, Deanery 1).

The trainees held different views on how representative the scenarios in the CSA were of general practice experience and also on whether the consultation style and skills necessary to pass the CSA were the same that were required in their day-to-day surgeries. However, the majority of the trainees who passed the CSA described finding it challenging:
I think it’s a good exam. I think they throw the kitchen sink at you. To be fair I came out of there thinking ‘God!’ I was actually fairly tired, I’d used lots of different parts of my brain but I think, overall, it was a good reflection on balance. I think that was slightly unlucky really wasn’t it to have all clinical and that’s unusual. I think that it should have been less like that. But I thought it was good; it was quite reflective of what you’re trying to be as a GP because there wasn’t a right or wrong. It was more how you dealt with things as they threw them at you (ST3, Deanery 4).

5.3.2 Conclusion

This year’s findings document a continuing assimilation of the new assessment system into the every day practice of general practice training. Trainers’ increased confidence in using CBDs and COTs effectively to provide teaching and learning opportunities and to ensure these are based on cases that provide challenge and opportunity for development across the areas of competence demonstrates the value of the ongoing provision of training in assessment. Formal training for trainers provided by the deaneries, and the training and support provided in local trainers’ groups is shown to have contributed to this improved understanding. The development of this understanding is seen to be a gradual process where training courses and learning sets build on the knowledge and experience already achieved and enable the sharing and development of best practice. This approach to training is in line with recommendations made by Fullan (2007) to ensure the intended meaning of educational change is understood on a personal level by practitioners and translated into practice.

Assessments carried out in the hospitals are generally viewed as less valuable and effective than those conducted in general practice. Training is available for hospital based clinical supervisors, and deaneries are expanding this work to ensure that both experience and assessment in hospital posts are relevant to general practice training. However, the findings show that high staff turn over and a high rate of assessments carried out by junior medical grades mean many assessors are untrained. Additionally, the pressure of the working environment means often there is insufficient time to conduct assessments properly and these are often reduced to a tick box exercise. Perhaps of greater concern are the findings from this year’s focus groups which show that despite a more standardised competence based approach to assessment across postgraduate medical training, some consultants are still reluctant to formally record concerns. The focus groups demonstrate this can be attributed to a culture that is still prevalent in medical training in hospitals where negative comments are avoided to prevent blighting a trainee’s records and subsequent career. This potentially undermines the value and reliability of the clinical supervisor’s report.

For many trainees, competence based assessments have become the norm with many who completed their general practice training in 2010 also having completed the Foundation Programme. This year, the findings suggest that a cultural shift is occurring among trainees so there is increasing recognition of the value of the formative approach to assessment used in general practice. Most
trainees now seemed to accept that being graded as ‘needs further development’ was not an indication of failure, but rather an essential part of the process of their development as general practitioners.

The AKT and CSA are generally viewed positively, although concerns remain about the expense of the CSA. Some trainees also raised concerns about a cultural bias in this exam which they considered resulted in lower pass rates among overseas trained doctors. This year’s findings also documented some concerns over ‘shock’ results of the CSA with some trainees either unexpectedly passing or failing the exam. Trainees’ descriptions of their experience of failure provide some insight into their expectations of the exam being focused on either clinical or social cases which suggests their preparation had been based on these assumptions and that a more holistic approach is required.

This year’s findings also demonstrate that the increased experience of using the assessments, in particular the experience gained from conducting panel reviews, has highlighted some difficulties in gauging trainees’ progress from individual sets of particular types of assessments. Whilst some concerns are raised about the validity and reliability of WPBAs, the value of the learning log and its link to the PDP was highlighted as a valuable assessment that could provide documentary evidence of progress and development. The learning log, MSF and PSQ were highlighted as valuable in identifying weaknesses in trainees. ‘Community orientation’ and ‘fitness to practice’ were identified as the most difficult areas of competence to achieve. Deanery leads considered that clinical leadership, clinical governance, critical reading and the business skills in running a practice were essential skills not sufficiently addressed by the assessments.

5.3.3 Recommendations

- Training in assessment for trainers needs to continue along with workshops where best practice can be shared.

- Training in assessment for hospital based assessors needs to continue and should be extended beyond clinical supervisors to include all those who conduct assessments.
6. FINDINGS: FITNESS FOR PURPOSE

This chapter presents findings on the fitness for purpose of the new training curriculum for general practice. Based on data from focus groups and interviews with deanery leads, trainers and trainees, their views and experience of the length and structure of general practice training and the preparedness for practice of newly qualified GPs are considered.

6.1 Length and Structure of Training

When the new training curriculum was first introduced, variation existed between and within deaneries in the balance of time spent in hospital and general practice training. However, for the majority of trainees their three-year specialty training programme provided two years experience in hospital posts and twelve months experience in general practice. Over the period of the study, most deaneries introduced an additional six-month post in general practice to replace one of the hospital placements during ST1 or ST2. This meant by 2010 most GP STs during their three-year training experienced 18 months in general practice and 18 months in hospital based training. This additional time spent in general practice was welcomed by trainers and trainees, although there were some concerns about the loss of experience in a hospital specialty. Hospital posts were viewed as variable in quality with many primarily oriented towards service provision rather than education and training, and often having poor relevance to general practice.

During the second year of the study, findings from the national survey and the focus groups showed a majority of trainers and trainees considered the overall period of general practice training should be extended to five years. Most felt this additional time should be spent in general practice to allow increased focus on practice management and administration and more time to consolidate learning to enable trainees to develop as independent practitioners.

6.1.1 Findings from Year 3

Deanery Leads’ Views

Deanery leads welcomed the changed structure of training which now allowed most trainees to gain 18 months experience in general practice. They also considered an extended period of training would be beneficial and ideally should allow a minimum of two years’ experience to be gained in general practice. In most of the deaneries, four and five year training schemes had been piloted and these had proved very successful in addressing areas that the leads currently felt were not allowed sufficient coverage; including skills in research and critical appraisal; and the business side of running a practice. Some considered a four year training programme would be adequate whilst others felt that five years’ training was required to consolidate learning and allow time for trainees to develop expertise:
Trainees don’t have the time to consolidate the clinical skills. There’s considerable evidence that it takes roughly 10,000 hours of practice to achieve excellence to expert. 10,000 hours is roughly five years; half of the three years is in hospital, not seeing GP specific patients, and certainly not working in a GP specific way. We’ve only got a year and a half of GP training (Deanery Lead).

Most of the leads considered that the additional training time should be spent in general practice. One lead considered it would also be useful during this additional period of training for trainees with gaps in their experience of key specialty areas to have “supernumerary clinical attachments” within hospitals to gain the “appropriate experience”.

Trainers’ Views

Most trainers considered the 18-month split between hospital based and general practice experience was “a reasonable balance”. They welcomed the additional post in general practice which they considered prepared the trainees well for their ST3 year and also for the AKT and CSA. Trainers identified how this additional six months enabled more rapid development of consultation skills and led to ST3s achieving ten-minute consultations much earlier than before. In some deaneries, this additional post in general practice allowed trainees to gain experience in a different practice to the one where they were based for their ST3 year. Trainers were generally positive about this wider exposure:

Certainly, I very much welcome the introduction of the 18 months in general practice. I think that’s very good for training. It allows, I think, from the trainees’ point of view, they get an exposure to two different practices and maybe three, four or five different trainers, and I think that’s very positive for their development (GP Trainer, Deanery 1).

As in previous years of the study, hospital posts were viewed as widely variable in quality with many being purely service oriented and where trainees often functioned mainly as “glorified clerking machines”. Whilst experience gained in acute care was seen as valuable for the trainee, many trainers considered that there was limited opportunities for GP STs to gain outpatient or community experience during their hospital posts, which they felt would be more relevant to general practice than ward based experience:

The ST1s and 2s enjoy being in hospitals, they do enjoy being on the wards, they do enjoy seeing acutely ill people. What they don’t enjoy is having a service commitment clerking on the wards. What they need, and what is closest to general practice, outside of A&E, is actually Out Patients experience, and what often happens is that the speciality trainees go to the Out Patients and they’re left on the wards. Nor are they going to community psychology, community paediatrics, community audiology and podiatrists (GP Trainer, Deanery 5).
Trainers expressed concern that often there was “little or no clinical supervisor input” in hospital posts. There was some evidence that educational supervisors could provide support and guidance to their trainees during their hospital placements. Communication via the ePortfolio could enable better coordination with general practice and ensure that relevant specialty experience was gained:

The difficulty before was trying to get over what they actually need to get out of the hospital posts that was relevant to general practice, I think now there’s a continuous process of education throughout the whole three years, I think they’re able to focus much more on that. And having an educational supervisor who is a GP trainer right from the start, you can help them to focus on what they need to get out of a hospital job that is relevant to general practice, which probably wasn’t happening before (GP Trainer, Deanery 5).

However, this was by no means widespread practice and there is certainly scope for further training to ensure this potential is realised in the educational supervisor’s role.

Most trainers were in favour of extending the overall period of general practice to either four or five years. Some felt that five years was required because of the complexity of general practice and also to bring GP training in line with other specialty training:

Well, they should all have ST4 and ST5 as well. I think that there is absolutely no justification whatsoever for 50% of doctors, who are doing the most difficult job in the health service, to have less training than other people. And I think that 18 months of hospital is fine, but they should have ST4 and ST5, all of them, because it is the most intellectually demanding branch of medicine, and they need more training. (GP Trainer, Deanery 5).

During the period of the evaluation, some deaneries had the opportunity to pilot four-year training. This experience was usually reported positively. However, in one of the regions where focus groups were held, trainers and programme directors reported none of the trainees had been interested in pursuing this. Most of these trainees were overseas trained and most wanted to qualify as GPs at the first possible opportunity:

They were offered an extra year for their training and none of ours took it up. We had one who was half interested who in the end didn’t, although it has been taken up elsewhere in ... (the deanery). So I think our view of what they need is different perhaps from the registrars. I think they feel they just want to get through the scheme and out the other side and earning as GPs really (GP Trainer and Programme Director, Deanery 3).

Most trainers considered that the extended period of training should be spent in general practice and this additional time would enable trainees to “get an overview of what general practice is all about”. 66
Trainers generally felt that during the ST3 year there was too much requirement to complete assessments and too little opportunity for ST3s to focus on their development as GPs:

I think the second six months in ST3 needs to be urgently looked at, because out of 16 weeks when they start that slot until they have their submission to the ARCP panel, they have to complete 12 assessments and we have to complete our educational supervisor’s report, and there is time for virtually no teaching at all in the second six months; and they are also trying to do their CSA at the same time, and I think it’s just got to change (GP Trainer and Programme Director, Deanery 3).

Some trainers discussed how this increased focus on assessments meant newly qualified GPs were often not sufficiently prepared for independent practice and therefore advocated an extended period of training was required:

I think that definitely, five years would be beneficial. … more and more now, it’s the newly qualified doctors that are having problems and complaints against them when they go out into general practice. And I think that some of it is that although they are completely focused on passing the exam, the CSA and the AKT, and doing their log, they are so sort of focused on what they’ve got to do, that they don’t see the bigger picture: the actual nuts and bolts of general practice; the role of the partnership; the role and the interaction of the staff; … actually, the practical things; and how you actually manage a practice; and how you have to learn to work with your staff in a certain way; and doing appraisals; learning about protocols; and knowing about running a practice: as well as all of the extra things; like being on call and being duty doctor; and doing visits. These require a bit more exposure. And so it’s like we’re teaching them the nuts and bolts of it, but we also need to immerse them in it a bit longer, so that they can experience it a bit more, before they go out into the world (GP Trainer and Programme Director, Deanery 4).

Longer training would also allow time for consolidation of learning. Many trainers felt that ideally, this extended period of training should be assessment free:

I think the most productive period I saw for my last registrar was that period after everything had been submitted. And he wasn’t under pressure and I wasn’t under pressure, and we were able to cover a lot of the things that I felt were important for us to cover that he couldn’t see the need for before, because the priority for him was getting all the ticks in the right places. And we actually spent a lot more time doing the free ranging experiential sharing of knowledge and broader overview of things, which don’t fit into the tick boxes (GP Trainer, Deanery 3).
Trainers and programme directors in the English deaneries agreed with deanery leads that an extended period of training would be necessary to address the planned major change for the role of general practitioners in commissioning services.

**Trainees’ Views**

Most of the trainees who participated in the study during the final year of the evaluation were completing their general practice training over a three year period and had spent 18 months in hospital based posts and 18 months in general practice. Some trainees had different patterns of training; these included those in academic posts who trained over a four-year period and trainees with Innovative Training Posts (ITPs). Whilst there were differences between the structures of ITP posts both within and between the deaneries, in general, these posts allowed trainees to gain a range of specialty experience that was tailored to meeting the needs of GP STs. Often these posts allowed trainees access to outpatient and community based clinics. Although four-year training was piloted in some of the participating deaneries, none of the trainees involved in these pilots participated in the study.

For the majority of trainees who had experienced an additional six months in general practice during ST1 or ST2, this opportunity to gain insight into general practice was welcomed and viewed as beneficial in enabling them to develop important skills. Many felt this improved understanding enabled them to ensure more general practice focus was achieved in any subsequent hospital posts:

> I thought the thing that was most valuable was obviously training in general practice. I was glad to get 18 months instead of a year (ST3, Deanery 1).

> I think the 18 months in general practice has definitely been beneficial ... you learn a lot more in general practice where you’re supernumerary; you’ve got a trainer who is usually well motivated too. My best training was in general practice with the protected learning time for actually learning about things. ... The thing that will help you become a good GP is spending a good bit of time in general practice (ST3, Deanery 1).

In one of the deaneries where most ST3s had not had this additional experience of general practice, they described experiencing a cultural shock at the start of their ST3 year because it was so different from their experience of working in secondary care:

> Starting a registrar year was quite tough when I hadn’t had any time in GP. ...If I had done six months of GP training initially in ST1 that would have given me an idea of what I was going to expect from the job. ... After doing two years of hospital, when I started my GP registrar year, I was in a totally different world and I wasn’t prepared for that (ST3, Deanery 2).
Whilst trainees found the additional post in general practice beneficial to their learning and development, many considered that there were gaps in the specialty experience they had gained in their hospital posts. As in previous years of the study, many trainees were generally dissatisfied with the balance of specialties within the hospital component of training, especially when placements in their general practice training were based in specialties they had worked in during their Foundation training:

I had already done psychiatry in the Foundation and again (in GP training) I was rotated through psychiatry. If you had a degree of choice you could identify other areas that you felt you needed so that you got a broader experience (ST3, Deanery 1).

I found that a lot of my jobs in the first two years I had done the same thing in Foundation, so they weren’t useful. I think there is no personalisation in terms of, ‘What are your weaknesses?’; ‘What do you need to learn personally to become a good GP?’ It’s a blanket system that doesn’t pick out any of your learning or any of the jobs you have done beforehand. You are straight into that system and come out at the end of it, and you haven’t done certain jobs which leaves deficiencies in some areas (ST3, Deanery 2).

The length of hospital rotations varied between the deaneries with some offering trainees three four-month posts in each of the first two training years while others offered just two six-month posts. Trainees generally considered four-month rotations were preferable as these allowed a wider range of specialty experience. There was also a general view that there should be exposure to certain core specialties, whether this was achieved through shorter rotations or through targeted placements in outpatient clinics.

As in previous years of the study, many of the trainees considered the quality of training in their hospital posts had been poor and that the focus had primarily been on service provision. In some hospitals this experience had been exacerbated by staff shortages and several trainees felt that this meant that GP trainees were “carrying a lot of the rotas”:

There are supposed to be six people on a rota and there is actually three and two of them are GP trainees, and they are left to do more than their fair share of work. Your educational needs are ignored in favour of, “Well actually, we need somebody to work overnight”. …that is one of the big problems there is not enough bodies to fill the jobs and your training is secondary to that (ST3, Deanery 1).

Many of the trainees valued the study days provided by their course organisers. ST3s considered these had been useful in addressing gaps in learning, particularly when they had not had the opportunity to gain experience in certain hospital specialties. Prior to one of the focus groups that
was held at an ST3 study day, the trainees had been given a talk by a physiotherapist on joint examinations, which they had found very useful:

In the session we had this morning we had a physiotherapist talking to us for two hours about shoulder joint examinations and that was at our request, because I don’t think that many of us had any orthopaedic experience (ST3, Deanery 1).

The study days and the responsiveness of course organisers to arrange sessions on topics requested by the trainees was described as “invaluable”:

One of the best educational things of the GP training is the Thursday days, because it can be so valuable. We’re a relatively small group and we can say to our programme director, ‘These are gaps in our knowledge, let’s address these’. Whereas in hospital that is never going to happen. You’re not going to get a half day a week, never mind a full day (ST3, Deanery 1).

The trainees held mixed views on whether training should be extended. Most felt additional time would be useful, but generally considered this should be spent in general practice. Similarly to last year, trainees felt strongly that any extension of training should be focused on their development as GPs and not on providing service during an extended period in hospital posts:

I would be concerned that actually you would just end up being used as cheap labour for the NHS when actually your learning needs wouldn’t be met anyway (ST2, Deanery 5).

Whilst there were different views on how much longer the training should be, most considered six to twelve months would be sufficient and agreed with trainers that this would be most beneficial after all the assessments had been completed:

If you did have six months time to consolidate things; maybe have an educational day a week in a certain type of clinic but still spending the rest of the time in general practice. ... I think so much of your year in GP is spent building up to your AKT and CSA that you have got very little time after that to do the other things to make sure you are where you need to be (ST3, Deanery 2).

You really develop your knowledge base of general practice in general practice. If the general practice part of things was extended for six months or a year you’re able to learn more. And if you were a senior registrar and were given more responsibility then ... I’d find that particularly helpful. I think you’ll then feel after four years you’d really want to be moving on (ST3, Deanery 4).

Several ST3s pointed out that they would have been unhappy if an extended period of training had been introduced part way through their training.
6.1.2 Conclusion

The additional post in general practice was valued by leads, trainers and trainees and examples were given of how this early experience of general practice meant ST3s were better prepared for their final year. Trainers and ST3s also agreed that most learning occurred during the time spent in general practice. Most trainees spoke positively about their ST3 year in general practice.

Three years into implementation of the new training curriculum the quality of some of the posts in the hospital-based component of training was still causing concern. Trainees and trainers considered many of these posts were focused on service provision and there were few opportunities to experience outpatient clinics where there would be greater relevance to general practice. There was also concern that not all trainees had the opportunity to gain experience in key specialties during their hospital based training. However, the value of study days in providing focused training to address these gaps was highlighted. The value of the role of the educational supervisor in ensuring that trainees were able to achieve a general practice focus during their hospital based training was also recognised.

Whilst there was a general consensus that the overall period of general practice training should be increased with a preference for an extended period in general practice after assessments have been completed, views differed on the length of this additional period.

6.1.3 Recommendations

- Work needs to continue to ensure greater consistency in the quality of posts in the hospital component of training.

- Training for educational supervisors should continue to ensure their effectiveness in providing guidance to their ST1s and ST2s on how to achieve a GP focus during their hospital based training.

6.2 Preparedness for Practice of Newly Qualified GPs

Last year, findings from the national survey, focus groups and interviews showed that most trainers did not feel that newly qualified GPs were fully prepared for independent practice. Trainers generally felt that there was insufficient time during the ST3 year to fulfil the assessment requirement and to consolidate learning. ST3s were generally more confident about their preparedness to practice independently, particularly in their clinical skills. However, many felt insufficiently prepared for the business side of general practice. Most newly qualified GPs commenced their careers either as salaried GPs or as locums. Whilst there were fewer opportunities than in the past to gain partnerships
immediately after qualifying, most of the ST3s considered salaried or locum posts were preferable first posts as they provided further opportunity to gain experience before taking on the business responsibilities associated with partnership.

6.2.1 Findings from Year 3

Deanery Leads’ Views

Findings from the interviews with deanery leads showed they considered that the new training curriculum was producing “clinically competent” doctors. However, most felt they were not sufficiently prepared to practice independently:

I’m not sure the curriculum is designed to produce independent practitioners. It’s certainly very good at looking at the clinical side of things and what is needed to be known, or how trainees can access information. However, working as a small business, which is what general practice is, is not really covered in the curriculum except in the management module, which is not really an element of the curriculum, unless I’ve completely missed it. So I think we’re training them to be good doctors with GP hats, but we’re not really preparing them for their real life as a small business person as well (Deanery Lead).

Several of the leads considered the new curriculum was producing better GPs than the previous system of training. The WPBAs and the CSA were generally viewed as better at testing necessary competences than either summative assessment or the old MRCGP:

I think they are overall better. I think we can be more assured that they are clinically competent. ... I would argue that the new system is better than the old in achieving going beyond competence (Deanery Lead).

I was an examiner before for the old membership exam and I don’t think that prepared them at all for being an independent GP. I think because all these assessments are workplace, they’re competency based, they’re looking at performance, and providing a trainer knows what they’re looking for – and they should do. ... You will have to see in a few years time how many complaints go to the GMC, but you have to think it’s going to be less than it was with the old system (Deanery Lead).

The leads generally thought that the new training programme had resulted in trainees having fewer gaps in knowledge as it ensured wider coverage of the curriculum. However, several gaps in learning were identified. These included insufficient exposure to relevant specialty experience; skills in management and clinical governance and delivering excellence beyond the end point of CCT.
The leads pointed out that it was no longer appropriate to view the career choices of newly qualified GPs as an indicator of their preparedness for independent practice. Changes to the structure of general practice have led to a considerable reduction of available partnerships. This has meant that for most newly qualified GPs this is no longer an option at the start of their careers:

Career choices are limited by the opportunities that are out there. At the moment partnerships aren’t freely available and a lot more trainees will go into out of hours or salaried positions, because that’s all they can get hold of. It’s hard to know if all these extra skills they are developing are going to have any benefit in the long run because at the moment there’s huge uncertainty in practice and people aren’t looking for extra doctors to help them. It’s likely that what we are producing now is much more able to cope with practice, but the jobs aren’t there at the moment (Deanery Lead).

Trainers’ Views

This year’s findings were similar to findings from Year 2 of the study. Most trainers felt that newly qualified GPs were not fully prepared to practice independently, but some pointed out that this was no different from in the past. Most trainers considered it was necessary for new GPs to gain further experience in their first posts whilst being supported by colleagues. Some held concerns about the increasing number of GPs working on a part time basis and felt this could lead to newly qualified GPs becoming isolated and less well supported by experienced colleagues. Similarly, concerns were raised about the increasing popularity of newly qualified GPs working in salaried posts in ‘Darzi’ style practices:

I worry about these salaried GPs that go into Darzi practices with people at the same level, you know, working side by side with someone with exactly the same experience without having someone senior to feed on (GP Trainer, Deanery 5).

Trainers held mixed views on the difference between the preparedness for independent practice of newly qualified GPs who had trained with the new curriculum and those trained under the previous system. Some highlighted that it was difficult to make comparisons between the two systems of training, because many factors, not just general practice training, needed to be taken into account. Some considered recent trainees were of a higher quality. Several trainers considered that the locums who were currently working in their practices were now of higher quality than those trained on the previous programme:

We had a locum in for six months, and that’s when you see where there are gaps. And I must say, compared to a locum who had come through the old system, there are less gaps with the new locum. … Locums coming out now are probably a lot more fit for purpose than they were, say, five years ago (GP Trainer, Deanery 1).
Some trainers pointed out that better-prepared trainees could, in part, be attributed to better selection procedures.

There is a compounding factor as well, which is the new selection process, which means that, you know, people seem to be coming in with a uniform standard of excellence, because they have been subjected to this incredibly rigorous selection process, which wasn’t in existence beforehand. So, my impression is that certainly, they’re getting better all of the time. And one actually wonders, you know, why are they getting better all of the time?: it might possibly be because of the selection method (GP Trainer, Deanery 4).

Changes within the structure of postgraduate medical training, particularly the introduction of the Foundation Programme, were generally viewed by trainers as having a positive impact on current trainees. Foundation training often meant that trainees had the opportunity to experience general practice at an early stage in their careers. They had a good understanding of a competence based approach to assessment and often had experience of recording their training on an electronic portfolio. This meant that GP STs who had trained on the Foundation Programme were better prepared for the transition to specialty training than career change doctors. Career change doctors often struggled to demonstrate their competence in a 12 month period with the new training programme:

I see it as more of a dichotomy between the ST3s that you’ve described that have come from Foundation matching those against the career changes, and that’s where I’m seeing the difference. We see more struggling doctors from the career change than before, and I think that’s brought it into a much more sharper focus the difficulties they have compared to the FY group. So I think I’m dealing, as a programme director, more this year with ST3s in difficulty, number 1; but also the background, as a general rule, is they are all career change. ... I think it is this competency based approach and they have difficulties with that if they’ve come from a previous surgical background. They’ve trained very much in the old way I was trained, through didactic humiliation, conquer and rule; and they find this very alien; the supportive, the mentorship, preceptorship, that type of thing, they struggle with (GP Trainer and Programme Director, Deanery 5).

However, some trainers considered changes in the nature of hospital training posts, particularly stemming from the European Working Time Directive (EWTD) meant trainees were “overprotected” and had less opportunity to take clinical responsibility or gain hands on experience:

And also the standard of training in the hospital now is a lot less than it used to be. I mean there’s some posts, you know, the SHOs don’t even have to work nights. And they’re so sheltered by
staff grades and consultants, and everything else, that the actual level of experience they’re getting in these posts now is a lot less. (GP Trainer, Deanery 1).

In some deaneries, hospital posts for GP STs were of four-month duration. Some trainers in these deaneries considered these shorter rotations reduced opportunities for the trainees to take on responsibility:

I think one of the problems is reducing posts down to four months. And the consultants really value that final two or three months where they have become more competent practitioners, and they never get there now. So therefore they feel they have to kind of nanny them more and give them less responsibility (GP Trainer, Deanery 4).

However, trainers identified several aspects of the new training programme which they considered prepared trainees better for their careers in general practice. Several trainers felt that GP STs did not get the opportunity to develop experience in decision making until they were working with a patient caseload in general practice. The additional time spent in general practice was viewed as providing greater breadth of experience and broader curriculum coverage:

In the old system in twelve months there were definitely things that people wouldn’t have come across, and I think in the eighteen months that’s less likely to be the case; and I think that the new curriculum does help identify areas that maybe haven’t been covered anywhere. So I think, you know, whilst you can think back to very good trainees who came in through the old system very well, I think if they’d have come through the new system, their training would have been more complete. So I think it is a definite benefit (GP Trainer, Deanery 1).

Trainers considered current trainees were better prepared as mature self directed learners by the new programme, in particular through their use of the ePortfolio. Preparation for lifelong learning also meant the trainees were equipped with the skills they would require for reflection, appraisal and revalidation:

I feel that one thing that is very different, is that they are so used to using the ePortfolio, and having to engage in that learning, they are much more mature learners, they’re more prepared to try and do the learning themselves, and make a record of it; and I would hope that with the revalidation and appraisal process coming, you know, it is very important, that they’ll be much better prepared for that than any registrars have, or we were in the past, and they will hopefully view learning as an ongoing thing, throughout the whole of their careers, rather than just for 3 years (GP Trainer, Deanery 4).
Another aspect of the new training curriculum recognised by trainers as contributing to newly qualified GPs being better prepared for independent practice was the improved structure of training that provided better coordination between hospital based and general practice training. The continuous assessment over the three years of training, contact with an educational supervisor in general practice and the ePortfolio were all seen to achieve an improved general practice focus:

The difficulty before was trying to get over what they actually need to get out of the hospital posts that was relevant to general practice, I think now there’s a continuous process of education throughout the whole three years, I think they’re able to focus much more on that. And having an educational supervisor who is a GP trainer right from the start, you can help them to focus on what they need to get out of a hospital job, that is relevant to general practice, which probably wasn’t happening before (GP Trainer, Deanery 5).

Although the new training curriculum was viewed as providing a more structured approach to training that ensured better curriculum coverage, trainers identified several areas where they considered there was insufficient focus. In both the second and final years of the study, practice management was one of the main areas where trainers identified a gap in learning among newly qualified GPs:

I find teaching general practice management quite difficult; it gets squeezed in at the end. And I think that if they are going to go into being a partner eventually, that it would be helpful to have more of an understanding of the business part of general practice management. And I think that for me, that’s an area that probably gets neglected a bit (GP Trainer, Deanery 4).

Another gap in learning identified by trainers in both years of the study was in academic skills and the ability to critically appraise evidence. Whilst this was addressed in the AKT, trainers generally felt that this was not in sufficient depth:

I think that the other thing is the sort of critical skills on research in general practice, because I tend to teach that to our registrars in our practice, and they all sort of just get a few bits and pieces, to get them ready for the AKT, because there is quite a section on that. And once the AKT has gone, then that’s it, they have moved on from that. ... Particularly, in a world where increasingly, there’s quite a lot of protocols and guidelines, they are not given the skills to actually critique those, and be as independent practitioners, make a decision on whether these things, which are often politically or financially driven, are actually sound (GP Trainer, Deanery 4).

Other gaps in learning identified by trainers in the final year of the study include experience of continuity of care and living with uncertainty. Generally, trainers considered an extended period of training, preferably with no further assessment, would be necessary to ensure this experience was gained because the ST3 year was currently too “bitty” because of its focus on assessment.
**Trainees’ Views**

Similar to last year, most ST3s who participated in the study this year planned to commence their careers in general practice in either salaried or locum posts. In part, this is because of the changing nature of careers in general practice, with partnership opportunities becoming rarer for newly qualified doctors. Just two of the ST3s who participated in the focus groups and interviews this year had gained partnerships. Some trainees felt that “partnerships have had their day”. Most ST3s considered that working as locums or salaried GPs would enable them to gain more experience before taking on the commitment of a partnership. Some ST3s also compared partnership to marriage, and considered that working as a locum allowed them to try out different practices and gain more experience before making a long term commitment:

I would like to see different practices and see a different angle of general practice and I don’t really want to be tied to the additional things around seeing the patient, and as a partner there are lots of additional administrative things and business things, that at the moment, I’m not really too interested in. So to me that would be more of a burden I think than a development for me (ST3, Deanery 4).

I think it is a commitment geographically and in terms of timeframes, it’s a bit like a marriage. There’s a lot to learn … I’m sure over the next couple of years our clinical confidence will continue to improve and that’s going to be my focus … I wouldn’t definitely want to be a partner until later on (ST3, Deanery 4).

However, some of the trainees considered the demands of locum work would be quite stressful as they would have to learn new systems of working:

The biggest apprehension is not the training or the amount of knowledge … it’s the working in new places with new people, new systems and when you actually start with the patients coming through before you know how to use the computer, or how you refer. People walking through the door doesn’t worry me; I can cope with that. … Yes, I will come up against things I’m not sure about or have problems with; and I may miss things or make mistakes, but I still feel most of my worries are knowing different working environments (ST3, Deanery 1).

Most of the ST3s felt confident and prepared to practice independently, although they acknowledged they still had a great deal to learn. Some highlighted how they had valued support from their trainers and experienced colleagues over the year and were uncertain of the support they could expect as newly qualified GPs:
After completing the rest of our training we will definitely be functioning very well as independent GPs … but it is only three years of training, so it gives you relatively less time to work independently on your own taking the responsibility and decisions. ... when you’re an ST3 working on your own you are always being supported and you can knock on a door and ask for some advice, but it’s not going to be that easy to ask for help (ST3, Deanery 3).

Most of the trainees recognised their careers as GPs would entail a commitment to lifelong learning and were confident their training had equipped them with the necessary skills for this:

I think medical knowledge wise, we’ve reached the stage where we can learn by ourselves by what we see and reading about it (ST2, Deanery 5).

The ST3s were generally confident in their clinical skills, and for most, the main gap in learning they identified was, as in the previous year, the business and administrative side of running a practice. Some acknowledged that it was too soon to know where any gaps existed and this would become more apparent once they embarked on their careers. One trainee pointed out that they had only just become aware that they had been carrying out shoulder examinations incorrectly, but despite this was confident he had learnt how to be a safe practitioner:

I think that we would all feel that we can manage an emergency, that you can recognise a significantly ill patient and deal with them appropriately. I do still feel that I have anxieties about if I’m doing things correctly and, for example, this morning when a physio came to teach a shoulder examination, several of us said ‘my goodness, I have just realised how deficient my shoulder examination is’. And I think there are lots of things that you maybe haven’t learnt as yet, but I think we would all probably feel that we’re safe to practice and if we don’t know things, we do know to ask (ST3, Deanery 1).

ST3s considered that the period in general practice after all assessments had been completed was valuable for addressing gaps in knowledge. For some this had been achieved through arranging visits to specialty outpatient clinics. Most trainees felt that an additional period at the end of their training would be valuable to consolidate learning and address and gaps.

6.2.2 Conclusion

Generally, deanery leads, trainers and trainees considered the new training curriculum prepared trainees well for clinical practice and most ST3s were fairly confident about embarking on their careers in general practice. The difficulties of comparing newly qualified GPs trained under the previous and current systems of training have to be acknowledged, and it is necessary to recognise the impact of other parallel changes that have occurred within the NHS. However, deanery leads and trainers were able to identify several aspects of the new training programme which they felt had
made a positive impact on better preparing trainees for independent practice. These include an additional period in general practice that prepares the trainee for their ST3 year and allows more focus on relevant learning. Trainees were considered to be better prepared for independent learning through their use of the ePortfolio. The written curriculum and assessment system ensured broader curriculum coverage than in the past. Whilst concerns remain over the quality of many of the hospital training posts, most trainers acknowledge that there is improved coordination between the two components of training through the educational supervisor’s role, the use of the ePortfolio and continuous assessment. The new system of assessment is also recognised as being more effective at identifying weaknesses in trainees.

Trainees considered their main gap in learning was a lack of knowledge of the business and administrative side of running a practice. This coincides with the concerns of GP registrars trained under the previous system that have been documented in other research. Whilst leads and trainers agreed that current training did not sufficiently prepare trainees for the business side of general practice, they also identified a more extensive list of areas where they felt there were gaps in learning. These include skills in clinical governance; academic and research skills; experience in continuity of care and living with uncertainty; and gaps in relevant experience of key specialties. Most considered there was a need to extend the overall period of training to ensure consolidation of learning and opportunity to develop these skills.

Most of the ST3s were confident about their clinical skills, but recognised the need for ongoing learning and development throughout their careers. Most planned to commence their careers as locums or as salaried GPs and considered this was preferable to taking on the commitment of a partnership at this early stage. However, changes to the structure of general practice and the limited availability of partnerships means that the opportunity to gain a partnership at the commencement of a career in general practice is now less likely. This means that the decision of ST3s to initially pursue other career options can no longer be viewed as a lack of their preparedness for practice.

6.2.3 Recommendations

- Ways need to be explored of ensuring that training includes coverage of skills in running a business, skills in clinical governance and academic and research skills; and also that experience is gained in continuity of care and living with uncertainty. However, it is possible that this might only be achievable if the overall period of training is extended.
7. ANALYSES OF NATIONAL ASSESSMENT DATA

7.1 Introduction

The ePortfolio is a key component of the new curriculum as it records workplace based assessments of GP trainees and logs curriculum coverage and other activities. The analyses described in this section use ePortfolio data provided by the RCGP; one of the main purposes is to address how well WPBAs function as assessments.

As this is a different emphasis from much of the report, we place the analysis within the context of van der Vleuten’s work on the assessment of professional competence. He described the utility of an assessment system as composed of reliability, validity, educational impact, stakeholder acceptability and cost (van der Vleuten, 1996) and the importance of these different components depends on the purpose of the assessment (van der Vleuten and Schuwirth, 2005). In that context, it is our understanding that reliability and validity are more important for AKT and CSA assessments as they primarily have a summative purpose, whereas educational impact and acceptability are more important for WPBAs as their purpose is primarily formative. It is also worth noting that in the second paper he sought to “achieve a conceptual shift so that instead of thinking about individual assessment methods, we think about assessment programmes” (p.310).

We take the view that the College designed WPBAs to provide formative learning experiences for trainees and perspectives on these are provided in other parts of this report. In this section, we analyse the extent to which WPBAs feed into ARCP judgements. To what extent do WPBAs indicate progression as represented in the ARCP outcomes and what is the association between WPBAs and summative assessment? In addressing these questions we apply a more ‘classical’ quantitative analysis. In doing so, we do not dismiss the formative aspects of WPBAs which, indeed, are paramount in terms of their routine use. It is, simply, not the focus here.

The methods used in these analyses are discussed in Methods, Chapter 3 in section 3.2.3. The data from the ePortfolio were ‘messy’ and required a good deal of ‘cleaning’ prior to analysis. These issues are described in the next section under ‘Data manipulation and Cleaning’ and lead to recommendations for improving the accuracy of the data if, in the future, the College wishes to use these data for further analysis.
7.2 Data manipulation and cleaning

The ePortfolio database is enormous, containing millions of records from over 11,000 trainees. As well as these trainees, data has been inputted by thousands of assessors, educational supervisors and other staff. The data is in a flexible structure: for example trainees can take different lengths of time to complete a training year and can undertake different numbers of WPBAs. In addition, the records are from the start of the ePortfolio. Taken together this means that the dataset is extremely large, complex and contains numerous inconsistencies.

Described below are the important steps undertaken to manipulate and clean the data before analysis could start. Many readers may wish to ignore this section! However, it contains information on the complexity of the data, indicates limitations in the analysis and may help the RCGP consider ways to improve the mechanics of the ePortfolio.

The age of trainees on 31st July 2010 was calculated. These ranged from 2 to 80! Ages younger than 20 (n=7) and over 65 (n=2) were treated as errors. On the database, only 17 (0.2%) trainees did not give their gender.

Ethnic data were provided for 11,697 (99%) of trainees. The ‘raw’ ethnic codes are shown in Table 7.1. ‘blank’, ‘*Please select...’ and ‘unknown’ were all treated as missing. The other codes were grouped into the main ethnic groups and are shown in Table 7.2. Just over half (53%) were White and 38% Asian. The percentage of Black (4%), Mixed (2%) and Other/Chinese (4%) were much smaller. Therefore for analysis, Black and Mixed ethnicities were combined with Other/Chinese to form a Black/Mixed/Chinese/Other group.
Table 7.1: ‘raw’ ethnic codes

<table>
<thead>
<tr>
<th>Code</th>
<th>Frequency</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asian – Bangladeshi</td>
<td>134</td>
<td>1.1</td>
</tr>
<tr>
<td>Asian – Indian</td>
<td>2266</td>
<td>19.2</td>
</tr>
<tr>
<td>Asian – Pakistani</td>
<td>945</td>
<td>8</td>
</tr>
<tr>
<td>Asian - Any other Asian Background</td>
<td>1044</td>
<td>8.8</td>
</tr>
<tr>
<td>Black – African</td>
<td>403</td>
<td>3.4</td>
</tr>
<tr>
<td>Black – Caribbean</td>
<td>28</td>
<td>0.2</td>
</tr>
<tr>
<td>Black - Any other Black Background</td>
<td>21</td>
<td>0.2</td>
</tr>
<tr>
<td>Mixed - White and Black African</td>
<td>29</td>
<td>0.2</td>
</tr>
<tr>
<td>Mixed - White and Black Caribbean</td>
<td>10</td>
<td>0.1</td>
</tr>
<tr>
<td>Mixed - Any other Mixed Background</td>
<td>205</td>
<td>1.7</td>
</tr>
<tr>
<td>White – British</td>
<td>5388</td>
<td>45.6</td>
</tr>
<tr>
<td>White – Irish</td>
<td>259</td>
<td>2.2</td>
</tr>
<tr>
<td>White - Any other White Background</td>
<td>511</td>
<td>4.3</td>
</tr>
<tr>
<td>Other – Chinese</td>
<td>181</td>
<td>1.5</td>
</tr>
<tr>
<td>Other - Any other Ethnic Group</td>
<td>273</td>
<td>2.3</td>
</tr>
<tr>
<td>Blank</td>
<td>21</td>
<td>0.2</td>
</tr>
<tr>
<td>*Please Select...</td>
<td>6</td>
<td>0.1</td>
</tr>
<tr>
<td>Unknown</td>
<td>88</td>
<td>0.7</td>
</tr>
<tr>
<td>Total</td>
<td>11812</td>
<td>100</td>
</tr>
</tbody>
</table>

Table 7.2: Ethnic groups

<table>
<thead>
<tr>
<th>Code</th>
<th>Frequency</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asian</td>
<td>4389</td>
<td>37.5</td>
</tr>
<tr>
<td>Black</td>
<td>452</td>
<td>3.9</td>
</tr>
<tr>
<td>Mixed</td>
<td>244</td>
<td>2.1</td>
</tr>
<tr>
<td>Other/ Chinese</td>
<td>454</td>
<td>3.9</td>
</tr>
<tr>
<td>White</td>
<td>6158</td>
<td>52.6</td>
</tr>
<tr>
<td>Total</td>
<td>11697</td>
<td>100</td>
</tr>
</tbody>
</table>

Table 7.3 displays the event codes tagged in the learning logs. Event codes 0 (lost records), 9 (presentation) and 14 (other) were deleted leaving 6,153,376 records; codes 1, 3 and 4 were merged into 36, 33 and 34 respectively.
<table>
<thead>
<tr>
<th>Event Code</th>
<th>Event name</th>
<th>Total entries</th>
<th>Action?</th>
<th>Kept entries</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Lost Records</td>
<td>27</td>
<td>deleted</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Lecture</td>
<td>11</td>
<td>to 36</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Tutorial</td>
<td>20</td>
<td>to 33</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Reading</td>
<td>3</td>
<td>to 34</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Presentation</td>
<td>3</td>
<td></td>
<td>deleted</td>
</tr>
<tr>
<td>14</td>
<td>Other</td>
<td>17</td>
<td></td>
<td>deleted</td>
</tr>
<tr>
<td>31</td>
<td>Clinical Encounter</td>
<td>1991260</td>
<td></td>
<td>1991260</td>
</tr>
<tr>
<td>32</td>
<td>Professional Conversation</td>
<td>463957</td>
<td></td>
<td>463957</td>
</tr>
<tr>
<td>33</td>
<td>Tutorial</td>
<td>1079117</td>
<td></td>
<td>1079137</td>
</tr>
<tr>
<td>34</td>
<td>Reading</td>
<td>530905</td>
<td></td>
<td>530908</td>
</tr>
<tr>
<td>35</td>
<td>Course/Certificate</td>
<td>347428</td>
<td></td>
<td>347428</td>
</tr>
<tr>
<td>36</td>
<td>Lecture/Seminar</td>
<td>731700</td>
<td></td>
<td>731711</td>
</tr>
<tr>
<td>37</td>
<td>Out Of Hours session</td>
<td>580075</td>
<td></td>
<td>580075</td>
</tr>
<tr>
<td>38</td>
<td>Audit/Project</td>
<td>91727</td>
<td></td>
<td>91727</td>
</tr>
<tr>
<td>39</td>
<td>Significant Event Analysis</td>
<td>137546</td>
<td></td>
<td>137546</td>
</tr>
<tr>
<td>40</td>
<td>eLearning Session</td>
<td>199627</td>
<td></td>
<td>199627</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td><strong>6153423</strong></td>
<td></td>
<td><strong>6153376</strong></td>
</tr>
</tbody>
</table>

These 6,153,376 events were also tagged to curriculum areas, competencies and DOPS as displayed in Table 7.4. Given the sheer number of codes, the results section only reports on curriculum area codes with more than 90,000 entries and competency codes with more than 50,000 entries; smaller entries are grouped as ‘other’.
<table>
<thead>
<tr>
<th>Code</th>
<th>Curriculum area</th>
<th>n</th>
<th>Code</th>
<th>Competency</th>
<th>n</th>
<th>Code</th>
<th>DOPS</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Being a General Practitioner</td>
<td>11</td>
<td>21</td>
<td>Communication and consultation skills</td>
<td>156019</td>
<td>41</td>
<td>Ability to take skin surface specimens for mycology</td>
<td>11</td>
</tr>
<tr>
<td>1.1</td>
<td>To manage primary contact with patients, dealing with unselected problems.</td>
<td>1677</td>
<td>22</td>
<td>Practising holistically</td>
<td>111693</td>
<td>42</td>
<td>Application of simple dressings</td>
<td>25</td>
</tr>
<tr>
<td>1.2</td>
<td>To cover the full range of health conditions.</td>
<td>1112</td>
<td>23</td>
<td>Data gathering and interpretation</td>
<td>128401</td>
<td>43</td>
<td>Aspiration of effusion</td>
<td>16</td>
</tr>
<tr>
<td>1.3</td>
<td>To co-ordinate care with other professionals in primary care, and with others</td>
<td>1583</td>
<td>24</td>
<td>Making a diagnosis/decisions</td>
<td>193792</td>
<td>44</td>
<td>Breast examination</td>
<td>36</td>
</tr>
<tr>
<td>1.4</td>
<td>To master effective and appropriate care provision and health service utilisation.</td>
<td>1070</td>
<td>25</td>
<td>Clinical management</td>
<td>315547</td>
<td>45</td>
<td>Cannulation</td>
<td>29</td>
</tr>
<tr>
<td>1.5</td>
<td>To make available to the patient the appropriate services within the healthcare system.</td>
<td>1263</td>
<td>26</td>
<td>Managing medical complexity</td>
<td>91639</td>
<td>46</td>
<td>Cauterisation</td>
<td>30</td>
</tr>
<tr>
<td>1.6</td>
<td>To act as an advocate for the patient.</td>
<td>823</td>
<td>27</td>
<td>Primary care admin and IMT</td>
<td>58204</td>
<td>47</td>
<td>Cervical cytology</td>
<td>55</td>
</tr>
<tr>
<td>2</td>
<td>The General Practice Consultation</td>
<td>358351</td>
<td>28</td>
<td>Working with colleagues and in teams</td>
<td>148685</td>
<td>48</td>
<td>Cryotherapy</td>
<td>31</td>
</tr>
<tr>
<td>2.1</td>
<td>To adopt a person-centred approach in dealing with patients and their problems, both in the context of patient’s circumstances.</td>
<td>1822</td>
<td>29</td>
<td>Community orientation</td>
<td>86159</td>
<td>49</td>
<td>Curettage/shave excision</td>
<td>41</td>
</tr>
<tr>
<td>2.2</td>
<td>To use the general practice consultation to bring about an effective doctor-patient relationship, always respecting the patient’s autonomy.</td>
<td>994</td>
<td>30</td>
<td>Maintaining performance, learning and teaching</td>
<td>201446</td>
<td>50</td>
<td>ECG</td>
<td>36</td>
</tr>
<tr>
<td>2.3</td>
<td>To communicate, to set priorities and to act in partnership.</td>
<td>962</td>
<td>31</td>
<td>Maintaining an ethical approach</td>
<td>93253</td>
<td>51</td>
<td>Excision of skin lesions</td>
<td>69</td>
</tr>
<tr>
<td>2.4</td>
<td>To provide long-term continuity of care as determined by the needs of the patient, referring to continuing and co-ordinated care management.</td>
<td>506</td>
<td>32</td>
<td>Fitness to practise</td>
<td>65587</td>
<td>52</td>
<td>Female genital examination</td>
<td>97</td>
</tr>
<tr>
<td>3</td>
<td>Personal and Professional responsibilities</td>
<td>10</td>
<td>83.1</td>
<td>To relate specific decision-making processes to the prevalence and incidence of illness in the community.</td>
<td>512</td>
<td>53</td>
<td>Hormone replacement implants of all types</td>
<td>10</td>
</tr>
<tr>
<td>3.1</td>
<td>Clinical Governance</td>
<td>118919</td>
<td>83.3</td>
<td>To adopt appropriate working principles (e.g. incremental investigation, using time as a tool), and to tolerate uncertainty.</td>
<td>673</td>
<td>54</td>
<td>ID Injection</td>
<td>4</td>
</tr>
<tr>
<td>3.2</td>
<td>Patient Safety</td>
<td>318851</td>
<td>83.4</td>
<td>To intervene urgently when necessary.</td>
<td>1179</td>
<td>55</td>
<td>IM Injection</td>
<td>26</td>
</tr>
<tr>
<td>3.3</td>
<td>Ethics and Values Based Medicine</td>
<td>167745</td>
<td>83.5</td>
<td>To manage conditions which may present early and in an undifferentiated way.</td>
<td>686</td>
<td>56</td>
<td>Incision and drainage of abscess</td>
<td>21</td>
</tr>
<tr>
<td>3.4</td>
<td>Promoting equality and valuing diversity</td>
<td>97198</td>
<td>83.6</td>
<td>To make effective and efficient use of diagnostic and therapeutic interventions.</td>
<td>790</td>
<td>57</td>
<td>IV Infusions</td>
<td>18</td>
</tr>
<tr>
<td>-------</td>
<td>----------------------------------------</td>
<td>-------</td>
<td>------</td>
<td>--------------------------------------------------------------------------------</td>
<td>------</td>
<td>----</td>
<td>---------------</td>
<td>----</td>
</tr>
<tr>
<td>3.5</td>
<td>Evidence-based Practice</td>
<td>261320</td>
<td>84.1</td>
<td>To simultaneously manage multiple complaints and pathologies, both acute and chronic health problems.</td>
<td>493</td>
<td>58</td>
<td>IV Injection</td>
<td>10</td>
</tr>
<tr>
<td>3.6</td>
<td>Research and Academic Activity</td>
<td>94426</td>
<td>84.2</td>
<td>To promote health and well-being by applying health promotion and disease prevention strategies appropriately.</td>
<td>595</td>
<td>59</td>
<td>Joint and peri-articular injections</td>
<td>70</td>
</tr>
<tr>
<td>3.7</td>
<td>Teaching, Mentoring and Clinical Supervision</td>
<td>169853</td>
<td>Total</td>
<td>1655353</td>
<td>60</td>
<td>Local Anaesthetics</td>
<td>34</td>
<td></td>
</tr>
<tr>
<td>4.1</td>
<td>Management in Primary Care</td>
<td>300263</td>
<td></td>
<td></td>
<td></td>
<td>Male genital examination</td>
<td>32</td>
<td></td>
</tr>
<tr>
<td>4.2</td>
<td>Information Management and Technology</td>
<td>109543</td>
<td></td>
<td></td>
<td></td>
<td>Perform and interpret spirometry and peak flow</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>4.3</td>
<td>To manage and co-ordinate health promotion, prevention, cure, care, rehabilitation and palliation</td>
<td>459</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Healthy People: promoting health and preventing disease</td>
<td>176042</td>
<td></td>
<td></td>
<td></td>
<td>Prostate examination</td>
<td>23</td>
<td></td>
</tr>
<tr>
<td>5.1</td>
<td>To reconcile the health needs of individual patients &amp; the health needs of the community in which they live, balancing these with available resources.</td>
<td>922</td>
<td></td>
<td></td>
<td></td>
<td>Rectal examination</td>
<td>26</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Genetics in Primary Care</td>
<td>37028</td>
<td></td>
<td></td>
<td></td>
<td>SC Injection</td>
<td>21</td>
<td></td>
</tr>
<tr>
<td>6.1</td>
<td>To use bio-psycho-social models, taking into account cultural and existential dimensions.</td>
<td>827</td>
<td></td>
<td></td>
<td></td>
<td>Suturing of skin wound</td>
<td>64</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Care of Acutely Ill People</td>
<td>244655</td>
<td></td>
<td></td>
<td></td>
<td>Testing for blood glucose</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Care of Children and Young People</td>
<td>266316</td>
<td></td>
<td></td>
<td></td>
<td>Urethral Catheterisation</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Care of Older Adults</td>
<td>167800</td>
<td></td>
<td></td>
<td></td>
<td>Veneupuncture</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>10.1</td>
<td>Women’s health</td>
<td>210484</td>
<td></td>
<td></td>
<td></td>
<td>total</td>
<td>894</td>
<td></td>
</tr>
<tr>
<td>10.2</td>
<td>Men’s health</td>
<td>65186</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Sexual Health</td>
<td>92337</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Care of People with Cancer &amp; Palliative Care</td>
<td>96943</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Care of People with Mental Health Problems</td>
<td>164406</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Care of people with Learning Disabilities</td>
<td>41801</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15.1</td>
<td>Cardiovascular problems</td>
<td>114183</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15.1</td>
<td>Skin problems</td>
<td>107667</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15.2</td>
<td>Digestive problems</td>
<td>91195</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15.3</td>
<td>Drug and Alcohol problems</td>
<td>59984</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15.4</td>
<td>ENT and facial problems</td>
<td>79859</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15.5</td>
<td>Eye problems</td>
<td>52277</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15.6</td>
<td>Metabolic Problems</td>
<td>90677</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The outcome scores of ARCPs are displayed in Table 7.5. ARCPs with Satisfactory Progress - academic (n=26) and OOP Clinical/Research/Career Break/Maternity (n=632) were deleted as they are not likely to have the same level of WPBAs. This left 21952 ARCP records for analysis.

Table 7.5: Outcome scores of ARCPs

<table>
<thead>
<tr>
<th>ARCP outcome</th>
<th>n</th>
<th>%</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Satisfactory Final – clinical</td>
<td>5415</td>
<td>23.9</td>
<td>5</td>
</tr>
<tr>
<td>Satisfactory Progress – clinical</td>
<td>14366</td>
<td>63.5</td>
<td>5</td>
</tr>
<tr>
<td>Unsatisfactory - Insufficient evidence</td>
<td>884</td>
<td>3.9</td>
<td>4</td>
</tr>
<tr>
<td>Unsatisfactory - No additional training</td>
<td>401</td>
<td>1.8</td>
<td>3</td>
</tr>
<tr>
<td>Unsatisfactory - additional training</td>
<td>843</td>
<td>3.7</td>
<td>2</td>
</tr>
<tr>
<td>Unsatisfactory - Released</td>
<td>43</td>
<td>0.2</td>
<td>1</td>
</tr>
<tr>
<td>Satisfactory Progress – academic</td>
<td>26</td>
<td>0.1</td>
<td>ignore</td>
</tr>
<tr>
<td>OOP Clinical/Research/Career Break/Maternity</td>
<td>632</td>
<td>2.8</td>
<td>ignore</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>22610</strong></td>
<td><strong>100.0</strong></td>
<td></td>
</tr>
</tbody>
</table>

Of the 22,610 ARCP entries on the database, 2334 were duplicates i.e. they had the same candidate number and ST year. Some of these duplicates seemed to be due to an initial unsatisfactory ARCP, which was then followed by another ARCP. Others duplicates looked to be simple errors, such as two entries both a year long in sequence, but both labelled ST1. Others were double entries, covering the same time period. Others looked as if an ARCP was undertaken part way through the year, then another at the end.

Of these duplicates, most (1994) gave the same ARCP score. Of these, 693 seemed to be multiple entries within the same ST year i.e. with identical start dates; Table 7.6 shows an example of this, with four satisfactory ST3 entries for the same trainee. In such cases, the latest end date was used and the duplicate entries deleted; this left 21,595 records. With other duplicates with identical ST year and score, the earliest start and latest end dates were used; this left 20,913 records.

Table 7.6: example ARCP coding ‘errors’ for one trainee

<table>
<thead>
<tr>
<th>Start</th>
<th>End</th>
<th>ARCP outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>ST3 01/10/2008</td>
<td>24/08/2010</td>
<td>Satisfactory Final – clinical</td>
</tr>
<tr>
<td>ST3 01/10/2008</td>
<td>30/09/2009</td>
<td>Satisfactory Progress - clinical</td>
</tr>
<tr>
<td>ST3 01/10/2008</td>
<td>30/09/2009</td>
<td>Satisfactory Progress - clinical</td>
</tr>
<tr>
<td>ST3 01/10/2008</td>
<td>02/11/2010</td>
<td>Satisfactory Progress - clinical</td>
</tr>
</tbody>
</table>
Duplicates were turned into single entries with multiple codes and ‘extension’ dates where it seemed appropriate e.g. if there were different ARCP outcomes with sequential dates or different start but the same end date, or the same start but different end date. Clear errors were dealt with as well as possible e.g. Satisfactory progress, then released, was interpreted as Unsatisfactory progress: additional training, then released (n=1). This led to 19929 ARCP entries as indicated in Table 7.7. The codes show that some trainees had more than one ARCP for a particular training year e.g. 1.2 means that these 20 trainees all had an ARCP ‘Unsatisfactory - additional training’ (i.e. code 2) and another ‘Unsatisfactory – released’ (i.e. code 1). Presumably the additional training ARCPs were given first, followed by ‘released’. This sequence is not included in these analyses, partly for simplicity, but mainly because in many cases the sequence was not clear. Also please note that some of the codes do not appear to make sense e.g. 1.25 and 3.5.

Table 7.7: Complete coding of multiple ARCP entries

<table>
<thead>
<tr>
<th>Outcome Code Description</th>
<th>Code</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unsatisfactory - released (total =41)</td>
<td>1</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>1.2</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>1.23</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>1.234</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1.25</td>
<td>1</td>
</tr>
<tr>
<td>Unsatisfactory - additional training (total = 697)</td>
<td>2</td>
<td>236</td>
</tr>
<tr>
<td></td>
<td>2.3</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td>2.345</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>2.35</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>2.4</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>2.45</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td>2.5</td>
<td>388</td>
</tr>
<tr>
<td>Unsatisfactory - no additional training (total = 355)</td>
<td>3</td>
<td>246</td>
</tr>
<tr>
<td></td>
<td>3.4</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>3.45</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>3.5</td>
<td>87</td>
</tr>
<tr>
<td>Unsatisfactory - insufficient evidence (total = 800)</td>
<td>4</td>
<td>466</td>
</tr>
<tr>
<td></td>
<td>4.5</td>
<td>334</td>
</tr>
<tr>
<td>Satisfactory Progress/ Final</td>
<td>5</td>
<td>18036</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>19929</td>
</tr>
</tbody>
</table>
Once each ST year had been given just one start and end date, the gaps and overlaps between years were investigated. A gap is when there are days between the recorded finish date of one ST year (e.g. ST1) and the start date of the next ST year (e.g. ST2). On overlap is when a trainee was simultaneously registered for two ST years. Table 7.8 shows that most trainees are recorded as starting ST2 the day after (82%) or on the same day (4%) as they ended ST1. However, for around 7% of these trainees, there is an overlap; we assume these are coding errors. There is a similar message regarding the gaps and overlaps between ST2 and 3.

Table 7.8: Gaps and overlaps between ST years

<table>
<thead>
<tr>
<th></th>
<th>ST1 and 2 (%)</th>
<th>ST2 and 3 (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n=4041</td>
<td>n=3758</td>
</tr>
<tr>
<td>More than a year gap</td>
<td>0.4</td>
<td>0.9</td>
</tr>
<tr>
<td>between 1 month and a year gap</td>
<td>1.5</td>
<td>3.0</td>
</tr>
<tr>
<td>between 1 day and 1 month gap</td>
<td>4.7</td>
<td>5.6</td>
</tr>
<tr>
<td>1 day gap</td>
<td>82.4</td>
<td>80.3</td>
</tr>
<tr>
<td>No gap</td>
<td>3.9</td>
<td>3.0</td>
</tr>
<tr>
<td>Between 1 day and 1 month overlap</td>
<td>0.7</td>
<td>1.0</td>
</tr>
<tr>
<td>Between 1 month and 1 year overlap</td>
<td>1.2</td>
<td>3.6</td>
</tr>
<tr>
<td>More than a year overlap</td>
<td>5.3</td>
<td>2.6</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>100.0</strong></td>
<td><strong>100.0</strong></td>
</tr>
</tbody>
</table>

The number of days for each training year is shown in Table 7.9. For each ST year, the majority of trainees take close to a year (between 360 and 370 days). This percentage falls from 92% in ST1, to 87% in ST2 and 78% in ST3.

Table 7.9: Duration of recorded ST year

<table>
<thead>
<tr>
<th>Days</th>
<th>ST1 (%)</th>
<th>ST2 (%)</th>
<th>ST3 (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n=7387</td>
<td>n=6358</td>
<td>n=6117</td>
</tr>
<tr>
<td>0 or less</td>
<td>1.2</td>
<td>2.3</td>
<td>5.3</td>
</tr>
<tr>
<td>1 to 299</td>
<td>1.2</td>
<td>1.9</td>
<td>2.9</td>
</tr>
<tr>
<td>300 to 359</td>
<td>0.8</td>
<td>0.6</td>
<td>1.1</td>
</tr>
<tr>
<td>360 to 370</td>
<td>91.8</td>
<td>88.6</td>
<td>77.9</td>
</tr>
<tr>
<td>371 to 430</td>
<td>0.5</td>
<td>0.6</td>
<td>3.5</td>
</tr>
<tr>
<td>431 to 730</td>
<td>1.2</td>
<td>1.8</td>
<td>6.2</td>
</tr>
<tr>
<td>731 or more</td>
<td>3.4</td>
<td>4.1</td>
<td>3.0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>100.0</strong></td>
<td><strong>100.0</strong></td>
<td><strong>100.0</strong></td>
</tr>
</tbody>
</table>

If we are looking at a particular year such as ST1, it makes sense to include data from all the trainees in that year. However, to consider the entire specialist training as a whole, we should only look at those trainees for whom we have reasonable data for ST1, 2 and 3. This complete data group of trainees was defined as follows:
1. Their ARCP is recorded for ST1, 2 and 3 (i.e. the 1851 trainees in the middle of Figure 1 in the results section)
2. And there is no overlap recorded between the ST years (n=1675)
3. And no ST year is recorded as being shorter than 300 days (n =1602)

Turning now to the coding of WPBAs, some COT entries contained 0 (n=499, 0.4%) or -1 (936, 0.8%) for the overall assessment. Looking at a few entries suggested many were regarded as good e.g. “excellent overall consultation” and a few clearly were not recorded e.g. “Same as above (this is a duplicate entry as forgot to mark overall status as competent)”. So these overall assessment scores were regarded as missing: these COTs with missing data were used to calculate the number of COTs and the ‘year decimal’, but not the average score.

Candidates could attempt the Applied Knowledge Test (AKT) as many times as they like (see Table 7.10). On the database, there were 9648 attempts by 8325 candidates. The analyses undertaken in this study required only one score per candidate; therefore all second and subsequent attempts were deleted. The data were for the sittings from Oct 2006 to April 2010 (Table 7.11), but the data from the first two dates were deleted as we only have pass/fail data.

Table 7.10: number of attempts at AKT by candidates

<table>
<thead>
<tr>
<th>Attempts</th>
<th>n</th>
<th>Candidates</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>7420</td>
<td>7420</td>
</tr>
<tr>
<td>2</td>
<td>1224</td>
<td>612</td>
</tr>
<tr>
<td>3</td>
<td>615</td>
<td>205</td>
</tr>
<tr>
<td>4</td>
<td>252</td>
<td>63</td>
</tr>
<tr>
<td>5</td>
<td>85</td>
<td>17</td>
</tr>
<tr>
<td>6</td>
<td>24</td>
<td>4</td>
</tr>
<tr>
<td>7</td>
<td>28</td>
<td>4</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>9648</strong></td>
<td><strong>8325</strong></td>
</tr>
</tbody>
</table>
Table 7.11: Number of Candidates at each AKT sitting: first attempts only

<table>
<thead>
<tr>
<th></th>
<th>Frequency</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oct-06</td>
<td>8</td>
<td>0.1</td>
</tr>
<tr>
<td>May-07</td>
<td>47</td>
<td>0.6</td>
</tr>
<tr>
<td>Oct-07</td>
<td>694</td>
<td>8.3</td>
</tr>
<tr>
<td>Jan-08</td>
<td>1072</td>
<td>12.9</td>
</tr>
<tr>
<td>May-08</td>
<td>733</td>
<td>8.8</td>
</tr>
<tr>
<td>Oct-08</td>
<td>1165</td>
<td>14</td>
</tr>
<tr>
<td>Jan-09</td>
<td>708</td>
<td>8.5</td>
</tr>
<tr>
<td>Apr-09</td>
<td>893</td>
<td>10.7</td>
</tr>
<tr>
<td>Oct-09</td>
<td>1312</td>
<td>15.8</td>
</tr>
<tr>
<td>Jan-10</td>
<td>694</td>
<td>8.3</td>
</tr>
<tr>
<td>Apr-10</td>
<td>999</td>
<td>12</td>
</tr>
<tr>
<td>Total</td>
<td>8325</td>
<td>100</td>
</tr>
</tbody>
</table>

As with the AKT, candidates could attempt the CSA as often as they like. Data from 6315 trainees are on the database, with 85% having one attempt (so far) and 11% had two attempts. Second and subsequent attempts were removed from the database.

Table 7.12: number of attempts at CSA by candidate

<table>
<thead>
<tr>
<th>Attempts</th>
<th>Candidates</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5358</td>
<td>85</td>
</tr>
<tr>
<td>2</td>
<td>716</td>
<td>11</td>
</tr>
<tr>
<td>3</td>
<td>169</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>48</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>17</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>6315</td>
<td>100</td>
</tr>
</tbody>
</table>

In the first CSA in October 2007, nine passes out of twelve were required (thereafter it was eight passes), so the results of the 36 trainees in this sitting were deleted, leaving 6279 trainees.

Once the AKT and CSA first attempts had been merged with the ARCP data, the ST year for these assessments was taken from the ARCP data (excluding extensions). There were 356 AKTs assigned to ST1, 2690 to ST2 and 4392 to ST3. However, as seen in Table 7.13, this assigning was problematic; 1058 AKTs were not assigned to any of these ST years and 241 were assigned to two ST years and 39 to all 3 ST years. There were 86 CSAs assigned to ST1, 179 to ST2 and 5616 to ST3. However, 629 were not assigned, 101 were assigned to two years and 72 to three.
Table 7.13: Assignment of AKT and CSA in ST years

<table>
<thead>
<tr>
<th>Number of ST Years</th>
<th>AKTs</th>
<th>CSA</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 (i.e. not assigned)</td>
<td>1058</td>
<td>629</td>
</tr>
<tr>
<td>1</td>
<td>6839</td>
<td>5463</td>
</tr>
<tr>
<td>2 (i.e. assigned to 2 ST years)</td>
<td>241</td>
<td>101</td>
</tr>
<tr>
<td>3</td>
<td>39</td>
<td>72</td>
</tr>
<tr>
<td>Total</td>
<td>8177</td>
<td>6265</td>
</tr>
</tbody>
</table>

During the analyses, numerous issues with the data were uncovered. Where the impact was minor, the decision was taken not to re-calculate previous tables. For example, five of the GP trainees were actually test entries, as a result of which the number of trainees is five too high in parts of this ‘data management and cleaning’ section. All analyses in the Results section, however, were re-calculated as required.

7.3 Aims

The primary aim of this analysis was to use inferential statistics to investigate the reliability, predictive validity and fairness of four of the Workplace based assessments (WPBA): Case-based discussion (CBD), the mini Clinical Evaluation Exercise (Mini-CEX), the Consultation Observation Tool (COT) and the Clinical Supervisors Report (CSR). A secondary aim was to explore the ‘tagged events’ that trainees include in their learning log.

Standard measures of reliability are difficult to apply to these data for four reasons. First, any number of each WPBA can be recorded for each trainee. Second, the WPBAs have a very limited scoring range. Third, due to variation in the cases considered, the assessments can be of very different levels of difficulty. Fourth, the identity of the assessors was not known. Given these difficulties, reliability was considered in terms of correlations between different types of WPBAs in the same training year and across training years for the same type of assessment.

Predictive validity was addressed by creating regression models to see how well WPBAs predicted the Annual Review of Competence Progression (ARCP) outcomes and the Applied Knowledge Test (AKT) and Clinical Skills Assessment (CSA) summative assessments.

Fairness was restricted to considering whether Place of training, Gender, Age and Ethnicity are associated with differential outcomes in these assessments.

---

2 The aims for this analysis were developed in discussion with many members of the RCGP, particularly Amar Rughani: we are extremely grateful for this advice.
These aims led to the following five research questions:

**RQ1 Tagged events**

a) How many tagged events have trainees put in their portfolios?
b) What is the distribution across categories?
c) Is there an association between the number of tagged events and performance at ARCPs?

**RQ2 Demographic data**

a) What are the Place of training, Gender, Age and Ethnicity of trainees?
b) What is the impact of these demographics on the assessments considered here?

**RQ3 Correlations between WPBAs (CBD, COT, mini-CEX) and CSR**

a) Within an ST year i.e. Do trainees who ‘do well’ on one of these assessments also do well on the others, compared with other trainees?
b) Across ST years i.e. for the same type of assessment, do trainees who do well in one ST year also do well in the next ST year?

**RQ4 Associations with ARCP outcome**

a) How well do the variables in RQ 1, 2 and 3 ‘predict’ ARCP performance?

**RQ5 Associations with summative assessment**

a) How well do the variables in RQ 1, 2, 3 and 4 ‘predict’ AKT and CSA outcomes?

‘Predict’ is used in RQ4 and 5 because regression analyses are used to model whether WPBAs and other variables ‘predict’ or ‘forecast’ outcomes at ARCP and summative assessment. As discussed at the end of this section, this does not imply direct causation.

### 7.4 Results

The final database contained 19,929 Specialty Training (ST) years from 11,812 trainees. Figure 7.1 displays the distribution of these ST years. There were 3224 who only completed ST1, presumably most started GP HST in 2009/10, although a few may have started earlier but not progressed onto ST2. The 2163 that only completed ST3 presumably were finishing their specialist training when this database started. Harder to explain are the 225 trainees who only completed ST2; these relatively few trainees may have left the course or missed years due to maternity or illness. Of those with two
recorded ST years, 2243 are from ST1 and 2, 2039 from ST2 and 3, and strangely 64 are from ST1 and 3 (presumably this last, small group are due to coding errors). Perhaps the most important group are those with ST1, 2 and 3 ARCPs (n=1851). As described in the Data Cleaning and Management section, above, the **complete data** group of trainees was defined as having ARCPs for ST1, 2 and 3, no overlap recorded between their ST years, and no ST year shorter than 300 days (n=1602).

These numbers are consistent with there being about 2000 to 3000 trainees in each cohort i.e. the five numbers in Figure 7.1 between 1851 and 3224 probably will largely consist of year cohorts; e.g. the cohort that started specialty training in 2005 will only have ST3 scores (n=2163) and the 2006 cohort will have ST2 and 3 scores (n=2039). Finally, 3 trainees on the database had only completed ST4. Not shown on the diagram is the pattern of ARCPs from the rest of the 66 ST4 trainees: this is due to the small numbers, complexity and apparent errors.

**Figure 7.1: Venn diagram of Specialty Training years**

As expected, there are fairly similar numbers of ARCPs across ST1 (7387), ST2 (6358) and ST3 (6117), with a few ST4s (67). In each training year, the vast majority are deemed satisfactory (without any unsatisfactory ARCPs); this figure was 93% for ST1, 92% for ST2, 86% for ST3 and 91% for ST4 (Table 7.14).
Table 7.14: lowest³ ARCP outcome by ST year

<table>
<thead>
<tr>
<th></th>
<th>ST1</th>
<th>ST2</th>
<th>ST3</th>
<th>ST4</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>%</td>
<td>n</td>
<td>%</td>
<td>n</td>
</tr>
<tr>
<td>Satisfactory Final – clinical</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4746</td>
</tr>
<tr>
<td>Satisfactory Progress - clinical</td>
<td>6868</td>
<td>93</td>
<td>5854</td>
<td>92</td>
<td>507</td>
</tr>
<tr>
<td>Unsatisfactory - Insufficient evidence</td>
<td>335</td>
<td>5</td>
<td>333</td>
<td>5</td>
<td>136</td>
</tr>
<tr>
<td>Unsatisfactory - No additional training</td>
<td>145</td>
<td>2</td>
<td>133</td>
<td>2</td>
<td>76</td>
</tr>
<tr>
<td>Unsatisfactory - additional training</td>
<td>27</td>
<td>0</td>
<td>36</td>
<td>1</td>
<td>625</td>
</tr>
<tr>
<td>Unsatisfactory - Released</td>
<td>11</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>27</td>
</tr>
<tr>
<td>Total</td>
<td>7387</td>
<td></td>
<td>6358</td>
<td></td>
<td>6117</td>
</tr>
</tbody>
</table>

The outcome codes for the ARCPs were correlated between the training years using Kendall’s non-parametric correlation. Given the large numbers involved, all the correlations were highly significant (p<0.001). However, the correlations between ST1, ST2 and ST3 varied from 0.11 to 0.16 i.e. they were small. The correlation between ST3 and ST4 was 0.51 for 47 trainees.

Table 7.15: Correlations across ST years for ARCP outcomes

<table>
<thead>
<tr>
<th>Kendall's tau_b</th>
<th>ST2</th>
<th>ST3</th>
<th>ST4</th>
</tr>
</thead>
<tbody>
<tr>
<td>ST1</td>
<td>tau</td>
<td>.149</td>
<td>.106</td>
</tr>
<tr>
<td></td>
<td>p</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>n</td>
<td>4094</td>
<td>1915</td>
</tr>
<tr>
<td>ST2</td>
<td>tau</td>
<td>.157</td>
<td>.</td>
</tr>
<tr>
<td></td>
<td>p</td>
<td>.000</td>
<td>.</td>
</tr>
<tr>
<td></td>
<td>n</td>
<td>3890</td>
<td>22²</td>
</tr>
<tr>
<td>ST3</td>
<td>tau</td>
<td>.157</td>
<td>.</td>
</tr>
<tr>
<td></td>
<td>p</td>
<td>.000</td>
<td>.</td>
</tr>
<tr>
<td></td>
<td>n</td>
<td>3890</td>
<td>.</td>
</tr>
</tbody>
</table>

³ i.e. if 2 or more ARCPs are recorded, the least good is displayed here

⁴ All ST1 ARCPs were satisfactory, so can’t compute a correlation

⁵ All ST4 ARCPs were satisfactory, so can’t compute a correlation
Table 7.16 displays the deaneries for the trainees when they were in ST1 (this is not the complete dataset, as some were in ST2 or 3 when the database started). Also, note the changes of deanery names in Yorkshire.

Table 7.16: Deaneries during ST1

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Armed Forces</td>
<td>76</td>
<td>1.0</td>
</tr>
<tr>
<td>East Midlands</td>
<td>343</td>
<td>4.6</td>
</tr>
<tr>
<td>East of England</td>
<td>538</td>
<td>7.3</td>
</tr>
<tr>
<td>East Scotland</td>
<td>79</td>
<td>1.1</td>
</tr>
<tr>
<td>Kent, Surrey, Sussex</td>
<td>717</td>
<td>9.7</td>
</tr>
<tr>
<td>London</td>
<td>835</td>
<td>11.3</td>
</tr>
<tr>
<td>Mersey</td>
<td>322</td>
<td>4.4</td>
</tr>
<tr>
<td>North Scotland</td>
<td>146</td>
<td>2.0</td>
</tr>
<tr>
<td>North Western</td>
<td>544</td>
<td>7.4</td>
</tr>
<tr>
<td>Northern</td>
<td>308</td>
<td>4.2</td>
</tr>
<tr>
<td>Northern Ireland</td>
<td>169</td>
<td>2.3</td>
</tr>
<tr>
<td>Oxford</td>
<td>246</td>
<td>3.3</td>
</tr>
<tr>
<td>Severn</td>
<td>263</td>
<td>3.6</td>
</tr>
<tr>
<td>South East Scotland</td>
<td>145</td>
<td>2.0</td>
</tr>
<tr>
<td>South West Peninsula</td>
<td>156</td>
<td>2.1</td>
</tr>
<tr>
<td>South Yorkshire and South Humber</td>
<td>73</td>
<td>1.0</td>
</tr>
<tr>
<td>Wales</td>
<td>353</td>
<td>4.8</td>
</tr>
<tr>
<td>Wessex</td>
<td>214</td>
<td>2.9</td>
</tr>
<tr>
<td>West Midlands</td>
<td>810</td>
<td>11.0</td>
</tr>
<tr>
<td>West Scotland</td>
<td>458</td>
<td>6.2</td>
</tr>
<tr>
<td>Yorkshire</td>
<td>168</td>
<td>2.3</td>
</tr>
<tr>
<td>Yorkshire and The Humber</td>
<td>419</td>
<td>5.7</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>7382</strong></td>
<td><strong>100.0</strong></td>
</tr>
</tbody>
</table>

The mean age of trainees on 31\textsuperscript{st} July 2010 was 32 from 11,728 (99\%) trainees; 59\% (6972) were female and 41\% (4823) male. To ease analysis, just three ethnic groups were used: White (n=6158, 53\%), Asian (4389, 38\%) and Black/Mixed/Chinese/Other (1150, 10\%). Also to ease analysis, just two places of undergraduate training were used: UK (n=8418, 72\%) and Outside the UK (3354, 29\%).
The age, gender, ethnicity and place of undergraduate training of trainees are linked. For the ‘complete data’ trainees, the average age of White trainees (30.4 years) was lower than for Asian (32.8) and Black/Mixed/Chinese/Other (33.1) trainees. Males’ average age was 32.6 compared with 30.5 for females, and UK trained graduates’ average age was 30.3, compared with 36.3 for those trained outside the UK. A higher percentage of Asian trainees (62%) were male compared with Black/Mixed/Chinese/Other (47%) and White (39%) trainees. A higher percentage of those with undergraduate training outside the UK were male (67%) compared with UK trained graduates (43%). Overall, 21% of this complete data group had completed undergraduate training outside the UK: 43% of Asians trained outside the UK, 4% of White and 38% of Black/Mixed/Chinese/Other trainees. These interactions mean that they should be considered together when considering their effect on WPBAs.

The tagged events on the trainees’ learning logs record both an event type (Table 7.17) and the learning achieved (Table 7.18). Perhaps of greatest note is the sheer number of events: 6,153,376 from 11,812 trainees, averages out at 521 events per trainee. For many of these trainees, this is from less than the full three years, so the actual average will be higher. For the ‘complete data’ trainees, the mean is 762. As shown in Figure 2, there is a huge range in the number of tagged events over the three ST years: the lowest was 108, and the highest, 4714! Clinical encounter (32%) was the most common event type, followed by Tutorials (18%) and Lecture/Seminar (12%).

<table>
<thead>
<tr>
<th>Event type</th>
<th>Number of entries</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinical Encounter</td>
<td>1991260</td>
<td>32</td>
</tr>
<tr>
<td>Professional Conversation</td>
<td>463957</td>
<td>8</td>
</tr>
<tr>
<td>Tutorials</td>
<td>1079137</td>
<td>18</td>
</tr>
<tr>
<td>Reading</td>
<td>530908</td>
<td>9</td>
</tr>
<tr>
<td>Course/Certificate</td>
<td>347428</td>
<td>6</td>
</tr>
<tr>
<td>Lecture/Seminar</td>
<td>731711</td>
<td>12</td>
</tr>
<tr>
<td>Out Of Hours session</td>
<td>580075</td>
<td>9</td>
</tr>
<tr>
<td>Audit/Project</td>
<td>91727</td>
<td>1</td>
</tr>
<tr>
<td>Significant Event Analysis</td>
<td>137546</td>
<td>2</td>
</tr>
<tr>
<td>eLearning Session</td>
<td>199627</td>
<td>3</td>
</tr>
<tr>
<td>Total</td>
<td>6153376</td>
<td>100</td>
</tr>
</tbody>
</table>

Table 7.17: Tagged events from the Learning Log
Figure 7.2: Histogram of the total tagged events by the complete data trainees

Table 7.18 displays the frequencies of the most common learning indicated on the tagged events; 73% of these are curriculum areas, 27% are competencies and less than 0.1% are DOPS. Since undertaking this analysis, we have been informed that the trainees tag to curriculum areas, indicating experience and the supervisors tag to competency areas. Therefore, 73% of entries are by the trainees, and 27% by their supervisors. Unfortunately we haven’t been able to incorporate this distinction in subsequent analyses.

Within the curriculum areas, the most common areas are broad ranging, such as the consultation, safety and children, women and older adults. Lower down Table 7.18 are more specific areas such as skin, neurological and digestive problems. Shown in Table 7.4, but not in Table 7.18, are those curriculum areas with fewer than 90,000 tagged events such as genetics (37,028 events) and “To reconcile the health needs of individual patients & the health needs of the community in which they live, balancing these with available resources” (922). It is not clear whether balancing patient and community needs is rarely considered in log entries or whether they are tagged to a broader area such as “Healthy People: promoting health and preventing disease” or the competency area of “Community orientation”.

97
<table>
<thead>
<tr>
<th>Curriculum area</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>The General Practice Consultation</td>
<td>358351</td>
<td>5.8</td>
</tr>
<tr>
<td>Patient Safety</td>
<td>318851</td>
<td>5.2</td>
</tr>
<tr>
<td>Management in Primary Care</td>
<td>300263</td>
<td>4.9</td>
</tr>
<tr>
<td>Care of Children and Young People</td>
<td>266316</td>
<td>4.3</td>
</tr>
<tr>
<td>Evidence-based Practice</td>
<td>261320</td>
<td>4.2</td>
</tr>
<tr>
<td>Care of Acutely Ill People</td>
<td>244655</td>
<td>4.0</td>
</tr>
<tr>
<td>Women's health</td>
<td>210484</td>
<td>3.4</td>
</tr>
<tr>
<td>Healthy People: promoting health and preventing disease</td>
<td>176042</td>
<td>2.9</td>
</tr>
<tr>
<td>Teaching, Mentoring and Clinical Supervision</td>
<td>169853</td>
<td>2.8</td>
</tr>
<tr>
<td>Care of Older Adults</td>
<td>167800</td>
<td>2.7</td>
</tr>
<tr>
<td>Ethics and Values Based Medicine</td>
<td>167745</td>
<td>2.7</td>
</tr>
<tr>
<td>Care of People with Mental Health Problems</td>
<td>164406</td>
<td>2.7</td>
</tr>
<tr>
<td>Clinical Governance</td>
<td>118919</td>
<td>1.9</td>
</tr>
<tr>
<td>Rheumatology and conditions of the musculoskeletal system (including trauma)</td>
<td>117682</td>
<td>1.9</td>
</tr>
<tr>
<td>Cardiovascular problems</td>
<td>114183</td>
<td>1.9</td>
</tr>
<tr>
<td>Respiratory problems</td>
<td>109990</td>
<td>1.8</td>
</tr>
<tr>
<td>Information Management and Technology</td>
<td>109543</td>
<td>1.8</td>
</tr>
<tr>
<td>Skin problems</td>
<td>107667</td>
<td>1.7</td>
</tr>
<tr>
<td>Neurological problems</td>
<td>100107</td>
<td>1.6</td>
</tr>
<tr>
<td>Promoting equality and valuing diversity</td>
<td>97198</td>
<td>1.6</td>
</tr>
<tr>
<td>Care of People with Cancer &amp; Palliative Care</td>
<td>96943</td>
<td>1.6</td>
</tr>
<tr>
<td>Research and Academic Activity</td>
<td>94426</td>
<td>1.5</td>
</tr>
<tr>
<td>Sexual Health</td>
<td>92337</td>
<td>1.5</td>
</tr>
<tr>
<td>Digestive problems</td>
<td>91195</td>
<td>1.5</td>
</tr>
<tr>
<td>Metabolic Problems</td>
<td>90677</td>
<td>1.5</td>
</tr>
<tr>
<td>Other</td>
<td>350176</td>
<td>5.7</td>
</tr>
<tr>
<td>Total</td>
<td>4497129</td>
<td>73.1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Competency</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinical management</td>
<td>315547</td>
<td>5.1</td>
</tr>
<tr>
<td>Maintaining performance, learning and teaching</td>
<td>201446</td>
<td>3.3</td>
</tr>
<tr>
<td>Making a diagnosis/decisions</td>
<td>193792</td>
<td>3.1</td>
</tr>
<tr>
<td>Communication and consultation skills</td>
<td>156019</td>
<td>2.5</td>
</tr>
<tr>
<td>Working with colleagues and in teams</td>
<td>148685</td>
<td>2.4</td>
</tr>
<tr>
<td>Data gathering and interpretation</td>
<td>128401</td>
<td>2.1</td>
</tr>
<tr>
<td>Practising holistically</td>
<td>111693</td>
<td>1.8</td>
</tr>
<tr>
<td>Maintaining an ethical approach</td>
<td>93253</td>
<td>1.5</td>
</tr>
<tr>
<td>Managing medical complexity</td>
<td>91639</td>
<td>1.5</td>
</tr>
<tr>
<td>Community orientation</td>
<td>86159</td>
<td>1.4</td>
</tr>
<tr>
<td>Fitness to practise</td>
<td>65587</td>
<td>1.1</td>
</tr>
<tr>
<td>Primary care admin and IMT</td>
<td>58204</td>
<td>0.9</td>
</tr>
<tr>
<td>Other</td>
<td>4928</td>
<td>0.1</td>
</tr>
<tr>
<td>Total</td>
<td>1655353</td>
<td>26.9</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>DOPS</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>894</td>
<td>0.0</td>
</tr>
<tr>
<td>Total</td>
<td>6153376</td>
<td>100.0</td>
</tr>
</tbody>
</table>
7.4.1 CBD

Descriptives for case-based discussions (CBD) are shown for all trainees in Table 7.19 and for the ‘complete data’ trainees in Table 7.20. In each year, the mean number of CBDs is slightly higher than minimum, with about seven CBDs in ST1 and 2 where the minimum is six, and about 13 in ST3 where the minimum is 12. This issue of how many trainees complete the mandatory minimum number of WPBAs was analysed in detail in our first report so is not considered here (Bedward et al., 2008). The average score is a little higher than 3 (competent) in all years, although not as high in ST3 as ST1 and 2. This is in line with the finding in our first report, which suggests that this fall is due to the ST3 CBDs being assessed by GP trainers whereas in ST1 and 2, the assessors are often in secondary care. The average ‘year decimal’ in the two tables varies between 0.52 and 0.60. This means that the mean time to record CBDs is just over half way through the training year.

Table 7.19: Number of CBDs with their average score and ‘year decimal’: all data

<table>
<thead>
<tr>
<th>ST Year</th>
<th>n</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of CBDs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>7338</td>
<td>1</td>
<td>46</td>
<td>7.29</td>
<td>3.12</td>
</tr>
<tr>
<td>2</td>
<td>6317</td>
<td>1</td>
<td>46</td>
<td>7.48</td>
<td>3.59</td>
</tr>
<tr>
<td>3</td>
<td>6061</td>
<td>1</td>
<td>36</td>
<td>13.08</td>
<td>3.32</td>
</tr>
<tr>
<td>Total</td>
<td>11789</td>
<td>1</td>
<td>51</td>
<td>16.33</td>
<td>6.89</td>
</tr>
<tr>
<td>Average CBD Score</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>7338</td>
<td>1.83</td>
<td>4.00</td>
<td>3.22</td>
<td>.35</td>
</tr>
<tr>
<td>2</td>
<td>6317</td>
<td>1.00</td>
<td>4.00</td>
<td>3.21</td>
<td>.35</td>
</tr>
<tr>
<td>3</td>
<td>6061</td>
<td>1.80</td>
<td>4.00</td>
<td>3.07</td>
<td>.31</td>
</tr>
<tr>
<td>Total</td>
<td>11789</td>
<td>1.80</td>
<td>4.00</td>
<td>3.15</td>
<td>.28</td>
</tr>
<tr>
<td>Average ‘year decimal’</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>7338</td>
<td>.20</td>
<td>1.00</td>
<td>.57</td>
<td>.09</td>
</tr>
<tr>
<td>2</td>
<td>6317</td>
<td>.14</td>
<td>1.00</td>
<td>.58</td>
<td>.10</td>
</tr>
<tr>
<td>3</td>
<td>6061</td>
<td>.13</td>
<td>1.00</td>
<td>.56</td>
<td>.12</td>
</tr>
</tbody>
</table>
Table 7.20: Number of CBDs with their average score and ‘year decimal’: complete data trainees

<table>
<thead>
<tr>
<th>ST Year</th>
<th>n</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of CBDs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1598</td>
<td>2</td>
<td>17</td>
<td>6.55</td>
<td>1.35</td>
</tr>
<tr>
<td>2</td>
<td>1600</td>
<td>3</td>
<td>22</td>
<td>6.90</td>
<td>1.63</td>
</tr>
<tr>
<td>3</td>
<td>1601</td>
<td>2</td>
<td>24</td>
<td>12.58</td>
<td>1.60</td>
</tr>
<tr>
<td>Total</td>
<td>1601</td>
<td>2</td>
<td>51</td>
<td>26.06</td>
<td>3.15</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Average CBD Score</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1598</td>
<td>2.00</td>
<td>4.00</td>
<td>3.31</td>
<td>0.34</td>
</tr>
<tr>
<td>2</td>
<td>1600</td>
<td>1.83</td>
<td>4.00</td>
<td>3.24</td>
<td>0.34</td>
</tr>
<tr>
<td>3</td>
<td>1601</td>
<td>1.85</td>
<td>4.00</td>
<td>3.07</td>
<td>0.31</td>
</tr>
<tr>
<td>Total</td>
<td>1601</td>
<td>2.50</td>
<td>3.88</td>
<td>3.18</td>
<td>0.22</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Average ‘year decimal’</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1598</td>
<td>.27</td>
<td>.90</td>
<td>.60</td>
<td>.08</td>
</tr>
<tr>
<td>2</td>
<td>1600</td>
<td>.30</td>
<td>.87</td>
<td>.56</td>
<td>.08</td>
</tr>
<tr>
<td>3</td>
<td>1601</td>
<td>.17</td>
<td>.96</td>
<td>.52</td>
<td>.06</td>
</tr>
</tbody>
</table>

One measure of the reliability of CBDs is whether the same trainees ‘do well’ in ST1 as in ST2, etc. These correlations across ST years are displayed in Table 7.21. Although all the correlations are significant, they are fairly modest: the number of CBDs, the average score and the ‘year decimal’ are only weakly correlated across training year. For the average score, the correlation was highest between ST1 and 2 at 0.20, 0.11 between ST2 and 3, and lowest between ST1 and 3 at 0.09. The correlation between the total count and average score across the three years was -0.055 (p=0.002), so those with higher average scores completed slightly fewer CBDs (not shown in Table 7.21).
Using the ‘complete data’ trainees (n=1582), multiple regression was used for the total number of CBDs although there was large positive skew i.e. a ‘tail’ of trainees recorded large numbers of CBDs. Using Gender, Age, Place of training and Ethnicity as predictors, Place of training was entered first, indicating that it was the most important predictor, then Asian, then Age, then Black/ Mixed/ Chinese/ Other; gender was not entered, indicating it was not statistically significantly associated with total number of CBDs, with Place of training, Age and Ethnicity already in the model. The total variance explained by the model was just 6.6%, so Place of training, Age and Ethnicity are not strongly predictive of the number of CBDs. All coefficients were positive, meaning that non-UK trained, Older, Asian, Black/ Mixed/ Chinese/ Other trainees completed slightly more CBDs.

Average CBD score (n=1582) was approximately normally distributed. Re-doing this regression, Place of Training, then Asian, then Age, then Black/ Mixed/ Chinese/ Other were entered; this model accounted for 9.1% of the variance. All coefficients were negative, indicating that UK trained, younger, white trainees tended to score slightly higher on CBDs.

The ST3 ‘year decimal’ (n=1582) was approximately normally distributed. Only Asian then Black/ Mixed/ Chinese/ Other were entered into the model and accounted for just 0.8% of the variance. Both coefficients were positive indicating that CBDs for these ethnic groups tend to be recorded very slightly later in the year than their White colleagues.
The mean number of COTs, average score and average ‘year decimal’ are similar for all trainees and ‘complete data’ trainees (Tables 7.22 and 7.23). The number of trainees recording any COTs and the mean number of those who did record any are both low in ST1 and 2; this is because not all ST1 and 2 trainees have a GP placement. In ST3, the mean is 13.1, which is a little higher than the minimum of 12. The average score rises slightly across the years: 2.6 in ST1, 2.7 in ST2 and 3.0 in ST3. The ‘year decimals’ indicate that the average recorded time for a COT is just over half way through the training year; this is because 0.5 is halfway through the training year.

Table 7.22: Number of COTs with their average score and ‘year decimal’: all data

<table>
<thead>
<tr>
<th>ST Year</th>
<th>n</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of COTs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>3222</td>
<td>1</td>
<td>39</td>
<td>3.75</td>
<td>2.60</td>
</tr>
<tr>
<td>2</td>
<td>3218</td>
<td>1</td>
<td>39</td>
<td>4.48</td>
<td>3.43</td>
</tr>
<tr>
<td>3</td>
<td>6044</td>
<td>1</td>
<td>62</td>
<td>13.16</td>
<td>4.04</td>
</tr>
<tr>
<td>Total</td>
<td>10221</td>
<td>1</td>
<td>65</td>
<td>11.30</td>
<td>6.66</td>
</tr>
<tr>
<td>Average COT Score</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>3220</td>
<td>.67</td>
<td>4.00</td>
<td>2.60</td>
<td>.51</td>
</tr>
<tr>
<td>2</td>
<td>3214</td>
<td>.67</td>
<td>4.00</td>
<td>2.72</td>
<td>.50</td>
</tr>
<tr>
<td>3</td>
<td>6044</td>
<td>1.08</td>
<td>4.00</td>
<td>2.96</td>
<td>.33</td>
</tr>
<tr>
<td>Total</td>
<td>10221</td>
<td>1.00</td>
<td>4.00</td>
<td>2.85</td>
<td>.39</td>
</tr>
<tr>
<td>Average ‘year decimal’</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>3222</td>
<td>.12</td>
<td>1.00</td>
<td>.55</td>
<td>.23</td>
</tr>
<tr>
<td>2</td>
<td>3218</td>
<td>.12</td>
<td>1.00</td>
<td>.59</td>
<td>.23</td>
</tr>
<tr>
<td>3</td>
<td>6044</td>
<td>.13</td>
<td>1.00</td>
<td>.56</td>
<td>.13</td>
</tr>
</tbody>
</table>
Table 7.23: Number of COTs with their average score and ‘year decimal’: complete data trainees

<table>
<thead>
<tr>
<th>Number of COTs</th>
<th>ST Year</th>
<th>n</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>464</td>
<td>1</td>
<td>12</td>
<td>3.18</td>
<td>1.76</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>795</td>
<td>1</td>
<td>17</td>
<td>3.86</td>
<td>2.23</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1600</td>
<td>2</td>
<td>55</td>
<td>13.17</td>
<td>3.49</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>1600</td>
<td>2</td>
<td>55</td>
<td>16.02</td>
<td>4.74</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Average COT Score</th>
<th>ST Year</th>
<th>n</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>462</td>
<td>.67</td>
<td>4.00</td>
<td>2.66</td>
<td>.61</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>794</td>
<td>1.00</td>
<td>4.00</td>
<td>2.72</td>
<td>.49</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1600</td>
<td>2.00</td>
<td>4.00</td>
<td>2.98</td>
<td>.31</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>1600</td>
<td>2.04</td>
<td>4.00</td>
<td>2.95</td>
<td>.28</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Average ‘year decimal’</th>
<th>ST Year</th>
<th>n</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>464</td>
<td>.15</td>
<td>.96</td>
<td>.54</td>
<td>.22</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>795</td>
<td>.12</td>
<td>1.00</td>
<td>.56</td>
<td>.22</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1600</td>
<td>.16</td>
<td>.97</td>
<td>.52</td>
<td>.08</td>
</tr>
</tbody>
</table>

Table 7.24 shows the correlations across ST years for COTs. For the number of COTs and the ‘year decimal’, these correlations are all below 0.14 and many are not significant. For average score, the correlation is 0.26 between ST1 and 2, 0.16 between ST2 and 3, and 0.15 between ST1 and 3. These average score correlations are a little higher than for CBDs; but they are still low and include few trainees.
Table 7.24: Correlations between ST years for COTs for ‘complete data’ trainees

<table>
<thead>
<tr>
<th></th>
<th>ST Year</th>
<th>ST2</th>
<th>ST3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of COTs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ST1</td>
<td>tau</td>
<td>.081</td>
<td>-.038</td>
</tr>
<tr>
<td></td>
<td>p</td>
<td>.102</td>
<td>.290</td>
</tr>
<tr>
<td></td>
<td>n</td>
<td>258</td>
<td>464</td>
</tr>
<tr>
<td>ST2</td>
<td>tau</td>
<td>.067</td>
<td></td>
</tr>
<tr>
<td></td>
<td>p</td>
<td>.015</td>
<td></td>
</tr>
<tr>
<td></td>
<td>n</td>
<td>795</td>
<td></td>
</tr>
<tr>
<td>Average score</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ST1</td>
<td>tau</td>
<td>.258</td>
<td>.145</td>
</tr>
<tr>
<td></td>
<td>p</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>255</td>
<td>462</td>
</tr>
<tr>
<td>ST2</td>
<td>tau</td>
<td>.157</td>
<td></td>
</tr>
<tr>
<td></td>
<td>p</td>
<td>.000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>n</td>
<td>794</td>
<td></td>
</tr>
<tr>
<td>Year decimal</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ST1</td>
<td>tau</td>
<td>.132</td>
<td>-.024</td>
</tr>
<tr>
<td></td>
<td>p</td>
<td>.002</td>
<td>.443</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>258</td>
<td>464</td>
</tr>
<tr>
<td>ST2</td>
<td>tau</td>
<td></td>
<td>.038</td>
</tr>
<tr>
<td></td>
<td>p</td>
<td></td>
<td>.106</td>
</tr>
<tr>
<td></td>
<td>n</td>
<td></td>
<td>795</td>
</tr>
</tbody>
</table>

With the ‘complete data’ trainees, multiple regression was used on the total number of COTs (n=1581), although there was some positive skew. After Place of training, Gender, then Asian, then Age were entered, and 3.8% of the variance was explained. Non-UK, older, Asian, male trainees completed slightly more COTs. Average COT scores were normally distributed; Place of training, Age and Gender were entered and accounted for 5.5% of the variance. UK trained, younger, female trainees scored slightly higher. With the ST3 ‘year decimal’ none of these variables was significant, so Place of training, Gender, Age and Ethnicity did not affect the average time in the training year for completing COTs.

7.4.3 Mini-CEX

The mean number of mini-CEX, average score and average ‘year decimal’ for all trainees and ‘complete data’ trainees are shown in Tables 7.25 and 7.26. It should be remembered that scoring is on a five point scale whereas COTs, CBDs and CSR use a four point scale: this explains the higher average scores. As mini-CEX assessments are undertaken in secondary care, far fewer trainees completed any in ST3. Also, in ST3, the mean number completed per trainee was lower and they tended to be completed earlier in the training year.
Table 7.25: Number of mini-CEXs with their average score and ‘year decimal’: all data

<table>
<thead>
<tr>
<th>ST Year</th>
<th>n</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of mini-CEXs</td>
<td>1</td>
<td>7223</td>
<td>1</td>
<td>30</td>
<td>5.69</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>6152</td>
<td>1</td>
<td>26</td>
<td>5.45</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1051</td>
<td>1</td>
<td>19</td>
<td>4.25</td>
</tr>
<tr>
<td>Average mini-CEX Score</td>
<td>1</td>
<td>7223</td>
<td>1.00</td>
<td>5.00</td>
<td>4.33</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>6152</td>
<td>1.00</td>
<td>5.00</td>
<td>4.33</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1051</td>
<td>1.00</td>
<td>5.00</td>
<td>4.22</td>
</tr>
<tr>
<td>Average ‘year decimal’</td>
<td>1</td>
<td>7223</td>
<td>.12</td>
<td>1.00</td>
<td>.59</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>6152</td>
<td>.12</td>
<td>1.00</td>
<td>.58</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1051</td>
<td>.11</td>
<td>1.00</td>
<td>.51</td>
</tr>
</tbody>
</table>

Table 7.26: Number of mini-CEXs with their average score and ‘year decimal’: complete data trainees

<table>
<thead>
<tr>
<th>ST Year</th>
<th>n</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of mini-CEXs</td>
<td>1</td>
<td>1579</td>
<td>1</td>
<td>16</td>
<td>5.69</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1576</td>
<td>1</td>
<td>18</td>
<td>5.17</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>212</td>
<td>1</td>
<td>12</td>
<td>3.30</td>
</tr>
<tr>
<td>Average mini-CEX Score</td>
<td>1</td>
<td>1579</td>
<td>1.75</td>
<td>5.00</td>
<td>4.36</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1576</td>
<td>1.00</td>
<td>5.00</td>
<td>4.34</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>212</td>
<td>1.00</td>
<td>5.00</td>
<td>4.22</td>
</tr>
<tr>
<td>Average ‘year decimal’</td>
<td>1</td>
<td>1579</td>
<td>.22</td>
<td>.99</td>
<td>.61</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1576</td>
<td>.14</td>
<td>.99</td>
<td>.57</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>212</td>
<td>.12</td>
<td>1.00</td>
<td>.46</td>
</tr>
</tbody>
</table>
Table 7.27 shows the correlations between ST years for mini-CEX. Once again, these correlations are low, with the largest being 0.16 for the average score between ST1 and 2.

**Table 7.27: Correlations between ST years for mini-CEX for ‘complete data’ trainees**

<table>
<thead>
<tr>
<th>ST Year</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Count</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>tau</td>
<td>.043</td>
<td>-.021</td>
</tr>
<tr>
<td></td>
<td>p</td>
<td>.035</td>
<td>.710</td>
</tr>
<tr>
<td></td>
<td>n</td>
<td>1555</td>
<td>206</td>
</tr>
<tr>
<td>2</td>
<td>tau</td>
<td>-.172</td>
<td></td>
</tr>
<tr>
<td></td>
<td>p</td>
<td>.002</td>
<td></td>
</tr>
<tr>
<td></td>
<td>n</td>
<td>204</td>
<td></td>
</tr>
<tr>
<td>Average score</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>tau</td>
<td>.162</td>
<td>.137</td>
</tr>
<tr>
<td></td>
<td>p</td>
<td>.000</td>
<td>.010</td>
</tr>
<tr>
<td></td>
<td>n</td>
<td>1555</td>
<td>206</td>
</tr>
<tr>
<td>2</td>
<td>tau</td>
<td>.019</td>
<td></td>
</tr>
<tr>
<td></td>
<td>p</td>
<td>.722</td>
<td></td>
</tr>
<tr>
<td></td>
<td>n</td>
<td>204</td>
<td></td>
</tr>
<tr>
<td>Year decimal</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>tau</td>
<td>.062</td>
<td>-.019</td>
</tr>
<tr>
<td></td>
<td>p</td>
<td>.000</td>
<td>.693</td>
</tr>
<tr>
<td></td>
<td>n</td>
<td>1555</td>
<td>206</td>
</tr>
<tr>
<td>2</td>
<td>tau</td>
<td>.053</td>
<td></td>
</tr>
<tr>
<td></td>
<td>p</td>
<td>.258</td>
<td></td>
</tr>
<tr>
<td></td>
<td>n</td>
<td>204</td>
<td></td>
</tr>
</tbody>
</table>

The distribution of the number of min-CEX was disjointed, meaning that we should be cautious interpreting the results from multiple regression. In ST1 the number of mini-CEX (n=1557) undertaken was only affected by Asian trainees completing more than average; this accounted for 0.7% of the variance. In ST2, Black/ Mixed/ Chinese/ Other, then Asian, then Gender were entered, but only accounted for 1.4% of the variance, with White males recording fewer mini-CEX. Given the low numbers, these analyses were not conducted on ST3.

With average scores, the data are approximately normal, but slightly disjointed. Place of training, then Black/ Mixed/ Chinese/ Other then Asian were entered for ST1, accounting for 3.9% of the variance with UK trained, white trainees scoring higher. In ST2, only Place of training was entered, accounting for 4.5% of the variance; UK trained scored higher.

The ‘year decimal’ data were normally distributed in ST1, but only 0.2% of the variance was accounted for by Age (older trainees recorded min-CEX very slightly later). In ST2, only Place of training was a significant predictor of the mini-CEX year decimal, accounting for 0.2% of the variance, with UK trained graduates completing them very slightly later.
7.4.4 CSR

Tables 7.28 and 7.29 display the number of Clinical Supervisors' Reports (CSRs), their average score and ‘year decimal’ for the three training years. On average, trainees in ST1 and 2 had just over two clinical supervisors reports (CSRs). Far fewer trainees had any CSRs in ST3 as they are usually only given at the end of placements in ST1 and 2. It is not known how many of these are coding errors, but it is noticeable that the ST3 CSRs for ‘complete data’ trainees were recorded very early in the year: presumably they were finishing off secondary care placements. In all years, the mean mark was between 3.2 and 3.4 i.e. between ‘competent’ and ‘excellent’.

Table 7.28: Number of CSRs with their average score and ‘year decimal’: all data

<table>
<thead>
<tr>
<th>Number of CSRs</th>
<th>ST Year</th>
<th>n</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6043</td>
<td>1</td>
<td>9</td>
<td>2.05</td>
<td>.91</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>5216</td>
<td>1</td>
<td>9</td>
<td>2.22</td>
<td>1.06</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1697</td>
<td>1</td>
<td>7</td>
<td>1.55</td>
<td>.86</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Average CSR Score</th>
<th>ST Year</th>
<th>n</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6043</td>
<td>1.00</td>
<td>4.00</td>
<td>3.34</td>
<td>.47</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>5216</td>
<td>1.00</td>
<td>4.00</td>
<td>3.35</td>
<td>.45</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1694</td>
<td>1.00</td>
<td>4.00</td>
<td>3.24</td>
<td>.56</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Average ‘year decimal’</th>
<th>ST Year</th>
<th>n</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6043</td>
<td>.01</td>
<td>1.00</td>
<td>.59</td>
<td>.17</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>5216</td>
<td>.01</td>
<td>1.00</td>
<td>.58</td>
<td>.18</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1697</td>
<td>.00</td>
<td>1.00</td>
<td>.55</td>
<td>.27</td>
<td></td>
</tr>
</tbody>
</table>
Table 7.29: Number of CSRs with their average score and ‘year decimal’: complete data trainees

<table>
<thead>
<tr>
<th>ST Year</th>
<th>n</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of CSRs</td>
<td>1</td>
<td>1594</td>
<td>1</td>
<td>6</td>
<td>2.19</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1596</td>
<td>1</td>
<td>7</td>
<td>2.33</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>164</td>
<td>1</td>
<td>2</td>
<td>1.02</td>
</tr>
<tr>
<td>Average CSR Score</td>
<td>1</td>
<td>1594</td>
<td>1.00</td>
<td>4.00</td>
<td>3.41</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1596</td>
<td>1.00</td>
<td>4.00</td>
<td>3.37</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>164</td>
<td>1.00</td>
<td>4.00</td>
<td>3.21</td>
</tr>
<tr>
<td>Average ‘year decimal’</td>
<td>1</td>
<td>1594</td>
<td>.23</td>
<td>.98</td>
<td>.65</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1596</td>
<td>.19</td>
<td>.99</td>
<td>.61</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>164</td>
<td>.00</td>
<td>.36</td>
<td>.21</td>
</tr>
</tbody>
</table>

Table 7.30 shows that the correlations between years for CSR. Once again these are low, with 0.16 being the highest positive correlation.

Table 7.30: Correlations between ST years for CSR for ‘complete data’ trainees

<table>
<thead>
<tr>
<th>ST Year</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Count</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>tau</td>
<td>.053</td>
</tr>
<tr>
<td></td>
<td>p</td>
<td>.019</td>
</tr>
<tr>
<td></td>
<td>n</td>
<td>1590</td>
</tr>
<tr>
<td>2</td>
<td>tau</td>
<td>.159</td>
</tr>
<tr>
<td></td>
<td>p</td>
<td>.034</td>
</tr>
<tr>
<td></td>
<td>n</td>
<td>163</td>
</tr>
<tr>
<td>Average score</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>tau</td>
<td>.114</td>
</tr>
<tr>
<td></td>
<td>p</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>n</td>
<td>1590</td>
</tr>
<tr>
<td>2</td>
<td>tau</td>
<td>.018</td>
</tr>
<tr>
<td></td>
<td>p</td>
<td>.788</td>
</tr>
<tr>
<td></td>
<td>n</td>
<td>163</td>
</tr>
<tr>
<td>Year decimal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>tau</td>
<td>.122</td>
</tr>
<tr>
<td></td>
<td>p</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>n</td>
<td>1590</td>
</tr>
<tr>
<td>2</td>
<td>tau</td>
<td>.157</td>
</tr>
<tr>
<td></td>
<td>p</td>
<td>.003</td>
</tr>
<tr>
<td></td>
<td>n</td>
<td>163</td>
</tr>
</tbody>
</table>
For the number of CSRs at both ST1 and 2, no variable had a significant effect. Place of training, Black/Mixed/Chinese/Other, and Asian were entered for ST1 average score accounting for 6.2% of the variance, suggesting that UK trained, White trainees scored slightly higher. The distribution of scores was approximately normal, but very disjointed (n=1575). There were similar results in ST2 (n=1577), but only accounting for 3.2% of the variance. For the ‘year decimal’, no variables were significant in ST1, and only Age had a small effect in ST2, accounting for just 0.3% of the variance, with the CSRs for older trainees recorded very slightly earlier.

7.4.5 Correlations between the Workplace Based Assessments

Correlations between WPBAs are shown for ST1 in Table 7.31, ST2 in Table 7.32 and ST3 in Table 7.33. In ST1 all the correlations are between 0.21 and 0.35 apart from the Correlations of COT with CSR (0.10) and mini-CEX (0.09). It should be noted that COTs are undertaken in general practice, whereas CSR and mini-CEX are done in secondary care. This pattern in repeated in ST2 and 3, although the correlation between COT and CBD was 0.53 in ST3.

Table 7.31: ST1 average score correlations between WPBAs for all trainees

<table>
<thead>
<tr>
<th></th>
<th>Mini-CEX</th>
<th>CBD</th>
<th>COT</th>
</tr>
</thead>
<tbody>
<tr>
<td>CSR</td>
<td>tau</td>
<td>.211</td>
<td>.324</td>
</tr>
<tr>
<td></td>
<td>p</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>n</td>
<td>1575</td>
<td>1594</td>
</tr>
<tr>
<td>Mini-CEX</td>
<td>tau</td>
<td>.323</td>
<td>.092</td>
</tr>
<tr>
<td></td>
<td>p</td>
<td>.000</td>
<td>.010</td>
</tr>
<tr>
<td></td>
<td>n</td>
<td>1578</td>
<td>442</td>
</tr>
<tr>
<td>CBD</td>
<td>tau</td>
<td></td>
<td>.353</td>
</tr>
<tr>
<td></td>
<td>p</td>
<td></td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>n</td>
<td></td>
<td>462</td>
</tr>
</tbody>
</table>

Table 7.32: ST2 average score correlations between WPBAs for all trainees

<table>
<thead>
<tr>
<th></th>
<th>Mini-CEX</th>
<th>CBD</th>
<th>COT</th>
</tr>
</thead>
<tbody>
<tr>
<td>CSR</td>
<td>tau</td>
<td>.243</td>
<td>.309</td>
</tr>
<tr>
<td></td>
<td>p</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>n</td>
<td>1573</td>
<td>1596</td>
</tr>
<tr>
<td>Mini-CEX</td>
<td>tau</td>
<td>.286</td>
<td>.087</td>
</tr>
<tr>
<td></td>
<td>p</td>
<td>.000</td>
<td>.001</td>
</tr>
<tr>
<td></td>
<td>n</td>
<td>1576</td>
<td>770</td>
</tr>
<tr>
<td>CBD</td>
<td>tau</td>
<td></td>
<td>.334</td>
</tr>
<tr>
<td></td>
<td>p</td>
<td></td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>n</td>
<td></td>
<td>794</td>
</tr>
</tbody>
</table>
Table 7.33: ST3 average score correlations between WPBAs for all trainees

<table>
<thead>
<tr>
<th></th>
<th>Mini-CEX</th>
<th>CBD</th>
<th>COT</th>
</tr>
</thead>
<tbody>
<tr>
<td>CSR</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tau</td>
<td>.263</td>
<td>.239</td>
<td>.036</td>
</tr>
<tr>
<td>p</td>
<td>.034</td>
<td>.000</td>
<td>.567</td>
</tr>
<tr>
<td>n</td>
<td>49</td>
<td>164</td>
<td>164</td>
</tr>
<tr>
<td>Mini-CEX</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tau</td>
<td>.285</td>
<td>.185</td>
<td></td>
</tr>
<tr>
<td>p</td>
<td>.000</td>
<td>.000</td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>212</td>
<td>211</td>
<td></td>
</tr>
<tr>
<td>CBD</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tau</td>
<td></td>
<td>.532</td>
<td></td>
</tr>
<tr>
<td>p</td>
<td></td>
<td>.000</td>
<td></td>
</tr>
<tr>
<td>n</td>
<td></td>
<td>1600</td>
<td></td>
</tr>
</tbody>
</table>

7.4.6 ARCP

As there are too few possible outcomes to create a linear regression model, the ARCP outcomes were converted to binary: satisfactory and unsatisfactory; this enables us to use logistic regression. First the descriptors were entered stepwise (forward conditional) as they are ‘fixed’; then CBD with mini-CEX (in ST1 or 2) or COT (in ST3); then CSR; and finally the total tagged events. CSR is entered after the WPBAs as it may well be derived from them. The tagged events are last, so we can see if they provide increased predictive power beyond that which is provided by the WPBAs.

Table 7.34 shows the final output for the regression model for ST1. Probably the most important column is Exp (B) which is the change in odds of achieving a satisfactory ARCP per unit change of the predictor. So, being female and younger increases the odds of a satisfactory ARCP but there is no effect of Place of training or Ethnicity. More CBDs and a higher average score increase the odds, while completing CBDs later in the year reduces the odds of a satisfactory ARCP. Similarly, completing mini-CEX later reduces the odds; more mini-CEX increases the odds slightly. More tagged events increase the odds of a satisfactory ARCP. CSR and mini-CEX average score were not significant and so were excluded from the model.
Table 7.34: logistic regression for predictors of ST1 ARCP (all trainees, n=5900)

<table>
<thead>
<tr>
<th></th>
<th>B</th>
<th>S.E.</th>
<th>Wald</th>
<th>df</th>
<th>Sig.</th>
<th>Exp(B)</th>
<th>CI Lower</th>
<th>CI Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>-.063</td>
<td>.011</td>
<td>31.903</td>
<td>1</td>
<td>.000</td>
<td>.938</td>
<td>.918</td>
<td>.959</td>
</tr>
<tr>
<td>Gender</td>
<td>-.385</td>
<td>.111</td>
<td>11.951</td>
<td>1</td>
<td>.001</td>
<td>.681</td>
<td>.547</td>
<td>.847</td>
</tr>
<tr>
<td>CBD count</td>
<td>.103</td>
<td>.033</td>
<td>9.818</td>
<td>1</td>
<td>.002</td>
<td>1.109</td>
<td>1.039</td>
<td>1.182</td>
</tr>
<tr>
<td>CBD average score</td>
<td>.617</td>
<td>.170</td>
<td>13.137</td>
<td>1</td>
<td>.000</td>
<td>1.854</td>
<td>1.328</td>
<td>2.588</td>
</tr>
<tr>
<td>CBD year decimal</td>
<td>-3.627</td>
<td>.570</td>
<td>40.517</td>
<td>1</td>
<td>.000</td>
<td>.027</td>
<td>.009</td>
<td>.081</td>
</tr>
<tr>
<td>Mini-CEX count</td>
<td>.096</td>
<td>.032</td>
<td>8.736</td>
<td>1</td>
<td>.003</td>
<td>1.101</td>
<td>1.033</td>
<td>1.173</td>
</tr>
<tr>
<td>Mini-CEX year decimal</td>
<td>-1.677</td>
<td>.394</td>
<td>18.112</td>
<td>1</td>
<td>.000</td>
<td>.187</td>
<td>.086</td>
<td>.405</td>
</tr>
<tr>
<td>Total tagged events</td>
<td>.001</td>
<td>.000</td>
<td>10.505</td>
<td>1</td>
<td>.001</td>
<td>1.001</td>
<td>1.000</td>
<td>1.001</td>
</tr>
<tr>
<td>Constant</td>
<td>4.612</td>
<td>.753</td>
<td>37.547</td>
<td>1</td>
<td>.000</td>
<td>100.699</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Female is the reference category In ST2, chance of successful ARCP outcome is significantly greater for female and younger trainees, but there is no effect of Place of training or Ethnicity (Table 7.35). A greater count, completed earlier in the year for both mini-CEX and CBDs increased the odds of a satisfactory ARCP outcome, but the mini-CEX count was of marginal significance. However, the CBD and mini-CEX average scores were not significant. More tagged events and higher CSR scores were also positively associated with ARCP outcome.

---

Note that the significances in these tables are derived from the Wald statistic, which are a good guide but can be prone to error (Field, 2000: 180); the decision by SPSS to include predictors in this table depends on the change in overall model fit. CBD and mini-CEX average scores are included in this table as they significantly improves the model (p<0.05), but their Wald statistics are not significant. In the text, we err on the side of caution by assuming they are not significant.
Table 7.35: logistic regression for predictors of ST2 ARCP (all trainees, n=5060)

<table>
<thead>
<tr>
<th></th>
<th>B</th>
<th>S.E.</th>
<th>Wald</th>
<th>df</th>
<th>Sig.</th>
<th>Exp(B)</th>
<th>CI Lower</th>
<th>CI Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>-.047</td>
<td>.011</td>
<td>16.929</td>
<td>1</td>
<td>.000</td>
<td>.954</td>
<td>.933</td>
<td>.976</td>
</tr>
<tr>
<td>Gender</td>
<td>-.295</td>
<td>.110</td>
<td>7.161</td>
<td>1</td>
<td>.007</td>
<td>.744</td>
<td>.600</td>
<td>.924</td>
</tr>
<tr>
<td>CBD average score</td>
<td>.353</td>
<td>.190</td>
<td>3.441</td>
<td>1</td>
<td>.064</td>
<td>1.423</td>
<td>.980</td>
<td>2.065</td>
</tr>
<tr>
<td>CBD year decimal</td>
<td>-1.943</td>
<td>.551</td>
<td>12.445</td>
<td>1</td>
<td>.000</td>
<td>.143</td>
<td>.049</td>
<td>.422</td>
</tr>
<tr>
<td>Mini-CEX count</td>
<td>.059</td>
<td>.030</td>
<td>4.025</td>
<td>1</td>
<td>.045</td>
<td>1.061</td>
<td>1.001</td>
<td>1.125</td>
</tr>
<tr>
<td>Mini-CEX average score</td>
<td>.281</td>
<td>.159</td>
<td>3.125</td>
<td>1</td>
<td>.077</td>
<td>1.324</td>
<td>.970</td>
<td>1.808</td>
</tr>
<tr>
<td>Mini-CEX year decimal</td>
<td>-1.661</td>
<td>.387</td>
<td>18.401</td>
<td>1</td>
<td>.000</td>
<td>.190</td>
<td>.089</td>
<td>.406</td>
</tr>
<tr>
<td>CBD count</td>
<td>.097</td>
<td>.026</td>
<td>14.395</td>
<td>1</td>
<td>.000</td>
<td>1.102</td>
<td>1.048</td>
<td>1.158</td>
</tr>
<tr>
<td>CSR average score</td>
<td>.569</td>
<td>.127</td>
<td>20.025</td>
<td>1</td>
<td>.000</td>
<td>1.766</td>
<td>1.377</td>
<td>2.266</td>
</tr>
<tr>
<td>Total tagged Events</td>
<td>.001</td>
<td>.000</td>
<td>10.462</td>
<td>1</td>
<td>.001</td>
<td>1.001</td>
<td>1.000</td>
<td>1.001</td>
</tr>
<tr>
<td>Constant</td>
<td>.775</td>
<td>.909</td>
<td>.726</td>
<td>1</td>
<td>.394</td>
<td>2.170</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

For ST3, mini-CEX and CSR were not entered into the regression model, but COTs were (Table 7.36). There was a higher probability of a satisfactory ARCP outcome for UK trained, younger, female, white trainees. Satisfactory outcome was also associated with high CBD count and average score as well as higher COT scores completed earlier in the year. Finally, more tagged events increased the odds.
Table 7.36: logistic regression for predictors of ST3 ARCP (n=5949)

<table>
<thead>
<tr>
<th></th>
<th>B</th>
<th>S.E.</th>
<th>Wald</th>
<th>df</th>
<th>Sig.</th>
<th>Exp(B)</th>
<th>CI Lower</th>
<th>CI Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>-.105</td>
<td>.009</td>
<td>123.036</td>
<td>1</td>
<td>.000</td>
<td>.900</td>
<td>.884</td>
<td>.917</td>
</tr>
<tr>
<td>Gender</td>
<td>-.497</td>
<td>.088</td>
<td>31.916</td>
<td>1</td>
<td>.000</td>
<td>.609</td>
<td>.512</td>
<td>.723</td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
<td></td>
<td>44.210</td>
<td>2</td>
<td>.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>.729</td>
<td>.111</td>
<td>43.212</td>
<td>1</td>
<td>.000</td>
<td>2.074</td>
<td>1.668</td>
<td>2.577</td>
</tr>
<tr>
<td>Black/Mixed/Chinese/Other</td>
<td>.089</td>
<td>.130</td>
<td>.462</td>
<td>1</td>
<td>.497</td>
<td>1.093</td>
<td>.846</td>
<td>1.411</td>
</tr>
<tr>
<td>Place of training*</td>
<td>-.590</td>
<td>.107</td>
<td>30.260</td>
<td>1</td>
<td>.000</td>
<td>.555</td>
<td>.450</td>
<td>.684</td>
</tr>
<tr>
<td>COT Average score</td>
<td>1.083</td>
<td>.169</td>
<td>41.250</td>
<td>1</td>
<td>.000</td>
<td>2.953</td>
<td>2.122</td>
<td>4.110</td>
</tr>
<tr>
<td>COT year decimal</td>
<td>-.3128</td>
<td>.316</td>
<td>97.705</td>
<td>1</td>
<td>.000</td>
<td>.044</td>
<td>.024</td>
<td>.081</td>
</tr>
<tr>
<td>CBD count</td>
<td>.035</td>
<td>.013</td>
<td>7.477</td>
<td>1</td>
<td>.006</td>
<td>1.035</td>
<td>1.010</td>
<td>1.061</td>
</tr>
<tr>
<td>CBD Average score</td>
<td>.921</td>
<td>.189</td>
<td>23.798</td>
<td>1</td>
<td>.000</td>
<td>2.511</td>
<td>1.735</td>
<td>3.635</td>
</tr>
<tr>
<td>Total tagged Events</td>
<td>.000</td>
<td>.000</td>
<td>6.582</td>
<td>1</td>
<td>.10</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>Constant</td>
<td>1.299</td>
<td>.590</td>
<td>4.838</td>
<td>1</td>
<td>.028</td>
<td>3.665</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*UK trained is the reference category

7.4.7 AKT

For these analyses, we require a single AKT score from trainees; therefore only their first attempts at AKT are considered. The final database contained 8177 first attempts at AKT; 6985 (85%) were passes and 1192 (15%) were fails (Table 7.37). Most (56%) first attempts were undertaken by ST3 trainees, with 40% by ST2 and 2% by ST1 trainees. The percentage of passes increases slightly in later ST years with 74% passes in ST1, 83% in ST2 and 88% in ST3.

Table 7.37: AKT passes and fails by ST year: first attempts only

<table>
<thead>
<tr>
<th></th>
<th>ST1</th>
<th>ST2</th>
<th>ST3</th>
<th>ST4</th>
<th>Other*</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>%</td>
<td>n</td>
<td>%</td>
<td>n</td>
<td>n</td>
</tr>
<tr>
<td>Fail</td>
<td>50</td>
<td>26</td>
<td>554</td>
<td>17</td>
<td>564</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td></td>
<td>0</td>
<td>23</td>
</tr>
<tr>
<td>Pass</td>
<td>146</td>
<td>75</td>
<td>2710</td>
<td>83</td>
<td>4046</td>
<td>82</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td>100</td>
<td>77</td>
</tr>
<tr>
<td>Total</td>
<td>196</td>
<td>100</td>
<td>3264</td>
<td>100</td>
<td>4610</td>
<td>106</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>8177</td>
<td>100</td>
</tr>
</tbody>
</table>

*Maternity, post CCT, sabbatical, sick leave and not available.
Figure 7.3 shows the distribution of AKT ‘marks minus the pass mark’; it is reasonably close to a normal distribution although with negative skew as there is a bigger tail to the left. Therefore multiple regression was considered appropriate.

**Figure 7.3: Histogram of mark minus pass mark for all candidates**

Table 7.38 shows the variables in the final regression equation when the Count, Average score, and Year decimal for ST3 COTS and CBDs were included along with Total tagged events for first attempts at AKT in ST3. UK-trained, younger, female and White trainees tended to score higher; higher average COT and CBD scores and recording COTs earlier in the year were all associated with higher AKT results. However, tagged events, the number of COTs and CBDs recorded, and the ‘Year decimal’ for CBDs were not significant. This regression equation accounted for 23.7% of the variance in AKT scores. When only place of training, age, gender and ethnicity were included, the model accounted for 20.8% of the variance, so the other factors only account for an extra 2.9%.
Table 7.38: Multiple regression of Mark minus pass mark with ST3 variables and all trainees (n=5873)

<table>
<thead>
<tr>
<th></th>
<th>B</th>
<th>Std. Error</th>
<th>Beta</th>
<th>t</th>
<th>Sig.</th>
<th>Lower CI</th>
<th>Upper CI</th>
<th>Tolerance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>16.900</td>
<td>1.406</td>
<td>12.019</td>
<td>.000</td>
<td>14.144</td>
<td>19.657</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Place of training</td>
<td>-2.524</td>
<td>.255</td>
<td>-.149</td>
<td>-9.893</td>
<td>.000</td>
<td>-3.024</td>
<td>-2.024</td>
<td>.574</td>
</tr>
<tr>
<td>Age</td>
<td>-.264</td>
<td>.024</td>
<td>-.150</td>
<td>-11.160</td>
<td>.000</td>
<td>-.311</td>
<td>-.218</td>
<td>.718</td>
</tr>
<tr>
<td>Black/ Mixed/ Chinese/ Other</td>
<td>-3.628</td>
<td>.326</td>
<td>-.138</td>
<td>-11.139</td>
<td>.000</td>
<td>-4.266</td>
<td>-2.989</td>
<td>.848</td>
</tr>
<tr>
<td>Male</td>
<td>-.763</td>
<td>.182</td>
<td>-.049</td>
<td>-4.193</td>
<td>.000</td>
<td>-1.120</td>
<td>-.406</td>
<td>.942</td>
</tr>
<tr>
<td>CBD average score</td>
<td>2.221</td>
<td>.392</td>
<td>.089</td>
<td>5.669</td>
<td>.000</td>
<td>1.453</td>
<td>2.989</td>
<td>.530</td>
</tr>
<tr>
<td>COT year decimal</td>
<td>-4.201</td>
<td>.663</td>
<td>-.073</td>
<td>-6.333</td>
<td>.000</td>
<td>-5.501</td>
<td>-2.901</td>
<td>.978</td>
</tr>
<tr>
<td>COT average score</td>
<td>1.205</td>
<td>.362</td>
<td>.052</td>
<td>3.329</td>
<td>.001</td>
<td>.496</td>
<td>1.915</td>
<td>.532</td>
</tr>
</tbody>
</table>

To consider the predictive power of ST1, 2 and 3 variables upon AKT performance, the ‘complete data’ trainees were used (Table 7.39). As usual, block 1 was place of training, gender, age and ethnicity. Block 2 was the average score for CBD (ST1, 2 and 3), mini-CEX and CSR (ST1 and 2) and COT (ST3). Also in block 2 was the date of the AKT exam as a year decimal including the ST year (e.g. 0.5 is halfway through ST1 and 2.75 is three quarters through ST3). Block 3 variables were the binary outcomes (satisfactory or unsatisfactory) for ARCP in all three years and tagged events. The reason for putting the ARCP outcomes last is that WPBA scores may influence the ARCP outcome. UK-trained, younger, White trainees scored higher, but the effect of gender was not significant. Higher average scores for CBD (ST2 and 3) and CSR (ST1), along with success at ARCP (ST2 and 3) and completing the AKT earlier all were associated with a higher AKT outcome. This model accounted for 30.7% of the variance (n=1568). With only place of training, age, gender and ethnicity, the model accounted for 20.1% of the variance, so 10.6% extra was explained by the other variables. A further note of caution is that the ARCP in ST3 may be influenced by passing AKT: with ARCP in ST3 removed, the block 2 variables account for 6% of the variance (total 26.1% including block 1).
Table 7.39: Multiple regression of Mark minus pass mark with variables from all years and ‘complete data’ trainees (n=1568)

<table>
<thead>
<tr>
<th></th>
<th>B</th>
<th>Std. Error</th>
<th>Beta</th>
<th>t</th>
<th>Sig.</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
<th>Tolerance</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Constant)</td>
<td>-3.909</td>
<td>3.279</td>
<td>-1.192</td>
<td>0.233</td>
<td>0.233</td>
<td>-10.341</td>
<td>2.523</td>
<td></td>
</tr>
<tr>
<td>Place of training</td>
<td>-2.268</td>
<td>0.584</td>
<td>-0.118</td>
<td>-3.881</td>
<td>0.000</td>
<td>-3.414</td>
<td>-1.121</td>
<td>0.496</td>
</tr>
<tr>
<td>Asian</td>
<td>-3.390</td>
<td>0.407</td>
<td>-0.207</td>
<td>-8.330</td>
<td>0.000</td>
<td>-4.188</td>
<td>-2.592</td>
<td>0.736</td>
</tr>
<tr>
<td>Black/ Mixed/ Chinese/ Other</td>
<td>-3.434</td>
<td>0.632</td>
<td>-0.125</td>
<td>-5.438</td>
<td>0.000</td>
<td>-4.673</td>
<td>-2.196</td>
<td>0.866</td>
</tr>
<tr>
<td>Age</td>
<td>-0.143</td>
<td>0.055</td>
<td>-0.072</td>
<td>-2.598</td>
<td>0.009</td>
<td>-0.251</td>
<td>-0.035</td>
<td>0.594</td>
</tr>
<tr>
<td>CBD ST3 Average score</td>
<td>3.351</td>
<td>0.563</td>
<td>0.132</td>
<td>5.956</td>
<td>0.000</td>
<td>2.247</td>
<td>4.455</td>
<td>0.929</td>
</tr>
<tr>
<td>AKT date in HST years</td>
<td>-1.988</td>
<td>0.451</td>
<td>-0.097</td>
<td>-4.405</td>
<td>0.000</td>
<td>-2.873</td>
<td>-1.103</td>
<td>0.939</td>
</tr>
<tr>
<td>CBD ST2 Average score</td>
<td>1.531</td>
<td>0.524</td>
<td>0.067</td>
<td>2.924</td>
<td>0.004</td>
<td>0.504</td>
<td>2.558</td>
<td>0.880</td>
</tr>
<tr>
<td>CSR ST1 Average score</td>
<td>.918</td>
<td>0.426</td>
<td>0.048</td>
<td>2.155</td>
<td>0.031</td>
<td>0.082</td>
<td>1.753</td>
<td>0.916</td>
</tr>
<tr>
<td>ARCP ST3 binary outcome</td>
<td>6.648</td>
<td>0.660</td>
<td>0.237</td>
<td>10.080</td>
<td>0.000</td>
<td>5.354</td>
<td>7.942</td>
<td>0.827</td>
</tr>
<tr>
<td>ARCP ST2 binary outcome</td>
<td>2.099</td>
<td>0.695</td>
<td>0.065</td>
<td>3.020</td>
<td>0.003</td>
<td>0.736</td>
<td>3.463</td>
<td>0.975</td>
</tr>
</tbody>
</table>

7.4.8 CSA

The Clinical Skills Assessment (CSA) is summative, intended to be taken in ST3, and is “an assessment of a doctor’s ability to integrate and apply clinical, professional, communication and practical skills appropriate for general practice”.

Figure 7.4: Distribution of CSA scores (n=6265)
Almost all trainees followed this advice to take the CSA in ST3, with 98% (6182) of first attempts coming from ST3 trainees. As with AKT, only first attempts are considered: of all first attempts, 89% (5593) were passes and 11% (686) were fails.

Figure 7.4 is the histogram of first attempt CSA scores, ranging from 19 to the maximum possible 48. The distribution is reasonably close to a normal distribution although with negative skew as there is a bigger tail to the left. Therefore multiple regression was considered appropriate.

Table 7.40 shows the variables in the final regression model when the Count, Average score, and Year decimal for ST3 COTS and CBDs were included along with Total tagged events and the ‘Year decimal’ for ST3 first attempts at CSA. UK-trained, younger, female and White trainees tended to score higher; higher average COT and CBD scores and recording COTs earlier in the year were all associated with higher CSA results. However, tagged events, the number of COTs and CBDs recorded, and the ‘Year decimal’ for both CBDs and the CSA were not significant. This regression equation accounted for 40.2% of the variance in CSA scores. When the regression was performed with only Place of training, Age, Gender and Ethnicity, these variables accounted for 38.8% of the variance, so only 1.4% was accounted by the additional variables.

Table 7.40: Multiple regression of CSA with ST3 variables and all trainees (n=5734)

<table>
<thead>
<tr>
<th>B</th>
<th>Std. Error</th>
<th>Beta</th>
<th>t</th>
<th>Sig.</th>
<th>Lower CI</th>
<th>Upper CI</th>
<th>Tolerance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>47.506</td>
<td>.792</td>
<td>59.946</td>
<td>.000</td>
<td>45.952</td>
<td>49.059</td>
<td></td>
</tr>
<tr>
<td>Place of training</td>
<td>-3.340</td>
<td>.144</td>
<td>-.320</td>
<td>-23.231</td>
<td>.000</td>
<td>-3.621</td>
<td>-3.058</td>
</tr>
<tr>
<td>Male</td>
<td>-1.343</td>
<td>.101</td>
<td>-.143</td>
<td>-13.293</td>
<td>.000</td>
<td>-1.541</td>
<td>-1.145</td>
</tr>
<tr>
<td>Asian</td>
<td>-2.210</td>
<td>.122</td>
<td>-.228</td>
<td>-18.086</td>
<td>.000</td>
<td>-2.450</td>
<td>-1.970</td>
</tr>
<tr>
<td>Black/ Mixed/ Chinese/Other</td>
<td>-2.047</td>
<td>.180</td>
<td>-.129</td>
<td>-11.367</td>
<td>.000</td>
<td>-2.401</td>
<td>-1.694</td>
</tr>
<tr>
<td>Age</td>
<td>-.133</td>
<td>.014</td>
<td>-.121</td>
<td>-9.782</td>
<td>.000</td>
<td>-.160</td>
<td>-.106</td>
</tr>
<tr>
<td>COT Average score</td>
<td>1.011</td>
<td>.209</td>
<td>.069</td>
<td>4.835</td>
<td>.000</td>
<td>.601</td>
<td>1.421</td>
</tr>
<tr>
<td>CBD Average score</td>
<td>.898</td>
<td>.223</td>
<td>.058</td>
<td>4.022</td>
<td>.000</td>
<td>.460</td>
<td>1.335</td>
</tr>
<tr>
<td>COT year decimal</td>
<td>-1.303</td>
<td>.379</td>
<td>-.036</td>
<td>-3.438</td>
<td>.001</td>
<td>-2.046</td>
<td>-.560</td>
</tr>
</tbody>
</table>

As with AKT, the ‘complete data’ trainees were used to consider the impact of scores obtained throughout specialty training. The multiple regression shown in Table 7.41 indicates that UK-trained, female, White trainees did better in the CSA. These three variables accounted for 38.7% of the
variance of CSA scores. Higher average scores of COT (ST3), CBD (ST1 and 2), mini-CEX (ST2) and CSR
(ST1) along with a positive ARCP ST3 outcome accounted for a further 7.5% of the variance as the
total variance explained by the model was 46.2%. With the ARCP ST3 outcome removed, the variance
explained by the WBPAs and tagged events is just 3.5%.

Table 7.41: Multiple regression of CSA with variables from all years and ‘complete data’ trainees
(n=1521)

<table>
<thead>
<tr>
<th></th>
<th>B</th>
<th>Std. Error</th>
<th>Beta</th>
<th>t</th>
<th>Sig.</th>
<th>Lower CI</th>
<th>Upper CI</th>
<th>Tolerance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>32.531</td>
<td>1.969</td>
<td></td>
<td>16.525</td>
<td>.000</td>
<td>28.670</td>
<td>36.393</td>
<td></td>
</tr>
<tr>
<td>Place of training</td>
<td>-3.049</td>
<td>.306</td>
<td>-.270</td>
<td>-9.959</td>
<td>.000</td>
<td>-3.650</td>
<td>-2.449</td>
<td>.489</td>
</tr>
<tr>
<td>Male</td>
<td>-1.422</td>
<td>.185</td>
<td>-.154</td>
<td>-7.706</td>
<td>.000</td>
<td>-1.785</td>
<td>-1.060</td>
<td>.900</td>
</tr>
<tr>
<td>Asian</td>
<td>-1.762</td>
<td>.213</td>
<td>-.184</td>
<td>-8.264</td>
<td>.000</td>
<td>-2.181</td>
<td>-1.344</td>
<td>.727</td>
</tr>
<tr>
<td>Black/ Mixed/ Chinese/Other</td>
<td>-1.715</td>
<td>.328</td>
<td>-.106</td>
<td>-5.222</td>
<td>.000</td>
<td>-2.359</td>
<td>-1.071</td>
<td>.866</td>
</tr>
<tr>
<td>Age</td>
<td>-.053</td>
<td>.030</td>
<td>-.045</td>
<td>-1.792</td>
<td>.073</td>
<td>-.111</td>
<td>.005</td>
<td>.564</td>
</tr>
<tr>
<td>COT ST3 Average score</td>
<td>1.287</td>
<td>.306</td>
<td>.084</td>
<td>4.204</td>
<td>.000</td>
<td>.686</td>
<td>1.887</td>
<td>.904</td>
</tr>
<tr>
<td>CBD ST1 Average score</td>
<td>.705</td>
<td>.297</td>
<td>.051</td>
<td>2.370</td>
<td>.018</td>
<td>.121</td>
<td>1.288</td>
<td>.766</td>
</tr>
<tr>
<td>Mini-CEX ST2 Average score</td>
<td>.710</td>
<td>.278</td>
<td>.053</td>
<td>2.558</td>
<td>.011</td>
<td>.165</td>
<td>1.254</td>
<td>.831</td>
</tr>
<tr>
<td>CSR ST1 Average score</td>
<td>.637</td>
<td>.238</td>
<td>.057</td>
<td>2.679</td>
<td>.007</td>
<td>.171</td>
<td>1.104</td>
<td>.799</td>
</tr>
<tr>
<td>CBD ST2 Average score</td>
<td>.600</td>
<td>.287</td>
<td>.044</td>
<td>2.091</td>
<td>.037</td>
<td>.037</td>
<td>1.163</td>
<td>.793</td>
</tr>
<tr>
<td>ARCP ST3 binary outcome</td>
<td>3.428</td>
<td>.345</td>
<td>.207</td>
<td>9.924</td>
<td>.000</td>
<td>2.750</td>
<td>4.105</td>
<td>.827</td>
</tr>
<tr>
<td>Total tagged events</td>
<td>.001</td>
<td>.000</td>
<td>.045</td>
<td>2.355</td>
<td>.019</td>
<td>.000</td>
<td>.001</td>
<td>.978</td>
</tr>
</tbody>
</table>

7.5 Discussion

Before considering the extent to which these analyses have addressed the research questions, a
number of issues should be considered.

This analysis has been undertaken on data from the first three years of the ‘new curriculum’. This
means that the system is changing fast, so there is a danger that the analysis is already dated. It is
possible that the system has already improved over these three years, for example by the
introduction of a modified CSR which is believed to have greater integrity. Our view is that there is
some value in these cautions, but they should not be exaggerated. The analyses here are the currently the best available evidence and should be treated as such, until superseded by subsequent research.

The aims were to explore the ‘tagged events’ and to investigate the reliability, validity and fairness of Case-based discussions, the Clinical Evaluation Exercise (Mini-CEX), the Consultation Observation Tool and the Clinical Supervisors Report.

These aims led to the following five research questions:

RQ1 Tagged events
   a) How many tagged events have trainees put in their portfolios?
   b) What is the distribution across categories?
   c) Is there an association between the number of tagged events and performance at ARCPs?

Over 6 million events were tagged in the ePortfolio; an average of 762 was recorded for each of the ‘complete data’ trainees i.e. about 21 per month. Most of these entries (73%) were to curriculum areas, tagged by the trainee, with the competency areas (27%) tagged by their supervisors. Most common entries were for Clinical encounters (32%) and tutorials (18%). In the logistic regression models, the total number of tagged events was a significant predictor of passing ARCP first time at ST1, 2 and 3. To obtain a sense of the size of the effect, in Table 7.34 for ST1 ARCP, the more exact figure for Exp(B) for tagged events is 1.0006; this is interpreted as the change in odds for a single additional tagged event. This means that for 100 extra tagged events, the odds are predicted to increase by 1.0618 (1.0006 to the power of 100). This could increase the chance of passing the ARCP from 90% to 90.53% (www.ats.ucla.edu/stat/stata/dae/logit.htm) i.e. the effect is modest. So there is a small additional chance of succeeding at ARCP with more tagged events once the effects of the other variables have been considered (the discussion around Table 7.42 provides more information on this issue). We could speculate that a greater number of tags indicates higher motivation or organisation, which feeds into successful ARCP outcome. With AKT and CSA, tagged events were not significant.

RQ2 Demographic data
   a) What are the Place of training, Gender, Age and Ethnicity of trainees?
   b) What is the impact of these demographics on the assessments considered here?

For these analyses, trainees were categorized into those trained in the UK (79%) and outside the UK (21%) and three broad ethnic groups were used: White (53%), Asian (38%) and Black/ Mixed/ Chinese/ Other (10%). For the ‘complete data’ group, 43% Asians, 4% White and 38% Black/ Mixed/ Chinese/ Other trained outside the UK; 67% who trained outside the UK were male compared with

---

7 90% pass rate is an odds ratio of 9 (90%/10%). New odds ratio is 9*1.0618=9.5562=90.5269%/9.4731%.
43\% of UK trained graduates. Asian trainees were more likely to be male (62\%) compared with Black/Mixed/Chinese/Other (47\%) and White (39\%) trainees. UK graduates were younger (30 years on average) than those who trained outside the UK (36 years). White trainees tended to be younger (30 years) compared with Asian and Black/Mixed/Chinese/Other (both 33 years). These interactions mean that their effects are partly confounded in the regression equations. Given this confounding, more weight should be given to the combined effect of these demographics e.g. in Table 7.43, than differences between them.

These demographics were entered first into all the regression models as they are fixed attributes of trainees that can’t be changed by the training system. There were three models for passing ARCP in ST1, ST2 and ST3; in all three models, younger, female trainees had a higher chance of passing the ARCP. UK-trained, White trainees had a higher ARCP pass rate in ST3, only. With both AKT and CSA, there was an ST3 model and a model looking at predictors from ST1, 2 and 3. In all four of these models, UK-trained, younger, female, White trainees performed better. These descriptives accounted for 20 to 21\% of the variance for AKT and 39\% for CSA. These are large and important effects.

It is well known that UK-trained, younger, White, female trainees tend to perform better in postgraduate medical examinations; for example Dewhurst et al. (2007) found that White candidates had higher pass rates for the MRCP (UK) Examination. These results are not evidence of discrimination as there are many possible explanations of these differences. The percentage variance could not be calculated for the prediction of ARCP outcomes as logistic regression was used; however, it is notable that these descriptives were much more important predictors of CSA than of AKT. Wakeford (2011) reports that in January to May 2010, 83\% of UK graduates passed AKT compared with 55\% non-UK graduates; this difference was greater for CSA with 94\% passes from UK and 52\% from non-UK graduates. This difference between AKT and CSA may be worthy of further investigation as it is *prima facie* evidence that overseas, non-White, older, male trainees have greater difficulty with CSA than with AKT.

RQ3 Correlations between WPBAs (CBD, COT, mini-CEX) and CSR

a) Within an ST year i.e. Do trainees who ‘do well’ on one of these assessments also do well on the others, compared with other trainees?

b) Across ST years i.e. for the same type of assessment, do trainees who do well in one ST year also do well in the next ST year?

This analysis is looking to see if those trainees who do particularly well in one assessment, also do well in others. Given the low correlations between one year and the next for CBDs (0.20 maximum), COTs (0.24 max), mini-CEX (0.16 max) and CSR (0.11 max) this seems unlikely. With CBDs, the correlation of
0.2 means that only 4% of the variance in ST2 results is predicted by how well trainees did in ST1. So how well a trainee performs in a WPBA in one year bears almost no relation to how well they will do in other years.

Correlations between WPBAs in the same year were higher than within WPBAs in different years. Between CBDs and mini-CEX the correlations were 0.32, 0.29 and 0.29 for ST1, 2, and 3 respectively; for CBDs and COTs they were 0.35, 0.33 and 0.53. This could be because trainees perform fairly consistently in one setting, but vary from placement to placement. Also, it might be argued that there will be better judgments in ST3 due the close personal supervision by the trainer. However, there is a simpler explanation: the same assessors are undertaking the assessments. As some evidence for this, the highest correlation is between CBDs and COTs in ST3, where it is likely that the same trainer will undertake most of these assessments. Also, correlations between COTs and mini-CEX were much lower at 0.09, 0.09 and 0.19; this could be because there must be different assessors.

These results raise the concern that differences between assessors may be of comparable importance to differences between trainees’ in determining assessment scores. This is not a unique finding. For example, in an assessment of mini-CEX reliability with 128 senior registrars, greater variance was estimated to be due to assessors (21%) and the assessor: trainee interaction (36%) than between trainees (15%) (Wilkinson et al., 2008). With secondary care assessments for general practice, there are three reasons why we might expect the reliability to be even lower than is found in hospital specialties. First, the assessors may not be as aware of the GP assessment procedures and criteria as their own specialty; second, there is a concern that peers, not trained assessors, may have conducted some of the assessments; third, the assessments will be undertaken in different specialties. This third reason is perhaps most interesting and unique to GP: a trainee’s performance may be quite different in (say) paediatrics than in trauma and orthopaedics.

The mini-CEX and CBDs are intended to feed into the CSR ratings, so it is perhaps surprising that the correlations are so low: 0.21 and 0.24 between CSR and mini-CEX for ST1 and 2 respectively, and 0.32 and 0.31 between CSR and CBD for ST1 and 2.

RQ4 Prediction of ARCP outcome: How well do the variables in RQs 1, 2 and 3 predict ARCP performance?

For this analysis, ARCP was deemed satisfactory if there had been no previous unsatisfactory ARCP in that training year; 93% of ST1 trainees were satisfactory, 92% of ST2 and 86% of ST3. For all three years, younger, female trainees who logged more tagged events were more likely to pass ARCP first time; in ST3, White, UK trainees were also more likely to pass. A higher number of WPBAs, higher
average score and taking them earlier in the training year were all associated with successful ARCP outcome. For CBD, this was true in all three years, except the time in the year in ST3. For Mini-CEX, which was only considered in ST1 and 2, it was true except for the average score in ST1. For COTS (ST3 only), higher average score and completing earlier but not the total number were positively associated with ARCP outcome.

To give a sense of the impact of these variables on ARCP performance, Table 7.42 shows the estimated change in chance of passing ARCP in ST1 for a ‘reasonable difference’ in the variable. For gender, the difference is between female and male. For the other variables, the ‘reasonable difference’ is taken as a one standard deviation increase, but rounded to an integer for the count variables.

<p>| Table 7.42: Estimated change in ST1 ARCP pass rate for a ‘reasonable difference’ in independent variables (n=5900) |
|---------------------------------|-----------|-----------------|-----------------|-----------------|-----------------|-----------------|-------------|</p>
<table>
<thead>
<tr>
<th></th>
<th>Sig.</th>
<th>Exp(B)</th>
<th>Original pass rate (%)</th>
<th>Overall odds</th>
<th>Reasonable difference</th>
<th>Odds with change</th>
<th>New pass rate (%)</th>
<th>Change (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>0.000</td>
<td>0.938</td>
<td>93.00</td>
<td>13.28571</td>
<td>4.5</td>
<td>9.960879</td>
<td>90.88</td>
<td>-2.12</td>
</tr>
<tr>
<td>Gender*</td>
<td>0.001</td>
<td>0.681</td>
<td>93.00</td>
<td>13.28571</td>
<td>1</td>
<td>9.047571</td>
<td>90.05</td>
<td>-2.95</td>
</tr>
<tr>
<td>CBD count</td>
<td>0.002</td>
<td>1.109</td>
<td>93.00</td>
<td>13.28571</td>
<td>3</td>
<td>18.12089</td>
<td>94.77</td>
<td>1.77</td>
</tr>
<tr>
<td>CBD Average score</td>
<td>0.000</td>
<td>1.854</td>
<td>93.00</td>
<td>13.28571</td>
<td>0.35</td>
<td>16.4901</td>
<td>94.28</td>
<td>1.28</td>
</tr>
<tr>
<td>CBD year decimal</td>
<td>0.000</td>
<td>0.027</td>
<td>93.00</td>
<td>13.28571</td>
<td>0.09</td>
<td>9.598595</td>
<td>90.56</td>
<td>-2.44</td>
</tr>
<tr>
<td>Mini-CEX count</td>
<td>0.003</td>
<td>1.101</td>
<td>93.00</td>
<td>13.28571</td>
<td>2</td>
<td>16.10496</td>
<td>94.15</td>
<td>1.15</td>
</tr>
<tr>
<td>Mini-CEX Year decimal</td>
<td>0.000</td>
<td>0.187</td>
<td>93.00</td>
<td>13.28571</td>
<td>0.15</td>
<td>10.33145</td>
<td>91.18</td>
<td>-1.82</td>
</tr>
<tr>
<td>Total tagged Events</td>
<td>0.001</td>
<td>1.001</td>
<td>93.00</td>
<td>13.28571</td>
<td>360</td>
<td>19.03938</td>
<td>95.01</td>
<td>2.01</td>
</tr>
</tbody>
</table>

*female is the reference category

The changes in pass rate in Table 7.42 are generally one or two percent. This may seem small, but a pass rate of 90% has twice the failure rate of a pass rate of 95%. However, there are three reasons to be extremely cautious about over-interpreting these figures. First, there are imperfections in the model e.g. count data have been treated as continuous. Second, the confidence intervals are not displayed. Third, and in our opinion, most importantly, there is no evidence of causality in these associations; for example, if trainees were encouraged to undertake twice as many CBDs, we do not know if this would increase the ARCP pass rate at all.
This analysis indicates associations between passing ARCP and positive WPBAs. These associations make intuitive sense: average score is some measure of success at the WPBAs; and the count is important as those with less than the mandatory minimum ought to be referred to the ARCP panel. Those with assessments undertaken earlier in the year may be more confident, organised and/or able; also, they may benefit more from the feedback given by their supervisors. From Table 7.42, the ‘count’ and ‘year decimal’ variables seem to be more important than ‘average score’ as predictors of passing ARCP in ST1. This raises the possibility that undertaking sufficient WPBAs, earlier in the training year may be more important that the scores awarded.

As this model uses logistic regression, it is problematic to calculate the proportion of ARCP variance associated with these descriptive, WPBA and tagged events and has not been attempted.

It should be noted that educational supervisors’ reports and qualitative interpretation of learning log entries are also used to make ARCP judgments; however, we were not able to incorporate these into our analyses.

RQ5 Prediction of summative assessment: How well do the variables in RQs 1, 2, 3 and 4 predict AKT and CSA outcomes?

Table 7.43 summarises the amount of variance in summative assessment outcome explained, drawn from the analyses relating to Tables 7.38 to 7.41. Place of training, age, gender and ethnicity have a substantial effect: about 20% for AKT and 39% for CSA. Adding the other ‘predictors’ only increases the explained variance by about 2% for the variables included in the ST3 analysis but by 9% for the all years variables. A reason for this difference may be the inclusion of the ST3 ARCP result as high summative assessment may lead to a satisfactory ARCP outcome. Removing ST3 ARCP reduces this variance to 6.0% for AKT and 3.5% for CSA. It is noteworthy that demonstrations of performance in WPBAs and ARCPs have tiny associations with the AKT and CSA scores when compared with the descriptives.

<table>
<thead>
<tr>
<th>Table 7.43: Percentage of summative assessment variance explained</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Predictors and trainees included</strong></td>
</tr>
<tr>
<td>AKT</td>
</tr>
<tr>
<td>ST3 only for all trainees</td>
</tr>
<tr>
<td>All years (average score only) for ‘complete data’ trainees</td>
</tr>
<tr>
<td>CSA</td>
</tr>
<tr>
<td>ST3 only for all trainees</td>
</tr>
<tr>
<td>All years (average score only) for ‘complete data’ trainees</td>
</tr>
</tbody>
</table>

*excluding ST3 ARCP
A number of criticisms can be levelled at the analyses regarding RQ4 and 5. They assume that ARCP, CSA and AKT are suitable ‘gold standards’. This can be challenged in two ways. First, if these assessments were not sufficiently robust, improving them would have a substantial impact on these predictive validities. Secondly, it can be argued that they are measuring quite different attributes from the WPBAs. Our judgement is that this claim has some force regarding AKT as it is more knowledge focussed than the WPBAs. However, ARCPs are intended to be some form of summation of the trainee’s year, and so should relate to the WPBAs. CSAs are likely to be assessing similar characteristics as many of the WPBAs, so we would expect the association between them to be higher than between AKT and the WPBAs.

In a recent review of the validity of WPBAs, Pelgrim et al. (2011) state “the validity of the mini-CEX and the ‘clinical encounter card’ appears to be supported by strong and significant correlations with other assessment instruments”. They cite a 0.7 correlation (i.e. about 49% of the variance is explained) between mini-CEX and the ‘written examination’ subscale of the Canadian medicine exam, showing that there can be a high correlation between tests that appear to be testing very different attributes. Therefore, we do not agree that the assessments in this current study are so dissimilar that it is not possible to have much stronger evidence of predictive validity.

7.6 Recommendations, limitations and potential further analyses

7.6.1 WPBAs and CSR

At the start of this section, we agreed with van der Vleuten that reliability, validity, educational impact, stakeholder acceptability and cost are important elements of an assessment system (van der Vleuten, 1996). As WPBAs are instruments for supporting formative assessment, we consider educational impact and acceptability to be more important than reliability and validity. This concurs with the RCGP view that WPBAs are qualitative judgments that evaluate trainees’ progress over time (www.rcgp-curriculum.org.uk/nmrcgp/wpba.aspx). Now that WPBAs are widely accepted, it may be time to consider the secondary issues of reliability and validity. As the correlations for the same WPBA from one year to the next were low, WPBAs were not reliable or good at evaluating progress. However, many of the factors that undermined the efforts to chart progress have been addressed during the development of WPBA. For example, the scoring range of the supervisor’s reports has increased; the reports are more evidence based; and assessors are better trained in the process. It is likely that the reliability of assessments has improved considerably since the start of the ePortfolio and a re-analysis with later data could test this possibility.

On-going moderation should be used to continue to improve the reliability of WPBAs and CSRs. Focus groups could explore possible solutions: for example, it is possible that ‘better’ trainees tend to be assessed on more challenging cases, and so may be assessed as needing further development. For
higher reliability, some standardisation of difficulty of WPBAs would be required, although this may be difficult to achieve whilst maintaining acceptability and educational impact.

This study looked at naturally occurring evidence and has not been designed as an experiment to estimate reliability. To assess inter-rater reliability more directly, perhaps within a generalisability framework, would require codes for individual assessors associated with each assessment; however, it may be prudent first to take steps to improve reliability.

- Recommendation 1: Analyse WPBAs from later cohorts to see if reliability has been improved.
- Recommendation 2: Consider ways to increase the reliability of WPBAs so that they can be used to evaluate trainee progress.
- Recommendation 3: Undertake research to directly estimate the reliability of WPBAs.

7.6.2 ‘Prediction’ of ARCP and summative assessments

From our regression models, the outcomes of ARCP and summative assessment depend far more on place of training, age, gender and ethnicity than on WPBAs. As reliability is a prerequisite for validity, the relatively low predictive validity of the WPBAs is unsurprising given the discussion in the preceding section. These models are determined by the bulk of trainees; however, the weakest trainees may be of more interest and concern. Therefore, it may be appropriate to re-analyse focusing on the weaker trainees.

Three measures of WPBAs were considered: average score, year decimal, and count. Perhaps surprisingly, the average score did not seem to be a better predictor of ARCP and summative assessments than the other two measures. As trainees must have considerable influence on the number and timing of WPBAs, it may be worth continuing to stress their importance.

GP educators suggest that ARCP judgements depend to a large extent on information from personal development plans, the quality of the tagged events and other naturally occurring evidence. An important research question is whether these judgements are robust.

As indicated above, these analyses have included all years of the ‘new’ GP curriculum. It is likely that understanding of assessment procedures and standards developed rapidly in the first year or two. Therefore, findings may be stronger after these early years.

- Recommendation 4: Undertake a reanalysis focusing on weaker trainees.
- Recommendation 5: (Continue to) stress to trainees the importance of completing sufficient WPBAs in good time.
• Recommendation 6: Explore qualitatively the ways in which ARCP judgments are made; also investigate the robustness of ARCP judgments.

7.6.3 Data

This analysis revealed many data issues, which should be resolved if the College wishes to undertake further analyses of these kinds in the future. In addition, changes may help educational supervisors and Deanery staff. For analyses of this type, the most important improvement is to ensure clear coding of ST years. Ideally, ST years should be contiguous, so ST2 starts the day after ST1 ends. To reduce coding errors and improve quality assurance, the college may decide minimum and maximum lengths for each ST year. It would be good also to record extensions and multiple ARCPs in the same place i.e. so each trainee only has one record of start and finish dates and ARCP result(s) for each ST year.

Another area that could be improved is the accuracy of entries for WPBAs. It would be helpful if trainers could modify entries (at least for a limited period), so that mistakes could be rectified: this should improve accuracy and stop some duplicate entries. In this analysis, 1435 COTs did not have a correct code for the overall assessment: presumably most of them simply were not recorded. So, to improve the accuracy of the ePortfolio, all parts of each WBPA entry should need to be completed with legitimate codes.

• Recommendation 7: Consider requiring legitimate entries for WPBAs to be recorded on the ePortfolio.
• Recommendation 8: Devise a system so that ST years and ARCP results are unambiguously recorded.

7.6.4 Descriptives

Throughout these analyses, UK-trained, younger, female, White trainees tended to perform better at ARCP, AKT and particularly at CSA. Although differences between groups are not in themselves evidence of bias, the large effect of these demographics on CSA performance suggests it is important that the College explores these issues. We understand that considerable work has already been undertaken in this area. With the current dataset, the most obvious additional analysis would be to explore the reasons for CSA failure i.e. exploring whether different demographic groups are given different reasons for failing CSA stations. Finally, on a statistical point, the regression models in this section did not take account of unreliability of the summative assessments: future analyses may adjust for this criterion unreliability.

• Recommendation 9: Explore the reasons for the large differences in CSA performance between demographic groups.
Recommendation 10: Incorporate estimates of criterion unreliability in future analyses of predictive validity for WPBAs and ARCP.
8. DISCUSSION

This evaluation has explored the experience of training with the new curriculum for general practice during the first three years of its implementation. The introduction of the new curriculum constitutes significant transitional educational change which required adaptation of previously established roles in general practice training. The phased approach adopted in the evaluation has allowed this change to be viewed as a process rather than an event (Fullan and Park, 1981). This final report, which constitutes Phase 4 of the study, has drawn together findings from all phases of the project. These will now be discussed in relation to the research questions to allow final conclusions to be made.

8.1 Curriculum Implementation: Phase 1

During this phase of the research, three main research questions were addressed:

1. What are trainer and trainee attitudes towards the new curriculum?
2. What support and resources would trainers and trainees value?
3. What operational issues have arisen from the implementation of the new curriculum?

Each of these will now be considered in relation to the findings.

8.1.1 What are trainer and trainee attitudes towards the new curriculum?

Findings over the three years of the study demonstrate increasing acceptance of the new curriculum and growing recognition and realisation of its potential. The main difficulties associated with initial implementation related to problems with the ePortfolio and the requirement for ST3s to meet the new assessment requirements during that initial year.

The findings over the three years of the study demonstrate increased acceptance of the ePortfolio as improvements were introduced and resources were developed to support its use. Over the three years, trainers and trainees developed more confidence and competence in its use and there was increased recognition of its potential as a recording tool; its role in assessment; as a tool that facilitates teaching and learning; as a tool for aiding cohesion between the hospital based and general practice components of training and ensuring the former has a general practice focus; its role in CPD, particularly in relation to appraisal and revalidation.

However, it is clear that its delayed launch in the 2007-8 training year, initial technological problems and the lack of training and resources available to support its early use has had a significant impact on its acceptance. The hostility and anger expressed by trainers and trainees in focus groups held during the first year of the study, clearly demonstrated the stressful nature of their experience of learning
how to use this essential component of the new curriculum. One of the comments made by ST3s in a focus group during the final year of the study was “the ePortfolio has united us in a community of moaners!” This comment reflects the impact of the initial experience, where to a certain extent the ePortfolio became the label associated with a discourse of criticism and opposition. This aspect of opposition to change frequently occurs during significant educational change, particularly during the implementation of new curricula (Hargreaves, 1981; Ball, 1990; 1994). It was evident in early focus groups when counter views were presented to this dominant discourse and these participants were quite hesitant in ‘admitting’ to either “quite liking” the ePortfolio or to not having experienced any major difficulties with its use. The significant improvements made to the ePortfolio since its launch and its ongoing refinement⁹, along with the development of resources and experience to support its use, should ensure greater acceptance among more recent trainees. Views expressed by deanery leads, trainers and trainees during the final year of the study suggest this is occurring.

The difficulties experienced by ST3s and their trainers during the first year of implementation were specific to that cohort of trainees and the short period of time available for adjustments to be made to accommodate the new assessment requirements. However, findings from the interviews with deanery leads held during the final year of the evaluation, suggest that there is a remaining impact of this; that some trainers hold a continued expectation for allowances to be made when their trainees fail to meet all the assessment criteria. It is likely that this expectation was also held by some of the next cohort of ST3s. The findings from the second year of the evaluation suggested that a culture of minimum effort and focus on low challenge cases in assessment was developing among some trainees. This could be attributed to the allowances made for the first cohort of ST3s. Trainees who described adopting this approach often referred to advice given by the previous year’s ST3s to put “minimum effort” into meeting the assessment requirement. In the final year of the study, there was no evidence of this culture prevailing. Both ST3s and trainers described approaches to assessment, including methods of selecting cases for assessment, that ensured challenge and learning value were maximised. Trainers identified the value of training for contributing to their increased effectiveness as assessors.

Apart from these difficulties, the new system of assessment was generally accepted at a relatively early stage. This could be partly attributed to the experience gained from the introduction of summative assessment the previous decade. Also, its similarity to assessment in the Foundation Programme meant less adaptation was required by trainees who had completed this for their basic training and for trainers who had worked with foundation doctors. However, trainers, particularly in the early stage of implementation, were critical that assessment dominated training, and did not allow time to provide traditional tutorials. Over the three years of the study, the findings show that

⁹ See [www.rcgp.curriculum.org.uk/eportfolio/previous_release_notes-1.aspx](www.rcgp.curriculum.org.uk/eportfolio/previous_release_notes-1.aspx)
trainers use the WPBAs to provide teaching opportunities and trainees value the formative use of these assessments. There is increased acceptance among trainees that the grade ‘needs further development’, along with constructive feedback, is far more meaningful than a ‘tick box excellent’.

DOPS were the most frequently criticised WPBA throughout the period of the evaluation. Most trainees considered the skills were set too low and resented having to repeat Foundation DOPS. During the final year of the study, trainees were no longer required to repeat DOPs achieved during their Foundation training. However, the findings show that this information had not reached all of the participants at this stage. The CSA and AKT were generally viewed positively and most viewed the CSA as preferable to the previous video assessment.

Meeting the assessment requirements and the implementation issues associated with the ePortfolio perhaps resulted in trainers and trainees focusing less on the written curriculum during the early stages of implementation. The findings over the three years show that trainees were generally overwhelmed by its size and disappointed that it did not provide a ‘quick fix’ guide to essential knowledge. Most made limited use of it during their training. Similarly, many trainers did not initially embrace the core curriculum and supporting statements. However, the findings show that over the three years, there is increasing acceptance of its value and relevance to general practice training. Training for trainers and the increased resources developed to support its use were shown to have increased the use of the written curriculum to support teaching and learning.

The approach taken by the evaluation to explore stakeholder views and experience of each of these three components separately has probably, to a certain extent, obscured their interrelatedness. However, findings from the final year of the study show increased recognition, particularly among programme directors and trainers of the potential of the new training curriculum, which perhaps has not been fully realised during this implementation period. The findings suggest that current trainees will benefit from the increased experience and resources that have been developed during the implementation period to support the delivery of the new training curriculum.

8.1.2 What support and resources would trainers and trainees value?

The evaluation explored what further support and resources would be valued by trainers and trainees. During the first year of the study, most of the suggestions made were in relation to the provision of additional support with the ePortfolio, including the provision of better guidance and training in its use. Trainees indicated they were generally very satisfied with the support received from their trainers.
During the second year of the study, findings from the national survey showed that trainees were generally satisfied with the support received from their programme directors, GP trainers and hospital supervisors, whilst only 38% felt satisfied with the support received from their deaneries. Trainees valued most highly the support received from their GP trainers and other GP STs. Trainers valued most the support received from trainers’ groups and programme directors. Further support that would be welcomed by trainers and trainees mainly related to improved systems of communication, including the provision of clear guidance on procedures relating to the curriculum; dates for completion of assessments; and a better flow of information from the College.

During the final year of the study, there was a general consensus that guidance and training for the ePortfolio had improved. Although dissatisfaction was expressed over poor communication and lack of guidance during the early stages of implementation, most considered that the deaneries were now better informed and more able to provide effective support. The main area highlighted where further support was required was in relation to assessments carried out in the hospital based component of training. Although the development of resources to support hospital training was ongoing, as was the training in assessment provided to hospital clinical supervisors, a clear need was identified for this training to be extended to other hospital based assessors.

8.1.3 What operational issues have arisen from the implementation of the new curriculum?

Findings from the evaluation over the three years of the study show that the main operational issues to arise from implementation of the new curriculum related to the introduction of the ePortfolio. Its late introduction, the technical problems that had not been addressed and the lack of resources and experience to support its use, all contributed to a hostile reaction from trainers and trainees, which tended to dominate the early phase of implementation. This negative and hostile response had a serious impact on the general acceptance of the new training curriculum for general practice. The study also documents the huge effort made to remedy these problems and the significant improvements made to the ePortfolio over the initial three years.

Findings from the final year of the study suggest the ePortfolio is beginning to bed in to the culture of general practice training with increasing acceptance evident among trainers and trainees. Whilst this is a significant achievement, it might have been preferable to have delayed its introduction and to have developed and refined the ePortfolio in a limited number of sites before extending its use. This type of approach to implementation fits with Schon’s (1971) proliferation-of-centres model, which he considered most effective for the diffusion of educational change. In this model, primary centres focus primarily on training trainers and providing support, monitoring and management of the change process, whilst secondary centres have responsibility for the diffusion of the innovation. However, adopting this approach might have entailed the delay of the ePortfolio’s introduction until the 2008-9
training year. This would have resulted in delaying the implementation of the new curriculum for a similar period as the ePortfolio is an essential constituent. This perhaps makes this option less feasible and the decision to go ahead with its premature launch more understandable.

8.2 How the Curriculum is Experienced – Phase 2

In Phase 2 of the study, the evaluation explored how the curriculum was experienced by trainers and trainees. The following main research questions were addressed during this phase:

1. How has the new GP training curriculum impacted on trainees’ learning experience?
2. What do trainees perceive as the strengths and weaknesses of the new curriculum?
3. How has the new GP training curriculum impacted on teaching practice and the role of the trainer?
4. How are workplace based assessment tools used by trainers and trainees in practice?
5. How has the new GP training curriculum impacted on teaching practice in the hospital-based setting?

These will now be considered in relation to the findings.

8.2.1 How has the new GP training curriculum impacted on trainees’ learning experience?

The findings from the final year of the study describe how the written curriculum, the ePortfolio and the new assessment system have impacted on the learning experience of trainees. Most trainees described making limited use of the written curriculum and trainers and deanery leads also considered it was underutilised by trainees. However, some trainees have found it helpful to use the curriculum statements to support their learning in hospital posts, and the work currently being done by deaneries in mapping these to job descriptions should ensure its wider impact on trainee’s learning during their hospital posts. Trainees also report using the written curriculum to guide AKT revision, but often other available resources are considered more helpful. The increased understanding and acceptance of the written curriculum among trainers, reflected in the findings, should ensure that they are better able to support and guide their trainees in its use. The development of resources, particularly by programme directors, to support the use of the written curriculum, and their focus on it with trainees during study days, should also contribute to an increased impact of the written curriculum on trainees’ learning experience. The findings suggest that ST1s during 2009-10 training year had access to this support that previously had not been widely available. Therefore, these and future trainees, will possibly be better equipped to use the written curriculum effectively to support their learning and development as GPs.

Trainees valued the ePortfolio for the record and evidence it provided of their learning. Whilst completing the learning log was considered time consuming, many trainees considered this had contributed to their learning. Mixed views were held on whether it aided reflection or whether it
entailed documenting reflection that had previously occurred. Feedback on entries made by their trainers was valued and considered to contribute to learning. Log entries were also shown to alert trainers to weaknesses in trainees that could be addressed. There was also evidence of its use by educational supervisors to support trainees’ learning during their hospital posts. Trainees valued the autonomy offered by the ePortfolio in allowing them to work independently on the learning log. Trainers tended to support this view and considered it helped to develop the trainees into mature self-directed learners.

Many trainers were critical that the new curriculum was assessment driven and that trainees were motivated to gain the required minimum number of assessments. However, the new competence based curriculum provided clear objectives from the start of training, and this was considered by some trainers and programme directors to provide increased motivation for trainees to learn during the hospital component of training. WPBAs, particularly in the general practice setting, were viewed by trainees and trainers as providing valuable opportunities for teaching and learning. Trainees showed increased confidence over the period of the evaluation in using assessments to provide opportunities for teaching and also in maximising the challenge and learning potential of these. Trainers considered the new assessment system was helpful in identifying gaps in learning and weaknesses in trainees allowing these to be addressed at an earlier stage. MSF and PSQ were highlighted as particularly useful in highlighting weaknesses. COTs and the use of video recorded consultations were viewed as helpful in developing consultation skills.

The findings demonstrate that, taken together, the three components of the new curriculum, are considered to have had a significant impact on enhancing the learning experience of trainees. The provision of clear aims and objectives ensure that learning relevant to general practice occurs throughout the three year programme. It is considered to be more effective in ensuring there are no gaps in learning and in identifying weaknesses in trainees that can be addressed. The new training curriculum is also seen to empower trainees as self-directed learners.

8.2.2 What do trainees perceive as the strengths and weaknesses of the curriculum?

For most trainees the strengths of the curriculum relate to how it was experienced in the general practice setting. The earlier exposure to general practice that most trainees experienced in the additional post during their ST1 or ST2 year was valued for ensuring they gained insight and experience that enabled them to focus on gaining experience and competence relevant to general practice in their subsequent hospital posts. Their experience of assessment “done properly” by GP trainers with time set aside provided valuable learning and often made them more critical of how assessments were conducted in the hospital setting. The opportunity to work with their own patient case load was recognised as providing valuable opportunities to develop skills in decision making, often not available in the context of the delivery of secondary care, because of the changes associated
with the introduction of Foundation training and the implementation of the European Working Time Directive (EWTD). There was general consensus among trainees that most learning occurred in the general practice setting.

Other than problems associated with the ePortfoli o, a major weakness of the curriculum for many trainees was how it was experienced during their hospital placements. Trainees often considered that there was little emphasis on learning and that posts were often primarily focused on service provision. They were generally dissatisfied with how most assessments were conducted and often reduced to rushed ‘tick box exercises’. Many trainees were dissatisfied when they were allocated posts in specialties previously experienced during foundation training and considered there should be a closer matching of placements to their learning needs. Whilst a majority of trainees welcomed an overall extension of the period of general practice training, most were emphatic that this should not be based in the hospital setting where they considered valuable opportunities to learn would be sacrificed as the main focus of these posts would be on service provision.

It is disappointing that these GP STs’ views on the hospital component of training mirror those of trainees trained under the previous system of general practice training. However, deanery leads, programme directors and trainers, whilst critical of some aspects of hospital based training, often point out how the new curriculum has ensured that significant improvements have occurred. These include the competence based curriculum now provides clear learning objectives for the whole period of general practice training. The role of the educational supervisor, who can maintain contact with trainees via the ePortfolio, can ensure that opportunities for learning relevant to general practice are increased. The findings also demonstrate that a main focus of deaneries is now on developing resources and training that will ensure the GP curriculum is delivered effectively to GP STs in these posts.

8.2.3 How has the new GP training curriculum impacted on teaching practice and the role of the trainer?

Over the period of the evaluation, trainers have discussed how the new curriculum has changed their role. Many consider they have a greater responsibility for assessment now than they had previously and, in particular, are critical of how assessment has come to dominate training. Of particular concern to them was the sacrifice of the traditional one-to-one tutorials, often viewed as the essence of good general practice training. There was no longer sufficient time for these because of the new assessment requirement. During the final year of the study, trainers in one of the focus groups discussed many positive aspects of the new training curriculum, but described feeling “bereft” at the loss of their traditional training role which they associated with tutorials. This coincides with Marris’ (1975) theory which considers the psychological impact of significant change on the individual and
describes this as a similar process to the experience of bereavement. The impact of this “loss” and the ongoing process of making the necessary adaptations to their roles prescribed by the new curriculum were still very much evident at this time.

The experience of new trainers is quite different. They are generally positive about the new curriculum and describe finding its structure helpful as they learn their new role.

Most trainers also considered that the new curriculum had significantly increased their workload. As well as having responsibility for their ST3s, many educational supervisors also have responsibility for trainees during their ST1 and ST2 training years, who are often referred to as “remote” trainees, because contact is most usual via the ePortfolio. Commenting on log entries and completing reviews were highlighted as time consuming with many trainers having to complete these in their own time.

The study also demonstrates the value of trainers’ groups in providing opportunities for collaborative learning and sharing experience. These have been helpful in enabling trainers to develop best practice and assimilate the meaning of the new training curriculum.

8.2.4 How are WPBAs used by trainers and trainees in practice?

The findings from the evaluation show that trainers and trainees value the formative use of assessment. During the final year of the study trainers demonstrated increased confidence in using the WPBAs to provide teaching opportunities. However, some were able to combine summative and formative approaches whilst others considered that the structure for completing formal CBDs did not allow opportunity for teaching until the formal aspect had been completed. Some trainers made a distinction between using WPBAs formally and informally, particularly CBDs. Formal CBDs were recorded on the ePortfolio whilst informal CBDs were not recorded, were viewed as purely formative and were used to facilitate teaching and learning.

Trainers also considered that WPBAs were useful in identifying or confirming trainees’ weaknesses. Examples were given of how weaknesses highlighted by the PSQ and MSF provided trainers with opportunities for these to be addressed.

Throughout the three years of the study trainees made a distinction between the value of WPBAs carried out in the general practice and hospital settings and preferred the experience in general practice where time with their trainer was set aside specifically for this purpose and the assessments provided valuable learning. There was no evidence in the findings from the final year that trainees were selecting low challenge case for assessment. The descriptions provided by trainers and trainees suggest that it was necessary to select challenging cases to ensure all the relevant competences were achieved.
8.2.5 How has the new GP curriculum impacted on teaching in the hospital based setting?

Over the period of the evaluation, there was little evidence that the new GP curriculum had made significant impact on the teaching that occurred in the hospital based setting. However, during the final year of the study, examples were given of the experience of good practice in relation to assessments, and this demonstrated that some supervisors were beginning to use the curriculum statements to identify skills and competences within their specialty that were relevant to general practice training. This suggests that the relatively recent work being undertaken in deaneries in mapping job descriptions to the statements is having some impact. This work should ensure that both clinical supervisors and trainees develop an improved understanding of how the written curriculum and assessment can be used to provide relevant learning in hospital posts.

8.3 Fitness for Purpose: Phase 3

8.3.1 To what extent is the training curriculum ‘fit for purpose’ in preparing GP specialty registrars for independent practice?

The findings from the final year of the evaluation show that most deanery leads, programme directors, trainers and trainers were generally positive about the effectiveness of the new training curriculum in preparing trainees well for clinical practice. Trainees also recognised the need for lifelong learning and development throughout their careers in general practice, which demonstrates their training had been effective in enabling them to develop this awareness.

Several aspects of the new curriculum were considered to have made a significant impact in better preparing trainees for independent practice. These include the additional period spent in general practice which prepares trainees better for their ST3 year; the ePortfolio better prepares them for independent learning; the written curriculum and assessments ensure broader curriculum coverage than in the past; and the new assessment system is more effective in identifying weaknesses in trainees. However, there was also consensus that the training curriculum does not sufficiently prepare trainees for the business side of general practice. Other gaps in learning include skills in clinical governance; academic and research skills; experience in continuity of care and living with uncertainty; and gaps in experience in key specialties.

There was a general consensus that there was a need to extend the overall period of general practice training to ensure consolidation of learning and opportunity for gaps in learning to be addressed.
8.3.2 What is the impact of the new curriculum on patient safety?

The evaluation was not able to address this question as the first cohort to train entirely with the new curriculum completed their training during the 2009-10 training year, which was the final year of the study. Trainers and deanery leads generally considered the new curriculum had prepared the trainees to be safe practitioners. This view was also reflected in the focus groups and interviews with trainees, where recognition of the need for continuous learning and awareness of when it was necessary to seek advice from more experienced colleagues were identified as essential features of safe practice.
9. CONCLUSION AND RECOMMENDATIONS

This evaluation has explored stakeholder views and experience of the implementation of the new training curriculum for general practice; how the curriculum is experienced; and the curriculum’s fitness for purpose. It has also included an analysis of the national assessment data that has been collated between the introduction of the ePortfolio in 2007 and the completion of the 2009-10 training year in July 2010. Final conclusions from each of these phases of the study will now be drawn and recommendations made for the continuing development of training curriculum for general practice.

9.1 Experience of implementation

The initial stages of implementation of the new training curriculum for general practice have now been completed. This has been a continuous process throughout the three years of the evaluation. The initial disruptive impact of this transitional educational change has now subsided and there is clear evidence that the new curriculum is beginning to bed in and become assimilated into the everyday practise of general practice training. However, it is clear that the process is still continuing. The scale of change involved suggests that it will be necessary for continued refinement, the continued development of resources to support its successful implementation and the continued provision of support and training over the next few years. In particular, focus needs to continue on ensuring the curriculum is delivered effectively during the hospital based component of training.

9.1.1 Recommendations

- The process of implementation should be viewed as ongoing and support and training is necessary to ensure the intended meaning of the new training curriculum is assimilated into practice both within the general practice and hospital settings.

9.2 Experience of the curriculum

The study has documented the phenomenology of change from the initial stage of implementation when trainers and trainees were engaged in constructing and assimilating the meaning of the changed curriculum and learning how to adapt to the new requirements. It has highlighted some of the tensions that developed between the new and previous systems of training; between confidence in established practices and uncertainty with new expectations. Trainers often demonstrated Schon’s (1971) concept of ‘dynamic conservatism’ when faced with these challenges to their established roles.

Learning the meaning of the written curriculum and translating how it can be used to support general practice training has been a gradual process. It is clear that the training and resources that have been developed to support its use have been helpful in facilitating this. Trainers in particular seem to have
gained a better understanding of its potential. It is likely that trainees who commenced their training during the 2009-10 training year will benefit from the improved resources, both in relation to the use of the curriculum during their hospital posts and also will develop increased awareness of how it underpins the whole training curriculum.

Initial tensions and hostility towards the ePortfolio seem to be easing and its potential as a tool to support teaching and learning is becoming increasingly recognised along with its effectiveness at providing a record of learning. Its potential in assessment to demonstrate progress and to allow weaknesses to be identified is also more evident. Its increased acceptance is evidenced also by recognition that its potential use extends beyond training to facilitate continuing professional development and revalidation for all GPs. However, continuing development and refinement is necessary to ensure its fitness for purpose.

There is continuing assimilation of the new assessment system into the everyday practice of general practice training within the context of general practice. Trainers demonstrate increased confidence in using WPBAs effectively to provide teaching and learning opportunities and to alert them to weaknesses in trainees. For trainees competence based assessment has become the norm and there is increased recognition of the value of the formative approach to assessment used in general practice. Assessment carried out by hospital based assessors continues to be considered less valuable and effective, particularly when carried out by staff other than the clinical supervisor. Concerns were raised over the reliability of the clinical supervisor’s review because of a reluctance to formally record negative comments. The AKT and CSA were generally viewed positively. Some concerns were raised over the reliability and validity of WPBAs and their effectiveness at providing evidence of progress. However, the learning log and its link to the PDP was valued for the documentary evidence it provides of progress and development.

Essential skills considered not to be sufficiently assessed include clinical governance, clinical leadership, critical reading and the business skills associated with running a practice. The areas of competence considered the most difficult to achieve were ‘community orientation’ and ‘fitness to practice’.

9.2.1 Recommendations

- Work on making the written curriculum more accessible for all GPs and those involved in general practice training should continue. This includes work currently being undertaken by the Curriculum Development Group and the Curriculum Guardians in refining and standardising the curriculum statements; and the work by deaneries in developing resources and training to support its effective use.
• Clearer guidance needs to be provided to trainees and educational supervisors on how log entries should be linked to curriculum headings.
• Further development is required to make the ePortfolio easier to navigate.
• Problems associated with user error with the ePortfolio need to be addressed.
• Training in assessment for trainers needs to continue along with workshops where best practice can be shared.
• Training in assessment for hospital based assessors needs to continue and should be extended to include all those who have involvement in assessment of GP STs.

9.3 Fitness for purpose

The current balance in training that offers most trainees 18 months’ experience in hospital training and 18 months based in general practice was viewed positively. The additional earlier exposure to general practice this provided was considered beneficial to trainees’ development as GPs. Concerns were raised about the quality of some hospital posts and gaps in trainees’ experience in key specialties.

The new training curriculum was generally considered to prepare trainees well for clinical practice. It was considered to be more effective than previous training in facilitating trainees’ development as independent learners and ensuring broader curriculum coverage. However, there was a consensus among stakeholders that the training did not fully prepare trainees for independent practice and this could only be achieved by extending the period of training to allow more time for consolidation of learning and for gaps in learning to be addressed. In particular, it was felt that trainees were not sufficiently prepared for the business side of general practice. Concerns were also raised that the new training curriculum did not prepare trainees in English deaneries for their proposed role in commissioning services. Whilst there was agreement that this additional period should be based in general practice, views differed on whether four or five years of training were required.

9.3.1 Recommendations

• Work needs to continue to ensure greater consistency in the quality of hospital posts.
• Training for educational supervisors needs to focus on developing their effectiveness in providing effective support to their ST1s and ST2s that ensures relevant general practice focus is achieved during their hospital based training.
• Ways need to be explored of ensuring training includes coverage of skills in running a business, skills in clinical governance, academic and research skills; and also that experience is gained in continuity of care and living with uncertainty. English deaneries also need to consider how training can prepare future GPs for their proposed role in the commissioning of
services. However, it is possible that these developments might only be achievable if the overall of period of general practice training is extended.

9.4 An Analysis of National Assessment Data

9.4.1 WPBAs and CSR

In Chapter 7, we agreed with van der Vleuten that reliability, validity, educational impact, stakeholder acceptability and cost are important elements of an assessment system (van der Vleuten, 1996). As WPBAs are instruments for supporting formative assessment, we consider educational impact and acceptability to be more important than reliability and validity. This concurs with the RCGP view that WPBAs are qualitative judgments that evaluate trainees’ progress over time (www.rcgp-curriculum.org.uk/nmr/cgp/wpba.aspx). Now that WPBAs are widely accepted, it may be time to consider the secondary issues of reliability and validity. As the correlations for the same WPBA from one year to the next were low, WPBAs were not reliable or good at evaluating progress. However, it is likely that the reliability of assessments has improved considerably since the start of the ePortfolio, in part because the scoring range of CSR has increased and assessors are both better trained and more familiar with the process. A re-analysis with later data could test this possibility.

On-going moderation should be used to continue to improve the reliability of WPBAs and CSRs. Focus groups could explore possible solutions: for example, it is possible that ‘better’ trainees tend to be assessed on more challenging cases, and so may be assessed as needing further development. For higher reliability, some standardisation of difficulty of WPBAs would be required, although this may be difficult to achieve whilst maintaining acceptability and educational impact. However, the increased consistency in the approach used by GP trainers and trainees in ensuring cases selected for assessment are sufficiently challenging, reported in Chapter 5, could lead to less variability in difficulty.

This study looked at naturally occurring evidence and has not been designed as an experiment to estimate reliability. To assess inter-rater reliability more directly, perhaps within a generalisability framework, would require codes for individual assessors associated with each assessment; however, it may be prudent first to take steps to improve reliability.

Recommendations

- Analyse WPBAs from later cohorts to see if reliability has been improved.
- Consider ways to increase the reliability of WPBAs so that they can be used to evaluate trainee progress.
- Undertake research to directly estimate the reliability of WPBAs.
9.4.2 ‘Prediction’ of ARCP and summative assessments

From our regression models, the outcomes of ARCP and summative assessment depend far more on place of training, age, gender and ethnicity than on WBPAs. As reliability is a prerequisite for validity, the relatively low predictive validity of the WBPAs is unsurprising given the discussion in Chapter 7. These models are determined by the bulk of trainees; however, the weakest trainees may be of more interest and concern. Therefore, it may be appropriate to re-analyse focussing on the weaker trainees.

Three measures of WBPAs were considered: average score, year decimal, and count. Perhaps surprisingly, the average score did not seem to be a better predictor of ARCP and summative assessments than the other two measures. As trainees must have considerable influence on the number and timing of WBPAs, it may be worth continuing to stress their importance.

GP educators suggest that ARCP judgements depend to a large extent on information from personal development plans, the quality of the tagged events and other naturally occurring evidence. An important research question is whether these judgements are robust. As indicated above, these analyses have included all years of the ‘new’ GP curriculum. It is likely that understanding of assessment procedures and standards developed rapidly in the first year or two. Therefore, findings may be stronger after these early years.

Recommendations

- Undertake a reanalysis focussing on weaker trainees.
- (Continue to) stress to trainees the importance of completing sufficient WBPAs in good time.
- Explore qualitatively the ways in which ARCP judgments are made; also investigate the robustness of ARCP judgments.

9.4.3 Data

This analysis revealed many data issues, which should be resolved if the College wishes to undertake further analyses of these kinds in the future. In addition, changes may help educational supervisors and Deanery staff. For analyses of this type, the most important improvement is to ensure clear coding of ST years. Ideally, ST years should be contiguous, so ST2 starts the day after ST1 ends. To reduce coding errors and improve quality assurance, the college may decide minimum and maximum lengths for each ST year. It would be good also to record extensions and multiple ARCPs in the same place i.e. so each trainee only has one record of start and finish dates and ARCP result(s) for each ST year.

Another area that could be improved is the accuracy of entries for WBPAs. It would be helpful if trainers could modify entries (at least for a limited period), so that mistakes could be rectified.
should improve accuracy and stop some duplicate entries. In this analysis, 1435 COTs did not have a
correct code for the overall assessment: presumably most of them simply were not recorded. So, to
improve the accuracy of the ePortfolio, all parts of each WBPA entry should need to be completed
with legitimate codes.

**Recommendations**

- Consider requiring legitimate entries for WPBAs to be recorded on the ePortfolio.
- Devise a system so that ST years and ARCP results are unambiguously recorded.

**9.4.4 Descriptives**

Throughout these analyses, UK-trained, younger, female, White trainees tended to perform better at
ARCP, AKT and particularly at CSA. Although differences between groups are not in themselves
evidence of bias, the large effect of these demographics on CSA performance suggests it is important
that the College explores these issues. We understand that considerable work has already been
undertaken in this area. With the current dataset, the most obvious additional analysis would be to
explore the reasons for CSA failure i.e. exploring whether different demographic groups are given
different reasons for failing CSA stations. Finally, on a statistical point, the regression models in this
section did not take account of unreliability of the summative assessments: future analyses may
adjust for this criterion unreliability.

**Recommendations**

- Explore the reasons for the large differences in CSA performance between demographic
groups.
- Incorporate estimates of criterion unreliability in future analyses of predictive validity for
WPBAs and ARCP.
REFERENCES


APPENDIX 1: Analysis of Case Study Eportfolios

Characteristics

Full access to their ePortfolio data was granted by nine case study trainees. All nine began their GP training in 2007, and data from their ePortfolios was obtained for the training period to August 2010. Participating trainees were from three deaneries: North Western; Wales; and Kent, Surrey and Sussex (KSS). Eight trainees completed their GP training within the three year study period, and one completed the first three years of a four-year academic study programme.

Learning log

The ePortfolio contains a learning log, in which trainees are encouraged to record examples of formal and experiential learning activity. Log entries are linked to particular curriculum statements, and the number of linked log entries for each area of the curriculum is shown in the table below. Low numbers of entries (less than 15 for each trainee) are shown for: genetics in primary care; men’s health; care of people with learning disabilities; and eye problems.

Table A1.1: Number of log entries for curriculum statements

<table>
<thead>
<tr>
<th>Curriculum statement</th>
<th>Number of linked log entries</th>
</tr>
</thead>
<tbody>
<tr>
<td>ID 1</td>
<td>ID 2</td>
</tr>
<tr>
<td>GP consultation</td>
<td>79</td>
</tr>
<tr>
<td>Clinical Governance</td>
<td>32</td>
</tr>
<tr>
<td>Patient safety</td>
<td>57</td>
</tr>
<tr>
<td>Ethics and values based medicine</td>
<td>37</td>
</tr>
<tr>
<td>Promoting equality and valuing diversity</td>
<td>16</td>
</tr>
<tr>
<td>Evidence based practice</td>
<td>83</td>
</tr>
<tr>
<td>Research and academic activity</td>
<td>15</td>
</tr>
<tr>
<td>Teaching, mentoring and clinical supervision</td>
<td>14</td>
</tr>
<tr>
<td>Management in primary care</td>
<td>46</td>
</tr>
<tr>
<td>Information management and technology</td>
<td>26</td>
</tr>
<tr>
<td>Healthy people: promoting health and preventing disease</td>
<td>32</td>
</tr>
<tr>
<td>Genetics in primary care</td>
<td>13</td>
</tr>
<tr>
<td>Care of acutely ill people</td>
<td>56</td>
</tr>
<tr>
<td>Care of children and young people</td>
<td>41</td>
</tr>
</tbody>
</table>
Table A1.2: Number and type of PDP entries

<table>
<thead>
<tr>
<th>ID</th>
<th>ID 2</th>
<th>ID 3</th>
<th>ID 4</th>
<th>ID 5</th>
<th>ID 6</th>
<th>ID 7</th>
<th>ID 8</th>
<th>ID 9</th>
<th>Total</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tutorial</td>
<td>40</td>
<td>27</td>
<td>174</td>
<td>31</td>
<td>31</td>
<td>95</td>
<td>91</td>
<td>50</td>
<td>24</td>
<td>563</td>
</tr>
<tr>
<td>E-learning</td>
<td>67</td>
<td>1</td>
<td>31</td>
<td>19</td>
<td>13</td>
<td>8</td>
<td>31</td>
<td>11</td>
<td>0</td>
<td>181</td>
</tr>
<tr>
<td>Lecture/seminar</td>
<td>52</td>
<td>12</td>
<td>3</td>
<td>40</td>
<td>23</td>
<td>16</td>
<td>40</td>
<td>65</td>
<td>1</td>
<td>252</td>
</tr>
<tr>
<td>Reading</td>
<td>12</td>
<td>5</td>
<td>187</td>
<td>64</td>
<td>28</td>
<td>70</td>
<td>39</td>
<td>17</td>
<td>3</td>
<td>425</td>
</tr>
<tr>
<td>Audit/project</td>
<td>4</td>
<td>1</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>10</td>
<td>2</td>
<td>6</td>
<td>34</td>
</tr>
<tr>
<td>Clinical encounter</td>
<td>71</td>
<td>17</td>
<td>104</td>
<td>72</td>
<td>14</td>
<td>109</td>
<td>168</td>
<td>39</td>
<td>71</td>
<td>665</td>
</tr>
<tr>
<td>Course / certificate</td>
<td>27</td>
<td>5</td>
<td>8</td>
<td>23</td>
<td>11</td>
<td>12</td>
<td>11</td>
<td>6</td>
<td>119</td>
<td>1</td>
</tr>
<tr>
<td>Sig. event analysis</td>
<td>4</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>7</td>
<td>2</td>
<td>1</td>
<td>27</td>
</tr>
<tr>
<td>Professional conversation</td>
<td>39</td>
<td>5</td>
<td>37</td>
<td>24</td>
<td>27</td>
<td>8</td>
<td>21</td>
<td>9</td>
<td>5</td>
<td>175</td>
</tr>
<tr>
<td>Out of hours</td>
<td>12</td>
<td>10</td>
<td>12</td>
<td>0</td>
<td>15</td>
<td>25</td>
<td>14</td>
<td>10</td>
<td>11</td>
<td>109</td>
</tr>
<tr>
<td>TOTAL</td>
<td>328</td>
<td>85</td>
<td>563</td>
<td>280</td>
<td>168</td>
<td>349</td>
<td>433</td>
<td>216</td>
<td>128</td>
<td>2550</td>
</tr>
</tbody>
</table>

Personal Development Plan (PDP)

The ePortfolio contains a Personal Development Plan (PDP), linked to the learning log, in which trainees record details of the learning experience and are encouraged to reflect on their learning, identify further learning needs and make a plan to address these.
As shown in the table above, there were a total of 2550 PDP entries in the nine case study ePortfolios. The total number of PDP entries varied widely between individuals, from 85 entries to 563. Over a quarter of entries (26%) were identified as being about a clinical encounter. Over a fifth (22%) were from a tutorial, 17% were linked to reading and 10% to a lecture or seminar. Lower percentages (less than 10%) were recorded for PDP entries linked to other types of learning: e-learning; audit or project; significant event analysis; course or certificate; professional conversation; or out of hours session.

Of the 2550 PDP entries, 2049 (80%) were shared with the educational supervisor or GP trainer and 501 (20%) were not shared. Characteristics of shared entries are shown in the tables below. Eight of the nine case study trainees had written entries that were not shared with their educational supervisor, although one trainee was responsible for 350 (70%) of the 501 unshared entries, indicating that this trainee had used the PDP as a private reflective tool more frequently than other trainees had done. All ten types of learning activity were represented in the 501 unshared entries, although there were larger numbers of unshared entries linked to reading (163), tutorials (157) and clinical encounters (97) than to other types.
Table A1.3: Characteristics of shared PDP entries / PDP entries with comments

<table>
<thead>
<tr>
<th>Type of learning activity</th>
<th>No. entries not shared</th>
<th>No. entries receiving comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tutorial</td>
<td>157</td>
<td>119</td>
</tr>
<tr>
<td>E-learning</td>
<td>22</td>
<td>34</td>
</tr>
<tr>
<td>Lecture/seminar</td>
<td>16</td>
<td>91</td>
</tr>
<tr>
<td>Reading</td>
<td>163</td>
<td>68</td>
</tr>
<tr>
<td>Audit/project</td>
<td>4</td>
<td>15</td>
</tr>
<tr>
<td>Clinical encounter</td>
<td>97</td>
<td>220</td>
</tr>
<tr>
<td>Course / certificate</td>
<td>10</td>
<td>29</td>
</tr>
<tr>
<td>Sig. event analysis</td>
<td>3</td>
<td>12</td>
</tr>
<tr>
<td>Professional conversation</td>
<td>27</td>
<td>60</td>
</tr>
<tr>
<td>Out of hours</td>
<td>2</td>
<td>40</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td><strong>501</strong></td>
<td><strong>688</strong></td>
</tr>
</tbody>
</table>

Table A1.4: Shared PDP entries / PDP entries with comments by trainee

<table>
<thead>
<tr>
<th>Participant ID</th>
<th>No. entries not shared (%)</th>
<th>No. entries receiving comments (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>82</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>18</td>
</tr>
<tr>
<td>3</td>
<td>350</td>
<td>79</td>
</tr>
<tr>
<td>4</td>
<td>59</td>
<td>48</td>
</tr>
<tr>
<td>5</td>
<td>13</td>
<td>60</td>
</tr>
<tr>
<td>6</td>
<td>32</td>
<td>47</td>
</tr>
<tr>
<td>7</td>
<td>23</td>
<td>197</td>
</tr>
<tr>
<td>8</td>
<td>22</td>
<td>114</td>
</tr>
<tr>
<td>9</td>
<td>0</td>
<td>43</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td><strong>501</strong></td>
<td><strong>688</strong></td>
</tr>
</tbody>
</table>

Of the 2049 shared PDP entries, written comments from the educational supervisor or GP trainer were given for 688 entries (34%). Characteristics of entries that received comments are shown in the tables above. All nine case study trainees received some written comments to PDP entries, although the number of entries commented on varied widely, from 18 entries for one trainee, to 197 for another. Entries linked to all types of learning activity received comments. The types of learning most commonly commented on were: clinical encounters (220 entries); tutorials (119 entries); and lectures or seminars (91 entries).

Comments provided educational supervisors and GP trainers with the opportunity to praise good practice, highlight further areas for consideration, direct them to further learning or suggest changes
to their practice. In many cases supervisors and trainers use the comments to push trainees to think more about the application of learning to their own practice. An example of such a comment is:

I wonder if this course has given you ideas how you might promote Chlamydia self-screening amongst young practice patients?

Another trainer noted: “Reading needs to be put into practice”, responding to another reading entry with:

Needs to determine what would advise parents re: bronchiolitis infectivity.

Some trainers provided comments on the quality of reflection evident in the PDP entries, praising good entries with: “Good example of reflective learning”, or suggesting improvements, such as:

A summary of what are the key new learning points from these guidelines for you are would demonstrate what you have learned rather than just stating that you have read them.

Some comments highlighted areas that trainees may have missed or failed to consider. Examples include:

Dizziness is a very common complaint in GP. Don’t forget other causes other than ENT though, e.g. arrhythmias.

Others suggested further learning activities for the trainee to undertake, either alone or with the trainer:

What are other ways in which patients get dependent on their doctors and how can this be prevented? Possible joint surgery doing telephone consultations and may role play some.

Useful if you could add some notes from your visit to the eye clinic. You might also wish to reflect on the next three cases of eye problems you see.

Comments also provided the opportunity for professional advice and encouragement, particularly when trainees reflected on something that had not gone well. One trainer noted:

Sometimes consultations are quite hard and frustrating. Important to realise that. Important to make sure you are ok before next consultation (house keeping).

Finally, trainers could praise good clinical practice through the comments section, for example writing:

Demonstrated excellent clinical skills in OOH setting, adapted consultation skills properly and referred appropriately.
In summary, there is evidence that the PDP has potential as a formative learning tool, through comments made by GP trainers and educational supervisors to PDP entries. In the case studies examined, just over a third (34%) of shared entries received comments, suggesting that there is scope for greater use of this aspect of ePortfolio functionality.
Workplace-based assessment

The ePortfolio records results from workplace-based assessments and provides the facility for assessors to provide feedback and recommendations, so that the assessments may act as a formative learning tool.

Table A1.5: Feedback on WPBAs

<table>
<thead>
<tr>
<th></th>
<th>Number of Assessments</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ID 1</td>
<td>ID 2</td>
</tr>
<tr>
<td>CBD (n)</td>
<td>25</td>
<td>25</td>
</tr>
<tr>
<td>Feedback and recommendations (%)</td>
<td>96</td>
<td>100</td>
</tr>
<tr>
<td>GP</td>
<td>15/15</td>
<td>12/12</td>
</tr>
<tr>
<td>Hospital</td>
<td>9/10</td>
<td>13/13</td>
</tr>
<tr>
<td>Agreed action (%)</td>
<td>96</td>
<td>100</td>
</tr>
<tr>
<td>GP</td>
<td>15/15</td>
<td>12/12</td>
</tr>
<tr>
<td>Hospital</td>
<td>9/10</td>
<td>13/13</td>
</tr>
<tr>
<td>COT (n)</td>
<td>15</td>
<td>14</td>
</tr>
<tr>
<td>Feedback and recommendations (%)</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Agreed action (%)</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>MiniCEX (n)</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>Anything especially good (%)</td>
<td>60</td>
<td>100</td>
</tr>
<tr>
<td>Suggestions for development (%)</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Agreed action (%)</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

Case Based Discussion (CBD)

As shown in the table above, the case study trainees had between 20 and 27 full CBD records in their ePortfolios. All records except one contained feedback and recommendations and agreed actions. The table shows the number of CBDs completed in general practice and hospital settings. Feedback in both settings varied widely in terms of the amount of feedback given and its nature, with difference as great within settings as it was between settings. All feedback is reviewed collectively below.

Of the 218 records, 132 (61%) contained positive feedback which detailed areas of practice where the trainee had performed well. This could be brief or fulsome, as the examples below demonstrate:
Very good with patient and colleague interaction.

Excellent exploration of psychosocial issues.


Detailed written record.

Ninety-nine records (45%) contained advice for changes to future clinical practice. This included advice to utilise time more efficiently, improve record keeping, improve communication skills or holistic care, and advice on how to respond appropriately to similar cases in the future. One assessor reflected the views of many others with the feedback:

**PSYCHOSOCIAL PSYCHOSOCIAL PSYCHOSOCIAL PSYCHOSOCIAL PSYCHOSOCIAL.**

Other examples included:

To try to work more in partnership with patient by identifying their concern. Identifying the issues that need to be addressed. Improving prescribing practice by looking up knowledge gaps in the consultation.

In future consultations – if pts state a diagnosis rather than symptoms... be prepared to check their understanding of this and what THEY mean by it. Do not be afraid to challenge the diagnosis made by others if clearly it seems unusual – seek clarification and seek to challenge (in non-confrontational way!) both the pts health beliefs and diagnoses of others if they are at odds with what you think.

To record all important factors in records. To enquire how symptoms are affecting daily life including psychological effects. To book patients follow up appts with you if you want to follow them up.

Seventy-one records (33%) set specific learning tasks for the trainee to complete, including writing a reflective account, reading more about a particular condition, attending a training day or completing e-learning. Examples included:

Further reading on alcoholic liver disease.

Suggest attend mental capacity training day – level 2. Read Making Decisions Mental capacity guide for health and social care.

Reflect on any patient you have seen during the day with yourself and or other colleagues.
Twenty-eight records (13%) indicated that the trainee should continue with their current practice, noting, for example:

- Very much satisfied with progress, keep working hard.
- Continue in the same vein [sic].
- Continue to provide this high level of care.
- No need to change.

In 27 records (12%) assessors noted that the trainee needed greater experience and to see more cases to improve their knowledge and skills. Comments included:

- Could gain from getting more experience prescribing antidepressents.
- Needs further exposure to medical cases to acquire more experience.
- Further exposure to primary care.

Twelve records (6%) suggested that trainees read specific guidelines or protocols, such as NICE guidelines or local protocols for review of patients with a particular condition. Seven records (3%) contained no constructive feedback, noting merely “none required”, “useful case discussion” or “no feedback required”.

**Consultation Observation Tool (COT)**

As shown in the table above, the case study trainees had between 12 and 22 COT records in their ePortfolios, and all records included comments under both “feedback and recommendations” and “agreed action”. The large majority of feedback included broad details of the case, which would aid the trainee in remembering the consultation. Of the 138 records, 99 (72%) contained positive comments about areas of good practice, noting specific areas where the trainee’s consulting skills were strong. Examples included:

- Excellent non-verbal attentive listening skills with open friendly position and good eye contact. Empathic responses to patient’s problems.

- Consults in a patient-centred manner, with open questions, excellent non-verbal communication and active listening.

- Good rapport with the patient exhibited. Patient was at ease with the doctor. Very much a thorough consultation – well done!
Ninety-six records (70%) included advice for changes to future clinical practice. Many trainers noted the need for trainees to explore the patient’s agenda, explore the patient’s ideas, concerns and expectations, explore the psychosocial context, and check the patient’s understanding. For example:

In consultations try to extract the patient’s agenda early or help them understand what they want from the consultation.

Focus on involving the patient more and checking their understanding.

Try to incorporate patient’s health concerns and share management options.

Remember to explore psychosocial aspects in more detail – particularly if relevant. “Do you think you need antibiotics?” “Do you think that this will get better on its own?” Can be useful and efficient ways of exploring patient health understanding. Be aware of cues and try to explore them.

The last example given above illustrates a practice used by a number of trainers, that of providing specific sentences that trainees could incorporate within their own practice.

Feedback on 27 records (20%) set specific learning tasks for trainees to complete. These included reading up about a particular condition pertinent to the consultation, such as management of acne or HRT, and referral of patients with a family history of breast cancer, and also reading about consultation skills, for example in the book by Silverman et al. Other learning tasks included reflection on future practice, through self-review of video-recorded consultations. Comments included:

Further experimentation with suitable ways of confirming the pts understanding – so videoing and self reflection (“mark” your own videos) and then using what you have noted in your next consultations.

Read around the concept of multiple presentations in the consultation and how to prioritise.

Nineteen records (14%) included comments that the trainee should continue their current good practice, saying for example:

Continue present excellent approach.

Continue with the timely use of open and closed questions.

Continue patient centred/unhurried approach when appropriate.
In seven records (5%) trainers suggested that the trainee would develop through further experience and seeing more cases, noting, for example, “Ongoing consultations, shared surgeries”. Two (1%) suggested that trainees read up on specific guidelines, such as NICE guidelines on smoking cessation.

Mini-CEX

As shown in the table above, the case study trainees had up to 12 full mini-CEX records in their ePortfolios. All records contained suggestions for development and agreed action from the assessor, and the large majority (88%) contained comments under the heading “anything especially good”. In this category, assessors made positive comments about areas of good practice. Examples included:

- Clinical judgement is above what would be expected at this stage.
- Put patient at ease.
- A calm and unflappable communication style and good pacing of interview.
- Systematic approach to complicated case.
- Has confidence in clinical examination.

In the categories “suggestions for development” and “agreed action”, feedback was more variable between assessors. Of the 89 records reviewed, 32 (36%) contained feedback about the need for more experience. Examples included:

- More exposure to similar cases.
- Continue gaining experience.
- Continue to gain/consolidate paediatric experience.

Twenty-five records (28%) set specific learning tasks for trainees, the majority suggesting further reading on particular conditions, although others recommended e-learning or reflection on a particular case. Examples included:

- To read more about bronchiolitis and childhood asthma.
- Further reading on Prada [sic] Willi syndrome.
Reflecting on the reasons for collusion in families and the potential implications for end of life decision making and planning.

Leg swelling \( \text{dvt} \) is a frequent presentation to A+E. Suggest look at elearning options.

Twenty-one records (24%) gave advice for future clinical practice, noting issues such as time management, confidence, or specific considerations for examination and history taking. Examples included:

Be confident in asking about consanguinity.

Suggest address positioning of self re: examination technique – no bending etc. Try to anticipate potential findings on examination after history taking. Brain drives the hand and not vice versa.

Make sure there is sufficient time to assess patient so there is not a rush to finish.

Ask about recent infections. Observe pattern of weightbearing.

Could acknowledge patients’ feelings as they occur in the room rather than trying to manage or avoid them.

Explore patient/parent concerns and involve these into the paediatric checks.

Thirteen records (15%) noted that the trainee should “continue with current practice” or “carry on”. Ten (11%) gave no constructive comment, noting merely “none” or “nil”. Five (6%) suggested the trainee read relevant guidelines or protocols, and three (3%) made other comments, specifically advice to: follow up the particular patient; arrange to go on home visits with other team members; and continue to seek feedback and advice from seniors.

The table below compares the nature of feedback between the three types of WPBA examined. As may be expected, the majority of feedback comments on COT assessments (70%) give advice for future practice, and this is also the most common form of feedback for CBDs. The most common feedback in the mini-CEXes studied was advice to gain more experience and see more cases. For all three types of WPBA, at least a fifth of comments set specific learning tasks for trainees to complete.
The results above summarise data from feedback and recommendations within the WPBAs of nine case study trainees. The data suggests that WPBAs are being used as opportunities for formative learning, with clear examples of constructive feedback to trainees.
Appendix 2: Focus Group Schedules Year 3

Evaluation of the RCGP Training Curriculum: Focus Groups 2008: Trainees

Background

Introduce ourselves / roles / institutions

Brief overview of the evaluation:

Purpose of today’s discussion:

The aim of today’s discussion is to find out about your experiences of training with the new training curriculum. As this is the final year of the evaluation, we are keen to explore your views on how well prepared you feel for general practice, particularly ST3s (who have trained entirely with the new curriculum).

Outline confidentiality and gain consent to audio-record:
Question Schedule (Trainees)

Introduction

1. Could each of you introduce yourselves and say which year of training you are in and could you sum up in one sentence your experience of training with the new curriculum?
2. How many of you have completed Foundation training?
3. Are any of you training on a part time basis – if so could you please describe how you are finding the experience of working with the new curriculum?

Structure and Length of Training

1. How many of you will have experienced 18 months of GP and 18 months hospital experience in your training? (Explore any different structures).
2. What are your views on this balance between GP and hospital experience? (Identify positive and negative experience: explore relevance to their development as GPs).
3. What are your views on the current length of GP training (explore views on plans to extend GP training to 5 years; any areas that require more focus?

Fitness for Purpose

1. How effective do you feel your training this year has been in preparing you for general practice? (Explore how prepared ST3s feel for independent practice; identify areas of confidence; identify any gaps).
2. How could general practice training be improved?
3. Do you have any further comments on the effectiveness of the new training curriculum?

Assessment

A new system of assessment was introduced in 2007, including WPBAs and the CSA and AKT:

1. What have been your experiences of WPBAs during the current training year? (Explore marking criteria – any differences between hospital / GP assessors? Comparison between experience last year and this)
2. How useful have you found these assessment tools? (Explore learning value and their formative use; effectiveness at testing competences).
3. Have you experienced any difficulty getting WPBAs done? (Explore views on the number they are required to complete).
4. Some of you will have completed the CSA and AKT; what have been your experiences of these? (explore impact on learning and development; any difficulties experienced).
5. How do you think the new assessments have influenced your training experience?
6. Do you have any further comments to make on assessments?

Eportfolio

Another feature of the new training has been the introduction of the e-Portfolio.

1. What have been your experiences of using the e-Portfolio? (explore positive and negative aspects).
2. What do you consider are the main functions of the e-Portfolio? (Explore record of learning as trainee and throughout GP career; the learning log as a reflective tool; communication tool;).
3. How do you think the e-Portfolio has affected your training and learning?
4. Do you have any further comments on the e-Portfolio?

Written Curriculum

There is also a new written curriculum, made up of a core statement and 31 supporting statements.

1. How familiar are you with these? (Explore understanding of fit between core statement and supporting statements).
2. How do you usually access these (RCGP website, published guide, via e-Portfolio?)
3. How do you use the written curriculum (Explore: self-study / exam revision, preparation for placements, completing learning log; E-learning modules; use as a teaching tool by trainers / half day release etc.)
4. How useful have you found the written curriculum? (Explore impact on learning and development as GPs).
5. Do you have any other comments on the written curriculum?

Support

I'd like to ask about the support you've received during the current training year.

1. Do you feel you have received sufficient support during your training this year from your supervisors / trainers / programme directors in general practice and in hospital / community posts? (Explore support received from different sources including the role of educational supervisor in coordinating GP and hospital experience).
2. Do you feel you have had sufficient support and guidance over the year in relation to different aspects of the new training programme (assessments, exams, e-Portfolio, written curriculum etc):
   Explore any gaps and who has provided support - clinical supervisors, educational supervisors, GP Trainers, programme directors, other GP STs, Deanery, College).
3. How would you prefer to receive information and guidance about the curriculum, eportfolio and assessments?
4. Do you have any further comments to make on the support and guidance you have received in your training?

Summing Up

Do you have any other comments to make about GP training – anything you feel you haven’t had the opportunity to say?

Thank Participants
Evaluation of the RCGP Training Curriculum: Focus Groups 2010: Trainers and Programme Directors

Background to Study

*Introduce ourselves / roles / institutions*

*Brief overview of the evaluation:*

*Outline purpose of today’s discussion:*

*The aim of today’s discussion is to find out about your experiences of working with the new training curriculum. As this is the final year of the evaluation, we are keen to explore you views on how well prepared current ST3s (who have trained entirely with the new curriculum) are for independent practice as GPs.*

*Outline confidentiality and gain consent to audio-record:*
Question Schedule (Trainers)

Introduction

1. Could you start by introducing yourselves, could you say what your role is in GP training and how long you have been in these roles?
2. It is now 3 years since the new GP training curriculum was introduced; could I ask each one of you in turn to sum up your views of the curriculum and what you feel its impact has been on GP training?

Structure and Length of Training

*Can we briefly talk about the structure and length of GP training?*

1. What are your views on the balance in training between GP and hospital experience?
2. What are your views on the plans to extend GP training to 5 years? (if in favour – what should extended period focus on; are there areas where you feel current newly qualified GPs are not fully prepared in?)

Fitness for Purpose

*Moving on to explore more fully your views on the new curriculum and its fitness for purpose in training competent GPs:*

1. How successful do you feel the new training programme is in enabling trainees develop the skills and experience necessary for general practitioners? (Explore views on how this compares to previous system of training).
2. This year’s ST3s will be the first cohort to have trained entirely with the new curriculum, how well prepared do you feel they are to practice independently as general practitioners? (Explore views on how this year’s ST3s compare to registrars who trained with the previous system of training; if any differences are identified try to establish what these are attributed to – different GP training; postgrad training (Foundation / SHO – etc.).
3. Are there any aspects of GP training that have improved significantly because of the new curriculum?
4. Do you feel that the new training programme has had any negative impact on GP training?
5. How do you feel the training curriculum could be improved?
6. Do you have any further comments on the fitness for purpose of the new training curriculum?
Written Curriculum

I’d like to explore your views on each of the 3 main components of the new training curriculum – the written curriculum, the assessment system, and the ePortfolio; so firstly, the written curriculum: The written curriculum made up of a core statement and 31 interpretive or supporting statements, was introduced in 2007; I’d like to explore your views and experience of working with this.

1. How familiar are you with the written curriculum? (explore understanding of fit between core statement ‘Being a general Practitioner and 31 supporting statements;
2. How frequently do you use it? (explore how it is accessed: electronically, printed version, published guide?)
3. What are your views on the written curriculum and its impact on general practice? (explore perceptions of its value: to GP training; to qualified GPs; to inform others of general practice).
4. How has the written curriculum influenced the training you deliver? (explore impact on trainers’ and programme directors’ – how it is used; its impact on teaching and learning).
5. How could the written curriculum be improved?
6. Do you have any other comments on the written curriculum?

Assessment

I would like to explore your experience of the new system of assessment, including WPBAs, the CSA and AKT:

1. What are your views of these assessments? (explore views on their effectiveness in testing competence; their potential for formative use; their impact on promoting learning; the number that have to be completed; impact on trainees; impact on trainer’s role).
2. Do you consider that these assessments ensure that newly qualified GPs are competent to practice independently? (explore CSA and AKT; summative use of WPBAs and reviews).
3. How could the assessment be improved?
4. Do you have any further comments on assessment?

E-Portfolio

Since its introduction in 2007 there have been quite a few changes to the e-Portfolio and I would like to explore your views and experience of using it:
1. What have been your experiences of using the ePortfolio? (Explore how it is used – assessment, communication; checking trainee progress etc; differences between this and previous years; any problems this year? identify positive aspects of the ePortfolio).
2. Do you feel that the ePortfolio has had an impact on training and learning? (explore and identify positive and negative impact).
3. How could the ePortfolio be improved?
4. Do you have any other comments on the ePortfolio?

Support, Training and Guidance

I would like to explore your experience of the support and training you have received in delivering the new programme:

1. What support, guidance and training have you received in delivering the new training curriculum. (Explore range of provision; whether this is viewed as sufficient; who has provided this - Deanery, College, other GP trainers?)
2. What has been the overall impact of the new curriculum on your role(s) as trainers / programme directors? (Explore whether they now feel confident in how they should be delivering it).
3. What further support, training and guidance would be useful to enable you to deliver the new training programme?
4. Do you have any further comments on support, training and guidance?

Summing Up

1. Do you have any other further comments to make about the new GP training programme?
2. Is there anything you feel that you haven’t had the opportunity to say?

Thank Participants.
APPENDIX 3: Questionnaire for GP Trainers

Evaluation of the RCGP Training Curriculum: Trainer Questionnaire – Year 3 June 2010

The aim of this questionnaire is to find out about your experiences of working with the new training curriculum. As this is the final year of the evaluation, we are keen to explore your views on how well prepared current ST3s (who have trained entirely with the new curriculum) are for independent practice as GPs.

Please answer each question as fully as you can.

Introduction

1. What role(s) do you have in GP training?

2. How many years have you been involved in GP training?

3. How would you sum up your views of the new training curriculum for general practice and its impact on training?

Structure and Length of Training

1. What are your views on the balance in training between GP and hospital experience?

2. What are your views on the proposal to extend GP training to 5 years?
3. If GP training is extended, how should the additional training time be spent?

Fitness for Purpose of the New Curriculum

1. How successful do you feel the new training programme is in enabling trainees develop the skills and experience necessary for general practice?

2. How well prepared for independent practice do you feel this year’s ST3s are? How do they compare with GP registrars trained with the previous system?

3. Are there any aspects of GP training that have improved with the new training programme?

4. How do you feel the training curriculum could be improved?
Written Curriculum

*The written curriculum made up of a core statement and 31 interpretive or supporting statements, was introduced in 2007; I’d like to explore your views and experience of working with this.*

1. How familiar are you with the written curriculum?

2. How do you use the written curriculum in your training role?
3. What impact do you feel the written curriculum has had on:

i) GP training

ii) qualified GPs

iii) others outside of general practice?

4. How could the written curriculum be improved?

Assessment

I would like to explore your experience of the new system of assessment, including WPBAs, the CSA and AKT:

1. What are your views on the new system of assessment?
2. Do these assessments ensure that newly qualified GPs are competent to practice independently?

3. What impact have WPBAs had on teaching and learning?

4. How could assessment be improved?

EPortfolio

Since its introduction in 2007 there have been quite a few changes to the e-Portfolio and I would like to explore your views and experience of using it:

1. What are your views on the ePortfolio?

2. What impact has the ePortfolio had on teaching and learning?
4. How could the ePortfolio be improved?

Support, Training and Guidance

1. What support, guidance and training have you received in delivering the new training curriculum?

2. What has been the impact of the new curriculum on your training role?

3. What further support training and guidance would be useful to enable you to deliver the new training programme?
Summing Up

1. Do you have any further comments to make about the new GP training programme?

Thank you for your time in completing this questionnaire.

Please return to:

Dr Julie Bedward
CRMDE
School of Education
University of Birmingham
Edgbaston
Birmingham.

If completed electronically return to: j.Bedward@bham.ac.uk
APPENDIX 4: Interview Schedules for case study trainees; educational supervisors; programme directors; clinical supervisors.

Year 3: Interview Schedule Case Studies: Trainees.

Section 1: Overview of Third Training Year

How have you found your third year of general practice training? (Ask for an overview of the year - has it all been spent in general practice?).

Relationship / role of GP trainer (educational supervisor)

Can you tell me about the relationship you have had with your GP trainer / educational supervisor this year?

Explore:

- Value of supervisor in supporting their development as GPs
- The frequency of meetings / contact.
- Provision of formal / informal teaching
- Provision of supervision
- Their role as assessor

GP Practice where based for ST3 Year

How have you found the GP practice where you've been based for the year?

Explore:

- its size
- number of GP trainers
- number of GPs
- number of trainees

How much involvement have other staff at the practice had in your training this year?

How valuable has their input been?

Explore in relation to:

- Other GPs
- Other trainees
• Practice manager
• Other staff (nurses, health visitors, etc).

Study days i.e. VTS scheme / ½ day release

Have you continued to attend study days during your ST3 year?
(If yes, Can you tell me more about this and how useful you’ve found these over the year?).

Explore:
• Frequency
• Ability to attend
• Content / activities – perceived usefulness
• Relationships with other trainees – is it for trainees from all years or just ST3s? Do they share information / experience?
• Relationships with Programme directors and their role in GP training

Section 2: Fitness for Purpose

How effective do you feel the new training programme has been in preparing you to practice independently as a GP?

Explore:
• how prepared they feel for independent practice; identify areas of confidence; identify any gaps;
• the relevance and value of their 3 years of training to their development;
• what they view as the most valuable aspects of their training;
• any aspects that haven’t been helpful
• any areas where further training would be helpful.

How could general practice training be improved?

Section 3: Main Components of New Training

The Written Curriculum
How familiar are you now with the written curriculum – that is the core statement ‘Being a General Practitioner’ and the 31 supporting statements?

How useful have you found it during your ST3 year?

Explore:
- How it has been accessed
- How it has been used
- by them (ask for example)
- their trainers and supervisors (ask for examples)
- Its impact on teaching and learning
- Its relevance to general practice training

Do you consider that as a qualified GP the written curriculum will have any further use to you in your career?

Are there any changes you would recommend to improve the written curriculum?

The Eportfolio

How have you found using the ePortfolio during your current training year?

Explore:
- Any improvements
- Whether any problems have been experienced
- Any areas where further support or clarification would be useful

How have you used the ePortfolio this year?

Explore:
- Completing the learning log (ask for an example)
- Use of private space
- recording assessments
- record of achievement
- communicating with educational supervisor
- communicating with others
- other uses
What do you consider is its main function? (assessment tool, recording tool, learning tool, communication tool).

What impact has the ePortfolio had on your teaching and learning and development as a general practitioner?

Explore (and where appropriate ask for examples)
- its impact on self study
- its use by educational supervisor / GP trainer
- its value to your future career in general practice

Assessment

What has been your experience of completing assessments over the year?

Explore
- experience of the CSA exam
- experience / perceptions of value of different workplace based assessments (WPBAs):
  a) DOPS
  b) CBDs
  c) COTs
  d) MSF
  e) PSQ (patient satisfaction questionnaire)

How effective do you feel these assessments are in measuring competence?

Have these assessments had any impact on your teaching and learning?

Explore
  a) whether WPBAs are used to provide teaching opportunities (if so, ask for example);
  b) the potential for the formative use of these tools – is feedback received; if so, what kind of feedback and from whom? Ask for an example of feedback from assessment that has been useful.
  c) Are there any changes you would recommend to improve the assessments particularly during the ST3 year?

Section 4: Career Plans
Have you secured a post for next year?

Explore type of post secured / seeking; whether this is as:
  • Salaried GP
  • Locum
  • Partner
  • other

What are your reasons for choosing to commence your career in this type of post?

Section 5: Summing Up

Overall, how would you sum up your final year of general practice training?

What would you identify as the main benefits of your training this year?

What aspect of your training this year do you consider has made the most significant contribution to your learning and development as a GP?

What would you identify as the main problems you’ve experienced during your training this year?

Have you got any further comments to make about your experience of general practice training?

End of Interview

Thank for their contribution and check the following:

1. During the indepth analysis of their eportfolio would it be OK to contact them by email if there are any specific queries.
2. If they could provide the contact details for their educational supervisor – you will request this information in a separate email.
3. If the evaluation were extended would they be willing to continue as a participant to provide an account of their experience as a newly qualified GP.
For Academic Trainee Only

Structure of training and how the year has been spent

How have you found your third year of general practice training? (Ask for an overview of the year).

Explore relationship / role of clinical / academic supervisor over the year.

Frequency of contact
Provision of formal / informal teaching
Provision of supervision
Their role as assessor
Value of supervisor in supporting their development as GPs

(Then ask questions from other sections that are appropriate).
RCGP Curriculum Evaluation August 2010: Interview Schedule Case Studies: GP Trainers, Educational Supervisors, Programme Directors

Section 1: Background

When did you qualify as a GP?
Can you tell me about your training role in general practice?

Explore:
- Nature and extent of role (i.e) trainer, educational supervisor, programme director:
- Length of time in this/these role(s):
- Reason for taking on this/these roles.

How big is your practice – how many GPs? How many trainees?

Section 2: Overview of experience of training with the new curriculum

What has been your experience of training with the new curriculum?

Explore:
- The number of trainees they have worked with and their year of training – i.e. St1, 2, 3; F2s.
- The impact of the new curriculum on their roles as trainers / educational supervisors/ programme directors; any changes they have had to make.
- The impact of the new curriculum on the training practice (other GPs; other professionals; practice manager; patients).
- The impact of the new curriculum on the trainees and their development as GPs: any differences in trainees’ approach to training as a result of the new curriculum?
- The impact of the new curriculum on learning. What are now the main drivers of learning? How are trainees’ learning needs identified and addressed?
- The support received during implementation from deanery, colleagues, RCGP, other sources – was this sufficient? If not, what further support is / was required?
Section 3: Fitness for Purpose

How successful do you feel the new training curriculum is in preparing trainees to practice independently as GPs?

Explore:
- differences between new and previous training;
- impact of other changes (selection; Foundation Programme; EWTD);
- gaps in learning
- early career choices for newly qualified
- any changes considered necessary

Section 4: Views and experience of the different components of training

The written curriculum and curriculum statements

How familiar are you now with the new RCGP written curriculum?
How useful have you found it?

Explore:
- How it has been accessed
- How it has been used (ask for examples)
- Its impact on teaching
- Its potential beyond specialty training (undergraduates; F2s; CPD; revalidation etc).

How familiar do you feel your trainee(s) is / are with the new written curriculum?
Do you feel it has had any impact on their learning and development as GPs?
Are there any changes you would recommend to improve the written curriculum?

The ePortfolio

How have you found using the ePortfolio this year?

Explore:
- Whether any problems have been experienced
- Whether sufficient support / guidance has been received
- Any improvements over the 3 years
- Any areas where further support or clarification would be useful
Any recommended changes
Its potential

Could you describe how you use the eportfolio with your trainee(s)

Explore:
- Its role in assessment
- Its impact on teaching and learning
- Its impact on relationships with trainees (including ST1s and 2s, if appropriate).
- Any recommended changes

Assessment
What has been your experience of working with the new workplace based assessments (WPBAs) (DOPS, CBDs, COTs, Mini-EX, MSF, PSQ)?

Explore:
- views on their effectiveness:
- their use in hospital based settings:
- the potential for the formative use of these tools (ask for examples):
- their impact on teaching (ask for examples):
- impact on learning:
- impact on relationship with trainee:
- the practicalities of completing the required number.

Are there any changes you would recommend to improve the WPBAs?
What are your views on the new CSA exam?
What are your views on the AKT?

Section 4: Summing Up
Overall, how would you sum up your experience of training with the new curriculum over the last 3 years?

What would you identify as the main benefits of the new curriculum?

What would you identify as the main problems you've experienced training with the new curriculum?

Have you got any further comments?
APPENDIX 5: Interview Schedule for Deanery Leads Year 3

Interview Schedule for Deanery Leads – September 2010

Background

What is your role at the deanery? (Explore what this involves).

Have there been any significant changes to the structure of how GP training is organised in your deanery over the last 3 years (Explore numbers of STs, trainers, PDs; time spent in general practice / hospital posts; types of trainees selected for GP training.)

How well prepared do you feel the deanery was for the introduction of the new curriculum in August 2007?

What impact has the new training curriculum had on your role?

Explore:
• the development of materials, training and systems;
• how these needs were identified;
• whether these have been sufficient;
• any further developments planned.

Fitness for purpose

How successful do you feel the new training curriculum is in preparing trainees to practice independently as GPs?

Explore:
• differences between new and previous training;
• impact of other changes (selection; Foundation Programme; EWTD);
• gaps in learning
• early career choices for newly qualified
• any changes considered necessary

Section 4: Views and experience of the different components of training

The written curriculum and curriculum statements
What are your views on the new RCGP written curriculum and its impact on training?

Explore:
- Its relevance to GP training
- The role of the deanery in facilitating improved understanding; how successful this has been.
- How it is used by trainers, GPSTs, hospital clinical supervisors.
- Its accessibility to trainers
- Its accessibility to GPSTs
- Its impact on trainees’ learning and development
- Its potential beyond specialty training (undergraduates; F2s; CPD; revalidation etc)

Are there any changes you would recommend to improve the written curriculum?

The ePortfolio
What are your views on the ePortfolio and its impact on training?

Explore:
- Its potential in training (as a tool for recording, assessment, teaching and learning, communicating)
- The role of the deanery in facilitating improved understanding; how successful this has been
- Whether any problems have been experienced

Are there any changes you would recommend to improve the ePortfolio?

Assessment
What are your views on the new system of assessment including the workplace based assessments (WPBAs) and the AKT and CSA?

Explore:
- views on their effectiveness:
- their use in hospital based settings:
- the potential for the formative use of these tools (ask for examples):
- their impact on teaching (ask for examples):
- impact on learning:
- the practicalities of completing the required number.
Are there any changes you would recommend to improve the WPBAs?

What are your views on the new CSA exam?
What are your views on the AKT?

Section 4: Summing Up

Overall, how would you sum up your experience of training with the new curriculum over the last 3 years?

What would you identify as the main benefits of the new curriculum?

What would you identify as the main problems you’ve experienced training with the new curriculum?

Have you got any further comments?