2009; 31: 320–339 MEDICAL TEACHER

BEME GUIDE

The effectiveness of portfolios for post-graduate assessment and education: BEME Guide No 12

CLAIRE TOCHEL, ALEX HAIG, ANNE HESKETH, ANN CADZOW, KAREN BEGGS, IAIN COLTHART & HEATHER PEACOCK

NHS Education for Scotland, UK

Abstract

Background: Portfolios in post-graduate healthcare education are used to support reflective practice, deliver summative assessment, aid knowledge management processes and are seen as a key connection between learning at organisational and individual levels. This systematic review draws together the evidence on the effectiveness of portfolios across postgraduate healthcare and examines the implications of portfolios migrating from paper to an electronic medium across all professional settings.

Methods: A literature search was conducted for articles describing the use of a portfolio for learning in a work or professional study environment. It was designed for high sensitivity and conducted across a wide range of published and unpublished sources relevant to professional education. No limits for study design or outcomes, country of origin or language were set. Blinded, paired quality rating was carried out, and detailed appraisal of and data extraction from included articles was managed using an online tool developed specifically for the review. Findings were discussed in-depth by the team, to identify and group pertinent themes when answering the research questions.

Results: Fifty six articles from 10 countries involving seven healthcare professions met our inclusion criteria and minimum quality threshold; mostly uncontrolled observational studies. Portfolios encouraged reflection in some groups, and facilitated engagement with learning. There was limited evidence of the influence of a number of factors on portfolio use, including ongoing support from mentors or peers, implementation method, user attitude and level of initial training. Confounding variables underlying these issues, however have not been fully investigated. A number of authors explored the reliability and validity of portfolios for summative assessment but reports of accuracy across the disparate evidence base varied. Links to competency and Quality Assurance frameworks have been demonstrated. There were conflicting reports about whether the different purposes of portfolios can be combined without compromising the meaningfulness of the contents. There was good evidence that the flexibility of the electronic format brought additional benefits to users, assessors and organisations, and encouraged more enthusiastic use. Security of data remained a high priority issue at all levels, and there was emerging evidence of successful transfer between electronic portfolio systems.

Conclusion: The evidence base is extensive, but contains few high quality studies with generalisable messages about the effectiveness of portfolios. There is, however, good evidence that if well implemented, portfolios are effective and practical in a number of ways including increasing personal responsibility for learning and supporting professional development. Electronic versions are better at encouraging reflection and users voluntarily spend longer on them. Regular feedback from a mentor enhances this success, despite competing demands on users' time and occasional scepticism about the purpose of a portfolio. Reports of inter-rater reliability for summative assessments of portfolio data are varied and there is benefit to be gained from triangulating with other assessment methods. There was insufficient evidence to draw conclusions on how portfolios work in interdisciplinary settings.

Introduction

Traditionally portfolios have been artistic (and then financial) compilations of documents for presentation, but more recently the term has come to encompass the collection, management and presentation of a far greater diversity of material for an increasing array of professions. Portfolios in healthcare education are used for a range of purposes, including supporting reflective practice, delivering summative assessment and aiding knowledge management processes. They are seen as a key connection between learning at organisational and individual levels. With portfolios' migration to the electronic medium the extent and depth of their usage continues to grow as they, for example, integrate with e-learning platforms and enable rapid analysis of data supporting learning.

Amongst the healthcare professions, nursing has a history of using portfolios for reflective practice and they are now

Correspondence: Claire Tochel, NHS Education for Scotland, Thistle House, 91 Haymarket Terrace, Edinburgh EH12 5HD, UK. Tel: +44 131 313 8085; fax: +44 131 313 8001; email: claire.tochel@nes.scot.nhs.uk

ISSN 0142–159X print/ISSN 1466–187X online/09/040320–20 © 2009 Informa Healthcare Ltd.
DOI: 10.1080/01421590902883056

Practice points

- It is key that portfolio implementation is well designed and sustained, with high-level organisational support, to ensure uptake.
- A well-informed mentor can have considerable impact on uptake, especially when regular feedback is given.
- Users can be simultaneously sceptical about a portfolio's intended purposes and appreciative of what it can deliver for them personally.
- There is agreement, and some evidence, that portfolio users feel increased responsibility for their learning.
- Summative assessment of portfolio contents can be reliable among multiple raters, but triangulation with other sources is desirable.
- Electronic portfolios have been demonstrated to have a number of benefits (flexibility of access and content, potential for links). Users spend longer with electronic versions.
- E-portfolios are more effective for feedback and encouraging reflection than paper ones, though assessments in both are well correlated.

required by the UK Nursing and Midwifery Council[1] But recent years have seen portfolios contributing to educational provision under the auspices of many regulatory bodies and professional organisations. For example, in the UK in the field of medicine they are used by some medical schools and following the introduction of Modernising Medical Careers,[2] required by the Postgraduate Medical Education and Training Board, medical schools and numerous Royal Colleges of Medicine.

The expanding and broadening use of portfolios in postgraduate healthcare education is being actively considered or used for both recertification/revalidation and continuing professional development. For high stake decisions in any setting, there is a clear need for validated assessment criteria against which to evaluate portfolio data (Tillema & Smith 2007). Alongside the rapid growth of portfolio usage has been corresponding publication of a diverse range of evidence and descriptions of the work; however, much of this is descriptive and there has been little attempt to aggregate or synthesise high quality findings.

Initial scoping work in 2005 established that no single study had comprehensively combined all evidence regarding the effectiveness of portfolio use. This systematic review draws together the evidence across postgraduate healthcare education and examines the implications of portfolios migrating from paper to an electronic medium, building on Challis' (1999) guide.

Review questions and objectives

The review aimed to answer three research questions in order to meet a number of objectives:

(1) Are portfolios effective and practical instruments for post-graduate healthcare education?

- establish how effective portfolios are as instruments to support reflective practice
- summarise the strengths and weaknesses of portfolios for conducting formative and summative assessment
- synthesise the evidence on portfolio usage in the work place and how they can further education
- ascertain whether portfolios can accurately support the educational needs of learners.

(2) What is the evidence that portfolios are equally useful across health professions, and can they be used to promote inter-disciplinary learning?

- determine any differences in the effectiveness of portfolio across the professions, and
- reveal how they can be used to support inter-professional education.

(3) What are the advantages and disadvantages in moving to an electronic format for portfolios?

- examine the impact and implications of migrating from paper to electronic format.

The terms 'effective' and 'practical' were discussed over the course of several group meetings from the team's broad experience of portfolios, and for the purposes of this review are defined as follows. An effective portfolio is one which meets the needs of the users, supports them to achieve the aim of the portfolio and delivers the required elements to an appropriate standard. A practical portfolio is one which is user-friendly, efficient in terms of the overall cost and time demands on both the user and the support team who maintain it.

Review methodology

The review was conducted by a team of seven (initially eight) employees of NHS Education for Scotland drawing on a range of experience in health services research, information science, systematic review, social science, medical educational research and development, nursing education, social anthropology and postgraduate medical education. It was carried out in line with established BEME methodology (www.bemecollaboration.org).

Literature search

The literature search was conducted across a wide range of sources relevant to professional education. The database search covered all relevant health as well as educational databases, and included: MEDLINE, British Education Index, ERIC, HMIC, EMBASE, CINAHL, British Nursing Index, TIMELIT and AMED.

The strategies were designed for high sensitivity to minimise the risk of missing potentially relevant articles. The search ran from the earliest available date in each database (e.g. 1966 MEDLINE) to January 2006 and did not limit by language, geography, or research methodology. An update search was conducted in October 2007 to include evidence published during the course of the group's first wave

of analysis. The full Medline search strategy is outlined in Appendix 1 (on BEME website: www.bemecollaboration.org). Additional strategies for the other databases were based upon this search using consistent syntax and terminology.

One member of the team conducted an initial filter of titles for clear irrelevance to the review, and then a list of titles and abstracts were distributed (where available) to randomly selected (and shuffled) pairs of team members. Reviewers read the available information on each citation independently and decided whether the full text should be ordered for appraisal. They compared their decisions and discussed anomalies, requesting the article if one or both reviewers were unsure.

Once reading full articles, the team were also asked to identify cited references that might be of importance to the review. A cited reference search was conducted in late 2007 on the highest rated articles and where appropriate these were obtained.

Grey literature. On an agreed date in September 2007 and then again in November three of the team independently searched www.google.co.uk for grey literature (evidence not formally or commercially published). A variety of search terms were used, related to the effectiveness of portfolio usage for education or learning (Box 1).

Each of the three team members reviewed retrieved citations for relevance to the review questions, and saved any potentially useful documents to a shared storage space, thus avoiding duplication. Each person committed 2–3 h to this search; the second date ensured results were as close to saturation point as reasonably possible.

Box 1. Combinations of search terms used.

Portfolio	Healthcare	Research
Eportfolio	Health professional	Evaluation
Personal development plan	Learning	Effectiveness

Selection of articles

Inclusion and exclusion criteria. In order to conduct a thorough and pragmatic review of the literature; broad criteria were set (Box 2). All study design types were included, as it was established by early scoping searches that in this field there was little experimental research. Letters, editorials and conference abstracts were obtained in case they referred to other work which may have provided some evidence.

Types of portfolio. The group discussed the boundaries and grey areas of what constituted a relevant portfolio during the early phase of the review. The type of portfolio of interest would include a collection of information to facilitate learning, and indicate engagement with the portfolio by the user, above and beyond a list of items; e.g. clinical procedures undertaken by the user. A precise definition was not pursued, as it was feared it may limit the generalisability of the review. Each article was considered on its own description of the tool used,

Box 2. Inclusion and exclusion criteria.

Research question	Inclusion criteria	Exclusion criteria
Questions 1 and 2	Articles which, both: • were about the use of a portfolio by a qualified professional group (in a healthcare setting) in an educational/learning/professional development context AND • described one or more of the following concepts: i. what you do with portfolios ii. what you learn by using them iii. how a portfolio is used iv. perceptions of effectiveness of portfolio usage (even if descriptive) v. informal evaluations i.e. perceptions, thoughts, views of users or others? vi. formal evaluation of portfolio as tool vii. portfolios contribution to career development	• articles including only undergraduate students (see question 2 exception) • articles where the portfolio was no more than a log-book or checklist of procedures or items
Question 3	Articles which described any aspect of the use of an electronic portfolio.	• articles where the portfolio was no more than a log-book or checklist of procedures or items • articles which only described the technical specification or implementation of a portfolio • articles where the portfolio was not used for learning e.g. as a teacher's planning tool/or for collation of pupil's work

Article types included – All questions
- any publicly distributed document (to include published and listed in a literature database, published in a print or electronic journal, or a publicly available website)
- any language (identifiable by English-language index terms)
- any country of origin

how it was used, and was included if enough information was provided to distinguish the interactive learning or reflection element which was of interest to this review. This meant that the same term, e.g. log book, may appear in one article representing a simple checklist tool (and thus be excluded) but in another it may incorporate a reflective element in which case the article would be included.

Types of participant. The main focus of our review was on articles involving postgraduate healthcare professionals; this was agreed in collaboration with another BEME systematic review group based at the University of Birmingham who were reviewing the literature to report on the effects of portfolio use on undergraduate student learning (Buckley et al. 2009). The term 'post-graduate' was defined as having passed the point at which professional status is achieved, i.e. when an individual is employable in their field. Outwith the UK, and across the health professions, however there are variations in the terminology for the status of an individual with a healthcare qualification or degree.

With regard to answering our question on electronic portfolios, an initial scoping search revealed little evidence. As this was an area of particular and growing interest, inclusion criteria were widened to include participants of all types (i.e. including teachers and students in all learning settings) for this part of the review. This constitutes an area of overlap with the Birmingham review (Buckley et al. 2009).

Types of outcome measure. Evidence on any reported outcome measure that addressed our research questions was included. Anticipated categories of outcomes which would inform on the effectiveness and practicality of portfolios in learning included:

- skill (e.g. communication, clinical examination, reflection/self-awareness)[3]
- attitude (e.g. views of learning and teaching, self-confidence, satisfaction);
- behaviour (e.g. level of portfolio usage, participation in further learning);
- efficiency (e.g. time taken to prepare portfolio).

Articles providing only procedural details of a portfolio implementation process rather than describing the learning involved were not included, as were articles which described only a portfolio product specification.

Assessment and appraisal of the evidence – Online form

An online form was developed to store citation information and details of critical appraisal and data abstraction by each reviewer. This was of considerable benefit as the team were based in four locations across Scotland, and therefore it was desirable to agree standardised formats for evaluating and managing information. This also facilitated the process of data checking and analysis. A software programmer was recruited to develop the form to the team's specifications; this was done as an ASP coded web application which stored form data in a Microsoft SQL Server 2000 database. Web access allowed users the ability to enter or check data at any internet-linked computer. Data was ultimately downloaded into another application (Microsoft Access) for synthesis and analysis.

Individual usernames were issued to the team, and everyone tested the system on several articles to identify technical bugs or elements which could be improved. A record was then created for every full-text article, and a link was made to a pair of reviewers so that they could click on it, and begin entering data when ready (more details are given further).

Quality assessment – All full-text articles. Firstly, the whole team read and scored five articles and discussed them in depth. This process allowed a common understanding of the elements required to achieve an acceptable standard for inclusion to be reached. These elements included study design (sample size and selection), execution of research elements, analysis and clear/fair reporting of results. The team preferred this method to a rigid points-based checklist to deal with the anticipated variety of study types. A quality score was applied on a scale of 1–5: 1 (very low), 2 (low), 3 (reasonable), 4 (high) and 5 (very high) and the team established a good level of consistency. These terms are used throughout the rest of this review to indicate the score applied to cited studies. For example, a study with a random selection of participants, achieving a representative sample of a population (if clearly stated e.g. including baseline characteristics) would score as 4 (high) or 5 (very high) depending on its size. A study using a convenience sample, or a sample whose characteristics were not described, would not score higher than a 3 (reasonable). The process revealed that an additional level of refinement was required as some of the literature fell within our inclusion criteria but could not directly answer our questions (listed in Section 1.2). Therefore a score for relevance to the research questions was added (on the same 1–5 scale). It was agreed that a minimum score (for relevance plus quality) of 7 out of the10 would be acceptable, but with a minimum of 3 on both measures (i.e. a score of 5 + 2 was not acceptable).

Each full text article identified by the literature searches was randomly distributed to two of the team, who read it in full, blinded, to identify whether it met our inclusion criteria (Box 3), and to score its methodological quality. This data was entered into their own record for that article on the online form. The pair then discussed each score and their reasoning for any discrepancies. If these could not be resolved to mutual satisfaction during this discussion the article was referred to a third party within the team. This happened on three occasions, and in one case the article was shared with the entire team to agree an appropriate decision. Pairs were shuffled, so each reviewer was paired with everyone else on the team during the review.

Critical appraisal and data abstraction – Included articles. Once the pair agreed that an article met minimum standards (i.e. scored 7 or more), it was assigned to one of them to fully appraise, and extract data which answered one of more of our research questions. The team member paired with them for scoring was available to check or clarify any issues, but as little complex data was retrieved, double extraction was not undertaken.

The online form comprised a detailed checklist for appraising different types of research method or analysis employed (including literature reviews) (Appendix 2 on BEME website: www.bemecollaboration.org). For every full text article, assigned reviewers were asked to:

- rate the appropriateness of the article design to answer their research questions;
- describe the design and methodology;
- rate how well the study was conducted;
- rate the quality of the analysis and reporting;
- record the main findings and conclusions
- assess the study's impact level (Box 4); and
- note any issues or concerns they had about the study quality or relevance to our review.

Study impact level. Kirkpatrick's hierarchy is used when reviewing evidence to indicate the extent to which a study reveals the impact of an intervention on participants (Hutchinson 1999). For example, a survey of users may report on their interaction or involvement with the portfolio, demonstrating a level one impact, in that they are engaged with the intervention. A before and after study may show that users' attitudes or knowledge level were changed by the portfolio (level two impact) or that users incorporated learning into their work (level three). A more detailed description of Kirkpatrick's hierarchy[4], adapted for medical education by the BEME Collaboration group (Harden et al. 1999), is given in Box 4.

Box 3. Abbreviations/definitions.

ACGME	Accreditation Council for Graduate Medical Education
AHP	Allied Health Professional
AMEE	Association for Medical Education in Europe
ASP	Active Server Pages
CPD	Continuing Professional Development
DEN	Doctor's Educational Needs (a self-directed learning tool)
FE	Further Education
GP	General Practice/Practitioner
KOALA	Computerized Obstetrics and Gynecology Automated Learning Analysis
NES	NHS Education for Scotland [www.nes.scot.nhs.uk]
PDP	Personal Development Plan
PGEA	Postgraduate Educational Allowance (an educational payment to GPs)
PRHO	Pre-Registration House Officer (a now obsolete term for the first postgraduate year of training)
PUN	Patients Unmet Needs (a self-directed learning tool)
RCT	Randomised Controlled Trial
SpR	Specialist Registrar (senior training grade)
SWOT	Strengths, Weaknesses, Opportunities and Threats
VT	Vocational Training/Trainee

Analytical procedure – Reporting the findings

The studies identified had insufficient homogenous or quantitative data to allow meta-analysis or formal synthesis. Reviewers individually identified all pertinent themes arising from each included article's findings. The team then discussed the evidence base in its entirety and themes were collated into related groups according to how they meaningfully answer or inform this review's research questions. These grouped themes form the structure of the results section in the form of a detailed narrative description of the evidence.

Results

Search results

From the electronic database searches 376 articles were found to meet this review's inclusion criteria. These were independently scored by pairs for quality and relevance to the review questions; 46 met minimum standards and were included.

After approximately 8 h spent on the grey literature search, a saturation point was not reached, but it was agreed it was impractical to keep searching. Forty six articles were identified of which four met our inclusion criteria. Citation follow-up and expert contact provided a further six articles which met the threshold. Therefore 56 articles were included in total (Figure 1).

Distribution of articles

Geographic location. Included studies were conducted (or written by people) in 10 countries (Figure 2). Almost half of the studies were conducted in the UK (46%) and almost a third from the USA (29%). There were four each from Canada and the Netherlands, and one each from six further countries.

Box 4. Adapted Kirkpatrick's hierarchy (Harden et al. 1999).

Level 1:
- Participation – covers learners' views on the learning experience, its organisation, presentation, content, teaching methods, and aspects of the instructional organisation, materials, quality of instruction.

Level 2:
a) Modification of attitudes/perceptions – outcomes relate to changes in the reciprocal attitudes or perceptions between participant groups toward intervention/simulation.
b) Modification of knowledge/skills – for knowledge, this relates to the acquisition of concepts, procedures and principles; or skills this relates to the acquisition of thinking/problem-solving, psychomotor and social skills.

Level 3:
- Behavioural change – documents the transfer of learning to the workplace or willingness of learners to apply new knowledge and skills.

Level 4:
a) Change in organisational practice – wider changes in the organisational delivery of care, attributable to an educational programme.
b) Benefits to patient/clients – any improvement in the health and well-being of patients/clients as a direct result of an educational programme.

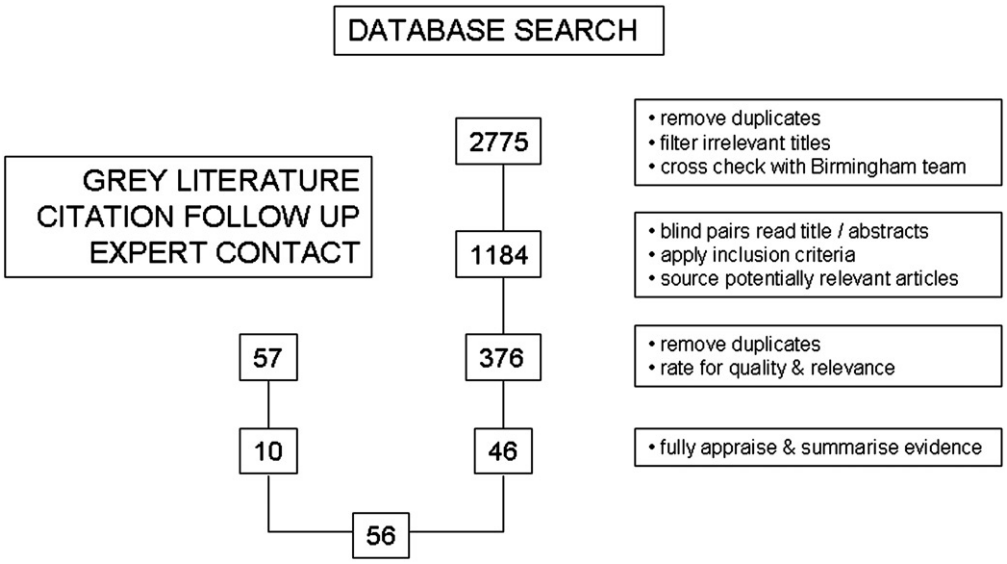

Figure 1. Flowchart of search and selection process showing number of included articles identified at each stage of the review.

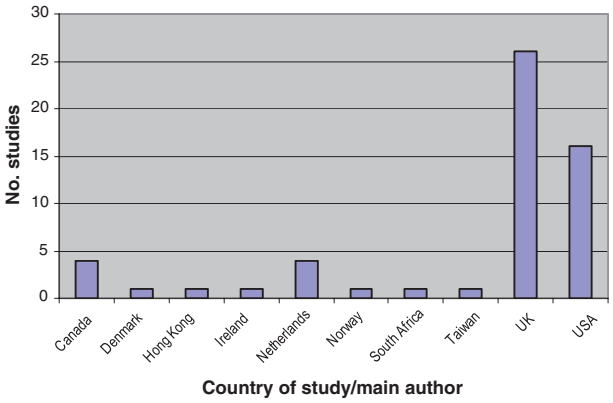

Figure 2. Location of included studies (or main author if not clearly stated).

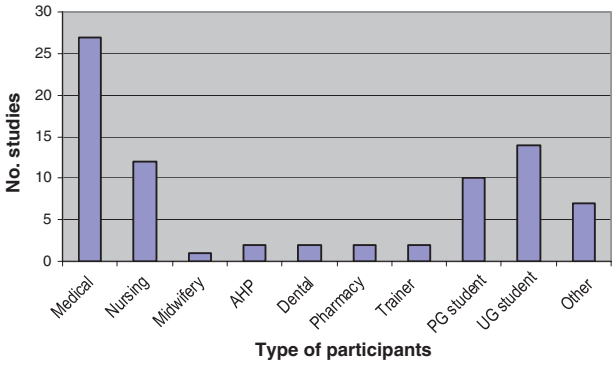

Figure 3. Professional group involved in included studies (UG students and non-healthcare setting participants included in 'other' relevant to question 2 – electronic portfolio only). AHP = Allied Health Professions.

Professional group participating in included articles. Among the 56 included articles, seven different healthcare professional groups were represented, most commonly medicine ($n = 27$) and nursing ($n = 12$) (Figure 3). Of the articles in medicine with a clearly stated setting, 13 were based in hospitals and 10 in general practice. Other groups of postgraduate portfolio users included trainees in counselling and educational technology.

Undergraduate students (included only for the electronic portfolio question) were predominantly medical and teaching students, and 'other' groups included school teachers, principals, and educational supervisors.

Description of included studies

On the basis of study design, execution and reporting more than half of the included articles just exceeded our quality threshold scoring 3 out of 5, and were therefore defined as 'reasonable' quality ($n = 32$). Twenty-four scored 4 (rated 'high' quality). None were rated 5, i.e. 'very high' quality.

Types of study design. The most common study design was uncontrolled observational ($n = 33$) (Figure 4). There were also 10 comparative studies (six observational and four experimental) and six literature reviews (three of which were described as systematic reviews). This categorisation was not always straightforward as some articles did not follow a recognisable methodology, or did not report it clearly (seven remained uncategorised – primarily descriptive reports).

Types of portfolio. The range of portfolio type used was very broad, and our review included all which involved the key element of user reflection or interaction with the contained information, for example a portfolio attempting to link learning to professional recertification, through to a very different one used to develop a counselling case profile. In many cases,

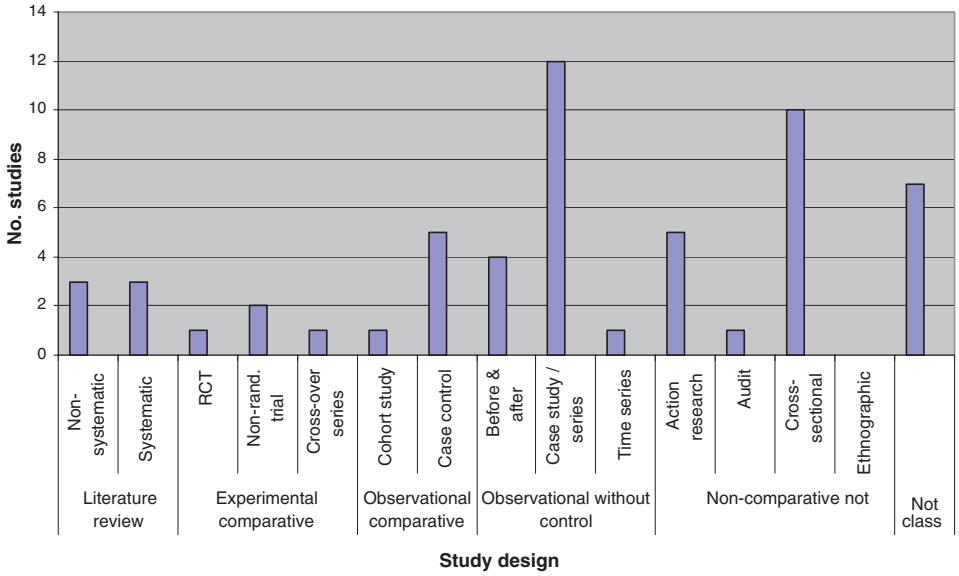

Figure 4. Study design of included articles.

descriptions of the content of the portfolio were scarce, therefore taking generalisable messages from the evidence base was not straightforward or justified.

Study impact level. According to Kirkpatrick's Hierarchy (Box 2), most included studies were found to impact on the learning of the portfolio user (a level 2 impact, $n = 26$), with fewer demonstrating effects on behaviour (level 3, $n = 10$) (Figure 5). Two were found to indicate some effect on organisational change or benefit beyond the portfolio user (level 4).

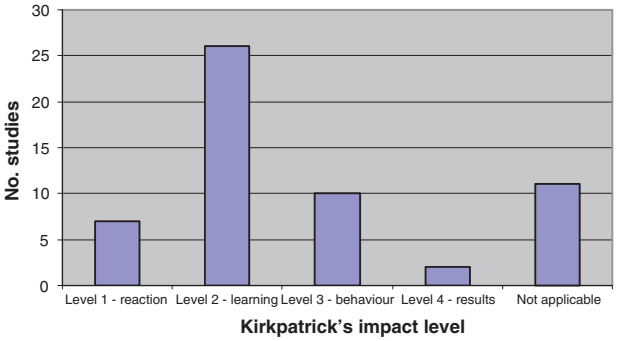

Figure 5. Impact level of included studies.

The evidence

This section reports relevant results from all 56 articles which met our minimum quality threshold. Under each theme, evidence from every relevant included article is presented. For the six included literature reviews which were found to meet our minimum standards for quality and relevance, evidence of relevance to our review questions and populations of interest are reported followed by additional primary evidence identified by our review. Higher quality studies (i.e. scoring 4 rather than 3) are given prominence in each section.

Are portfolios effective and practical instruments for post-graduate healthcare education? The review team identified 38 articles which describe or test various aspects of the effectiveness and practicality of portfolio use. The evidence is grouped under the following themes:

- factors influencing portfolio use;
- use of portfolios for assessment;
- outcomes of portfolio use.

Factors influencing portfolio use. The evidence brought together in this section demonstrates the extent to which the effectiveness and practicality of a portfolio (to an individual or an organisation) are influenced by a range of factors. These include users' positive or negative attitudes; gender; different levels of organisational support during implementation; early or sustained support/mentoring and the challenges of the time and cost involved in portfolio use. This section examines the evidence for factors influencing use in general, obtained from 23 relevant articles, but where authors specifically examine the electronic medium, or compare electronic with traditional portfolios, the topic is discussed in the later electronic section.

User attitude

A UK study of personal development plan (PDP) usage in general practice (GP) medicine reported somewhat contradictory attitudes in users (Cross & White 2004). Whilst 64% of respondents ($n = 277$ in total) reported submitting their PDP as a means to obtain PostGraduate Educational Allowance (PGEA) accreditation and 53% agreed a PDP was a '*hoop-jumping*' exercise, their attitudes to the educational value of PDPs were simultaneously quite positive – depending on the educational tool. Only 42% found the portfolio (referred to as a 'regional workbook') of use and 36% valued SWOT analysis; however, 61% valued the use of Patient Unmet Needs

and Doctor's Educational Needs (self-directed learning tools), 74% valued the reflective practice and 81% thought the Significant Event Analysis component was valuable. These survey results, based on a strong postal response rate (81%), convey wide variation in what general practitioners value in their PDP with the high rating of some tools seeming to contradict the notion that the PDP is merely a form-filling exercise. The potential cynicism expressed by many in completing PDPs was also balanced by the fact that 82% of respondents saw the PDP as forming a substantial part of their revalidation.

A small, well-conducted two-part study (focus-group, semi-structured interview) of UK general dental practitioners also reported that portfolios could be well received in revalidation in this sector (Maidment et al. 2006a, b). Feedback from the volunteer group was largely very positive about the potential for revalidation. They felt that including a system for appraisal would be beneficial, although the small ($n = 10$) size of this study within primary care dentistry may limit the generalisability of the findings.

Gender

Murray's UK study used quantitative and qualitative analysis of e-portfolio data from Pre-Registration House Officers (PRHOs), their educational supervisors, nurses, nurse supervisors and two cohorts of further education (FE) students (the latter group out with the inclusion criteria of this review question), about engaging with portfolios (grey literature, 2007). The authors compared portfolio use by gender, and showed that a greater proportion of female users accessed the portfolio following training (64% vs. 55%), but were less likely to progress from being a 'reader' to a 'poster'. Once using the portfolio, females were more likely to remain users and qualitative analysis indicated that they were more likely to perceive and describe positive educational effects. However, these analyses did not incorporate the effect of being a voluntary or mandatory user.

Implementation method

Low initial compliance rates were reported by a USA surgical study, after implementing their Surgical Learning and Instructional Portfolio, a case-based portfolio that included self-assessment and reflection (Webb et al. 2006). Although the programme director and coordinator actively tried to improve compliance, the rate remained <50% and residents ($n = 40$ in total, but early numbers are not clearly reported) did not rate it highly. No detail was provided of the implementation process to this point. The processes were revised in 2004 to include monthly feedback, topic collation and coded discussion as new resources, e-mail contact with the supervisor and quarterly notification of incompletion to all relevant parties. Once put into practice, the lessons learned from the initial implementation saw compliance rise to 100% and considerable higher appreciation from residents. The article cited '*dedicated faculty review*' and '*perceived importance of the project*' as critical factors in successful implementation. This article would have benefited from the provision of more detail, particularly on the initial implementation work, but does provide reasonable evidence on how embedded communication and feedback during implementation can influence uptake.

Other articles reported similar limited evidence from doctors in other specialties. Snadden and Thomas (1998) conducted a qualitative action research study in a geographically diverse area across the north of Scotland on portfolios in GP vocational training. This revealed doubts regarding the introduction of portfolio learning without '*intensive support at a one to one level*'. Their work, which included extensive interviews and focus groups with 20 pairs of trainers and their trainees (four were unavailable and one pair refused to participate), concluded the implementation process for portfolios might be more important than the structure of the portfolio itself. Kjaer et al. (2006) reported on the implementation of an e-portfolio for GP medicine in Denmark ($n = 90$ GP trainees). Similarly, this article did not set out to measure the implementation process, but cited proper time and scheduling, consideration and provision of information about the portfolios use to users, and a '*practical technical demonstration*' as being key to proper implementation.

Murray's UK study demonstrated that from a user's initial contact with a portfolio to their full engagement with it, the key factor in uptake is its relevance to the individual (Murray, grey literature 2007). As previously mentioned, in this study the portfolio was used by PRHOs, nurses, and two cohorts of further education (FE) students. Use was compulsory for PRHOs ($n = 33$) and voluntary for other users ($n = 171$) and this is reflected the proportion who accessed it (88% vs. 55%) and made entries (88% vs. 23%).

Mentoring/support

The impact of constructive interaction with a mentor or supervisor on portfolio use has been explored in a number of studies. Driessen et al. (2007b) state in their recent review of the effectiveness of portfolios in medicine (30 included articles, of which nine were in the postgraduate sector, five Continuing Medical Education) that mentoring made an important contribution to the success of the portfolio, but a definition of this success was not clear.

The following evidence describes the influence of mentoring on the process of portfolio use, but less on how mentoring affects outcomes. Considering the initial uptake of portfolios among potential learners, Webb et al. (2006) found that compliance among surgeons in training increased from less than 50% to 100% when, as previously described, monthly feedback from a dedicated supervisor was introduced. Snadden and Thomas (1998), in a qualitative study of portfolio use among 44 trainees in general practice, reported that the portfolio was '*usually not adopted where there was no support from the trainer*' or where tensions existed in the trainee/trainer relationship. This was illustrated by means of a few case studies which did not explore possible confounding variables. Pearson and Heywood (2004) achieved a good response rate (77%) of registrars in a UK deanery when evaluating a pilot portfolio for 92 GPs. Authors reported that users with a

supportive trainer more commonly used their portfolio for reflection on their practice.

Few studies looked at the potential impact of mentoring on sustained portfolio use, but Snadden and Thomas (1998) found that the majority of their study group had stopped using the portfolio by months 6–8 of the training year, '*despite the intense effort to support portfolios in the region*'. In his study looking at uptake and subsequent level of use of electronic portfolios among cohorts of PRHOs, nursing students (under- and postgraduate) and sixth form school pupils, Murray found a relationship between the provision of feedback on the portfolio from a mentor and the frequency and level of use by the learner (grey literature, 2007). A comparison of 46 learners who received feedback with 22 who did not, showed that 57% of those who received feedback went on to become classified by the author as '*continuous users*' versus 0% of those who had not received feedback on their initial postings. However, it is not reported which of the cohorts these learners were from and this finding should be interpreted with caution as the terms of use and purpose of the portfolio were very different for each cohort. Likewise, the timescale of the project was unclear, so that the term 'continuous' does not give any indication of actual duration of sustained use.

There is some suggestion in the literature that, for some individuals, mentor support was needed for reflection. The assessors in a previously cited study of dentists (Maidment et al. 2006a, b), expressed this opinion, although the dentists themselves had mixed views. Tiwari and Tang (2003) made the observation that some of their learners (12 postgraduate nurses in Hong Kong) appeared to lack sufficient cognitive and reflective skills to make best use of the portfolio. They recommended that support be tailored according to need.

Users have also reported concerns regarding supervisors with insufficient knowledge or understanding of the portfolio. Ryland et al. (2006) conducted a pilot study into portfolio use amongst second year Foundation doctors (i.e. doctors in the first 2 years of postgraduate training) in the UK ($n = 147$) in 2005/2006. Using qualitative analysis of free text questionnaire responses (response rate: 65%), the article stated one of two emergent themes as educational supervisors '*needed more guidance on how to use the portfolio*'. Although the study was relatively simple, the deanery that conducted it used the evidence the basis for the roll-out of consequent supervisor training as they believed there was a '*continuing need to emphasise the educational value of the portfolio by both Foundation trainees and their educational supervisors*'. Hrisos et al. (2008) in a UK study noted that over half of Foundation trainees ($n = 182$) felt their educational supervisors ($n = 108$) were not sufficiently knowledgeable about the portfolio.

Lack of support was identified as a factor which was considered to limit the potential of the portfolio from a survey involving 121 nurses in the UK (Richardson 1998) and in another survey of 90 GP trainees (Kjaer et al. 2006). One outcome of focus groups conducted by Chabeli (2002) with 20 postgraduate nursing students in South Africa required to complete a portfolio for a semester for assessment purposes was, that they felt that teachers should, '*constantly monitor and provide support and guidance to the learners during the preparation and compilation of the portfolio*'. Similarly, Coffey's (2005) survey of nurses ($n = 22$) using a portfolio for assessment for a diploma in gerontological nursing in Ireland, found that respondents felt more support was needed in completing the portfolio. It was implied that mentoring should be the vehicle for this support. McMullan et al. (2003) concluded in their review of the use of portfolios in the assessment of learning and competence for nursing, that it was important for the tutor to provide regular support and feedback, '*as this helps them build their portfolio*', likewise, Bowers and Jinks (2004) reported (from a limited evidence base) that UK practitioners needed guidance and support.

Peer support

A small number of studies explored the influence of peer support on portfolio users. Mathers et al. (1999) conducted a crossover study comparing traditional and portfolio method of PGEA, and used a model of three facilitated meetings of groups of UK GPs ($n = 32$) compiling portfolios for PGEA purposes during a 6-month study period. Authors reported that this process provided a supportive stimulus to learning and was an appropriate use of time by the GPs. A survey conducted by Austin et al. (2005) of 1415 Canadian pharmacists highlighted the value of an information-sharing session, allowing participants to discuss experiences with colleagues in a facilitated environment. It was reported that after this session, the feedback from subjects indicated that they were, '*far more informed, aware and supportive of the portfolio concept*'. In Tiwari and Tang's small study of nursing students, portfolio users spontaneously developed collaborative learning strategies and gave each other support, apparently as a result of being involved in the portfolio process.

Time

Many authors reported time as a factor that had a negative influence on portfolio use by healthcare practitioners (e.g. Jensen & Saylor 1994; Keim et al. 2001; Dornan et al. 2002; Dagley & Berrington 2005; Maidment et al. 2006; Duque et al. 2006), as they had difficulty adding portfolio use to their already busy schedules. Kjaer et al. (2006) had doubts that the 10–15 minutes allocated protected time could be worked into the existing trainee/trainer interaction. In the GP PDP study, Cross and White (2004) reported 73% (of 204) respondents as 'disagreeing' or 'strongly disagreeing' they had enough protected or unprotected time to undertake their PDP. Seventy four percent of this group also 'agreed' or 'strongly agreed' that the PDP study competed with the time reserved for their socialising and family. No studies objectively tested the implication that time was a barrier to the practicality of portfolio usage.

Mathers et al. (1999) crossover study cited above, demonstrated that portfolios take a considerable and very varied amount of time (at least for new users). The time involved in preparing a portfolio for PGEA was 24.5 ± 12 h (range 10.5–64 h): much more than the fifteen hours which

could be claimed for the process. The implications of this additional time on the relative efficiency or '*amount of educational gain*' of traditional pattern of PGEA was discussed by authors as it does not allow for practical elements e.g. travel time to courses, preparation and follow up. Authors report that these issues make comparison with the portfolio approach more equivocal.

Keim et al. (2001) showed that dietetics professionals assigned to use a portfolio ($n=661$) conducted learning needs assessments significantly quicker than a control group ($n=714$) following the traditional route (2.7 ± 2.6 h vs. 4.4 ± 5.1 h, $p=0.002$). They were also quicker in developing learning plans (4.0 ± 4.9 h vs. 2.4 ± 1.9 h, $p=0.018$).

Cost

Although many studies allude to savings made by adopting portfolios (particularly electronic versions) such as reduced administration cost or printing, a single small study substantiated the claim. Moyer (2002) reported feedback from four of thirteen nurses who used a portfolio in the USA, and compared the traditional cost of nurse credentialing (>$40,000 per examination) with the cost of portfolio evaluation of the same content ($14,752). Among our retrieved articles there were none examining finance and its potential influence on individuals' portfolio use. However, note that we did not search specifically for economic articles or have cost-effectiveness as part of our inclusion criteria, therefore we do not draw further conclusions.

Use of portfolios for assessment. Twenty two articles reported on the use of portfolios around the assessment of healthcare professionals at work exploring the ways in which they have been used for formative or summative types of assessment, and exploring the boundaries of reliability and validity.

Reliability for summative assessment

Several articles reported on the reliability of using portfolio assessment for summative decisions about healthcare professionals – sometimes referred to as 'high stakes decisions'. Six articles examined by Driessen et al. (2007b) in their systematic review of portfolios in medical education, gave an 'average' reliability of 0.63, although the range of scores of the six studies cited was unclear. Increasing the number of raters raised the reliability towards a value of 0.8 as usually required for high stakes decisions. Also reported were a number of measures which had positive impact on inter-rater agreement, i.e. training, rater discussion, global criteria with rubrics. Lynch (2004), whose literature review included portfolio assessment as part of a wider focus on practice based learning for residents and physicians, and who cited similar articles to Driessen, reported a slightly more negative view. A key focus was on studies by Pitts et al. (2002) who looked at portfolio assessment with eight GP trainers. They achieved poor to moderate inter-rater reliability of 0.1–0.41 which increased to 0.5 with rater criteria discussion. McCready (2007) carried out a literature review on portfolios as a tool for assessing competence in nursing and also reported the literature as ambiguous with regard to reliability ($n=15$ included studies). She questioned whether conventional tests of reliability and validity can be brought to bear on the holistic data presented in portfolios (referring to Pitts et al. 2002). The literature review by McMullan et al. (2003) focussed on the use of portfolios in nursing and concluded that there were difficulties in assessing portfolios using purely quantitative methods.

Enhancing reliability

As already highlighted, Driessen et al. (2007b) reported some successful strategies to improve reliability; use of small groups of trained assessors and discussion amongst raters before (and sometimes after) the assessment. These findings were supported in McCready's literature review. Jasper and Fulton (2005), although reporting on the development of marking criteria for practice-based portfolios, tested their new criteria on 30 portfolios at two UK sites where Masters courses in nursing and healthcare disciplines were offered. They concluded that the use of double marking with an external examiner along with explicit descriptive criteria against which portfolio content could be judged, was the way forward.

Alternative strategies to improve reliability were raised by other authors. Melville et al. (2004) reported ratings of all paediatric Specialist Registrars' (SpRs) portfolios in one UK deanery ($n=76$). In the first year portfolios were assessed by a single rater, and the following year by two raters. They concluded that although their method of portfolio assessment could not be used as a single assessment method for high stake decisions, without multiple observers (assessors) or observations, it had a place as part of a triangulation process with other assessment methods. In two studies identified in McCready's review, tri-partite meetings during the portfolio assessment process were used. In the first study this tri-partite assessment was between the academic supervisor, practice mentor and subject (post-registration nurse). It reported the subjects as having valued this approach ($n=15$, 75% participants). The other article, although there was little detail provided, suggested the tri-partite meeting was crucial. Another study by Jarvis et al. (2004) looked at portfolio entries representing thirteen psychiatric skills from eighteen psychiatry residents in the USA. A total of 80 entries were examined in the light of the six ACGME (Accreditation Council for Graduate Medical Education) general competencies. They found five out of six competencies represented in the portfolio and similarly concluded that whilst it was desirable for a single evaluation method to assess competencies, it was reasonable and realistic to use more than one form of evaluation to examine performance. Maidment et al. (2006a) reported on a portfolio developed with a range of specific sections to meet dental professional body requirements with regard to providing evidence of fitness to practice. Based on the study sample of 10 general dental practitioners, they concluded that when using the portfolios for revalidation the scheme would be significantly enhanced

by using it as the basis for an appraisal interview, thus triangulating the data and its interpretation.

Validity for assessment of competence

The validity and reliability of portfolios assessment are often combined in the literature making it difficult to distil clear messages. There would however, seem to be tension between balancing both reliability and validity of portfolio assessment with learning.

On the positive side some studies found portfolio assessment valid for specific criteria. For example, in Mathers et al. (1999) comparison of traditional route to PGEA accreditation with a portfolio-based learning route for GPs, the breadth of topics covered in the portfolio was extremely wide and entries were seen to be appropriate for the claimed educational objectives. Jarvis et al. (2004) as described earlier, examined portfolio entries in the light of the six ACGME general competencies. Although all general competencies bar one were represented, they concluded that all the competencies could be covered with some revision of the portfolio guidelines. O'Sullivan (2004) tested the reliability and validity of eighteen psychiatry residents' portfolios in the USA. Scores were compared with another cognitive performance measure and global faculty ratings on clinical performance. Authors concluded that portfolios provided valid evidence of competency although the evidence was not strong.

Other authors expressed more uncertainty or concerns. Smith and Tillema (2001) looked at portfolio use in the Netherlands among different types of professionals and in different settings which included senior nurses (unit leaders, $n=26$) and nursing staff ($n=33$. Interviewees ($n=12$ unit leaders) highlighted the perception that the evidence found in the portfolios was considered to have questionable validity, especially when it is used for assessment and is no longer a working portfolio: *'if the evidence is original, who chooses it and what is the quality of the various portfolio entries?'* The literature review by Carraccio and Englander (2004) focused on portfolio assessment in medicine and reported the difficulty in striking a balance between the creative, reflective aspects of the portfolio which is learner focused with a structure that is reliable and valid. Finally the small scale pilot by Maidment et al. (2006a, b) found significant concerns about the use of a portfolio for revalidation to meet dental professional body requirements: *'revalidation* [using a portfolio] *doesn't prove you are a good or a safe dentist, it proves you can fill a book'*.

Linking portfolio to quality assessment frameworks

A small pilot study (Dagley & Berrington 2005) evaluated the way in which a portfolio was used by UK GPs ($n=5$). This included logging critical incidents and attempting to link revalidation categories to elements of their PDP and CPD actions. These links, however, were found by the authors to include some inconsistencies, and they proposed this area required further training. PDPs were quality assured against two published CPD frameworks: Rughiani's and Cromarty Eastern Deanery matrix.[5] They were found to have evidence of a continuous learning 'spiral' and to contain rich material. However audit, and the more objective elements were underused.

Compliance

It seems evident that when portfolios are required for summative assessment, compliance is greater. Driessen noted that if portfolios were not formally assessed, their use tailed off (based on Snadden & Thomas 1998; Pearson & Heywood 2004). Smith and Tillema (2001) inferred from user comments that because keeping a portfolio was not required, participants did not find time for it in their daily work. McMullan et al. also identified a study in which participants were less likely to use the portfolio if assessment was not present although no data on this was presented.

However, the point was also confirmed by Murray (grey literature, 2007) as previously noted in his study of implementing e-portfolios with mostly healthcare professionals in four colleges. He found that after training for all, only 23% ($n=171$) of users who were given the choice of using the portfolio (for others it was compulsory), actually used the system.

Formative assessment

Reviews by both McMullan et al. (2003) and Kjaer et al. (2006) found there was considerable support for portfolios to be used for formative assessment. Kjaer et al. carried out a study with a cohort of GP trainees and an on-line portfolio ($n=79$ portfolio users, 11 non users) and used two evaluation questionnaires (one for users, one for non users) which had been validated for construct and content validity and which collected both quantitative and qualitative data. They found that the portfolio was a good basis for formative assessment and recommended that a part of the portfolio should be kept exclusively for formative feedback. Although not distinguishing between formative and summative assessment, the article by Tiwari and Tang (2003) reported on the qualitative data collected through semi structured interviews with twelve of the study participants, selected according to criteria specified in the article. They found that portfolio assessment can have a positive effect on learning and users reported a distinct preference for the portfolio form of assessment over the standard approach (written assignment and end of term test). Webb et al. (2006) in a study with a cohort of surgical residents, concluded from the user survey (40 residents) that the most beneficial aspects of portfolios was the educational aspect e.g. the faculty interaction and feedback. Similarly a study by Coffey (2005) with 22 postgraduates from a nursing programme reported findings were mainly '*positive regarding the effect of the assessment on their learning'* and gave some quotes to back this finding. Finally, Smith and Tillema (2001) identified the importance of feedback provided by the portfolio regardless of whether it was formative or

summative – it gave an opportunity for subsequent improvement of actions.

Influence of assessment on portfolio contents

Driessen et al. reported (based on two studies by Driessen (2005) (not included in our review) and Mathers et al. (1999)) that there was no conflict between using the portfolio for summative assessment and learning in the postgraduate sector and that they can be successfully combined. However, there is some evidence to the contrary. McMullan, Endacott, Gray *et al* concluded in their literature review that portfolios become assessment led, resulting in a reduction in learning value. Three primarily qualitative studies also addressed the formative/assessment conflict. Snadden et al. (1996) through an action research project with 20 pairs of GP trainers and registrars, reported that participants perceived that formal assessment would inhibit the type of material collected in the portfolio, but it must be noted that these perceptions were not substantiated by any differences in portfolio content. In the latter part of the Webb et al. (2006) study, when 40 surgical residents (100%) complied with the use of the portfolio, only 20% felt that their portfolio should be used for resident assessment although no reasons were given. Kjaer et al. (2006) with 56 (71%) of portfolio users showed that GP trainees feared they would be less honest and avoid showing shortcomings, if their notes were used for assessment purposes. On a similar point Murray found that assessment impacted on the type of engagement displayed by the users: 55% of assessed users only submitted entries to the required sections compared to 41% who used it continuously.

Outcomes of portfolio use. Many articles alluded to outcomes of portfolio use, however, as will be discussed in more detail later, most of them failed to clearly or objectively demonstrate that self-reported or measurable effects are in fact due to portfolio usage. The following sections describe some evidence from seventeen articles which did attempt to demonstrate true outcomes.

Promoting reflection

The encouragement of reflection is a commonly cited purpose of a portfolio, and there is some evidence that this is facilitated by portfolio use. In one study, simply providing a portfolio appeared to have a positive impact on users' attitudes to completing activities that were previously unsupported by portfolios. Keim et al. (2001), randomly assigned dietetics professionals to either a portfolio group or a control group who reported Continuing Professional Educational activities in the traditional format (Cronbach's $\alpha > 0.75$). At the 2-year follow-up (79% response rate from 1082 surveys), a significantly greater proportion of the portfolio group (79% vs. 46%) reported that they had completed considerably more self-reflective entries in the previous 12 months ($p < 0.001$). A 5-year evaluation of portfolio use by six to ten surgeons per year (total $n = 40$) indicated that 72% of users felt the portfolio should be used for self-reflection (Webb et al. 2006). This contrasts to 42% of GP trainees in a study by Pearson and Heywood (2004), who actually reported using their portfolio for reflection; and 56% of educational supervisors who felt their trainees were encouraged to reflect by use of a portfolio (Hrisos et al. 2008).

Other authors have reported some adverse effects. Swallow et al. (2006) found some negative views among 25 UK community pharmacists, some of whom felt that the portfolio could actually inhibit reflection if there was a lack of confidence about how the information may be used '*against them*', a view echoed by Pearson and Heywood (2004). Austin et al. (2005), pointed out that some users already described themselves as being reflective, and believed that being forced to use a tool for this purpose interfered with their own approach to their professional development.

Some authors state that users can use portfolios to reflect but few describe how reflection is defined or measured making it difficult to determine whether it is a meaningful outcome or one which has a knock on effect on professional practice. Dagley and Berrington (2005) found that some records showed evidence of users completing a reflective cycle – this was shown by electronic links with recorded incidents from their practice, their PDPs and CPD activities. The two-part study by Maidment et al. (2006a, b) also reported on the potential for portfolios to support reflective practice. Among the ten participants however, there were reports that reflective practice took place regardless and therefore portfolios were an artificial and unnecessary imposition. The concept of the portfolio as a '*burden*' was also raised in Hrisos et al.'s (2008) study cited earlier, with two-thirds of the trainees reported that the collection of required paperwork was difficult to manage in busy hospital wards.

Learning/knowledge

Tiwari and Tang's (2003) controlled study is probably more usefully considered as a case study, as the two groups of users are at different stages of learning – with the control arm being undergraduate students following traditional assessment methods, whereas the group of interest to this review were postgraduate nurses using a portfolio. Ten of the 12 participants interviewed reported positive academic effects of the portfolio, including a deeper understanding of study topics, and the process of learning itself. The attitudes of users were cited as explanation though as the remaining two participants were reported to be '*only interested in getting a degree*'. Webb et al. found that 75% of users (30 out of the 40) felt that the portfolio had improved their understanding of a topic they were studying.

Coleman et al. (2006) conducted a controlled study in the USA using two cohorts of graduate multicultural counsellors ($n = 28$) who were assigned to use a portfolio or case formulation method to demonstrate their competence. The final exams were rated blind to group allocations, and showed a significant difference with the case formulation group rated higher than the portfolio group. The lack of detail on participant characteristics and randomisation procedure for the study however, makes this comparison somewhat unsafe.

There was a high inter-rater agreement (0.67–0.79) on the quality of portfolio contents.

Engagement with learning

Mathers et al. (1999) crossover study of GPs using traditional or portfolio PGEA methods undertook an experimental study design but presented the analysis in a qualitative narrative style, not taking into account any effect of the crossover itself on outcomes. It states that there was evidence of completion of a learning cycle by portfolio users who reported a mean number of seven (±4 SD) critical incidents which subsequently modified their learning objectives i.e. evidence that portfolio caused people to adopt principles into practice more than PGEA route. The method of analysis and reporting unfortunately mean it is not possible to determine when the effects happened in relation to the method being used by the user at that point; any lasting effect beyond the 6-month period on each approach and any effect of which came first.

Keim et al. (2001) showed that, compared to control, their portfolio group produced more learning needs assessments (71–22%, $p<0.001$), and more learning plans (70–12%, $p<0.001$). Overall though, measures such as attitude towards professional development, self-efficacy to conduct a learning needs assessment were reasonably positive at baseline but did not change significantly by 2-year follow up (paired *t*-tests, $p>0.05$) The perception that portfolio maintains competence was not rated positively by either group and again did not change significantly between baseline and follow up. In Mathers et al. (1999) portfolio users were found to tackle a much wider breadth of learning activities and study topics.

Fung et al. (2000) conducted a multi-centred non-randomised trial in Canadian obstetrics and gynaecology departments, giving an advanced year of exposure to a prototype portfolio (described in Walker et al. 1997) to residents at one school, and then a comparison of measures with three other schools as they embark on usage of the full internet-linked version. Compared to control, the residents using KOALA (Computerised Obstetrics and Gynaecology Automated Learning Analysis) reported increased awareness of their self-directed learning ($p<0.05$), were more inclined to learn on their own ($p<0.015$), had a positive attitude toward life-long learning ($p<0.000$) and expressed strong interest in taking on new learning ($p<0.018$). This well-cited study also reports the impact on their perceptions of their future learning. They felt a clinical experience portfolio would now contribute to their residency ($p<0.011$) and that didactic lectures would not be sufficient to support their future learning ($p<0.028$). This study was limited however, by a number of factors, including lack of information about group comparability at baseline, insufficient detail on the timing of data collection and the fact that the intervention consisted of a year's exposure to the portfolio's prototype. Although authors concluded that the internet-linked portfolio has positive effects, it may have been the advance year of the (non-internet linked) prototype which had these effects.

In Keim et al. (2001), both the portfolio and control groups demonstrated generally positive attitudes towards assessing learning needs and developing learning plans across the 2-year follow up: ratings showed no significant difference between groups (*t*-tests). Both groups were slightly less positive however that the portfolio maintained competence (scores around 52–54, on a scale where the midline is 55). Tiwari and Tang (2003) found that all 12 portfolio users reported a high level of satisfaction in using the portfolio, once the initial lack of confidence about the process was dealt with. They expressed pleasure in the freedom afforded by this method of assessment.

The evaluation ($n=147$) conducted by Ryland et al. (2006) concluded that a portfolio (used by UK Foundation doctors) did support educational processes; trainees reported the positively on the role of the portfolio in supporting assessments and enhancing reflective practice. The size and response rate of the study were relatively low, however, and the study was reported in brief.

Supporting learning into practice

Coffey (2005) evaluated a clinical learning portfolio for gerontological nursing by means of a postal survey of the programme's first graduates. The author reported an unexpected and tangible result in that respondents' use of the portfolio continued on to their subsequent clinical practice. However, the study had inherent weaknesses, including a small sample ($n=22$) of a single cohort, the survey instrument not being tested for reliability and validity, and there was no description of the qualitative analysis used. In Austin et al.' (2005) study of 1415 pharmacists using a portfolio in Canada; users completed a mean of 5.6 learning objectives per year (range of 0–10). Almost two thirds of self-identified learning objectives were achieved (63% ± 25%) which resulted in a mean of 2.2 changes to practice, facilitated by the portfolio. Campbell et al. (1996) found that two thirds of study participants ($n=152$ Canadian physicians) reported that portfolio use made them reflect on patient care, and to take note of which educational activities enhanced their expertise.

Are portfolios equally useful across health professions; can they be used to promote inter-disciplinary learning? No evidence was identified to allow us to answer this question – a small number of studies were found which included for example nursing and midwifery, or postgraduate and undergraduate medical students, but no sub-group analysis was conducted to allow understanding of the relative needs of the different groups or the different ways in which they engaged with the portfolio.

It is likely that this reflects the traditional divisions between the healthcare professions where each works independently from undergraduate level through to continuing professional development. Although some organisations are beginning to promote multi-disciplinary learning it may be some time before the commonalities between the professions are recorded in any standardised or comparable way.

What are the advantages and disadvantages in moving to an electronic format for portfolios? The team identified nineteen articles which provided evidence on this question. Note that as

electronic portfolios were of special interest to the review, wider inclusion criteria were adopted, to include undergraduate students and articles conducted in a non-healthcare setting.

The main messages extracted from the evidence were grouped under the following themes:

- factors influencing e-portfolio use
- outcomes of e-portfolio use.

Factors Influencing E-Portfolio Use

The electronic medium

One good quality study directly tested the effect of the electronic format on portfolio use. Driessen et al. (2007a) conducted a randomised trial of two types of portfolio format with year one medical students in Maastricht. Five of the 17 mentors were randomly selected to participate (all agreed) and the two groups of students each was responsible for, were randomly allocated to either paper ($n=47$) or web ($n=45$) based portfolio. Although the comparativeness of groups was not described, it is assumed that the (unspecified) randomisation procedure adequately minimised bias. Pairs of raters independently scored the portfolio content for quality of evidence and reflection (coefficients 0.71–0.91). The scores were very similar with the notable exception of the 'additional effort' of the web-based population with the perceived effort they applied to creating their portfolios. This manifested in more personal approaches to the look and content of students work. There was strong evidence that the medium of the portfolio influences the amount of time users are willing to spend with it. There was a moderate effect size of 0.46 indicating that the web group spent more time on developing their portfolio (15.4 vs. 12.2 h, $p=0.05$). Both groups were similarly satisfied with their portfolio. The article's discussion refuted the notion that extra time was required for the web versions, and hypothesised the electronic medium motivated the users to spend more time with the portfolio. There was unanimous agreement from mentors ($n=5$) that web-portfolios are easier to use as they allow faster retrieval of evidence through hyperlinks, and enabled access from a variety of sites at the mentor's convenience.

Chang (2001) conducted an evaluation of an electronic portfolio used by (an unspecified number of) undergraduate teachers assessing its functions and impact on students' educational progress. Most respondents felt it was beneficial to use the electronic medium to access others' e-portfolios. A finding was also described by Clegg et al. (2005). The vast majority (93%) of Chang's students believed they could improve the standard of their own work by having the option to view their peers'. Students found the feedback from peers more helpful than that of their instructors, which authors speculate may be due to higher expectations of instructors and demands on their time to provide extensive information. There was 80% agreement that using peers' portfolios enhances communication with those peers. The electronic medium therefore enabled sharing and exchange of information that would not be possible in paper format.

Fung et al. (2000), is an often quoted study cited as demonstrating the positive effect of the electronic medium, however as previously mentioned it appears that the comparisons made are between residents at one school exposed to a prototype e-portfolio for a year ahead of three other schools who all then used an internet-linked version of the same tool. The additional positive learning effects may therefore be attributable to the advanced exposure to the tool rather than the electronic medium.

Banister et al. (2006) highlighted the importance of piloting new e-portfolio systems, in their study which revealed that an in-house system was better suited to their purpose (teacher education in the USA) than a commercially available one. This is echoed by Scott and Howes (grey literature, 2007) who reported learning important lessons about improvements required to the interface of a new portfolio system, following a pilot with UK medical students.

Data transfer/accuracy across systems

A portfolio's ability to support an individual's life-long learning necessitates the transfer of the relevant records and information through one's educational and professional transitions. In theory, the electronic medium would be an ideal medium to ensure one could have continuous access to all relevant past items. In reality, this Horner et al. (grey literature 2007) in a series of case studies illustrated the difficulty in transporting data between different e-portfolio systems in further and higher education institutions across England. Concerns regarding the security or confidentiality of data contained within electronic portfolios emerge in many studies (e.g. Carney & Jay 2002).

Dorn and Sabol (2006) demonstrated in a multi-site before and after study conducted in the USA, that paired rating scores correlated well for artistic portfolios assessed in both paper and digital formats. Assessment scores for the digital portfolios were slightly higher than those on paper, but were a good predictor (significant at $\alpha=0.05$ level, confidence interval 0.96–1.03).

Users' IT experience/skill

Students' experience in information technology correlates positively with their perceptions of learning through an electronic medium and therefore, use of the portfolio model. Hauge (2006) measured this in their Norwegian interview of five student teachers and survey of 76 students ($\beta=0.38$, $p<0.05$). Dornan et al. (2003) conducted a qualitative case study which describes the evaluation of a web-based portfolio, demonstrating that students appreciated the design, for example, the ease of navigation.

Kjaer et al. (2006) developed and validated a questionnaire to evaluate the use of a new online portfolio by 90 Danish GP trainees (79 of whom had used the portfolio and 11 had not). The response rate was over 70% for both groups. Whilst two fifths of respondents (39%, $n=56$) stated that they would not have started using an e-portfolio if given the choice, after the study, 87% agreed that they preferred the electronic medium. With regard to post-study use, 50% agreed that they would

continue using the portfolio the same amount, and 46% expected to increase their use. Some portfolio users were wary of the perceived potential for external control of their learning. It was described as being more appropriate for formative than summative assessment, in that it could be used as a prompt for discussion points with a trainer.

Whilst the electronic medium requires support and training especially for those less familiar with the technology, any portfolio system would require this from an educational perspective. '*It is frustrating when the trainers are not completely familiar with the use of the portfolio. The time spent with the trainer should be used to discuss educational issues – not technical issues*' (Kjaer et al. 2006). Trainees noted that the hospital setting may make the use of an electronic portfolio problematic (with access to computers) unless a PDA version was available. Non-users of the portfolio related common responses to why they felt unable to use the portfolio including: lack of information; protected time and support from trainers; access to ICT and personal motivation.

Training/support for e-portfolios

The training and support that users receive was frequently cited as a factor that influences their uptake of portfolios. Redish et al. (2006) in their description of the migration of a paper to web based portfolio in a graduate education programme, exemplify what many articles relate by concluding, '*training for both faculty and students is critical to successful implementation and ongoing technical support should be given careful consideration*'. Unfortunately they do not substantiate this sentiment by linking it to research.

Similar to the other factors influencing portfolio use, training and support were not directly evaluated as an intervention in most studies. Duque et al. (2006) provide the single instance we found of evaluation of training against a control in this Canadian study of 133 medical trainees on a geriatric rotation, though they do not measure the training's influence directly against usage. The study evaluated students use of an e-portfolio divided into control (no training) and intervention (introductory hands-on session) groups, surveying both students and tutors. Students' comfort with the e-portfolio was surveyed immediately post rotation and at the conclusion of the clerkship year (response rates 98% and 55%). The first survey revealed 66% felt they 'strongly/somewhat' agreed they felt comfortable, compared to 48% of the control ($p < 0.05$). The survey at the end of the clerkship year found that the difference between the groups comfort levels had disappeared, following a significant increase in the control group and decrease in the training group (both $p < 0.04$) (final scores: 57% and 56%). Tutors in the Duque *et al.* study were surveyed once, and were asked to rate training as a limiting factor in use of the e-portfolio. None saw it as a strong limitation, 30% as moderate and 60% saw training as having no limitation on their e-portfolio use. Support was viewed in a largely similar way with the helpdesk availability seen as strongly limiting by 10%, moderately by 20% and of no limitation by 40%. From these results it would appear that most of these tutors did not regard training and support as significant factors influencing use, but the size of the sample ($n = 18$) and (critically) the fact the results were not measured against actual usage by the tutors, would call into question how much the tutors' results should be generalised.

Outcomes of e-portfolio use

Engagement with learning

The potential for the portfolio to capture the dynamic aspects of learning, particularly in relation to the student/tutor relationship was illustrated by Duque et al. (2006). Their case control study of 133 undergraduate medical students found that the e-portfolio was perceived to be a more effective feedback tool than more traditional methods ($p < 0.04$). These perceptions were given further weight by a demonstrable increase in the number of portfolio entries made by both students and tutors. Portfolio entries were only validated if they included comments and action plans, illustrating a quantifiable ongoing record of self-reflection with an average of 30 entries in 1 month. From this limited evidence they concluded that the inclusion of comments and action plans, and the engagement of both the student and the tutor in these evaluative entries showed that the portfolio was more than an information repository, but a dynamic account of learning, reflection and supervision.

Chang (2001) reported that a web-based portfolio was perceived to have had a positive impact on learning processes across a number of areas, with 47% of students 'strongly' agreeing and 42% agreeing. These positive findings were echoed by Bartlett and Sherry (2006) on their USA study of 34 undergraduate and postgraduate teaching students.

Learning into practice

The potential of the portfolio to bridge the perceived gap between the curriculum and the individual learner, or between teaching and practice, was examined by a number of studies including Avraamidou and Zembal-Saul (2003) and Jensen and Saylor (1994). In Jensen and Saylor's study (may be $n = 49$ but not clear) of physical therapy and nursing students in the USA. Students identified that the process of portfolio completion allowed them to structure their learning and reflection as well as place learning in the context of completed practice. The authors advised against measuring or assessing portfolios, stating that the aim of portfolios should be to inform, not to measure. They conclude that portfolios are '*more valuable for what they do than what they are*', suggesting (as Duque et al. 2006) that the very process of portfolio completion can be a learning experience but only with the support of mentors, tutors and the organisation as a whole. However, the evidence to support this conclusion was meagre.

Cotterill et al.'s (grey literature 2007) study of electronic portfolio implementation in two UK medical schools highlighted the potential contribution portfolios can make to organisational practice. They contrasted experiences in introducing portfolios to undergraduate medical students in two medical schools using questionnaire feedback from

around 500 students. Around 80% of students from one medical school thought that the portfolio was a useful learning experience, and as well as helping students plan and organise their learning there is some evidence that portfolio use prompted reflection (72% spent time reviewing what they had learned). However in the second medical school, only 39% reported that recording their learning helped them to think about the process of learning. The portfolio appeared to be perceived as somewhat separate to the 'real work' of the curriculum, indicating that perceptions of the role and purpose of the portfolio may affect the ability of students to engage fully in portfolio use to develop learning. Swallow et al. (2006) showed that portfolio use was beneficial in the planning and organisation of nine UK pharmacists' professional activities.

Discussion

This review has taken a broad and pragmatic look at all types of evidence regarding the effectiveness of portfolios across post-graduate healthcare education (and beyond for electronic formats). While it is important not to lose sight of common sense when attempting to evaluate an evidence-base with potential recommendations for decision makers or practitioners (Smith and Pell 2003), it is unavoidable to conclude that there remains a lack of objective examination of the effectiveness of portfolios. Although exploratory and uncontrolled investigations can be informative, there was a tendency towards reporting statements not backed up by evidence. The same unsubstantiated opinions of an author (or portfolio users and trainers) sometimes then repeated as fact in subsequent publications. This along with insufficient studies being conducted with due consideration of study size or sampling, failure to use an appropriate and clear intervention, no consideration or reporting of characteristics of participants and non-participants, make the body of evidence less than robust. With substantial funding going into widespread, and sometimes mandated, portfolio use, coupled with high expectations of what those portfolio systems can deliver, it would seem highly desirable that every opportunity be taken to properly investigate and test how portfolios are implemented, designed and supported allowing generalisable messages for other users and providers. Proportionate evaluation should be built in as a key feature of new portfolio projects, but research which generates generalisable findings will be of most value.

Portfolios: Practical instrument for education?

The evidence base contained many examples of portfolio in regular use by professional groups in the workplace across the healthcare and educational sectors. It was apparent that planned, supportive implementation of a portfolio was a vital step in enhancing its uptake and use by the target group. Evidence from successful implementations have incorporated buy-in at an organisational or faculty level, perhaps to create a purposeful and clear driving force as users begin to invest time in the portfolio.

There was good evidence to indicate that the support of a well-informed mentor can be a crucial factor in the uptake of portfolios. There was also evidence to suggest that it can influence the extent of portfolio use, particularly when specific regular feedback was provided. However, even when this kind of input was present, it was not always sufficient to ensure long-term sustained portfolio use. Competing demands on time often intervened and portfolio learners reported needing more support from faculty.

Other factors have been demonstrated to influence whether uptake and use of portfolios is achieved, including the characteristics, attitudes, experience and learning preferences of the users, however this evidence is less substantial in some cases e.g. gender of user. Many others are alluded to in the evidence base, but have not been objectively tested: including the availability and flexibility of users' time, access to computers, relevance and quality of the individual constituent parts of a portfolio. Unfortunately, there is no substantiated evidence that specifically examines portfolios' attributes (components, functions, linkages, core purposes) against how well that portfolio is used. Measuring a portfolio's use by altering the attributes and features that comprise it would be a comparatively simple task, and one that could be done retrospectively.

The status of the portfolio – voluntary or mandatory – is a crucial defining feature which directly influences user attitude, uptake, and the amount of time they are willing to spend on it. Therefore it should also influence the way in which evaluations or research should be interpreted. Clearly if professional registration hinges on its completion, users will put in the time required for this even to their own personal cost. However, they are likely to report concerns about use of their data, its security and suspicions regarding the purpose of monitoring. There is evidence that users may be simultaneously cynical about the purpose of a portfolio, but positive about its potential to them individually – this conflicting feeling by users has to be managed. Unless compulsory or an embedded part of the organisation's ethos, there is likely to be an uphill struggle to achieve compliance.

Portfolios: effective instrument for education?

If well implemented, portfolios have been demonstrated to effectively further both personal and professional learning in a number of ways. There was evidence of increased responsibility for learning: i.e. portfolio users have been shown to be less passive about their own learning needs and plans for future learning (but without a baseline measure in most studies, this assertion is not robust). There is overall agreement that portfolios aid learning processes and outcomes. There are mixed views of whether portfolios aid or hinder reflection, with evidence on both sides – this may come down to the individual's learning preferences, or some aspect of the portfolio itself. Although some authors suggest that a mentor may be beneficial to support reflection, this hypothesis has not been directly tested. A small number of studies describe users' views of the benefits of peer support. These include a more positive attitude towards portfolios and as a stimulus for learning. But in virtually all studies a substantial minority of users fail to engage with the portfolio. No studies were found which thoroughly investigated reasons for non-compliance or

resistance to portfolio use. Future research work on portfolios would benefit from taking these (and other) important confounding variables into account, and may allow refinement of successful portfolios already in use.

The outcomes occurring as a result of portfolio use are a direct way of assessing their effectiveness. However few articles were found which tested a meaningful control between, or within groups of users, or looked at a comparison intervention in order to reliably reveal outcomes of portfolio use. Many were cross-sectional or case studies of one particular portfolio, evaluating users and/or supervisors' feelings and experiences of the portfolio or the supporting processes after a fixed period of use. These articles were an often-quoted source of beneficial effects or positive reports of portfolio use in the literature. While looking at these provided an insight into the range of ways portfolios are used, and how successful they were individually, the generalisable messages were limited. These 'snapshots' of portfolio use failed to measure baseline characteristics of users (or give any indication of characteristics of non-users), meaning that positive or negative outcomes were impossible to attribute confidently to the portfolio. Few made attempts to identify confounding variables and incorporate them into the presentation of results e.g. the level of experience with portfolio or self-directed learning, ability to use or access appropriate technology, attitudes to learning, learning style – which were all alluded to as reasons why a portfolio was or was not successful.

Portfolios for assessment?

The meaningfulness of attempts to rate portfolios have been questioned in the literature, and there remains a lack of evidence in terms of inter-rater reliability. There was wide variation in published studies on the level of reliability of portfolios for summative assessment (principally conducted in medicine). It is clear that reliability increases with more raters or discussion between raters, but this incorporates additional time/cost, and it is unclear what size or direction of impact this would have on the ultimate scores. Evidence from both medicine and dentistry described the importance, to both practitioners and assessors, of triangulating portfolio data with other assessment methods.

Quantifying portfolio content and use may be too simplistic to capture professional learning and engagement and some authors reported that portfolios should not be used for summative judgements but instead for more qualitative and less structured personal development. It may be that more structured portfolios can and should be assessed, particularly for students and newly qualified professionals. However, as individuals progress through their career, qualitative methods of judging the portfolio may be more appropriate to allow the less tangible learning outcomes such as professional values and judgements to be captured. This depends on the type of portfolio, and attempting to generalise from a range of types may be unhelpful. These however lose the potential for individualised features which allow users to focus on developing their own needs and learning.

There was more positive, but weaker quality, evidence that portfolios are effective and useful for formative assessment. However, to date this mainly comes from a theoretical understanding of the potential analysis of the information obtained within portfolios, rather than objective tests that this process works well or is meaningful.

Advantages and disadvantages of the electronic format?

By definition a portfolio in the electronic medium offers the advantage of additional flexibility in a number of ways. This included flexibility of access to the information for users and supervisors, and virtually unlimited potential variation in content. This appeared to inspire or motivate users: good quality evidence was found to show that electronic portfolio users were willing to spend longer on it than those using paper-based portfolios, although ultimately self-reported satisfaction was similar between the two groups. A longer term analysis of these groups may be interesting to determine if the additional time spent provides a benefit. Ready access to peer's portfolio work was rated by some users as a particular advantage. We found a small amount of good evidence that electronic portfolios were more effective than a direct comparator in paper format both as a feedback mechanism, and for encouraging reflection in users.

An electronic portfolio may be readily linked to competency or quality assurance frameworks, or to users' PDP/CPD activities. These links can be automated and updated far more simply in the electronic format. Such links, however, particularly with mandated portfolios and those used for sensitive assessments or high stakes decisions may trigger security concerns.

Many authors cite training as being important when implementing an electronic portfolio, and this is likely to be a requirement when implementing an electronic portfolio system, as there was evidence that users' technical ability and knowledge significantly affect how they interacted with it. Technophobia remained an issue for many users, and if portfolio content is to be assessed, users must be adequately equipped to enter appropriate information, and not disadvantaged by their lack of confidence. However few have investigated this: e.g. the frequency, duration, format or content of training, to identify the key elements. The provision of technical support should be distinct from education support to contend with such issues.

There was reasonable evidence that moving from paper to electronic can be done accurately and that assessments of the same material in both formats are well correlated. The transferability of data between e-portfolio systems (required to facilitate life-long learning) is tentatively successful at the moment with some pilot projects now published but the process is far from straightforward.

True (and safe) interoperability has to be achieved before the full potential of e-portfolios to support lifelong learning is realised. Nevertheless the evidence indicated that progress was being made towards the realisation of

standards that will sustain the transfer of data between e-portfolios.

Strengths of our review. This BEME systematic review was based on a broad, sensitive search including all healthcare professional settings. All available articles were read, blind by two team members; non-English language articles were translated and a thorough grey literature search was undertaken. Good internal consistency was achieved for quality scoring and critical appraisal.

Limitations of our review. The time-consuming systematic processes involved in undertaking this review have been challenging for the team, and subsequently another (albeit with a narrower, medical focus) has published before us, therefore some of this evidence has recently been evaluated.

While the online data entry form was extremely valuable and has been of interest to other review groups, it may be worth noting that considerably more time would have been required, in collaboration with the programmer, to develop it into a fully functional and user-friendly system.

Future research. There are many gaps in the evidence, much of which appears to have been produced as a result of short-term local projects, e.g. rapid evaluations on specific portfolio projects. Several areas of research are urgently required to provide generalisable evidence:

- identifying genuine outcomes of portfolio use;
- identifying confounding variables underlying the variation in portfolio use among different learner types and professional groups;
- identifying the types of portfolio which are appropriate for the range of purposes they may be employed for: summative/formative assessment; creative/self-directed learning;
- assessing the cost-effectiveness of different approaches to portfolio implementation and the necessary support mechanisms;
- determining the differences in the effectiveness of portfolio across the professions, and revealing how they can be used to support education between the professions.

Portfolios are increasingly expected to support education and training and many organisations, professional bodies and academic institutions are investing significant resource (financial and time) in introducing them to students, trainees and staff. Given the lack of high quality evidence, and gaps identified above, this may be premature. The ambitious and ever changing expectations attached to portfolios, particularly electronic portfolios, may risk losing sight of the fundamental purpose of the educational environment which portfolios were introduced to support. The portfolio should perhaps be seen as a 'tool' to support education, not an educational instrument in itself. It is likely that the most appropriate portfolio to support summative assessment is different in nature and function to that best suited to self-directed learning. Anecdotal evidence may be useful to organisations selecting a portfolio to use, but a solid evidence-base relating to effectiveness, confounding variables, costs and outcomes would better support such decisions.

Back in 1994 Jensen and Saylor stated '*we believe portfolios should be a recognized legitimate aspect of a course or program, not a busywork activity*' – a sentiment that has been echoed consistently by many authors since – both the belief in portfolios, and the concept of embedding them in study or work. It would appear that with substantial and sustained commitment at all levels when implementing a portfolio (organisational, faculty, mentor/peer/supervisor *and* user) it can facilitate a range of learning and work-based development.

Conclusions

There is extensive and rapidly expanding evidence relating to the use of portfolios for assessment and education in post-graduate medical education. The heterogeneity of design and data, as well as questions around quality, makes formal synthesis impossible. However, it is possible to draw a number of conclusions:

(1) It is key that portfolio implementation is well designed and sustained, with high-level organisational support, to ensure uptake.
(2) A well-informed mentor can have considerable impact on uptake, especially when regular feedback is given.
(3) Users can be simultaneously sceptical about a portfolio's intended purposes and appreciative of what it can deliver for them personally.
(4) There is agreement, and some evidence, that portfolio users feel increased responsibility for their learning.
(5) Summative assessment of portfolio contents can be reliable among multiple raters, but triangulation with other sources is desirable.
(6) Electronic portfolios have been demonstrated to have a number of benefits (flexibility of access and content, potential for links). Users spend longer with electronic versions.
(7) E-portfolios are more effective for feedback and encouraging reflection than paper ones, though assessments in both are well correlated.

Acknowledgements

We thank Tim Brown for development and ongoing maintenance of the online data entry system throughout the review process; David McColl for support in use of Access database to retrieve data; Susan Morrow, who was part of the review team until end of initial filter of literature and Paola Solar & Thierry Boucheny for translation. The review was supported entirely by NHS Education for Scotland. All of the team are employed by NES, and additional funds for dissemination and group meetings were funded by the organisation's Educational Research Governance Committee.

Conflicts of Interest

NES is responsible for the development of a number of e-portfolios for trainee health professionals and two group members manage these projects. One other member monitors the use of e-portfolios by Foundation doctors (doctors in the first 2 years of postgraduate training).

Notes

1. www.nmc-uk.org (formerly the United Kingdom Central Council for Nursing, Midwifery and Health Visiting (UKCC))
2. http://www.foundationprogramme.nhs.uk/pages/home/training-and-assessment
3. There is some debate over whether reflection/self-awareness should be considered as 'skills'.
4. http://www.facs.org/education/technicalskills/kirkpatrick/kirkpatrick.html
5. www.londondeanery.ac.uk/. . ./files/primary-care-development/eastern-deanery-the-quality-of-a-pdp.pdf

Notes on contributors

[redacted].

References

Austin Z, Marini A, Desroches B. 2005. Use of a learning portfolio for continuous professional development: A study of pharmacists in Ontario (Canada). Pharm Educ 5:175–181.

Avraamidou L, Zembal-Saul C. 2003. Exploring the influence of web-based portfolio development on learning to teach elementary science. J Technol Teach Educ 11:415–442.

Banister S, Vannatta RA, Ross C. 2006. Testing electronic portfolio systems in a teacher education: finding the right fit. Action Teach Educ 27:81–90.

Bartlett A, Sherry AC. 2006. Two views of electronic portfolios in teacher education: non-technology undergraduates and technology graduate students. Int J Instruct Media 33:245–253.

Bowers SJ, Jinks AM. 2004. Professional issues. Issues surrounding professional portfolio development for nurses. Brit J Nurs 13:155–159.

Buckley SG, Coleman J, Davison I, Khan KS, Zamora J, Malick S, Morley D, Pollard D, Ashcroft T, Popovic C, et al. 2009. The educational effects of portfolios on undergraduate student learning: a Best Evidence Medical Education (BEME) systematic review. Med Teach 31(4):340–355.

Campbell CM, Parboosingh JT, Gondocz ST, Babitskaya G, Lindsay E, De Guzman RC, Klein LM. 1996. Study of physicians' use of a software program to create a portfolio of their self-directed learning. Acad Med 71((10 Sup):S49–S51.

Carney JM, Jay JK. 2002. Translating theory into practice: the dilemmas of teacher portfolios, Annual Meeting of the Southwest Educational Research Association.

Carraccio C, Englander R. 2004. Evaluating competence using a portfolio: a literature review and web-based application to the ACGME competencies. Teach Learn Med 16(4):381–387.

Chabeli MM. 2002. Portfolio assessment and evaluation: implications and guidelines for clinical nursing education. Curationis: S Afr J Nursing 25:4–9.

Challis M. 1999. AMEE Medical Education Guide No. 11 (revised): portfolio-based learning and assessment in medical education. Med Teach 21:370–386.

Chang C.-C. 2001. A study on the evaluation and effectiveness analysis of web-based learning portfolio (WBLP). Brit J Educ Technol 32:435–458.

Clegg S, Hudson A, Mitchell A. 2005. The personal created through dialogue: enhancing possibilities through the use of new media. ALT-J Res Learn Technol 13:3–15.

Coffey A. 2005. The clinical learning portfolio: a practice development experience in gerontological nursing. J Clin Nurs 14:75–83.

Coleman HLK, Morris D, Norton RA. 2006. Developing multicultural counselling competence through the use of portfolios. J Multicultural Counsel Dev 34:27–37.

Cotterill S, Aiton JF, Bradley PM, Hammond GR, McDonald AM, Struthers J, Whiten S. 2007. A flexible component-based e-portfolio. Embedding in the curriculum expert referral.

Cross M, White P. 2004. Personal development plans: the Wessex experience. Educ Prim Care 15(2):205–212.

Dagley V, Berrington B. 2005. Learning from an evaluation of an electronic portfolio to support general practitioners' personal development planning, appraisal and revalidation. Educ Prim Care 16:567–574.

Dorn CM, Sabol FR. 2006. The effectiveness and use of digital portfolios for the assessment of art performances in selected secondary schools. Stud Art Educ: J Issues Res Art Educ 47:344–362.

Dornan T, Carroll C, Parboosingh J. 2002. An electronic learning portfolio for reflective continuing professional development. Med Educ 36:767–769.

Dornan T, Maredia N, Hosie L, Lee C, Stopford A. 2003. A web-based presentation of an undergraduate clinical skills curriculum. Med Educ 37:500–508.

Driessen EW, Muijtjens AM, Van Tartwijk J, van der Vleuten CP. 2007a. Web- or paper-based portfolios: is there a difference? Med Educ 41:1067–1073.

Driessen E, Van Tartwijk J, van der Vleuten C, Wass V. 2007b. Portfolios in medical education: why do they meet with mixed success? A systematic review. Med Educ 41:1224–1233.

Duque G, Finkelstein A, Roberts A, Tabatabai D, Gold SL, Winer LR, Members of the Division of Geriatric Medicine M.U. 2006. Learning while evaluating: The use of an electronic evaluation portfolio in a geriatric medicine clerkship, BMC Medl Educ 6: 1–7.

Fung KFM, Walker M, Fung KFK, Temple L, Lajoie F, Bellemare G, Bryson P. 2000. An Internet-based learning portfolio in resident education: the KOALA-super(TM) multicentre programme. Med Educ 34:474–479.

Harden RM, Grant J, Buckley G, Hart IR. 1999. Best evidence medical education – BEME Guide No 1. Med Teach 21:553–562.

Hauge TE. 2006. Portfolios and ICT as means of professional learning in teacher education. Stud Educ Eval 32:23–36.

Horner AP, Cotterill SJ, Ingraham B, Thompson J, Gill S, Ayestaran H, Webster D, Ollerenshaw B, McDonald AM, Taylor L, et al. 2007. epics – Outcomes of a regional eportfolio initiative to support life-long learning (October 2007 grey literature).

Hrisos S, Illing J, Burk J. 2008. Portfolio learning for foundation doctors: early feedback on its use in the clinical workplace. Med Educ 42:214–223.

Hutchinson L. 1999. Evaluating and researching the effectiveness of educational interventions. BMJ 318:1267–1269.

Jarvis RM, O'Sullivan PS, McClain T, Clardy JA. 2004. Can one portfolio measure the six ACGME general competencies? Acad Psychiatry 28:190–196.

Jasper MA, Fulton J. 2005. Marking criteria for assessing practice-based portfolios at masters' level. Nurse Educ Today 25:377–389.

Jensen GM, Saylor C. 1994. Portfolios and professional development in the health professions. Eval Health Prof 17:344–357.

Keim KS, Gates GE, Johnson CA. 2001. Dietetics professionals have a positive perception of professional development. J Am Diet Assoc 101:820–824.

Kjaer N, Maagaard R, Wied S. 2006. Using an online portfolio in postgraduate training. Med Teach 28:708–712.

Lynch DC. 2004. Assessing practice-based learning and improvement. Teach Learn Med 16:85–92.

Maidment YG, Rennie JS, Thomas M. 2006a. Revalidation of general dental practitioners in Scotland: the results of a pilot study. Part 1 – Feasibility of operation. Brit Dental Journal 7:399–402.

Maidment YG, Rennie JS, Thomas M. 2006b. 'Revalidation of general dental practitioners in Scotland: the results of a pilot study. Part 2 – acceptability to practitioners'. British Dent J 8:455–458.

Mathers NJ, Challis MC, Howe AC, Field NJ. 1999. Portfolios in continuing medical education – Effective and efficient? Med Educ 33:521–530.

McCready T. 2007. Portfolios and the assessment of competence in nursing: a literature review. Int J Nurs Stud 44:143–151.

McMullan M, Endacott R, Gray MA, Jasper M, Miller CML, Scholes J, Webb C. 2003. Portfolios and assessment of competence: a review of the literature. J Adv Nurs 41:283–294.

Melville C, Rees M, Brookfield D, Anderson J. 2004. Portfolios for assessment of paediatric specialist registrars. Med Educ 38:1117–1125.

Moyer JE. 2002. The APNG(c): a preliminary look at credentialing nurses through portfolio review. Newborn Infant Nurs Rev 2:254–258.

Murray C. 2007. Is anybody out there? Developing the skills for lifelong learning: towards a model for engagement in the e-portfolio process. (grey literature).

O'Sullivan PS. 2004. Demonstration of portfolios to assess competency of residents. Adv Health Sci Educ 9:309–323.

Pearson DJ, Heywood P. 2004. Portfolio use in general practice vocational training: a survey of GP registrars. Med Educ 38:87–95.

Pitts J, Coles C, Thomas P, Smith F. 2002. Enhancing reliability in portfolio assessment: discussions between assessors. Medl Teach 24:197–201.

Redish T, Webb L, Jiang B. 2006. Design and implementation of a web-based portfolio for aspiring educational leaders: a comprehensive, evidence-based model. J Educ Technol Syst 34:283–295.

Richardson A. 1998. Personal professional profiles. Nurs Stand 12:35–40.

Ryland I, Brown J, O'Brien M, Graham D, Gillies R, Chapman T, Shaw N. 2006. The portfolio: how was it for you? Views of F2 doctors from the Mersey Deanery Foundation Pilot. Clin Med 6:378–380.

Scott H, Howes A. 2007. Using eportfolios to assess the reflective capabilities of medical students, http://www.eportfolios.ac.uk/docs/fdtl4_docs/leedsPaperFinal.rtf (grey literature October 2007).

Smith GCS, Pell JP. 2003. Parachute use to prevent death and major trauma related to gravitational challenge: systematic review of randomised controlled trials. BMJ 327:1459–1461.

Smith K, Tillema H. 2001. Long-term influences of portfolios on professional development. Scand J Educ Res 45:183–202.

Snadden D, Thomas ML. 1998. Portfolio learning in general practice vocational training – Does it work? Med Educ 32:401–406.

Snadden D, Thomas ML, Griffin EM, Hudson H. 1996. Portfolio-based learning and general practice vocational training. Med Educ 30:148–152.

Swallow V, Clarke C, Iles S, Harden J. 2006. Work based, lifelong learning through professional portfolios: challenge or reward? Pharm Educ 6:77–89.

Tillema H, Smith K. 2007. Portfolio appraisal: in search of criteria. Teach Teach Educ: Int J Res Stud 23:442–456.

Tiwari A, Tang C. 2003. From process to outcome: the effect of portfolio assessment on student learning. (Qualitative and quantitative research in Hong Kong). Nurse Educ Today 23:269–277.

Walker M, Fung KFM, Ash K. 1997. Advances in resident education: the introduction of a computerized learning portfolio – KOALA. J Soc Obstet Gynaecol Can 19:992–997.

Webb TP, Aprahamian C, Weigelt JA, Brasel KJ. 2006. The surgical learning and instructional portfolio (SLIP) as a self-assessment educational tool demonstrating practice-based learning. Curr Surg 63:444–447.