Measuring clinical practice

Rakhee Shah1,2, David Edgar2 and Bruce J. W. Evans1,2

1The Neville Chappell Research Clinic, The Institute of Optometry, 56-62 Newington Causeway, London SE1 6DS and 2Henry Wellcome Laboratories for Vision Sciences, Department of Optometry and Visual Science, City University, Northampton Square, London, UK

Abstract
The goals of this paper are to review techniques for measuring clinical practice within healthcare professions and to discuss possible applications of these techniques to primary care optometry. A review of the literature suggests a lack of systematic research investigating standards of clinical practice within optometry. It is argued that evidence-based research to determine the content of typical optometric eye examinations would be valuable for several reasons: to evaluate the service provided to the public by the profession; setting priorities and assessing the outcomes of continuous education and training; to influence governmental and professional policy decisions; National Health Service General Ophthalmic Services issues; the equitable management of clinicolegal matters and consumer complaints; setting appropriate professional guidelines and developing undergraduate training. Evidence-based studies within other healthcare professions have evaluated the content of clinical consultations. The literature reviewed reveals three main approaches: (1) abstraction of medical records, (2) use of clinical vignettes and (3) use of standardized patients (SPs) who present unannounced to clinics. In this review, we compare and contrast the use of these different methods in assessing the content of clinical consultations. It is clear from the literature reviewed that the use of SPs is the ‘gold standard’ methodology. Clinical vignettes can also provide useful data, especially if computerized.

Keywords: clinical practice, optometry, primary eyecare, quality of care, standardized patient, standards

Introduction
There are about 9200 optometrists in the UK [National Health Service (NHS), 2006], which is more than all the other eyecare professions added together. Nearly all primary eyecare in the UK is practised by optometrists. The scope of optometry is wider now than ever before, with a growing number of specialities within optometry and with optometrists playing an important role in many secondary and tertiary care hospitals. However, still the overwhelming majority of optometrists are engaged in routine primary eyecare examinations.

The fact that there is a fairly standard item of service that optometrists provide means that, compared with other healthcare professions, it should be relatively straightforward to determine the typical standard of care and range of standard of care within the optometric profession. For example, when presented with a patient aged 45 years who attends for their first eye examination, what are the contents of a typical eye examination? What proportion of optometrists would test visual fields? How many would undertake tonometry, test ocular motility, test pupil reactions, or even carry out ophthalmoscopy? Optometric practice is legislated in the Opticians Act (1989) (Office of Public Sector Information, 1989), and professional guidelines are also provided within optometry by the General Optical Council (GOC) (2005) and College of Optometrists (2006) (Table 1). These are clearly valuable as they provide a plan for standards of professional practice. Various funding bodies also provide additional regulations or terms of service (Association of Optometrists, 2006a), most notably within the
Table 1. Optometric practice is legislated in the Opticians Act (1989), professional guidelines are provided by the General Optical Council (GOC) and the College of Optometrists, and regulations are issued by the National Health Service (NHS) [General Ophthalmic Services (GOS)]. This table, while not a complete list of these documents, concentrates on those that affect the standards of professional practice relevant to this review.

Opticians Act 1989

The statutory duties imposed by the Opticians Act 1989 (Office of Public Sector Information, 1989) include categories of persons who can carry out eye examinations and fit contact lenses. The Opticians Act also stipulates duties to be fulfilled when examining a patient's eyes (College of Optometrists, 2006). The regulatory background to the eye examination (whether performed privately or under the GOS) contained in the Sight Testing (Examination and Prescription) (No. 2) Regulations 1989 (GOC, 1989) states:

When a doctor or optician tests the sight of another person, it shall be his duty

a) to perform, for the purpose of detecting signs of injury, disease or abnormality in the eye or elsewhere
   1. an examination of the external surface of the eye and its immediate vicinity,
   2. an intra-ocular examination, either by means of an ophthalmoscope or by such other means as the doctor or optician considers appropriate,
   3. such additional examinations as appear to the doctor or optician to be clinically necessary and
b) immediately following the test to give the patient a written statement
   1. that he has carried out the examinations required by sub-paragraph (a) of this section, and
   2. that he is or (as the case may be) is not referring him to a registered medical practitioner

The Act also requires that the statement should say if the patient is being referred to a registered medical practitioner and if s/he is being referred, the reason for referral.

General Optical Council

The GOC Code of Conduct concentrates on general professional conduct and does not give details of clinical standards. It states, ‘The GOC recognizes that other bodies have issued detailed guidance with regard to the matters covered in this Code. Practitioners are therefore expected to be familiar with the relevant guidance and advice issued by other organizations and, in particular, that of the professional and representative bodies. Reference may be made by the GOC to the guidance and advice of other bodies in the exercise of its functions.’ This is generally taken as referring to the College of Optometrists’ Code of Ethics and Guidelines (see below).

College of Optometrists

The guidelines and advice start with the general statement: ‘The optometrist has a duty to place the welfare of his/her patient before all other considerations, to apply to each patient the full extent of his/her knowledge and skill, and to maintain and develop his/her professional competence throughout his/her life.’

The full Code of Ethics and Guidelines for Professional Conduct is very detailed (College of Optometrists, 2006), specifying general ethics (e.g. patient practitioner relationships) and detailing the types of clinical tests that may be appropriate for specific categories of patients. However, it is stressed that the guidance document represents the College's view of good practice, this being defined by the College Council as being ‘what a competent optometrist is able to do in practical and achievable terms and within existing training and skills’. This implies the optometrist has a duty to carry out whatever tests are necessary to determine the patient's need for vision care as to both sight and health. It is not a set of instructions and does not constitute a ‘check list’ of clinical or professional procedures that must be carried out. The content is to be determined by both the practitioner's professional judgement and the minimum legal requirements (College of Optometrists, 2006).

NHS (General Ophthalmic Services; GOS) and other funding bodies

Optometrists carrying out NHS sight tests are bound by the NHS (GOS) regulations in addition to the above (Association of Optometrists, 2006a). These allow for a greater scope of practice within Scotland than the rest of the UK. In 2005–2006, 12 million of the 17.5 million primary care eye examinations were paid for by the NHS (2006).

Optometrists providing eyecare that is paid for by other funding bodies [e.g. by Primary Care Trusts (PCTs) as part of a co-management (shared-care) scheme or vocational eyecare by corporate organizations] are likely to be bound by other service contracts or agreements.

NHS, where regional variations now exist. A recent government review of the regulation of non-medical healthcare professionals has been welcomed by the GOC and this review discusses effective regulation of healthcare staff (Department of Health, 2006). However, none of these documents tell us what actually happens inside the optometric consulting room. There have been attempts to gain an insight into the clinical activities of optometrists through questionnaires (O'Leary and Evans, 2003), most notably those administered by the College of Optometrists (Stevenson, 1999). These are useful, but it is probable that there will be a sampling bias as conscientious practitioners are more likely to respond. Additionally, there is a further source of bias with human nature resulting in replies which indicate higher standards of practise than may actually pertain.

Another method of gaining an insight into standards within optometric practice is to study practitioners’ clinical records. This approach is routinely used in clinicolegal cases (e.g. GOC disciplinary hearings, civil litigation), but is subject to a number of problems. These are discussed in detail later in this paper and include errors of over- and of under-reporting of clinical tests. Furthermore, clinical records often give very little information about how thoroughly and appropriately a test was carried out.

This issue of determining standards of clinical care is common to all the healthcare professions. Assessing quality of care by health outcome measures (e.g. number of cases diagnosed or referred) is very limited and processes that measure quality of care are increasingly being used (Peabody et al., 2000), as detailed below.
Clinical competence in primary eyecare (optometry)

As clinicians we could describe clinical competence as ‘the degree to which a clinician can use their associated knowledge, aptitude, attitude and good judgement in the course of their professional practise and be able to work in an effective way in all situations that correspond to their field of practice’ (Miller, 1990). The different levels of clinical competence can be illustrated in a simple and elegant conceptual model, Miller’s pyramid (Figure 1).

The base of the pyramid consists of factual knowledge, such as that learnt during lectures in undergraduate training or lectures on continuing education and training (CET) courses.

One level up, Miller describes the ability to use knowledge in a particular context as ‘knows how’. This comes close to clinical reasoning and problem solving and might be assessed, for example, by the type of examination where the practitioner is given a clinical scenario and asked to write down which procedures they would carry out. At a higher level, ‘shows how’ reflects the person’s ability to act appropriately in a practical situation and describes hands-on behaviour in a simulated or practice situation. For optometrists, this is tested in the final assessment at the end of the pre-registration period (PRP) and usually never again in that practitioner’s professional career.

The ‘does’ level refers to actual performance in habitual practice. As can be seen, the higher the skills being tested in the pyramid, the more clinically authentic the assessment needs to be. The ‘action’ component of professional behaviour is the most difficult to measure reliably and accurately (Miller, 1990). Our literature review (summarized below) indicates a lack of systematic research that aims to investigate the upper level of the pyramid within optometry.

Why do we need to measure clinical practice?

Valid measures of the practice of clinicians are the basis of efforts to improve quality of care (Peabody et al., 2000): ‘practice can be improved, but only if it is measured’ (Peabody et al., 2004). There are several key reasons why we feel that the standard of optometric care in the UK needs to be determined and these are summarized below.

To evaluate the service provided to the public by the optometric profession

We feel that it is valuable for the profession of optometry to obtain objective data on the service that the profession provides to the public. The profession needs to guarantee that the procedure of an eye examination will identify any ocular abnormality; will use resources appropriately to identify ocular and systemic health problems (using tests which have adequate sensitivity and specificity); will result in the prescribing of functional corrections for defects of sight; will determine the need for remedial eye exercises where appropriate; and will provide advice to the patient on all aspects of visual efficiency. This guarantee is important because it allows patients to acknowledge the service they receive and to have confidence in the profession. Objective data on the eye examination will help to demonstrate the profession’s commitment to promote high standards and to ameliorate low standards. An eye examination is an important health check (RNIB, 2006), yet many members of the public may still not realize this, and research of this nature will help to demonstrate this point.

Determining priorities for continuing education and training

The GOC introduced compulsory CET for optometrists in 2005 to encourage high professional standards (GOC, 2004). The NHS makes a contribution towards the cost of this CET. Knowledge of the strengths and weaknesses of contemporary clinical practice will help to determine priorities for future CET.

Evaluating outcomes of continuing education and training

The GOC states that the purpose of optometric CET is to maintain high standards of professional knowledge and skills (GOC, 2004). A related question is whether CET can go further and bring about an improvement in standards of clinical practice? The research that is proposed below will measure contemporary standards of optometric practice. If this research is repeated in the future then it may allow changes in standards of practice to be detected.

Governmental and professional policy decisions

Governments sometimes have targeted campaigns on healthcare issues, for example cataract treatment and glaucoma detection in people of African ethnic origin. These campaigns rely on the assumption that appropriate clinical services are available in primary eye care to detect and manage these conditions. Research into the
content of an eye examination using specific methods described below will contribute towards the provision of this information.

Implications for NHS General Ophthalmic Services

The NHS provides the General Ophthalmic Services (GOS) and in 2005–2006, 12 million of the 17.5 million primary eyecare examinations were funded by the NHS (2006). Schedule 1 of the NHS (GOS) Regulations (1986) Terms of Service states ‘A contractor shall, having accepted pursuant to the regulations an application for the testing of sight, test the sight of a patient to determine whether the patient needs to wear or use an optical appliance, and on so doing shall fulfil any duty imposed on him by, or in Regulations made under, section 20B of the Opticians Act 1958’ (Association of Optometrists, 2006a). The Opticians Act does not define the content of an eye examination in detail (see next section). The Association of Optometrists has produced a document defining the contents of the GOS sight test (Association of Optometrists, 2006b), and the College of Optometrists guidelines are also relevant in this context. However, the authors are not aware of any research that investigates what actually occurs during a typical GOS sight test.

It is perhaps surprising that the GOS funds an item of service, the contents of which are poorly specified and poorly quantified. Information on the typical content and length of a GOS sight test might be useful for a number of reasons, including negotiations on fees and discussions on expanding the role of the optometrist. Research of this nature using standardized patients (SPs) (described below) presenting for a GOS sight test, would provide the first solid data on the actual content of the GOS sight test.

Clinicolegal issues

Section 26 (1a) of The Opticians Act (1989) (Office of Public Sector Information, 1989) states that when a registered optometrist tests the sight of another person, it shall be their duty to perform such examinations of the eye for the purpose of detecting injury, disease or abnormality in the eye or elsewhere. There are also Statutory Instruments, most recently SI 1999/3267 relating to patient referral (GOC, 1999).

The Sight Testing (Examination and Prescription) (No. 2) Regulations 1989 give statutory force to the common law duty of care which optometrists must meet, namely that whenever a person’s sight is tested within the meaning of Section 24 as defined by Section 36(2) of the Opticians Act 1989 (Office of Public Sector Information, 1989), a full eye examination must be carried out: this requirement holds whether the eye examination is carried out under the NHS or under private contract. The Regulations require the optometrist to perform an examination of the external eye and its immediate vicinity, an intra-ocular examination, either by means of an ophthalmoscope or other appropriate means, and any additional examinations as appear to the optometrist to be clinically necessary.

It is becoming increasingly common for healthcare practitioners to be accused of malpractice. For optometrists, malpractice accusations are defended in disciplinary hearings before the GOC or in civil litigation. It is an acceptable defence if it can be demonstrated that there is a body of reasonably competent optometrists who would have acted in a similar way to the practitioner (the Bolam and Bolitho tests (Jones, 1996; Herring, 2006)). Unfortunately, there is a lack of information on the content of eye examinations carried out by reasonably competent optometrists, as discussed above. This impacts on both the prosecuting and defending counsels, and means that they have to rely to a great extent on expert witnesses. The lack of factual information on the contents of a typical eye examination means that the experts also often have to base their advice on anecdotal experience, rather than factual data. For example, in recent GOC disciplinary cases experts have commented that they felt that a certain test should have been in an eye examination because it was a test that would have been carried out by a pre-registration optometrist in their final examinations. This reflects confusion about the different levels of assessment of clinical knowledge, as illustrated in Figure 1.

We believe that research is necessary to provide rigorous data on the content of eye examinations of randomly selected optometrists. This factual information will make it much easier for the Bolam and Bolitho tests to be applied in a fair and consistent way.

Professional guidelines

The College of Optometrists has published a code of ethics and guidance for professional conduct in accordance with the College’s objective: ‘the maintenance for the public benefit of the highest standards of professional competence and conduct’ (College of Optometrists, 2006). The College stresses that the guidelines are by no means a set of instructions or a checklist, allowing each practitioner to exercise his or her professional judgment (College of Optometrists, 2006).

A detailed understanding of the nature of current optometric practice and of typical standards within optometric practice would be helpful for evolving professional guidelines. Evidence-based research on the content of typical optometric eye examinations would help to develop guidelines that differentiate between

© 2007 The Authors. Journal compilation © 2007 The College of Optometrists
realistic minimum standards of competence (e.g. an important test that nearly all optometrists carry out) and aspirational goals (best practice, that may not necessarily be achieved by a significant body of reasonably competent optometrists).

**Consumer complaints**

Consumer complaints about optical services in the UK are typically dealt with by the Optical Consumer Complaints Service (OCCS), an independent body set up to settle complaints from members of the public who are not satisfied with the optical services received in an optical practice (Optical Consumer Complaints 2006). When these complaints relate to the eye examination, it would be useful for OCCS to know the content of a typical eye examination.

**Setting priorities for undergraduate training**

The GOC approves institutions for the training of optometrists. This approval is based upon the institutions demonstrating that their training secures to the student adequate knowledge and skill for the practice of their profession. An investigation of the content of typical optometric examinations is likely to identify priorities not just for CET, but also for undergraduate training.

**Methods of measuring clinical practice**

The authors believe that the points listed above indicate the need for objective, evidence-based, data on the content of typical optometric eye examinations in the UK. This raises the question of ‘What is the most appropriate method for measuring clinical practice?’ In other words, how can we measure the peak of Miller’s pyramid (Figure 1)? This question is applicable to all healthcare professions, so a literature review was carried out to determine the answer to this question by analysing work in other healthcare professions. The literature revealed that clinical practice is commonly assessed by three methods: (1) abstraction of medical records, (2) using clinical vignettes and (3) use of SPs who present unannounced to clinics. Each of these methods is discussed in detail and is summarized in Table 2 below. Another method of measuring clinical practice, assessing billing claims for various procedures, is not well-suited to optometry in the UK because the NHS is the major funding source and in most regions presently funds only one item of service.

**Record abstraction**

Record abstraction has been described as the most widely used method of measuring quality of clinical care (Rubin et al., 1992; Gilbert et al., 1996; McDonald et al., 1997). Record abstraction is sometimes performed simultaneously with use of SPs (Luck et al., 2000; Dresselhaus et al., 2002) and/or clinical vignettes (Dresselhaus et al., 2000; Peabody et al., 2000) (described below). Records generated during a clinical encounter are retrieved at the end of the visit for abstraction by a skilled expert. The requirement of skilled expertise means record abstraction is expensive to perform. Information generated during the abstraction is recorded on a pre-designed checklist. The individual scoring items in a checklist are categorized into four domains of clinical performance: history and symptoms, physical examination, diagnosis and treatment. The checklist is filled out using the information abstracted from each record and scores assigned for each of the four domains and added to obtain an overall score.

There are widespread concerns regarding the use of this method due to the validity and reliability of results obtained (Norman et al., 1985; Rethans et al., 1994; McLeod et al., 1997). One of the main limitations to consider is that record abstraction is subject to false negative results, i.e. tests carried out but not documented in the record (Dresselhaus et al., 2000; Luck et al., 2000). Busy practitioners may not record everything that was examined during the consultation. On the other hand, good record keepers may not necessarily be good physical examiners. The opposite form of bias can also occur, i.e. concern over medico-legal attention might lead some practitioners to record tests that they have not completed (Dresselhaus et al., 2000; Luck et al., 2000). These limitations could therefore skew the results leading to an overestimation or underestimation of the quality of care (Lawthers et al., 1995; Katz et al., 1996).

Other problems associated with record abstraction include illegibility, incomplete or unavailable record cards and differing skills between abstractors (Dresselhaus et al., 2000, 2002; Peabody et al., 2000). Record abstraction only provides a limited insight into the practitioner’s clinical skills and practitioner–patient interactions. The usefulness of record abstraction is further limited by the fact that a skilled (and costly) expert must collect the data (Norman et al., 1993; Ashton et al., 1995).

**Vignettes**

Vignettes are written or computerized case simulations that have been widely used by educators and health service researchers to measure processes in a range of practice settings (Sriram et al., 1990; O’Neill et al., 1995; Glassman et al., 1997). Vignettes are set up to simulate patient visits and have been used to measure a practitioner’s ability to evaluate, diagnose and treat specific medical conditions (Peabody et al., 2004). A
practitioner is exposed to the presenting problem and asked to provide open-ended responses identifying the most important element(s) of history for each case scenario. A similar stepwise process is repeated for the physical examination, diagnostic testing and treatment plan (Dresselhaus et al., 2000). Practitioners are not allowed to return to and modify previously completed answers as new information is provided at each stage. Skilled experts score the completed vignettes in a similar way to that described above for record abstraction.

Clinical vignettes are not only designed to simulate a range of medical conditions but also to evaluate skills required in the care of the patient. Each practitioner

<table>
<thead>
<tr>
<th>Method</th>
<th>Limitations</th>
<th>Level in Miller’s pyramid</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unannounced standardized patient</td>
<td>Selection bias, avoided by encouraging high participation rate&lt;br&gt;Hawthorne effect if SP is expected, avoided by fairly long duration of study&lt;br&gt;Normally only used with new patients (first visit) (&lt;em&gt;Luck and Peabody, 2002&lt;/em&gt;)&lt;br&gt;Expensive</td>
<td>Does</td>
</tr>
<tr>
<td>Announced standardized patient (e.g., at training session)</td>
<td>Selection bias if low participation rate&lt;br&gt;Hawthorne effect inevitable (practitioner will behave differently as being observed)&lt;br&gt;Expensive</td>
<td>Shows how</td>
</tr>
<tr>
<td>Direct observation of patient consultation by an expert</td>
<td>Selection bias if low participation rate&lt;br&gt;Hawthorne effect (&lt;em&gt;Luck et al., 2000&lt;/em&gt;) inevitable (practitioners perform better when observed)&lt;br&gt;(&lt;em&gt;Franco et al., 1997&lt;/em&gt;)&lt;br&gt;Very expensive because involves expert observer (&lt;em&gt;Luck et al., 2000&lt;/em&gt;)</td>
<td>Shows how</td>
</tr>
<tr>
<td>Clinical vignettes (response to written case scenarios)</td>
<td>Selection bias if low participation rate&lt;br&gt;Hawthorne effect inevitable&lt;br&gt;Vignettes over-estimate the quality of an examination and are inaccurate for reporting treatment plans (&lt;em&gt;Peabody et al., 2000&lt;/em&gt;)</td>
<td>Knows how</td>
</tr>
<tr>
<td>Abstracted clinical records</td>
<td>Selection bias if low participation rate&lt;br&gt;Hawthorne effect if abstraction expected&lt;br&gt;Practitioners ‘under-record’ tests actually done (&lt;em&gt;Dresselhaus et al., 2000; Luck et al., 2000; Peabody et al., 2000&lt;/em&gt;)&lt;br&gt;Practitioners ‘over-record’ tests not actually done (&lt;em&gt;Luck et al., 2000; Dresselhaus et al., 2002&lt;/em&gt;)&lt;br&gt;May be illegible&lt;br&gt;Expensive because a skilled expert must abstract the data (&lt;em&gt;Peabody et al., 2000&lt;/em&gt;)&lt;br&gt;A less accurate method of measuring quality than vignettes (&lt;em&gt;Peabody et al., 2000; Peabody et al., 2004&lt;/em&gt;)</td>
<td>Knows how</td>
</tr>
<tr>
<td>Interview with practitioners</td>
<td>Selection bias if low participation rate&lt;br&gt;Hawthorne effect&lt;br&gt;Practitioners state that they do more than they actually do (&lt;em&gt;Franco et al., 1997&lt;/em&gt;)&lt;br&gt;Reveal knowledge, not performance (&lt;em&gt;Franco et al., 1997&lt;/em&gt;)</td>
<td>Knows</td>
</tr>
<tr>
<td>Questionnaire about current practice</td>
<td>Selection bias if low participation rate&lt;br&gt;Hawthorne effect inevitable&lt;br&gt;Reveal knowledge, not performance</td>
<td>Knows</td>
</tr>
</tbody>
</table>
could be asked to complete several vignettes to simulate diverse medical conditions (Peabody et al., 2004). Vignettes are a cost-effective way of assessing levels of clinical care and can be easily administered and therefore used in a great variety of settings.

Peabody et al. (2004) used computerized vignettes. The practitioner completing the vignette ‘sees the patient’ on a computer. Vignettes are well suited to large scale (Epstein et al., 2001; Morita et al., 2002) quality assessments for cross-system comparisons (O’Connor et al., 1996; Nordske, 2002) or if ethical issues preclude the use of patients or their records (Rosen et al., 1995; Gould, 1996; Aitken et al., 1998).

Despite widespread use of vignettes, there is uncertainty and controversy about whether vignettes reflect actual clinical practice or merely practitioners’ competence at the vignette task (Peabody et al., 2000). Some researchers feel vignettes only reflect what practitioners are competent or knowledgeable enough to do (Everitt et al., 1990; Rethans et al., 1991). For example, it seems likely that practitioners will give their ‘best answer’ when responding to a vignette as they are in an assessment scenario. This best answer may not reflect the tests that they would actually have carried out if they had been presented with such a patient during everyday clinical practice when they would be unaware that their clinical performance was being observed or assessed.

**Standardized patients**

During a clinical consultation, only two people are usually present: the practitioner and the patient. So, the most appropriate way of determining what the practitioner does is to ask the patient, in particular to ask a patient who has been trained to be an expert observer. There are numerous descriptors of the roles played by individuals during simulated healthcare encounters. Examples of these descriptors include programmed patients, prepared patients, trained patients, standard patients, actors, patient instructors and pseudo-patients. Each term however has a specific meaning depending on the clinical setting and the encounter being simulated in that setting. Nevertheless, the term *standardized patient* is a well-accepted term in the literature, with only one author using the term differently (Barrows, 1993), and as can be seen from its definition below it describes an approach that is ideally suited to determining an optometrist’s performance in a clinical setting.

A *simulated patient* encounter occurs when practitioners examine people who are simulating real patients. In optometry, this occurs during the final assessment at the end of the PRP. The most rigorous form of simulated patient is a SP who is trained to give consistent verbal and behavioural responses to the examiner (Adamo, 2003) in order to accurately portray a specific patient (Ebbert and Connors, 2004). Typically, the SP is a highly trained actor.

The SP approach has been used in several healthcare professions for 40 years (Whelan et al., 2005) and a search on PubMed (06-Oct-06) for the key phrase ‘standardized patient’ found 462 references. The literature on the use of SPs will now be summarized.

**Literature review of the use of standardized patients**

**Comparisons of methods of measuring clinical care**

Franco et al. (1997) compared three methods of assessing the performance of healthcare practitioners: direct observation of patient consultation, interviews with practitioners and SPs. They found that SP data are probably the best in reflecting normal practice and that during direct observation the practitioner is likely to give better than normal levels of quality of care. The authors cautioned that data from interviews with practitioners may reflect practitioner knowledge and not necessarily performance (Franco et al., 1997): the base of Miller’s pyramid rather than the top (Figure 1). Concerning SP study design, their data suggested that practitioners’ behaviour is not consistent across several patients and so SP testing should ideally be repeated with more than one patient if a more accurate reflection of a practitioner’s typical practice is to be obtained (Franco et al., 1997).

Three methods of assessing practitioners’ compliance with preventative care guidelines were compared by Dresselhaus et al. (2000): abstracted medical records, SPs and responses to written case scenarios (vignettes). Clinical record abstraction under-estimated performance by 16%, compared with SP check-lists which were taken to be the gold standard. Depending on the aspect of clinical performance that was measured, vignettes were either superior to or not different from record abstraction. The authors concluded that relying on clinical records is misleading (Dresselhaus et al., 2000). It was subsequently shown that practitioners’ clinical records over-estimate the quality of care in some cases (Dresselhaus et al., 2002).

Luck et al. (2000) showed that medical records were neither a sensitive nor a specific report of the clinical encounter. Moreover, as the differences in scores between record abstraction and SP checklists ranged from −10% to +23% for different aspects of the consultation, it is not possible to apply a ‘correction factor’ to convert scores based on record abstraction to an equivalent for SP data.

Clinical vignettes, record abstraction and SPs for four common outpatient medical conditions were compared by Peabody et al. (2000). A three-way comparison of
methods used to assess quality of care for all cases combined, revealed SP scores (76.2%) to be consistently higher than vignettes scores (71.0%) and record abstraction (65.6%). Vignettes were superior to record abstraction for most measures, but compared with SPs, vignettes over-estimated the quality of examinations and were inaccurate at reporting treatment plans (Peabody et al., 2000). A later study used computerized clinical vignettes, which were significantly superior to record abstraction (Peabody et al., 2004). Indeed, it was suggested in this study that, if appropriately designed, vignettes can achieve greater accuracy than previous authors had suggested. However, Peabody and colleagues still used SPs as their gold standard measure.

**SPs: the gold standard for measuring clinical practice**

Standardized patients are not the only method that has been used to investigate clinical practice and standards, but unannounced SPs (and completed SP checklists) are regarded as the gold standard for quality measurement in clinical practice (Dresselhaus et al., 2000, 2002; Luck et al., 2000; Peabody et al., 2000, 2004; Luck and Peabody, 2002; Peabody et al., 2004). Luck and Peabody (2002) demonstrated the validity of SPs to measure the quality of physicians’ practice, as the gold standard, by covertly tape recording the SP visit. At the end of the visit, the SPs reported on the physician’s performance by completing a checklist to ‘score’ the consultation in the usual way, but the tape recordings were also independently ‘scored’ by experts. The level of agreement was very high (sensitivity 95%, specificity 85%) and the authors concluded that SP assessment is a valid measure of the quality of care (Luck and Peabody, 2002). This supports the assertion by many authors that SPs are the gold standard method of measuring the quality of practitioner’s practice. To the present authors, this assertion has considerable face validity. In attempting to measure clinical practice, it seems likely that the major confounding variable will be the tendency for people to change their habits when they know that they are being observed or assessed. Of all the methods of measuring clinical practice, it is only unannounced SPs that can determine what practitioners do without alerting them to the fact that they are being assessed.

To measure everyday clinical practice, it is important for the SPs to be unannounced: the practitioner must not believe that the SP is there to assess their clinical practice. Several authors have provided detailed summaries of the use of SPs (Ramsey et al., 1998; Glassman et al., 2000; Adamo, 2003). The various methods of investigating clinical practice are contrasted in Table 2.

**Use of SPs in training**

A common use of SPs is in clinical skills training and assessment of medical students, where they are often used in an objective structured clinical examination (OSCE) (Barrows, 1993; Adamo, 2003; Major, 2005). The United States Medical Licensing Examination now uses SPs (Adamo, 2003) and SPs have also been used in dentistry (Maupome and Sheiham, 2000) and nursing (Ebbert and Connors, 2004).

The structure of the College of Optometrists’ pre-registration training year and professional qualifying examinations has recently been significantly modified. The new scheme for registration is based on continuous assessment in the workplace during the PRP, followed by a final assessment. Although some components of the optometric pre-registration final assessment are a form of OSCE, they do not use SPs. Rather, they use simulated patients with in-room examiners carrying out the assessment.

A simulated patient is someone who pretends to be a patient but who, in contrast with a SP, has not been trained to complete a check-list that allows an assessment of the examination. Also, SP encounters tend to be unannounced whereas the practitioners usually know when they are examining a simulated patient. Use of simulated patients for educational purposes avoids mistreatment of real patients and allows students to work without embarrassment about their novice status (Barrows, 1993). Working with SPs allows trainees to build their confidence and learn from actual patients without the trainee worrying about their ability or technique (Barrows, 1993).

Standardized patients used in training can be manipulated for educational purposes unlike a real patient and can therefore be used to directly assess behaviours that are required in a competent clinical performance. SPs could be used to monitor the progress of pre-registration training optometrists as part of their continuous assessment and in the final assessment.

**Use of SPs in assessing clinical care**

In addition to their use in training and for examinations, SPs have also been widely used to assess the quality of clinical care given by qualified practitioners (Ramsey et al., 1998; Barragan et al., 2000; Glassman et al., 2000; Peabody et al., 2000; Luck and Peabody, 2002; Bachmann et al., 2004; Dresselhaus et al., 2004). They can be used not just to assess clinical criteria, but also to investigate history taking (Ramsey et al., 1998), compliance with preventative care guidelines (Dresselhaus et al., 2000), and advice/counselling given to the patient (Russell et al., 1983; Ramsey et al., 1998).
When SPs are used for quality assessment with qualified practitioners, the practitioners and staff in most studies are unaware of when they are seeing an SP (unannounced SPs), although in one or two studies the SPs have been seen in a special clinic on a special day (announced SPs) (Ramsey et al., 1998).

Rethans and Saebu (1997) used SPs to establish the consistency in performance of general practitioners when they examine the same patient twice. This study also assessed inter-examiner variability when the same SP is examined by more than one clinician. Although there was no significant difference in the performance of general practitioners between the first and second consultation of the same patient, the results showed significant variation in performance between physicians (Rethans and Saebu, 1997).

**SP recruitment, training and quality assurance**

The selection and training of SPs is crucial in studies measuring clinical care (Ramsey et al., 1998; Luck and Peabody, 2002; Adamo, 2003; Dresselhaus et al., 2004; Peabody et al., 2004). Individual SPs are usually selected on the basis of age, gender, ethnicity, physique, current and previous medical history and level of education and/or language. Certain other characteristics (Table 3) are important and are usually assessed during and after SP training.

Throughout the course of the research, each SP’s performance is monitored for quality assurance (Adamo, 2003; Peabody et al., 2004). This is usually achieved either by video-taping or by directly monitoring a clinical encounter (Luck et al., 2000).

The SP is usually matched to a case requirement (e.g. of an appropriate age, race and possibly with symptoms or signs of relevant pathology) and trained to reliably portray a clinical scenario and accurately recall both the details of the conversation between the practitioner and the SP and the tests performed during the encounter. The SPs usually report their clinical encounter by providing an accurate, written and objective report in the form of a checklist.

One of the main drawbacks of using the same SP for several clinical encounters is the need for them to continue to portray themselves as a ‘non-expert’ patient. Having undergone several examinations, SPs begin to volunteer information during the course of the examination thereby prompting the practitioner (Adamo, 2003). It is therefore important to monitor SP consistency by video recording or directly observing a clinical encounter for quality assurance purposes (Adamo, 2003; Peabody et al., 2004).

**Checklists for SP based assessments**

At the end of an SP encounter, it is usual practice for the SP actor to report a practitioner’s performance by completing a checklist. These checklists can either be case specific or generic (Gorter et al., 2000). Generic checklists are used in assessing general skills whereas case specific checklists provide detailed information about a practitioner’s skills during history taking, physical examination, case management and/or communication (Gorter et al., 2000). Case specific checklists are therefore tailored to the content of the consultation.

The information recorded in the checklists should be accurate, reliable and valid. Therefore, it is important that the actors are carefully trained to complete the checklists in a consistent manner. Checklist development is crucial in order to obtain a valid and reliable record of a practitioner’s performance. At the end of the checklist development stage it is essential to state who developed the checklist, whether or not the development of the checklist was based on reviewed literature or data resulting from consensus procedures, and the scoring system used (Gorter et al., 2000).

**Choice of standardized patient profiles**

The reviewed literature reveals the importance of the need to carefully develop scenarios incorporating observable evidence-based criteria into realistic scripts and objective checklists, and to carry out extensive pre-testing of the scripts (Ramsey et al., 1998; Glassman et al., 2000; Luck and Peabody, 2002). In SPs used with physicians, actors have been shown to be able to cope with completing, immediately after the consultation, checklists of 35–45 items that might have been performed or discussed (Luck and Peabody, 2002).

A typical number of scenarios to select are 3 or 4 (Luck et al., 2000; Dresselhaus et al., 2002; Bachmann et al., 2004). Typically, the case scenarios and scoring criteria for the SPs are selected based on evidence-based reviews and clinical guidelines, and are reviewed by a panel of experts during the development phase of the study (Luck et al., 2000; Peabody et al., 2000; Luck and Peabody,

---

**Table 3. Important characteristics of standardized patients (SPs)**

<table>
<thead>
<tr>
<th>Characteristics assessed during training</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ability to adapt to varying practitioner and/or interviewer styles</td>
</tr>
<tr>
<td>Ability to effectively portray a case requirement</td>
</tr>
<tr>
<td>Demonstration of active listening and communication skills</td>
</tr>
<tr>
<td>Demonstration of promptness and preparedness</td>
</tr>
<tr>
<td>Demonstration of ability to adapt behaviour as a result of coaching/feedback</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Characteristics assessed at the end of training</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stable findings during clinical examination</td>
</tr>
<tr>
<td>Ability to deliver constructive feedback from a patient perspective</td>
</tr>
<tr>
<td>Recording accuracy</td>
</tr>
</tbody>
</table>

© 2007 The Authors. Journal compilation © 2007 The College of Optometrists
Scoring is usually broken down into domains of the consultation (Luck and Peabody, 2002; Whelan et al., 2005) such as: history and symptoms, examination, further tests requested, diagnosis and management (prescription and discussion) (Peabody et al., 2000).

Sample size

Sample size in SP research can be described by stating (1) the number of scenarios investigated (one scenario might be played by more than one actor), (2) the number of practitioners who received SP visits and (3) the total number of SP encounters (visits). This latter variable is usually the product of (1) and (2).

In a comparison of different methods of measuring quality of clinical care, Franco et al. (1997) carried out SP assessments of 20 practitioners. In another evaluation of different methods of measuring quality in clinical encounters, Peabody et al. (2000) invited 101 physicians to participate in the study. Ninety-seven physicians consented to be randomized for the study. Twenty of these 97 practitioners were randomly selected for inclusion in the study.

Ramsey et al. (1998) used SPs to investigate the ability of primary care physicians to take a complete and accurate history from their patients using a sample size of 134. Of those who were originally asked to participate only 53% consented, but this low participation rate compared with other published SP studies may have been because practitioners were required to take their history from the (announced) SPs in a special clinic on Saturdays.

A study to assess clinical competence in primary care evaluated 22 doctors using three SP visits for each doctor (66 in total) (Barragan et al., 2000). However, these authors noted that their conclusions were limited by the small sample size. In a study of primary care physicians, 79% of those approached gave consent to participate, giving a sample size of 232 (Bowman et al., 1992). In another study of 101 primary care physicians, 97% agreed to participate, from which a sample of 20 practitioners were chosen, each of whom saw eight SPs (played by 27 actors), resulting in a total of 160 SP encounters (Dresselhaus et al., 2000, 2002; Glassman et al., 2000; Luck et al., 2000).

In a study to validate SP use in clinical settings, Luck and Peabody (2002) found that 88% of 163 eligible physicians consented to participate and recorded data were obtained from 40 of these. This population also indicates that between 53% and 98% of invited practitioners decline to participate. The studies reviewed above indicate that between 53% and 98% of invited practitioners accept the invitation to participate.

Hawthorne effect. The Hawthorne effect is the positive impact on behaviour that sometimes occurs in a study as a result of the interest shown by the experimenter in humans who are being treated, studied, or observed (Lied and Kazandjian, 1998). From a scientific point of view, the ideal SP study would be of practitioners who are more confident of their clinical skills may be more likely to volunteer (Ramsey et al., 1998), which could result in a bias likely to discover a higher standard of practice than that which is typical. This is a problem common to all SP research; therefore, it is surprising that the authors are not aware of this issue being raised in the literature. However, this is only likely to be a major problem if a high proportion of invited practitioners decline to participate. The studies reviewed above indicate that between 53% and 98% of invited practitioners accept the invitation to participate.

Selection bias. In SP research practitioners are the research participants and it is therefore appropriate to afford them the same rights as are applied to any research participant. This means that practitioners should only be included as participants if they have given informed consent; in other words, if they volunteer. Practitioners who are more confident of their clinical skills may be more likely to volunteer (Ramsey et al., 1998), which could result in a bias likely to discover a higher standard of practice than that which is typical. This is a problem common to all SP research; therefore, it is surprising that the authors are not aware of this issue being raised in the literature. However, this is only likely to be a major problem if a high proportion of invited practitioners decline to participate. The studies reviewed above indicate that between 53% and 98% of invited practitioners accept the invitation to participate.

Limitations of SP research

From the literature, we conclude that it is typical for SP case scenario(s) to be tested on 20–232 practitioners and the mean number of practitioners that were assessed in previous studies is 58. The number of SP visits per practitioner varies from one to eight.
the chances of a significant Hawthorne effect are reduced if participating practitioners selected are those who normally examine a fairly high number of new patients and if practitioners are informed that the SPs will visit the practice at any time over a reasonably long time period.

SP research in optometry

The authors carried out a literature search to identify research that employed SP methodology or other methods of measuring practitioner standards in optometry. The search was carried out in April 2006 using PubMed MeSH terms: standardized patient OR simulated patient OR vignette OR record abstraction OR clinical scenario AND optometry OR primary eye care OR optician. This search revealed no references that contained any of the above combination of words.

Conclusions

This review has concentrated on SPs as these are recognized as the gold standard method of assessing clinical practice and assess ‘real life’ clinical practice. As long as practitioners do not detect the SP then clinicians’ true behaviour will be observed and will not be modified by the awareness that they are being assessed. Although this is not true of assessment by clinical vignettes, recent research demonstrates that carefully constructed computerized clinical vignettes can also obtain valuable data and compare fairly well with the gold standard of SPs. Indeed, we would stress that all methods of measuring clinical practice have advantages and disadvantages (Table 2).

In many respects, it should be easier to apply SP methodology and clinical vignette methodology to optometric practice than to medical practice. This is because, compared with, for example, a general medical practitioner, a community optometrist has a much narrower scope of practice and a more limited potential test battery. For all the reasons outlined in this paper, it is surprising that there appear to have been no published scientific attempts to obtain an evidence-based assessment of clinical practice within primary care optometry. The present authors have started a programme of research using SPs and clinical vignettes with consenting practitioners to obtain such an assessment. The ethos of the study is to be descriptive rather than judgemental. The research will clearly define the current scope of routine optometry, so that appropriate goals and plans for the future can be made.

Acknowledgements

Rakhee Shah, Dave Edgar and Bruce Evans are members of EyeNET, the London primary care eye research network funded by the Department of Health. The present work was funded in part by EyeNET, and also by the Association of Optometrists, Royal National Institute of the Blind (RNIB), Central (LOC) Fund, College of Optometrists and the American Academy of Optometry (British Chapter). The authors received a proportion of their funding from the Department of Health; the views expressed in this publication are those of the authors and not necessarily those of the Department of Health.

References


© 2007 The Authors. Journal compilation © 2007 The College of Optometrists