UNIVERSITY of BRADFORD Library

University of Bradford eThesis

This thesis is hosted in Bradford Scholars – The University of Bradford Open Access repository. Visit the repository for full metadata or to contact the repository team

© University of Bradford. This work is licenced for reuse under a Creative Commons Licence.

REFERRALS FROM PRIMARY EYE CARE: AN INVESTIGATION INTO THEIR QUALITY, LEVELS OF FALSE POSITIVES AND PSYCHOLOGICAL EFFECT ON PATIENTS

Christopher James DAVEY

PhD

BRADFORD SCHOOL OF OPTOMETRY AND VISION SCIENCE
SCHOOL OF LIFE SCIENCES
UNIVERSITY OF BRADFORD

2011

Referrals from Primary Eye Care:

An Investigation into their quality, levels of false positives and psychological effect on patients.

Christopher J Davey

Supervisors:

Professor David B. Elliott

Dr Clare Harley

Miss Clare Green

Abstract

Previous research into the accuracy of referrals for glaucoma has shown that a large number of referrals to the Hospital Eye Service are false positive. Research in areas of healthcare other than ophthalmology has shown that psychological distress can be caused by false positive referrals. The present study aimed to evaluate the quality of referrals to the HES for all ocular pathologies, and also to quantify the proportion of these referrals that were false positive. Any commonality between false positive referrals was investigated. The psychological effect of being referred to the HES was also evaluated using the Hospital Anxiety and Depression Scale (HADS) and State-Trait Anxiety Inventory (STAI). Both scales were validated in this population with Rasch analysis before use. A final aim was to develop an improvement to the present referral pathway in order to reduce numbers of false positive referrals.

The accuracy of referrals to the HES appears to improve as clinicians become more experienced, and greater numbers of false positive referrals are generated by female clinicians. Optometrists refer patients with a wide range of ocular diseases and in most cases include both fundus observations and visual acuity measurements in their referrals. GPs mainly refer patients with anterior segment disorders, particularly lid lesions, based on direct observation and symptoms. Illegibility and missing clinical information reduce the quality of many optometric referrals. Patients referred to the HES experience raised levels of anxiety as measured by the STAI and raised levels of depression as measured by the HADS-Depression subscale. As a method of assessing psychological distress, the questionnaires HADS-T (all items), STAI-S (State subscale) and STAI-T (Trait subscale)

show good discrimination between patients when administered to a population of new ophthalmic outpatients, despite all having a floor effect. Subsequently a referral refinement service was developed which reduced numbers of unnecessary referrals and reduced costs for the NHS.

Acknowledgements

There are two people who have had seemingly endless patience during the writing of this thesis, my wife Fahmidah and my main supervisor Dave.

Fahmidah has tolerated four and a half years of missed deadlines, impingement on our personal life and reduced income. There have also been quite a few conversations that she must have found mind numbingly tedious. She has remained tirelessly encouraging and surprisingly optimistic, but the main reason I have to thank her is for still marrying me despite my promise that I would definitely have finished before the wedding.... in reality it turned out to be our first anniversary. You will probably never read this.... but thank you.

Dave has also had a superhuman tolerance for stupid questions, missed deadlines and repetition of the same stupid questions. Either that, or he keeps forgetting how frustrating our previous meeting was and agrees to another one. He should be thanked for dragging me over the finishing line when it was so close but seemed eternally just out of reach. His experience and breadth of knowledge was invaluable and I have learned a vast amount about research and academia from him. One similarity between Fahmidah and Dave is that eventually they seem to land at the same final solution; a trip to the pub.

My other supervisors, Clare and Clare were both absolutely essential in the planning and implementation of the study. It would not exist without either of them. It was equally convenient and confusing that they had the same name.

And a few more people that I feel have been invaluable over the last four and half years: Jim Hanson, for the fresh green chilli. Ian Pacey, for the FC. My parents, for the roof. Mark Holloway, for the ~~squash and beer~~ data entry. Stewart Mitchell, for the SOAP. Naveen, for the mangocellular. Monika, for the feuerzangenbowle. Luis, for the urbex. Marta and Esti, for the Rasch. John, Alison, Pedro, Vanda, t'other Alison, Atiq and all the other post-grads who have passed through G39 for the shared knowledge and support.

That's enough now.

Table of contents Page

Chapter 1: Introduction

Referrals to Ophthalmology Departments in the United Kingdom

As there are no national screening programs in the UK for ophthalmic pathologies, apart from the diabetic retinopathy screening service, most eye problems are detected by optometrists performing routine eye examinations in a primary eye care setting. Only optometrists, medical practitioners and ophthalmic medical practitioners (OMPs) are permitted to perform eye examinations in the UK. The Opticians Act 1989 (SI 1999 No. 3267) states that optometrists have a duty of care to refer patients with ocular injury or disease to a medical practitioner if, in their professional opinion, they think it is justified. Prior to Statutory Instrument (SI) 3267 in 1999, optometrists were required to refer every ocular abnormality regardless of their professional opinion, although very few closely followed this requirement.

The normal referral pathway (illustrated by Figure 1.1) from primary to secondary eye care in England and Wales at the present time involves the optometrist writing a referral letter, or completing a specialised referral form (General Ophthalmic Services, GOS 18 form), to the patients' General Medical Practitioner (GP), and advising the patient to book an appointment with them. The letter should include all relevant detail from the eye examination, and a tentative diagnosis with suggested urgency if possible. After the patient has seen the GP, the GP generates a referral to the Hospital Eye Service (HES) including a copy of the optometrist's letter and any additional relevant information. At the hospital a senior clinician in the ophthalmology department reads all referrals and prioritises them before

appointments are allocated. This prioritising of cases at the eye department currently differs from most other specialities within the hospital. Particularly since the Department of Health introduced the 'Choose and Book Policy Framework' in 2004, GPs can prioritise most cases themselves and book appointments directly from their practice after offering patients a choice of providers. Consultant led prioritisation is necessary in Ophthalmology due to the complex nature of referrals, although a recent study suggests that, with guidelines and protocols, specially trained non-medical practitioners could safely prioritise most referrals (Hodi, 2007).

If the optometrist believes the patient needs immediate ophthalmic opinion, then the GP can be bypassed by sending the patient directly to an Accident and Emergency clinic with a referral letter. A copy of the letter is also sent to the GP. Some Primary Care Trusts (PCT's) have initiated 'Direct Referral' schemes for certain pathologies in order to reduce waiting times by removing the initial appointment with the GP. These services involve the optometrist taking on more responsibility for the referral by, for example, performing further diagnostic tests, counselling the patient or offering choice. As this is not part of the GOS contract, appropriate remuneration is given to the optometrist. The GP receives a fee to forward some necessary clinical information, (Dinah et al., 2010), but the initial assessment appointment is at the hospital. Most direct referral schemes have been set up for cataract patients, however some PCTs have similar schemes in place for glaucoma (Henson et al., 2003) and posterior capsular opacification (Menon et al., 2004). The effectiveness of direct referral services is discussed in more detail later.

Optometrists are reported to refer up to 6 % of all their eye examinations for further investigation by a medical practitioner (Hobley et al., 1992, Port, 1989, Port and Pope, 1988, 2008) although most of these studies were performed some time ago in response to the Opticians Act 1989. Specifically the authors were concerned that the abolition of NHS (free) sight tests for everyone and introduction of ready-made spectacles for presbyopia ('ready readers') may have an adverse affect on the number of people annually diagnosed with treatable sight threatening pathology. This prediction was later confirmed by data from secondary care showing a decrease in diagnosis rates for glaucoma (Laidlaw et al., 1994). With an estimated 17.5 million eye examinations being performed in the UK per year (NHS, 2006) there are probably over 1 million HES referrals from optometric practice.

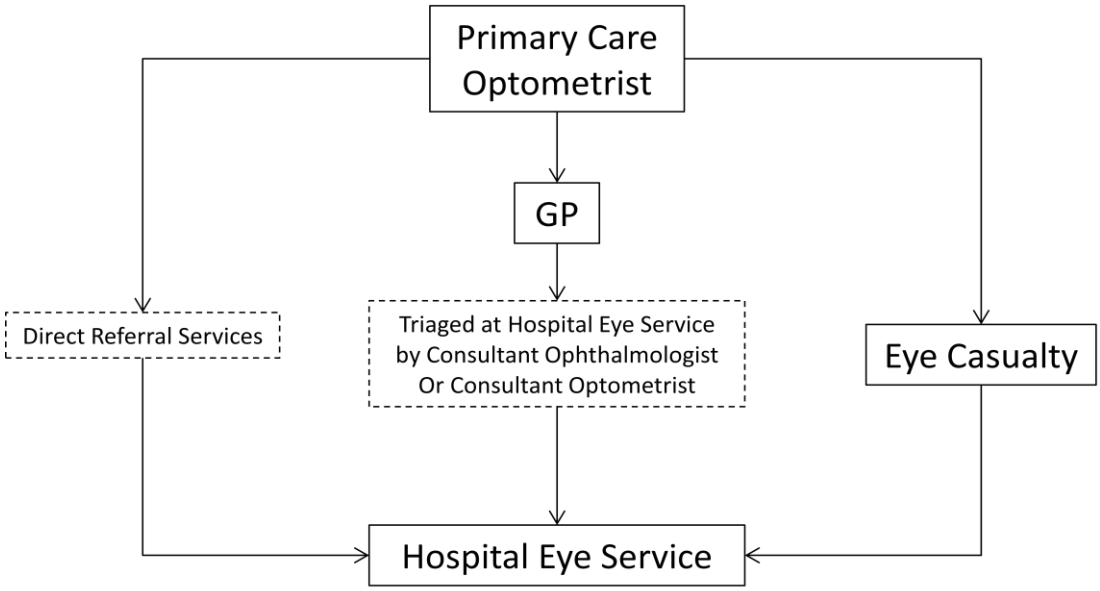

Figure 1.1. The referral pathway from primary care optometrists to the Hospital Eye Service in Bradford and Airedale.

Currently (1st April 2010-2011) optometrists in England receive a fee of £20.70 for every NHS sight test, otherwise known as a GOS sight test, performed on a

qualifying patient. This fee is normally incrementally increased every April but is significantly below the true cost of providing a sight test, which was quoted at an average of £37 in 2005, in a report by Bonsaquet (2006). Costs of a sight test will vary depending on the costs of rent, rates, council tax, equipment, staff etc. Personal communication of more recent unpublished calculations of the true cost of eye examinations may be even higher, at an average of £130 per hour (Llewellyn, 2008) and thus £65 for 30-minute sight tests and £43 for 20-minute sight tests. Patient groups that qualify for an NHS test include those who are over 60, under 16, receiving benefits, diabetic, diagnosed with glaucoma, over 40 with a family history of glaucoma, or blind or partially sighted. If a patient has a GOS sight test the optometrist is obliged to perform a sight test according to the Opticians Act, and determine a prescription for refractive correction. If the optometrist is suspicious of any ocular pathology, and supplementary or repeat measurements on a separate occasion would aid diagnosis, such as repeating visual field examination or tonometry to help in the possible diagnosis of glaucoma (Gardiner et al., 2006; Ang et al., 2009), no extra fee is received. Conversely, the optometrist incurs very little extra cost by referring the patient to the hospital. This reason, along with the fact that our society is becoming ever more litigious, makes referral increasingly tempting compared to further investigation. Despite not being funded by the GOS, College of Optometrists (CoO) guidelines deem repeat measurements of intra-ocular pressure as a necessary part of the eye examination and especially vital when referring (College of Optometrists, 2005). The value of repeat testing will be discussed in more detail later.

Private eye examinations and supplementary tests are priced at the discretion of the optometric practice and are paid for directly by the patient. Private examinations must include mandatory diagnostic tests (1989), but may include other more advanced techniques. Due to the current climate in the industry the average cost of a private examination (£22.90 according to the Federation of Ophthalmic and Dispensing Opticians, (2008) is still well below what it realistically costs to provide a sight test, and is similar to the NHS fee. Many English optical practices thus tend to subsidise examination fees by increasing the price of spectacles and other products (Calver, 2010).

Royal National Institute of Blind People (RNIB) research suggests that a significant proportion of patients do not go for regular eye examinations due to the fear of the cost of spectacles (Chitty, 1997), rather than the fear of the cost of the examination. Therefore treatable ocular pathology could be unnecessarily going undiagnosed due to spectacle prices subsidising professional costs.

Despite having an increased likelihood of having visual impairment (Lavery et al., 1988, Wormald et al., 1992, Klein et al., 1991), elderly people make insufficient use of eye care facilities (Smeeth and Iliffe, 1998). The paucity of usage may be due to a variety of factors, including decreased expectations in old age, failure by the patient to recognise visual loss, the presence of another handicap that dominates the perception of difficulties, fears about surgical treatment and costs and the stigma of blindness (Smeeth and Iliffe, 1998).

Optometry Scotland and the Scottish Executive Health Department have recently developed a new Optometry contract for Scotland to replace the GOS. One of its

aims was to reduce inappropriate referrals to the Hospital Eye Service by improving the standard of the NHS sight test, and funding repeat or supplementary appointments where clinically necessary. The fee for the primary eye examination in 2010-2011 was agreed at £45 for those aged 60 and over or £37 for those aged under 60, and if a supplementary appointment was necessary the practitioner receives an additional £21.50 (Scottish Government Health Directorate, 2008). Free eye examinations through the NHS in Scotland are now also available for everyone in an attempt to make the community optometrist the first point of call for all acute eye problems, and each optometrist is limited to 20 appointments per day with a minimum examination duration of 30 minutes. Scottish optometrists are now more able to refer pathology directly to the hospital, depending on the nature of the condition. Co-managed care between hospitals and community optometry has been encouraged with optometrists able to manage more conditions in their local practices in what is a largely rural country. In conclusion the new contract aimed to increase choice and convenience for the patient while reducing waiting lists and costs for the NHS. The traditional NHS "sight test", primarily designed for those who require spectacles, has been replaced by a comprehensive eye examination appropriate to a patient's needs, symptoms and general health and which may not necessarily include a refraction (Optometry Scotland, 2006). To aid with upgrading examinations to the required standard an equipment grant of £8000 was allocated to every optometric practice in Scotland, and mandatory training was provided (Optometry Scotland, 2006). Similar co-management schemes between GPs and Optometrists have also been previously trialled in England and have generally proved to be successful (Austen, 2003).

1.1 Referral Patterns and reasons for referral

Although specifying a reason for referring a patient, in terms of a tentative diagnosis, is not a legal requirement, nearly all referrals from optometrists include this information. As previously mentioned referrals are triaged for urgency by an ophthalmologist in secondary care, therefore clinical findings in the referral are important. From this information a number of studies have ascertained the mix of pathologies being referred to the hospital eye service. It is natural to expect that the most prevalent reasons for referral to the HES will differ according to whether the referral source is optometric or medical practice.

Cataract and lens disorders have been repeatedly identified as being the most common reason for referral by optometrists (Port and Pope, 1988, Pooley and Frost, 1999, Lash, 2003, Pierscionek et al., 2009). Glaucoma is usually the second most commonly referred pathology (Port and Pope, 1988, Pooley and Frost, 1999, Lash, 2003) and was found to be more common than cataract by one study (Harrison et al., 1988) but the third most common, after retinal problems and cataract, by another (Pierscionek et al., 2009). The most commonly referred condition by GPs was disorders of the lids and adnexae (Pierscionek et al., 2009, Pooley, 1996, Harrison et al., 1988).

Pierscionek et al.'s results are somewhat surprising, with the proportion of glaucoma referrals being less than expected for optometrists, but more than expected for GPs (31%). The remainder of the literature has reported values of 0-10% of referrals originating from GPs (Pooley and Frost, 1999, Harrison et al., 1988). This could be explained by the fact there is no mention of GP referrals that

originated in optometric practice but did not enclose the optometrists letter in Pierscionek et al.'s paper. This has been cited as a significant issue previously (Pooley and Frost, 1999) and has been reported to be as high as 22% of referrals (Pooley, 1996). Other explanations could include the presence of a GP with special interest in ophthalmology within the catchment area, although this was not specified in the report or acknowledged as being different from previous studies. Another difference worth highlighting is the absence of any referrals from GPs for binocular vision disorders in the Pierscionek study, whereas previous studies have found GPs to be the main source of these patients (Harrison et al., 1988, Pooley and Frost, 1999).

1.2 Referral Quality and forms used

Pooley and Frost assessed the quality of 172 referrals by Optometrists during 1997 (Pooley and Frost, 1999). The content of the referrals was described in part, with all of them containing visual acuity measurements and 51% containing a diagnosis. A number of referrals had legally mandatory details missing; 6 (~3%) of the referrals had no practitioner name or address, and 30 (17%) had neither patient age nor date of birth. The other legally required information that was frequently missed was the date, which was missed on 48 (28%) referrals, 47 of these were on GOS18 forms. In total, 69% of the referrals were on a GOS18 form, with the remainder using letters or other forms, for example practice-specific forms. 29% of these GOS18 forms had signatures confirming that the patient consented to their optometrist receiving a copy of all correspondence. In addition to missing dates, the authors identified another problem with all forms being the very poor legibility of optometric

referrals, with GP referrals being significantly more legible. The reason for this was not discussed, but it could be partially attributed to the GOS18 form, which is a triplicate carbon copy form, whereas GP handwritten referrals were on paper and not carbon copy forms. Any further content from the referrals was not commented on.

An analysis of 444 GOS18 forms was performed by Lash in 2003 although the main focus of the paper was directed towards cataract and glaucoma (Lash, 2003). The presence of the following details from the referral was recorded: practitioner name, practitioner address, date of birth, spectacle prescription, visual acuities, intraocular pressures, cup:disc ratios, visual fields and patient consent. The reason for referral was used to allocate the referral to one of ten categories with the majority of referrals being due to cataract (37%) or glaucoma (18%). Only 7% of the cataract referrals included information regarding an effect on patient lifestyle and willingness for surgery, which are both deemed as being essential according to "Action for Cataracts" published by the Department of Health in 2000 (2000). Legally mandatory information was also absent on some referrals as 31% had no legible optometrist name and 6% had no practice address. Signed patient consent was obtained in only 5% of referrals, which could be in part due to lack of attention, but also due to it not increasing the likelihood of getting a response from the hospital (Whittaker et al., 1999). It would have been of particular interest to see if the referral visual acuities and intra-ocular pressures (IOP) differed from those taken in hospital.

As a follow up to the study above, Lash investigated the quality of 412 referrals for cataract in 2006 (Lash et al., 2006). Three referral techniques were compared; GOS18, letter and direct referral using a purpose designed form. All 143 referrals made directly included the three parts of information deemed essential by the Department of Health in 2000 (2000), namely 'Presence of cataract', 'Effect on lifestyle' and 'Willingness for surgery'. Only 10% of GOS18 forms and 17% of letters contained this information with 72% of GOS18 forms and 76% of letters merely confirming the presence of cataract with no mention of symptoms or surgery. It is clear that direct referral is associated with higher quality referrals, and indeed most direct referral schemes pay an extra fee to optometrists to discuss the risks and benefits of surgery and submit all the necessary referral information.

Tattersall and Sullivan investigated the appropriateness of 412 cataract referrals originating from Optometrists (Tattersall and Sullivan, 2008) but did not explicitly reveal the content of the referrals. The authors did state that there was no association between lifestyle changes being recorded on the referral, and the final outcome of the referral, which suggests that some referrals must not have commented on effect on lifestyle but contradicts the conclusions made by Lash and colleagues (Lash et al., 2006).

The quality of glaucoma referrals was investigated by Scully and colleagues (2009) who also found that a large proportion of referrals did not include legally mandatory non-clinical information, such as referral date (45% of referrals), patient date of birth (32%) and referrer name (26%).

The GOS18 form, despite being designed specifically for the purpose of optometric referrals, seems to have a number of flaws and frequently results in an unacceptable proportion of poor quality and unlawful referrals. Its triplicate carbon-copy handwritten nature reduces the copy legibility compared to a word processed letter or a photocopied handwritten letter. Lash and colleagues (Lash et al., 2006) concluded in 2006 that, when compared to letters or direct referrals, patients referred by GOS18 were less likely to be listed for surgery. The GOS18 referrals were also less likely to contain as much essential information, this could be due to the lack of space for optometrists to write in which a significant number of optometrists commented on in a survey by Whittaker and colleagues (Whittaker et al., 1999). Despite these flaws, and authors advocating the use of letter writing (Clarke, 2008), the GOS18 is still used for the vast majority of referrals (Pooley and Frost, 1999, Scully et al., 2009).

The space for the patient to sign to consent to disclosure of medical information to the optometrist has been shown by Whittaker (Whittaker et al., 1999) to have a perverse effect on whether optometrists receive any reply from the hospital, with 12% getting a response if the form was signed and 17% getting a response if the form was unsigned. Communication of medical information between clinicians is lawful without patient consent, if it is in the interests of the patient, therefore the patient signature space is unnecessary. It could be replaced by a request to ophthalmologists to copy correspondence to the optometrist, or abandoned

completely, which is the route taken by some locally adapted GOS18 forms (Naru and Green, 2009).

In Bradford and Airedale the locally adapted GOS18 form was developed by Naru and Green in collaboration with local ophthalmologists and GP alliances (Naru and Green, 2010). In addition to the lack of patient consent, there was also no section for the GP to complete as local alliances suggested that this was not used. This freed up a significant amount of space which was filled by a number of tick boxes for the most commonly referred ocular pathologies, allowing the triaging ophthalmologist to quickly determine what urgency to allocate. To address the other problems of missing legally required information, there are more prominent areas for referrer name, referrer signature and date. The problem of poor legibility and self carbon replicating remained, therefore in July 2010 an electronic version of the form was made available, which can be completed on a computer and printed, but in all other respects is the same (Naru and Green, 2010). As computer generated letter writing does not seem to have been adopted by the majority of optometrists as their referral method of choice, perhaps this is an adequate compromise that will improve the quality, legibility and legality of referrals.

1.3. Referral accuracy

A false positive is commonly known as a false alarm. In healthcare, the occurrence of a false positive generally means a patient has been initially diagnosed with a disease, typically at primary healthcare level, that turns out to be absent after referral for further investigation, typically at secondary healthcare level. Conversely a true positive is a patient whose diagnosis is confirmed as correct. More false positives are generated when the initial diagnosis is influenced by diagnostic techniques with low specificity that are poor at identifying normals. Some studies calculate a Positive Predictive Value (PPV) to show levels of false positives. The PPV is simply the number of true positives divided by the sample size, thereby giving the probability of a referral being a true positive.

1.3a. Accuracy of referrals to Ophthalmology for all ocular pathology

Only two studies have reported what proportion of referrals for all pathologies were correct, and what proportion were false positive (Pooley and Frost, 1999, Harrison et al., 1988). There is, however, a large body of research covering the levels of correct and false positive referrals for glaucoma patients, which is covered in detail below. Most research into referral accuracy in Ophthalmology departments is performed retrospectively as an audit. Audits are performed simply by recalling all the records from a certain period of time from the archives, and extracting relevant information from them. Retrospective studies are usually easier to perform than prospective studies as they generally require no additional work from

examining clinicians, and less planning. Difficulties arise when the records are ambiguous about final diagnoses or more clarification is needed. Occasionally patient records have also been known to be incomplete or even missing and this introduces inconsistencies in these data. Prospective studies can record similar information to retrospective studies, but the examining clinician and patient are aware of the study prior to the appointment. It is possible that this could bias the outcome of the appointment, however it means that the clinicians' input can be sought to record the results as accurately as possible. The data are obtained during, or immediately after the appointment, which minimises the time during which the record can be lost.

Pooley and Frost (1999) conducted a retrospective analysis of referral correspondence to St Georges Ophthalmic Department and Sutton Eye Department during a three month period in 1997. At this time, prior to Statutory Instrument 3267 (1999), optometrists were required by the Opticians Act 1989 to refer every patient with any ocular abnormality, even if referral was not justified in their professional opinion. 433 patients were included in the study, of which 172 were originally referred by optometrists and 190 referrals appeared to originate from GPs. An additional 28 referrals were from ophthalmic practitioners of some sort, but it was impossible to identify whether it was from an optometrist or Ophthalmic Medical Practitioner (OMP), generally due to their correspondence not being included with the GP's letter. The authors analysed the quality and the contents of the referral letters and this is discussed in more detail under the relevant heading below.

The accuracy of referrals originating from optometrists was assessed after the main part of the study using a separate cohort of 89 patients attending St. Georges Hospital only. The authors do not state why a separate cohort was used although it was perhaps to allow a prospective evaluation using a single ophthalmologist. However whether it was performed retrospectively or prospectively is not explicitly explained. Using a separate group of referrals gathered over a shorter period of time to the first cohort, and from one hospital only, introduces an unnecessary inconsistency when evaluating all the results. There are also fewer diagnosis categories than used earlier in the study, making comparison difficult. This is probably due to the fact that fewer referrals were included. This lack of numbers critically reduces the reliability of any conclusions made in the study. Examples of accuracies obtained range from 0% correct for 'visual disturbance' and 'nerve and visual pathways' (a total of 4 patients and 3 patients respectively) to 100% correct for five pathology categories, all with a total of three patients or fewer. 26% of the 19 glaucoma referrals were finally diagnosed with glaucoma.

Pooley and Frost (1999) challenged the necessity of optometrists having to refer patients to the HES via their GP unless in an emergency, arguing that it was of low added value and many GP appointments could be made available by removing this step. The authors also identified that vital correspondence from the referring optometrist would be guaranteed to be received by the HES and delays would be reduced. The authors raised the point discussed earlier that the optometrist was not legally required to do any more than make the decision to refer. This may have reduced both the accuracy and quality of referrals, despite College of Optometrists guidelines at the time stating as much information should be included in the referral

as possible (1991). The authors acknowledged that referral accuracy should have improved with the introduction of SI 3267, which was being finalised as the article was written. An important conclusion was made that at the time there was a financial disincentive for optometrists to improve referrals, they were not remunerated for further clinical tests and the sight test fee was paid whether the referral was necessary or not. In light of this study the local health authority undertook a programme of incentives to improve the referral process, including guidelines and training, but the financial disincentive remained.

A larger study was conducted in 1986 and 1987 by Harrison and colleagues (1988) in response to the possible impending abolishment of the NHS sight test for all patients. The medical records of 1113 new patients referred to a consultant ophthalmologist were reviewed retrospectively, of which 49% appeared to be from GPs and 39% from optometrists. The majority of referrals were accurate, with an overall false positive rate of 25% for referrals from optometrists and 40% from GPs. These accuracies were calculated simply by comparing the first and second referral diagnoses (if present) with the final hospital diagnosis. 120 patients were referred by optometrists for suspected glaucoma and of these 96 (80%) were either subsequently diagnosed with the condition or required further ophthalmological follow up. The authors were primarily concerned that 16% of patients had asymptomatic disease, and would be unlikely to attend for eye examinations if they were not free through the NHS. As patients over 60 are now entitled to an NHS sight test, it would be interesting to know what proportion of this group with asymptomatic disease were over 60.

A number of studies have looked at the accuracy of cataract referrals, specifically to evaluate whether an enhanced service involving direct referral would increase the proportion of those referred going on to have surgery. In addition to direct referral, which avoids the GP having to see the patient (and the patient having to attend the GP's practice), these enhanced services usually involve the optometrist offering a choice of providers to the patients and completing a proforma to gather information recommended by the Department of Health paper; Action on Cataracts (2000). This includes noting the presence of cataract, the effect on the patient's lifestyle and the patient's willingness for surgery. The latter would involve a discussion of the pros and cons of surgery. As this is a form of referral refinement, and not part of the GOS contract, the referring optometrist receives a fee which varied between £35 and £44 in 2004 according to the Association of Optometrists (2004a). Perhaps unsurprisingly all studies found that the direct referral service resulted in an increased proportion of referred patients being listed for surgery and therefore fewer false positives (Lash et al., 2006, Park et al., 2009, Newsom et al., 2005, Sharp et al., 2003).

1.3b Accuracy of referrals for Glaucoma

The inaccuracy of referrals to Ophthalmology for suspected glaucoma has repeatedly been studied by various groups over the past twenty-five years. The vast majority of studies have identified this inaccuracy as a cause for concern, and many authors have made suggestions to improve accuracy. The National Eye Care Steering Group's first report (2004b) recommended changes to the current glaucoma referral pathway in an attempt to utilise the expertise of community

Optometrists in monitoring suspect glaucoma cases that do not need treatment. The report stated that local PCTs would have to approve and fund individual schemes as there would be no change to the General Ophthalmic Service. Due to the relatively high prevalence in the UK, a national screening programme for glaucoma has been suggested in the past, however a recent systematic review and cost-evaluation for a glaucoma screening programme deemed that it would not currently be cost effective for the whole population (Burr et al., 2007) but may be if targeted at high risk groups.

Frequently studies present the data in differing ways, and have different definitions for a positive referral; therefore their results are not directly comparable. Each study has been individually summarised below in chronological order, along with its definition of a positive referral. Although we are primarily concerned with false positive referrals, and their causes, most of the studies have presented their data in terms of proportion of true positive referrals, and so that is how it has been summarised below. Table 1.1 and Figure 1.2 attempt to show the change in accuracy of glaucoma referrals over time, along with the definition of a correct positive referral used for comparison. Where many different diagnosis categories were given an attempt has been made to standardise the definition to 'diagnosed with glaucoma or follow up required' to allow graphical comparison. Some studies give only the final diagnoses and no indication of whether follow up was required and therefore may give high false positive levels. Equally, sometimes essential diagnostic tests, for example visual fields, may not be done at the first appointment due to staffing constraints. Therefore the positive definition of 'diagnosed with glaucoma or follow up required' potentially gives artificially low false positive levels.

The ideal false positive definition would be based on whether the examining ophthalmologist believes that the referral was warranted and is therefore dependent on the assumption that the examining ophthalmologist is the gold standard. This is sometimes impossible to ascertain in a retrospective study. The development of the ideal false positive definition for HES referrals is discussed in more detail in chapter 3.2.

A small early study by Clearkin and Harcourt (1983) identified 43 new referrals for glaucoma to Leeds General Infirmary during the 12 month period from May 1980 to April 1981. Chronic open angle glaucoma was confirmed in 9 patients (21%) and normal tension glaucoma was confirmed in 2 (5%). Of the patients with suspect glaucoma, 2 had ocular hypertension and 2 had suspicious discs. 28 (65%) patients had no suspicion of glaucoma and were therefore false positive (Clearkin and Harcourt, 1983). The authors also revealed that all the patients finally diagnosed with glaucoma were originally referred by Ophthalmic Opticians, with all nine referrals from GPs turning out to be false positive.

Brittain and colleagues (1988) used a prospective survey specifically to determine the diagnostic accuracy of glaucoma referrals to Leicester Royal Infirmary. Ophthalmologists examining cases of suspected glaucoma were required to complete a small questionnaire after every appointment for a period of six months in 1987. The questionnaire required the clinician to classify each patient as; Glaucoma confirmed, glaucoma suspected (follow up required), no evidence of glaucoma or missed glaucoma (cases referred with another diagnosis but found to have glaucoma). 93 referrals for suspected glaucoma were received from

Optometrists, 29 from GPs and three from other sources (Ophthalmic Medical Practitioners or A&E). Referrals were found to be 66% positive (glaucoma or follow up required) from Optometrists and 21% positive from GPs (Brittain, 1988).

Brittain concluded that as almost two thirds of referrals were justified, Ophthalmic Opticians in Leicestershire appeared to diagnose glaucoma quite accurately. In 1987 a significant number of patients were referred directly by the GP with suspect glaucoma. This study was critical of this referral route for chronic open angle glaucoma, as GPs do not have the equipment or generally the expertise to accurately diagnose this condition.

Tuck (1991) conducted a prospective survey of optometrists from a 5% sample of socio-economically representative practices from England and Wales on behalf of the International Glaucoma Association. The study lasted six months, spanning 1988 and 1989 with 1103 referrals being included in total. 71.4% of referrals were classified as correct as they were diagnosed with glaucoma or required follow up. As this study did not contact the hospital directly, in the cases where no correspondence was received by the optometrists from the hospital, the patient had to be asked what the result of the consultation was, which likely introduced errors. However, if only the referrals with replies are included (n=704) the true positive rate was still very similar at 71.7% (Tuck and Crick, 1991).

A nine-month prospective study was performed from October 1990 to June 1991 at the Ophthalmology Department at Leicester Royal Infirmary. This study was partially in response to the abolition of free Optometric eye examinations in April 1989 and was attempting to determine any resultant change in referral patterns. In

total 213 referrals were seen and of these 31.5% were diagnosed with glaucoma, 22.5% were diagnosed with OHT, 29% had no abnormality detected and 17% had other diagnoses (Sheldrick et al., 1994). The authors commented that abnormalities of the optic disc and/or visual field were reported by the optometrist in only a minority of referrals. New guidelines for referral when abnormalities are suspected were suggested as a way to improve referral accuracy. A final point was made that only 35% of those with glaucoma and 31% of ocular hypertensives were eligible for free eye examinations. Old age, the biggest risk factor for glaucoma, was not grounds for a free sight test in 1991.

Bell and O'Brien (1997b) performed a retrospective cohort study of all referrals for suspected glaucoma between October 1993 and March 1994 at the Royal Infirmary of Edinburgh. They wrote two papers from the study and published them at the same time in the same journal. The first paper gave in depth diagnosis details along with definitions for the most common diagnostic definitions used. In total 271 referrals were evaluated and the diagnostic groupings in order of size were 42% ocular hypertension, 29% discharged, 17% glaucoma and 5% with large physiological cupping (Bell and O'Brien, 1997b). Sixty-one of the OHT patients, 22.5% of the total cohort, were placed on treatment to reduce the likelihood of progression to glaucoma. The proportion of patients diagnosed with glaucoma was significantly less than in previous studies whereas relatively more patients had OHT. The authors suggested an explanation of either stricter diagnostic criteria, or Optometrists becoming more prone to refer at an earlier (OHT) stage of disease. A shared care approach with Optometrists was again suggested for lower risk OHT patients.

The second paper gave a detailed analysis of the diagnostic tests performed in primary care, and their comparison to the hospital examination. The accuracy of referrals was compared between those who only measured IOP (17%), those who measured IOP and Visual Fields (7%), those who measured IOP and performed disc assessment (26%) and those who measured all three (48%). The authors therefore recommended that all three tests be performed prior to referral, and that any abnormality needed retesting. If raised IOP was found with non-contact tonometry, repeat measurement with applanation tonometry was suggested. Indeed, the authors believed applanation tonometry should be more widespread in optometric practice. Finally, improved quality of visual field testing and interpretation was also advised (Bell and O'Brien, 1997a).

Vernon documented how referrals had changed in the five-year period between 1988 and 1993 at Queens Medical Centre in Nottingham by retrospectively analysing the records of every referral to the glaucoma clinic in 1988 and 1993 and comparing the findings. 95 referrals were seen in 1988, of which 75 had a GOS18 form enclosed and could be identified as originating from an Optometrist. Vernon defined a 'positive' screen as having glaucoma or treated Ocular Hypertension (OHT), and on this basis 56% were 'positive' in 1988. A more detailed breakdown of the diagnoses was also given: 48% Glaucoma, 8% OHT (treated), 28% OHT (not treated), 12% suspicious disc, 3% normal, 1% other disease. A random sample of 95 referrals from 1993 were analysed, of these 71 has a GOS18 form enclosed and could be identified as originating from an Optometrist. The referral accuracy was only 37% positive, with a breakdown of 34% glaucoma, 3% OHT (treated), 20% OHT (untreated), 15% suspicious disc, 23% normal and 5% other disease. Other

significant differences found between the cohorts were that the presenting IOP was lower, and the number of visual field tests performed by Optometrists had increased in 1993. Incidentally, the average IOPs of patients diagnosed with glaucoma were not significantly different. Vernon concluded that there was a significant reduction in referral accuracy between 1988 and 1993. As the presenting IOP also reduced over this period, the author suggested that Optometrists were relying more on disc assessment and Visual Field screening. He also observed that in 1993 many Optometrists arranged for field screening to be performed prior to the eye examination, frequently by a colleague, and therefore Ophthalmoscopy may have been influenced by the results of field screening (Vernon, 1998).

As a result of the previous study glaucoma referral guidelines were distributed to all Optometrists in the catchment area of Queens Medical Centre, Nottingham. A second study attempted to assess the impact of this intervention on referral accuracy. An audit was performed on the records of all new glaucoma referrals for the 12 months preceding intervention (1997, 105 referrals). These results were compared to 12 months after intervention (mid 1998 to mid 1999, 102 referrals). The same definition of a positive screen was used, glaucoma or treated OHT, giving an accuracy of 40% 'positive' in 1997. The proportion of each diagnosis in detail was 37.2% glaucoma, 2.8% OHT (treated), 13.3% OHT (not treated), 8.6% suspicious disc, 0.9% other diagnoses and 37.2% normal. After the intervention the proportion of positive screens had dropped to 32.3% with a breakdown of 31.4% glaucoma, 0.9% OHT treated, 17.6% OHT not treated, 6.8% suspicious disc and 43.1% normal (Vernon and Ghosh, 2001). The authors therefore concluded that dissemination of glaucoma screening guidelines did not appear to have improved diagnostic accuracy

of Optometrists. The data suggest that after guidelines were introduced the Optometrists generated fewer false positive referrals on grounds of IOP, but significantly more on grounds of visual field and optic disc interpretation.

Newman, Anwar and Jordan (1998) completed a three month retrospective study at West Suffolk Hospital in 1994, with the aim of assessing the positive predictive value (PPV) of visual field testing by optometrists. 86 referrals were received for suspected glaucoma, all initiated by Optometrists, of which 82 patients attended. 40% of these patients were diagnosed with glaucoma, 20% with OHT, 2% with NTG, and 1% with pigment dispersion syndrome. 30% had no abnormality and 7% had other pathology unrelated to glaucoma. After analysing the diagnostic tests performed by optometrists, the authors concluded that ophthalmoscopy had poor validity as a screening test for glaucoma. Tonometry was more effective, but limited by a variable cut-off. Visual field assessment had similar limitations but had the greatest efficacy as a screening test. Again it was concluded that using visual field testing, ophthalmoscopy and tonometry in combination was best. The authors also believed that the optometrists may not have been following validated screening methodology when performing and interpreting visual field tests (Newman et al., 1998).

A study was initiated by Theodossiades and Murdoch (1999) at Moorfields community eye clinic at Ealing Hospital to determine the PPV of optometric referrals for suspected glaucoma. There were 87 eligible referrals received between September 1996 and February 1997. Of these 22% had glaucoma, 21% were glaucoma suspects and 57% received a negative diagnosis, consisting of 14% OHT,

5% large discs, 3% other diagnoses and 36% no abnormality (Theodossiades and Murdoch, 1999). Similar to previous studies, the authors found that the highest PPV was demonstrated when information on IOP, optic disc assessment and visual fields was included in the referral.

In an attempt to improve the poor PPV observed in the previous study, an intervention trial was performed by Theodossiades and colleagues from August 1999 to March 2001. Thirty-two local practices were chosen as the study population, half acted as controls, and half were offered training in optic disc assessment, given standardised referral criteria and feedback from ophthalmologists. Optometrists taking part in the intervention referred 210 patients compared to 119 from the control group. A similar definition of a positive outcome was used as in the last study. Specifically 'a confirmed or suspected diagnosis of glaucoma' where a suspect was 'a patient felt to warrant repeat examination or follow up'. Referrals from the intervention group resulted in 49% with a positive outcome, and the control group had a 46% positive outcome. More detailed data on diagnoses were not released. The authors hailed the intervention as a success because, although accuracy was not significantly improved, the PPV remained the same yet significantly more patients were referred. They also concede that a larger trial was required to provide conclusive evidence of an effect. (Theodossiades et al., 2004).

The intervention was subsequently extended to all optometric practices in the area. Patel and colleagues examined the effect of this by analysing the result of every glaucoma referral, 376 in total, between June 2002 and May 2003. This was

significantly more than an equivalent 12 month period in the original study. The same definitions were used as the previous studies, resulting in positive outcomes for 45% of referrals. Referrals from optometric practices that were part of the original intervention group were 51% positive, whereas from the original control group they were 41% positive. Referrals from practices outside the control area were 37% positive and from unidentified optometrists, 44% positive (Patel et al., 2006). The authors concluded that the intervention had a positive effect on numbers of referrals for glaucoma, but referral accuracy remained the same. This therefore had a positive effect on the total number of patients diagnosed with glaucoma during the 12 month period.

Bowling and colleagues (2005) initiated a ten-year long prospective study, currently the largest of this nature, in 1994. A proforma was used to collect data relating to the first appointment of every new referral from optometric practice for suspected glaucoma until 2004. This resulted in data from 2505 referrals. Each case was allocated a provisional diagnosis from one of five groups based on the initial HES assessment. 45.8% were reported to have no glaucoma or OHT, 20.4% Glaucoma, 6% Normal Tension Glaucoma, 29.8% OHT and 5% glaucoma suspect (Bowling et al., 2005). The authors felt that although the false positive rates were high, Optometrists were under pressure to detect every case of glaucoma, and frequently referred patients with OHT who only required monitoring.

A retrospective audit by Salmon and colleagues (2007) at Oxford Eye Hospital covered three years of new glaucoma referrals. Results were given separately for each year, with the positive referral rate in 2003 being 28% (n=278), in 2004 it was

54% (n=418) and in 2005 it was 53% (n=410). A positive referral was again classified as the patient being diagnosed with glaucoma or requiring ophthalmological review. Similarly to previous studies, the authors concluded that is was essential to combine tonometry with ophthalmoscopy and perimetry, as most false positives came from referrals with only one abnormal parameter. Approximately one-fifth of the false positive referrals had been referred with suspiciously high IOPs, yet had normal IOPs when measured by Goldman tonometry in the hospital. The authors therefore agree with Bell and O'Brien's (1997b) view that optometrists should repeat high IOP measurements with applanation tonometry. The authors concede that optometrists are under considerable pressure to detect every case, however they conclude that false positive referrals could be significantly reduced.

To summarise, all studies have identified the accuracy of glaucoma referrals as an area for potential improvement. When considering all studies since 1980, Figure 1.2 and Table 1.1 do not appear to show a trend, however they may show a positive trend over the studies published since 1990. Another review of the literature in the near future, as more data are published from studies after publication of the NICE glaucoma guidelines, would be of interest.

Table 1.1. Summary of studies that have reported Glaucoma referral accuracy since 1980

Authors	Location	Start	End	Duration	Referrals	%Correct	Definition of correct (true positive) used	Notes
Clearkin & Harcourt	Leeds General Infirmary	1980	1981	12 months	43	35%	Confirmed glaucoma or observation required	
Harrison, Wild & Hobley	Burton General Hospital	1986	1987	14 months	120	80%	Diagnosed with glaucoma or follow up required	Discussed in Chapter 1.1.3a
Brittain	Leicester Royal Infirmary	1987	1987	6 months	93	66%	Diagnosed with glaucoma or follow up required	
Vernon	Queens Medical Centre, Nottingham	1988	1988	12 months	75	56%	Glaucoma or treated OHT	
Tuck & Crick	England and Wales	1988	1989	6 months	1103	71%	Diagnosed with glaucoma or follow up required	
Sheldrick et al.	Leicester Royal Infirmary	1990	1991	9 months	213	32%	Diagnosed with glaucoma	
Vernon	Queens Medical Centre, Nottingham	1993	1993	12 months	71	37%	Glaucoma or treated OHT	
Bell & O'Brien	Royal Infirmary, Edinburgh	1993	1994	6 months	271	64%	Treated for glaucoma or OHT or observed.	
Newman et al.	West Suffolk Hospital	1994	1994	3 months	82	40%	Diagnosed with glaucoma	
Theodossiades & Murdoch	Ealing Hospital	1996	1997	6 months	87	43%	Confirmed diagnosis of glaucoma or follow up required	
Vernon & Ghosh	Queens Medical Centre, Nottingham	1997	1997	12 months	105	40%	Glaucoma or treated OHT	
Pooley & Frost	St. Georges Hospital & Sutton Eye Dept.	1997	1997	1 month	19	26%	Diagnosed with glaucoma	Discussed in Chapter 1.1.3a
Vernon & Ghosh	Queens Medical Centre, Nottingham	1998	1999	12 months	102	32%	Glaucoma or treated OHT	After intervention.
Bowling et al.	Oxford Eye Hospital	1994	2004	10 years	2505	54%	Glaucoma, NTG, OHT or glaucoma suspect	
Theodossiades et al.	Ealing Hospital	1999	2001	20 months	210	49%	Confirmed diagnosis of glaucoma or follow up required	After intervention (training, referral criteria and feedback).
Theodossiades et al.	Ealing Hospital	1999	2001	20 months	119	46%	Confirmed diagnosis of glaucoma or follow up required	Control group – no intervention.
Patel et al.	Ealing Hospital	2002	2003	12 months	376	45%	Confirmed diagnosis of glaucoma or follow up required	Intervention extended to all Optometrists in area.
Salmon et al.	Oxford Eye Hospital	2003	2003	12 months	278	48%	Diagnosed with glaucoma or OHT	
Salmon et al.	Oxford Eye Hospital	2004	2004	12 months	418	54%	Diagnosed with glaucoma or OHT	
Salmon et al.	Oxford Eye Hospital	2005	2005	12 months	410	53%	Diagnosed with glaucoma or OHT	

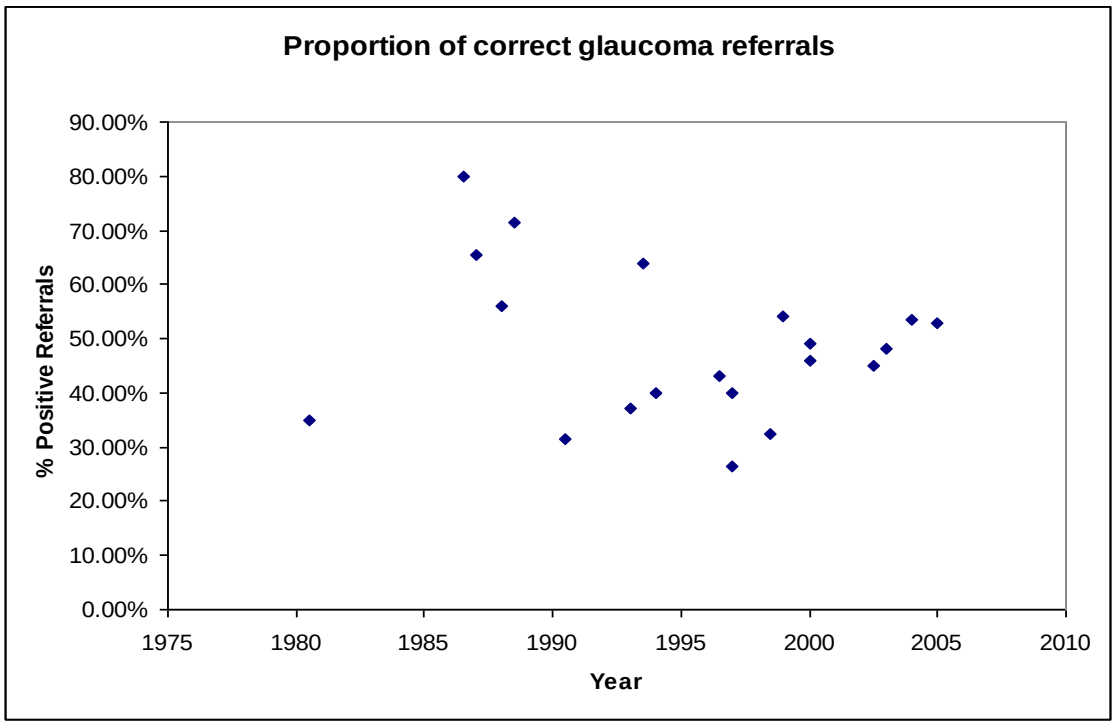

Figure 1.2. The proportion of correct (true positive) glaucoma referrals as reported by the literature between 1980 and 2005.

1.3c Reasons for false positive referrals

i) Money (and time)

A recent study by Myint and colleagues (2010) found that lack of time to repeat measurements, or remuneration for doing such, as the most commonly reported barriers to effective glaucoma detection in the UK. Fewer optometrists in Scotland reported this, which is coincident with a study by Ang and colleagues (2009) investigating the effect of the new GOS contract in Scotland on glaucoma referrals. This study found that after the introduction of the new contract, which is discussed

in section 1.1.1, there was a significant reduction in false positive referrals and a significant increase in true positive referrals.

ii) Litigation

It is true that litigation against clinicians in the UK is steadily increasing (The National Health Service Litigation Authority Report and Accounts 2010–2011), and therefore the sensitivity of optometrists to the risk of litigation may possibly have increased in line with this. This trend is not reflected by an increase in numbers of Fitness to Practice complaints, which have been variable in recent years (GOC annual reports, 2009-2010:158, 08-09: 150, 07-08: 172, 06-07: 129, 05-06: 146). Fear of litigation, and an increase in modern screening equipment in practice (Myint et al., 2011), may increase the likelihood of Optometrists screening their patients for as many pathologies as possible whereas the decision to screen should take into account risk factors and the social cost (Stewart-Brown and Farmer, 1997).

iii) Optometrists: Their accuracy of measurement and clinical decision making

A small number of published papers have assessed the quality of optometrists' clinical decision making, two did so retrospectively (Banes et al., 2006, Ho and Vernon, 2011), and most prospectively (Azuara-Blanco et al., 2007, Hau et al., 2007, Banes et al., 2000, Spry et al., 1999, Gray et al., 1997, Oster et al., 1999).

An optometry led new patient clinic at Moorfields was the subject of a six month study by Oster and colleagues (1999). Although few data are published, a consultant ophthalmologist assessed 152 patients that were seen the same day by a trained hospital optometrist (exact level of experience was not specified). Measurement

validity was not compared, but diagnosis was, for which accuracy was high at 79.6%, with only 3.3% of patients being incorrectly assessed, and 17.1% were partially correct (slight errors or misclassifications).

Hau and colleagues (2007) were concerned only with accident and emergency patients and, despite being a prospective study, only compared the ability to diagnose and manage correctly and did not include validity of measurement (i.e. whether the clinical measurements obtained by optometrists were correct according to the gold standard, a consultant ophthalmologist). Patients were assessed by one of two optometrists (both had >3 years extended role hospital experience) and immediately afterwards assessed by one consultant ophthalmologist. A high level of agreement in diagnosis and management between the two clinicians was found, at 89.3% and 79.3% respectively (n=150 patients). Although immediate reassessment minimised the likelihood of any signs or symptoms manifesting, the accuracy of individual measurements (i.e. the difference between the optometrist's measurements and the ophthalmologist's measurements) was not published.

The remaining studies were concerned only with glaucoma patients. Gray (1997), Spry (1999) and colleagues published only the reliability and validity of various clinical measurements relating to glaucoma during the Bristol Shared Care Glaucoma Study (Spencer et al., 1995). This involved patients being seen by at least three clinicians within two months, an optometrist in the community, an ophthalmologist at the HES, and members of the research team as part of a comprehensive Gold Standard reference assessment. Neither study found

significant differences in measurement or outcome across the three groups, and therefore concluded that trained community optometrists can reliably take measurements and make management decisions.

The study by Banes and colleagues (2000) also implied that clinical measurements were performed by the optometrist and repeated by an ophthalmologist, but does not specify the length of time between. Both clinical measurements and decision making are recorded and compared. Results were again very good, with high measurement accuracy and 98.9% agreement for management recommendations. In response to the new glaucoma referral pathway in Scotland, Azuara-Blanco and colleagues (2007) assessed the accuracy of the diagnosis and management decisions made by accredited glaucoma optometrists in community. Measurements were repeated by an ophthalmologist within one month, clinical decisions were compared, but unfortunately measurement accuracy was not published. Although unlikely, a significant time gap between repeating the measurements may allow signs or symptoms to develop. Agreement in diagnosing glaucoma and treatment recommendation was again substantial, at 89% and 88% respectively.

Banes and colleagues (2006) used a 'paper only' method where two consultants reviewed the records of glaucoma patients seen by optometrists and made decisions based on the results recorded. The consultants' decisions were compared to the optometrists' and each others. This has the benefit of allowing a large number of cases to be assessed conveniently, but inaccuracies are introduced as the consultants are limited to the history and measurements recorded by the optometrist. This method is fine for studying clinical decision making, but cannot

quantify the clinical skill of the practitioner or the accuracy of measurements. The authors partially addressed this by asking the optometrists to grade 134 pairs of previously published optic disc stereo photographs as glaucomatous or non-glaucomatous. The level of sensitivity in diagnosing the discs as glaucomatous was 77.8-88.2% and specificity was 76-89%, but unfortunately the consultants were not asked to do the same for comparison. The accuracy of other vital measurements, such as Goldmann tonometry was not validated. Agreement between consultants and optometrists was 79% for medical management, 78% for next appointment schedule, 72-98% for other aspects of patient management but 55% for evaluation of visual field status. Banes and colleagues concluded that within appropriate environments, and possibly with additional visual field evaluation training, optometrists can safely work as part of the hospital glaucoma team.

A recent study found agreement to be greater than 88% for visual field interpretation, medical and surgical management decisions, timing of next appointment and ordering of visual field tests (Ho and Vernon, 2011). The paper's main purpose was to validate decision making for a glaucoma shared care service, and again the optometrists all had at least five years of hospital experience in addition to specialist training provided by a consultant ophthalmologist.

The optometrists in all of these studies were specially trained, with most having worked in hospital clinics for at least two years, and attended consultant ophthalmologist led training sessions. In summary the literature concludes that measurement and diagnostic agreement between optometrists and ophthalmologists is generally good. As all optometrists in these studies were

specially trained, it is not wise to extrapolate these conclusions as evidence supporting similar enhanced services nationwide without incorporating mandatory training into the service protocol. Alternatively, it would be interesting to conduct a similar study using optometrists who were not specially trained, and who relied on core competencies learned as part of the undergraduate course, everyday practice and continuing education.

1.4 Referral Reduction Schemes

As discussed above, many ways of reducing the numbers of false positive glaucoma referrals have been repeatedly identified but dissemination of these has not resulted in improvements. The central problem of being asked to do more work for no remuneration has recently been successfully addressed by varying methods (Parkins and Edgar, 2011, Devarajan et al., 2011, Bourne et al., 2009, Henson et al., 2003, 2006).

The community glaucoma management scheme set up by Henson and colleagues in Manchester was suggested as a way in which levels of false positive referrals could be reduced (Henson et al., 2003). Similar improvements to the glaucoma referral pathway were also suggested in the National Eye Care Steering Groups' first report (2004b).

1.5 Psychological effect of referral

No research has yet considered the adverse psychological effects of false positives generated by referrals from optometrists, but evidence suggests this may be the case in other areas such as screening for congenital conditions and oncology (Stewart-Brown and Farmer, 1997). False positives, and possible adverse effects, are particularly notorious when they are generated by national screening programs, for example, the national screening program for breast cancer. Numerous studies have investigated the levels of false positives generated by screening mammography and their adverse psychological and economic effects, and they are discussed in more detail below.

Although no direct research has yet been performed into the effect of false positive referrals from optometrists, there is proof that depression and anxiety are more prevalent in visually impaired people (Evans et al., 2007). Raised levels of anxiety are also present in patients attending optometric practice for an eye examination (Margrain et al., 2003).

Breast Cancer and other areas

In the UK all women aged 50 to 70 are invited to have mammography performed as part of the NHS Breast Screening Programme. Up to 10% of mammography screens can result in false positive referral for further investigation (Brown et al., 1995), therefore there is a large body of research into the psychological effect of this. It

has been shown that breast cancer screening does not raise anxiety if patients receive a clear result and are not prematurely recalled (Cockburn et al., 1994, Lerman et al., 1991, Brett et al., 2005). Two systematic reviews of the studies investigating the effect of mammography have been performed, one in the UK (Brett et al., 2005) including 54 papers, and one in the USA (Brewer et al., 2007) including 23 papers. The papers used a wide variety of measures of psychological distress, both questionnaires and interview. Some used test-retest strategies to detect the baseline and time scale of any raised levels of distress, and others used a cohort study to compare with controls, or both. The UK review by Brett and colleagues concluded that women definitely experience significant anxiety in the short term and also possibly in the long term (Lerman et al., 1991) up to 18 months after the screening (Gram et al., 1990). The review by Brewer and colleagues was less conclusive, stating that some women may have persistent small effects on psychological well being, and included some studies that showed no effect. In addition to the psychological effect, false positive cancer screens also have an economic impact, with Lafata and colleagues finding the average cost to be $1024 per woman (2004).

Research on the psychosocial impact of screening in other areas is less prevalent. Studies on screening for congenital hypothyroidism (Tymstra, 1986) and pre-natal screening for Down's syndrome (Marteau et al., 1988) both showed there was psychological distress caused to the parents by false positive screens.

1.6 Summary

The literature reviewed in sections 1.1-1.3 reveals that a large number of referrals to the HES are false positive. It has also been shown that there is potentially room for improvement in the current referral pathway which would allow for a reduction of these unnecessary referrals. As almost all of the research in this area has included only glaucoma referrals, the present study aimed to evaluate the quality of referrals across all ocular pathologies, and also to quantify numbers of false positive referrals for all ocular pathologies. Once these data were gathered and analysed, a further aim was to develop an improvement to the present referral pathway in order to reduce numbers of false positive referrals.

Reducing false positive referrals is desirable to reduce wasted NHS resources, but it may also be of benefit in reducing the number of patients who may be having an adverse psychological reaction to being referred. Chapter 1.5 shows that referrals to secondary care in other specialities results in raised psychosocial distress, however it is unknown whether this effect is present in Ophthalmology departments. The present study therefore aims to detect whether referral to the HES also has a similar impact.

Chapter 2. General Methodology

In order to determine referral quality and accuracy a random sample of records of newly referred patients was audited from Bradford Royal Infirmary Ophthalmology Department. More specific details of the audit methodology are given in Chapters 3 and 4. Questionnaires were sent to a random sample of new ophthalmology outpatients in advance of their appointments to elicit whether referral to the HES resulted in raised levels of psychological distress. Questionnaire choice is very important and is discussed in detail below. As the questionnaires had not yet been validated in this population, Chapter 3 attempted to evaluate the suitability of the questionnaires used before Chapter 4 goes on to detect whether referral had a psychological impact on the study cohort. The sampling and methodology is discussed in more detail in the respective Chapters.

2.1 Questionnaire Methodology

In order to measure levels of psychosocial distress, the present study used self report questionnaires (psychometric scales), which are individually discussed later. Questionnaires are very widely used in healthcare and research, however the gold standard measure of psychological effect is an interview conducted by a trained clinician or researcher. Using interviews was obviously desirable but was unfortunately outside the financial constraints of the present study. In addition to using appropriate questionnaires, the reliability of questionnaire based research can be improved by using specific statistical techniques to validate the questionnaires and improve the scoring, such as Rasch analysis (discussed in more detail below) and factor analysis. Factor analysis allows the dimensionality of a questionnaire to

be assessed, i.e. do all questions elicit the same symptom (dimension), or does the questionnaire assess more than one symptom. If a questionnaire has more than one dimension then the questions eliciting different symptoms should be separated and analysed separately.

2.2 Rasch analysis

Rating scale questionnaires contain a number of questions, with a number of response categories for each and each response category usually has a different score. This technique is called a Likert scoring scale and the scores for each response category are usually the same for every question (item). Unless every item elicits exactly the same level of symptom, this does not seem valid. For example, on a depression questionnaire that has a 4-point Likert scoring scale (0 to 3), if a depressed patient fully endorses an item "I feel suicidal" and also the item "I laugh less than I used to", they will receive the same score of 3 for both. Yet surely one refers to more severe symptoms than the other. Questionnaires have items measuring different points lying along a symptom continuum.

Rasch analysis is an item response theory (IRT) based model that attempts to address this problem. Rasch analysis uses the responses from all patients for all items to rank the questions in terms of symptom severity (sometimes termed “item difficulty”) and then also to rank the patients according to the severity of symptoms they exhibit. This provides an item map (e.g. Figure 5.1), which shows all items and participants ranked on the same continuum according to symptom severity, with the top of the graph being the most severe. From the item map it is easy to see ‘at a glance’ whether the questionnaire elicits an appropriate level of symptom severity

for the population. If all the items are towards the top of the map, and participants towards the bottom, the questionnaire has questions that are too severe (or too difficult) which is also known as a floor effect. If the items are at the bottom and the participants are at the top then the questions are too easy and a ceiling effect is present.

Subsequently misfitting items can be identified, which do not conform to the model, for example, an item which patients with mild symptoms (according to the general trend demonstrated by their answers to the other questions) endorse but patients with severe symptoms do not. Infit and Outfit are terms that identify how well an individual item fits the Rasch model. Items with values greater than 1 exhibit more variability than expected by the model, and if an item is too variable then information from it will be unreliable or may measure a different symptom. Items with values below 1 exhibit less variation in the observed response pattern meaning they are more predictable and if too predictable will provide limited information. The limits used for infit and outfit need to take into consideration both the type of test being used and the sample size (Bond & Fox, 2004, (Linacre, 2003), with 0.6-1.4 being suggested for a Likert survey (Bond & Fox, 2004, Wright & Linacre, 1994) and limits becoming stricter as sample sizes increase over 100 (Linacre, 2003). Infit is the more important measure as it is less affected by the occasional outlier, unlike outfit values (e.g., Bond & Fox, 2004; Pesudovs et al., 2007). Misfitting items can therefore be removed from the analysis in order to improve the questionnaire for this population. Likert based scoring involves a number of response categories being available for the participant to choose from. Another use of Rasch analysis is in identifying the appropriateness of the response categories, for example in a

population with mild symptoms the third and fourth response categories could be combined if the analysis revealed they offer no improvement in patient separation. Indeed, collapsing response categories like this can improve the separation of the patients i.e. the discriminative ability of the questionnaire.

Rasch analysis also contains a form of factor analysis called Principle Components Analysis (PCA) which detects whether a scale is multi dimensional. As stated previously, if more than one dimension is found within a scale then the questions for each dimension should be separated and analysed separately.

2.3 Questionnaire Choice

The battery of questionnaires chosen for use in Chapters 3.3 and 3.4 all have merits and problems, as detailed below. However the combination of the two will hopefully compensate for their individual failings. They are all validated and well established with thousands of published uses between them, allowing us to compare our data with other populations. None of the scales has been used with a population of ophthalmology patients before. Copies of each questionnaire are in appendix B.

2.3 i) Hospital Anxiety and Depression Scale

The Hospital Anxiety and Depression Scale (HADS) was developed by Zigmond and Snaith (1983) as a self assessment instrument for detecting states of depression and anxiety over the past seven days in non psychiatric hospital outpatients. It is probably the most widely used self report anxiety questionnaire in research over the past two decades. A common problem with mood disorder questionnaires prior

to the HADS was that scores were regularly affected by the physical illness of the patient. The items were therefore initially selected to address this failing (Zigmond and Snaith, 1983, Snaith, 2003). As the name suggests, the HADS consists of two subscales, one for measuring anxiety (A-scale) and one for measuring depression (D-scale), each containing seven items. Each item is scored on a four point Likert scale (0-3).

During its construction the HADS initially had eight items composing the depression subscale, and eight items composing the anxiety subscale. The items for the depression subscale were mainly statements or questions based on the anhedonic state (lack of pleasure response). The items for the anxiety subscale were mainly derived from the authors' personal research (Snaith et al., 1982). This questionnaire was given to a number of patients in a general medical outpatient clinic, who were also interviewed by the researchers after their appointment, to assess their overall severity of anxiety and depression. The questionnaire responses were analysed and compared to the results of the psychiatric assessment interview. The weakest correlating item from each subscale was removed.

Analysis of scores of both subscales from a new clinical sample allowed the authors to nominate a scale score of 0-7 to be normal. A score of 11 or higher could be regarded as indicating the probable presence of anxiety or depression and a score of 8-10 being suggestive of the presence of anxiety or depression. Subsequent experience seems to have resulted in the development of four scoring bands, however the English test manual (Snaith and Zigmond, 1994) states the research

concerning the development of these ranges is unpublished. The scoring ranges currently in use are 0-7 'normal', 8-10 'mild', 11-14 'moderate' and 15-21 'severe'.

There have been two published systematic reviews of the literature regarding the psychometric properties of the HADS (Herrmann, 1997, Bjelland et al., 2002). Herrmann reported information from approximately 250 studies and Bjelland et al found 747 that had reported using HADS. Both reviews summarize the available validation data and give an overview of how different translations of the scale perform.

Sensitivity, specificity and cut off points for identifying 'cases'

In the 2002 review by Bjelland and colleagues, the large majority of papers specified the optimum balance between specificity and sensitivity to be very close to a score of 8 for both subscales. The mean cut off point for HADS-A was 7.87 and for HADS-D was 8.13. This is remarkably similar to the cut off of 8 suggested for detecting 'possible cases' in the original HADS paper and test manual (Zigmond and Snaith, 1983, Snaith and Zigmond, 1994). When a cut-off of 8 was used the sensitivities were generally found to be between 0.70 and 0.93 with an outlier for HADS-A at 0.64 and HADS-D at 1. The specificity varied between 0.68 and 0.95. For these studies the average sensitivity and specificity for HADS-A was 0.84 and 0.86, and for HADS-D 0.86 and 0.82 respectively.

In contrast, Herrmann's review (1997) correctly states that it is 'not possible to give unequivocal general values for sensitivity and specificity' due to the fact that

investigators had used so many different cut-offs and several different 'gold standards' for comparison. Herrmann did report however that seventeen studies of the English HADS report sensitivities and specificities of 0.8 or greater, which is satisfactory.

The only published research into identifying cases since the 2002 review was by Poole and Morgan (2002) and the results were very good. For a cut-off of 8 the sensitivity and specificity for HADS-A was found to be 0.96 and 0.79 and for HADS-D 1.00 and 0.87 respectively (Poole and Morgan, 2006). A cut off of 8 has therefore been used by the present study in order to detect clinical significance.

2.3 ii) State-Trait Anxiety Inventory (STAI)

The original version of the STAI (Form X) was developed by Spielberger et al. (1970) as a self report questionnaire to assess two constructs; State anxiety, and Trait anxiety. Trait anxiety is described as "relatively stable individual differences in anxiety proneness", whereas State anxiety is the level of anxiety in an individual's current emotional state (Spielberger, 1983). Up to 1989, the STAI had appeared in over 3000 widely varying studies and has been translated into over 30 languages (Spielberger, 1989). Despite being used in thousands of studies, not many of these have used their data to comment on the reliability of the instrument (Barnes et al., 2002).

It was revised in 1983 (Form Y), with 30% of the items being replaced in an attempt to give a purer measure of anxiety that was less affected by depression. The

revision also replaced items that were found to be unsuitable or ambiguous for younger persons, and to give a better balance between anxiety-present and anxiety-absent items (Spielberger, 1983). The instrument contains 40 items, 20 aimed at State anxiety, followed by 20 for Trait anxiety. 21 of the items are anxiety-present items (e.g. “I feel nervous and restless) and 19 are anxiety-absent items (e.g. “I feel pleasant”). Items are scored on a four point Likert scale, scored 1-4, with the anxiety-absent items being reverse scored. Forty items is a large inventory when time is limited or multiple scales are being administered, therefore a number of studies have successfully reduced the number while retaining adequate reliability for certain populations (Chlan et al., 2003, Koizumi et al., 1998, Marteau and Bekker, 1992). Rasch analysis could be used for this purpose, however these studies used factor analysis, multiple regression and item remainder correlations respectively.

Every study cited below acknowledges that the STAI definitely has clinical usefulness in measuring overall psychological distress. Differing views are given as to whether the instrument does specifically what it was intended to do. The large majority of the following studies use populations with psychological disorders, whereas for the purposes of the current study more notice should be taken of validation in normal populations.

Cut off points for identifying ‘cases’

Other scales, such as the HADS, provide cut-off points for identifying and grading patients based on normative data. Normative data tables are provided in the STAI

manual (1983) for a number of different demographic groups, but cut off points are not. Kabacoff and colleagues (1997) investigated the possibility of having cut off scores, but concluded that the tradeoffs between sensitivity and specificity made it unfeasible. As the present study does not need to diagnose psychological problems, cut off scores are not required.

Cross cultural equivalence

As discussed elsewhere, due to differing interpretations of anxiety, the robustness of a psychometric instrument when applied to participants of various ethnicities is particularly important in an ethnically diverse city such as Bradford. The psychometric properties of the STAI have been shown to hold up well when administered to most cultures (Hishinuma et al., 2000), however caution must still be exercised. In the study by Hishinuma and colleagues participants of Filipino ethnicity were the only demographic group whose answers did not correlate well. Further investigation into the cross-cultural equivalence in the UK is required.

2.3 iii) General Health Questionnaire – 28 (GHQ-28)

The GHQ-28 is a 28 item questionnaire designed to have four 7 item subscales; A – somatic symptoms, B – anxiety and insomnia, C – social dysfunction, and D – severe depression. These subscales were identified by principal components analysis of a longer version; the 60 item General Health Questionnaire (GHQ-60). Other shorter versions of this questionnaire have also been developed by other methods; GHQ-12

and GHQ-30 (Goldberg and Williams, 1988). The original work upon which the GHQ-60 was based sought to find the features that distinguished between psychiatric patients and healthy individuals (Goldberg, 1972).

Similarly to the previously discussed questionnaires, a four-point response scale is used. As with the previous scales this is usually scored with a different score for each response category (0-3), however Goldberg suggested an alternative way of scoring the scale that he called GHQ-scoring (Goldberg and Williams, 1988). GHQ-scoring allocates the first two categories a score of 0, and the final two categories a score of 1 (0-0-1-1) as the first two correspond to 'absence of symptoms' and the final two 'presence of symptoms'. This was reported to improve the validity of the questionnaire by slightly improving the sensitivity and specificity when used to diagnose mental illness. Although some items are based on symptom-absent questions, there are no reverse scored items. This means for every question the first response category always indicates the strongest symptom-absent response, and the last category always indicates the strongest symptom-present response. This runs the risk of the participants not reading the questions properly and endorsing all the answers in the same column to save time.

Cross cultural equivalence

All research into cross cultural stability has been favourable. The GHQ-28 has been described as being highly stable across cultural backgrounds (Iwata and Saito, 1992) and a suitable screening test in elderly Iranians (Malakouti et al., 2007). The Urdu translation by Riaz and Reza (1998) was found to be comparable to the English version in a population of bilingual Pakistani students.

Factor structure and internal consistency

The majority of studies that have investigated the factor structure of the GHQ-28 have agreed with the original four factor structure developed by Goldberg (Vallejo et al., 2007, Aderibigbe and Gureje, 1992, Iwata and Saito, 1992). The largest study to date that reported details of factor analysis was by Werneke, Goldberg and colleagues (2000) with 5273 participants across 15 international centres. Between four and six factors were found, with the median being four. Two of these factors were very robust across all centres, the others were more variable. When reported, internal consistency is high or acceptable (Iwata and Saito, 1992, Vallejo et al., 2007, Riaz and Reza, 1998).

Sensitivity, specificity and Cut off points for identifying 'cases'

A number of studies have reported ideal cut-off points for identifying individuals with psychiatric disorder, with corresponding sensitivities and specificities. All the studies have been performed in different populations, and have been detailed in Table 3.1, with the original authors' research suggesting a cut off of 4/5. Also in the table are the various interview techniques used to validate the GHQ-28, and define which patients are 'cases'. These include the Clinical Interview Schedule (CIS), Present State Examination (PSE), Diagnostic Interview Schedule (DIS), the Mini Mental State Examination (MMSE), the Psychiatric Assessment Schedule (PAS) and

the Composite International Diagnostic Interview (CIDI). Standardisation of the scoring technique is needed to allow comparison.

As can be seen, the concurrent validity of the GHQ-28 is generally good. Other studies have supported this view without specifying cut-offs or sensitivities and specificities: Benjamin and colleagues (1991) validated the scale against the CIS, and Ormel (1989) reported that it performed “remarkably well” when compared to the PSE, both in cross sectional studies and longitudinally in detecting changes over time.

The lower sensitivity reported by Makowska and colleagues (2002) was attributed to the influence of the somatic subscale, and its probable lack of correlation with psychological disorders. Andersen (2002) commented that the validity was affected by the inappropriateness of some items in certain populations. For example, “Have you recently been satisfied with the way you have carried out your tasks?” in a prison population.

Rasch analysis

The GHQ-28 has not yet been analysed by Rasch techniques.

Author (Year)	Population	Interview	Cutoff	Sensitivity	Specificity
(Goldberg and Hillier, 1979)	GP patients: London Manchester	CIS CIS	4/5 5/6	88 80	84.2 88.8
(Rabins and Brooks, 1981)	Neurological outpatients	PSE	4/5	92	92
(Banks, 1983)	UK adolescents	PSE	5/6	100	84.5
(Mann et al., 1983)	UK 15 year old schoolgirls	CIS	4/5	48.5	83.1
(Medina-Mora et al., 1983)	Mexican outpatients	CIS	5/6	72.5	73.8
(Bridges and Goldberg, 1986)	UK Neurological in-patients	MMSE	11/12	80.2	80.7
(Lobo et al., 1986)	Spanish outpatients	CIS	6/7	76.9	90.2
(Goldberg and Bridges, 1987)	UK GP patients	PAS	4/5	87.1	75.4
(Romans-Clarkson et al., 1989)	New Zealand women	PSE	3/4	~87	~63
(Lim and Chew, 1991)	Singapore, VDU operators	CIS	5/6		
(Goldberg et al., 1997)	15 International centres	CIDI	Ave. 5/6	79.7	79.2
(Makowska et al., 2002)	Poland, working population	CIDI	5/6	59	75
(Andersen et al., 2002)	Denmark, prison population	PSE	9/10 10/11	65	69
(Malakouti et al., 2007)	Iran, elderly	CIDI	19/20	83	76

Table 2.1 Proposed cut-off values and associated sensitivity and specificity for GHQ-28 along with the populations and interview techniques used (Clinical Interview Schedule (CIS), Present State Examination (PSE), Diagnostic Interview Schedule (DIS), the Mini Mental State Examination (MMSE), the Psychiatric Assessment Schedule (PAS) and the Composite International Diagnostic Interview (CIDI))

2.4 Ethics

The protocol for all chapters complied with the Declaration of Helsinki. Ethical approval was given by the Bradford NHS Research Ethics Committee (REC), and local approval was given by Bradford Royal Infirmary Research and Development (R&D) Office. Two substantial amendments were made to the protocol at the request of the REC; the first requirement was that the GHQ-28 was not used as they viewed the items in the "Severe Depression" subscale as potentially distressing, and the second was that the initial letter to the patient had to come from a member of the care team at Bradford Royal Infirmary (CG) instead of the principal investigator (CD). The REC and R&D office disagreed with each other over other small aspects of the protocol, but these were ultimately resolved for data collection to begin in earnest in December 2007.

2.4a The effect of assessing anxiety/depression on participants in studies

Although we are attempting to assess levels of psychological distress through the use of questionnaires, we must also consider the possible psychosocial impact of the questionnaire itself. Anxiety and depression questionnaires, similar to most mental health questionnaires, can contain some potentially sensitive items.

Previous research on mental health surveys concluded that very few participants reported any adverse effect on emotional state caused by the interview (Henderson and Jorm, 1990, Jorm et al., 1994). The study by Jorm and colleagues in 1994 found that only 4% of participants were distressed by the survey, 1% were depressed by it and 2% said it intruded on their privacy. Conversely, 52% of respondents said it

made them feel good about themselves. A more recent study by Jacomb and colleagues, again using a mental health survey, found similar results. 5% of respondents felt it had caused them to feel distressed, 3% depressed and 3% were concerned about privacy. A large proportion (35%) of respondents again reported feeling good about themselves after completing the survey (Jacomb et al., 1999). All the surveys used in these studies contain items focusing on dementia, cognitive decline, depression, and current life circumstances which are potentially more upsetting than our anxiety and depression instruments.

Ways to reduce psychological harm to participants include informing them that they can decline to answer any problematic questions and that they can withdraw at any time (Jorm et al., 1994, Evans et al., 2002). Subjects also need to be fully informed, both about the purpose of the research and any potential benefits and risks, before giving consent. They also need to be aware of the significance of research for their own care and be informed of the results of the research at the end of the study (Mayberry, 2002, Evans et al., 2002). Capacity to consent also needs to be checked according to Department of Health guidelines (2001). Many studies on the effect of questionnaires have been done using questionnaires about medical conditions concerning the population being assessed, for example, questionnaires about breast cancer symptoms in mammography patients.

2.4b Non-response bias

With large scale questionnaire and survey based research response rates are frequently rather low, therefore possibly introducing a bias in the study population.

In longitudinal studies with follow up questionnaires, attrition can have a further effect. The sample may not be representative of the whole population; it is limited to the responders, who may be a non-random distribution. A small number of studies have investigated commonality in non-responders in an attempt to see what type of person this research may be unknowingly excluding.

A postal survey of women's experience of childbirth (Cartwright, 1986) reported significant differences in response rates for a large number of variables. The study found that mothers who described themselves as Caucasian, were breast feeding their baby, or had a normal birth weight baby that was not admitted to special care were more likely to respond. If the mothers were in paid employment, did not delay in registering the birth, or did not attend late for ante-natal appointments, they were also more likely to be responders. The author identified that relationships probably exist between some of these variables, and with the language of the mother, although no data were available regarding this. No bias was associated with age, smoking habits or previous obstetric history.

Due to the study population in Bradford, which is discussed in more detail later, the most relevant of these variables will be ethnicity, with language possibly as the true underlying cause of non-response. Interestingly though, Cartwright does discuss that among the patients of Asian ethnicity, Hindu mothers had a higher response rate (76%) than Muslim mothers (44%), which the author suggests may be due to the Hindu immigrant group being established for a longer period of time in the study area. Cartwright's questionnaires were distributed over 25 years ago,

therefore it is reasonable to believe this bias may be somewhat diminished in present day Bradford, where the community has been established much longer.

Two studies have looked at survey non-response in a population of family physicians. The latest in Norway (Bjertnaes et al., 2008) found six variables improved the likelihood of response, for example, longer experience as a GP, being registered with the Norwegian Medical Association and regional variations. None of the variables is directly applicable to our study population. Interestingly a similar study of physicians in the US (McFarlane et al., 2007) found only gender to affect response rates, with men being initially more likely to respond.

Attempts to statistically assess the impact of non-response and attrition have suggested that despite it leading to non-random sampling, relationships between variables targeted by the surveys are unaffected (Goodman and Blum, 1996, Goldberg, 2003). The family physician studies (Bjertnaes et al., 2008, McFarlane et al., 2007) also found that their survey estimates changed little when they increased response using reminders. The only improvement was a reduction in gender bias, with men more likely to be early responders and women responding well to reminders.

2.4c Improving response rates of questionnaire studies

A way to minimise any possible bias would be to improve the response rate as much as possible. An excellent review of 292 randomised controlled trials with a total of 258315 participants (Edwards et al., 2002) suggested that a number of simple

methods can significantly improve response rates. Some of relevance to the current study include; using coloured ink, shorter questionnaires, personalised letters and first class or recorded post. Participants responded well if they were contacted before they received the questionnaires, and after if they did not respond. Questionnaires with questions of a sensitive nature were less likely to be returned, but thankfully surveys from universities were more likely to be completed than those from commercial organisations.

Electronic response vs. paper response – validity & response rates

One strategy not evaluated by Edwards and colleagues is to include an internet response option, allowing patients to answer the questionnaires via a webpage instead of on paper. As this would allow participants to respond without visiting a post box, is it reasonable to assume the increased convenience would improve response rates? Contrary to this, Brogger and colleagues (2007) found that having a web response option did not improve response rate at all, despite 18.5% of participants responding in this way. Earlier studies found that having a web response option significantly decreased the likelihood of response (Griffin et al., 2001, Kaplowitz et al., 2004) with the authors hypothesising that this may be due to a perceived reduction in privacy or frustration caused by using the webpage (Griffin et al., 2001). Kaplowitz and colleagues (2004) acknowledged that a web based response was substantially cheaper than a postal response.

If web response options are included, we need to be sure that it is possible to group together web responses and postal responses. Once again there is conflicting evidence regarding this. A small number of studies have concluded that, although it

is recommended to keep administration format constant there was a high correlation between web based and paper questionnaires, with both usually showing similar psychometric properties (Carlbring et al., 2007, Vallejo et al., 2007, Andersson et al., 2003). A review of three projects by Buchanan (2003) warned that using available normative data (from paper questionnaires) to analyse data from internet administration would introduce errors. Some authors found that online questionnaires resulted in higher scores (Andersson et al., 2003), some found similar scores (Carlbring et al., 2007), and others found lower scores (Vallejo et al., 2007). This is clearly an area that requires further randomised control trials before firm conclusions can be drawn, however no study has yet strongly denounced the combining of internet and paper questionnaire data.

2.4d Special communication needs

A significant number of patients at Bradford Royal Infirmary may be unable to read or write in English but are able to read Urdu (Tuffnell et al., 1994). To accommodate these individuals as best we could, we have sourced validated Urdu translations of the Hospital Anxiety and Depression Scale, and the General Health Questionnaire-28. Since all translated questionnaires must be validated prior to use, we were unable to source verified Urdu translations of the State Trait Anxiety Inventory. All written information documents were available on an audio CD in both Urdu and English. We were unable to provide translators for this study as they would have been required intensively for a long period of time, incurring costs that are many times larger than available funding.

2.4e Confidentiality of data collected

Until consent was given, the patients were contacted only by a member of the clinical care team (CG). Only after consent was given were the University of Bradford researchers allowed access to participant details.

For chapters 3.1 and 3.2, all personal data were fully anonymised. Date of birth, gender and ethnicity were recorded but names and addresses were not. The data from the second part of the study were fully anonymised and codes have been used to link medical details with questionnaire responses. Names and contact details stored separately for administering the follow up questionnaires will be destroyed from University records by shredding once the study has been completed. During the study, hard copies of data were stored in a locked filing cabinet in a private office in the Richmond Building, University of Bradford. Data were only stored electronically on university computers (not laptops) and files were coded and password protected.

Chapter 3. False positive referrals

and the accuracy of referrals to secondary eye care

3.1. Introduction

A discussion of the research literature of false positive referrals to secondary eye care and the accuracy of such referrals is provided in Chapter 1.1.3. The aims of this study were to determine:

i) The levels of false positive referrals by optometrists and general practitioners (GPs) to the Hospital Eye Service (HES).

ii) The levels of accuracy of referrals by optometrists and GPs to the HES provided in terms of proportion discharged at their first visit and diagnosis agreement between primary and secondary eye care clinician.

iii) The factors that influence false positive referrals and accuracy of referral. Factors considered for inclusion were patient age, patient gender, patient ethnicity, referral format, referral diagnosis, type of referring clinician,

3.2. Methods

A random sample of 431 new referrals to Bradford Royal Infirmary (BRI) ophthalmology department during 2007 and 2008 were retrospectively analysed. The presence of the following information was recorded from the referral: patient

name and address, date, referrer name, referral format, referrer address, any diagnosis given or alluded to and final diagnosis at the hospital (classified based on the International Classification of Diseases-10 (ICD-10), World Health Organisation). A significant proportion of referrals to ophthalmology in the Bradford PCT area go to the optometric led Shipley Ophthalmic Assessment Project (SOAP) where a specially trained hospital optometrist assesses the patient and either discharges or lists the patient for appropriate treatment or review in secondary care. According to their information services department, BRI sees approximately 1400 new ophthalmic outpatients per year whereas approximately 350 are seen by SOAP, therefore to get an accurate impression of referrals in this district we also included approximately 25% (n=106) SOAP patients.

In order to assess false positive levels and accuracy of referral, definitions of each of these were first determined.

3.2a. Proportion of False Positive referrals

An attempt to define a false positive referral to the hospital eye service is given below, although any single definition will have problems accurately representing the data. It is therefore of value to also report the proportion of patients discharged at the first appointment and the proportion of referrals where the referral diagnosis matched the hospital diagnosis. Considering all three of these values should help to give a more balanced view of the data. As detailed in Chapter 1.1.3, previous studies have used differing methods of classifying the accuracy of referrals. Those

occasions where patients were not easily classified into these groups, these patients were discussed with the examining ophthalmologist when possible.

3.2b Definition of False Positive Referral

A false positive referral was identified by either of the following:

1. The ophthalmologist examined the patient, and subsequently discharged the patient due to the absence of significant ocular pathology. The ophthalmologist's decision to discharge must not have been *solely* influenced by clinical techniques that were *not currently commonly available* to the referring practitioner.

2. The examining ophthalmologist diagnosed the patient with, or was suspicious of, pathology that was unrelated to the diagnosis given or implied by the referring practitioner. The ophthalmologist was happy that the pathology for which the patient was referred for was not present, with this decision not being influenced *solely* by clinical techniques that were *not currently commonly available* to the referring practitioner.

Fundoscopy, either direct or indirect using a non-contact lens, tonometry and central visual field screening are examples of techniques that should all be available in UK optometric practices according to College of Optometrists guidelines (2007) and previous literature (Myint et al. 2011). Examples of techniques not currently widespread in UK optometric practices are pachymetry, gonioscopy, optical coherence tomography and fluorescein angiography. General Medical Practitioners, unless they have a special interest in Ophthalmology, tended to only have case

history, direct observation, fundoscopy, pupil assessment and visual acuity measurements at their disposal.

3.2c. Definition of Correct Referral

Correct referrals were identified by either of the following:

1. The examining ophthalmologist diagnosed a condition that necessitated referral and it was the same as or closely related to, the condition diagnosed or inferred by the referring practitioner.

2. The examining ophthalmologist deemed it necessary for the patient to be seen in the Ophthalmology department for an additional appointment at some point in the future due to suspicion of the referral pathology. Any appointment that is more than 48 hours after the first is classified as an additional appointment.

3.2d. Limitations of these definitions

As stated above, any definition will have limitations and may misrepresent the data. The major limitation of the false positive definition is that it depends on what clinical techniques are commonly available to the referring practitioner. This is necessary in order to compensate for the expectations and requirements of different clinicians, however it is naturally weighted in favour of those offering few ophthalmic techniques and may conceal some poor referrals. For example, any glaucoma referral by a GP would not have been deemed as false positive regardless of the outcome, because GPs do not have the equipment required to accurately make a diagnosis of glaucoma. It is also only possible to validate diagnostic decisions made by the referring clinician, as opposed to the validity of the referral

itself. In other words, a practitioner may refer simply for a second opinion, which may be for a valid reason, without having arrived at a diagnostic decion.

To allow practical analysis each referral pathology was classified into one of 18 groups:

3.2e. Classification of Ocular Pathology

Ocular pathology was classified using a simplified version of the International Statistical Classification of Diseases and Related Health Problems, 10th version (ICD-10), by the World Health Organization © Copyright WHO/DIMDI (2006). The equivalent ICD-10 classification codes are given for reference. Our intention was for all ocular pathology to be classified according to the following 18 groups.

H00-H06	Disorders of eyelid, lacrimal system and orbit
H10-H13	Disorders of conjunctiva
H15	Disorders of Sclera
H16-H19	Disorders of the cornea
H20-H22	Disorders of iris and ciliary body
H25-H28	Disorders of lens
H30-H32	Disorders of choroid
H33	Retinal detachments and breaks
H34	Retinal vascular occlusions
H35-H36.8	Other disorders of retina

H36.0 Diabetic retinopathy

H40.0 Suspect Glaucoma

H40.1 Primary open angle glaucoma

H40.2-H42 Angle closure, secondary and other Glaucoma

H46-H48 Disorders of Optic nerve and visual pathway

H43-H45 Disorders of vitreous body and globe

H46-H48 Disorders of Optic nerve and visual pathway

H49-H53 Disorders of ocular muscles, binocular movement, binocular vision, amblyopia, accommodation and refraction

H53-H59 Visual disturbances and other disorders of eye and adnexa

3.3. Data analysis

Data were analysed with the help of statistician Mr Andy Scally (School of Health, University of Bradford) with a logistic regression model using Stata version 9.0 statistical programme (Stat Corp., College Station, USA). Variables of interest were incorporated sequentially and their statistical significance was assessed. The predictor variables were; type of referring clinician, referrer gender, years the referrer has been registered, type of practice, pathology classification, format of referral, legibility, age of patient, gender of patient, and ethnicity of patient. The outcome variables were; defined as false positive, diagnosis agreement, and

discharged at first visit. Significance of the two-level factors was determined by the 'Z'-statistic, while the significance of a higher number of factors was tested using a likelihood ratio (χ^2) test after dropping individual factors from the model. Factors with a *p*-value less than 0.1 were provisionally retained, whereas those above 0.1 were dropped. The final model adopted was the most parsimonious one that was felt to adequately explain the data, with the final level of significance set at $p < 0.05$. Factors were first considered in a multiple logistic regression model. When collinearity or missing data were a problem, univarite logistic regression analyses were used.

The results have been described using odds ratios (OR). The odds of an event taking place in a group is the number of times it occurs in that group divided by the number of times that it does not occur in that same group. Odds ratios are the odds of an effect being found in a study group divided by the odds of it being found in a comparison group. For example, the odds ratio of a false positive referral being made by an optometrist would be the odds of a false positive referral being made by an optometrist divided by the odds of a false positive referral being made by the other primary care referrers. The odds of an event taking place is the number of times it takes place in a group divided by the number of times that it does not occur in the same group. Using the same example the odds of an optometrist making a false positive referral would be the number of false positive referrals optometrists make divided by the number of true positive referrals optometrists make. An OR greater than one means that the 'event' (e.g., A false positive referral) occurred more often in the group being investigated than in the comparison group. An OR of

one means it occurred equally in the two groups whereas an OR between 0 and 1 means the event occurred more often in the comparison group (Cockburn, 2006).

3.4. Results

The number (and percentage) of false positives and accurate referrals from all primary care clinicians, with some figures given for significant factors, are shown in Table 3.1 below.

Source of referral	False positive	Discharged	Medication	Review	Surgery	*Diagnosis agreement
GP (n=131)	4 (3%)	37 (28%)	26 (20%)	29 (22%)	39 (30%)	92 (70%)
Optometrist (n=366)	105 (29%)	114 (31%)	30 (8%)	141 (39%)	81 (22%)	253 (69%)
Pre-reg Optom (n=26)	11 (42%)	11 (42%)	1 (4%)	9 (35%)	5 (19%)	18 (69%)
DRSS technician (n=9)	0	4 (44%)	0	5 (56%)	0	7 (78%)
Female Optoms (n=122)	47 (39%)	47 (39%)	7 (6%)	48 (40%)	21 (17%)	81 (66%)
Male Optoms (n=159)	36 (23%)	41 (26%)	13 (8%)	63 (40%)	41 (26%)	121 (76%)
Independent practice (n=169)	38 (22%)	46 (27%)	14 (8%)	68 (40%)	41 (24%)	123 (73%)
Multiple practice (n=206)	74 (36%)	76 (37%)	16 (8%)	72 (35%)	42 (20%)	136 (66%)

Table 3.1. The accuracy of all referrals from all sources by the three different measures of referral accuracy
*For the purposes of diagnosis agreement glaucoma and glaucoma suspect have been classified as the same.

It was not possible to ascertain the practice type in 16 referrals (DRSS technician referrals were not included in practice-type analyses, but pre-reg student referrals were). It was not possible to ascertain the Optometrists gender in 85 referrals (pre-

reg student referrals were not included in the gender analyses as we were unaware of the gender of the supervisor who was legally responsible for the referral).

3.4a. **False positive:**

The relationship between false positive referrals and eye disease category is shown in Table 3.2.

Diagnosis Category	False positive referrals
H00-H06 Disorders of eyelid, lacrimal system and orbit (n=13)	2 (15%)
H10-H13 Disorders of conjunctiva (n=3)	1 (33%)
H16-H19 Disorders of the cornea (n=15)	2 (13%)
H25-H28 Disorders of lens (n=93)	15 (16%)
H30-H32 Disorders of choroid (n=4)	3 (75%)
H33 Retinal detachments and breaks (n=7)	3 (43%)
H34 Retinal vascular occlusions (n=14)	3 (21%)
H35.3.1 Age related macular degeneration (n=25)	7 (28%)
H35-H36.8 Other disorders of retina (n=23)	11 (48%)
H36.0 Diabetic retinopathy (n=30)	6 (20%)
H40.0 Glaucoma suspect (n=79)	22 (28%)
H43-H45 Disorders of vitreous body and globe (n=18)	12 (67%)
H46-H48 Disorders of Optic nerve and visual pathway (n=7)	5 (71%)
H49-H53 Disorders of ocular muscles, binocular movement, binocular vision, amblyopia, accommodation and refraction (n=7)	2 (29%)
H53-H59 Visual disturbances and other disorders of eye and adnexa (n=25)	11 (44%)
None (n=3)	0

Table 3.2. Number and percentage of referrals from Optometrists defined as false positive in each referral diagnosis category

A multiple logistic regression was performed to detect the differences in false positive levels for all variables and associated significance values. As shown in Table 3.1, almost all GP referrals were not false positive and therefore these data had to be removed from the multiple regression due to this almost perfect prediction and the resulting distortion of the remaining analyses.

No significant effects were found for patient gender, practice type, legibility, type of referring clinician or age of patient (all $p>0.10$). Referrals using direct methods were 2.7 times less likely to be false positive than other referral formats ($p=0.007$). When compared individually, direct referral methods were 3.2 times less likely to result in a false positive referral than new GOS18 referrals, 2.8 times less likely than old GOS18 referrals and 3.5 times less likely than a letter.

The proportion of false positive referrals generated by primary care clinicians decreases with experience at a rate of 9.1% per year since registration ($Z=-5.62$, $p<0.0001$). When considering only results from optometric practice the effect reduces slightly to 6.2% per year ($p<0.0001$).

Univariate analysis showed there was a significant difference between proportions of false positive referrals generated by independent and multiple optometric practices, with independent practices generating about half the number of false positives as multiple practices (Odds Ratio, $OR=0.52$, $p=0.005$, $N=376$). When controlling for years since registration ($N=294$ as registration date was not ascertained for 82 referrals) the effect reduced, with independents generating 30% fewer false positive referrals but the difference became not significant ($OR=0.7$, $Z=-1.28$, $p=0.20$, $N=294$).

Male optometrists were about half as likely to generate a false positive referral than females (OR=0.51, Z=-2.64, p=0.008, N=305). Female optometrists were younger and a greater proportion worked in multiple practices. The effect still remained and was still significant when controlling for years since registration (OR=0.57, Z=-2.18, p=0.029, N=298). However when including practice type and years registered as confounders the effect only approached significance (OR=0.62, Z=-1.79, p=0.073, N=294).

To allow for statistical analysis the diagnosis categories were further condensed into the five biggest groups, which were; disorders of lids/lashes, disorders of lens, glaucoma, visual disturbance/other and the remainder were grouped together. A just statistically significant link between false positives and diagnosis category (Likelihood Ratio, LR χ^2=9.7 p=0.046) was found. The rank order from lowest to highest false positive proportion was; 1. lens, 2. lid/lashes, 3. glaucoma, 4. everything else, and finally 5. visual disturbance/other, which had the most false positives. On further investigation the patients that were referred with this diagnosis category were found to have the following diagnoses at the hospital (Table 3.3).

Hospital diagnosis category		N=60
H00-H06	Disorders of eyelid, lacrimal system and orbit	3
H10-H13	Disorders of conjunctiva	3
H16-H19	Disorders of the cornea	1
H25-H28	Disorders of lens	7
H30-H32	Disorders of choroid	1
H34	Retinal vascular occlusions	3
H35.3.1	Age related macular degeneration	2
H35-H36.8	Other disorders of retina	2
H36.0	Diabetic retinopathy	1
H40.0	Glaucoma suspect	2
H43-H45	Disorders of vitreous body and globe	6
H46-H48	Disorders of Optic nerve and visual pathway	2
H53-H59	Visual disturbances and other disorders of eye and adnexa	13
None		14

Table 3.3 The ophthalmological diagnoses of patients referred with "Visual disturbances and other disorders of eye and adnexa".

Tables 3.4 and 3.5 show that "Visual disturbances and other disorders of eye and adnexa" was the referral reason for 6.8% of optometrists referrals and 22.9% of GP referrals. No relationship was found between levels of false positive referrals and patient ethnicity.

Hospital diagnosis		GP reason for referral																Total	% Agree
		H00-06	H10-13	H15	H16-19	H20-22	H25-28	H33	H34	H35.3.1	H35-36.8	H36	H40	H43-45	H46-48	H49-53	H53-59		
	None	1												1			8	10	
Lid/lacrimal	H00-06	**51**	4			1											2	58	88%
Conjunctiva	H10-13		**5**	1													2	8	63%
Sclera	H15																	0	
Cornea	H16-19				**2**													2	100%
Iris/Ciliary	H20-22																	0	
Lens	H25-28						**13**	1									5	19	68%
Retinal det/breaks	H33																	0	
Vascular occlusions	H34						1										3	4	0%
ARMD	H35.3.1						1			**4**							1	6	67%
Other retina	H35-36.8										**1**							1	100%
Diabetic retinop.	H36	1																1	0%
Suspect glaucoma	H40	1															1	2	0%
Vitreous/globe	H43-45													**3**			3	6	50%
Visual pathway	H46-48														**1**		1	2	50%
BV/muscles/ref.	H49-53															**8**		8	100%
Vis. disturb./other	H53-59																**4**	4	100%
	Total	54	9	1	2	1	15	1	0	4	1	0	0	4	1	8	30	131	
	% Agree	94%	56%	0%	100%	0%	87%	0%		100%	100%			75%	100%	100%	13%		

Table 3.4. A comparison of the GP referral diagnosis and the final hospital diagnosis.
For the purposes of diagnosis agreement glaucoma and glaucoma suspect have been classified as the same. A diagnosis category of 'None' for GPs means no specific reason for referral was ascertainable from the referral letter/form eg. due to illegibility or missing information. When two reasons for referral were given, either of them agreeing with the primary hospital diagnosis counted as an agreement.

HES diagnosis		Optometrist reason for referral																Total	% Agree
		None	H00-06	H10-13	H16-19	H25-28	H30-32	H33	H34	H35.3.1	H35-36.8	H36	H40	H43-45	H46-48	H49-53	H53-59		
	None		1			2	2	2	2	2	7	2	23	4	2		5	54	
Lid/lacrimal	H00-06	1	**11**		2						1		1	1	1		1	19	58%
Conjuntiva	H10-13		1	**3**	3						1	1					1	10	30%
Cornea	H16-19				**9**						1						1	11	82%
Lens	H25-28	1			1	**89**			1	1			1			2	2	98	91%
Choroid	H30-32						**2**				1							3	67%
Ret. det./breaks	H33							**4**			1							5	80%
Vascular occlusions	H34								**10**	1								11	91%
ARMD	H35.3.1									**19**	1						1	21	90%
Other retina	H35-36.8					2			1	2	**8**	1		1			2	17	47%
Diab. Retinop.	H36										1	**24**					1	26	92%
Sus. Glaucoma	H40					1					1	2	**51**	1			1	57	89%
Vitreous/globe	H43-45							1						**11**			3	15	73%
Visual pathway	H46-48												1		**4**			5	80%
BV/muscles/ref.	H49-53															**5**		5	100%
Vis. Distu./other	H53-59	1											1				**7**	9	78%
Total		3	13	3	15	94	4	7	14	25	23	30	78	18	7	7	25	366	
% Agree			85	100	60	95	50	57	71	76	35	80	65	61	57	71	28		

Table 3.5. A comparison of the Optometrist referral diagnosis and the final hospital diagnosis.
For the purposes of diagnosis agreement glaucoma and glaucoma suspect have been classified as the same. A diagnosis category of 'None' for Optometrists means no specific reason for referral was ascertainable from the referral letter/form eg. due to illegibility or missing information. When two reasons for referral were given, either of them agreeing with the primary hospital diagnosis counted as an agreement.

3.4b. Diagnosis agreement:

A multiple regression was performed on all variables and diagnosis agreement between referral and hospital diagnosis. Many variables exhibited strong co-linearity and therefore practice type, clinician and referral format were removed from the analyses.

Table 3.6 shows the number of referral diagnoses that agreed with the hospital diagnosis for the condensed pathology classifications. Univariate analysis showed that lid/lacrimal referrals had no significant difference in diagnostic agreement to glaucoma (OR=1.13, p>0.10) and the 'everything else' category (OR=0.93, p>0.10). Lens disorders were significantly easier to diagnose correctly compared to lids/lacrimal (OR=5.4, p<0.0001) whereas visual disturbances/other gave the worst agreement (OR=0.14, p<0.0001) when compared to lids/lacrimal.

Referral diagnosis category	Diagnosis agreement
H00-H06 Disorders of eyelid, lacrimal system and orbit (n=69)	44 (64%)
H25-H28 Disorders of lens (n=115)	104 (90%)
H40.0 Glaucoma suspect (n=84)	56 (67%)
H53-H59 Visual disturbances and other disorders of eye and adnexa (n=60)	12 (20%)
Everything else (n=209)	130 (62%)

Table 3.6. Diagnosis agreement between referrer and hospital clinician for condensed diagnosis groups.

Similar to the false positive analysis, referrals from independents were 52% more likely to have a correct diagnosis, although this effect only approached statistical significance (OR 1.52, Z=1.89, p=0.059, N=376). The effect was not significant when

controlling for years since registration (OR 1.34, Z=1.06, p=0.29, N=294). When the effect that years since registration had on diagnosis agreement was investigated by itself, this was also found to be approaching significance (OR=1.02, Z=1.82, p=0.069, N=298).

A significant effect still existed for referrer gender. Male optometrists were more likely to provide a correct diagnosis (OR=1.71, Z=2.15, p=0.031, N=305). This effect was reduced when controlling for years since registration (OR=1.67, Z=2.02, p=0.043, N=298) and only approached significance when years since registration and practice type were included as confounders (OR=1.60, Z=1.80, p=0.071).

3.4c. Further Investigation (Discharged at first appointment)

Univariate analysis found no significant effect for patient age, patient ethnicity, patient gender, and type or gender of referring clinician. Years since registration again showed a significant effect (Z=3.96, p=0.0001).

The only other significant effect was found with referral diagnosis category (χ^2=36.3, p<0.0001). Further analysis showed that there was no significant difference between referrals for disorders of lids/lacrimal and disorders of lens (p=0.96), however the three remaining categories were all significantly different. When compared to lids/lacrimal, glaucoma referrals gave an effect of z=-2.17 (p=0.03), visual disturbance referrals gave an effect of z=-3.58 (p<0.001) and the group

containing everything else had an effect of z=-3.12 (p=0.002). These groups are shown in rank order in Table 3.7.

Disorders of eyelid, lacrimal system and orbit	Least discharged after first appointment
Disorders of lens	
Glaucoma suspect	
Everything else	
Visual disturbances and other disorders of eye and adnexa	Most discharged after first appointment

Table 3.7. Condensed referral diagnosis categories shown in order from least discharged at first appointment to most discharged at first appointment.

3.4d. Optometrist gender, practice type and years since registration

Combining gender of optometrist, years since registration and type of practice were significantly linked with levels of false positive referrals (χ^2=24.9, p<0.0001) but further analysis was required to find the variable(s) driving this link. Multiple regression analysis showed years since registration as the strongest predictor (Z=-3.47, p=0.001) although this effect was reduced when only including male optometrists and was less significant (Z=-1.79, p=0.073). When using gender and years since registration as confounders the effect of practice type was not significant (p=0.38) and can therefore be dropped from the analysis which still leaves significant effects for both gender (Z=-2.02, p=0.043) and years registered (Z=-3.9, p<0.0001). There was no interaction effect for years since registration and gender (p=0.63) therefore the gender effect does not appear to be age related. In summary, years since registration is the most important variable that drives an increase in false positive referrals.

When looking at the three variables and diagnosis agreement there were no significant effects. When practice type was removed as the weakest predictor a gender effect was present (Z=-1.99, p=0.046) and a non-significant effect was present for years registered (Z=-1.65, p=0.099).

When considering whether the referral resulted in further investigation a significant effect was only found for 'years since registration' (Z=3.96, p<0.0001). Practice type (z=0.19, p=0.88) and optometrist gender (z=1.43, p=0.15) both being not significant.

3.4e. Effects of Shipley Ophthalmic Assessment Project (SOAP)

Table 3.1.8 shows the referrals directed to SOAP appear to be less precise across all three methods of assessing accuracy. Pearson χ^2 test for significance showed all differences to be significant; for false positive χ^2=69.5 (p<0.0001), discharged at first visit χ^2=66.4 (p<0.0001) and diagnosis agreement χ^2=17.2 (p<0.0001).

Destination of Referral	False positive	Discharged	Medication	Review	Surgery	*Diagnosis agreement
Hospital (n=431)	65 (15%)	100 (23%)	58 (13%)	146 (34%)	127 (29%)	296 (69%)
SOAP (n=106)	56 (53%)	68 (64%)	0	38 (36%)	0	50 (47%)

Table 3.8. A comparison of the accuracy of referrals to the hospital eye service and Shipley Ophthalmic Assessment Project (SOAP).
*For the purposes of diagnosis agreement glaucoma and glaucoma suspect have been classified as the same

	Hospital		SOAP	
Referral diagnosis category	Discharged	Surgery/Review/ Medication	Discharged	Review at hospital
H00-H06 Disorders of eyelid, lacrimal system and orbit (n=69)	9 (14% of 64)	55 (86% of 64)	2 (40% of 5)	3 (60% of 5)
H25-H28 Disorders of lens (n=115)	14 (14% of 98)	84 (86% of 98)	4 (24% of 17)	13 (76% of 17)
H40.0 Glaucoma suspect (n=84)	13 (20% of 66)	53 (80% of 66)	15 (83% of 18)	3 (17% of 18)
H53-H59 Visual disturbances /other disorders of eye and adnexa (n=60)	15 (38% of 39)	24 (62% of 39)	13 (62% of 21)	8 (38% of 21)
Everything else (n=209)	49 (30% of 164)	115 (70% of 164)	34 (76% of 45)	11 (24% of 45)

Table 3.9. A comparison of the management of patients referred to the Hospital Eye Service and Shipley Ophthalmic Assessment Project (SOAP) for each condensed referral diagnosis category.

	GP	Multiple	Independent	Unable to ascertain
Hospital (n=431)	119 (28%)	149 (35%)	145 (34%)	18 (4%)
SOAP (n=106)	23 (22%)	58 (55%)	24 (23%)	1 (1%)

Table 3.10. A comparison of the source of referrals that reach the Hospital Eye Service and Shipley Ophthalmic Assessment Project (SOAP).

Table 3.10 shows that there was a significantly higher proportion of referrals from multiple practices going to SOAP (χ^2=15.9, p=0.001).

3.5. Discussion

3.5a. Referring clinician: GPs and Optometrists

According to the definition in 3.1.2b, there were very few false positive referrals among the GP (and DRSS) referrals. This means that nearly all of these referrals were appropriate within the remit of that clinician's speciality, but this probably says more about the limited ophthalmic clinical techniques available to GPs (and DRSS technicians). If these patients had attended their optometrist it is possible that they may not have been referred, but investigated and managed appropriately. Despite few GP referrals being deemed as false positive, 28% were discharged at the first appointment after investigation with ophthalmic clinical techniques, and 20% were initiated on medication (11% being ocular lubricants, antibiotics or a combination of both which are commonly prescribed by GPs). 70% of the GP referrals agreed with the eventual diagnosis at the hospital, similar to the figure for optometrists (69%), however (as discussed in the following chapter 3.2), the majority of GP referrals are for lids/lashes/lacrimal disorders (classification H00-H06), which may be easy to diagnose within the H00-H06 diagnosis category even if the diagnoses are not exactly the same. For example, if a patient is referred with a "stye", but the ophthalmologist diagnoses blepharitis, the referral would still be counted as agreeing with the hospital.

Without the H00-H06 referrals the diagnosis agreement for GPs drops to 53%. If more stringent rules were used to classify the referrals in this category, whereby

the referrer must get the diagnosis exactly correct, or very similar, instead of just the correct category, then agreement for GPs drops markedly to 56% and slightly for optometrists to 67% (table 3.11). For example, if a patient was referred with 'dry eye' but 'blepharitis' was diagnosed at the hospital, there was no exact agreement. For conjunctivitis, the type of conjunctivitis was not important, however merely referring with 'sticky red eyes' without diagnosing conjunctivitis did not result in exact agreement. If the referring practitioner had not given an exact diagnosis, eg. 'lid lesion', then it was impossible to get an agreement. A simple referral diagnosis of cataract was accepted as agreeing with the more specific ophthalmological diagnosis which usually provided the morphology of the cataract.

Source of referral	Exact diagnosis agreement
GP (n=131)	73 (56%)
Optometrist (n=366)	244 (67%)

Table 3.11. Exact diagnosis agreement for GPs and Optometrists: The referrer must get the exact diagnosis correct instead of within the same diagnosis category to count as an agreement.

Despite 29% of referrals from optometrists being defined as false positive compared to 3% from GPs, the proportion of patients discharged, and the proportion of agreeing diagnoses were very similar. Harrison and colleagues (1988) also compared primary and secondary care diagnoses with an agreement of 60% from GPs and 75% from Optometrists, which is comparable to the current study ('Exact diagnosis agreement' in Table 3.11 above). It should be noted that (as detailed in chapter 3.2) over the same time period the overall proportion of referrals from optometrists has increased by 33% relative to GPs. Sheldrick et al (1992) looked at only GP referrals and found slightly poorer diagnosis agreement, with non-accordance in diagnosis being 42%. The main problem with comparing these three studies is that it is hard

to determine from the papers exactly how stringent each author has been when determining diagnosis agreement.

As discussed in the introduction, we are mainly assessing clinical diagnostic decision making. It is possible for a clinician to have made a valid decision to refer for a second opinion and not have given a tentative diagnosis. This is another reason why all three measures of referral accuracy have been presented as “discharged at first appointment” may discriminate least in these instances. Another alternative measure of referral validity that could be used in a prospective study would involve a consultant ophthalmologist grading every new patient they saw in clinic as “valid referral” or “not-valid referral” according to their clinical judgement. Of course this would also still be subject to bias.

3.5b. Referral format

The direct referral schemes, for example the cataract choice service, were significantly less likely to result in false positive referrals. The reasons for this are discussed in detail in chapter 4, and this is further validation of direct referral schemes which usually require additional funding. Fewer false positive referrals reaching the hospital may result in an overall saving for the NHS dependent on the fee paid for enhanced services.

3.5c. Visual disturbances /other disorders of eye and adnexa

The indiscriminate diagnosis category of visual disturbance/other results in the highest proportion of false positive referrals, lowest diagnosis agreement and highest proportion of discharged patients. The nature of the category means it is probably a 'dumping ground' for patients that have non-specific symptoms where it is not possible to come up with a better diagnosis and therefore could be a reason for lower referral accuracy. The most common outcome for a patient being referred with this diagnosis category was that no pathology was found (13 referrals, 21%). Understandably a greater proportion of GP referrals fall into this category (23% of GP referrals and 7% of Optometrist referrals) as they have less facility, experience and confidence (Featherstone et al., 1992) to investigate eye pathology. Previous studies have regarded the ophthalmology content of undergraduate medical training as inadequate (Featherstone et al., 1992, Sheldrick et al., 1993, Shuttleworth and Marsh, 1997, Dayan et al., 1997). This raises the question as to whether optometrists are better placed as the first point of contact for primary eye care, which is a similar situation to the recently initiated GOS contract in Scotland (2006).

3.5d. Pre-registration optometrists

Pre-reg optometrists had the highest false positive referral rate (Table 3.1) and this makes some sense as they are inexperienced and likely to be more cautious.. However, the referrals from pre-registration optometrists are ultimately the responsibility of their supervisor and the amount of control that the supervisor wishes to impose and this complicates interpretation of the results. Diagnosis

agreement, for example, is the same for pre-reg students and qualified optometrists (Table 3.1). In addition, not enough pre-reg referrals were received to make firm conclusions.

3.5e. Years since registration

Univariate analysis showed that a more experienced clinician was significantly less likely to generate a false positive referral, and was significantly more likely to get their diagnosis correct. This seems logical, and improvement in diagnostic proficiency with increasing experience has been shown before across various medical disciplines (Meyer et al., 2010, Moss et al., 2005, Morton and Mackie, 1998), but it is a novel finding for optometry. If an inexperienced optometrist is unsure of a diagnosis, it would be unfair and potentially dangerous to criticise or discourage referral as there is a natural learning curve with experience in any profession. Of course to fully understand the situation it is desirable to quantify the numbers of patients that are not referred but should have been (false negative).

The multiple regression performed with data of optometrist gender, years since registration and practice type found a significant reduction in false positive referrals with experience, however there was no significant effect with diagnosis agreement. This indicates less experienced clinicians are generating more false positive referrals, but not necessarily providing referrals with less correct diagnoses. This again may be a case of referring pathology that is insignificant and does not need hospital attention according to the hospital clinician. An example from these data is a patient who was referred on the basis of “vitreous changes” with no symptoms,

and therefore was in the category H43-H45 Disorders of vitreous body and globe. The hospital clinician diagnosed a posterior vitreous detachment (PVD), using only techniques that are commonly used by the referring clinician, and therefore this was defined as a false positive referral because the pathology was not significant. Both clinicians had diagnosed pathology in the same category. This issue could possibly be addressed with training.

3.5f. Gender

There was a difference between referrals made by optometrists of different genders, with female optometrists generating significantly more referrals defined as false positive. There was also a greater proportion discharged at the first appointment and having fewer diagnoses agreeing with the hospital clinician. The analysis also showed no interaction between years since registration and referrer gender, which means the gender effect is not related to experience.

It has been previously documented that behaviour and decision making is different between male and female physicians, with the rates of screening (Lurie et al., 1993, Franks and Clancy, 1993, Kreuter et al., 1995, Bertakis et al., 1995, Shokar et al., 2009), referral (Boulis and Long, 2004) and the likelihood of initiating or intensifying treatment (Schmittdiel et al., 2009, Arouni and Rich, 2003) being higher for female doctors in the majority of studies. This appears to hold true in primary care optometry. Lurie and colleagues (1997) attempted to find a reason for these differences and discovered that female physicians felt more personal responsibility for ensuring that their patients received screening, and reported more comfort in performing Pap smears and breast examinations. Similarly, female physicians were

seen as more caring by patients and wrote longer referral letters (Parker and Hyett, 2009). A recent survey of 808 eyecare practitioners, of which 54% were women, found that women were more likely to agree with the statement that 'I feel vulnerable to the possibility of litigation in relation to my work' (31% vs 24%, Ewbank, 2010). Similar investigation into the referral behaviour of optometrists is required in order to ascertain why this difference exists.

3.5g. Type of optometric practice: Multiple and Independent

Referrals from independent practice result in fewer false positive referrals than those from multiples. Section 3.1.4d indicates that this is because multiples tend to employ younger staff and, to a lesser extent, more female staff. There has been some investigation into differences between high and low volume medical practices with mixed results. Curran and colleagues (2005) found higher-volume medical practices to be more likely to screen for prostate cancer, whereas a previous study found high-volume practices less likely to schedule well (preventative) care (Zyzanski et al., 1998). Zyzanski et al also found high-volume physicians to naturally have 30% shorter visits and lower up-to-date rates of preventative services. The contradictions within the literature indicate that this inequality requires further investigation. The business structures of both modes of practice also differ and therefore the commercial pressures on clinicians may be different. As previously discussed, there is no financial disincentive to generating an inaccurate referral, indeed in an increasingly litigious modern society there is a potential financial incentive to refer whenever there is any element of doubt.

3.5h. The influence of the Shipley Ophthalmic Assessment Project (SOAP)

The referrals to SOAP had to be included in the analysis as the aim was to obtain an accurate representation of what happened to the referrals being generated by primary care practitioners in Bradford and Airedale and similar services have existed nationally for some time (Henson et al., 2003). It could be argued that these referrals should be considered separately to the ones that go directly to the hospital as the presence of such a triaging service may influence referral behaviour. All GPs within the SOAP catchment area are aware of the service and a sizeable proportion of the optometrists are also. Having such a service available may make primary care clinicians more inclined to refer rather than investigating further or managing in practice as, compared to the hospital, SOAP is cheaper for the NHS and more convenient for the patient. Being inclined to refer more 'borderline' cases would theoretically increase the proportion of false positives, and this is supported by the data (Table 3.7) with referrals to SOAP being significantly more likely to be deemed false positive, discharged at the first appointment and have a diagnosis different to the one given by the SOAP clinician (all $p<0.0001$).

SOAP, and similar services nationally, work to strictly defined protocols for restricted pathologies and is only a triage rather than a treatment service. This means that any patients that definitely need treatment should not really end up there and will be referred direct to hospital, whereas more borderline referrals may go to SOAP. This is an additional reason why patients referred there are more likely to have been referred unnecessarily. 'Filtering out' these unnecessary referrals is exactly why the service was set up and therefore it appears to be working well.

With these data, it is also saving the NHS money, as every SOAP appointment costs £67, whereas first hospital appointments are £110. It should be pointed out that no false negative assessment of SOAP has been made. The question arises as to whether this could be done in practice? The clinicians working at SOAP have always been either experienced optometrists or GPs with a special interest in Ophthalmology and the equipment available to them is no different to normal optometric practice.

3.6. Limitations of the study.

A major limitation of the study was the lack of investigation of false negative decisions. i.e., patients who were not referred who should have been. Low false positive referrals must only be strived for without generating an increase in false negative decisions. This omission was due to logistical and financial difficulties of sourcing a significant sample of patients not referred, obtaining agreement for their re-assessment and arranging for an ophthalmological examination (the gold standard). Due to the low prevalence of referable eye disease within an optometric patient population of approximately 2-6%, a very large sample size is needed to evaluate false negative decisions. There is a minimal literature on false negative decisions within optometry and the only figure that was found via a PubMed Literature search was 1.13% (Newcomb & Potter, 1981) which was based on retrospective clinical judgement rather than a more robust re-examination strategy. Using Peduzzi and colleagues' (1996) sample size formula of N=10k/p, where k is the number of covariates accounted for and p is the likely proportion of positive cases,

this suggests a sample size of 885 assuming one covariate is assessed. Another limitation is the lack of a gold standard measure for accuracy of referral and this makes it very difficult to compare results to previous studies and to get agreement on the factors influencing accuracy. We have attempted to counteract this by providing three assessments of accuracy, but agreement on a gold standard would be greatly beneficial to the research in this area. Secondly and as discussed above, the presence of SOAP in this study will likely have affected referral behaviour of the clinicians in the catchment area and this will likely make the results not representative of all PCT areas. Of course, this is a problem for many areas given the huge variability of provision of enhanced ophthalmic services across England. Other limitations include that 'years since registration' was used as a measure of experience, yet this does not consider working practices (full time or part-time) and career breaks. It may have been preferable to have documented 'full-time equivalent years of practice since qualification' or similar instead of or in addition to 'Years since qualified'. Gender was not ascertainable in 85 cases as the name was illegible or not provided, but lack of referrer on optometric referral forms is unfortunately reasonably common (as discussed in chapter 1.2).

3.7. Conclusions

Clinician experience has the greatest effect on referral accuracy, particularly in terms of false positive referrals (sections 3.4d, 3.5e) and this seems logical. As practitioners become more experienced, they appear to become more confident about their ability to monitor patients rather than refer them. We assume that

more experienced clinicians will not make more false negative decisions, but this needs to be determined. Because multiple practices tend to recruit a greater number of younger clinicians, they tend to give more false positive referrals, but this effect appears to be purely due to the demographics of their staff. There is also a significant effect of gender on referral accuracy with women tending to refer more false positives and this appears to be due to a different approach to patient care (section 3.5f) and possibly a greater sensitivity to litigation (Ewbank, 2010). Once again, possible differences in false negative decisions with gender need to be determined.

Chapter 4. Assessment of referrals to the hospital eye service by optometrists and GPs in Bradford and Airedale

Published as: Davey CJ, Green C & Elliott DB. Assessment of referrals to the hospital eye service by optometrists and GPs in Bradford and Airedale. *Ophthalmic and Physiological Optics* 2011, **31:** 23–28.

4.1. Introduction

This study investigated the content of referrals to the Bradford Royal Infirmary Hospital Eye Service (HES) and describes differences between referring clinician (optometrist or general practitioner, GP) and referral formats. New patients attending the HES in the United Kingdom are referred there from optometric practice (39-57% of referrals) or by GPs (43-49% of referrals), (Pierscionek et al., 2009, Pooley and Frost, 1999, Harrison et al., 1988) unless they have self-referred to Accident and Emergency clinics. The quality of referrals from optometrists (Pooley and Frost, 1999, Harrison et al., 1988, Lash, 2003) particularly those for suspect glaucoma (Scully et al., 2009, Theodossiades and Murdoch, 1999, 2007, Newman et al., 1998, Bell and O'Brien, 1997a, Bowling et al., 2005, Sheldrick et al., 1994, Vernon, 1998) have been analysed periodically over the last 30 years. Researchers have used these reports to highlight inadequacies in the HES referral system that they subsequently attempted to improve (Henson et al., 2003, Bourne et al., 2009, Vernon and Ghosh, 2001, Theodossiades et al., 2004) and it is therefore useful to

continue this periodic review. In addition, very limited information is available that compares referrals, from optometrists and GPs, in terms of the types of condition referred, information provided or accuracy of tentative diagnoses. A very recent study by (Pierscionek et al., 2009) reported such information for referrals to a (non-hospital based) ophthalmology practice in Northern Ireland, but the results may not be typical of those found for the UK HES (this is discussed in detail in the discussion), where the great majority of patients are referred.

GP referral letters are usually computer generated and printed, whereas optometrist referrals can be written on a General Ophthalmic Services (GOS) 18 form, a direct referral form, a shared care form or a typed or handwritten letter, and these can differ depending on local schemes. In Bradford there is also an adapted GOS 18 form developed in collaboration with local ophthalmologists and aimed at improving optometric referral quality (Naru and Green, 2009). It differed in that there was no area for the GP to complete as local GP alliances informed the development team that this was not used anymore. It also included a number of tick boxes for the most commonly referred pathologies, so the hospital clinician can see at first glance what the referral reason is. Despite some authors advocating the use of letter writing to refer effectively (Clarke, 2008) the GOS 18 form remains the most popular referral method (69 - 73% of referrals; (Pooley and Frost, 1999, Scully et al., 2009). The study also compared the content of referrals based on the referral format and whether the adapted GOS 18 form led to improvements in referral quality.

4.2. Methods

Ethical approval was gained from the Bradford Research Ethics committee and the Bradford Hospitals NHS Foundation Trust Research Office. A random sample of 445 new referrals to Bradford Royal Infirmary ophthalmology department during 2007 and 2008 were retrospectively analysed. The presence of the following information was recorded from the referral: patient name and address, date, referrer name, referral format, referrer address, refraction, visual acuity, ophthalmoscopy, visual field assessment, tonometry, pupils, co-morbidity, slit lamp examination, symptoms, family history and any diagnosis given or alluded to (classified based on the International Classification of Diseases-10 (ICD-10), World Health Organisation). If the referral was handwritten then the legibility of the referral was graded by one person (CD) as; fully legible, illegible in part but understandable overall, or not legible enough to understand the reason for referral.

Although there are many aspects of a good referral that are essential for all pathologies, the content will vary depending on the referral reason. For example, the Department of Health paper “Action on Cataracts” (2000) recommends that referrals should contain a confirmation of the presence of cataract that is affecting vision, an indication that the reduced vision was having a detrimental effect on the patient’s lifestyle and an indication that the patient was willing to undergo surgery. However, the literature concerning glaucoma referrals agrees that the best referrals contain a triad of information including assessments of visual fields, optic discs and intra-ocular pressures (Scully et al., 2009, Theodossiades and Murdoch, 1999, 2007, Newman et al., 1998, Bell and O'Brien, 1997a, Bowling et al., 2005, Sheldrick et al.,

1994, Vernon, 1998). For this reason, referrals of certain categories of ocular disease were assessed using disease specific recommendations.

4.3. Results

4.3a. Referring Clinician

12 of the 445 referrals were excluded as the GP letter suggested that they originated from an optometric practice, but either the GP had not enclosed the optometrists' referral letter or it had been lost. Of the remaining 433 referrals, 311 (72%) from optometric practice and 122 (28%) were from general practice. Of the 311 referrals from optometric practice, 231 were from optometrists and 17 were from pre-registration graduates. It was not possible to identify whether 59 of the referrals from optometric practice were from Optometrists or pre-registration graduates. It is unlikely, but not impossible, that the referrals are from Ophthalmic Medical Practitioners (OMPs) as none are registered with the local Primary Care Trust as performers. As the supervising optometrist takes legal responsibility for all referrals from a pre-registration graduate, all these referrals were combined as "optometric referrals". It was not possible to ascertain the type of clinician from 11 referrals, which were all Diabetic Retinal Screening Service (DRSS) referrals and were probably from DRSS technicians and not included in the GP/optometrist comparison.

4.3b. Referral Format

All 114 referrals from GPs were typed and computer generated. The seven DRSS referrals from General Practice were on the relevant DRSS proforma. The format of the 311 referrals from optometric practice is shown in Table 4.1.

Referral Format	N (%)
Old GOS18	124 (40%)
New GOS18	115 (37%)
DRSS form	33 (11%)
Optometrist letter – typed	17 (5%)
Cataract CHOICE	13 (4%)
Optometrist letter – handwritten	4 (1%)
Glaucoma Monitoring Scheme	3 (1%)
Practice Specific Proforma	2 (1%)

Table 4.1. The referral format of 311 referrals from optometrists to the Hospital Eye Service in Bradford (Key. GOS18: General Ophthalmic Services 18 form. DRSS: Diabetic Retinal Screening Service).

4.3c. Legibility

All GP referrals were fully legible as they were all typed and either securely transmitted electronically or faxed. Of the 311 referrals from Optometric practice, 29 had an illegible referrer name and 26 had no referrer name present. Of the 275 handwritten referrals from Optometric practice, 202 (73%) were fully legible, 71 were illegible in part but with the general meaning intact (26%), and two were fully

illegible. Illegibility was similar for the different types of form used (old or new GOS18 or individual letter).

4.3d. Referral diagnoses

The range of primary diagnoses given in referrals from optometric and general practice is given in tables 4.2 and 4.3 respectively using categories from the ICD-10 of the World Health Organisation.

Diagnosis in referral:	N (%)
Disorders of lens	84 (27%)†
Primary open angle glaucoma suspect	61 (20%)
Diabetic retinopathy	32 (10%)
Age related macular degeneration	23 (7%)
Visual disturbances and other disorders of eye and adnexa	20 (6%)
Other disorders of retina	17 (5%)
Disorders of the cornea	14 (5%)
Disorders of eyelid, lacrimal system and orbit	12 (4%)
Disorders of vitreous body and globe	11 (4%)
Retinal vascular occlusions	10 (3%)
Disorders of muscles, binocular movement/vision, amblyopia, accommodation and refraction	8 (3%)
Retinal detachments and breaks	7 (2%)
Disorders of Optic nerve and visual pathway	3 (1%)
Disorders of conjunctiva	3 (1%)
Angle closure, secondary and other Glaucoma	2 (1%)
Disorders of choroid	2 (1%)
No diagnosis given in referral	2 (1%)

Table 4.2. Primary diagnoses given in referrals from Optometric practice (n=311) and classified using the International Classification of Diseases (ICD) 10th revision.

† includes 13 referrals for posterior capsular opacification and 1 for lens subluxation

Diagnosis in referral:	N (%)
Disorders of eyelid, lacrimal system and orbit	52 (46%)
Visual disturbances and other disorders of eye and adnexa	21 (19%)
Disorders of lens	12 (11%)
Disorders of conjunctiva	10 (9%)
Disorders of muscles, binocular movement/vision, amblyopia, accommodation and refraction	7 (6%)
Age related macular degeneration	4 (4%)
Disorders of the cornea	2 (2%)
Disorders of vitreous, globe, sclera, optic nerve, visual pathway, iris, ciliary body, retinal breaks and other disorders of retina.	6 (5%)

Table 4.3. Primary diagnoses given in referrals from GPs in General Practice (n=114) and classified using the International Classification of Diseases (ICD) 10th revision.

4.3e. Ophthalmoscopy

Five of 114 GP letters (4%) commented on the fundus. Two of them cited direct ophthalmoscopy as the method used whereas the remaining three did not comment on the technique used. 291 of 311 optometric letters (94%) commented on the fundus. In most cases the technique used was unspecified (179, 58%), but the following techniques were reported: Dilated indirect ophthalmoscopy (53, 17%), direct ophthalmoscopy (34, 11%), retinal photography only (22, 7%) and undilated indirect ophthalmoscopy (3, 1%).

4.3f. Visual Acuity (VA)

GPs included VA results in 9 referrals (8%, n=114) and VAs were included in 299 referrals from optometric practice (96% n=311).

4.3g. Lids and Lacrimal Disorder Referral Quality

Fifty-two referrals for lids and lacrimal disorders were from General Practice, and 12 were from Optometric practice. GP referrals mainly provided symptoms (N=49, 94%), with VA provided in 2 cases (4%) and no reports of any examination beyond direct observation. Topical drugs were prescribed in 8 cases (15%). Optometric referrals typically provided VA results (12, 100%) and symptoms (11, 92%) and often indicated that a slit-lamp examination had been performed (9, 75%), but did not report prescribing drugs.

4.3h. Cataract Referral Quality

There were 12 referrals from GPs, although two had been previously diagnosed with early cataract by ophthalmologists. GP referral letters for cataract usually included information about the effect on the patient's lifestyle (11, 92%), occasionally that they were willing to have surgery (4, 33%), but rarely a report of the cataract (1, 8%) or fundus (1, 8%) and no reports contained VA or refractive error information. One of the 12 referrals (8%) could be said to have completely followed the Department of Health "Action on Cataracts" (2000) recommendations for referrals, yet 9 of the 12 referrals (75%) led to cataract surgery.

There were 61 referrals for cataract from optometrists (table 4). Optometrist referral letters usually included information about the cataract (61, 100%), VA (60, 98%), refractive error (57, 93%) and fundus (59, 97%), but less often information about the effect on the patient's lifestyle (39, 64%) or willingness for surgery (30, 49%). Thirty of the 61 referrals (49%) followed completely the Action on Cataracts recommendations and 45 of the 61 referrals (74%) led to cataract surgery.

Referrer/Format	Lens report	Visual acuity	Affecting lifestyle	Fundal report	Refractive error	Resulted in surgery
Cataract CHOICE	8	8	8	8	8	8
(n=8) †	(100%)	(100%)	(100%)	(100%)	(100%)	(100%)
Old GOS18	32	31	21	31	30	23
(n=32)	(100%)	(97%)	(66%)	(97%)	(94%)	(72%)
New GOS18	16	16	5	15	14	10
(n=16)	(100%)	(100%)	(31%)	(94%)	(88%)	(63%)
Optometrist letter	5	5	5	5	5	4
(n=5)	(100%)	(100%)	(100%)	(100%)	(100%)	(80%)

Table 4.4. Contents of 61 cataract referrals from optometrists to the hospital eye service by referral format (Key. GOS18: General Ophthalmic Services 18 form).

† In 6 cases, an older cataract CHOICE form, which included data obtained by phone by a patient care co-ordinator was used and these data are not presented. Three other referrals were submitted on Diabetic Retinal Screening Service (2) or Glaucoma Monitoring Service (1) forms.

4.3i. Glaucoma Referral Quality

No new referrals were made by GPs for suspect glaucoma. Two glaucoma referrals from GPs were excluded as the patients had been previously diagnosed with glaucoma, but had been lost to recall by the hospital. The two patients already on the Glaucoma Monitoring Enhanced Service were excluded as they had also been previously diagnosed. There were 61 referrals from Optometric practice, of which 47 (77%) gave the recommended triad of assessment for suspected chronic open-angle glaucoma of disc assessment with IOP and visual field measurement, 12 (20%) gave disc assessment with tonometry readings and two (3%) gave IOP measurements only. Non-contact tonometry was reported in 33 cases (54%) and contact tonometry in 8 cases (13%), with the remainder being unspecified. Of the 31 patients with defective visual fields, 3 had obviously had visual fields repeated. A positive family history was reported in 13 cases (21%). It should be noted that all these referrals occurred prior to the National Institute for Health and Clinical Excellence or NICE guidelines on referral of patients with suspect chronic open-angle glaucoma or ocular hypertension (April 2009).

4.4. Discussion

4.4a. Referring Clinician

Compared to studies in the 1980s and 1990s, the proportion of referrals to the HES from optometrists appears to be increasing significantly (1988: 39%,(Harrison et al., 1988) 1999: 48%,(Pooley and Frost, 1999) present study 72%). There is no obvious

reason to suspect that GP and optometrists referrals to the HES in the Bradford and Airedale region would be different to other parts of England, but will probably differ from Scotland and Wales due to the enhanced GOS contracts agreed there. The recent study from Pierscionek, Moore & Pierscionek described data from a non-HES ophthalmology practice in Northern Ireland (Pierscionek et al., 2009) and may be atypical as discussed below. Even then, they report optometric referrals at 57%. This relatively increased role of optometry in the referral process is likely due to the increased links between optometry and ophthalmology over the last 20 years and highlighted by the various direct referral and co-management schemes throughout the UK.

The majority of the 114 referrals from GPs (table 4.3) were for symptomatic disorders of eyelid, lacrimal system and orbit (N=52, 44%), of which 20 were lid cysts or lesions, 17 were symptoms of a lid/lacrimal disorder and 6 were entropion/ectropion or trichiasis. Other ocular abnormalities diagnosed were either symptom-based (such as “poor” or “distorted” vision, N=21, 19%), cataract (N=12, 11%) and conjunctival diseases (N=10, 9%). In the majority of cases, the referral letters suggest that GPs diagnosed anterior segment diseases on the basis of direct observation (i.e. without slit-lamp or ophthalmoscopy) and symptoms, with rare provision of visual acuity data (N=9, 8%) or mention of the use of a direct ophthalmoscope (N=5, 4%). This summary differs from the view provided by Pierscionek, Moore & Pierscionek (Pierscionek et al., 2009) who reported referrals from GPs and optometrists to a non-HES ophthalmology clinic in Northern Ireland (NI). For example, we found no referrals for suspect glaucoma originating from GPs, which is similar to the majority of the literature with reported values from 0-10%

(Bowling et al., 2005, Newman et al., 1998, Sheldrick et al., 1994, Theodossiades and Murdoch, 1999, Bell and O'Brien, 1997a). However, Pierscionek, Moore & Pierscionek (Pierscionek et al., 2009) reported that 31% of glaucoma referrals in their study originated from GPs. In the current study there were 12 referrals where the GP indicated that they had originated in optometric practice but did not have the original optometrists' referral included in the GPs' correspondence. This problem was also identified by Pooley and Frost (Pooley and Frost, 1999) and could be an explanation for the disparity between the data from the NI study and the rest of the literature. i.e. perhaps some of the glaucoma referrals from GPs in the NI study were originally referred by optometrists. Alternatively, the NI study could have included referrals from a GP with a special interest in ophthalmology, and therefore with access to visual field analysers and tonometers and with significant experience of disc assessment. However, one would have expected this information to have been provided in the report and it was not. The NI study also seems atypical in that the third most common cause for referral by GPs was retinal problems (N=28, 12%) and 19 were reported to have provided a correct diagnosis. In this study, six patients (5%) were referred for retinal problems by GPs. Three of these patients had previously been diagnosed with age-related maculopathy (AMD; one by their optometrist who had provided an Amsler chart for home monitoring and two by the HES who had subsequently discharged the patients), two were referred on the basis of symptoms only ("flashes and floaters" with a diagnosis of posterior vitreous detachment (PVD), and "reduced vision" with a suggested diagnosis of AMD) and one was referred for a retinal detachment with PVD based on symptoms and ophthalmoscopy, for whom the HES subsequently diagnosed as cataract only.

Optometric referrals (see Table 4.2) were principally for disorders of the lens (27%), glaucoma or suspect glaucoma (20%) and diabetic retinopathy (10%). These referral patterns have been described before (Lash, 2003, Pooley and Frost, 1999, Pierscionek et al., 2009) and do not appear to have changed substantially. The optometric cataract referrals were all grouped together, but should perhaps have been divided into two groups: one set of referrals arising from a GOS 'sight test' as funded by the NHS and a smaller group from a CHOICE direct referral pathway where additional funding was provided to accredited optometrists by the local Primary Care Trust (PCT) for performing a dilated fundus examination, discussing cataract surgery with the patient and completing a cataract referral form. This is a form of referral refinement requiring a higher level of referral quality and uses a cataract-specific proforma. Perhaps not surprisingly, 100% of the eight referrals via the CHOICE pathway (the service was introduced during 2008 so numbers were relatively low) provided the information indicated by the Action on Cataracts recommendations and all eight underwent surgery. Referrals arising from GOS tests were more variable. GOS referral letters virtually always included information about the cataract, VA, refractive error and fundus (all above 88%), but less often information about the effect on the patient's lifestyle (26/48, 54%) or willingness for surgery (23/48, 48%). Part of the reason for this is that some optometrists clearly did not view a discussion of the need for surgery as part of the remit of the GOS sight test. Indeed, one GOS referral for cataract included the following: "Please arrange referral to ophthalmology to discuss the pros and cons of cataract surgery". This latter approach is in agreement with the obligations of the GOS contract, but is clearly not followed by those optometrists who reported whether the patient was

willing to undergo surgery. There is a need for all optometrists to be clear about what their requirements are for cataract referral as part of a GOS sight test.

The recommended information provided in a referral for suspect glaucoma includes an optic disc assessment and measurements of intra ocular pressure and visual fields in both eyes (Vernon and Ghosh, 2001, Crick and Tuck, 1995, Bell and O'Brien, 1997b, Sheldrick et al., 1994, Bell and O'Brien, 1997a, Newman et al., 1998, Theodossiades and Murdoch, 1999) and ideally, abnormal intra ocular pressures and visual field assessments should be repeated to avoid unnecessary false positive referrals. (Salmon et al., 2007, Crick and Tuck, 1995, Bell and O'Brien, 1997a, Vernon and Ghosh, 2001, Henson et al., 2003) Information regarding all three glaucoma assessments were provided in 47 of 61 optometric referrals (77%), with 12 referrals including just tonometry and disc assessment (20%) and two (3%) with just tonometry readings. Visual field measurements were repeated for 3 of 31 defective fields (10%) and 15% of referrals included contact tonometry readings, although it was not clear whether these were initial or repeated measurements. There is therefore a range of opinion about what a referral for suspect glaucoma from a GOS sight test should include, although the vast majority deem it to be either the triad of disc assessment, visual field and tonometry (77%) or just tonometry and disc assessment (20%). A small number (~ 10%) repeated visual field or tonometry measurements as part of the sight test and a small number just provided tonometry readings only (3%). Once again, there is a need for all optometrists to be clear about what their requirements are for referrals of patients with suspect glaucoma or ocular hypertension (OHT) as part of a GOS sight test.

In this regard, it is pertinent to highlight that a referral refinement service has recently been piloted in the Bradford area between 1/2/09 and 31/7/09 (Davey et al., 2010). This service funded repeat measurement of IOP by applanation and threshold visual field assessment for patients that would have otherwise been referred for OHT or suspect primary open angle glaucoma. The pilot significantly reduced unnecessary referrals to secondary care and has since been rolled out over the rest of the PCT

4.4b. Referral Format

The majority of optometric referrals were made using GOS 18 forms, with 40% being made using the original GOS 18 and 37% with the newer locally adapted GOS 18. There appeared to be no difference in the information provided within the different formats and the locally adapted GOS 18 form showed no significant improvements in the information provided over the old GOS 18 (e.g., Table 4.4). The adapted form remained one that was attempting to be useful for all types of referral and it may be that improvements (such as those provided by the disease specific CHOICE forms (see Table 4.4) will only come about by the introduction of disease specific referral forms for the very common referrals of cataract and suspected glaucoma or ocular hypertension, with a generic GOS 18 being used for other referrals. Fully illegible forms are not a major problem, however over a quarter of handwritten referrals were illegible in part, which may cause delay or inaccuracy when the referral gets to the HES. It is not clear if reduced legibility was due to poor handwriting or poor quality due to scanning, faxing or self-carbonating

copies. This represents an improvement over 1999 findings which reported that only 33% of optometric referrals were read without difficulty (Pooley and Frost, 1999). Word processing would naturally improve legibility and all GP referrals were fully legible as they were computer generated. A surprisingly small proportion of optometric referrals were made using a word-processed letter (5%). The number of optometric practices using electronic patient records in this area is unknown, but as more practices switch to a paperless system perhaps a nationally adopted electronic referral template (or series of templates for different referral types as discussed above) could gradually replace the ageing GOS 18.

4.5. Conclusion

The proportion of patients attending a HES department that have been referred by an optometrist rather than a GP is increasing. Optometrists refer patients with a wide range of ocular diseases and in most cases include fundus observations and visual acuity measurements in their referrals. GPs mainly refer patients with anterior segment disorders, particularly lid lesions, based on direct observation and symptoms. GP referral letters include all relevant non-clinical information and are all perfectly legible, whereas illegibility and missing clinical information remains a problem in optometric referrals. This could be minimised by word-processed referrals, and in the future, direct electronic referral. The introduction of disease-specific referral forms for suspect glaucoma and cataract, rather than the standard GOS 18 referral form, could also help to improve referral quality. There is significant variability in the provision of information in optometric referral letters to the HES

for the common conditions of cataract and suspect chronic open-angle glaucoma or ocular hypertension and there is a need to determine and disseminate what exactly should be provided as part of a GOS sight test.

Chapter 5. Rasch analysis of the Hospital Anxiety and Depression Scale (HADS) and State-Trait Anxiety Inventory (STAI) in a population of new ophthalmic outpatients

5.1. Introduction

As discussed in section 1.4, the Hospital Anxiety and Depression Scale (HADS) and State Trait Anxiety Inventory (STAI) are two questionnaires aimed at detecting levels of psychosocial distress. These instruments have not been evaluated before on a population of new ophthalmic outpatients. In other populations, the usefulness of these questionnaires has been determined using evaluations of dimensionality (does the questionnaire measure one factor or several?), internal consistency of results (usually assessed using Cronbach's alpha), concurrent validity (how well results correlate with gold standard measurements) and assessments of results using Rasch analysis.

5.1a Hospital Anxiety and Depression Scale

Factor structure and internal consistency of the HADS

Two published literature reviews suggested that when using factor analysis, the large majority of studies discovered two main factors within the HADS (Herrmann, 1997, Bjelland et al., 2002), although a few studies reported three (Razavi et al., 1989, Brandberg et al., 1992, Leung et al., 1993, Dunbar et al., 2000, Lewis, 1991,

Martin and Thompson, 2000) or even four separate factors (Andersson, 1993, Martin and Thompson, 1999). The large majority of studies from both reviews found the fourth anxiety item to have relatively low loadings on anxiety and unexpectedly high loadings on depression. Combining the two reviews gives a range of internal consistency (Cronbach's alpha coefficient) of 0.68 to 0.93 for HADS-A and 0.67 to 0.90 for HADS-D, which is acceptable.

Recent research, published since the Bjelland review, has contradicting opinions on the number of factors present. Half of the studies conclude that the HADS fits a two-factor model (Smith et al., 2002, Johnston et al., 2000, Mykletun et al., 2001, Flint and Rifat, 2002) and half argue that their results fit a three-factor model (Jomeen and Martin, 2004, Martin et al., 2004, Caci et al., 2003b, Friedman et al., 2001). Caci et al. (2003b) for example, used factor analysis to find that the original Anxiety subscale could be split into two components that they labelled 'Anxiety' and 'Restlessness', while the original Depression subscale was slightly modified. Cronbach's alpha coefficients for these studies ranged between 0.73 and 0.84 for both HADS-A and HADS-D, although HADS-D had one outlier at 0.49 (Caci et al., 2003b). Internal consistency for HADS-T varied between 0.81 and 0.86.

Concurrent validity of HADS

Bjelland, Dahl, Haug & Neckelman (2002) revealed that when compared to other questionnaires for anxiety and depression (e.g., Beck's Depression Inventory, STAI, various GHQ versions, etc.) the correlations to HADS-D and HADS-A, respectively,

were between 0.60 and 0.80. The authors therefore concluded that the concurrent validity of HADS is 'good to very good'.

Rasch analysis of HADS

Only four studies to date have published the results of Rasch analysis on the HADS, therefore they have been individually summarised below.

Smith and colleagues performed Rasch analysis using Winsteps software on both the total score (HADS-T) and the anxiety (HADS-A) and depression (HADS-D) subscales (Smith et al., 2006) with 1855 patients. The study reported that all three scales were unidimensional although they used a cut-off figure of 3 or below for the eigenvalue of the second factor. This value is occasionally used (e.g., Forjaz et al., 2009), but the more commonly used cut-off figure is 2.0 (Pesudovs et al., 2007) which includes an earlier paper by Smith (2002). Pallant and Tennant performed Rasch analysis using RUMM2020 software on the HADS-T to measure the appropriateness of using it as a measure of generalised psychological distress in 296 outpatients undergoing musculoskeletal rehabilitation (Pallant and Tennant, 2007). For comparative purposes, brief results for HADS-A and HADS-D were also given. The questionnaire was reported to split into two dimensions of essentially anxiety and depression, but the authors argue that the two scores can be validly combined. Tang and colleagues performed Rasch analysis using Winsteps software on the HADS-D subscale to examine the optimal scoring scheme for 100 Chinese stroke survivors (Tang et al., 2007). The low sample size used limits the findings of this

study, plus their assessment of unidimensionality was limited to the fit of the items within the questionnaire and not factor analysis.

The different studies reported different individual items misfitting. For example, Pallant and Tennant (2007) found that HADS-D had no misfitting items whereas HADS-T and HADS-A did not fit the Rasch model for one item. Smith et al (2006) found three items misfitting in HADS-T, but these were different items to those reported by Forjaz et al. (2009). This could be due to different types of subject involved in the studies and possibly different levels of anxiety and/or depression. Smith et al. (2006) suggested that the HADS-T scale showed a floor effect, meaning lower levels of psychological distress could be underestimated. However, Pallant and Tennant (2007) reported that the residual mean value for persons for HADS-T indicated no serious misfit among the respondents in this study, and therefore no floor effect.

5.1b State Trait Anxiety Inventory

Factor structure and internal consistency for STAI

In the manual for the revised STAI (1983) Spielberger identified that, due to an artefact of the structure of the instrument, “anxiety absent” and “anxiety present” factors were introduced by the “anxiety absent” and “anxiety present” questions. Higher scores tend to be more frequent on anxiety absent items (Kvaal et al., 2001, Mook et al., 1991, Mook et al., 1992) especially in Asian (Japanese) populations (Iwata and Higuchi, 2000).This is because confirming the presence of anxiety is not

the same as not confirming the presence of calmness. After factor analysis, Spielberger concluded that a four factor model fitted the data best; State Anxiety Present, State Anxiety Absent, Trait Anxiety Absent, and Trait Anxiety Present. A number of studies since have concurred with this (Hishinuma et al., 2000) or were able to differentiate between state-anxiety present and trait-anxiety present, but all anxiety absent items loaded on a single factor (Iwata et al., 1998). Vigneau and Cormier have argued that this may not be the best way to represent the data, because although anxiety-absent and anxiety-present items appear to load on different factors, they both should load on the same construct, it is only the method that differs. Alternatively, they propose a two construct, two method model as being statistically advantageous (Vigneau and Cormier, 2008).

Other authors have also conflicted with Spielbergers' proposed four factor model, with some proposing only the two method factors are present, meaning therefore that STAI does not differentiate between State and Trait anxiety (Kabacoff et al., 1997). Two non-method factors were found by Kvaal and colleagues (2001) in the State subscale, described by the authors as "well being" and "nervousness". Factor analysis has led a number of studies to conclude that the STAI loads significantly on depression as well as anxiety, both for the whole questionnaire (Gros et al., 2007), and the Trait subscale (Bieling et al., 1998, Caci et al., 2003a). This was supported by Kennedy and colleagues (2001), who found that the STAI does not differentiate anxiety disorders from depressive disorders. This is to be expected and almost impossible to avoid, as although anxiety and depression are distinct constructs, they have somewhat overlapping features (Endler et al., 2003).

A recent review by Barnes and colleagues (2002) found average internal consistency to be acceptable where reported. This is supported throughout the literature (Kabacoff et al., 1997, Iwata and Mishima, 1999, Bieling et al., 1998).

Concurrent validity of STAI

There are many examples of the STAI being used as a 'gold standard' of anxiety questionnaires when assessing the concurrent validity of other psychometric instruments (Michopoulos et al., 2008, Soury et al., 2005, Deborde et al., 2004, Cleemput et al., 2004, Brouwers et al., 2001). Rarely is a reason given as to why authors use the STAI as a gold standard, with only occasional brief comments about it being "previously validated" (Brouwers et al., 2001). It must be assumed that it is used in this capacity simply because of a precedent being set, and its relatively good psychometric properties. Bieling and colleagues (1998) found that the depression and anxiety factors that they identified from the Trait subscale correlated well with Becks Depression and Anxiety Inventories and the Depression Anxiety Stress Scales. Concurrent validity has also been reported as good when compared to other measures of anxiety, namely HADS and Hamilton Anxiety Scale (Mondolo et al., 2007).

Rasch analysis of STAI

Very little is published on Rasch analysis of the STAI. Tenenbaum and colleagues published two reports in 1985 which contained Rasch analysis of STAI answers

completed by young athletes. Both studies used the old version (Form X) of the STAI. In the first study (Tenenbaum et al., 1985) the state (n=55) subscale was completed 30 minutes prior to an athletic event, and the trait (n=100) subscale was completed by a different cohort in a group setting. The authors drew similar conclusions for both cohorts; (a) the items are not spread equally enough along the anxiety continuum (b) no items are present to adequately assess participants with low anxiety (c) six items from each scale misfitted the model (d) many items are located at the same points on the continuum. Although this is all true, the study suffered from a small sample size, therefore reducing the credibility of their conclusion that "the STAI questionnaires are not accurate enough to differentiate the state and trait anxiety levels of the subjects".

Tenenbaum and Fursts second publication (1985) used Rasch analysis to see if it was possible to use the state subscale to retrospectively measure state anxiety. One cohort (n=55) completed the scale 30 minutes prior to an athletic event, and a second cohort (n=113) completed it retrospectively (timescale is not given) so that the low sample size again limits the usefulness of the results.

In summary, the STAI is influenced by depression, not purely anxiety, giving a measure of overall psychological well being, which is satisfactory for the planned study. We are not aiming to diagnose anxiety disorders or differentiate between psychological pathologies. It would however be very useful to assess whether a participant has high trait-anxiety or whether any raised levels of anxiety are being generated purely by their current situation. From the evidence provided, it is

arguable whether the STAI adequately discriminates between the two, although it remains the most widely used instrument that attempts to quantify trait-anxiety.

The aim of this study was to use Rasch analysis plus traditional statistical evaluation in order to assess their suitability to evaluate levels of anxiety and depression in a population of new ophthalmic outpatients. This was determined using the following assessments:

1) The distribution of responses to the categories of each item. i.e., floor and ceiling effects and percentage of missing data. This was further assessed using skewness and kurtosis measurements.
2) Rasch analysis of the response scales.
3) Rasch analysis of individual items in terms of their fit to the Rasch model.
4) Rasch analysis of the discriminative ability of each questionnaire.
5) Principal components analysis (PCA) of the residuals to determine the dimensionality of the questionnaires.

5.2. Methods

Potential participants were new patients who had an outpatient appointment booked at Bradford Royal Infirmary Eye Service between January 2008 and December 2008. The identification of these potential participants was via the hospital booking system. These patients were sent a covering letter, an information sheet, contact details of the research team and the two anonymised questionnaires (State Trait Anxiety Inventory, Hospital Anxiety and Depression Scale, see chapter

2.1 and Appendix B). The two questionnaires were placed in a random order within the envelope to attempt to eliminate order bias and were coded with a 4 digit number. A note on the information sheet indicated that it was also available in English and Urdu on an audio CD, although none of the participants requested the audio version. The information arrived by post at least 24 hours in advance of their appointment at the hospital. If the patients read the information and subsequently consented to participate they were requested to bring the anonymised questionnaires on the day of their appointment. The consenting patients were asked to complete the questionnaires, which should take less than 15 minutes, while they were waiting for their appointment. Completion of the questionnaires was taken as written implication of consent and a private room was available for patients if they wished to use it. When the participants were called for their appointment they were asked to hand the completed questionnaires, in the sealed envelope provided, to the clinician as they entered the consulting room. The clinician passed the questionnaires to a researcher at the end of the clinic, or at another convenient time. Identifying codes were used, instead of names, which were cross referenced at a later stage to access the results of their ophthalmic appointment, to determine whether they received a true or false positive referral.

The patients were sent a letter thanking them for their participation and informed that if they had ticked the relevant box on the questionnaire we would post them a copy of the study if and when published.

5.2a Inclusion criteria

Inclusion criteria for the referred cohort included:

- Initial referral from a GP or optometrist to the hospital eye service, within the proposed testing schedule of the study.
- Newly referred, so by definition, a patient who had been called back for review or requiring further investigation was not included.
- Aged over 16, as legally at this age a young person can be treated as an adult and can be presumed to have capacity to decide.

5.3. Results

321 pairs of questionnaires were returned from 1,854 posted, giving a response rate of 17%.

5.3a. Demographic information.

Demographic information regarding the patients who completed the questionnaires prior to their hospital eye appointment (patients referred) are displayed in Table 5.1.

Characteristic	Patients referred (n=321)
Mean age (years ± SD)	61 ± 19
Gender: Female	169 (53%)
Male	144 (45%)
Unspecified	8 (2%)
Ethnicity*: White	187 (58%)
Asian	39 (12%)
Black	6 (2%)
Not Stated	88 (27%)
Any other group	1(<1%)

Table 5.1 Demographics of the main cohort.

*White ethnicity included White (British), White (Irish) and white (other) and was predominantly White (British). Asian ethnicity included Asian (Indian), Asian (Pakistani), Asian (Bangladeshi) and Asian (other). A very small percentage were Black (African), Black (Caribbean), Black (other) and Chinese.

5.3b. HADS-T data

321 patients completed the HADS-Total. The item map for HADS-T is shown in Figure 5.1 and descriptive statistics for each individual item of the HADS-T showing reliability indicators and proportion of responses within each category are shown in Table 5.2. Analysing this dataset gave a Participant Separation Index (PSI) of 2.18, a Participant Reliability Index (PRI) of 0.82, an Item mean of 51.1 and a Participant mean of 35.6.

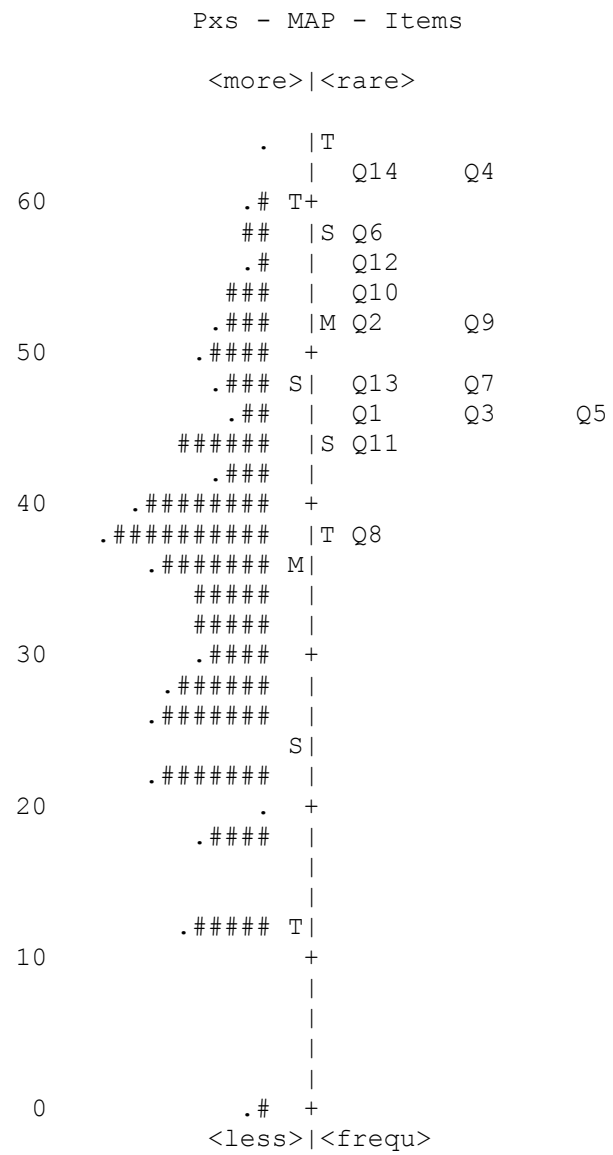

Figure 5.1. Item-Participant map for HADS-T. EACH '#' IS 3.

Items		Skew	Kurtosis	Missing data (%)	Response category 0 (Floor, %)	Response category 1 (%)	Response category 2 (%)	Response category 3 (Ceiling, %)	Infit	Outfit
1	I feel tense or 'wound up'	0.75	0.35	2	30	52	13	6	0.71	0.71
2	I still enjoy things I used to enjoy	0.98	0.75	1	48	42	8	2	1.03	1.06
3	I get a sort of frightened feeling as if something awful is about to happen	0.70	-0.66	2	45	28	21	6	1.14	1.08
4	I can laugh and see the funny side of things	**2.05**	**4.34**	2	72	22	4	2	**1.41**	**1.60**
5	Worrying thoughts go through my mind	0.73	-0.23	2	36	41	15	8	0.89	0.93
6	I feel cheerful	1.57	**2.22**	2	65	28	6	2	0.78	0.58
7	I can sit at ease and feel relaxed	0.42	-0.53	3	37	48	15	1	0.68	0.74
8	I feel as if I am slowed down	0.30	-0.61	2	19	44	27	10	1.05	1.07
9	I get sort of frightened feeling like 'butterflies' in the stomach	0.87	0.54	2	45	44	9	2	0.89	0.92
10	I have lost interest in my appearance	1.32	0.83	2	61	24	11	4	**1.46**	**1.45**
11	I feel restless as I have to be on the move	0.33	-0.73	3	31	41	25	4	1.13	1.19
12	I look forward with enjoyment to things	1.44	1.67	2	62	29	7	2	1.03	0.88
13	I get sudden feelings of panic	0.92	0.14	2	45	37	13	5	0.86	0.87
14	I can enjoy a good book or radio or TV program	**2.02**	**3.99**	1	73	20	5	2	**1.35**	**1.96**

Table 5.2. Descriptive statistics of the HADS-T showing the proportion of responses within each category and Rasch item fit values. Fit values outside the range 0.7 to 1.3 and skew/kurtosis values outside the range -2 to +2 are in bold.

5.3b.i) HADS-T Conversion table from traditional scoring to ordinal scoring

If the HADS-T is used in its original form in a similar population it is possible to rescore the items according to table 5.3 in order to make the data ordinal. Each item's response categories have been given in ascending order, even for items that are reverse scored, meaning response category 0 was originally scored as zero, and response category 3 was originally scored as three.

Item	Response category 0	Response category 1	Response category 2	Response category 3
1	19.8	38.3	53.9	69.8
2	27	45.5	61.1	77
3	21.2	39.7	55.2	71.2
4	35.8	54.3	69.8	85.8
5	19.7	38.2	53.8	69.7
6	32.3	50.8	66.4	82.3
7	22.8	41.3	56.9	72.8
8	13.2	31.7	47.3	63.2
9	25.7	44.2	59.8	75.8
10	28.7	47.2	62.8	78.7
11	18	36.5	52.1	68
12	30.6	49.1	64.7	80.6
13	23.2	41.7	57.3	73.2
14	36.1	54.6	70.2	86.1

Table 5.3. Values used to rescore HADS-T in its original form when used on a population of new ophthalmic outpatients.

5.3b.ii) Response Scale Assessment – Category reduction

Response Scale	Participant mean	Item mean	Participant Separation Index (PSI)	Participant Separation Reliability
0123	35.6	51.1	2.18	0.83
0122	40.2	50.0	2.30	0.84
0011	31.6	50.4	1.17	0.58
0111	49.3	49.9	1.79	0.76

Table 5.4. Table summarising the effect of category reduction for the HADS-T on the person separation index and person separation reliability

Table 5.4 shows combining response categories 2 and 3 gives the best Person Separation Index (PSI). This causes all the skew and kurtosis values to fall within the accepted range of +2 to -2 (Table 5.5), the Item map to improve so that participants' scores more closely match the items (Figure 5.2) and the item and participant means are closer and within 10 units from near 15 with 4 response categories (table 5.4). When response categories are combined in this way the values used to rescore the questionnaire in order to create ordinal data are shown in Table 5.6.

Items		Skew	Kurtosis	Missing data (%)	Response category 0 (Floor, %)	Response category 1 (%)	Response category 2 (Ceiling, %)	Infit	Outfit
1	I feel tense or 'wound up'	0.15	-0.88	2	30	52	19	0.70	0.73
2	I still enjoy things I used to enjoy	0.61	-0.65	1	48	42	10	1.01	1.06
3	I get a sort of frightened feeling as if something awful is about to happen	0.36	-1.46	2	45	28	27	1.25	1.15
4	I can laugh and see the funny side of things	1.56	1.41	2	72	22	6	1.25	**1.48**
5	Worrying thoughts go through my mind	0.22	-1.23	2	36	41	23	0.86	0.87
6	I feel cheerful	1.18	0.26	2	65	28	8	0.71	**0.56**
7	I can sit at ease and feel relaxed	0.30	-0.90	3	37	48	16	0.77	0.80
8	I feel as if I am slowed down	-0.28	-1.09	2	19	44	37	0.97	0.98
9	I get sort of frightened feeling like 'butterflies' in the stomach	0.52	-0.74	2	45	44	11	0.96	0.98
10	I have lost interest in my appearance	0.98	-0.51	2	61	24	15	**1.42**	**1.51**
11	I feel restless as I have to be on the move	0.03	-1.32	3	31	41	29	1.21	1.29
12	I look forward with enjoyment to things	1.06	-0.07	2	62	29	9	0.98	0.86
13	I get sudden feelings of panic	0.47	-1.08	2	45	37	18	0.82	0.84
14	I can enjoy a good book or radio or TV program	1.62	1.55	1	73	20	6	1.22	**1.98**

Table 5.5. Descriptive statistics of the HADS-T showing reliability indicators and proportion of responses within each category after combining response categories 2 and 3. Infit and outfit values outside the range 0.7-1.3 are in **bold**.

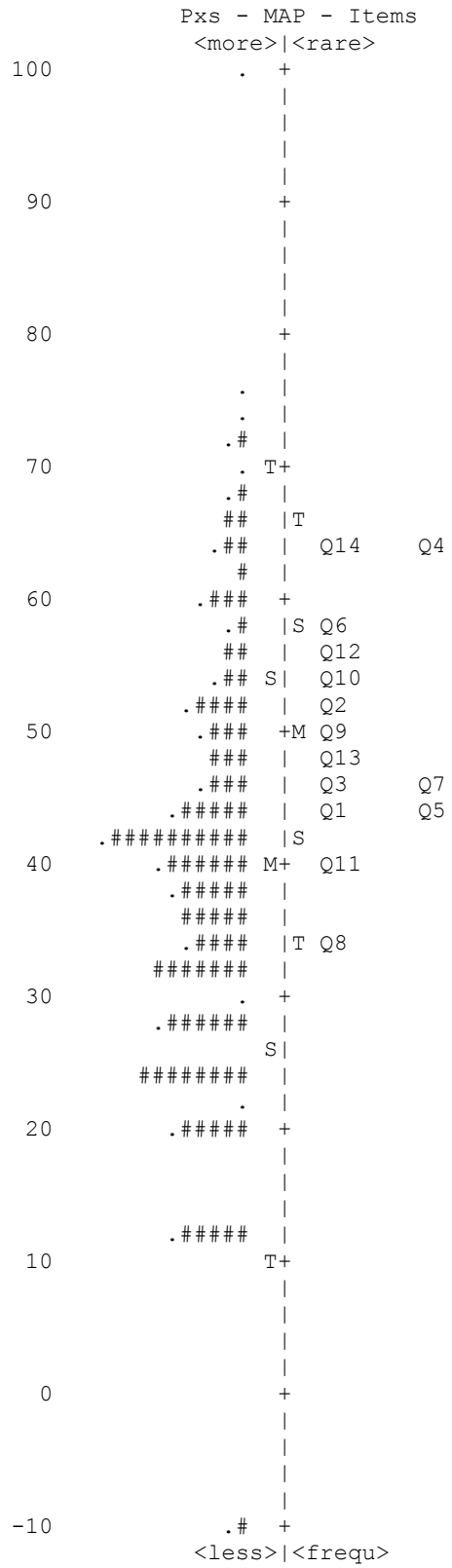

Figure 5.2. Item-Participant map for HADS-T after combining response categories 2 and 3. EACH '#' IS 3.

Item	Reponse category 0	Response category 1	Response category 2 & 3
1	21.8	43.2	64.5
2	30.6	51.9	73.3
3	23.9	45.3	66.7
4	41.9	63.3	84.7
5	22.2	43.6	65
6	37.5	58.8	80.2
7	24.5	45.9	67.2
8	12.9	34.3	55.7
9	28.9	50.2	71.6
10	33.5	54.8	76.2
11	18.5	39.9	61.3
12	35.4	56.8	78.1
13	26.5	47.8	69.2
14	42.2	63.6	85

Table 5.6. Table to convert traditional scoring to ordinal scoring for HADS-T after combining response categories 2 and 3.

5.3b.iii) Principal Components Analysis (PCA) of HADS-T.

HADS-T has two subscales; HADS-A (anxiety, even question numbers or lower case letters on figure 5.3) and HADS-D (depression, odd question numbers or capital letters on figure 5.3). PCA identifies these two factors as being separate, with the items forming separate strata (Table 3.3.7 and Figure 3.3.3). The raw variance explained by the measures was 49.3%, which is well below the 60% suggested as indicating unidimensionality (Smith, 2002; Linacre, 2009). In addition, the eigenvalue of the first contrast was 2.2 and greater than 2.0 suggesting that another significant dimension exists within the data (Smith, 2002; Linacre, 2009). The standardised residual data plot in Figure 5.3 also shows a clear differentiation into two groups of data and these two groups almost perfectly match the split of the items of the HADS-T into the A and D subscales.

Item	Map code	Loading	Measure
Q3	A	0.55	45.3
Q9	B	0.53	50.2
Q13	C	0.52	47.8
Q1	D	0.52	43.2
Q5	E	0.33	43.6
Q11	F	0.06	39.9
Q7	G	-0.15	45.9
Q8	g	-0.18	34.3
Q14	f	-0.23	63.6
Q6	e	-0.26	58.8
Q10	d	-0.29	54.8
Q4	c	-0.44	63.3
Q12	b	-0.50	56.8
Q2	a	-0.51	51.9

Table 5.7. Standardised residual loadings for HADS-T items after category reduction.

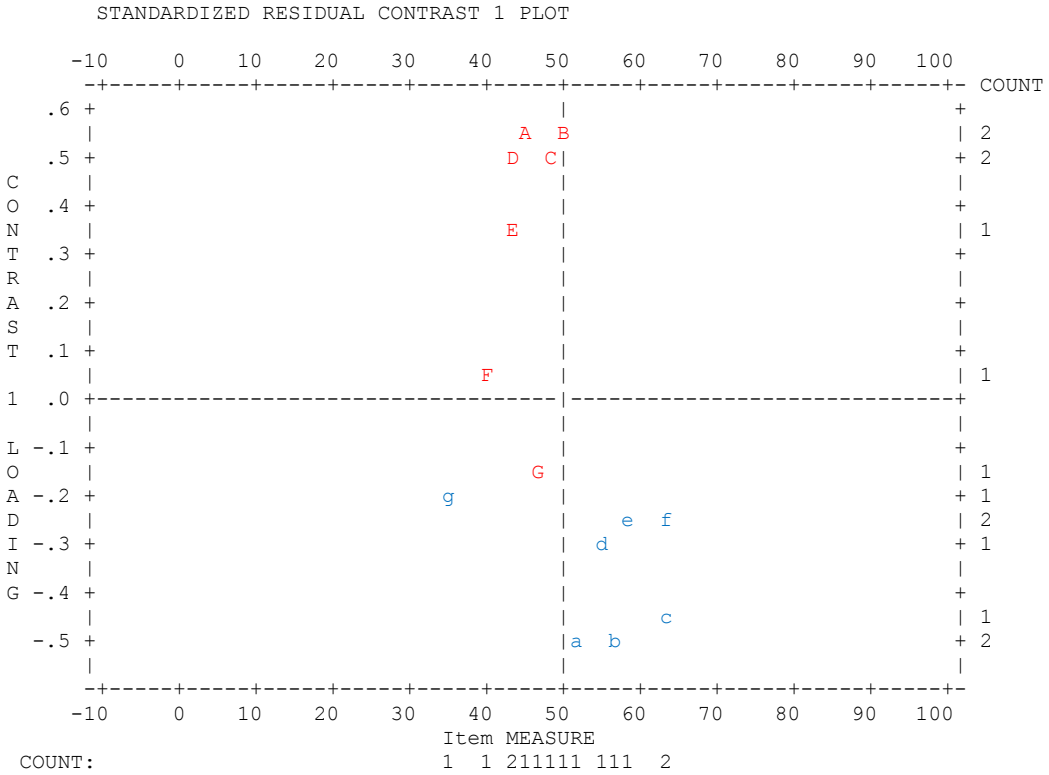

Figure 5.3. The Rasch standardised residual data plot for HADS-T data with the individual items from table 5.7. Anxiety items are in red, depression items are in blue.

5.3b.iv) HADS-T item removal

As discussed in section 2.2 items can be removed from a questionnaire if they have been shown to not provide useful information. This may be because they provide data that is too variable or does not fit the model (it may be measuring something else, infit and/or outfit values are too high) or data that is too predictable so that its removal does not affect the usefulness of the overall questionnaire (infit and/or outfit values are too low). An advantage of removing items is that they can reduce respondent burden and improve the discriminative ability of the instrument. The limits on infit and outfit vary with sample size (Linacre, 2003), and despite these data originating from a rating scale, as this sample size is reasonably large it is fair to use strict limits of 0.7 to 1.3 (Bond & Fox, 2004). Of course fit statistics must be used with an understanding of the items and data, and a misfitting item will not be removed if removal decreases person separation or if no other item measures a similar severity of symptom. Guidelines for item removal were followed (Pesudovs et al., 2007) and items were removed sequentially.

As can be seen by comparing Table 5.2 with Table 5.5, combining response categories 2 and 3 improved the fit statistics so that only item 10 (I have lost interest in my appearance) had infit outside the range of 0.7 to 1.3 and no items had skewness and kurtosis values outside the range -2 to +2. Item 10 also had a large floor effect (61%).

Item 10 was removed and the PSI decreased very slightly to 2.28 and Person Separation Reliability (PSR) remained at 0.84. Item and participant means were

relatively unchanged at 49.9 and 40.4 respectively. Fit statistics for the remaining items are given in Table 5.8.

Items		Infit	Outfit
Q1	I feel tense or 'wound up'	0.70	0.72
Q2	I still enjoy things I used to enjoy	1.04	1.08
Q3	I get a sort of frightened feeling as if something awful is about to happen	1.27	1.17
Q4	I can laugh and see the funny side of things	**1.32**	**1.54**
Q5	Worrying thoughts go through my mind	0.87	0.88
Q6	I feel cheerful	0.76	**0.59**
Q7	I can sit at ease and feel relaxed	0.79	0.81
Q8	I feel as if I am slowed down	1.02	1.02
Q9	I get sort of frightened feeling like 'butterflies' in the stomach	0.97	1.01
Q11	I feel restless as I have to be on the move	1.26	**1.31**
Q12	I look forward with enjoyment to things	1.03	0.92
Q13	I get sudden feelings of panic	0.83	0.84
Q14	I can enjoy a good book or radio or TV program	1.25	**2.29**

Table 5.8. Infit and outfit of HADS-T items after category reduction and removal of item 10. Values outside the range 0.7-1.3 are in **bold**.

After removal of item 10, the infit mean square of item 4 increases and becomes greater than 1.3. Item 4 also has the second largest floor effect (72%) therefore this item was removed and the PSI and PSR remain at 2.28 and 0.84 respectively whilst the participant and item means are 41.4 and 49.8 respectively. Fit statistics for the remaining items are given in Table 5.9.

Items		Infit	Outfit
Q1	I feel tense or 'wound up'	**0.69**	0.72
Q2	I still enjoy things I used to enjoy	1.10	1.13
Q3	I get a sort of frightened feeling as if something awful is about to happen	1.28	1.16
Q5	Worrying thoughts go through my mind	0.87	0.87
Q6	I feel cheerful	0.80	**0.63**
Q7	I can sit at ease and feel relaxed	0.83	0.85
Q8	I feel as if I am slowed down	1.03	1.02
Q9	I get sort of frightened feeling like 'butterflies' in the stomach	0.97	1.01
Q11	I feel restless as I have to be on the move	1.28	1.29
Q12	I look forward with enjoyment to things	1.09	0.97
Q13	I get sudden feelings of panic	0.82	0.84
Q14	I can enjoy a good book or radio or TV program	1.30	**2.39**

Table 5.9. Infit and outfit of HADS-T items after category reduction and removal of items 10 and 4. Values outside the range 0.7-1.3 are in **bold**.

After removal of items 4 and 10, item 1 has an infit value slightly less than 0.7, item 6 has an outfit below 0.7 and the outfit of item 14 remains above 1.3. As items 4 and 14 are measuring very similar levels of symptom (as shown by Rasch map Figure 5.2) and item 14 also has the highest floor effect and outfit value, the analysis was rerun excluding item 14, but including item 4.

Items		Infit	Outfit
Q1	I feel tense or 'wound up'	0.70	0.73
Q2	I still enjoy things I used to enjoy	1.08	1.11
Q3	I get a sort of frightened feeling as if something awful is about to happen	1.28	1.18
Q4	I can laugh and see the funny side of things	**1.37**	**1.60**
Q5	Worrying thoughts go through my mind	0.89	0.89
Q6	I feel cheerful	0.77	0.77
Q7	I can sit at ease and feel relaxed	0.81	0.81
Q8	I feel as if I am slowed down	1.03	1.03
Q9	I get sort of frightened feeling like 'butterflies' in the stomach	0.97	0.97
Q11	I feel restless as I have to be on the move	1.29	1.29
Q12	I look forward with enjoyment to things	1.07	1.07
Q13	I get sudden feelings of panic	0.84	0.86

Table 5.10. Infit and outfit of HADS-T items after category reduction and removal of items 10 and 14. Values outside the range 0.7-1.3 are in **bold**.

After removal of item 14 (and item 10) PSI and PSR are 2.26 and 0.84 whilst the participant and item means are 41.7 and 49.8 respectively. Fit values (Table 5.10) show that now item 4 is the only item with fit values outside the range 0.7-1.3. The analysis was therefore rerun without items 4, 10 and 14.

Items		Infit	Outfit
Q1	I feel tense or 'wound up'	0.70	0.72
Q2	I still enjoy things I used to enjoy	1.15	1.19
Q3	I get a sort of frightened feeling as if something awful is about to happen	1.29	1.18
Q5	Worrying thoughts go through my mind	0.89	0.89
Q6	I feel cheerful	0.82	**0.66**
Q7	I can sit at ease and feel relaxed	0.86	0.90
Q8	I feel as if I am slowed down	1.04	1.08
Q9	I get sort of frightened feeling like 'butterflies' in the stomach	0.97	1.01
Q11	I feel restless as I have to be on the move	**1.32**	**1.34**
Q12	I look forward with enjoyment to things	1.14	1.03
Q13	I get sudden feelings of panic	0.84	0.86

Table 5.11. Infit and outfit of HADS-T items after category reduction and removal of items 4, 10 and 14. Values outside the range 0.7-1.3 are in **bold**.

Item	Reponse category 0	Response category 1	Response category 2 & 3
Q1	22	45.6	69.1
Q2	31.9	55.5	79
Q3	24.5	48	71.6
Q5	22.6	46.1	69.7
Q6	39.7	63.2	86.8
Q7	25.1	48.6	72.2
Q8	11.9	35.5	59.1
Q9	30	53.6	77.1
Q11	18.4	42	65.5
Q12	37.3	60.9	84.5
Q13	27.3	50.9	74.5

Table 5.12. Table to convert traditional scoring to ordinal scoring for HADS-T after combining response categories 2 and 3 and removal of items 4, 10 and 14.

Removal of items 4, 10 and 14 gave a PSI and PSR of 2.24 and 0.83, and the participant and item means were 43.0 and 50.0. Item 11 now has infit and outfit values slightly higher than 1.3 and item 6 has an outfit value which is below 0.7 (Table 5.11). These items were kept as the Rasch map (Figure 5.4) shows they measure unique levels of symptom and their fit values were only slightly out of the strict range 0.7-1.3. Values to convert traditional scoring to ordinal scoring are slightly different, and given in table 5.12. A summary of the category reduction and item removal process for HADS-T is given in table 5.13.

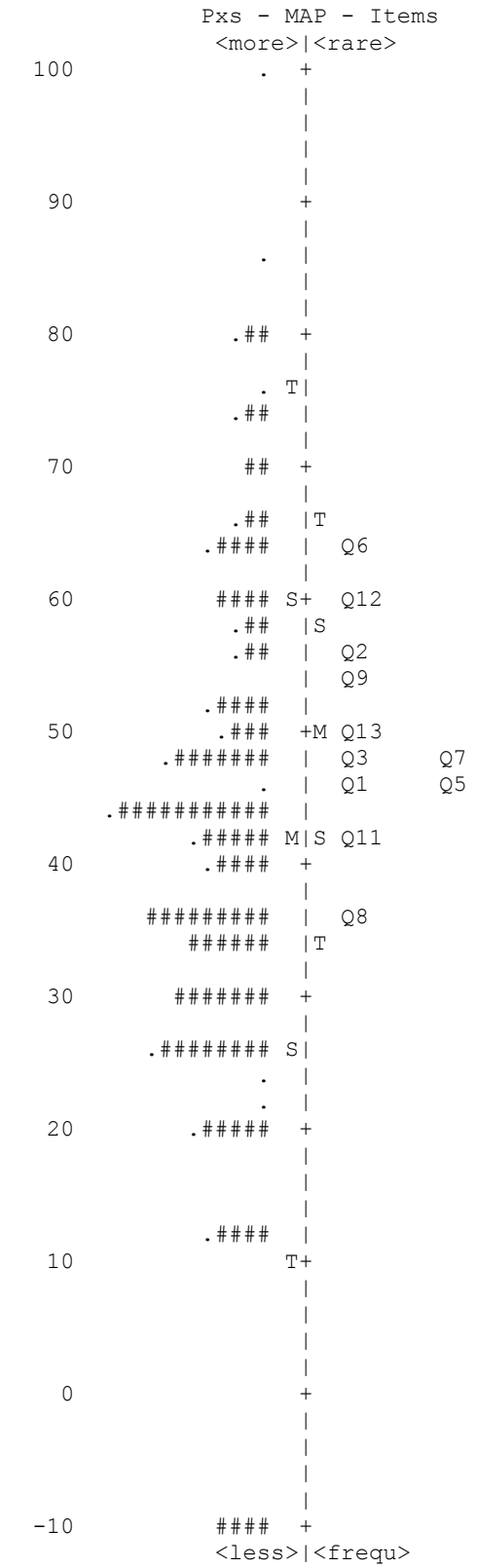

Figure 5.4. Item-Participant map for HADS-T after combination of response categories 2 and 3, and removal of items 4, 10 and 14. EACH '#' IS 3.

		All items		Q10 removed		Q4 and Q10 removed		Q10 and Q14 removed		Q4, Q10 and Q14 removed	
Items		Infit	Outfit	Infit	Outfit	Infit	Outfit	Infit	Outfit	Infit	Outfit
Q1	I feel tense or 'wound up'	0.70	0.72	0.70	0.72	**0.69**	0.72	0.70	0.73	0.70	0.72
Q2	I still enjoy things I used to enjoy	1.04	1.08	1.04	1.08	1.10	1.13	1.08	1.11	1.15	1.19
Q3	I get a sort of frightened feeling as if something awful is about to happen	1.27	1.17	1.27	1.17	1.28	1.16	1.28	1.18	1.29	1.18
Q4	I can laugh and see the funny side of things	**1.32**	**1.54**	**1.32**	**1.54**			**1.37**	**1.60**		
Q5	Worrying thoughts go through my mind	0.87	0.88	0.87	0.88	0.87	0.87	0.89	0.89	0.89	0.89
Q6	I feel cheerful	0.76	**0.59**	0.76	**0.59**	0.80	**0.63**	0.77	0.77	0.82	**0.66**
Q7	I can sit at ease and feel relaxed	0.79	0.81	0.79	0.81	0.83	0.85	0.81	0.81	0.86	0.90
Q8	I feel as if I am slowed down	1.02	1.02	1.02	1.02	1.03	1.02	1.03	1.03	1.04	1.08
Q9	I get sort of frightened feeling like 'butterflies' in the stomach	0.97	1.01	0.97	1.01	0.97	1.01	0.97	0.97	0.97	1.01
Q10	I have lost interest in my appearance	**1.42**	**1.51**								
Q11	I feel restless as I have to be on the move	1.26	**1.31**	1.26	**1.31**	1.28	1.29	1.29	1.29	**1.32**	**1.34**
Q12	I look forward with enjoyment to things	1.03	0.92	1.03	0.92	1.09	0.97	1.07	1.07	1.14	1.03
Q13	I get sudden feelings of panic	0.83	0.84	0.83	0.84	0.82	0.84	0.84	0.86	0.84	0.86
Q14	I can enjoy a good book or radio or TV program	1.25	**2.29**	1.25	**2.29**	1.30	**2.39**				
	Participant Separation Index	**2.30**		**2.28**		**2.28**		**2.26**		**2.24**	
	Participant Separation Reliability	**0.84**		**0.84**		**0.84**		**0.84**		**0.83**	
	Participant mean	**40.2**		**40.4**		**41.4**		**41.7**		**43.0**	
	Item mean	**50.0**		**49.9**		**49.8**		**49.8**		**50.0**	

Table 5.13. Summary of category reduction and item removal process for HADS-T.

5.3c. HADS-A results

Responses from 321 participants were analysed. The PSI was 1.94, PSR was 0.79. The item mean was 51.3 and participant mean was 37.6. Infit and outfit values are given in table 5.14 and the item map is given in figure 5.5.

Items		Skew	Kurtosis	Missing data (%)	Response category 0 (Floor, %)	Response category 1 (%)	Response category 2 (%)	Response category 3 (Ceiling, %)	Infit	Outfit
Q1	I feel tense or 'wound up'	0.75	0.35	2	30	52	13	6	0.74	0.73
Q3	I get a sort of frightened feeling as if something awful is about to happen	0.70	-0.66	2	45	28	21	6	1.14	1.08
Q5	Worrying thoughts go through my mind	0.73	-0.23	2	36	41	15	8	1.00	1.01
Q7	I can sit at ease and feel relaxed	0.42	-0.53	3	37	48	15	1	0.97	1.06
Q9	I get sort of frightened feeling like 'butterflies' in the stomach	0.87	0.54	2	45	44	9	2	0.88	0.95
Q11	I feel restless as I have to be on the move	0.33	-0.73	3	31	41	25	4	**1.36**	**1.38**
Q13	I get sudden feelings of panic	0.92	0.14	2	45	37	13	5	0.84	0.81

Table 3.3.14. Descriptive statistics of the HADS-A showing reliability indicators and proportion of responses within each category. Fit values outside the range 0.7-1.3 are in **bold**.

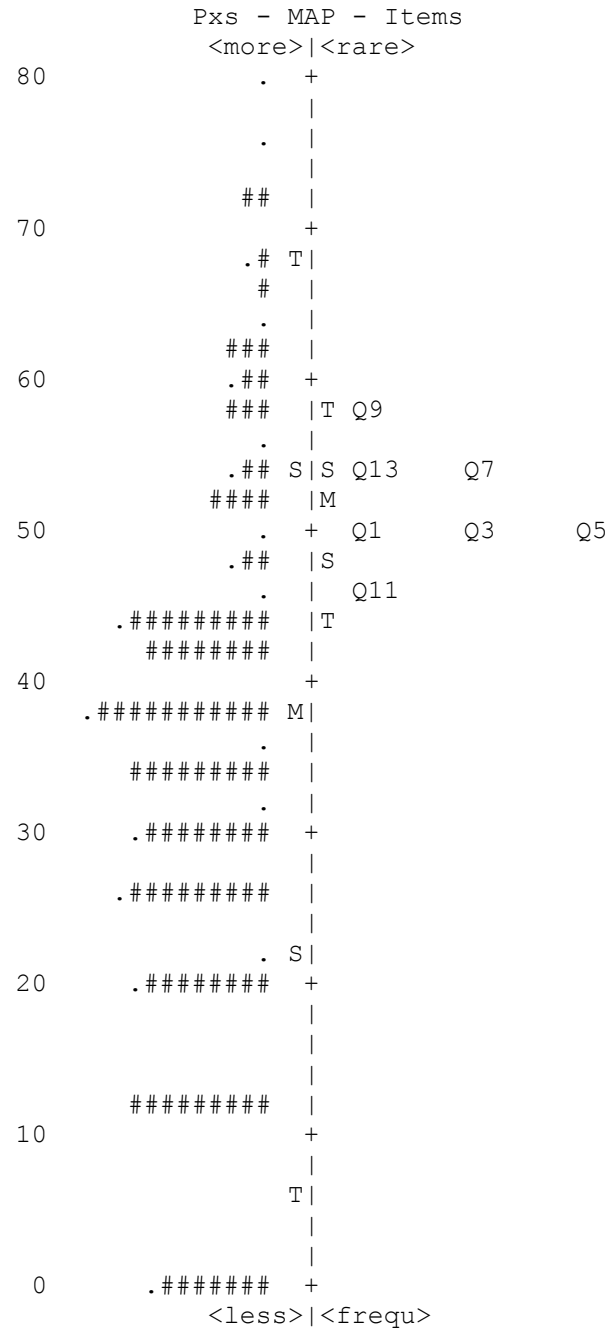

Figure 5.5. Item map for HADS-A. EACH '#' IS 3.

5.3c.i) HADS-A conversion table from traditional scoring to ordinal scoring

If the HADS-A is used in its original form in a similar population it is possible to rescore the items according to Table 5.15 in order to make the data ordinal. Each items response categories have been given in ascending order, even for items that are reverse scored, meaning response category 0 was originally scored as zero (least presence of anxiety), and response category 3 was originally scored as three (most presence of anxiety).

Items		Response Category 0	Response Category 1	Response Category 2	Response Category 3
Q1	I feel tense or 'wound up'	16.2	38.9	60.5	79.6
Q3	I get a sort of frightened feeling as if something awful is about to happen	18	40.7	62.3	81.5
Q5	Worrying thoughts go through my mind	16.2	38.9	60.5	79.6
Q7	I can sit at ease and feel relaxed	20.2	42.9	64.5	83.7
Q9	I get sort of frightened feeling like 'butterflies' in the stomach	24.2	46.9	68.5	87.7
Q11	I feel restless as I have to be on the move	13.7	36.5	58	77.2
Q13	I get sudden feelings of panic	20.8	43.5	65.1	84.2

Table 5.15. Table to convert traditional scoring to ordinal scoring for HADS-A.

5.3c.ii) HADS-A Principal Components Analysis

The raw variance explained by the measures was 53.2%, which is below the 60% suggested as indicating unidimensionality. However, the eigenvalue of the first contrast was 1.6 which is less than 2.0 suggesting that any other dimensions within

the data were not significant. The HADS-A data therefore appear to be unidimensional.

5.3c.iii) HADS-A combining response categories

Combining response categories 2 and 3 resulted in a decrease in the PSI to 1.78, and the PSR to 0.76, although the Item mean and participant means were closer at 50.0 and 44.5 respectively. Due to the decrease in PSI response categories were not combined.

5.3c.iv) HADS-A item removal

Item 11 (I feel restless as I have to be on the move, Table 5.14) was the only item with fit values outside the range 0.7-1.3 indicating it had too much variability. This item was therefore removed and the analysis rerun resulting in a PSI of 1.93, PSR of 0.79, item mean of 52.1 and participant mean of 37.6. The fit statistics of all items were now within the range 0.7-1.3 (Table 5.16). As this item measures the lowest level of symptom (Figure 5.5), and removal did not result in an increase in PSR, which was already below the cut-off value of 2.0, it was not removed.

Items		Infit	Outfit
Q1	I feel tense or 'wound up'	0.77	0.75
Q3	I get a sort of frightened feeling as if something awful is about to happen	1.15	1.10
Q5	Worrying thoughts go through my mind	1.02	1.04
Q7	I can sit at ease and feel relaxed	1.11	1.21
Q9	I get sort of frightened feeling like 'butterflies' in the stomach	0.95	0.99
Q13	I get sudden feelings of panic	0.91	0.89

Table 5.16. Infit and outfit values of HADS-A after removal of Q11.

5.3d. HADS-D results

Responses from 321 patients were received. The PSI was 1.31, PSR was 0.63. The item mean 52.3 and participant mean was 32.6. Infit and outfit values are given in table 5.17 and the item map is shown in figure 5.7.

Items		Skew	Kurtosis	Missing data (%)	Response category 0 (Floor, %)	Response category 1 (%)	Response category 2 (%)	Response category 3 (Ceiling, %)	Infit	Outfit
Q2	I still enjoy things I used to enjoy	0.98	0.75	1	48	42	8	2	0.87	0.82
Q4	I can laugh and see the funny side of things	**2.05**	**4.34**	2	72	22	4	2	1.23	1.17
Q6	I feel cheerful	1.57	**2.22**	2	65	28	6	2	0.76	**0.67**
Q8	I feel as if I am slowed down	0.30	-0.61	2	19	44	27	10	1.04	1.06
Q10	I have lost interest in my appearance	1.32	0.83	2	61	24	11	4	**1.33**	1.21
Q12	I look forward with enjoyment to things	1.44	1.67	2	62	29	7	2	0.80	**0.69**
Q14	I can enjoy a good book or radio or TV program	**2.02**	**3.99**	1	73	20	5	2	**1.34**	**1.38**

Table 5.17. Descriptive statistics of the HADS-D showing reliability indicators and proportion of responses within each category. Fit values outside the range 0.7-1.3 are in **bold**. Skew and kurtosis values outside the range -2 to 2 are in **bold**.

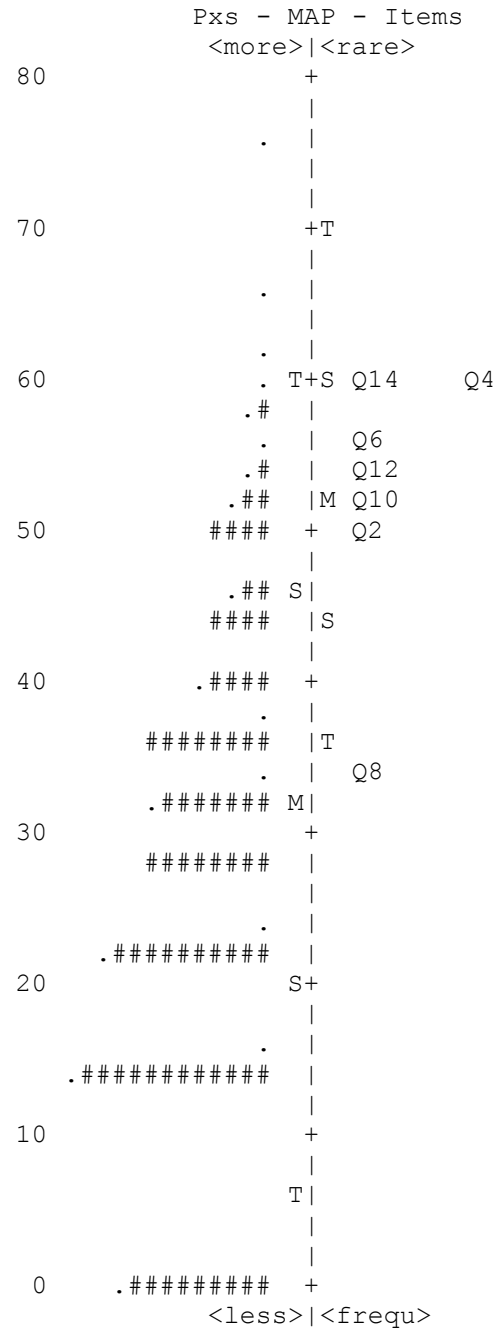

Figure 5.7. Item-Participant map for HADS-D. EACH '#' IS 4.

5.3d.i) HADS-D Principal Components Analysis

The raw variance explained by the measures was 51.0%, which is below the 60% suggested as indicating unidimensionality. However, the eigenvalue of the first contrast was 1.5 which is less than 2.0 suggesting that any other dimensions within the data were not significant. The data within HADS-D therefore appear to be unidimensional.

5.3d.ii) HADS-D combining response categories

Due to the large floor effect, in order to improve skew, kurtosis and fit values, response categories 2 and 3 were combined, meaning that the four response categories were scored 0, 1, 2, 2 in ascending order. This resulted in a slight increase in the PSI to 1.40, and the PSR to 0.66. The item and participant means were closer at 51.96 and 38.19 respectively. New fit values are given in table 5.18 and the item map is shown in figure 5.8.

Items		Infit	Outfit	Skew	Kurtosis
Q2	I still enjoy things I used to enjoy	0.91	0.87	0.61	-0.65
Q4	I can laugh and see the funny side of things	1.08	1.14	1.56	1.41
Q6	I feel cheerful	0.76	**0.68**	1.18	0.26
Q8	I feel as if I am slowed down	0.98	1.09	-0.28	-1.09
Q10	I have lost interest in my appearance	**1.31**	1.25	0.98	-0.51
Q12	I look forward with enjoyment to things	0.81	**0.69**	1.06	-0.07
Q14	I can enjoy a good book or radio or TV program	1.21	**1.41**	1.62	1.55

Table 5.18. Infit, outfit, skew and kurtosis of HADS-D items after combination of response categories 2 and 3. Fit values outside the range 0.7-1.3 are in **bold**.

As the PSI was still not over 2, further collapsing of response categories was attempted. Combining response categories 1, 2 and 3, meaning that the four response categories were scored 0, 1, 1, 1 in ascending order, resulted in a drop of the PSI to 0.77 and PSR to 0.37. The participant mean was 49.1 and the item mean was 55.0.

Combining response categories 0 and 1, and response categories 2 and 3, meaning that the four response categories were scored 0, 0, 1, 1 in ascending order, resulted in a drop of the PSI and PSR both to 0. The participant mean was 35.8 and the item mean was 53.9.

5.3d.iii) HADS-D Item removal

After combining response categories 2 and 3, outfit of item Q14 was still 1.41, and item map (Figure 5.7) shows that Q14 measures a similar level of symptom to Q4, therefore this item was removed and the analysis rerun. This resulted in a PSI of 1.35, a PSR of 0.65, item mean of 52.0 and participant mean of 39.7. All items now had fit values within the range 0.7-1.3 apart from Q10, therefore this item was removed and the analysis rerun. This resulted in a PSI of 1.32, PSR of 0.64, item mean of 52.7 and participant mean of 41.09. As removing these items did not result in a substantial increase in PSI, which was well below the cut-off value of 2.0, they were retained.

5.3e. STAI-S results

Responses from 318 participants were analysed. PSI was 2.75, PSR was 0.88, Item mean was 49.9, and Participant mean was 38.3. Reliability indices are given in table 5.19 and the item map is shown in figure 5.11.

Items		Skew	Kurtosis	Missing data (%)	Response category 1 (Floor, %)	Response category 2 (%)	Response category 3 (%)	Response category 4 (Ceiling, %)	Infit	Outfit
Q1	I feel calm	0.83	0.21	1	35	46	12	7	0.74	0.91
Q2	I feel secure	1.13	0.52	1	52	32	11	5	0.74	0.68
Q3	I am tense	0.7	-0.8	2	47	24	20	9	**1.34**	**1.44**
Q4	I feel strained	1.04	-0.2	2	58	18	17	7	**1.38**	1.29
Q5	I feel at ease	0.6	-0.6	2	35	38	19	8	0.75	0.74
Q6	I feel upset	1.86	**2.33**	3	75	11	9	4	**1.38**	1.27
Q7	I am presently worrying over possible misfortunes	1.04	-0.2	1	55	22	14	9	1.12	1.07
Q8	I feel satisfied	0.56	-0.5	2	33	41	20	7	0.90	1.05
Q9	I feel frightened	**2.02**	**3.16**	2	77	12	8	3	**1.31**	1.10
Q10	I feel comfortable	0.74	-0.3	1	38	39	16	7	0.88	0.97
Q11	I feel self-confident	0.6	-0.5	3	33	41	18	8	0.93	0.97
Q12	I feel nervous	0.95	-0.1	2	51	29	15	5	1.13	1.16
Q13	I am jittery	1.68	1.7	3	72	14	10	4	**1.34**	1.28
Q14	I feel indecisive	1.42	1.04	4	63	22	10	5	**1.31**	**1.41**
Q15	I am relaxed	0.56	-0.5	1	30	42	19	9	**0.61**	**0.63**
Q16	I feel content	0.47	-0.7	4	32	38	23	8	0.78	0.86
Q17	I am worried	1.09	0.17	1	52	29	11	8	1.15	1.28
Q18	I feel confused	1.92	**2.62**	1	76	11	9	3	1.19	0.98
Q19	I feel steady	0.72	-0.5	3	40	35	17	8	1.03	1.11
Q20	I feel pleasant	0.73	-0.4	1	42	35	18	5	0.89	0.92

Table 5.19. Descriptive statistics of the STAI-S showing the proportion of responses within each category. Fit values outside the range 0.7-1.3 are in **bold**. Skew and kurtosis values outside the range -2 to 2 are in **bold**.

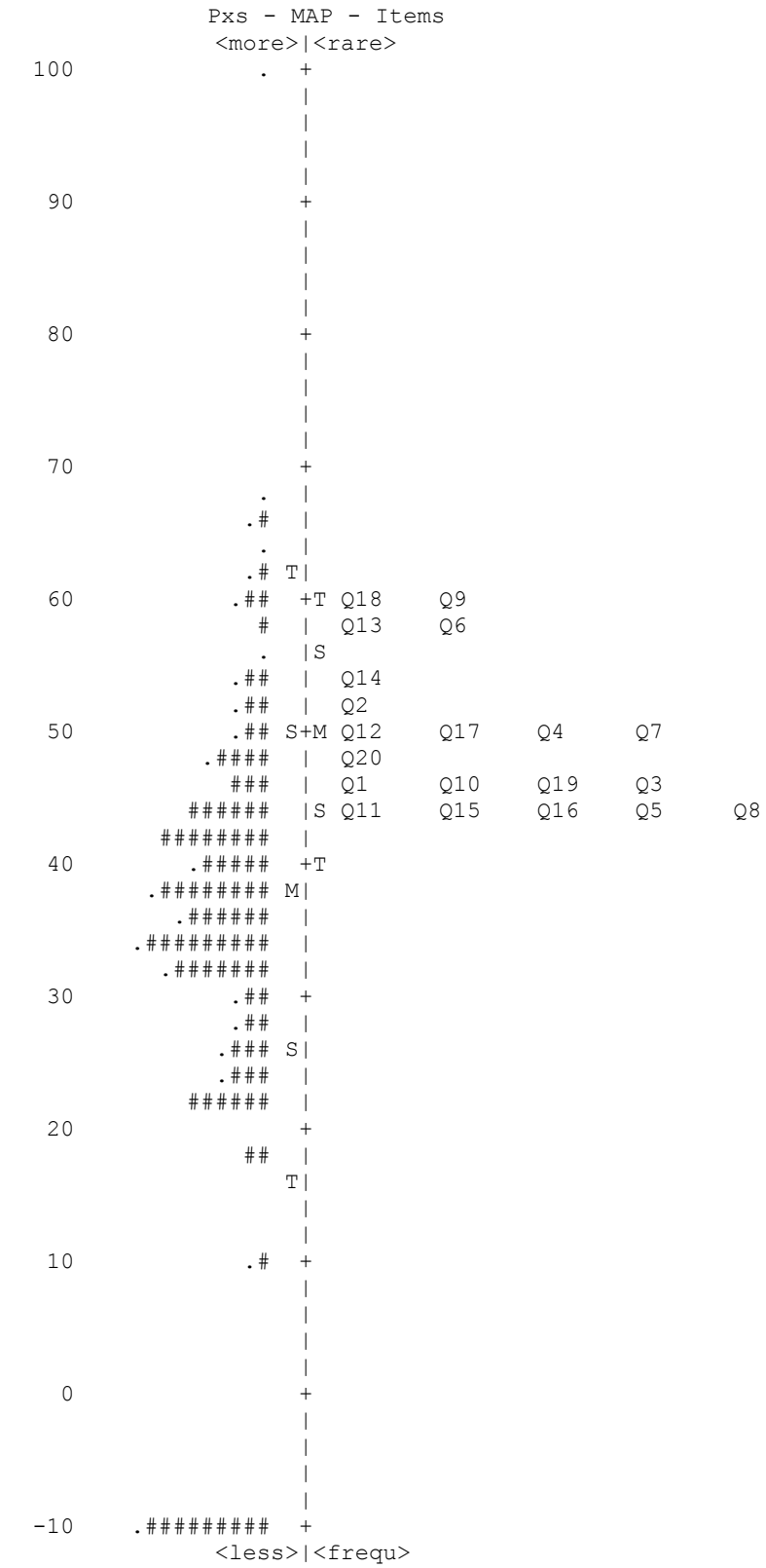

Figure 5.11. Item map for STAI-S. Each '#' IS 3.

5.3d.i) STAI-S combining response categories

Due to the large floor effect and less than 10% of participants endorsing response category 4 for any item, response categories 3 and 4 were combined, meaning that the four response categories were scored 1, 2, 3, 3 in ascending order. This resulted in a slight increase in PSI to 2.77, a PSR of 0.88, and participant and item means that were closer together at 43.67 and 49.70 respectively. Fit values, skew and kurtosis are given in Table 5.20 and the item map is shown in figure 5.13.

Items			Infit	Outfit	Skew	Kurtosis
Q1	J	I feel calm	0.78	0.92	0.26	-1.03
Q2	G	I feel secure	0.72	**0.66**	0.69	-0.87
Q3	d	I am tense	1.28	**1.35**	0.35	-1.54
Q4	g	I feel strained	1.27	1.15	0.73	-1.18
Q5	F	I feel at ease	0.82	0.78	0.13	-1.37
Q6	c	I feel upset	1.26	**1.32**	1.55	0.76
Q7	j	I am presently worrying over possible misfortunes	1.12	1.08	0.67	-1.18
Q8	E	I feel satisfied	0.95	1.11	0.10	-1.29
Q9	f	I feel frightened	1.18	1.07	1.69	1.34
Q10	C	I feel comfortable	0.85	0.94	0.26	-1.25
Q11	I	I feel self-confident	0.91	1.02	0.10	-1.30
Q12	a	I feel nervous	1.11	1.11	0.59	-1.16
Q13	b	I am jittery	1.22	1.26	1.37	0.25
Q14	i	I feel indecisive	**1.35**	**1.53**	1.05	-0.39
Q15	B	I am relaxed	**0.64**	**0.67**	0.04	-1.29
Q16	A	I feel content	0.82	0.86	0.03	-1.39
Q17	h	I am worried	1.10	1.28	0.64	-1.06
Q18	e	I feel confused	1.11	0.99	1.62	1.01
Q19	H	I feel steady	0.93	1.04	0.28	-1.37
Q20	D	I feel pleasant	0.91	0.93	0.36	-1.29

Table 5.20. Infit, outfit, skew and kurtosis of STAI-S items after combination of response categories 3 and 4. Fit values outside the range 0.7-1.3 are in **bold**. Anxiety present items are in red and anxiety absent items are in blue.

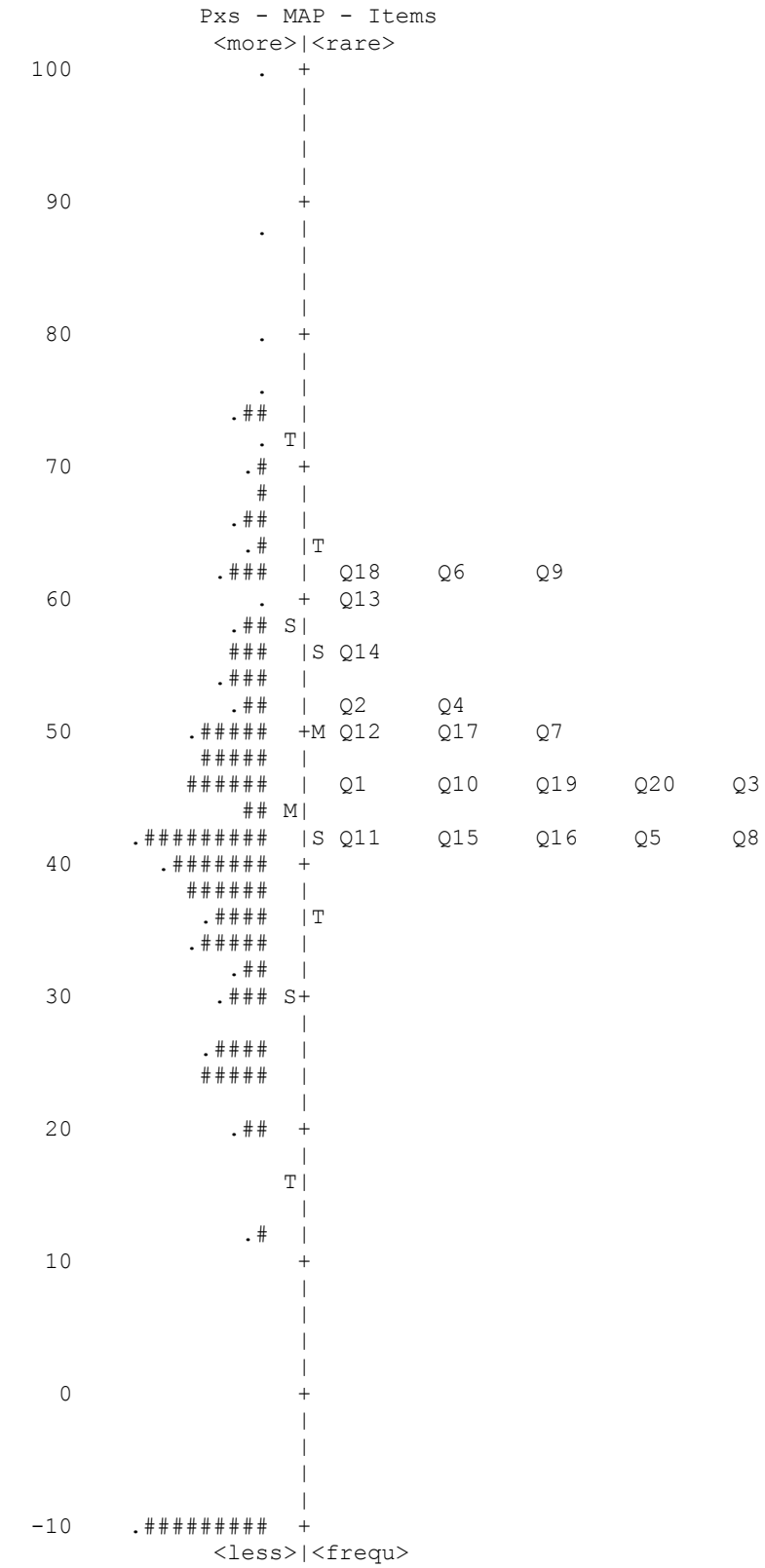

Figure 5.13. Item map for STAI-S after combining response categories 3 and 4. Each '#' IS 3.

5.3d.ii) STAI-S Principal components analysis

The raw variance explained by the measures after combination of response categories 3 and 4 was 47.8%, which is well below the 60% suggested as indicating unidimensionality and the eigenvalue of the first contrast was 3.0 and greater than 2.0 suggesting that another significant dimension within the data. The 2nd contrast had an eigenvalue of 1.7 indicating just two factors within the data. The standardised residual data plot in Figure 5.14 shows a clear differentiation into two groups of data and these two groups match the split of the items into state anxiety-present and state anxiety-absent factors. Despite both contributing towards the same construct, these two factors are clearly separate, therefore the items were split into two subscales and re-analysed.

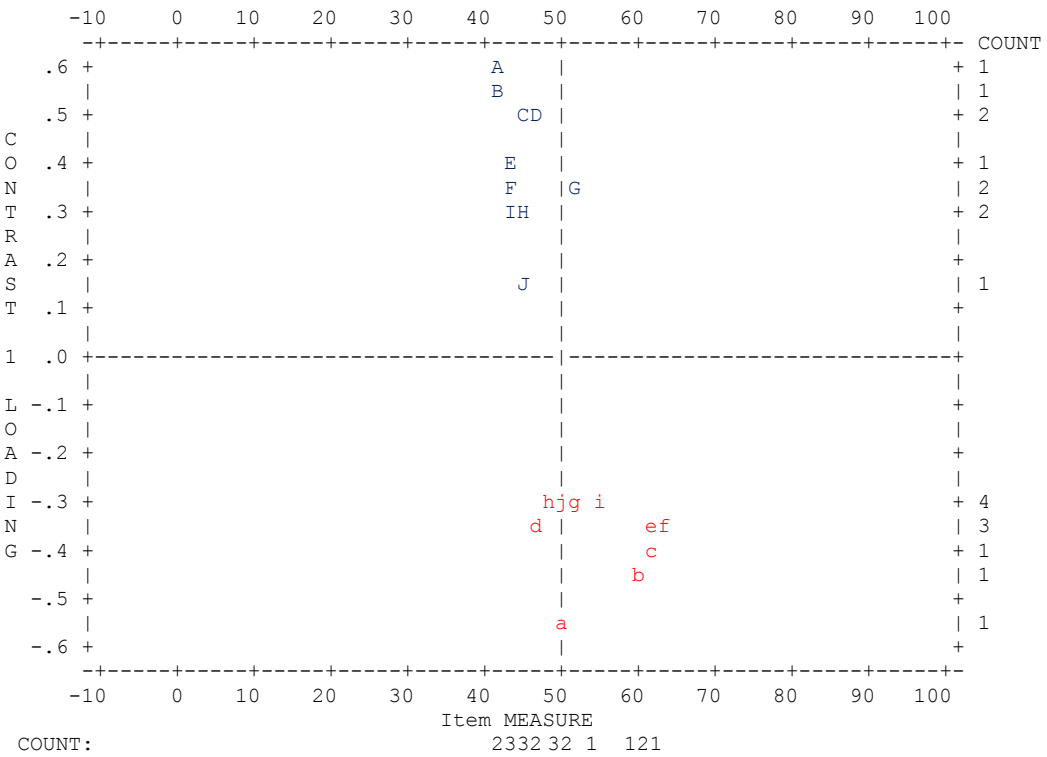

Figure 5.14. Standardised residual plot. State anxiety absent items are in blue and state anxiety present items are in red. Letter to item conversion key is given in table 5.20.

5.3e STAI-S Anxiety Absent Results

317 participants were included, which gave a PSI of 2.30, PSR of 0.84, participant mean of 47.52 and item mean of 49.87. Reliability indices are given in table 5.21 and the item map is shown in figure 5.15.

Items		Skew	Kurtosis	Missing data (%)	Response category 1 (Floor, %)	Response category 2 (%)	Response category 3&4 (%)	Infit	Outfit
Q1	I feel calm	0.26	-1.03	1	35	46	19	1.09	1.20
Q2	I feel secure	0.69	-0.87	1	52	32	16	0.90	0.81
Q5	I feel at ease	0.13	-1.37	2	35	38	27	1.00	0.95
Q8	I feel satisfied	0.10	-1.29	2	33	41	27	1.14	1.23
Q10	I feel comfortable	0.26	-1.25	1	38	39	23	0.88	0.93
Q11	I feel self-confident	0.10	-1.30	3	33	41	27	1.17	**1.33**
Q15	I am relaxed	0.04	-1.29	1	30	42	28	**0.69**	0.70
Q16	I feel content	0.03	-1.39	4	32	38	30	0.84	0.87
Q19	I feel steady	0.28	-1.37	3	40	35	25	1.20	1.17
Q20	I feel pleasant	0.36	-1.29	1	42	35	23	1.02	0.95

Table 5.21. Descriptive statistics of the STAI-S Anxiety Absent items showing reliability indicators and proportion of responses within each category. Fit values outside the range 0.7-1.3 are in **bold**.

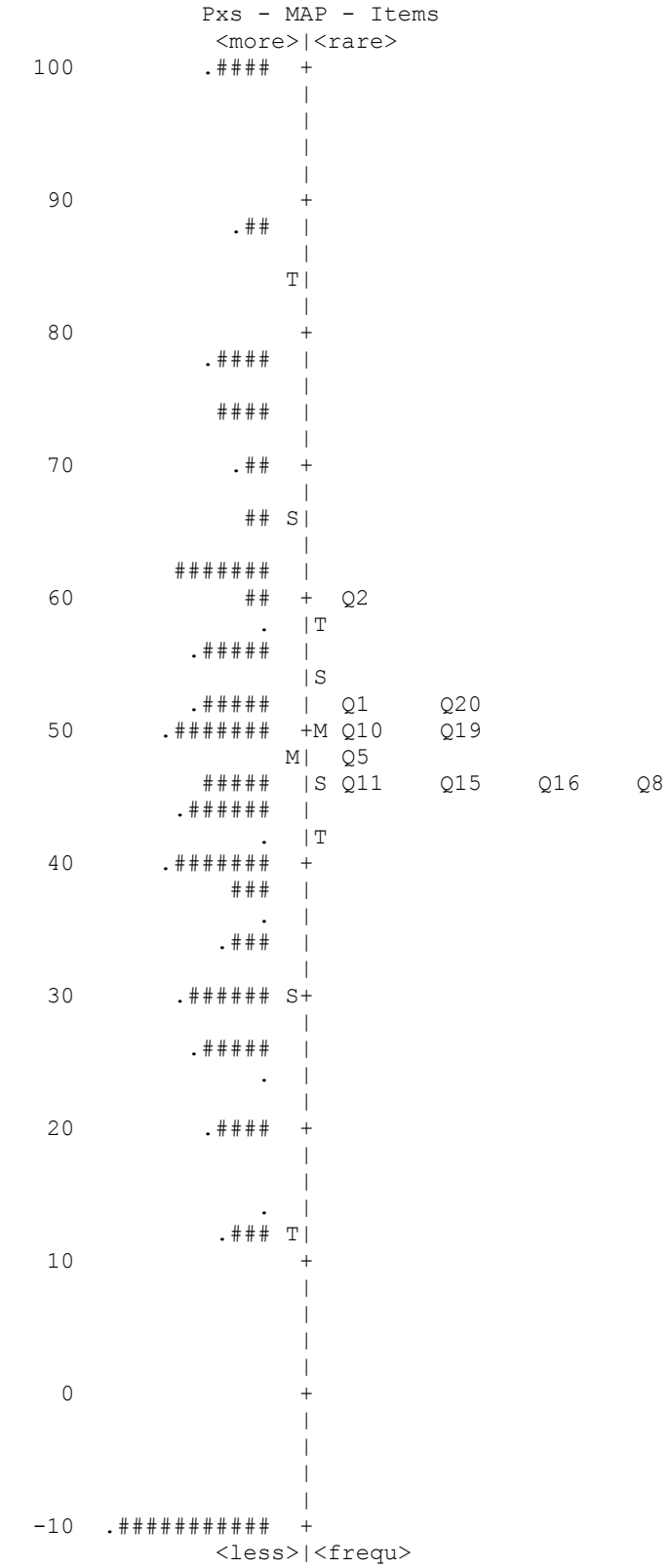

Figure 5.15. Item map for STAI-S Anxiety Absent items after combining response categories 3 and 4. Each '#' is 3 participants.

5.3e.i) Principal Components Analysis for STAI-S Anxiety Absent items

The raw variance explained by the measures was 50.4%, which is below the 60% suggested as indicating unidimensionality. However, the eigenvalue of the first contrast was 1.8 which is less than 2.0 suggesting that any other dimensions within the data were not significant. The data within STAI-S Anxiety Absent therefore appear to be unidimensional.

5.3f STAI-S Anxiety Present Results

Responses were analysed from 317 patients which gave a PSI of 1.69, PSR of 0.74, participant mean of 40.93 and item mean of 49.97. Reliability indices are given in table 5.22 and the item map is shown in figure 5.16.

Items		Skew	Kurtosis	Missing data (%)	Response category 1 (Floor, %)	Response category 2 (%)	Response category 3 & 4(%)	Infit	Outfit
Q3	I am tense	0.35	-1.54	2	47	24	29	1.07	1.21
Q4	I feel strained	0.73	-1.18	2	58	18	24	1.07	1.03
Q6	I feel upset	1.55	0.76	3	75	11	14	1.06	0.87
Q7	I am presently worrying over possible misfortunes	0.67	-1.18	1	55	22	22	1.00	1.00
Q9	I feel frightened	1.69	1.34	2	77	12	11	1.02	0.97
Q12	I feel nervous	0.59	-1.16	2	51	29	21	0.80	0.79
Q13	I am jittery	1.37	0.25	3	72	14	15	0.94	0.97
Q14	I feel indecisive	1.05	-0.39	4	63	22	15	1.21	**1.41**
Q17	I am worried	0.64	-1.06	1	52	29	19	0.99	1.06
Q18	I feel confused	1.62	1.01	1	76	11	13	0.94	0.89

Table 5.22. Descriptive statistics of the STAI-A Anxiety Present items showing reliability indicators and proportion of responses within each category. Fit values outside the range 0.7-1.3 are in **bold**.

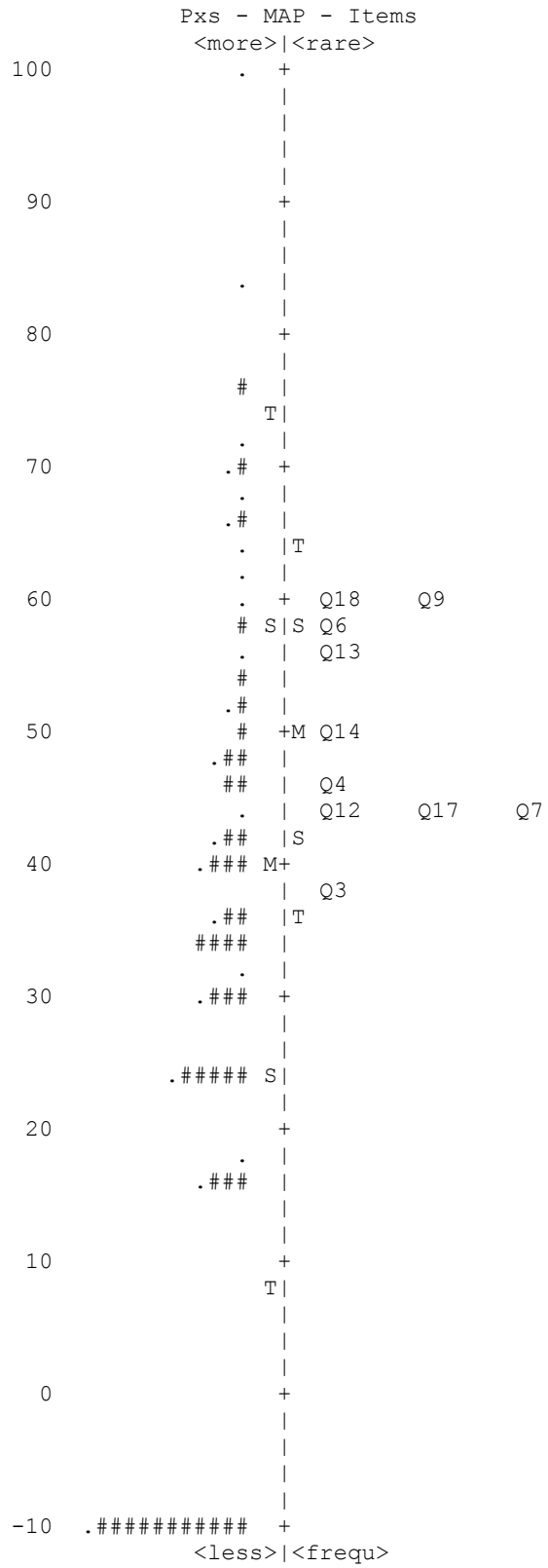

Figure 5.16. Item map for STAI-S Anxiety Present items after combining response categories 3 and 4. Each '#' is 6 participants.

5.3f.i) Principal Components Analysis for STAI-S Anxiety Present items

The raw variance explained by the measures was 47.1%, which is below the 60% suggested as indicating unidimensionality. However, the eigenvalue of the first contrast was 1.6 which is less than 2.0 suggesting that any other dimensions within the data were not significant. The data within STAI-S Anxiety Present therefore appear to be unidimensional.

5.3g. STAI-T results

Responses from 280 patients were analysed. PSI was 3.01, PSR was 0.90, the participant mean was 38.96 and the item mean was 50.04. Fit values, skew and kurtosis are given in table 5.23 and the item map is shown in figure 5.17.

Items		Skew	Kurtosis	Missing data (%)	Response category 1 (Floor, %)	Response category 2 (%)	Response category 3 (%)	Response category 4 (Ceiling, %)	Infit	Outfit
Q21	I feel pleasant	0.37	-0.96	1	37	36	24	2	1.02	1.09
Q22	I feel nervous and restless	0.85	0.86	0	34	53	9	4	0.76	0.82
Q23	I feel satisfied with myself	0.13	-1.09	4	30	30	34	6	0.95	1.02
Q24	I wish I could be as happy as others seem to be	0.71	-0.57	1	38	35	15	11	**1.91**	**2.18**
Q25	I feel like a failure	1.30	1.16	1	59	30	9	2	0.81	0.67
Q26	I feel rested	-0.05	-0.93	3	23	29	38	10	1.20	1.30
Q27	I am “calm, cool and collected”	0.04	-1.06	1	27	31	36	5	0.73	0.74
Q28	I feel that difficulties are piling up so that I cannot overcome them	1.15	0.81	1	50	36	9	5	1.08	1.11
Q29	I worry too much over something that really doesn’t matter	0.74	-0.11	0	37	42	15	6	0.91	0.88
Q30	I am happy	0.61	-0.55	1	43	34	20	3	0.80	0.79
Q31	I have disturbing thoughts	1.03	0.55	1	52	36	10	2	0.97	1.04
Q32	I lack self confidence	0.92	0.12	1	41	40	12	8	1.19	1.12
Q33	I feel secure	0.62	-0.83	2	46	26	23	5	0.85	0.77
Q34	I make decisions easily	0.16	-1.09	2	30	29	33	8	1.26	**1.51**
Q35	I feel inadequate	1.05	1.00	3	49	42	7	2	0.92	0.98
Q36	I am content	0.58	-0.84	3	44	28	24	4	0.79	0.71
Q37	Some unimportant thought runs through my mind and bothers me	0.76	0.22	0	39	46	13	3	0.75	0.81
Q38	I take disappointments so keenly that I can’t put them out of my mind	0.84	0.04	0	41	40	14	6	1.27	**1.67**
Q39	I am a steady person	0.46	-1.04	1	43	28	26	3	0.81	0.75
Q40	I get in a state of tension or turmoil as I think over my recent concerns and interests	0.94	0.56	1	40	45	10	5	0.79	0.92

Table 5.23. Descriptive statistics of the STAI-T showing reliability indicators and proportion of responses within each category. Fit values outside the range 0.7-1.3 are in **bold**.

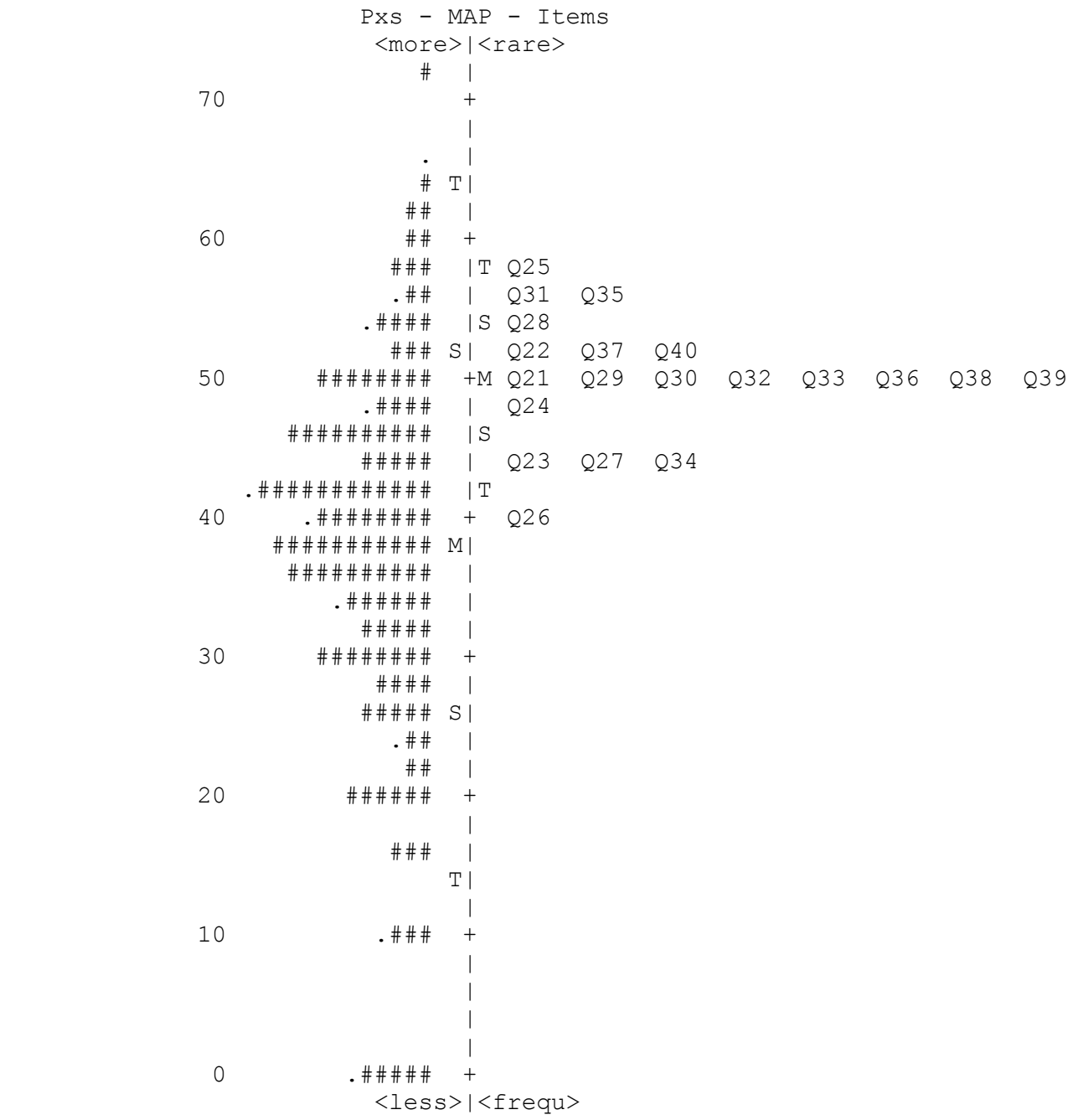

Figure 5.17. Item map for STAI-T. Each '#' IS 2 participants.

5.3g.i). STAI-T combining response categories

Although skew and kurtosis values are all in the range -2 to +2, suggesting a normal distribution of data, there is a large floor effect present for most items. Only one item has more than 10% of respondents endorsing response category 4, therefore combination of categories 3 and 4 was attempted to discover if there was an improvement in PSI or PSR. The resulting PSI was increased from 3.01 to 3.10, the PSR from 0.90 to 0.91 and the participant and item means were much closer at 45.93 and 50.00 respectively. New fit statistics are given in table 5.24.

Items			Infit	Outfit
Q21	E	I feel pleasant	1.14	1.16
Q22	f	I feel nervous and restless	0.80	0.90
Q23	B	I feel satisfied with myself	1.07	1.03
Q24	g	I wish I could be as happy as others seem to be	**1.59**	**1.92**
Q25	J	I feel like a failure	0.86	0.72
Q26	H	I feel rested	1.24	**1.48**
Q27	G	I am "calm, cool and collected"	0.88	0.80
Q28	d	I feel that difficulties are piling up so that I cannot overcome them	0.99	1.05
Q29	a	I worry too much over something that really doesn't matter	0.90	0.92
Q30	C	I am happy	0.80	0.73
Q31	h	I have disturbing thoughts	1.05	1.10
Q32	i	I lack self confidence	1.07	1.00
Q33	D	I feel secure	0.84	0.73
Q34	I	I make decisions easily	**1.34**	**1.65**
Q35	j	I feel inadequate	0.88	0.96
Q36	A	I am content	0.79	**0.68**
Q37	b	Some unimportant thought runs through my mind and bothers me	0.83	0.87
Q38	c	I take disappointments so keenly that I can't put them out of my mind	1.13	**1.42**
Q39	F	I am a steady person	0.87	0.75
Q40	e	I get in a state of tension or turmoil as I think over my recent concerns and interests	0.80	0.95

Table 5.24. Infit and outfit values of STAI-T after combining response categories 3 and 4. Fit values outside the range 0.7-1.3 are in **bold.** Trait anxiety present items are in red and trait anxiety absent items are in blue.

Combining the response categories resulted in an increase in PSI and a significant reduction in the difference between participant mean score and item mean score from about 11 to 4. Although the fit statistics of items 26 and 34 increased, those for items 24 and 38 decreased. The combination of response categories was therefore retained.

5.3g.i) STAI-T Principal Components Analysis

The raw variance explained by the measures after combination of response categories 3 and 4 was 50.9%, which is well below the 60% suggested as indicating unidimensionality and the eigenvalue of the first contrast was 3.2 and greater than 2.0 suggesting that another significant dimension within the data. The 2nd contrast had an eigenvalue of 1.7 indicating just two factors within the data. The standardised residual data plot (Fig 5.18) showed a clear differentiation into two groups of data and these two groups perfectly match the split of the items of the STAI into anxiety-present and anxiety-absent factors as shown previously for STAI-S.

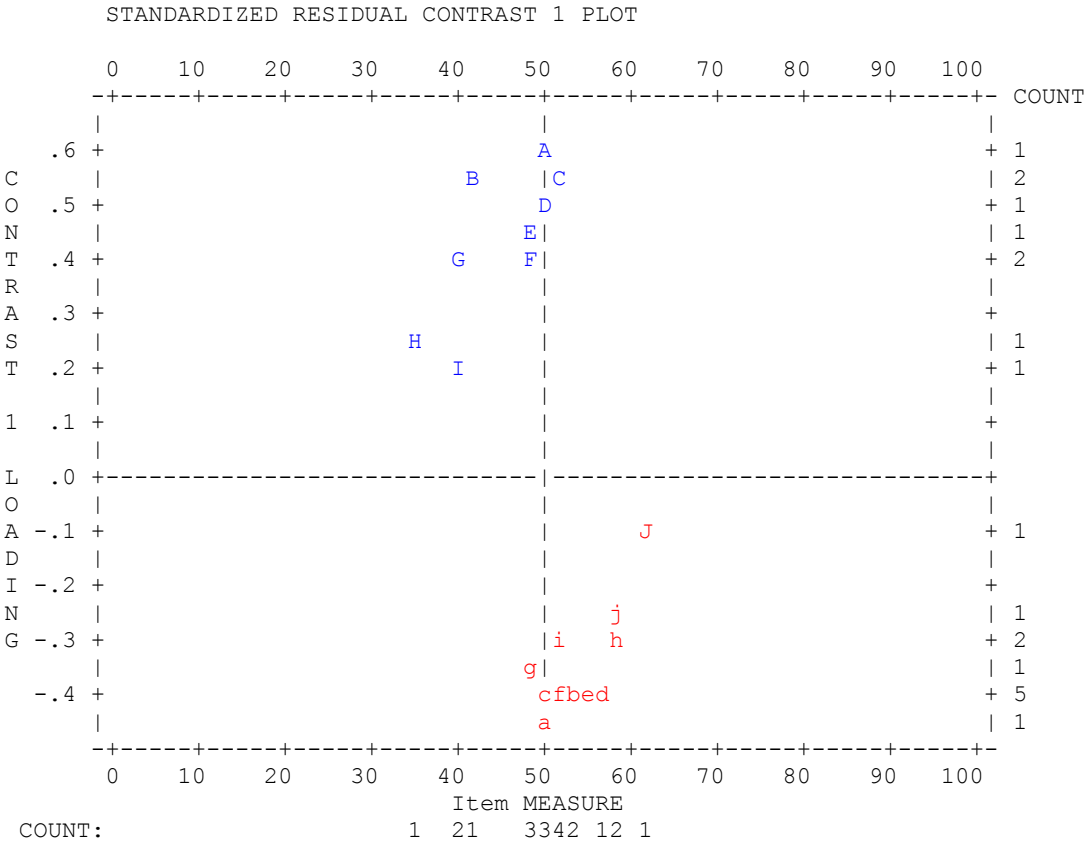

Figure 5.18. Standardised residual plot. Trait anxiety absent items are in blue and trait anxiety present items are in red. Letter to item conversion key is given in table 5.24.

5.3h STAI-T Anxiety Absent Results

After combination of response categories 3 and 4, 280 patients responded which gave a PSI of 2.07, PSR of 0.81, participant mean of 48.78 and item mean of 50.17.

Items		Skew	Kurtosis	Missing data (%)	Response category 1 (Floor, %)	Response category 2 (%)	Response category 3&4 (%)	Infit	Outfit
Q21	I feel pleasant	0.20	-1.38	1	37	36	26	1.10	1.10
Q23	I am content	-0.18	-1.54	4	30	30	40	0.95	0.95
Q26	I make decisions easily	-0.51	-1.18	3	23	29	48	**1.34**	**1.60**
Q27	I feel secure	-0.27	-1.46	1	27	31	42	0.86	0.83
Q30	I am happy	0.33	-1.25	1	43	26	28	0.76	0.75
Q33	I am “calm, cool and collected”	0.34	-1.52	2	46	26	28	0.87	0.79
Q34	I feel rested	-0.21	-1.55	2	30	29	28	**1.52**	**1.78**
Q36	I feel satisfied with myself	0.32	-1.48	3	44	28	28	**0.68**	**0.66**
Q39	I am a steady person	0.26	-1.54	1	43	28	29	0.90	0.82

Table 5.25. Descriptive statistics of the STAI-T Anxiety Absent items showing reliability indicators and proportion of responses within each category. Fit values outside the range 0.7-1.3 are in **bold**.

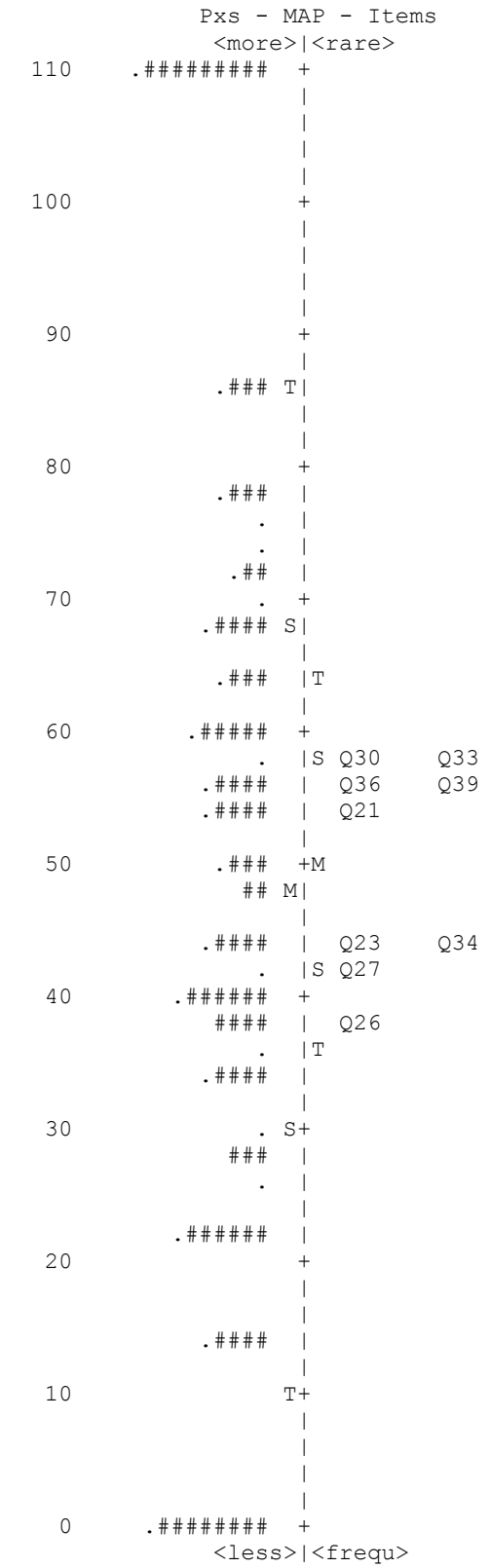

Figure 5.19. Item map for STAI-T Anxiety Absent items. Each '#' IS 3 participants.

5.3h.i) Principal Components Analysis for STAI-T Anxiety Absent items

The raw variance explained by the measures was 54.4%, which is below the 60% suggested as indicating unidimensionality. However, the eigenvalue of the first contrast was 1.5 which is less than 2.0 suggesting that any other dimensions within the data were not significant. The data within STAI-T Anxiety Absent therefore appear to be unidimensional.

5.3h.ii) STAI-T Anxiety Absent item removal

Q34 (I feel rested) misfits more than any other, therefore this was removed and the data reanalysed. The resulting PSI was 2.00, the PSR was 0.80 and the participant and item means are almost unchanged at 49.01 and 50.38 respectively. As there is a reduction in PSI and the item and participant means are already very close, it is probably best to retain Q34.

Items		Infit	Outfit
Q21	I feel pleasant	1.12	1.28
Q23	I am content	1.00	0.96
Q26	I make decisions easily	**1.38**	**1.89**
Q27	I feel secure	0.96	1.01
Q30	I am happy	0.81	0.79
Q33	I am “calm, cool and collected”	0.91	0.85
Q36	I feel satisfied with myself	0.72	0.70
Q39	I am a steady person	1.04	1.06

Table 5.26 Fit values for STAI-T Anxiety absent items after removal of Q34.

After removal of Q34, Q26 is the only remaining item to misfit, therefore this was removed next. This did not have a beneficial effect on the PSI which was lowered to 1.88 and is below the threshold of 2.

5.3i STAI-T Anxiety Present Results

280 participants responded which gave a PSI of 2.31, PSR of 0.84, participant mean of 42.76 and item mean of 49.93.

Items		Skew	Kurtosis	Missing data (%)	Response category 1 (Floor, %)	Response category 2 (%)	Response category 3&4 (%)	Infit	Outfit
Q22	I feel nervous and restless	0.24	-0.73	0	34	53	13	0.83	0.82
Q24	I wish I could be as happy as others seem to be	0.22	-1.40	1	38	35	26	**1.58**	**1.65**
Q25	I feel like a failure	0.96	-0.31	1	59	30	11	0.94	0.83
Q28	I feel that difficulties are piling up so that I cannot overcome them	0.66	-0.79	1	50	36	14	0.96	0.93
Q29	I worry too much over something that really doesn't matter	0.27	-1.17	0	37	42	21	0.86	0.89
Q31	I have disturbing thoughts	0.73	-0.63	1	52	36	12	1.08	1.07
Q32	I lack self confidence	0.38	-1.12	1	41	40	19	1.12	1.13
Q35	I feel inadequate	0.62	-0.63	3	49	42	10	0.93	0.94
Q37	Some unimportant thought runs through my mind and bothers me	0.36	-0.93	0	39	46	15	0.81	0.83
Q38	I take disappointments so keenly that I can't put them out of my mind	0.37	-1.13	0	41	40	19	1.07	1.13
Q40	I get in a state of tension or turmoil as I think over my recent concerns and interests	0.38	-0.92	1	40	45	15	0.80	0.80

Table 5.27. Descriptive statistics of the STAI-T Anxiety Present items showing reliability indicators and proportion of responses within each category. Fit values outside the range 0.7-1.3 are in **bold**.

5.3i.i) Principal Components Analysis for STAI-T Anxiety Present items

The raw variance explained by the measures was 48.2%, which is below the 60% suggested as indicating unidimensionality. However, the eigenvalue of the first contrast was 1.8 which is less than 2.0 suggesting that any other dimensions within the data were not significant. The data within STAI-T Anxiety Present items therefore appear to be unidimensional.

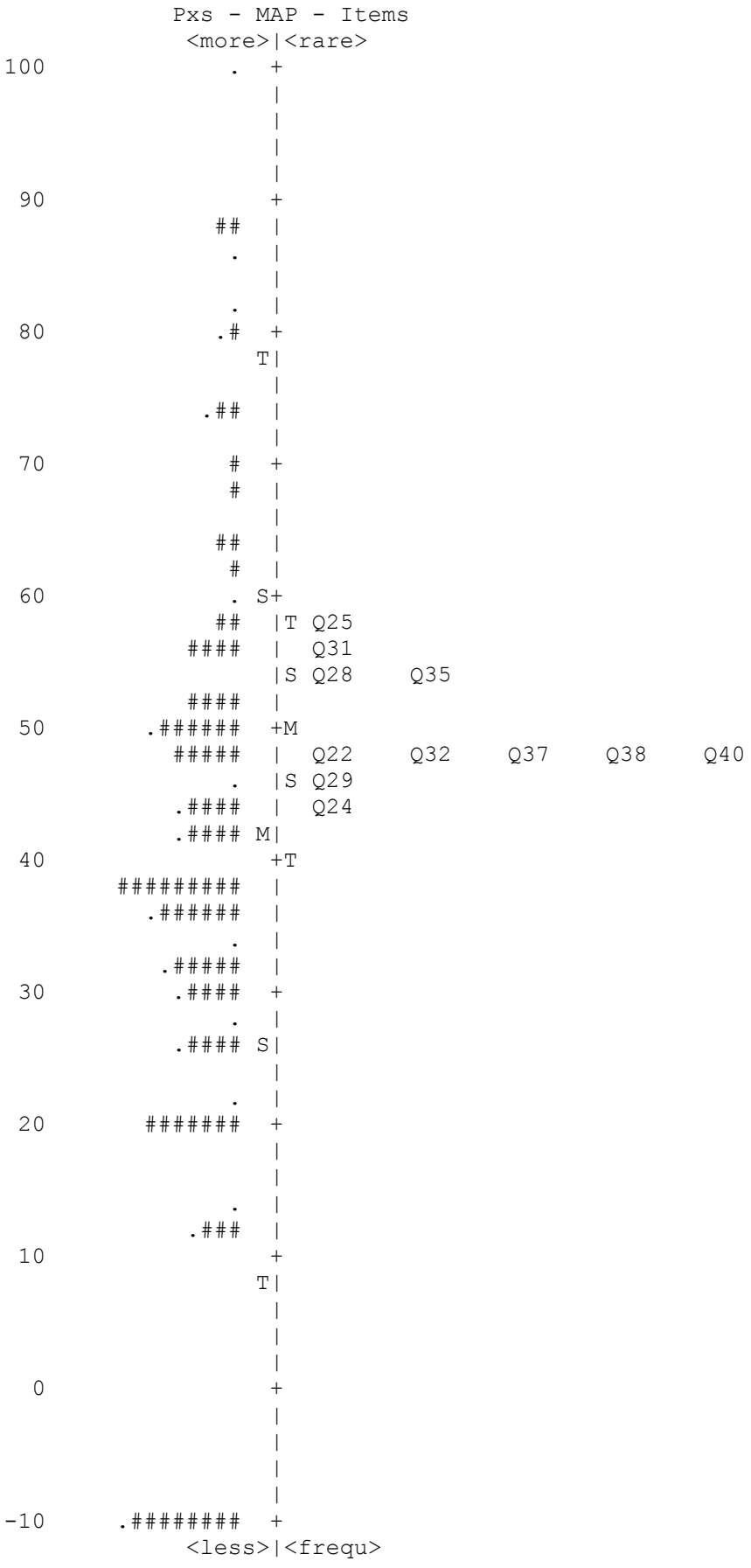

Figure 5.20. Item map for STAI-T Anxiety Present items. Each '#' IS 3 participants.

5.3i.ii) STAI-T Anxiety Present items item removal

Q24 is the only item to misfit, therefore this was removed and the data re-analysed resulting in a PSI of 2.27, PSR of 0.84, participant mean of 42.27 and item mean 49.93. The fit values of the remaining items are in table 5.28. Removal of Q24 results in a slight deterioration of PSI and a marginal increase in separation of participant and item means therefore it should probably be retained if maximum PSI is required.

Items		Infit	Outfit
Q22	I feel nervous and restless	0.92	0.93
Q25	I feel like a failure	0.99	0.89
Q28	I feel that difficulties are piling up so that I cannot overcome them	1.02	0.98
Q29	I worry too much over something that really doesn't matter	0.88	0.92
Q31	I have disturbing thoughts	1.15	1.15
Q32	I lack self confidence	1.20	1.19
Q35	I feel inadequate	0.98	0.99
Q37	Some unimportant thought runs through my mind and bothers me	0.84	0.85
Q38	I take disappointments so keenly that I can't put them out of my mind	1.17	1.23
Q40	I get in a state of tension or turmoil as I think over my recent concerns and interests	0.85	0.83

Table 5.28. Fit values for STAI-T Anxiety Present items after removal of Q24.

5.4 Discussion

This is the first time that the HADS and STAI questionnaires have been used in an ophthalmic outpatient population.

5.4a HADS

The HADS and its subscales would be better matched to a population with higher levels of psychological distress than new ophthalmic outpatients, as demonstrated by the Rasch maps (Figures 5.2, 5.5, 5.7) and presence of floor effects for most items (Tables 5.2, 5.14, 5.17). Previous studies have also found a floor effect using HADS-T and interpreted this to mean that it is targeted at more severe anxiety and depression levels than those present in mixed oncology and psychology patient hospital populations (Smith et al., 2006) or patients with Parkinsons Disease (Forjaz et al., 2009). Using a questionnaire that shows a floor effect in a new ophthalmology outpatient population raises the risk of not accurately quantifying or discriminating between the patients with lower levels of anxiety and/or depression. As the response category for the severest symptoms was rarely used, combination of the top two response categories improved the suitability of this questionnaire, and its subscales, to this population.

Different populations appear to respond quite differently to the HADS, with all previous Rasch studies (Smith et al., 2006, Gough and Hudson, 2009, Pallant and Tennant, 2007, Tang et al., 2008, Lambert et al., 2010), along with the present study, reporting different questions as easiest and hardest for all subscales. At the same time, the same body of research, with a wide variety of populations has supported the removal of item 11 from the Anxiety subscale (I feel restless as I have to be on the move) in order to improve reliability and fit to the Rasch model (Smith

et al., 2006, Gough and Hudson, 2009, Pallant and Tennant, 2007, Tang et al., 2008, Lambert et al., 2010) and the present study would support this (Table 5.14). The symptom of restlessness is clearly not a good fit to the other anxiety questions and this could be due to restlessness being feasibly caused by other physical or psychological issues other than anxiety. In addition to the present study, two previous studies (Smith et al., 2006, Lambert et al., 2010) have found misfit for item 14 (I can enjoy a good book or radio or TV program) which is meant to elicit depression but could also be affected by a number of physical or psychological problems. With the exception of item 11, it is possibly premature to recommend removal of items prior to future administration of HADS on the basis of reducing respondent burden. Although, using common sense, if any questions appear to conflict with the population, for example "I can enjoy a good book or radio or TV program" in a severely sight impaired cohort, then it should be removed to improve reliability. In cohorts that are as obviously biased as severely sight impaired, it should be questioned whether scales such as the HADS should be administered, or whether something more fitting to the population should be used or developed. As more studies are published that use Rasch analysis on the HADS it is possible that more trends may emerge.

HADS-T was found to be not unidimensional using a cut-off for the eigenvalue of the second contrast of 2.0. We found the first factor to only explain 49 % of the variance in the data and an eigenvalue of 2.2 for the second contrast. Many researchers view this as indicating a non-unidimensional instrument (Smith et al., 2002, Gothwal et al., 2010, Linacre, 2009, Marella et al., 2010). The eigenvalue of 2.2 for the second factor agrees with findings in previous studies (Smith et al., 2006,

Forjaz et al., 2009) who found eigenvalues of 2.4 and 2.11 respectively. However, neither of these studies provided details of the percentage variance explained by the first factor and most importantly, they used a value of 3.0 for the cut-off eigenvalue and argued that their data and analyses suggested that the HADS-T *is* unidimensional. The issue of unidimensionality should not be based solely on simple numbers such as a cut-off eigenvalue and consideration must be made of the size of the variance explained by the principal factor and whether any subgrouping of the items makes sense (Linacre, 2008, Vianya-Estopa et al., 2010, Smith et al., 2002). In the case of the HADS, the questionnaire was developed to assess both anxiety and depression, so that any subgrouping by data analysis into these two domains would make sense. In our analyses, PCA split the items perfectly into those assessing anxiety and those assessing depression (Figure 5.3) which indicated logically that patients answered the two different groups of questions in different ways. Previous studies concur with this apparently sensible split (Lambert et al., 2010) including Pallant and Tennant (2007) who found a similar distribution of items using Rasch PCA on 296 patients attending an out-patient musculoskeletal rehabilitation program (they found one misfitting item, which was Anxiety 7, "I can sit at ease and feel relaxed", but the rest split perfectly into anxiety and depression subscales). Once again, they argue that the questionnaire can be considered unidimensional as the scores from each subscale were similar to each other, and to HADS-T, and the discriminative ability of the HADS-T (as shown by PSI) was very good. We also found a good PSI for HADS-T (PSI of 2.18, PSR of 0.82 and similar to Pallant and Tennant's 0.89), but the PSI of both HADS-A (1.94) and particularly HADS-D (1.31) was poorer and below the oft-used cut-off value to suggest useful discriminative ability of 2.0

(Pesudovs et al., 2007). In addition, mean values of person and items means for HADS-A, HADS-D and HADS-T were all similar. Whether that means that the HADS-T score is viable despite not being unidimensional is questionable and not answerable with these data. However it can be concluded that HADS-T is much better than its subscales on all measurements of discriminative ability and therefore if any one of the scales is to be used in this population, it would appear to be HADS-T.

5.4b STAI

Both state and trait scales of the STAI showed good discriminative ability (PSI>2.0) and for both anxiety absent and present items subscales, apart from STAI-S Anxiety Present items which only achieved a maximum PSI of 1.69.

The PCA assessment of unidimensionality for both state and trait scales that showed two factors within the data agrees with the original author's two factor model for anxiety present and anxiety absent items (Spielberger, 1983). Multiple studies have since agreed that higher scores are provided by respondents for anxiety absent items such as "I feel…calm, at ease, satisfied, comfortable etc" compared to anxiety present items such as "I feel…..strained, upset, frightened, jittery etc" (e.g., Spielberger, 1983; Mook et al., 1992; Kvaal et al., 2001). This is because confirming the presence of anxiety is not psychologically equivalent to not confirming the presence of calmness.

A recent study has performed Rasch analysis on the STAI and suggested that the questionnaire had many poorly fitting items and response category disorders and suggested a reduced questionnaire that was unidimensional and had good

psychometric properties (Kaipper et al., 2010). However, their rationale for category reduction seems illogical in that it does not consider the distribution of responses and combines the middle two response categories leaving end categories containing as little as 1.4% of the data. In these cases where an end category has less than 10% of the data (in 8 of the 23 items that had response categories combined), combining the middle two categories effectively provides a two category scale. The logical and standard technique (Pesudovs et al., 2007) is to combine the end categories if it unites the categories with the least amount of data, which would have produced a viable and discriminative three-category scale. The paper is further limited by important typos in the tables (e.g., Table 5 provides data for two 'trait' scales and not a 'trait' and 'state' scale, some percentages are incorrect and there is no acknowledgment of the reverse scoring required for approximately half the items). The removal of items from the questionnaire makes no consideration of the well known separation of the items into anxiety-absent and anxiety present (Spielberger, 1983) as well as using an illogical response category reduction. The finding of unidimensionality for both State and Trait scales is in contradiction to the literature on the questionnaire (Spielberger, 1983, Hishinuma et al., 2000, Kvaal et al., 2001, Mook et al., 1992) including the findings of the original author and the present study. The cause of this is unclear, but it is even possible that the reverse scoring required for the anxiety-absent items was not performed as this is not mentioned and the items are listed without reverse scoring in Table 5.

5.5. Limitations of the study

The larger proportion of missing data for the second half of the STAI (STAI-T) was presumably due to patients failing to turn over the sheet of paper, despite "Questions continue overleaf" and an arrow being printed and highlighted at the bottom of the page. The questionnaires were administered (placed in the envelope) in a random order to neutralise order bias, however this order was not recorded. The HADS is only a single sided sheet therefore it is possible that the patient, having completed the HADS first, may have assumed the STAI was also single sided. The large proportion of participants that didn't complete the second half of the STAI (STAI-T) could also be an indication of respondent burden. Numerous attempts at shortening the State subscale of the STAI have been made in order to reduce the burden on participants (Marteau and Bekker, 1992, Chlan et al., 2003) and the present study supports the suggestion that some items assess very similar levels of symptom in this population. A Rasch shortened STAI with optimum participant separation would be useful.

5.6 Conclusion

Despite having a floor effect, HADS-T, STAI-S and STAI-T show good discrimination between patients when administered to a population of new ophthalmic outpatients. All three scales are not unidimensional, but split into well-established and logical subscales with PCA.

6. Levels of psychosocial distress in patients who have been referred to secondary eye care

6.1 Introduction

As discussed in detail in section 1.1, when a patient is referred to the hospital the NHS has to pay the hospital for the attention they receive, therefore high levels of false positive referrals will naturally result in potentially unnecessary spending of public money. From research in breast cancer and genetic screening of young women and neonates screening it has been shown that referral for further investigation raises levels of psychosocial distress (Gram et al., 1990, Brewer et al., 2007, Stewart-Brown and Farmer, 1997, Cockburn et al., 1994, Marteau et al., 1988, Tymstra, 1986, Keyzer-Dekker et al., 2011, Gøtzsche and Nielsen, 2011) therefore there is an additional psychological burden that patients have to bear if they are referred. It is unknown whether referral from primary care to secondary eye care similarly results in raised levels of distress. The terms psychological distress and psychosocial distress tend to be used interchangeably and are over arching terms referring to psychogenic pain which can be caused by anxiety, stress, depression and a number of other (often overlapping) psychological symptoms.

The present study aimed to determine whether being referred as a new ophthalmic outpatient raises levels of psychosocial distress. The questionnaires that were evaluated in a population of new ophthalmic outpatients in the previous chapter were used. The State-Trait Anxiety Inventory (STAI) plays an important role in this

study because it allows the quantification of the propensity of a patient to be anxious or distressed (trait anxiety) and can separate this from anxiety caused by the patient's current situation (state anxiety). In addition, the data from the Hospital Anxiety and Depression Scale (HADS) allowed us to use the original authors cuts off values (Snaith and Zigmond, 1994), which have been subsequently validated (Bjelland et al., 2002), to quantify the proportion of patients that had levels of distress raised above normal levels to determine the significance of any increases in anxiety and/or depression caused by hospital referral.

6.2 Method

In order to determine whether hospital patients had raised levels of psychological levels of distress, the level of distress in a control group also had to be determined. The most suitable control group was patients that had an eye examination in primary care but had not been referred. Local optometric practices were approached via a local optical committee meeting and invited to recruit patients on our behalf. Seven optometry practices (4 independent, 3 multiple) agreed to participate. The optometrists asked all patients within the inclusion criteria (over 16 years of age and not needing referral to secondary eye care) if they would participate in the study and those who were interested were given information sheets and the questionnaires. Completion of the questionnaires was taken as implied written consent.

Both traditional scoring and Rasch converted scores were analysed. In order to compare the control group to the main cohort the results from both had to be combined, and Rasch was performed as above to optimise the scales and create a conversion key.

When patients missed items, as long as enough questions had been completed to retain validity according to the instrument manual (Spielberger, 1983, Snaith and Zigmond, 1994) the patients' answers were perorated to give a mean total score. An example of peroration would be if a patient had missed one question from STAI-S, an average of the remaining 19 was taken and multiplied by 20 to give a mean total score.

6.3 Results

6.3a Demographics

Characteristic	Patients referred (n=322)	Control (n=80)
Mean age (years ± SD)	61 ± 19	61 ± 16
Gender: Female	170 (53%)	43 (53.75%)
Male	144 (45%)	30 (37.5%)
Unspecified	8 (2%)	7 (8.75%)
Ethnicity: White	188 (58%)	71 (89%)
Asian	39 (12%)	2 (3%)
Black	6 (2%)	1(1%)
Not Stated	88 (27%)	6 (8%)
Any other group	1(<1%)	

Table 6.1. Demographic data for the main cohort (new ophthalmic outpatients) and the control group (patients who have had an eye examination, but have not been referred).

The age and gender of the main cohort were similar to the control group however there was a difference in ethnicity. As discussed in Chapter 2.1, although more investigation is required into cross cultural equivalence, the literature suggests that the psychometric properties of the scales hold up well across different ethnicities.

6.3b HADS-T

401 responses were Rasch analysed (321 main cohort and 80 control). The resulting Participant Separation Index (PSI) was 2.20 and Participant Separation Reliability (PSR) was 0.83. In order to improve participant separation, as reported in the previous chapter, the top two response categories (2 and 3) were combined which resulted in an increased PSI of 2.30, PSI of 0.84, participant mean of 39.62 and item mean of 49.91. Mean Rasch item scores and mean Rasch total scores are given in table 6.2. Removal of any misfitting items resulted in a decrease is PSI, therefore all items were retained.

Even after response categories 2 and 3 were combined, the data were not normally distributed (Kolmogorov-Smirnov test $p<0.001$), therefore a Mann-Whitney U Test (using a null hypothesis of both populations having the same score) was used to detect whether scores from the main cohort were significantly higher than the control (Armstrong et al., 2011). The Mann-Whitney U test is a non-parametric test to detect whether the difference between scores from two non-normally distributed and unrelated populations is significant. This gave $p=0.11$, therefore the null hypothesis was accepted meaning the two groups do not have significantly different total scores.

Items		Control (n=80)		Main cohort (n=321)	
		Mean	SD	Mean	SD
Q1	I feel tense or 'wound up'	40.4	12.7	39.6	14.8
Q2	I still enjoy things I used to enjoy	39.7	11.7	44.3	14.3
Q3	I get a sort of frightened feeling as if	40.1	17.4	40.7	17.9

	something awful is about to happen				
Q4	I can laugh and see the funny side of things	46.4	8.2	49.7	12.5
Q5	Worrying thoughts go through my mind	38.1	16.2	40.7	16.3
Q6	I feel cheerful	43.8	10.2	47.2	13.7
Q7	I can sit at ease and feel relaxed	39.8	14.8	40.9	14.9
Q8	I feel as if I am slowed down	33.7	12.8	38.5	15.8
Q9	I get sort of frightened feeling like 'butterflies' in the stomach	41.6	13.4	42.4	14.4
Q10	I have lost interest in my appearance	41.3	12.8	45.6	16.0
Q11	I feel restless as I have to be on the move	38.8	17.1	38.7	16.6
Q12	I look forward with enjoyment to things	43.0	12.8	45.8	14.3
Q13	I get sudden feelings of panic	39.9	13.4	42.0	16.1
Q14	I can enjoy a good book or radio or TV program	47.7	11.8	48.9	12.7
Mean item score for all questions		41.0	8.7	43.3	9.9
Perorated mean total score		574.2	121.4	605.6	138.5

Table 6.2. Mean item scores (± SD, Standard Deviation) for HADS-T after combining response categories 2 and 3 and rescoring to create ordinal data.

Traditional Likert scoring gave the average mean total scores for the main cohort and for the control group as 10.2 (±7.1) and 8.5 (±6.2) respectively. As there was a non-normal distribution (Kolmogorov-Smirnov test $p<0.001$) a Mann-Whitney Test (Wilcoxon Rank Sum Test) was used and showed no significant difference between the two groups ($p=0.08$). 27% of those referred and 19% of the control group had HADS-T scores of 14 or above, which was found as the mean cut-off value to identify patients as cases with higher than normal emotional distress (Bjelland et al., 2002).

6.3c HADS-A

In order to make the data ordinal, both the results from the cohort and the control group were combined in order to perform Rasch analysis, and then rescored.

Similarly to the analysis of the main cohort alone, the PSI was 1.95, PSR was 0.79, the participant mean was 37.71 and the item mean was 51.54. Combination of response categories or item removal did not improve the PSI. Mean Rasch item scores and mean Rasch total scores are given in table 6.3.

The data were not normally distributed (Kolmogorov-Smirnov test $p<0.001$), therefore a Mann-Whitney U Test (using a null hypothesis of both populations having the same score) was used which showed that there was no significant difference between the total scores of the two groups ($p=0.62$).

Items		Control (n=80)		Main cohort (n=321)	
		Mean	SD	Mean	SD
Q1	I feel tense or 'wound up'	37.2	14.4	36.9	17.7
Q3	I get a sort of frightened feeling as if something awful is about to happen	36.6	20.1	37.4	20.8
Q5	Worrying thoughts go through my mind	34.6	19.4	37.6	19.7
Q7	I can sit at ease and feel relaxed	37.0	15.6	38.3	15.9
Q9	I get sort of frightened feeling like 'butterflies' in the stomach	39.0	16.1	39.4	16.0
Q11	I feel restless as I have to be on the move	36.3	19.0	36.1	18.5
Q13	I get sudden feelings of panic	36.2	15.4	38.9	19.0
Mean score for all questions		36.7	13.3	37.8	13.7
Perorated mean total score		256.8	93.1	265.0	96.1

Table 6.3. Mean scores (± SD, Standard Deviation) for HADS-A after rescoring to create ordinal data.

Traditional Likert scoring also resulted in a non-normal distribution (Kolmogorov-Smirnov test $p<0.001$) and the average mean total scores for the main cohort and for the control group are 6.1 (±4.4) and 5.7 (±4.2) respectively. The total scores were not significantly different (Mann-Whitney U test $p=0.54$). 32% of those

referred and 26% of the control group had HADS-A scores above 7, which is the cut-off value used to identify patients as cases with higher than normal anxiety (Snaith and Zigmond, 1994, Bjelland et al., 2002).

6.3d HADS-D

To make the data ordinal, both the results from the cohort and the control group were combined in order to perform Rasch analysis, and then rescored. Combining response categories 2 and 3 gave the highest PSI of 1.37, a PSR of 0.65, participant mean of 37.16 and item mean of 52.01. Mean Rasch item scores and mean Rasch total scores are given in table 6.4. Removal of any misfitting items further reduced the PSI, which is already well below the cut-off of 2, therefore all items were retained to keep participant discrimination as high as possible (as discussed in 5.3d).

Items		Control (n=80)		Main cohort (n=321)	
		Mean	SD	Mean	SD
Q2	I still enjoy things I used to enjoy	34.3	13.3	39.6	16.2
Q4	I can laugh and see the funny side of things	42.7	9.4	46.5	14.2
Q6	I feel cheerful	39.3	11.7	43.2	15.6
Q8	I feel as if I am slowed down	25.4	14.5	30.8	17.9
Q10	I have lost interest in my appearance	36.2	14.5	41.1	18.2
Q12	I look forward with enjoyment to things	38.3	14.7	41.4	16.3
Q14	I can enjoy a good book or radio or TV program	44.0	13.5	45.3	14.4
Mean score for all questions		37.2	8.5	41.2	11.2
Mean total score		260.2	59.6	288.2	78.5

Table 6.4. Mean item scores for HADS-D (± SD, Standard Deviation) after combining response categories 2 and 3, and rescoring to create ordinal data.

Kolmogorov-Smirnov testing showed these data to not be normally distributed ($p<0.001$), therefore a Mann-Whitney U Test (using a null hypothesis of both populations having the same score) was used which showed that scores from the main cohort were significantly higher than the control ($p=0.007$)

Traditional Likert scoring was also not normally distributed (Kolmogorov-Smirnov $p<0.001$), and the average mean total scores for the main cohort and for the control group are 4.1 (±3.5) and 2.8 (±2.5). Mann-Whitney U testing also showed this difference to be significant ($p=0.004$). 17% of those referred and 5% of the control group had HADS-D scores above 7, which is the cut-off value used to identify patients as cases with higher than normal depression (Snaith and Zigmond, 1994, Bjelland et al., 2002).

6.3e STAI-S

As discussed in the previous chapter, the STAI State and Trait anxiety subscales each have two factors within them; items containing Anxiety Absent questions and items containing Anxiety Present questions. Therefore as the State and Trait subscales were not unidimensional, the two factors from each subscale had to be split before analysis. As some previous studies have not split the subscales into their constituent factors prior to analysis, to allow comparison this was also performed in addition to the correct method:

Combining response categories 3 and 4 gave the highest PSI of 2.75, a PSR of 0.88, a participant mean of 42.88 and an item mean of 49.68. Kolmogorov-Smirnov testing showed these data to not be normally distributed ($p<0.001$), therefore a Mann-Whitney U Test (using a null hypothesis of both populations having the same score) was used which gave $p=0.008$ therefore the null hypothesis is rejected meaning that scores from main cohort are significantly higher than the control. Traditional Likert scoring was also not normally distributed (Kolmogorov-Smirnov $p<0.001$), and the average mean total scores for the main cohort and for the control group were 35.6 (±12.7) and 32.0 (±11.4) which were also significantly different (Mann Whitney U $p=0.012$).

6.3e.i) STAI-S Anxiety Absent items

To make the data ordinal, both the results from the cohort and the control group were combined in order to perform Rasch analysis, and then rescored. As concluded in the previous chapter, response categories 3 and 4 were combined which give the highest PSI of 2.28 and PSR of 0.84, but no items had to be removed. The participant and item means were 47.03 and 49.92 respectively. Mean Rasch item scores and mean Rasch total scores are given in table 6.5.

Kolmogorov-Smirnov testing showed these data to not be normally distributed ($p<0.001$), therefore a Mann-Whitney U Test was used which gave $p=0.039$

therefore null hypothesis is rejected meaning scores from main cohort are significantly higher than the control.

Items		Control (n=80)		Main cohort (n=318)	
		Mean	SD	Mean	SD
Q1	I feel calm	44.9	18.6	47.4	18.4
Q2	I feel secure	51.8	18.6	51.2	19
Q5	I feel at ease	40.3	18.6	45.5	20.1
Q8	I feel satisfied	39.3	19.3	45.4	19.7
Q10	I feel comfortable	42.8	20.2	47	19.6
Q11	I feel self-confident	42.8	18.5	45.4	19.7
Q15	I am relaxed	39.5	20.8	44.9	19.5
Q16	I feel content	38.2	19.5	44.7	20.2
Q19	I feel steady	42.2	20	47.1	20.3
Q20	I feel pleasant	42.5	17.6	47.9	20
Mean score for all questions		42.4	14.8	46.6	15.3
Mean total score		424	148	466	153

Table 6.5. Mean item scores for STAI-S Anxiety Absent items (± SD, Standard Deviation) after combining response categories 3 and 4, and rescoring to create ordinal data.

Traditional Likert scoring was also not normally distributed (Kolmogorov-Smirnov $p<0.001$), and the average mean total scores for the main cohort and for the control group were 18.7 (±6.0) and 17.1 (±8.8) respectively. These were also significantly different (Mann Whitney U $p=0.035$).

6.3e.ii) STAI-S Anxiety Present items

To make the data ordinal, both the results from the cohort and the control group were combined in order to perform Rasch analysis, and then rescored. This combined group also behaved in a similar way to the previous chapter; response categories 3 and 4 were combined which give the highest PSI of 1.69 and PSR of

0.74, but no items had to be removed. The participant and item means were 40.54 and 49.94 respectively. Mean Rasch item scores and mean Rasch total scores are given in table 6.6.

Kolmogorov-Smirnov testing showed these data to not be normally distributed ($p<0.001$), therefore a Mann-Whitney U Test was used which gave $p=0.01$ therefore the null hypothesis is rejected meaning scores from main cohort are significantly higher than the control.

Items		Control (n=80)		Main cohort (n=318)	
		Mean	SD	Mean	SD
Q3	I am tense	29	15.5	35.5	18.1
Q4	I feel strained	35.3	16.2	37.7	17.8
Q6	I feel upset	42.3	11.4	45.4	15.1
Q7	I am presently worrying over possible misfortunes	34.5	15.5	36.9	17.4
Q9	I feel frightened	44.4	12.7	45.8	14.3
Q12	I feel nervous	32.2	15	37.6	16.8
Q13	I am jittery	41.2	14.4	42.8	15.5
Q14	I feel indecisive	37.6	13.8	40.2	15.8
Q17	I am worried	34	15.5	37.3	16.4
Q18	I feel confused	43.4	12.8	46.1	14.8
Mean score for all questions		37.6	10.9	40.6	11.6
Mean total score		375.6	109	406.2	116.3

Table 6.6. Mean item scores for STAI-S anxiety present items (± SD, Standard Deviation) after combining response categories 3 and 4, and rescoring to create ordinal data.

Traditional Likert scoring was also not normally distributed (Kolmogorov-Smirnov $p<0.001$), and the average mean total scores for the main cohort and for the control

group were 15.6 (±5.5) and 14.2 (±5.2) respectively. These were also significantly different (Mann Whitney U p=0.009).

6.3f STAI-T

STAI-T was again split into Anxiety Absent and Anxiety Present items but to allow comparison with previous literature these data were first analysed together; The optimum PSI was obtained by combining response categories 3 and 4 but not removing any items. After the data was made ordinal Kolmogorov-Smirnov testing showed these data to not be normally distributed (p=0.007), therefore a Mann-Whitney U Test was used which gave p=0.096 therefore null hypothesis is accepted meaning that the main cohort did not have significantly different scores to the control. Traditional Likert scoring also resulted in a non-normal distribution and the average mean total scores for the main cohort and for the control group were 37.8 (±12.0) and 34.9 (±10.8) respectively. Mann-Whitney U testing showed the total scores for the two groups to not be significantly different (p=0.074).

6.3f.i) STAI-T Anxiety Absent items

To make the data ordinal, both the results from the cohort and the control group were combined in order to perform Rasch analysis, and then rescored. As concluded in the previous chapter, response categories 3 and 4 were combined which gave the highest PSI of 2.04 and PSR of 0.81 and participant and item means of 48.24 and 50.14 respectively. Mean Rasch item scores and mean Rasch total scores are given in table 6.7. Similarly to the previous chapter item Q34 misfitted (infit 1.48, outfit

1.69) but when removed it resulted in a PSI of 1.96, below the threshold of 2, and was therefore retained.

After the data was made ordinal, Kolmogorov-Smirnov testing showed these data to not be normally distributed ($p<0.001$), therefore a Mann-Whitney U Test (using a null hypothesis of both populations having the same score) was used which gave $p=0.16$ therefore null hypothesis is accepted meaning that the main cohort did not have significantly different scores to the control.

Items		Control (n=74)		Main cohort (n=280)	
		Mean	SD	Mean	SD
Q21	I feel pleasant	48.8	17.9	51.5	19.2
Q23	I feel satisfied with myself	45.2	20.9	46.4	20.3
Q26	I feel rested	39.4	19.4	44.2	19.5
Q27	I am “calm, cool and collected”	43.1	20.1	46.6	19.9
Q30	I am happy	50.5	17.1	53.3	19.2
Q33	I feel secure	50.1	18.7	52.5	20.6
Q34	I make decisions easily	44	19.9	46.7	20.4
Q36	I am content	49.8	19.6	53	20.3
Q39	I am a steady person	49.1	19.6	52.4	20.5
Mean score for all questions		46.8	15.4	49.8	16.1
Perorated mean total score		421.2	139	448	144

Table 6.7. Mean item scores for STAI-T Anxiety Present items (± SD, Standard Deviation) after combining response categories 3 and 4, and rescoring to create ordinal data.

Traditional Likert scoring also resulted in a non-normal distribution (Kolmogorov-Smirnov $p<0.001$) and the average mean total scores for the main cohort and for the control group were 17.8 (±5.9) and 16.8 (±5.7) respectively. Mann-Whitney U testing showed the total scores for the two groups to not be significantly different ($p=0.16$).

6.3f.ii) STAI-T Anxiety Present items

To make the data ordinal, both the results from the cohort and the control group were combined in order to perform Rasch analysis, and then rescored. As concluded in the previous chapter, response categories 3 and 4 were combined which gave the highest PSI of 2.31 and PSR of 0.84 and participant and item means of 42.02 and 49.93 respectively. Mean Rasch item scores and mean Rasch total scores are given in table 6.8. Similarly to the previous chapter item Q24 was the only item to misfit (infit 1.56, outfit 1.64) but when removed it resulted in a reduction of the PSI to 2.28, and a marginal increase in the separation of the participant and item means to 41.49 and 49.92 respectively. Q24 was therefore retained.

After the data was made ordinal Kolmogorov-Smirnov testing showed these data to not be normally distributed ($p<0.001$), therefore a Mann-Whitney U Test (using a null hypothesis of both populations having the same score) was used which gave $p=0.103$ therefore null hypothesis is accepted meaning that the main cohort did not have significantly different scores to the control.

Items		Control (n=74)		Main cohort (n=280)	
		Mean	SD	Mean	SD
Q22	I feel nervous and restless	37.1	14	42.4	16.7
Q24	I wish could be as happy as others seem to be	36.7	17.8	41.1	20.3
Q25	I feel like a failure	43.2	14.4	46.7	17.4
Q28	I feel that difficulties are piling up so that I cannot overcome them	40.6	15.3	44.4	18.1
Q29	I worry too much over something that really doesn't matter	38.1	18.9	40.9	19
Q31	I have disturbing thoughts	43.3	15.2	44.3	17.5
Q32	I lack self confidence	40.5	17.2	41.1	18.9
Q35	I feel inadequate	41.6	16.5	44.5	16.7
Q37	Some unimportant though runs through my mind and bothers me	39	16.1	42.3	17.8
Q38	I take disappointments so keenly that I can't put them out of my mind	39.9	19.5	41.1	19
Q40	I get in a state of tension or turmoil as I think over my recent concerns and interests	37.6	16.4	42.9	17.8
Mean score for all questions		39.9	12.1	42.9	13.2
Perorated mean total score		439.2	133	472	145

Table 6.8. Mean item scores for STAI-T Anxiety Present items (± SD, Standard Deviation) after combining response categories 3 and 4, and rescoring to create ordinal data.

Traditional Likert scoring also resulted in a non-normal distribution (Kolmogorov-Smirnov $p<0.001$) and the average mean total scores for the main cohort and for the control group were 19 (±5.7) and 17.7 (±5.7) respectively. Mann-Whitney U testing showed the total scores for the two groups to not be significantly different ($p=0.094$).

6.3g STAI-S with participants that also completed STAI-T

STAI-S and STAI-T had different sample sizes; n=276 completed both subscales, n=42 completed only STAI-S, n=4 completed only STAI-T. Therefore the difference in significance of STAI-S and STAI-T could simply have been an artefact of the difference in sample sizes. In order to be able to fairly compare the results, STAI-S Anxiety Absent and Anxiety Present subscales were re-analysed using only the results from patients who completed both subscales to see if this had a detrimental effect on the level of significance.

To allow comparison with previous literature these data were also analysed without splitting into Anxiety Absent and Present subscales; Using this subset, the highest PSI of 2.79 was obtained by combining response categories 3 and 4 and not removing any items. This gave a PSR of 0.88, participant mean of 42.87 and item mean of 49.68. Kolmogorov-Smirnov testing revealed that these data were not normally distributed ($p<0.001$). A Mann-Whitney U test of the participants total scores gave $p=0.006$ therefore scores from main cohort are still significantly higher than the control. A similar result was found with traditional Likert scoring (Kolmogorov-Smirnov $p<0.001$, Mann-Whitney U test $p=0.007$).

6.3g.i) STAI-S Anxiety Absent with participants that also completed STAI-T

Using this subset, the highest PSI of 2.29 was obtained by combining response categories 3 and 4 and not removing any items. This gave a PSR of 0.84, participant

mean of 47.15 and item mean of 49.93. Mean Rasch item scores and mean Rasch total scores are given in table 6.9.

Items		Control (n=80)		Main cohort (n=276)	
		Mean	SD	Mean	SD
Q1	I feel calm	44.3	18.6	47.2	18.8
Q2	I feel secure	50	18.7	49.9	19.4
Q5	I feel at ease	40.4	18.7	46.3	20.3
Q8	I feel satisfied	39.9	19.4	46.4	19.6
Q10	I feel comfortable	42.9	20.3	47.5	19.8
Q11	I feel self-confident	41.7	18.6	45	19.6
Q15	I am relaxed	40.2	20.9	45.8	19.6
Q16	I feel content	39.5	19.6	46.1	20.1
Q19	I feel steady	43.3	20.1	47.9	20.2
Q20	I feel pleasant	42.4	17.7	48.7	20.2
Mean score for all questions		42.5	14.9	47.1	15.5
Mean total score		424.7	148.9	471.03	155.25

Table 6.9. Mean item scores for STAI-S Anxiety Absent items (± SD, Standard Deviation) using only data from participants who completed both State and Trait Subscales. Response categories 3 and 4 have been combined, and the scale rescored to create ordinal data.

Kolmogorov-Smirnov testing revealed that these data were not normally distributed ($p<0.001$). Even with this reduced cohort, a Mann-Whitney U test of the participants total scores gave $p=0.019$ therefore scores from main cohort are still significantly higher than the control. A similar result was found with traditional Likert scoring (Kolmogorov-Smirnov $p<0.001$, Mann-Whitney U test $p=0.021$)

6.3g.ii) STAI-S Anxiety Present with participants that also completed STAI-T

Using this subset, the highest PSI of 1.69 was obtained by combining response categories 3 and 4 and not removing any items. This gave a PSR of 0.74, participant mean of 40.46 and item mean of 49.92. Mean Rasch item scores and mean Rasch total scores are given in table 6.10.

Items		Control (n=80)		Main cohort (n=276)	
		Mean	SD	Mean	SD
Q3	I am tense	29.6	15.5	35.8	18.1
Q4	I feel strained	34.7	16.2	37.4	17.8
Q6	I feel upset	42.5	11.4	45.5	15
Q7	I am presently worrying over possible misfortunes	34.9	15.5	37	17.4
Q9	I feel frightened	44	12.7	45.6	14.3
Q12	I feel nervous	32.8	15.1	37.8	16.8
Q13	I am jittery	39.9	14.5	42.1	16
Q14	I feel indecisive	37.1	13.8	39.9	15.9
Q17	I am worried	34	15.6	37.2	16.6
Q18	I feel confused	43.4	12.9	46.1	14.9
Mean score for all questions		37.5	11.0	40.6	11.7
Mean total score		374.7	109.6	406	117

Table 6.10. Mean item scores for STAI-S anxiety present items (± SD, Standard Deviation) using only data from participants who completed both State and Trait Subscales. Response categories 3 and 4 have been combined, and the scale rescored to create ordinal data.

Kolmogorov-Smirnov testing revealed that these data were not normally distributed ($p<0.001$). Even with this reduced cohort, a Mann-Whitney U test of the participants

total scores gave p=0.009 therefore scores from main cohort are still significantly higher than the control. A similar result was found with traditional Likert scoring (Kolmogorov-Smirnov p<0.001, Mann-Whitney U test p=0.009)

6.4 Discussion

Levels of depression as measured by the HADS-D subscale, were significantly higher in the HES cohort whether scored by Likert or Rasch. In addition, 5% of the controls were above the cut-off figure for HADS-D of 7 and thus were deemed as having 'mild' or worse levels of depression, while 17% of the HES cohort were above this cut-off figure. This is despite the fact the HADS-D showed poor person separation using Rasch analysis, with a PSI of well below 2.0, and removal of misfitting items such as Q14 reduced the PSI further (as discussed in chapter 5.3d). A PSI of 2.0 is typically used to signify a useful level of discriminative ability for a questionnaire (Pesudovs et al., 2007) as it represents the ability to distinguish three distinct strata of person ability. Given that so few of the participants had anything other than normal or mild levels of depression according to HADS-D (4.5% were scored as having moderate symptoms and 0.7% severe), so that there was little opportunity for the HADS-D to show discrimination between three distinct strata of depression symptoms, but it could still clearly discriminate between two levels of depression. This highlights the limitations of using Rasch analysis on a sample of subjects with minimal variation in the symptom level or state being measured.

HADS-A showed no difference in levels of anxiety between the HES cohort and control group. Virtually all items of HADS-A were similar for cohort and control, except for the item "worrying thoughts go through my mind" (question 5, control

34.6, SE 2.2 vs. 37.6 SE 1.1 cohort). Considering that the questionnaire was completed whilst waiting for their first outpatient appointment in ophthalmology, this may be quite telling and this will be discussed further in the discussion of the STAI results. Due to its multidimensionality the original authors of the HADS advised against combining the subscale scores to give a nebulous scale of psychological distress (Snaith and Zigmond, 1994), but despite this it has been done in many previous studies (Bjelland et al., 2002) and therefore results for HADS-T have been given by the present study. The results of HADS-T were not surprisingly mid-way between those for HADS-D and HADS-A in that the difference between HADS-T scores for HES cohort and control approached significance (p=0.08).

Analysis of STAI-T (both Anxiety Absent and Anxiety Present item subsets) showed there was no significant difference in trait anxiety between the control cohort and the cohort that had been referred to the HES. This means that the main cohort were not significantly more prone to being anxious (ie. Trait anxiety). Analysis of STAI-S (both Anxiety Absent and Anxiety Present item subsets) showed that levels of state anxiety, ie. how anxious the patient is "right there and then", were significantly higher in the patients who had been referred to the HES. This indicates that whenever a patient is referred to secondary eye care there is a psychological burden which is a similar finding to other areas of healthcare such as dentistry, oncology or screening for congenital syndromes (Corah et al., 1988, Stark and House, 2000, Marteau et al., 1988, Tymstra, 1986, Cockburn et al., 1994, Keyzer-Dekker et al., 2011, Gøtzsche and Nielsen, 2011). In this instance this conclusion would also have been made using traditional scoring. There are no cut-off levels for

the STAI because when studies have tried to develop such figures, they have found that the trade offs between sensitivity and specificity provided no clear cut-off value (Kabakoff et al., 1997). The levels of STAI-State scores for HES cohort (35.6) and control (32.0) were similar to those with low perceived susceptibility to breast cancer (~31.5) and moderate to high perceived susceptibility (moderate ~34, high ~37) in a breast cancer screening study (Absetz et al., 2003) and thus appears to present a clinically significant as well as a statistically significant increase in scores. The mean score of 35.6 for STAI-State in the HES cohort is also above the 95% confidence limits for normative working adult data for the 60-69 age group of 34.6 (mean ~32.2, SE 2.2, Spielberger, 1983).

Management of suitable patients within primary care instead of secondary care would be another way of reducing the psychological burden on patients and successful shared care schemes have been set up to do this (Ho and Vernon, 2011, Vernon and Adair, 2010, Spencer et al., 1995). Previous studies (Court et al., 2008, Court et al., 2009, Margrain et al., 2003) have shown that patients also experience anxiety when attending primary eye care, therefore a direct comparison is required for patients being managed for similar conditions in primary and secondary care to see which induces the least anxiety. A comparison between secondary care NHS and private patients may also be of interest.

Raised levels of anxiety are a barrier to effective healthcare, in ways such as reducing attention (Easterbrook, 1982), disrupting recall of information (Kent, 1984) and increasing non-compliance (Corah, 1988). If referral to secondary care results in raised levels of anxiety then ways should be sought to minimise the number of false

positive referrals reaching the hospital, such as the service developed in Chapter 3.5, and those described in several recently published studies (Parkins and Edgar, 2011, Devarajan et al., 2011, Bourne et al., 2009, Henson et al., 2003). Additionally, the patient journey at the hospital could be amended to reduce the psychological impact it currently has. Methods of reducing anxiety caused by other areas of healthcare have been investigated, for example; the perceived behaviour of clinicians (Corah et al., 1988), music (Holm and Fitzmaurice, 2008, Bampton and Draper, 1997, Cooke et al., 2005) and hypnosis (Simon, 1999).

6.5 Limitations

The original protocol involved the use of three questionnaires, including the GHQ-28 (Goldberg and Hillier, 1979), but ethical approval difficulties meant that the questionnaire was not used. One of the subscales of the GHQ-28 elicited general health problems (somatic symptoms) of the patient, which would have allowed us to control for this factor in subsequent analyses (or at least determine whether the intervention and control groups had similar levels of general health problems).

Any psychometric instrument (questionnaire) has flaws, but they are used because they are vastly easier and less expensive to administer than the gold standard. The gold standard is generally regarded as an interview with an appropriately trained clinician, and a variety of interview techniques have been used as gold standards (Bjelland et al., 2002). Despite HADS-T being better than its subscales at discriminating between participants, the combination of the scores from the two

subscales was not recommended by the scales authors (Snaith and Zigmond, 1994). It could be argued that due to these limitations the HADS should not be used in this population but it remains very widely used (Bjelland et al., 2002) and allows comparison with existing literature.

Although the study appeared to show a psychological burden on the patients, it appears to be mild in the majority of patients and the timescale of this effect was not investigated. The clinical significance of the effect would be limited if it returned to baseline very soon after the appointment, or after being informed of the absence of pathology in patients who have had false positive referrals. Further work into the duration and clinical significance of this effect is required.

6.6 Conclusion

Patients referred for further investigation in ophthalmology departments experience raised levels of state anxiety as measured by the State-Trait Anxiety Inventory and raised levels of depression as measured by the HADS-D scale.

Chapter 7. Development of a referral refinement service to reduce the number of unnecessary optometric referrals to secondary eye care

Presented as a poster at the College of Optometrists Research Symposium 2010. Poster reproduced in Appendix E.

Abstract published as: Davey CJ, Naru R, Bonner E, Elliott DB. An optometric referral refinement pilot study in Bradford and Airedale. *Ophthalmic and Physiological Optics* 2010 30: 865–866.

7.1. Introduction

Optometrists are reported to refer between about 3 to 6% of all their patients for further investigation by a medical practitioner (Hobley et al., 1992, Port, 1989, Port and Pope, 1988). Data from 1997 suggested that 46% of new referrals to secondary care ophthalmology services originated following abnormal findings at the patients' optometric examination (Pooley and Frost, 1999). The study described in Chapter 3 suggests that this figure is rising, up to 71%, and a greater proportion of referrals are from optometrists. Many UK studies have shown that a significant number of referrals from primary care to secondary eye care are unnecessary (false positive) across all ophthalmic pathologies (Pooley and Frost, 1999, Harrison et al., 1988), and especially for glaucoma (Bell and O'Brien, 1997b, Bowling et al., 2005, Brittain, 1988, Clearkin and Harcourt, 1983, Newman et al., 1998, Patel et al., 2006, Salmon et al., 2007, Tuck, 1991, Tuck and Crick, 1991, Vernon, 1998, Theodossiades and

Murdoch, 1999). In the study described in Chapter 3, the false positive referral rate from optometrists was found to be 29% and the main causes of these referrals were posterior segment diseases and glaucoma suspects.

Methods of reducing numbers of unnecessary referrals have been identified by previous studies, as detailed below. Repetition of the measurement is advised if an abnormality is noted from visual field assessment (Crick and Tuck, 1995, Bell and O'Brien, 1997a, Vernon and Ghosh, 2001, Salmon et al., 2007, Henson et al., 2003) or tonometry (Tuck and Crick, 1994, Bell and O'Brien, 1997a, Henson et al., 2003). The majority of practitioners use non-contact tonometry as their first choice (Willis et al., 2000, Myint et al., 2011) as it is quicker, does not require instillation of anaesthetic and can be performed by non-optometrists. When repeating tonometry it is generally advised that contact tonometry be used (Bell and O'Brien, 1997a, Tuck and Crick, 1994) as this is the gold standard, and is used in secondary eye care. Multiple studies have found that significant numbers of patients referred on the basis of abnormally high intra-ocular pressures (IOPs) turn out to have IOPs within normal limits when repeated at a later date with contact tonometry (Salmon et al., 2007, Sheldrick et al., 1994, Bell and O'Brien, 1997a, Henson et al., 2003). In addition, biomicroscopy with mydriasis (dilatation) has been shown to have better accuracy at diagnosing, and discriminating between, retinal pathologies when compared to ophthalmoscopy without mydriasis (Parisi et al., 1996, Siegel et al., 1990).

Refinement of referral in this way is not part of a routine NHS sight test in the UK outside of Scotland and Wales as defined by the General Ophthalmic Services (GOS,

2011). Therefore to reduce unnecessary false positive referrals, an enhanced service under General Ophthalmic Services Legislation 2008 is required. The aim of the present study was to assess the feasibility and cost-effectiveness of a referral refinement service in Bradford and Airedale PCT. At the time of this study, only one research publication assessing referral refinement of glaucoma suspects had been published (Henson et al., 2003). Since that time, three other papers on referral refinement of glaucoma have been published (Bourne et al., 2009; Deverajan et al., 2011; Parkins & Edgar, 2011), although there are no other publications regarding referral refinement using mydriasis and binocular indirect ophthalmoscopy.

7.2 Methodology

In late 2008, the Bradford and Airedale PCT agreed to fund a six-month study (February to July 2009) to assess the effectiveness of referral refinement, and specified that a 20% sample of Optometric practices (providers) would be funded.

7.2a Study protocol

A protocol was developed in collaboration with the Optometrist Consultant at Bradford Royal Infirmary (BRI; Clare Green) and the optometric clinical leads at the PCT (Ravi Naru and Edwin Bonner). Given that the greatest false positive referrals appear to be in cases of glaucoma suspects as discussed above and also posterior segment disease, protocols were considered that particularly refined referral in these two cases. The protocol was based on the College of Optometrists Guidelines (2005) and a copy of the original protocol given to optometrists is provided in

Appendix D. It stated that appropriate patients could be brought back to the practice on a separate occasion within two weeks to have a repeat examination. The repeat examination could consist of:

i) Contact applanation tonometry and threshold visual fields

ii) Binocular indirect ophthalmoscopy with mydriasis

iii) All of the above tests.

The conditions under which patients could be brought back for contact applanation tonometry and threshold visual fields were defined as:

- IOP of 25mmHg or higher in either eye (or in the case of non contact instruments the average reading should be 25mmHg or higher) with a normal visual field.
- Difference of 5mmHg or more between the two eyes (or in the case of non contact instruments the difference between the average readings should be 5mmHg or higher) with a normal visual field.
- At least 3 points missed in either eye using a suprathreshold screening programme on a visual field machine, with IOP's of less than 25mmHg.

If a patient had a high or asymmetrical IOP in combination with an abnormal visual field screening, as described above, the optometrist was advised to use their professional judgement as to whether it would be in the patient's interests to refer or to repeat measurements.

Referral to the HES via the GP was advised after the referral refinement appointment if any of the following were found:

- The IOP read 25mmHg or above in either eye,
- The difference in pressures between the two eyes was 5mmHg or more,
- The threshold visual fields highlighted a reduction or depression of the visual field.
- A combination of the above abnormalities was present.

7.2b Influence of NICE guidelines

On April 22nd 2009 NICE issued guidelines on the diagnosis and management of glaucoma and ocular hypertension (OHT)(2009). As a result the Association of Optometrists issued advice to its members recommending that all patients with IOPs over 21mmHg be referred (AOP, 2009). The research protocol therefore was required to be amended (Appendix D) so that referral to the HES via the GP was indicated if the IOP at the refinement appointment was 22mmHg or above in either eye, instead of 25mmHg or above.

The conditions under which patients could be brought back for binocular indirect ophthalmoscopy with mydriasis were less strictly described, and the optometrist was left to use their own professional judgement. Examples of applicable suspicious symptoms were suggested as flashing lights, sudden increase in floaters, sudden decrease in vision or visual distortion. Examples of applicable suspicious clinical

findings were loosely given as a decrease in visual acuity or a suspicious part of the retina. If suspicion was confirmed then the patient was to be referred by the relevant pathway. Dilation and biomicroscopy for the purpose of differentiating wet and dry age-related macular degeneration was not covered by this service. This was a decision by the PCT given that a pathway was already in operation for these conditions and overlap may result in patients being referred via the wrong pathway or at the wrong urgency.

In addition to the protocol there were also three general inclusion criteria which had to apply to patients for them to be eligible for participation in the study:

- Old enough to give informed consent (16 or over according to Fraser guidelines)
- Patients registered with a GP practice in contract with Bradford and Airedale PCT.
- Patients not already under the care of any Hospital Eye Services (HES).

7.2c Optometric practice recruitment

All practices in the eligible areas (N=65) were sent letters of information with a copy of the protocol and given the opportunity to express an interest in participating in the study. The practices needed to have contact tonometry and threshold visual field equipment in good working order. The reimbursement for each patient that attended any referral refinement appointment was £42. This reimbursement was

calculated by the PCT. Twenty eight practices (43%) expressed an interest in participating and a sample of thirteen of these consenting providers (i.e., 20% of the total of 65) was selected using a randomisation procedure (random number generator).

7.2d PCT analysis of patient satisfaction

The patient was given a copy of the Patient Information Leaflet on Referral Refinement at the end of the first visit (Appendix D). This also contained a simple patient satisfaction questionnaire to be completed after the referral refinement appointment, and asked the patient to estimate the time spent at the appointment. The PCT wished to ascertain whether the patients were satisfied with the service they received and developed a very simple questionnaire for this purpose. This consisted of three items, which are in appendix D and questioned the patient whether they were satisfied with the service, whether they would have preferred to have been referred to the hospital and whether they were satisfied with the time they had to wait for the appointment. There was no opportunity for any direct involvement in the questionnaire development, as the PCT wished to have independent control of this aspect of the study.

Providers were required to attend a meeting at the start of the trial to introduce the study and allow the protocol to be verbally clarified. If all requirements were met the providers were given remuneration of £42 for each patient that attended a referral refinement appointment. A referral refinement form was completed for

each appointment. Providers were expected to report to the PCT via e-mail the following parameters every two months to allow payment to be authorised and ongoing data analysis to be performed.

1. Number of patients who had a referral refinement appointment.

2. Number of patients who declined a referral refinement appointment and were referred on first findings.

3. Number of patients referred to GP/ HES following referral refinement appointment.

4. Number of patients not referred to GP/HES following referral refinement appointment.

5. Number of patients who did not attend for booked appointments for referral refinement (such patients were referred to the GP/ HES as they would have been prior to the service).

7.2e Analysis of cost effectiveness

These data were collated and forwarded to the author. Subsequently the referral refinement forms were reviewed and further information regarding the type of assessments made were recorded. An economic analysis of any savings made by the service was performed by comparing the cost of the service versus the potential cost of referral to the HES. The cost of each referral refinement appointment was £42. The tariff for a first appointment at the BRI HES was £110 with a tariff for each follow-up appointment of £53. The percentage of patients discharged after referral from optometrists in the study described in chapter 3.1 was only determined after the first appointment (13/66 or 20% of all suspect glaucoma patients and X/Y or Z% visual disturbances/other). To provide a conservative estimate of cost savings, we have used the Royal College of Ophthalmologists' estimate of 50% of referred patients who are subsequently discharged, rather than the slightly higher figure of 55% used in a previous study (Parkins and Edgar, 2011). It has been estimated that each discharged glaucoma suspect has between 2.10 (Parkins and Edgar, 2011) and 2.33 (Henson et al., 2003; Devarajan et al., 2011) HES appointments (or 1.10 to 1.33 follow-up visits) and the lower 1.10 figure was used here to again provide the most conservative figure for any cost savings.

7.3 Results

Between February and July there were 134 patients that were eligible for referral refinement and Figure 7.1 shows the distribution of eligible patients across the providers. Their mean age was 59 years (SD±15.7) with a median of 60 (range 18-

94). Over this timescale 78 (58%) patients were eligible for refinement by contact tonometry and threshold visual fields, 48 (36%) were eligible for refinement using mydriasis and biomicroscopy and 8 (6%) needed all three techniques performing. All of the patients booked a referral refinement appointment within two weeks, and none failed to attend. 83 patients (62%) did not subsequently require referral after further investigation and the breakdown of patients referred or not referred according to the techniques performed at the refinement appointment is shown in table 7.1. Of the patients that had tonometry and visual fields performed at the refinement appointment, 47 (60%) did not require referral. A slightly greater proportion of patients (69%, 33 patients) did not require referral if they had mydriasis and biomicroscopy at the return appointment.

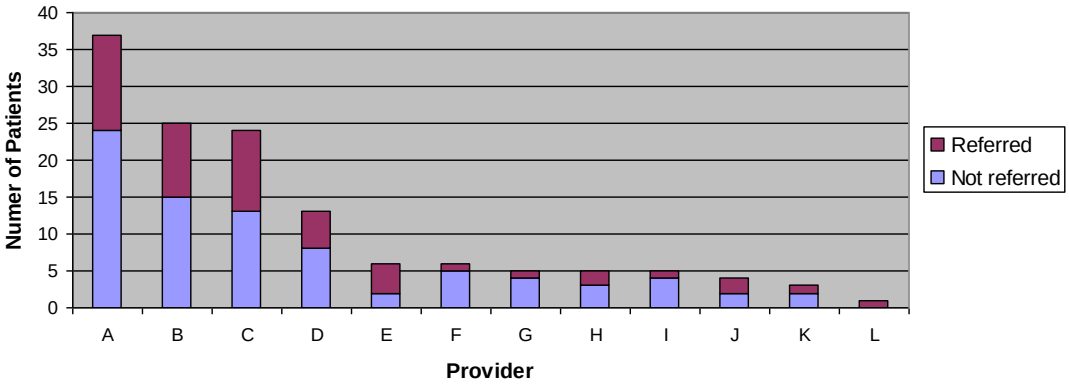

Figure 7.1. Distribution of referral refinement patients across the participating practices between February and July (n=134).

	N	Referred	Not Referred
Tonometry and Visual Fields	78	31 (40%)	47 (60%)
Mydriasis and Biomicroscopy	48	15 (31%)	33 (69%)
All three techniques	8	5 (63%)	3 (38%)
Total	134	51 (38%)	83 (62%)

Table 7.1. Outcome of referral refinement.

7.3a Response to the NICE guidelines

As the NICE guidelines were published on the 22nd of April 2009, the results for pre (1st February to 22nd April) and post publication (23rd April to 31st July) are compared in Table 7.2.

	Pre NICE (1/2/09-22/4/09)			**Post NICE** (23/4/09-31/7/09)		
	N	Referred	Not Referred	N	Referred	Not Referred
Tonometry and Visual fields	25	12 (48%)	13 (52%)	53	19 (36%)	34 (64%)
Mydriasis and Biomicroscopy	24	7 (29%)	17 (71%)	24	8 (33%)	16 (67%)
All three techniques	3	2 (67%)	1 (33%)	5	3 (60%)	2 (40%)
Total	52	21 (40%)	31 (60%)	82	30 (37%)	52 (63%)

Table 7.2. Outcome of referral refinement. Pre and post NICE glaucoma guideline publication.

7.3b Patient Satisfaction

Of the 134 patients seen, 48 completed questionnaires have been received. Figures 7.2 to 7.4 show patient agreement with the statements "I am happy having these procedures performed at my opticians rather than at the hospital", "I am happy with the time I waited between appointments" and "I am satisfied overall with the service I have received" respectively. The time spent at the referral refinement appointment was reported by 43 patients, with a mean of 34.2 minutes (SD±14.1).

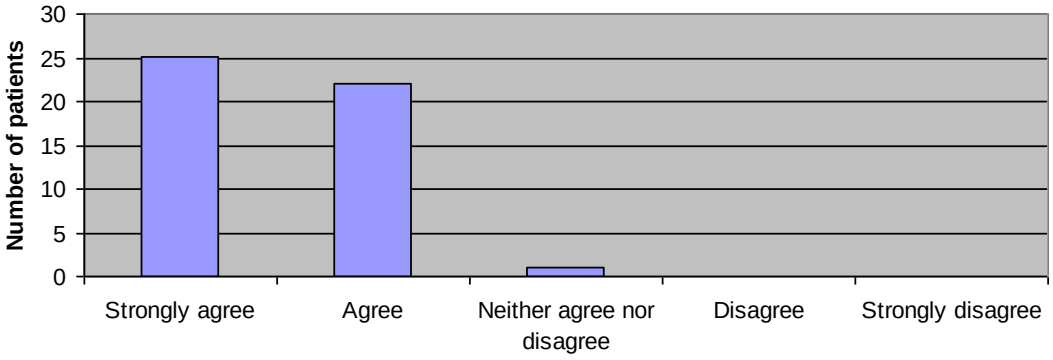

Figure 7.2. Agreement with the statement **"I am happy having these procedures performed at my opticians rather than at the hospital.**"

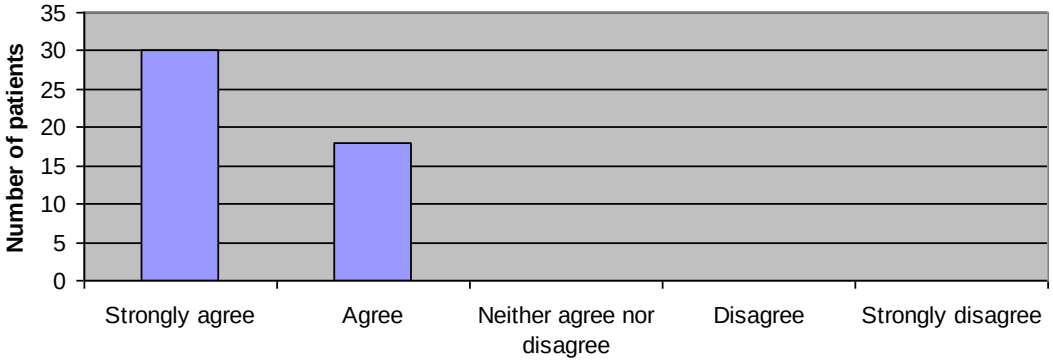

Figure 7.3. Agreement with the statement **"I am happy with the time I waited between appointments.**"

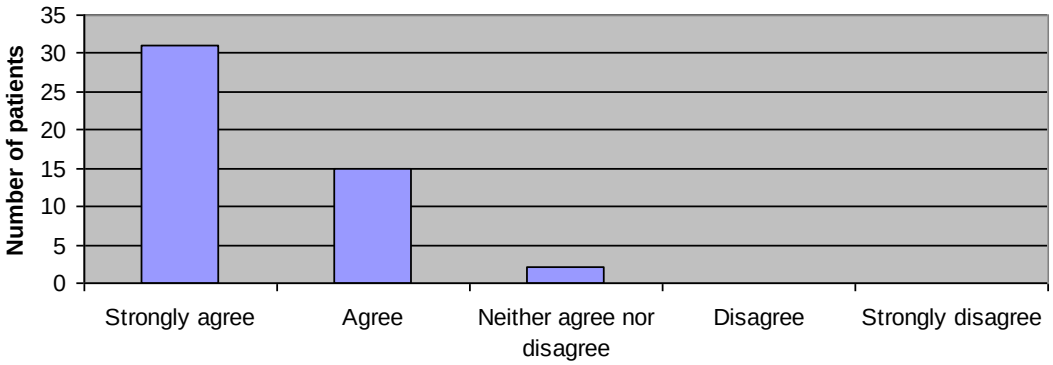

Figure 7.4. Agreement with the statement **"I am satisfied overall with the service I have received."**

7.3c Economic evaluation

The results of the analysis of cost-effectiveness of the study is provided in Table 7.3.

Costs without Service		**Costs with Service**	
Cost of HES 1st appointment	134 x £110 = £14,740	Cost of Ref. Ref. appointments	134 x £42 = £5628
(Cost of HES follow up)	(134 x 50% x 1.1 x £53 = £3,906)	Cost of HES appointments (51 x £110) + (51 x 50% x 1.1 x £53)	£7097
Total Cost	£18, 646 (minimum)	Total cost	£12,724
Saving made by service (20% sample of providers) over 6 months			£5,922
Potential annual saving if service is offered by all providers in PCT			£59, 220

Table 7.3. Approximate cost saving made by the Referral Refinement study.

7.4 Discussion

The results from this study suggested that if an enhanced service is in place for optometrists to recall patients and perform certain techniques within their competency, then the proportion of unnecessary referrals can be significantly reduced. Of the 134 patients that were going to be referred but satisfied the inclusion criteria of the referral refinement protocol, 62% turned out to not require referral after further investigation. A modest ongoing saving was made by this enhanced service as indicated by Table 7.3. Therefore a conservative estimate of about £6K was saved during the six months of this study, which extrapolates to an annual PCT saving of approximately £60K. Although reducing costs was not a primary aim of the service, it is an additional benefit. Given the current economic climate, even small savings are of importance, particularly if this or similar services could be rolled out over the rest of the country. The 62% figure of saved referrals is within the range of previous findings (35%, Bourne et al., 2009; 76% Parkins & Edgar, 2011), although the previous studies only assessed the role of referral refinement of glaucoma suspects. As the NICE guidelines for glaucoma were published on the 22nd of April, the results from post-publication may be a more accurate reflection of the potential success of a referral refinement service in the current situation. Although the limited numbers mean that any subset analysis may be inaccurate, when data from before and after the publication of the NICE guidelines are compared, some interesting differences are present. Table 7.2 illustrates that the proportion of patients requiring refinement appointments on the basis of tonometry and visual fields increased by 17% after the guidelines (from 48%

to 65%). In addition, of these patients eligible for refinement by tonometry and visual fields, a greater proportion did not require subsequent referral, 64% after the guidelines compared to 52% before (Table 7.2). This trend appears to be specific to tonometry and visual fields, and therefore must be attributable to the NICE guidelines, as the proportion of patients not referred after mydriasis and biomicroscopy over the same timescale were similar at 71% and 67%. This suggests that when the NICE guidelines were introduced, and the referral refinement protocol was modified in response, more patients were eligible on grounds that IOPs were greater than 21mmHg. As the proportion that were not referred also increased between the two timescales, it is possible that a significant proportion of the patients initially with IOPs over 21mmHg, turned out to have IOPs of 21mmHg or under after repeat measurement with contact tonometry, which is in agreement with the literature (Salmon et al., 2007, Sheldrick et al., 1994, Bell and O'Brien, 1997a, Henson et al., 2003). The results agree with recently published data from Moorfields Eye Hospital (Murdoch & Shah, 2011). Since the NICE guidelines for glaucoma were issued, ophthalmology departments in England have struggled to cope with the increased numbers of new outpatients (Naru and Green, 2009). The referral refinement enhanced service is one way in which to lower numbers of false positive referrals generated by the NICE guidelines, and reduce the burden on ophthalmology departments.

This is the first study that has assessed referral refinement with mydriasis and binocular indirect ophthalmoscopy. The results are limited by the relatively small sample size (N=48) and the lack of an assessment of false negatives, but show a high reduction in potential referrals of 69%. These findings strongly support the value of

a dilated fundus examination in certain cases (Sigel et al., 1990; Parisi et al., 1996). In many cases, the reason for the need for mydriasis was not reported, but some cases indicated the presence of suspicious naevi and posterior vitreous detachment (PVD)/ retinal detachment. The assessment of PVDs within a GOS or referral refinement process is complicated: Should a dilated fundus examination be performed in a patient with flashes and floaters and possible PVD within the GOS sight test as it could be deemed 'clinically necessary'? Or, could such symptoms warrant referral without a dilated fundus examination within the GOS sight test, so that a dilated fundus examination could be performed as a referral refinement? There is a clear need for this to be explicitly defined.

All 48 patients who completed a questionnaire were happy with the time between appointments. 47 of the 48 patients were happy having the procedures performed at their opticians and 46 of the patients were satisfied overall with the service received. It appears therefore that this limited cohort was happy with all aspects of this enhanced service, and it will be interesting to see if this trend is maintained if the service is extended. This referral refinement enhanced service should also have a positive impact on patient psychological well being (see Chapter 6) as it has reduced levels of inappropriate referrals, and moved care closer to home.

7.5 Limitations of the study

The main limitation of this study is the lack of assessment of false negative referrals. These are patients that are not referred through the referral refinement process,

but should have been according to the opinion of the gold standard assessment, that of the consultant ophthalmologist. Ideally, a sample of the patients not referred through the service should have been assessed by a consultant ophthalmologist with a interest in glaucoma and medical retina (Bourne et al., 2009). Other studies have used an assessment of a sample of patients not referred using clinical notes and disc photographs (Devarajan et al., 2011). However, neither of these approaches was possible in the present study due to financial constraints. More minor limitations of the study include that only 20% of the eligible optometry practices in the PCT area could take part in the study due to cost restrictions imposed by the PCT. However, the choice of which practices were included was randomised, so that the results should not have been in any way biased. The sample size, particularly if the data are separated into glaucoma suspects (N=78) and suspicious fundus anomalies (N=48) is relatively small (e.g., N=512, Bourne et al., 2011), but similar to some studies (e.g., N=100, Devarajan et al., 2011). One criticism of the patient satisfaction questionnaire is that the items are slightly leading because agreement means endorsement of the service for all three statements. If the service is to continue then some of the items could be rephrased to reduce the likelihood of bias, for example “I would prefer to have these procedures performed at the hospital” instead of “I am happy having these procedures performed at my opticians rather than at the hospital”.

7.6 Conclusions

In conclusion, this enhanced service is easy to implement and appears to be achieving its aims within the constraints of the study methodology. As the dataset in the study is limited in size, if the service is offered to all providers in the PCT then ongoing audit is advisable to see if its success is maintained. In particular, an assessment of false negative referrals is necessary, as any decrease in false positives needs to not result in an increase in false negatives. In addition, separating the IOP and visual field measurements in referral refinement appointments would be useful going forward. In the present study, patients attending referral refinement clinics with high IOP but full visual fields by screening are required to have a threshold visual field test and contact tonometry at a referral refinement appointment. As their fields have been screened negative, it should not be necessary to further investigate fields unless their IOP proves to be high after repetition by contact tonometry. The change could potentially be more cost effective. Finally, there is no clinical reason why the referral refinement appointment has to be on a separate day, and this should be considered if the service is extended to all providers. As the vast majority of initial IOP measurements are performed using non-contact tonometry (Myint et al., 2011, Willis et al., 2000), then contact tonometry could be performed on the same day. This implies that there is potentially some inherent proneness to inaccuracy of non-contact tonometers, or lack of skill of those that use them. However, if a patient requires repeat visual fields, then performing it on a separate occasion will reduce the effect of fatigue, which can be a significant factor in visual field unreliability (Hudson et al., 1994). Similarly, if contact tonometry was

performed initially, then the repeat contact tonometry should be performed at a later date (Salmon et al., 2007, Sheldrick et al., 1994, Bell and O'Brien, 1997a, Henson et al., 2003).

8. Overall Conclusions

The accuracy of referrals to the HES appears to improve as clinicians become more experienced, however the proportion of false negative decisions has not yet been evaluated. Greater numbers of false positive referrals are also generated by female clinicians. It is imperative that lower numbers of false positive referrals should not increase numbers of false negative decisions, and more work is required in this area. Optometrists refer patients with a wide range of ocular diseases and in most cases include both fundus observations and visual acuity measurements in their referrals. GPs mainly refer patients with anterior segment disorders, particularly lid lesions, based on direct observation and symptoms. GP referral letters include all relevant non-clinical information and are all perfectly legible, whereas illegibility and missing clinical information remains a problem in optometric referrals.

Patients referred for further investigation in ophthalmology departments experience raised levels of state anxiety as measured by the State-Trait Anxiety Inventory and raised levels of depression as measured by the HADS-D scale although the clinical significance of this effect is not fully known. As a method of assessing psychological distress, the questionnaires HADS-T, STAI-S and STAI-T show good discrimination between patients when administered to a population of new ophthalmic outpatients, despite having a floor effect. All three scales are not unidimensional, but split into well-established and logical subscales with PCA. The HADS Anxiety and Depression subscales are not recommended in this population due to inadequate participant discrimination.

The present study agrees with the National Eye Care Steering Group (2004b) that direct referral methods and enhanced services should be encouraged. The referral refinement service described in chapter 7 proved to be a cost effective way of reducing false positive referrals and improving referral quality.

References:

1989. Opticians Act.

1991. *Code of ethics and guidelines for professional conduct* [Online]. British College of Optometrists. Available: http://college-optometrists.org/wccoo/filemanager/root/site_assets/guidelines_archive/1991.zip

2000. *Action on Cataracts. Good Practice Guidance* [Online]. London: Department of Health. Available: http://www.dh.gov.uk/asset.Root/04/01/34/49/04013449.pdf.

2004a. AOP Guidance on Direct Referral of Cataract. Association of Optometrists.

2004b. *National Eye Care Steering Group First Report* [Online]. Available: http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/documents/digitalasset/dh_4080999.pdf

2005. *Framework for referrals* [Online]. College of Optometrists.

2006. *Free Sight Tests For All* [Online]. Optometry Scotland. Available: http://www.optometryscotland.org.uk/includes/download.php?id=34 [Accessed 30th March 2006].

2007. *Members Handbook* [Online]. College of Optometrists. Available: http://www.college-optometrists.org/coo/download.cfm?uuid=5D174FF5-1079-D4FC-897B9AD125A6A399&type=ethics_guidelines.

2008. *Optics at a Glance* [Online]. Federation of Ophthalmic and Dispensing Opticians. Available: http://www.fodo.com/fckextras/File/OAAG%202008,%20FINAL%20FINAL.pdf [Accessed June 2010 2008].

2009. Glaucoma: diagnosis and management of chronic open angle glaucoma and ocular hypertension. National Institute for Health and Clinical Excellence.

ADERIBIGBE, Y. A. & GUREJE, O. 1992. The Validity of the 28-item General Health Questionnaire in a Nigerian antenatal clinic. *Social Psychiatry and Psychiatric Epidemiology,* 27**,** 280-283.

ANDERSEN, H. S., SESTOFT, D., LILLEBAEK, T., GABRIELSEN, G. & HEMMINGSEN, R. 2002. Validity of the General Health Questionnaire (GHQ-28) in a prison population: data from a randomized sample of prisoners on remand. *Int J Law Psychiatry,* 25**,** 573-80.

ANDERSSON, E. 1993. The hospital anxiety and depression scale - homogeneity of the subscales. *Social Behavior and Personality,* 21**,** 197-204.

ANDERSSON, G., KALDO-SANDSTROM, V., STROM, L. & STROMGREN, T. 2003. Internet administration of the Hospital Anxiety and Depression Scale in a sample of tinnitus patients. *Journal of Psychosomatic Ressearch,* 55**,** 259-62.

AOP. 2009. *Advice on NICE glaucoma guidelines* [Online]. Association of Optometrists. Available: http://www.aop.org.uk/uploaded_files/glaucoma_advice_and_faqs_24.04.09.pdf [Accessed 24th April 2009].

ARMSTRONG, R. A., DAVIES, L. N., DUNNE, M. C. & GILMARTIN, B. 2011. Statistical guidelines for clinical studies of human vision. *Ophthalmic and Physiological Optics,* 31**,** 123-36.

AROUNI, A. & RICH, E. 2003. Physician gender and patient care. *J Gend Specif Med,* 6**,** 24-30.

AUSTEN, D. P. 2003. Primary eyecare in the community: GP ophthalmic referrals to optometrists. *Quality in Primary Care,* 11**,** 293-304.

AZUARA-BLANCO, A., BURR, J., THOMAS, R., MACLENNAN, G. & MCPHERSON, S. 2007. The accuracy of accredited glaucoma optometrists in the diagnosis and treatment recommendation for glaucoma. *British Journal of Ophthalmology,* 91**,** 1639-43.

BAMPTON, P. & DRAPER, B. 1997. Effect of relaxation music on patient tolerance of gastrointestinal endoscopic procedures. *J Clin Gastroenterol,* 25**,** 343-5.

BANES, M. J., CULHAM, L. E., BUNCE, C., XING, W., VISWANATHAN, A. & GARWAY-HEATH, D. 2006. Agreement between optometrists and ophthalmologists on clinical management decisions for patients with glaucoma. *British Journal of Ophthalmology,* 90**,** 579-85.

BANES, M. J., CULHAM, L. E., CROWSTON, J. G., BUNCE, C. & KHAW, P. T. 2000. An optometrist's role of co-management in a hospital glaucoma clinic. *Ophthalmic and Physiological Optics,* 20**,** 351-9.

BANKS, M. H. 1983. Validation of the General Health Questionnaire in a young community sample. *Psychological Medicine,* 13**,** 349-53.

BARNES, L. L. B., HARP, D. & JUNG, W. S. 2002. Reliability generalization of scores on the Spielberger state-trait anxiety inventory. *Educational and Psychological Measurement,* 62**,** 603-618.

BELL, R. W. D. & O'BRIEN, C. 1997a. Accuracy of referral to a glaucoma clinic. *Ophthalmic and Physiological Optics,* 17**,** 7-11.

BELL, R. W. D. & O'BRIEN, C. 1997b. The diagnostic outcome of new glaucoma referrals. *Ophthalmic and Physiological Optics,* 17**,** 3-6.

BENJAMIN, S., LENNON, S. & GARDNER, G. 1991. The validity of the General Health Questionnaire for 1st-stage screening for mental-illness in pain clinic patients. *Pain,* 47**,** 197-202.

BERTAKIS, K., HELMS, L., CALLAHAN, E., AZARI, R. & ROBBINS, J. 1995. The influence of gender on physician practice style. *Med Care,* 33**,** 407-16.

BIELING, P. J., ANTONY, M. M. & SWINSON, R. P. 1998. The State-Trait Anxiety Inventory, Trait version: structure and content re-examined. *Behaviour Research and Therapy,* 36**,** 777-88.

BJELLAND, I., DAHL, A. A., HAUG, T. T. & NECKELMANN, D. 2002. The validity of the Hospital Anxiety and Depression Scale - An updated literature review. *Journal of Psychosomatic Research,* 52**,** 69-77.

BJERTNAES, O. A., GARRATT, A. & BOTTEN, G. 2008. Nonresponse bias and cost-effectiveness in a Norwegian survey of family physicians. *Eval Health Prof,* 31**,** 65-80.

BOSANQUET, N. 2006. Developing a new partnership contract for community eye care in England. London: Imperial College.

BOULIS, A. & LONG, J. 2004. Gender differences in the practice of adult primary care physicians. *J Womens Health (Larchmt),* 13**,** 703-12.

BOURNE, R. R., FRENCH, K. A., CHANG, L., BORMAN, A. D., HINGORANI, M. & NEWSOM, W. D. 2009. Can a community optometrist-based referral refinement scheme reduce false-positive glaucoma hospital referrals without compromising quality of care? The community and hospital allied network glaucoma evaluation scheme (CHANGES). *Eye,* 24**,** 881-7.

BOWLING, B., CHEN, S. D. M. & SALMON, J. F. 2005. Outcomes of referrals by community optometrists to a hospital glaucoma service. *British Journal of Ophthalmology,* 89**,** 1102-1104.

BRANDBERG, Y., BOLUND, C., SIGURDARDOTTIR, V., SJODÉN, P. & SULLIVAN, M. 1992. Anxiety and depressive symptoms at different stages of malignant melanoma. *Psycho-Oncology,* 1**,** 71-78.

BRETT, J., BANKHEAD, C., HENDERSON, B., WATSON, E. & AUSTOKER, J. 2005. The psychological impact of mammographic screening. A systematic review. *Psycho-Oncology,* 14**,** 917-938.

BREWER, N. T., SALZ, T. & LILLIE, S. E. 2007. Systematic review: The long-term effects of false-positive mammograms. *Ann Intern Med,* 146**,** 502-510.

BRIDGES, K. W. & GOLDBERG, D. P. 1986. The validation of the GHQ-28 and the use of the MMSE in neurological inpatients. *British Journal of Psychiatry,* 148**,** 548-553.
BRITTAIN, G. 1988. A prospective study to determine sources and diagnostic accuracy of glaucoma referrals. *Health Trends,* 20**,** 43-4.
BROGGER, J., NYSTAD, W., CAPPELEN, I. & BAKKE, P. 2007. No increase in response rate by adding a web response option to a postal population survey: a randomized trial. *J Med Internet Res,* 9**,** e40.
BROUWERS, E. P., VAN BAAR, A. L. & POP, V. J. 2001. Does the Edinburgh Postnatal Depression Scale measure anxiety? *J Psychosom Res,* 51**,** 659-63.
BROWN, M. L., HOUN, F., SICKLES, E. A. & KESSLER, L. G. 1995. Screening mammography in community practice: positive predictive value of abnormal findings and yield of follow-up diagnostic procedures. *AJR Am J Roentgenol,* 165**,** 1373-7.
BUCHANAN, T. 2003. Internet-based questionnaire assessment: appropriate use in clinical contexts. *Cogn Behav Ther,* 32**,** 100-9.
BURR, J. M., MOWATT, G., HERNANDEZ, R., SIDDIQUI, M. A., COOK, J., LOURENCO, T., RAMSAY, C., VALE, L., FRASER, C., AZUARA-BLANCO, A., DEEKS, J., CAIRNS, J., WORMALD, R., MCPHERSON, S., RABINDRANATH, K. & GRANT, A. 2007. The clinical effectiveness and cost-effectiveness of screening for open angle glaucoma: a systematic review and economic evaluation. *Health Technol Assess,* 11**,** 1-190.
CACI, H., BAYLE, F. J., DOSSIOS, C., ROBERT, P. & BOYER, P. 2003a. The Spielberger Trait Anxiety Inventory measures more than anxiety. *Eur Psychiatry,* 18**,** 394-400.
CACI, H., BAYLE, F. J., MATTEI, V., DOSSIOS, C., ROBERT, P. & BOYER, P. 2003b. How does the Hospital and Anxiety and Depression Scale measure anxiety and depression in healthy subjects? *Psychiatry Research,* 118**,** 89-99.
CARLBRING, P., BRUNT, S., BOHMAN, S., AUSTIN, D., RICHARDS, J., OST, L. G. & ANDERSSON, G. 2007. Internet vs. paper and pencil administration of questionnaires commonly used in panic/agoraphobia research. *Computers in Human Behavior,* 23**,** 1421-1434.
CARTWRIGHT, A. 1986. Who responds to postal questionnaires? *Journal of Epidemiology and Community Health,* 40**,** 267-73.
CHITTY, A. 1997. *The eye examination - a test of quality?* [Online]. RNIB. Available: http://www.rnib.org.uk/xpedio/groups/public/documents/visugate/public_eyexam.hcsp.
CHLAN, L., SAVIK, K. & WEINERT, C. 2003. Development of a shortened state anxiety scale from the Spielberger State-Trait Anxiety Inventory (STAI) for patients receiving mechanical ventilatory support. *J Nurs Meas,* 11**,** 283-93.
CLARKE, A. 2008. Writing effective referrals. *Optician,* 236**,** 12-13.
CLEARKIN, L. & HARCOURT, B. 1983. Referral pattern of true and suspected glaucoma to an adult ophthalmic outpatient clinic. *Transactions of the Ophthalmological Societies of the United Kingdom,* 103**,** 284-287.
CLEEMPUT, I., KESTELOOT, K., MOONS, P., VANRENTERGHEM, Y., VAN HOOFF, J. P., SQUIFFLET, J. P. & DE GEEST, S. 2004. The construct and concurrent validity of the EQ-5D in a renal transplant population. *Value Health,* 7**,** 499-509.
COCKBURN, D. M. 2006. How to make clinical decisions from statistics. *Clinical and Experimental Optometry,* 89**,** 176-183.
COCKBURN, J., STAPLES, M., HURLEY, S. F. & DE LUISE, T. 1994. Psychological consequences of screening mammography. *Journal of Medical Screening,* 1**,** 7-12.
COOKE, M., CHABOYER, W. & HIRATOS, M. A. 2005. Music and its effect on anxiety in short waiting periods: a critical appraisal. *J Clin Nurs,* 14**,** 145-55.
CORAH, N. L. 1988. Dental anxiety. Assessment, reduction and increasing patient satisfaction. *Dent Clin North Am,* 32**,** 779-90.

CORAH, N. L., O'SHEA, R. M., BISSELL, G. D., THINES, T. J. & MENDOLA, P. 1988. The dentist-patient relationship: perceived dentist behaviors that reduce patient anxiety and increase satisfaction. *J Am Dent Assoc,* 116**,** 73-6.

COURT, H., GREENLAND, K. & MARGRAIN, T. H. 2008. Evaluating patient anxiety levels during contact lens fitting. *Optom Vis Sci,* 85**,** 574-80.

COURT, H., GREENLAND, K. & MARGRAIN, T. H. 2009. Predicting state anxiety in optometric practice. *Optom Vis Sci,* 86**,** 1295-302.

CRICK, R. P. & TUCK, M. W. 1995. How can we improve the detection of glaucoma. *British Medical Journal,* 310**,** 546-547.

CURRAN, V., SOLBERG, S., MATHEWS, M., CHURCH, J., BUEHLER, S., WELLS, J. & LOPEZ, T. 2005. Prostate cancer screening attitudes and continuing education needs of primary care physicians. *J Cancer Educ,* 20**,** 162-6.

DAVEY, C. J., NARU, R., BONNER, E. & ELLIOTT, D. B. 2010. An Optometric Referral Refinement pilot study (Optometry Tomorrow 2010 conference abstract). *Ophthalmic and Physiological Optics,* In press.

DAYAN, M., FITT, A. & BOSANQUET, R. 1997. General practitioner registrars' opinions of general practice training in ophthalmology: a questionnaire survey in the northern region. *Eye (Lond),* 11 (Pt 4)**,** 496-9.

DEBORDE, A. S., BERTHOZ, S., PERDEREAU, F., GODART, N., CORCOS, M. & JEAMMET, P. 2004. [Validity of the BVAQ: a study in eating disorder patients and controls]. *Encephale,* 30**,** 464-73.

DEPARTMENT OF HEALTH 2001. *Reference guide to consent for examination or treatment*, Department of Health. [Online] Available: http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/documents/digitalasset/dh_103653.pdf

DEVARAJAN, N., WILLIAMS, G. S., HOPES, M., O'SULLIVAN, D. & JONES, D. 2011. The Carmarthenshire Glaucoma Referral Refinement Scheme, a safe and efficient screening service. *Eye (Lond),* 25**,** 43-9.

DINAH, C., SIM, D. & MENON, G. 2010. Ensuring the quality of cataract referrals. *Eye (Lond),* 24**,** 745; author reply 745.

DUNBAR, M., FORD, G., HUNT, K. & DER, G. 2000. A confirmatory factor analysis of the Hospital Anxiety and Depression scale: Comparing empirically and theoretically derived structures. *British Journal of Clinical Psychology,* 39**,** 79-94.

EASTERBROOK, J. A. 1982. Citation classic - The effect of emotion on cue utilisation and the organisation of behaviour. *Current Contents/Social & Behavioral Sciences***,** 20-20.

EDWARDS, P., ROBERTS, I., CLARKE, M., DIGUISEPPI, C., PRATAP, S., WENTZ, R. & KWAN, I. 2002. Increasing response rates to postal questionnaires: systematic review. *British Medical Journal,* 324**,** 1183-1185.

ENDLER, N. S., MACRODIMITRIS, S. D. & KOCOVSKI, N. L. 2003. Anxiety and depression: Congruent, separate, or both? *Journal of Applied Biobehavioral Research,* 8**,** 42-60.

EVANS, J. R., FLETCHER, A. E. & WORMALD, R. P. L. 2007. Depression and anxiety in visually impaired older people. *Ophthalmology,* 114**,** 283-288.

EVANS, M., ROBLING, M., RAPPORT, F. M., HOUSTON, H., KINNERSLEY, P. & WILKINSON, C. 2002. It doesn't cost anything just to ask, does it? The ethics of questionnaire-based research. *Journal of Medical Ethics,* 28**,** 41-44.

FEATHERSTONE, P., JAMES, C., HALL, M. & WILLIAMS, A. 1992. General practitioners' confidence in diagnosing and managing eye conditions: a survey in south Devon. *Br J Gen Pract,* 42**,** 21-4.

FLINT, A. J. & RIFAT, S. L. 2002. Factor structure of the Hospital Anxiety and Depression Scale in older patients with major depression. *International Journal of Geriatric Psychiatry,* 17**,** 117-123.

FORJAZ, M. J., RODRIGUEZ-BLÁZQUEZ, C., MARTINEZ-MARTIN, P. & GROUP, L. P. S. D. P. S. 2009. Rasch analysis of the hospital anxiety and depression scale in Parkinson's disease. *Mov Disord,* 24**,** 526-32.

FRANKS, P. & CLANCY, C. 1993. Physician gender bias in clinical decisionmaking: screening for cancer in primary care. *Med Care,* 31**,** 213-8.

FRIEDMAN, S., SAMUELIAN, J. C., LANCRENON, S., EVEN, C. & CHIARELLI, P. 2001. Three-dimensional structure of the Hospital Anxiety and Depression Scale in a large French primary care population suffering from major depression. *Psychiatry Research,* 104**,** 247-257.

GOLDBERG, C. B. 2003. Who responds to surveys? Assessing the effects of nonresponse in cross-sectional dyadic research. *Assessment,* 10**,** 41-8.

GOLDBERG, D. & BRIDGES, K. 1987. Screening for psychiatric-illness in General Practice – The General Practitioner versus the screening questionnaire. *Journal of the Royal College of General Practitioners,* 37**,** 15-18.

GOLDBERG, D. P. 1972. *The Detection of Psychiatric Illness by Questionnaire,* Oxford, Oxford University Press.

GOLDBERG, D. P., GATER, R., SARTORIUS, N., USTUN, T. B., PICCINELLI, M., GUREJE, O. & RUTTER, C. 1997. The validity of two versions of the GHQ in the WHO study of mental illness in general health care. *Psychological Medicine,* 27**,** 191-7.

GOLDBERG, D. P. & HILLIER, V. F. 1979. SCALED VERSION OF THE GENERAL HEALTH QUESTIONNAIRE. *Psychological Medicine,* 9**,** 139-145.

GOLDBERG, D. P. & WILLIAMS, P. 1988. *A Users Guide To The General Health Questionnaire,* London, nferNelson.

GOODMAN, J. S. & BLUM, T. C. 1996. Assessing the non-random sampling effects of subject attrition in longitudinal research. *Journal of Management,* 22**,** 627-652.

GOTHWAL, V. K., WRIGHT, T. A., ELLIOTT, D. B. & PESUDOVS, K. 2010. The refractive status and vision profile: Rasch analysis of subscale validity. *J Refract Surg,* 26**,** 912-5.

GOUGH, K. & HUDSON, P. 2009. Psychometric properties of the Hospital Anxiety and Depression Scale in family caregivers of palliative care patients. *J Pain Symptom Manage,* 37**,** 797-806.

GRAM, I. T., LUND, E. & SLENKER, S. E. 1990. Quality of life following a false positive mammogram. *British Journal of Cancer,* 62**,** 1018-22.

GRAY, S. F., SPENCER, I. C., SPRY, P. G., BROOKES, S. T., BAKER, I. A., PETERS, T. J., SPARROW, J. M. & EASTY, D. L. 1997. The Bristol Shared Care Glaucoma Study--validity of measurements and patient satisfaction. *J Public Health Med,* 19**,** 431-6.

GRIFFIN, D. H., FISCHER, D. P. & MORGAN, M. T. 2001. Testing an Internet Reponse Option for the American Community Survey. *American Association for Public Opinion Research 56th Annual Conference.* Montreal, Quebec.

GROS, D. F., ANTONY, M. M., SIMMS, L. J. & MCCABE, R. E. 2007. Psychometric properties of the State-Trait Inventory for Cognitive and Somatic Anxiety (STICSA): Comparison to the State-Trait Anxiety Inventory (STAI). *Psychological Assessment,* 19**,** 369-381.

GØTZSCHE, P. C. & NIELSEN, M. 2011. Screening for breast cancer with mammography. *Cochrane Database Syst Rev***,** CD001877.

HARRISON, R. J., WILD, J. M. & HOBLEY, A. J. 1988. Referral patterns to an Ophthalmic outpatient clinic by general practitioners and ophthalmic opticians and the role of these professionals in screeninf for ocular disease. *British Medical Journal,* 297**,** 1162-1167.

HAU, S., EHRLICH, D., BINSTEAD, K. & VERMA, S. 2007. An evaluation of optometrists' ability to correctly identify and manage patients with ocular disease in the accident and emergency department of an eye hospital. *British Journal of Ophthalmology,* 91**,** 437-40.

HENDERSON, A. S. & JORM, A. F. 1990. Do mental health surveys disturb. *Psychological Medicine,* 20**,** 721-724.
HENSON, D. B., SPENCER, A. F., HARPER, R. & CADMAN, E. J. 2003. Community refinement of glaucoma referrals. *Eye,* 17**,** 21-26.
HERRMANN, C. 1997. International experiences with the hospital anxiety and depression scale - A review of validation data and clinical results. *Journal of Psychosomatic Research,* 42**,** 17-41.
HISHINUMA, E. S., MIYAMOTO, R. H., NISHIMURA, S. T., NAHULU, L. B., ANDRADE, N. N., MAKINI, G. K., JR., YUEN, N. Y., JOHNSON, R. C., KIM, S. P., GOEBERT, D. A. & GUERRERO, A. P. 2000. Psychometric properties of the state-trait anxiety inventory for Asian/Pacific-islander adolescents. *Assessment,* 7**,** 17-36.
HO, S. & VERNON, S. A. 2011. Decision making in chronic glaucoma--optometrists vs ophthalmologists in a shared care service. *Ophthalmic and Physiological Optics,* 31**,** 168-73.
HOBLEY, A. J., WOODWARD, E. G. & PORT, M. J. A. 1992. Retrospective study of optometric referrals. *Ophthalmic and Physiological Optics,* 12**,** 395-399.
HODI, S. 2007. Management of ophthalmology referral letters in the United Kingdom: are traditional methods the best? *Ophthalmic and Physiological Optics,* 27**,** 394-8.
HOLM, L. & FITZMAURICE, L. 2008. Emergency department waiting room stress: can music or aromatherapy improve anxiety scores? *Pediatr Emerg Care,* 24**,** 836-8.
HUDSON, C., WILD, J. M. & O'NEILL, E. C. 1994. Fatigue effects during a single session of automated static threshold perimetry. *Investigative Ophthalmology & Visual Science,* 35**,** 268-80.
IWATA, N. & HIGUCHI, H. R. 2000. Responses of Japanese and American university students to the STAI items that assess the presence or absence of anxiety. *J Pers Assess,* 74**,** 48-62.
IWATA, N. & MISHIMA, N. 1999. Reliability of the State-Trait Anxiety Inventory, Form Y in Japanese samples. *Psychol Rep,* 84**,** 494-6.
IWATA, N., MISHIMA, N., SHIMIZU, T., MIZOUE, T., FUKUHARA, M., HIDANO, T. & SPIELBERGER, C. D. 1998. Positive and negative affect in the factor structure of the State-Trait Anxiety Inventory for Japanese workers. *Psychol Rep,* 82**,** 651-6.
IWATA, N. & SAITO, K. 1992. The factor structure of the 28-item General Health Questionnaire when used in Japanese early adolescents and adult employees – Age-cultural and cross-cultural comparisons. *European Archives of Psychiatry and Clinical Neuroscience,* 242**,** 172-178.
JACOMB, P. A., JORM, A. F., RODGERS, B., KORTEN, A. E., HENDERSON, A. S. & CHRISTENSEN, H. 1999. Emotional response of participants to a mental health survey. *Social Psychiatry and Psychiatric Epidemiology,* 34**,** 80-84.
JOHNSTON, M., POLLARD, B. & HENNESSEY, P. 2000. Construct validation of the hospital anxiety and depression scale with clinical populations. *Journal of Psychosomatic Research,* 48**,** 579-584.
JOMEEN, J. & MARTIN, C. R. 2004. Is the Hospital Anxiety and Depression Scale (HADS) a reliable screening tool in early pregnancy? *Psychology & Health,* 19**,** 787-800.
JORM, A. F., HENDERSON, A. S., SCOTT, R., MACKINNON, A. J., KORTEN, A. E. & CHRISTENSEN, H. 1994. Do mental health surveys disturb - further evidence. *Psychological Medicine,* 24**,** 233-237.
KABACOFF, R. I., SEGAL, D. L., HERSEN, M. & VAN HASSELT, V. B. 1997. Psychometric properties and diagnostic utility of the Beck Anxiety Inventory and the State-Trait Anxiety Inventory with older adult psychiatric outpatients. *J Anxiety Disord,* 11**,** 33-47.

KAIPPER, M. B., CHACHAMOVICH, E., HIDALGO, M. P., TORRES, I. L. & CAUMO, W. 2010. Evaluation of the structure of Brazilian State-Trait Anxiety Inventory using a Rasch psychometric approach. *J Psychosom Res,* 68**,** 223-33.

KAPLOWITZ, M. D., HADLOCK, T. D. & LEVINE, R. 2004. A comparison of Web and mail survey response rates. *Public Opinion Quarterly,* 68**,** 94-101.

KENNEDY, B. L., SCHWAB, J. J., MORRIS, R. L. & BELDIA, G. 2001. Assessment of state and trait anxiety in subjects with anxiety and depressive disorders. *Psychiatric Quarterly,* 72**,** 263-276.

KENT, G. 1984. *The Psychology of Dental Care,* Bristol, Wright and Sons.

KEYZER-DEKKER, C. M., DE VRIES, J., VAN ESCH, L., ERNST, M. F., NIEUWENHUIZEN, G. A., ROUKEMA, J. A. & VAN DER STEEG, A. F. 2011. Anxiety after an abnormal screening mammogram is a serious problem. *Breast*.

KLEIN, R., KLEIN, B. E., LINTON, K. L. & DE METS, D. L. 1991. The Beaver Dam Eye Study: visual acuity. *Ophthalmology,* 98**,** 1310-5.

KOIZUMI, N., FUJITA, D., NINOMIYA, R. & NAKAMOTO, N. 1998. [Screening test by statistically reducing the number of the State-Trait Anxiety Inventory (STAI) items]. *Sangyo Eiseigaku Zasshi,* 40**,** 107-12.

KREUTER, M., STRECHER, V., HARRIS, R., KOBRIN, S. & SKINNER, C. 1995. Are patients of women physicians screened more aggressively? A prospective study of physician gender and screening. *J Gen Intern Med,* 10**,** 119-25.

KVAAL, K., LAAKE, K. & ENGEDAL, K. 2001. Psychometric properties of the state part of the Spielberger State-Trait Anxiety Inventory (STAI) in geriatric patients. *International Journal of Geriatric Psychiatry,* 16**,** 980-986.

LAFATA, J. E., SIMPKINS, J., LAMERATO, L., POISSON, L., DIVINE, G. & JOHNSON, C. C. 2004. The economic impact of false-positive cancer screens. *Cancer Epidemiology Biomarkers & Prevention,* 13**,** 2126-2132.

LAIDLAW, D., BLOOM, P., HUGHES, A., SPARROW, J. & MARMION, V. 1994. The sight test fee: effect on ophthalmology referrals and rate of glaucoma detection. *BMJ,* 309**,** 634-6.

LAMBERT, S., PALLANT, J. F. & GIRGIS, A. 2010. Rasch analysis of the Hospital Anxiety and Depression Scale among caregivers of cancer survivors: implications for its use in psycho-oncology. *Psychooncology*.

LASH, S. C. 2003. Assessment of information included on the GOS 18 referral form used by optometrists. *Ophthalmic and Physiological Optics,* 23**,** 21-23.

LASH, S. C., PRENDIVILLE, C. P., SAMSON, A., LEWIS, K., MUNNEKE, R. & PARKIN, B. T. 2006. Optometrist referrals for cataract and "Action on Cataracts" guidelines: are optometrists following them and are they effective? *Ophthalmic and Physiological Optics,* 26**,** 464-7.

LAVERY, J. R., GIBSON, J. M., SHAW, D. E. & ROSENTHAL, A. R. 1988. Vision and visual acuity in an elderly population. *Ophthalmic and Physiological Optics,* 8**,** 390-3.

LERMAN, C., TROCK, B., RIMER, B. K., JEPSON, C., BRODY, D. & BOYCE, A. 1991. PSYCHOLOGICAL SIDE-EFFECTS OF BREAST-CANCER SCREENING. *Health Psychology,* 10**,** 259-267.

LEUNG, C. M., HO, S., KAN, C. S., HUNG, C. H. & CHEN, C. N. 1993. Evaluation of the Chinese version of the Hospital Anxiety and Depression Scale. A cross-cultural perspective. *Int J Psychosom,* 40**,** 29-34.

LEWIS, G. 1991. Observer bias in the assessment of anxiety and depression. *Social Psychiatry and Psychiatric Epidemiology,* 26**,** 265-272.

LIM, L. C. C. & CHEW, S. J. 1991. Valiadation of the General Health Questionnaire in female video display unit operators in Singapore. *SMJ Singapore Medical Journal,* 32**,** 143-145.

LINACRE, J. M. 2003. Rasch Power Analysis: Size vs. Significance: Standardized Chi-Square Fit Statistic. *Rasch Measurement Transactions,* 17**,** 918.
LINACRE, J. M. 2008. *RE: Personal communication.* Type to VIANYA-ESTOPA, M.
LINACRE, J. M. 2009. A User's Guide to Winsteps. Chaicago, IL: Winsteps.com.
LLEWELLYN, R. 2008. *RE: Eyecare Fee Consultancy.*
LOBO, A., PEREZECHEVERRIA, M. J. & ARTAL, J. 1986. Validity of the Scaled version of the General Health Questionnaire (GHQ-28) in a Spanish population. *Psychological Medicine,* 16**,** 135-140.
LURIE, N., MARGOLIS, K., MCGOVERN, P., MINK, P. & SLATER, J. 1997. Why do patients of female physicians have higher rates of breast and cervical cancer screening? *J Gen Intern Med,* 12**,** 34-43.
LURIE, N., SLATER, J., MCGOVERN, P., EKSTRUM, J., QUAM, L. & MARGOLIS, K. 1993. Preventive care for women. Does the sex of the physician matter? *N Engl J Med,* 329**,** 478-82.
MAKOWSKA, Z., MERECZ, D., MOSCICKA, A. & KOLASA, W. 2002. The validity of general health questionnaires, GHQ-12 and GHQ-28, in mental health studies of working people. *Int J Occup Med Environ Health,* 15**,** 353-62.
MALAKOUTI, S. K., FATOLLAHI, P., MIRABZADEH, A. & ZANDI, T. 2007. Reliability, validity and factor structure of the GHQ-28 used among elderly Iranians. *International Psychogeriatrics,* 19**,** 623-634.
MANN, A. H., WAKELING, A., WOOD, K., MONCK, E., DOBBS, R. & SZMUKLER, G. 1983. Screening for abnormal eating attitudes and psychiatric morbidity in an unselected population of 15-year-old schoolgirls. *Psychological Medicine,* 13**,** 573-580.
MARELLA, M., PESUDOVS, K., KEEFFE, J. E., O'CONNOR, P. M., REES, G. & LAMOUREUX, E. L. 2010. The psychometric validity of the NEI VFQ-25 for use in a low-vision population. *Invest Ophthalmol Vis Sci,* 51**,** 2878-84.
MARGRAIN, T. H., GREENLAND, K. & ANDERSON, J. 2003. Evaluating anxiety in patients attending optometric practice. *Ophthalmic and Physiological Optics,* 23**,** 287-293.
MARTEAU, T. M. & BEKKER, H. 1992. The development of a six-item short-form of the state scale of the Spielberger State-Trait Anxiety Inventory (STAI). *British Journal of Clinical Psychology,* 31 (Pt 3)**,** 301-6.
MARTEAU, T. M., KIDD, J., COOK, R., JOHNSTON, M., MICHIE, S., SHAW, R. W. & SLACK, J. 1988. Screening for Downs-syndrome. *British Medical Journal,* 297**,** 1469-1469.
MARTIN, C. R. & THOMPSON, D. R. 1999. Utility of the Hospital Anxiety and Depression Scale in patients with end-stage renal disease on continuous ambulatory peritoneal dialysis. *Psychol Health Med,* 4**,** 369-76.
MARTIN, C. R. & THOMPSON, D. R. 2000. A psychometric evaluation of the Hospital Anxiety and Depression Scale in coronary care patients following acute myocardial infarction. *Psychol Health Med,* 5**,** 193-201.
MARTIN, C. R., TWEED, A. E. & METCALFE, M. S. 2004. A psychometric evaluation of the Hospital Anxiety and Depression Scale in patients diagnosed with end-stage renal disease. *British Journal of Clinical Psychology,* 43**,** 51-64.
MAYBERRY, J. F. 2002. The cost of questionnaire based research. *Journal of Epidemiology and Community Health,* 56**,** 956-957.
MCFARLANE, E., OLMSTED, M. G., MURPHY, J. & HILL, C. A. 2007. Nonresponse bias in a mail survey of physicians. *Eval Health Prof,* 30**,** 170-85.
MEDINA-MORA, M. E., PADILLA, G. P., CAMPILLOSERRANO, C., MAS, C. C., EZBAN, M., CARAVEO, J. & CORONA, J. 1983. The factor structure of the GHQ – A scaled version for a hospitals general practice service in Mexico. *Psychological Medicine,* 13**,** 355-361.
MENON, G., FARIDI, U. & GRAY, R. 2004. Direct referral of posterior capsular opacification by optometrists. *Ophthalmic and Physiological Optics,* 24**,** 106-10.

MEYER, M., KEITH-ROKOSH, J., REDDY, H., MEGYESI, J. & HAMMOND, R. R. 2010. Sources of error in neuropathology intraoperative diagnosis. *Can J Neurol Sci,* 37**,** 620-4.

MICHOPOULOS, I., DOUZENIS, A., KALKAVOURA, C., CHRISTODOULOU, C., MICHALOPOULOU, P., KALEMI, G., FINETI, K., PATAPIS, P., PROTOPAPAS, K. & LYKOURAS, L. 2008. Hospital Anxiety and Depression Scale (HADS): validation in a Greek general hospital sample. *Ann Gen Psychiatry,* 7**,** 4.

MONDOLO, F., JAHANSHAHI, M., GRANA, A., BIASUTTI, E., CACCIATORI, E. & DI BENEDETTO, P. 2007. Evaluation of anxiety in Parkinson's disease with some commonly used rating scales. *Neurol Sci,* 28**,** 270-5.

MOOK, J., KLEIJN, W. C. & VANDERPLOEG, H. M. 1991. Symptom positively and symptom negatively worded items in 2 popular self-report inventories of anxiety and depression. *Psychol Rep,* 69**,** 551-560.

MOOK, J., VANDERPLOEG, H. M. & KLEIJN, W. C. 1992. SYMPTOM-POSITIVE AND SYMPTOM-NEGATIVE ITEMS IN THE STATE-TRAIT ANXIETY INVENTORY - A COMPARISON AND REPLICATION. *Anxiety Stress and Coping,* 5**,** 113-123.

MORTON, C. A. & MACKIE, R. M. 1998. Clinical accuracy of the diagnosis of cutaneous malignant melanoma. *Br J Dermatol,* 138**,** 283-7.

MOSS, S. M., BLANKS, R. G. & BENNETT, R. L. 2005. Is radiologists' volume of mammography reading related to accuracy? A critical review of the literature. *Clin Radiol,* 60**,** 623-6.

MYINT, J., EDGAR, D. F., KOTECHA, A., MURDOCH, I. E. & LAWRENSON, J. G. 2011. A national survey of diagnostic tests reported by UK community optometrists for the detection of chronic open angle glaucoma. *Ophthalmic and Physiological Optics,* 31**,** 353-9.

MYKLETUN, A., STORDAL, E. & DAHL, A. A. 2001. Hospital Anxiety and Depression (HAD) scale: factor structure, item analyses and internal consistency in a large population. *British Journal of Psychiatry,* 179**,** 540-544.

NARU, R. & GREEN, C. 2009. Optometric referrals - An alternative approach. *Optician,* 237**,** 36-37.

NARU, R. & GREEN, C. R. 2010. *Bradford and Airedale GOS18 - An electronic version* [Online]. NHS Bradford and Airedale. Available: http://www.bradfordloc.org.uk/22/Forms-for-Practice-Use.php [Accessed July 2010.

NEWCOMB, R. D., POTTER J. W. 1981. Part II. Diagnostic outcomes of optometric referrals in a VA medical center. *J Am Optom Assoc*. 1981 Jul;52(7):599-604.

NEWMAN, D. K., ANWAR, S. & JORDAN, K. 1998. Glaucoma screening by optometrists: positive predictive value of visual field testing. *Eye,* 12**,** 921-924.

NEWSOM, W., GIBBONS, H., FLANAGAN, D. & HINGORANI, M. 2005. Direct referral of cataracts in Huntingdon. Sharing good practice *Optometry Today,* 22**,** 21-24.

NHS. 2006. *Sight Tests Volume and Workforce Survey 2005-06* [Online]. Available: http://www.ic.nhs.uk/webfiles/publications/sighttestvolandwforce/SightTestsVolumeAndWorkforceReport241006_PDF.pdf.

ORMEL, J., KOETER, M. W. J., VANDENBRINK, W. & GIEL, R. 1989. CONCURRENT VALIDITY OF GHQ-28 AND PSE AS MEASURES OF CHANGE. *Psychological Medicine,* 19**,** 1007-1013.

OSTER, J., CULHAM, L. E. & DANIEL, R. 1999. An extended role for the hospital optometrist. *Ophthalmic and Physiological Optics,* 19**,** 351-6.

PALLANT, J. F. & TENNANT, A. 2007. An introduction to the Rasch measurement model: An example using the Hospital Anxiety and Depression Scale (HADS). *British Journal of Clinical Psychology,* 46**,** 1-18.

PARISI, M. L., SCHEIMAN, M. & COULTER, R. S. 1996. Comparison of the effectiveness of a nondilated versus dilated fundus examination in the pediatric population. *Journal of the American Optometric Association,* 67**,** 266-72.

PARK, J., ROSS, A., TOLE, D., SPARROW, J., PENNY, J. & MUNDASAD, M. 2009. Evaluation of a new cataract surgery referral pathway. *Eye (Lond),* 23**,** 309-13.

PARKER, G. & HYETT, M. 2009. Management of depression by general practitioners: impact of physician gender. *Aust N Z J Psychiatry,* 43**,** 355-9.

PARKINS, D. J. & EDGAR, D. F. 2011. Comparison of the effectiveness of two enhanced glaucoma referral schemes. *Ophthalmic and Physiological Optics,* 31**,** 343-52.

PATEL, U. D. M., MURDOCH, I. E. & THEODOSSIADES, J. 2006. Glaucoma detection in the community: does ongoing training of optometrists have a lasting effect? *Eye,* 20**,** 591-594.

PEDUZZI P., CONCATO J., KEMPER E., HOLFORD T. R., FEINSTEIN A. R. 1996. A simulation study of the number of events per variable in logistic regression analysis. J Clin Epidemiol. Dec;49(12):1373-9.

PESUDOVS, K., BURR, J. M., HARLEY, C. & ELLIOTT, D. B. 2007. The development, assessment, and selection of questionnaires. *Optom Vis Sci,* 84**,** 663-74.

PIERSCIONEK, T. J., MOORE, J. E. & PIERSCIONEK, B. K. 2009. Referrals to ophthalmology: optometric and general practice comparison. *Ophthalmic and Physiological Optics,* 29**,** 32-40.

POOLE, N. A. & MORGAN, J. F. 2006. Validity and reliability of the Hospital Anxiety and Depression Scale in a hypertrophic cardiomyopathy clinic: the HADS in a cardiomyopathy population. *General Hospital Psychiatry,* 28**,** 55-58.

POOLEY, J. E. 1996. *Referrals from the Community to the Hospital Eye Service.* PhD thesis, City University.

POOLEY, J. E. & FROST, E. C. 1999. Optometrists' referrals to the Hospital Eye Service. *Ophthalmic and Physiological Optics,* 19**,** S16-S24.

PORT, M. J. A. 1989. Referrals and notifications by Optometrists within the UK - 1988 survey. *Ophthalmic and Physiological Optics,* 9**,** 31-35.

PORT, M. J. A. & POPE, C. A. 1988. Referrals and notifications by British optometrists. *Ophthalmic and Physiological Optics,* 8**,** 323-326.

RABINS, P. V. & BROOKS, B. R. 1981. EMOTIONAL DISTURBANCE IN MULTIPLE-SCLEROSIS PATIENTS - VALIDITY OF THE GENERAL HEALTH QUESTIONNAIRE (GHQ). *Psychological Medicine,* 11**,** 425-427.

RAZAVI, D., DELVAUX, N., FARVACQUES, C. & ROBAYE, E. 1989. VALIDATION OF THE FRENCH VERSION OF HADS IN A SAMPLING OF HOSPITALIZED CANCER-PATIENTS. *Revue De Psychologie Appliquee,* 39**,** 295-308.

RIAZ, H. & REZA, H. 1998. The evaluation of an Urdu version of the GHQ-28. *Acta Psychiatrica Scandinavica,* 97**,** 427-432.

ROMANS-CLARKSON, S. E., WALTON, V. A., HERBISON, G. P. & MULLEN, P. E. 1989. VALIDITY OF THE GHQ-28 IN NEW ZEALAND WOMEN. *Australian and New Zealand Journal of Psychiatry,* 23**,** 187-196.

SALMON, N. J., TERRY, H. P., FARMERY, A. D. & SALMON, J. F. 2007. An analysis of patients discharged from a hospital-based glaucoma case-finding clinic over a 3-year period. *Ophthalmic and Physiological Optics,* 27**,** 399-403.

SCHMITTDIEL, J., TRAYLOR, A., URATSU, C., MANGIONE, C., FERRARA, A. & SUBRAMANIAN, U. 2009. The association of patient-physician gender concordance with cardiovascular disease risk factor control and treatment in diabetes. *J Womens Health (Larchmt),* 18**,** 2065-70.

SCULLY, N. D., CHU, L., SIRIWARDENA, D., WORMALD, R. & KOTECHA, A. 2009. The quality of optometrists' referral letters for glaucoma. *Ophthalmic and Physiological Optics,* 29**,** 26-31.

SHARP, J., MORIARTY, A., WARBURTON, T. & DIXON, J. 2003. Cataract assessment and direct referral. Stockport optometrists take the initiative *Optometry Today,* 25**,** 16-17.

SHELDRICK, J., VERNON, S. & WILSON, A. 1992. Study of diagnostic accord between general practitioners and an ophthalmologist. *BMJ,* 304**,** 1096-8.

SHELDRICK, J., WILSON, A., VERNON, S. & SHELDRICK, C. 1993. Management of ophthalmic disease in general practice. *Br J Gen Pract,* 43**,** 459-62.

SHELDRICK, J. H., NG, C., AUSTIN, D. J. & R, R. A. 1994. An analysis of referral routes and diagnostic accuracy in cases of suspected glaucoma. *Ophthalmic Epidemiology,* 1**,** 31-9.

SHOKAR, N., NGUYEN-OGHALAI, T. & WU, H. 2009. Factors associated with a physician's recommendation for colorectal cancer screening in a diverse population. *Fam Med,* 41**,** 427-33.

SHUTTLEWORTH, G. & MARSH, G. 1997. How effective is undergraduate and postgraduate teaching in ophthalmology? *Eye (Lond),* 11 (Pt 5)**,** 744-50.

SIEGEL, B. S., THOMPSON, A. K., YOLTON, D. P., REINKE, A. R. & YOLTON, R. L. 1990. A comparison of diagnostic outcomes with and without pupillary dilatation. *Journal of the American Optometric Association,* 61**,** 25-34.

SIMON, E. P. 1999. Hypnosis using a communication device to increase magnetic resonance imaging tolerance with a claustrophobic patient. *Mil Med,* 164**,** 71-2.

SMEETH, L. & ILIFFE, S. 1998. Effectiveness of screening older people for impaired vision in community setting: systematic review of evidence from randomised controlled trials. *BMJ,* 316**,** 660-3.

SMITH, A. B., SELBY, P. J., VELIKOVA, G., STARK, D., WRIGHT, E. P., GOULD, A. & CULL, A. 2002. Factor analysis of the Hospital Anxiety and Depression Scale from a large cancer population. *Psychology and Psychotherapy-Theory Research and Practice,* 75**,** 165-176.

SMITH, A. B., WRIGHT, E. P., RUSH, R., STARK, D. P., VELIKOVA, G. & SELBY, P. J. 2006. Rasch analysis of the dimensional structure of the Hospital Anxiety and Depression Scale. *Psychooncology,* 15**,** 817-27.

SNAITH, R. P. 2003. The Hospital Anxiety And Depression Scale. *Health Qual Life Outcomes,* 1**,** 29.

SNAITH, R. P., BAUGH, S. J., CLAYDEN, A. D., HUSAIN, A. & SIPPLE, M. A. 1982. The Clinical Anxiety Scale: an instrument derived from the Hamilton Anxiety Scale. *British Journal of Psychiatry,* 141**,** 518-23.

SNAITH, R. P. & ZIGMOND, A. S. 1994. *HADS: Hospital Anxiety and Depression Scale,* Windsor, NFER Nelson.

SOURY, S., MAZAUX, J. M., LAMBERT, J., DE SEZE, M., JOSEPH, P. A., LOZES-BOUDILLON, S., MCCAULEY, S., VANIER, M. & LEVIN, H. S. 2005. [The neurobehavioral rating scale-revised: assessment of concurrent validity]. *Ann Readapt Med Phys,* 48**,** 61-70.

SPENCER, I. C., SPRY, P. G., GRAY, S. F., BAKER, I. A., MENAGE, M. J., EASTY, D. L. & SPARROW, J. M. 1995. The Bristol Shared Care Glaucoma Study: study design. *Ophthalmic and Physiological Optics,* 15**,** 391-4.

SPIELBERGER, C. D. 1983. *State-Trait Anxiety Inventory (Form Y) Manual,* Palo Alto, CA, Mind Garden.

SPIELBERGER, C. D. 1989. *State-Trait Anxiety Inventory: A comprehensive bibliography,* Palo Alto, CA, Consulting Psychologists Press.

SPIELBERGER, C. D., GORSUCH, R. L. & LUSHENE, R. D. 1970. *STAI: manual for the State-Trait Anxiety Inventory,* Palo Alto, CA, Consulting Psychologists Press.

SPRY, P. G., SPENCER, I. C., SPARROW, J. M., PETERS, T. J., BROOKES, S. T., GRAY, S., BAKER, I., FURBER, J. E. & EASTY, D. L. 1999. The Bristol Shared Care Glaucoma Study:

reliability of community optometric and hospital eye service test measures. *British Journal of Ophthalmology,* 83**,** 707-12.
STARK, D. P. & HOUSE, A. 2000. Anxiety in cancer patients. *Br J Cancer,* 83**,** 1261-7.
STEWART-BROWN, S. & FARMER, A. 1997. Screening could seriously damage your health - Decisions to screen must take account of the social and psychological costs. *British Medical Journal,* 314**,** 533-534.
TANG, W. K., WONG, E., CHIU, H. F., LUM, C. M. & UNGVARI, G. S. 2008. Examining item bias in the anxiety subscale of the Hospital Anxiety and Depression Scale in patients with chronic obstructive pulmonary disease. *Int J Methods Psychiatr Res,* 17**,** 104-10.
TANG, W. K., WONG, E., CHIU, H. F. & UNGVARI, G. S. 2007. Rasch analysis of the scoring scheme of the HADS Depression subscale in Chinese stroke patients. *Psychiatry Res,* 150**,** 97-103.
TATTERSALL, C. & SULLIVAN, S. 2008. Audit of referrals for cataract extraction: are they appropriate? *Br J Nurs,* 17**,** 974-7.
TENENBAUM, G., FURST, D. & WEINGARTEN, G. 1985. A statistical reevaluation of the STAI anxiety questionnaire. *Journal of Clinical Psychology,* 41**,** 239-44.
TENENBAUM, G. & FURST, D. M. 1985. Similarities between retrospective and actual anxiety states. *J Psychol,* 119**,** 185-90.
THEODOSSIADES, J., MURDOCH, I. & COUSENS, S. 2004. Glaucoma case finding: a cluster-randomised intervention trial. *Eye,* 18**,** 483-490.
THEODOSSIADES, J. & MURDOCH, L. 1999. Positive predictive value of optometrist-initiated referrals for glaucoma. *Ophthalmic and Physiological Optics,* 19**,** 62-67.
TUCK, M. W. 1991. Referrals for suspected glaucoma - An International Glaucoma Association survey. *Ophthalmic and Physiological Optics,* 11**,** 22-26.
TUCK, M. W. & CRICK, R. P. 1991. Efficiency of referral for suspected glaucoma. *British Medical Journal,* 302**,** 998-1000.
TUCK, M. W. & CRICK, R. P. 1994. Non-contact tonometry: optometrists' current practice in England and Wales. *Ophthalmic and Physiological Optics,* 14**,** 347-50.
TUFFNELL, D. J., NUTTALL, K., RAISTRICK, J. & JACKSON, T. L. 1994. USE OF TRANSLATED WRITTEN MATERIAL TO COMMUNICATE WITH NON-ENGLISH SPEAKING PATIENTS. *British Medical Journal,* 309**,** 992-992.
TYMSTRA, T. 1986. False positive results in screening-tests - Experiences of parents of children screened for congenital hypothyroidism. *Family Practice,* 3**,** 92-96.
VALLEJO, M. A., JORDAN, C. M., DIAZ, M. I., COMECHE, M. I. & ORTEGA, J. 2007. Psychological assessment via the internet: a reliability and validity study of online (vs paper-and-pencil) versions of the General Health Questionnaire-28 (GHQ-28) and the Symptoms Check-List-90-Revised (SCL-90-R). *J Med Internet Res,* 9**,** e2.
VERNON, S. A. 1998. The changing pattern of glaucoma referrals by optometrists. *Eye,* 12**,** 854-857.
VERNON, S. A. & ADAIR, A. 2010. Shared care in glaucoma: a national study of secondary care lead schemes in England. *Eye (Lond),* 24**,** 265-9.
VERNON, S. A. & GHOSH, G. 2001. Do locally agreed guidelines for optometrists concerning the referral of glaucoma suspects influence referral practice? *Eye,* 15**,** 458-463.
VIANYA-ESTOPA, M., ELLIOTT, D. B. & BARRETT, B. T. 2010. An evaluation of the Amblyopia and Strabismus Questionnaire using Rasch analysis. *Invest Ophthalmol Vis Sci,* 51**,** 2496-503.
VIGNEAU, F. & CORMIER, S. 2008. The factor structure of the state-trait anxiety inventory: an alternative view. *J Pers Assess,* 90**,** 280-5.
WERNEKE, U., GOLDBERG, D. P., YALCIN, I. & USTUN, B. T. 2000. The stability of the factor structure of the General Health Questionnaire. *Psychological Medicine,* 30**,** 823-829.

WHITTAKER, K. W., IKRAM, K., ANDERSON, D. F., KIEL, A. W. & LUFF, A. J. 1999. Non-communication between ophthalmologists and optometrists. *J R Soc Med,* 92**,** 247-8.

WILLIS, C. E., RANKIN, S. J. & JACKSON, A. J. 2000. Glaucoma in optometric practice: a survey of optometrists. *Ophthalmic and Physiological Optics,* 20**,** 70-5.

WORMALD, R. P., WRIGHT, L. A., COURTNEY, P., BEAUMONT, B. & HAINES, A. P. 1992. Visual problems in the elderly population and implications for services. *BMJ,* 304**,** 1226-9.

ZIGMOND, A. S. & SNAITH, R. P. 1983. The Hospital Anxiety and Depression Scale. *Acta Psychiatrica Scandinavica,* 67**,** 361-370.

ZYZANSKI, S., STANGE, K., LANGA, D. & FLOCKE, S. 1998. Trade-offs in high-volume primary care practice. *J Fam Pract,* 46**,** 397-402.

Appendices

A. Letters of ethical approval

NHS
National Research Ethics Service

Bradford Research Ethics Committee
Top Floor
Extension Block
St Lukes Hospital
Little Horton Lane
Bradford
BD5 0NA

Chairman: **Professor Alan C Roberts**
OBE TD DL MPhil PhD DSc DTech LLD FLS FIBiol
Administrator: Sue Bell

Tel: 01274 365508
Fax: 01274 365509
Email: sue.bell@bradfordhospitals.nhs.uk
Email: alan.roberts@bradfordhositals.nhs.uk

18 June 2007

Mr Christopher J Davey
Post-Graduate (MPhil/PhD)
University of Bradford

Dear Mr Davey

Full title of study: **False positive referrals to Ophthalmology and levels of associated anxiety**
REC reference number: **07/Q1202/41**

Thank you for your letter of 05 June 2007, responding to the Committee's request for further information on the above research and submitting revised documentation.

The further information has been considered on behalf of the Committee by the Chairman.

Confirmation of ethical opinion

On behalf of the Committee, I am pleased to confirm a favourable ethical opinion for the above research on the basis described in the application form, protocol and supporting documentation.

Conditions of approval

The favourable opinion is given provided that you comply with the conditions set out in the attached document. You are advised to study the conditions carefully.

Approved documents

The final list of documents reviewed and approved by the Committee is as follows:

Document	*Version*	*Date*
Application	1	05 April 2007
Investigator CV		30 March 2007
Protocol	2	
Covering Letter		04 April 2007

NHS
National Research Ethics Service

Bradford Research Ethics Committee
Top Floor
Extension Block
St Lukes Hospital
Little Horton Lane
Bradford
BD5 0NA

Chairman: **Professor Alan C Roberts**
OBE TD DL MPhil PhD DSc DTech LLD FLS FIBiol
Administrator: **Susan Jude**

Tel: 01274 365508
Fax: 01274 365509
Email: susan.jude@bradfordhospitals.nhs.uk
Email: alan.roberts@bradfordhositals.nhs.uk

25 July 2007

Mr Christopher J Davey
Post-Graduate (MPhil/PhD)
Richmond Road
Bradford
West Yorkshire
BD7 1DP

Dear Mr Davey

Study title: **False positive referrals to Ophthalmology and levels of associated anxiety**
REC reference: **07/Q1202/41**
Amendment number: **1**
Amendment date: **01 July 2007**

The above amendment was reviewed at the meeting of the Committee held on 17 July 2007.

Ethical opinion

The members of the Committee present gave a favourable ethical opinion of the amendment on the basis described in the notice of amendment form and supporting documentation.

Approved documents

The documents reviewed and approved at the meeting were:

Document	Version	Date
Protocol	3	25 June 2007
Participant Information Sheet	3	25 June 2007
Printed versions of web pages	1	25 June 2007
Covering letter 6 months	2	25 June 2007

NHS
National Research Ethics Service

Bradford Research Ethics Committee
Top Floor
Extension Block
St Lukes Hospital
Little Horton Lane
Bradford
BD5 0NA

Chairman: **Professor Alan C Roberts**
OBE TD DL MPhil PhD DSc DTech LLD FLS FIBiol
Administrator: **Susan Jude**

Tel: 01274 365508 **Email:** susan.jude@bradfordhospitals.nhs.uk
Fax: 01274 365509 **Email:** alan.roberts@bradfordhositals.nhs.uk

26 November 2007

Mr Christopher J Davey
Department of Optometry
University of Bradford
Post-Graduate (MPhil/PhD)
Richmond Road
Bradford
BD7 1DP

Dear Mr Davey

Study title: **False positive referrals to Ophthalmology and levels of associated anxiety**
REC reference: **07/Q1202/41**
Amendment number: **2**
Amendment date: **02 November 2007**

The above amendment was reviewed at the meeting of the Committee held on 20 November 2007.

Ethical opinion

The members of the Committee present gave a favourable ethical opinion of the amendment on the basis described in the notice of amendment form and supporting documentation.

Approved documents

The documents reviewed and approved at the meeting were:

Document	Version	Date
Protocol	5	02 November 2007
Participant Information Sheet	5	02 November 2007
Notice of Substantial Amendment (non-CTIMPs)	2	02 November 2007
Letter of invitation to participant	5	02 November 2007

Appendix B: Questionnaires

State Trait Anxiety Inventory

SELF EVALUATION QUESTIONNAIRE

CODE: «Code»

Please read this carefully:

A number of statements which people have used to describe themselves are given below. Read each statement and then circle the appropriate answer to the right of the statement to indicate how you feel *right* now, that is *at this moment*. There are no right or wrong answers. Do not spend too much time on any one statement but give the answer which seems to describe your *present* feelings best.

It is important you try to answer all the questions.

Thank you very much for your co-operation.

1	I feel calm	Not at all	Somewhat	Moderately so	Very much so
2	I feel secure	Not at all	Somewhat	Moderately so	Very much so
3	I am tense	Not at all	Somewhat	Moderately so	Very much so
4	I feel strained	Not at all	Somewhat	Moderately so	Very much so
5	I feel at ease	Not at all	Somewhat	Moderately so	Very much so
6	I feel upset	Not at all	Somewhat	Moderately so	Very much so
7	I am presently worrying over possible misfortunes	Not at all	Somewhat	Moderately so	Very much so
8	I feel satisfied	Not at all	Somewhat	Moderately so	Very much so
9	I feel frightened	Not at all	Somewhat	Moderately so	Very much so
10	I feel comfortable	Not at all	Somewhat	Moderately so	Very much so
11	I feel self-confident	Not at all	Somewhat	Moderately so	Very much so
12	I feel nervous	Not at all	Somewhat	Moderately so	Very much so
13	I am jittery	Not at all	Somewhat	Moderately so	Very much so
14	I feel indecisive	Not at all	Somewhat	Moderately so	Very much so
15	I am relaxed	Not at all	Somewhat	Moderately so	Very much so
16	I feel content	Not at all	Somewhat	Moderately so	Very much so
17	I am worried	Not at all	Somewhat	Moderately so	Very much so
18	I feel confused	Not at all	Somewhat	Moderately so	Very much so
19	I feel steady	Not at all	Somewhat	Moderately so	Very much so
20	I feel pleasant	Not at all	Somewhat	Moderately so	Very much so

Questions continue overleaf →

SELF EVALUATION QUESTIONNAIRE

A number of statements which people have used to describe themselves are given below. Read each statement and then circle the appropriate answer to indicate <u>how you *generally* feel</u>.

21	I feel pleasant	Almost never	Sometimes	Often	Almost always
22	I feel nervous and restless	Almost never	Sometimes	Often	Almost always
23	I feel satisfied with myself	Almost never	Sometimes	Often	Almost always
24	I wish I could be as happy as others seem to be	Almost never	Sometimes	Often	Almost always
25	I feel like a failure	Almost never	Sometimes	Often	Almost always
26	I feel rested	Almost never	Sometimes	Often	Almost always
27	I am “calm, cool and collected”	Almost never	Sometimes	Often	Almost always
28	I feel that difficulties are piling up so that I cannot overcome them	Almost never	Sometimes	Often	Almost always
29	I worry too much over something that really doesn’t matter	Almost never	Sometimes	Often	Almost always
30	I am happy	Almost never	Sometimes	Often	Almost always
31	I have disturbing thoughts	Almost never	Sometimes	Often	Almost always
32	I lack self confidence	Almost never	Sometimes	Often	Almost always
33	I feel secure	Almost never	Sometimes	Often	Almost always
34	I make decisions easily	Almost never	Sometimes	Often	Almost always
35	I feel inadequate	Almost never	Sometimes	Often	Almost always
36	I am content	Almost never	Sometimes	Often	Almost always
37	Some unimportant thought runs through my mind and bothers me	Almost never	Sometimes	Often	Almost always
38	I take disappointments so keenly that I can’t put them out of my mind	Almost never	Sometimes	Often	Almost always
39	I am a steady person	Almost never	Sometimes	Often	Almost always
40	I get in a state of tension or turmoil as I think over my recent concerns and interests	Almost never	Sometimes	Often	Almost always

Hospital Anxiety and Depression Scale

CODE: «Code»

Please read this carefully:
Read each item and circle the reply that comes closest to how you have been feeling in the *past week*. Don't take too long over your replies; your immediate reaction to each item will probably be more accurate than a long thought out response
Thank you very much for your co-operation.

1	I feel tense or 'wound up'	Most of the time	A lot of the time	Occasionally	Not at all
2	I still enjoy things I used to enjoy	Definitely as much	Not quite so much	Only a little	Hardly at all
3	I get a sort of frightened feeling as if something awful is about to happen	Very definitely and quite badly	Yes, but not too badly	A little, but it doesn't worry me	Not at all
4	I can laugh and see the funny side of things	As much as I always could	Not quite so much now	Definitely not so much now	Not at all
5	Worrying thoughts go through my mind	A great deal of the time	A lot of the time	From time to time, but not too often	Only occasionally
6	I feel cheerful	Not at all	Not often	Sometimes	Most of the time
7	I can sit at ease and feel relaxed	Definitely	Usually	Not often	Not at all
8	I feel as if I am slowed down	Nearly all the time	Very often	Sometimes	Not at all
9	I get sort of frightened feeling like 'butterflies' in the stomach	Not at all	Occasionally	Quite often	Very often
10	I have lost interest in my appearance	Definitely	I don't take as much care as I should	I may not take quite as much care	I take just as much care as ever
11	I feel restless as I have to be on the move	Very much indeed	Quite a lot	Not very much	Not at all
12	I look forward with enjoyment to things	As much as I ever did	Rather less than I used to	Definitely less than I used to	Hardly at all
13	I get sudden feelings of panic	Very often indeed	Quite often	Not very often	Not at all
14	I can enjoy a good book or radio or TV program	Often	Sometimes	Not often	Very seldom

HADS copyright © R.P.Snaith and A.S.Zigmond, 1983, 1992, 1994. Record forms originally published in Acta Psychiatrica Scandinavica, 67, 361-70, copyright © Munksgaard International Publishers Ltd, Copenhagen 1983. Reproduced by permission of the Publishers, nferNelson Publishing Company Ltd., of The Chiswick Centre, 414 Chiswick High Road, London W4 5TF UK. All rights reserved including translation. nferNelson is a division of Granada Learning Limited.

«salutation» «forename» «surname»
«address_1»,
«address_2»,
«town»,
«county»,
«postcode»

Clare Green
Ophthalmology Department
Our ref:«Code»

<<Date>>,

Dear «salutation» «forename» «surname»,

We are writing to you regarding your upcoming visit to the Eye Department at Bradford Royal Infirmary. The University of Bradford is working with Bradford Royal Infirmary to improve the quality of service that patients receive, and as part of this research we would like to ask all new patients to take part in a study. We thought we'd take this opportunity to give you some information about the research before your visit so you can decide whether you would like to participate.

Enclosed with this letter are an information sheet and 2 questionnaires in an envelope. If you decide to take part, please bring the questionnaires and the envelope with you to your appointment and complete them while you are sat waiting for your appointment (after you have reported to reception or seen the nurse). Do not worry if some of the questions sound very similar, please try to answer them all. When you are called for your appointment, please give the completed questionnaires to the nurse or doctor who calls you. If you have any questions, feel free to talk to the research team on the day of your appointment or contact them in advance using the contact details enclosed. Unfortunately we are unable to pay you for participation.

Kind regards,
Yours sincerely,

Clare Green BSc (Hons). MCOptom. DipTp (AS), DipTp (SP).
Optometrist Consultant, Bradford Royal Infirmary.

Patient Information Sheet

Study Title: Referrals to Ophthalmology

Who am I?

I am an Optometrist Consultant working at the Eye department at Bradford Royal Infirmary. I am researching the eye care screening and referral process, to see whether it is possible to improve the services that patients receive. This research is being done in collaboration with researchers from the University of Bradford.

What is the purpose of the study?

We hope to find out how accurate the screening tests and referrals from doctors and optometrists (opticians) are. We will also use questionnaires to try to determine how people feel before their appointments. This will help improve the quality of referrals for future patients.

Why have you been chosen?

Because you are of suitable age, and have recently been referred to see an Ophthalmologist in the Hospital Eye Service.

What will happen if you agree to take part?

We ask you to complete two simple questionnaires. The questionnaires are enclosed with this leaflet, and are meant to be completed while you are sat waiting for your appointment in the hospital. These should take about 10 minutes to complete and should be given to the doctor or nurse who calls you for your appointment. We will record a few relevant details from your medical records after your appointment with the Ophthalmologist. These include your date of birth, gender, details of the optician or doctor who first referred you to the hospital, the reason why you have been referred and the results of your appointment.

How long will the questionnaires take and where will they be completed?

The questionnaires should take less than 15 minutes to complete. They will be completed in the hospital waiting room while you are waiting for your appointment. Please bring them to the hospital on the day of your appointment.

What happens after this study?

The questionnaire responses from all patients will be gathered together and analysed to see how any improvement to referrals can be made. We will not record your name with your questionnaire responses. Each questionnaire will be marked with a unique code which can be cross referenced with your hospital records. This means that your responses to the questionnaires will

be recorded anonymously. **Your details will not be passed on to anyone else or used by us for any other purpose than this study.**

Is there any risk of harm to myself?

No harm will come to you from taking part in this study. No clinical procedures will be carried out on you by the researchers. The questionnaires that we ask you to complete are commonly used by doctors and researchers. All the information we require for the study will be taken from your hospital medical record and the questionnaires that you complete. You can withdraw from the study at any time. If you decide not to take part in the study, don't complete the questionnaires. This will not affect the standard of care you receive from the hospital. The results of this study will be used for research purposes. If published, all data will remain anonymous. If you would like to be notified of when and where the research is published, please tick the relevant box on the questionnaire.

Why should I be involved?

You are under no obligation to take part in this study. However, your participation will be greatly valued and will contribute to the improvement of Bradford's eye screening and referrals process.

Thank you for taking the time to read this information sheet.

FURTHER INFORMATION IS **AVAILABLE FROM MEMBERS OF THE RESEARCH TEAM**:

- Christopher Davey BSc (Hons). MCOptom. Post Graduate, Optometry, University of Bradford. BD7 1DP: Tel. 01274 236230 Email: c.j.davey@bradford.ac.uk
- Clare Green BSc (Hons). MCOptom. DipTp (AS), DipTp (SP). Optometrist Consultant, Bradford Royal Infirmary. BD9 6RJ Tel. 01274 382386 Email: Clare.Green@bradfordhospitals.nhs.uk
- John Bradbury MBChB, FRCS, FRCOphth. Consultant Ophthalmologist, Bradford Royal Infirmary. BD9 6RJ
- David B Elliott PhD. BSc (Hons). MCOptom. FAAO. Professor of Clinical Vision Science, Optometry Dept, University of Bradford. Tel. 01274 234640
- Clare Harley PhD. BSc (Hons). Psychology. Post-doctoral Research Assistant, Optometry Dept, University of Bradford. Tel. 01274 234640

Referral Refinement Protocol

A second appointment should be made within 2 weeks if any of the following findings are noted at the patient's initial eye examination.

Increased Intra Ocular Pressures and / or Suspicious Visual Fields

- IOP of 25mmHg or higher in either eye (or in the case of non contact instruments the average reading should be 25mmHg or higher) with a normal visual field.
- Difference of 5mmHg or more between the two eyes (or in the case of non contact instruments the difference between the average readings should be 5mmHg or higher) with a normal visual field..
- At least 3 points missed in either eye on a screening mode on a visual field screening machine, with IOP's of less than 25mmHg

Notes

1. Ensure that at least 3 readings are taken for non contact instruments. These should all be recorded.
2. Also ensure that instrument type and time is recorded on the patient record card.
3. CD ratio should also be recorded on the patients' record card.
4. Optometrists are advised to always work within the guidelines on glaucoma referral in the College of Optometrists Framework of Referrals 2005.
5. A print out of the deficient visual field should be retained in the patients' record.
6. Should there be 3 or more points missed on visual field screening in combination with the IOP anomalies above, then a practitioner should exercise his/her judgement and act appropriately regarding referral.

Repeat Examination will consist of applanation tonometry and threshold visual field examination.

- Should the IOP read 25mmHg or above in either eye, the patient should be referred to their GP.
- Should the difference in pressures between the two eyes be 5mmHg or more, the patient should be referred to their GP.
- Should the threshold visual field highlight a reduction or depression of the visual field the patient should be referred to their GP.

Should the IOP readings revert to within normal ranges and the visual field proves to be intact then the patient will be put onto an appropriate recall and seen as normal either one or two years later.

Suspicion of Retinal or Macula Anomaly

Symptoms may include (but are not confined to)

- Flashing lights
- Sudden increase in floaters
- Sudden decrease in vision
- Visual Distortion

Findings may include (but are not confined to)

- Suspicious appearance of any part of the retina.
- Decreased visual acuity.

Repeat Examination will consist of dilated fundus examination with slit lamp or other indirect biomicroscopy.

Should any suspicion be confirmed the patient should be referred to their GP or (in the case of wet macular degeneration) direct to the macular assessment unit at St Lukes Hospital, using the appropriate referral form.

Should the original suspicion be eliminated then the patient will be put onto an appropriate recall and seen as normal either one or two years later.

Summary

Initial Findings	Repeat Examination(s)
IOP of 25mmHg or higher in either eye	Applanation tonometry and threshold visual field examination
Difference of 5mmHg between the two eyes.	Applanation tonometry and threshold visual field examination
3 or more points missed on a visual field screening program in either eye.	Applanation tonometry and threshold visual field examination
Suspicion of retinal or macula anomaly	Dilated slit lamp bio microscopy

Notice of amendment to protocol

7th May 2009

Dear

Re: Referral Refinement and recently issued NICE guidelines

Thank you for your continued involvement in the referral refinement study.

You will no doubt be aware of the recently issued guidance regarding glaucoma and referral of patients demonstrating pressures of over 21mmHg.

In light of this, it would seem sensible to alter the parameters for referral refinement, with regard to the IOP measurements.

Please adopt the referral refinement pathway for any patient that demonstrates IOP's of over 21mmHg in either eye. If, at the referral refinement appointment the pressures remain at this level, then refer the patient, even in the absence of any other glaucomatous signs as per NICE guidance.
Should applanation tonometry reveal lower pressures than 21mmHg then the patient need not be referred.

Please keep all other parameters for the referral refinement study the same.

Could I also please remind you to ensure that your electronic activity sheets are submitted at the end of May and July.

Kind regards,

Referral Refinement leaflet/questionnaire

General Information

Should you not be able to attend your second appointment please telephone the practice to let us know. We will be happy to rearrange the appointment at a suitable time, but this must be within 2 weeks of the original appointment.

Should you not attend your appointment and the practice has not been informed we will write to your GP with our original findings and you will probably be given an appointment at the hospital. The waiting time for a hospital appointment in the eye department is currently about 15-18 weeks.

Practice Stamp

Dilated Retinal Examination

When you call back you will have some drops put into your eyes. These will take about 15-20 minutes to take effect and cause your pupils to dilate.
Once your pupils have dilated the optometrist will be able to examine the back of your eye very thoroughly using a special lens and a special eye microscope called a slit lamp.
Should anything abnormal be found during this examination we will again refer you to your GP with a recommendation that you are seen by an eye specialist at the hospital.

If there are no abnormalities found then we will be able to see you as normal when you return for your regular scheduled eye examination.

Please Note:

These dilating drops may cause some blurring of vision for some time (typically 2-3 hours) following instillation.

For this reason you should not drive afterwards

Please make alternative travel arrangements if you were intending to drive to and from this appointment.
Also as the drops cause the pupils to dilate it is common to experience a degree of light sensitivity following the appointment and it would be advisable to bring some dark glasses in case it is sunny on the day of your appointment.

REFERRAL REFINEMENT INFORMATION LEAFLET

THIS LEAFLET IS ALSO AVAILABLE IN LARGE PRINT

Thank you for coming for your eye examination today.

Your optometrist has requested that you come back and see him/her again for a second examination. The tests that will be carried out at this examination are ticked below.

- ○ Intra Ocular Pressures & Visual Field Assessment
- ○ Dilated Retinal Examination

Intra Ocular Pressure Measurement.

The pressure inside your eye can be dependent on many factors. It may vary according to the time of day or it can be raised simply due to slight apprehension of the procedure itself.
More seriously it can be a sign of glaucoma, a condition that is virtually symptom free to begin with, but can damage ones eyesight if it is not caught and controlled at an early stage.

When you return for your repeat examination the optometrist will check your pressures again, this time with a different instrument which is more consistent with methods used at the hospital.

Should the pressures still be high we will refer you to your GP with a recommendation that you are seen by an eye specialist so that the cause of your raised intra ocular pressure can be determined.

Should the pressures be back within the normal range we will be able to just keep a regular check here at the practice when you come for your regular eye examinations.

Visual Field Assessment.

Your visual field can also vary for many reasons. Sometimes facial features, spectacle frames, and even tiredness can all flag up a deficient visual field.

Such problems are easily explained and hence these are not reasons that would require referral to a hospital.

To eliminate these factors we will repeat the field test using a more detailed setting on the machine. This will give us a very accurate assessment of your visual field. Since you will now be used to the instrument, any errors that were initially caused by unfamiliarity will be eliminated too.

Should there be a reduction or a problem with your visual field this is always something that should be investigated further by a specialist in the hospital and again we will refer you to your GP in the first instance with a recommendation that you are referred appropriately to the hospital.

Please find opposite a short set of questions that we would like you to answer for us. Your responses will enable us to evaluate this scheme and make improvements.

Please have a look through the questions and answer them following your second appointment. Please detach this and hand it in before you leave.

Thank you for taking the time to do this.

Patient Questionnaire

I am happy having these procedures performed at my opticians rater than at the hospital

- ○ Strongly agree
- ○ Agree
- ○ Neither agree nor disagree
- ○ Disgree
- ○ Strongly disagree

I am happy with the time I waited between appointments

- ○ Strongly agree
- ○ Agree
- ○ Neither agree nor disagree
- ○ Disgree
- ○ Strongly disagree

I am satisfied overall with the service that I have received

- ○ Strongly agree
- ○ Agree
- ○ Neither agree nor disagree
- ○ Disgree
- ○ Strongly disagree

What date was your first appointment?

_____

What date was your second appointment?

_____

How long were you at the opticians for your second appointment?

_____

Appendix E: Poster Presentation for Chapter 7

NHS Bradford and Airedale

An Optometric Referral Refinement pilot study

© NHS Bradford and Airedale

Chris Davey, David B. Elliott, Bradford School of Optometry and Vision Science.
Ravi Naru, Edwin Bonner. NHS Bradford and Airedale.

BRADFORD SCHOOL OF OPTOMETRY AND VISION SCIENCE

1 INTRODUCTION

- A significant number of referrals from optometrists in primary eye care to secondary eye care are false positive [1, 2].
- The present study assessed the economic feasibility and patient preference of a referral refinement scheme in which optometrists repeated certain techniques on patients that would otherwise have been referred, in order to reduce false positive referrals.
- The NICE guidelines for glaucoma were introduced approximately halfway through the pilot [3] meaning that any effect on the service could be identified.

2 METHODS

- A 20% sample of optometric practices in Bradford and Airedale were randomly selected to participate in a six month pilot study between 1/2/09 and 31/7/09.
- Patients were offered a referral refinement (RR) appointment if they would otherwise have been referred on the basis of:
  - IOP >24mmHg in either eye (changed to >21mmHg after the NICE guidelines).
  - IOP asymmetry of >4mmHg.
  - 3 or more points missed on a suprathreshold visual field screening in either eye.
  - Suspicion of fundus abnormality where dilatation would aid diagnostic accuracy.
- At the RR appointment, either contact tonometry and a threshold visual field assessment, or a dilated indirect ophthalmoscopy examination were performed. The practice was paid £42 per appointment.
- If findings were within normal limits then the patient was placed on normal recall, otherwise they were referred to secondary eye care.
- Patients were also asked to complete a brief satisfaction questionnaire.

3 RESULTS

- 134 patients were eligible for the Referral Refinement Service over the 6 month timescale.
- 83 patients (62%) did not require referral to secondary care (Figure 1).

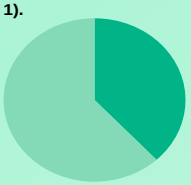

Figure 1. Outcome of all referral refinement appointments between 1/2/09 and 31/7/09 (n=134)

- After the NICE guidelines the number of patients eligible for referral refinement increased (Table 1).
- The proportion of patients not requiring referral on the grounds of IOP/Visual fields also increased (Table 1/Figure 3).

Table 1.

	Pre NICE (1/2/09-22/4/09)			Post NICE (23/4/09-31/7/09)		
	N	Referred	Not Referred	N	Referred	Not Referred
Tonometry and VFs	25	12 (48%)	13 (52%)	53	19 (36%)	34 (64%)
Mydriasis and Volk	24	7 (29%)	17 (71%)	24	8 (33%)	16 (67%)
All three techniques	3	2 (67%)	1 (33%)	5	3 (60%)	2 (40%)
Total	52	21 (40%)	31 (60%)	82	30 (37%)	52 (63%)

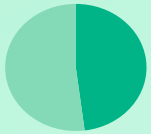

Figure 2. Outcome of Tonometry/VF referral refinement appointments **before the NICE guidelines** (n=52).

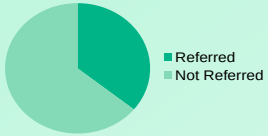

Figure 3. Outcome of Tonometry/VF referral refinement appointments **after the NICE guidelines** (n=82).

- Cost savings were £3,502, calculated from 83 x £110 (hospital tariff cost for patients that would have been referred) less the cost of the service (134 x £42), therefore the potential savings per year across the area would be approximately £35K.
- Forty-eight of the 134 patients (36%) completed satisfaction questionnaires. Results are given in table 2.

Table 2. Results of the three item satisfaction questionnaire.

	Strongly agree	Agree	Neither agree nor disagree	Disagree	Strongly disagree
I am happy having these procedures performed at my opticians rather than at the hospital	52% (25)	46% (22)	2% (1)	0	0
I am happy with the time I waited between appointments	63% (30)	37% (18)	0	0	0
I am satisfied overall with the service I have received	65% (31)	31% (15)	4% (2)	0	0

5 CONCLUSIONS

- The pilot referral refinement service reduced false positive referrals to secondary eye care by 62%.
- Not only did this make a significant cost saving, it was viewed very positively by patients.
- Potential secondary benefits include increased capacity within the hospital eye service, and reduced waiting times.

6 REFERENCES

1. Pooley JE & Frost EC. Optometrists' referrals to the Hospital Eye Service. *Ophthalmic & Physiological Optics*, 1999; 19: S16-S24.
2. Bowling B, Chen SDM & Salmon JF. Outcomes of referrals by community optometrists to a hospital glaucoma service. *British Journal of Ophthalmology*, 2005; 89: 1102-1104.
3. Glaucoma: diagnosis and management of chronic open angle glaucoma and ocular hypertension. 2009, National Institute for Health and Clinical Excellence (NICE).

7 ACKNOWLEDEMENTS

- The pilot study was funded by NHS Bradford and Airedale.
- The authors would like to thank the practices that volunteered to participate in this pilot study.