DELAY IN ACCESSING HEALTHCARE AFTER TRANSIENT ISCHAEMIC ATTACK AND MINOR STROKE: THE ROLE OF PRIMARY CARE IN THE PROBLEM AND THE SOLUTION

DANIEL SASCHA LASSERSON M.A.
Fitzwilliam College

Department of Primary Care Health Sciences
University of Oxford
2nd Floor 23-38 Hythe Bridge Street
Oxford OX1 2ET
SUMMARY

Background

Transient ischaemic attack (TIA) and minor stroke are associated with a high risk of recurrent stroke which can be predicted with a clinical rule and reduced with urgent treatment. Delay in accessing assessment and vascular risk factor modification should therefore be as short as possible, yet little is known in the UK about where patients seek care and the key influences of the time to contact healthcare services. However, using cohort studies to answer questions on healthcare access requires an assessment of how well such cohorts represent the wider population. Within the primary care consultation, the recognition of TIA is an important step in the care pathway as definitive treatment is initiated by specialists, yet TIA presentations are not common for individual GPs and difficulties in diagnosis may be due to low clinical exposure in routine practice or inadequacies in training. For patients where GPs suspect that TIA may be the cause of symptoms, inaccurate risk prediction and diagnosis of TIA can result in delay to definitive care and the existing tools for prognosis and diagnosis have been exclusively derived from clinical assessments in secondary care rather than primary care.

RESEARCH QUESTIONS

1. Is the OXVASC population representative of the national population in terms of age and deprivation?
2. Do differences in age and deprivation between OXVASC practices affect the presentation of patients with TIA to healthcare?
3. Is healthcare access after TIA or minor stroke influenced by the following factors; choice of healthcare provider, time of symptom onset, clinical features of the event, deprivation and previous history of cerebrovascular disease?
4. Did the change in general practitioner contract influence healthcare access after TIA or minor stroke?
5. Do GP trainees fail to recognise high risk TIA cases that have been missed by established GPs?
6. What are the recruitment and completion rates for a web-based vignette study of TIA recognition by GP trainees?
7. Does alteration of one case vignette parameter influence TIA recognition and management decisions?
8. Do GPs’ and specialists’ records of the same patients with suspected TIA agree about the clinical features?
9. Does GP/specialist disagreement vary according to final clinic diagnosis?
10. What impact does GP/specialist disagreement have on the performance of the ABCD2 score for accurate triage of patients with TIA?

11. What impact does GP/specialist disagreement have on the discrimination metrics of existing clinical prediction tools for TIA diagnosis?

12. Which clinical predictors are included in a prediction rule for TIA diagnosis derived from primary care records in suspected TIA?

13. What are the calibration and discrimination metrics for a model of TIA diagnosis derived from primary care records in suspected TIA?

14. Does choice of statistical model affect discrimination metrics for TIA diagnostic models?

METHODS

Questions 1 and 2 were answered with a comparison of the age and deprivation structure of the Oxford Vascular Study (OXVASC) population and the national population, as well as inter-practice variation in age and deprivation and its effect on the ratio of TIA to major stroke. Questions 3 and 4 were answered by determining the routes to healthcare and delay to call for medical attention for patients with TIA and minor stroke ascertained between 2002 and 2006 in OXVASC. Questions 5, 6 and 7 were answered by recruiting GP trainees to a pilot study of a web-based vignette questionnaire. This presented, in random order, missed high risk TIA cases (patients ascertained with stroke who had presented to primary care in the previous 30 days with transient neurological symptoms), matched cases of TIA where one parameter was altered to test its influence on decision making and distractor cases without a cerebrovascular diagnosis. Questions 8, 9, 10 and 11 were answered with an analysis of referrals to the OXVASC TIA clinic from primary care with GP and specialist records compared to assess disagreement over clinical features. The ABCD2 score was calculated for all referred patients from primary care records and compared with the ABCD2 score from hospital records for the accuracy of detection of high risk status in primary care. The discriminating ability for TIA diagnosis of the only existing TIA recognition tool in the literature and the ABCD2 score were compared. The effect of using primary care records rather than secondary care records to populate the scores was assessed using receiver operator characteristic curves. Questions 12, 13 and 14 were answered by qualitatively analysing primary care clinical records from consultations and referral letters, to determine groups of similar symptoms for potential predictors for a decision tool, and then deriving diagnostic models using three alternative modelling techniques – logistic regression, classification trees and random forest analysis.
RESULTS

1. Compared with the population of England, the registered population at OXVASC practices were less deprived across all age ranges, although the population age structure and change in relative deprivation across age ranges were equivalent.

2. The age and deprivation structure of practices in the OXVASC study varied significantly but there was no relationship between these parameters and the ratio of ascertained TIA to major stroke.

3. Of 359 patients with TIA and 434 with minor stroke, 25 were excluded who were outside the study area at the time of their event. The majority of patients (73%) sought care from general practice. The median (IQR) time to call for medical attention was longer for patients choosing primary care rather than the emergency department (11.0 (46.0) hrs vs 1.0 (4.0) hrs, P<0.001). In the primary care population those calling for attention following events occurring outside office hours was longer than for those calling following in hours events (12.0 (41.0) hrs vs 4.0 (44.5) hrs, P <0.001). For the cohort of patients who sought care in general practice, those using an on call primary care service had significantly shorter delays to call than those (the majority, 72%) who waited until their general practice was open (1.0 (2.46) hrs vs 24.8 (48.5) hrs, P<0.001). There was no effect of deprivation, or prior history of cerebrovascular disease on choice of provider or delay to call. Weakness and speech disturbance were associated with shorter times to call for medical attention, and visual and sensory disturbance with longer times to call.

4. The only significant effect of the new GMS contract was a small increase in those using an emergency primary care service for events occurring in the out of hours period (20% vs 32%, P = 0.034)

5. Among GP trainees who responded to the invitation to take part in a questionnaire survey of clinical diagnosis and management from vignettes there was variable recognition of TIA in the high risk TIA cases that had been missed in routine clinical practice. For the three actual missed cases TIA was diagnosed in 9%, 72% and 27% (n=11)

6. Of 54 Oxford deanery GP trainees invited to participate in the high risk TIA diagnosis study, the recruitment rate was 35% (n=19) and of those who responded, the completion rate of all cases was 58% (n=11).

7. Recognition of TIA appeared to be influenced by the alteration of single parameters but a study powered to detect parameter influences on recognition would require 300 trainees and therefore collaboration across multiple deaneries.
8. There was disagreement in records of clinical histories between primary and secondary care for the symptoms that are used in the ABCD2 risk prediction score. Agreement for motor symptoms and speech symptoms showed kappa values of 0.58 (S.E. 0.04) and 0.68 (S.E. 0.04) respectively.

9. Altman-Bland plots demonstrated wide disagreement in ABCD2 scores and this varied with diagnosis. Comparing ABCD2 score disagreements showed that specialists would tend to score TIA patients higher and non-TIA patients lower than GPs (mean (S.E.) difference, 0.1 (0.07) for TIA vs -0.29 (0.07) for non-TIA patients p<0.001).

10. Primary care based ABCD2 risk scores detected 89.7% of high risk anterior circulation TIA patients and 61.5% of high risk posterior circulation TIA patients as judged from secondary care ABCD2 scores. Inaccurate triage in primary care based on these parameters is likely to result in one case of stroke every four years in a population of 91,000.

11. The existing potential decision rules for TIA diagnosis among patients with suspected TIA had higher discrimination metrics in secondary care records than in primary care records. Area under the receiver operator characteristic curve (AUC) (S.E.) from secondary care assessments for discriminating all TIA, anterior circulation TIA and posterior circulation TIA from non-TIA were 0.71 (0.02), 0.75 (0.02) and 0.59 (0.03) respectively using the ABCD2 score and 0.82 (0.02), 0.85 (0.02) and 0.70 (0.03) using the Dawson TIA recognition tool. AUC values for basing scores on primary care records for all TIA, anterior TIA and posterior TIA discrimination from non-TIA were 0.62 (0.03), 0.65 (0.03) and 0.48 (0.04) for the ABCD2 score and 0.70 (0.02), 0.75 (0.03) and 0.52 (0.04) for the Dawson TIA recognition tool.

12. Primary care clinical records in 496 referred patients were analysed and 17 predictors were defined for model derivation. Unilateral limb weakness was not a predictor of TIA diagnosis, and focal sensory limb deficit was predictive of non-TIA diagnoses in a logistic regression model. The derived score had nine variables requiring calculation.

13. For the logistic regression model on internal validation, calibration was acceptable (Hosmer Lemeshow goodness of fit test p = 0.65) and AUC values (S.E.) for discriminating all TIA, anterior TIA and posterior TIA were 0.81 (0.02), 0.82 (0.02) and 0.77 (0.04) respectively.

14. A classification tree model using the entire dataset had AUC (S.E.) values for discriminating all TIA, anterior TIA and posterior TIA of 0.80 (0.02), 0.80 (0.02) and 0.80 (0.03) respectively. The random forest analysis although more complex was inferior in
discriminating ability in all diagnostic groups. A simplified classification tree using five questions had a negative predictive value for TIA diagnosis of 96%.

**CONCLUSION**

The OXVASC registered population does not represent the national population in terms of deprivation, although the age structures are similar, and so the strength of associations with deprivation in the cohort may not be generalisable. The majority of patients after TIA and minor stroke seek care in general practice and will wait to be seen at their registered practice if their symptoms start in the out of hours period. Delay in calling for medical attention is strongly influenced by timing of symptom onset and availability of routine primary care at that time. The organisation and delivery of primary care influences healthcare seeking behaviour and is a barrier to achieving the National Stroke Strategy targets, although the GMS contract has had a small positive impact on out of hours access. GP trainees may show similar difficulties in recognising high risk TIA cases as established GPs but a larger definitive study is required to determine this as well as the influence of single alterations in vignettes on diagnosis. Disagreement exists between GPs and specialists over the clinical histories of events in patients with suspected TIA, which affects triage with the ABCD2 score and results in poor validations of diagnostic tools that have been derived from secondary care records. The use of primary care datasets in patients referred to secondary care produces complex logistic regression scores and simpler classification trees with high rule out function to aid referral decisions but predictors do not match the classical TIA phenotype by excluding weakness, due to a selection bias that results from using patients that have been referred rather than initially presenting in primary care. Inclusion of all high risk phenotypes in derivation datasets, avoiding selection bias from referred populations and external validation are required. This could be achieved by deriving and validating decision rules in the general population with transient symptoms that presents in primary care, rather than in the selected population of patients where a GP has determined the need for a TIA clinic assessment, although there are barriers to recruitment in such studies.

**KEY WORDS** Transient ischaemic attack, delay, primary care, diagnosis
I would like to acknowledge the support and mentorship of Professor David Mant and Professor Peter Rothwell during the research that has culminated in this thesis. Their counsel and critical insights have taught me how to think like a scientist.

The thesis is dedicated to Amelie, Joseph and Desi who collectively but unknowingly have given me strength to finish it with their love and humour - our family life has spanned the timeframes of this work.

It is a great sadness that there are two GPs who have been my friends, confidents and teachers that did not live to see this thesis finished – my father John and my uncle Mike.
# CONTENTS

<table>
<thead>
<tr>
<th>SECTION 1</th>
<th>Introduction, review and research questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chapter 1</td>
<td>Introduction</td>
</tr>
<tr>
<td>Chapter 2</td>
<td>Literature Review</td>
</tr>
<tr>
<td>Chapter 3</td>
<td>The Evidence Gap</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SECTION 2</th>
<th>The Oxford Vascular Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chapter 4</td>
<td>The Oxford Vascular Study – methods and generalisability</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SECTION 3</th>
<th>Initial Interface Between Patient and Healthcare System</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chapter 5</td>
<td>Methods</td>
</tr>
<tr>
<td>Chapter 6</td>
<td>Choice of Provider after TIA and minor stroke</td>
</tr>
<tr>
<td>Chapter 7</td>
<td>Delay in seeking care after TIA and minor stroke</td>
</tr>
<tr>
<td>Chapter 8</td>
<td>Missed TIA in Primary Care</td>
</tr>
<tr>
<td>Chapter 9</td>
<td>Summary and Discussion</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SECTION 4</th>
<th>GP and Specialist Disagreement – risk tools and diagnosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chapter 10</td>
<td>Methods</td>
</tr>
<tr>
<td>Chapter 11</td>
<td>Difference in histories and ABCD2 score</td>
</tr>
<tr>
<td>Chapter 12</td>
<td>ABCD2 for diagnosis</td>
</tr>
<tr>
<td>Chapter 13</td>
<td>Validation of a TIA Recognition Tool</td>
</tr>
<tr>
<td>Chapter 14</td>
<td>Summary and Discussion</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SECTION 5</th>
<th>New Prediction Rules for TIA in Primary Care</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chapter 15</td>
<td>Methods</td>
</tr>
<tr>
<td>Chapter 16</td>
<td>Coding Predictors and Logistic Regression</td>
</tr>
<tr>
<td>Chapter 17</td>
<td>Classification Tree</td>
</tr>
<tr>
<td>Chapter 18</td>
<td>Random Forest</td>
</tr>
<tr>
<td>Chapter 19</td>
<td>Summary and Discussion</td>
</tr>
<tr>
<td>Chapter 20</td>
<td>Final Conclusion</td>
</tr>
</tbody>
</table>

Reference List

Page
Appendix 1  
Information for MD Examiners

Appendix 2  
Search Strategy

Appendix 3  
TIA case vignettes
Distractor (non TIA) vignettes

Appendix 4  
Materials for GP Trainee High Risk TIA Study
Section 1  Introduction, review and research questions
1.1 Observations from Clinical Practice

As a Specialist Registrar in General and Stroke Medicine, I was struck by how often patients with completed stroke would arrive in the Emergency Department (ED), but rarely would the ED have the opportunity to assess patients with its precursor, transient ischaemic attack (TIA). This would contrast with patients who had transient symptoms of chest pain who would very frequently attend during or after an episode of pain. It seemed from these observations that healthcare seeking behaviour after transient symptoms potentially due to cerebral ischaemia may be very different from healthcare seeking behaviour after transient symptoms potentially due to cardiac ischaemia. The underlying pathological processes can be quite similar but the actions taken by patients can be very different.

After re-training in General Practice, I encountered the related problem of deciding which patients with transient or mild persisting neurological symptoms should be referred for further specialist assessment. As a GP, one of my main clinical roles is to accurately identify low prevalent but serious disease from a background that includes a range of presentations overlapping with those of serious conditions. I am also mindful of the need to use resources appropriately and not to expose patients to the unnecessary harms of over investigation and treatment.

These issues have become increasingly important with the realisation that delays in receiving risk reducing interventions after the onset of symptoms of TIA or minor stroke can determine rates of recurrent stroke. I became interested in investigating how primary care could be contributing to the problem of delays in optimal treatment for patients at high risk of stroke and also what changes could be implemented so that the problem could be alleviated.

The burden of cerebrovascular disease is heavy as it causes disability and death. Stroke causes the greatest number of years lived with disability of all brain diseases in Europe (1) and it is the third most common cause of death in the United Kingdom, resulting in 137,000 deaths from 2007 - 2009 (2). The total cost to the UK economy from direct treatment, ongoing social care and loss of productivity in people of working age has been estimated at £9 billion per year (3).

Reducing the burden of cerebrovascular disease will therefore have benefits for the individual and their families, the healthcare system and the economy as a whole. I wanted to study how primary care could contribute to this healthcare goal. Working in Oxford gave me the opportunity to examine how the Oxford Vascular Study (OXVASC) could help in
answering research questions concerning the improvement of the patient pathway from symptom onset to definitive treatment.

1.2 Reducing cerebrovascular disease burden – treatment or prevention of acute ischaemic stroke?

The majority of strokes are due to acute ischaemia and this has been documented consistently across countries, using varying methods of case ascertainment, in a recent review of 56 prospective population based studies which span across national income, recruiting cohorts from 1970 to 2008 (4). Ischaemia results from in-situ atherothrombosis or thrombo-embolic disease from a distal site, usually intra-cardiac thrombus, thrombus forming on a ruptured plaque in the arterial tree of large arteries or systemic thrombus via an intra-cardiac shunt. The major risk factors for acute ischaemic stroke alongside genetic susceptibility are either chronic non-communicable diseases or risk factors for these diseases: diabetes, atrial fibrillation, ischaemic heart disease/peripheral vascular disease, chronic kidney disease, obesity, hypertension, hypercholesterolaemia, smoking and obesity.

Treatment of established ischaemic stroke has benefited from the increasing evidence base for intravenous and intra-arterial thrombolysis (5) with subsequent demonstrations from prospectively collected audit data of the safety of thrombolysis in routine practice (6). Although deaths are increased with the use of thrombolysis (mainly early after treatment) there is a significant reduction in the numbers of patients who are dependent long term after stroke and the commonly chosen composite outcome of death or dependency at three to six months reflects the net benefit of acute stroke treatment.

The burden of stroke is also influenced by the structure of health care delivery as well as specific medical interventions. Admission to an acute stroke unit rather than a general medical ward also affects stroke outcome with a reduction in death or dependency at one year (7). Variations in coordinated stroke-specific care may also offer additional benefits at crucial times after stroke, as greater reductions in long term dependency and institutionalisation for selected patients have been demonstrated using early supported discharge from stroke units (8).

The scope for reducing total cerebrovascular disease burden from acute treatment is limited. Extrapolating from randomised controlled trials, the majority of patients treated with thrombolysis will not actually benefit from it, as on average 20 patients need to be treated to prevent one patient being either dead or dependent at six months after stroke (5). Furthermore, as there is a limited time window for the safety of thrombolysis, recently extended to 4.5 hours post stroke (9), factors need to be addressed that influence time to
presentation to medical services after stroke as well as time to deliver thrombolysis after presentation (e.g. transfer to a specialist unit).

One nationwide study from Sweden has assessed the numbers of patients with ischaemic stroke who received thrombolysis and although improvements were seen from 2003 to 2010, only 8.7% of acute stroke patients received this intervention (10). By comparison, in the UK in 2010, 5% of all patients with stroke received thrombolysis (11) and this was not limited by pre-hospital delay, as this figure represents only 25% of all eligible patients, i.e. patients who attended within the correct timeframe and had a proven ischaemic stroke. Thus given the limited gain at population level from acute stroke treatment, due either to efficacy or access, improved stroke prevention strategies are required to reduce disease burden.

1.3 Stroke prevention – primary and secondary strategies

There are essentially two preventative strategies to reduce cerebrovascular burden; prevention of either a first ever event (primary prevention) or prevention of a recurrent event (secondary prevention).

Primary prevention involves identifying modifiable risk factors and then controlling them with lifestyle alteration or medication. Identification in this setting could take place as a result of screening e.g. the current Department of Health Vascular Checks programme (12) or opportunistically if patients present to healthcare services for another reason and either the patient requests, or the clinician suggests, some form of screening test for a vascular risk factor e.g. blood pressure or a blood sample.

Secondary prevention is initiated after a first cerebrovascular event, in order to reduce the probability of future cerebrovascular events occurring. Delivery of such a strategy requires the recognition that an event has occurred and then implementation of factor control, addressing the currently known cardiovascular risk factors of blood pressure, blood lipids, glycaemic control in diabetes, lifestyle alteration and evidence-based interventions of anti-platelet agents or anticoagulants.

In order to reduce the overall burden of disease, it is particularly important for patients and the healthcare system to recognise and appropriately react to mild or transient cerebrovascular events. This will maintain individuals in a disabling stroke-free status for as long as possible.
1.4 The importance of transient ischaemic attack (TIA) and minor stroke for secondary prevention

TIA and minor stroke are cerebrovascular events which result in either very limited or no disability. Their recognition and treatment presents an opportunity to reduce subsequent death or disability due to recurrent events.

The clinical definition of TIA, used most frequently in epidemiological studies, is a focal brain or monocular dysfunction, that resolves within 24 hours, presumed to be due to reduced blood flow (13). Given that over a third of patients with clinical TIA have an infarct on magnetic resonance imaging (MRI) (14), a new imaging-based definition has been proposed (15). This adds a further specification to the TIA diagnosis, namely that imaging rules out a structural brain lesion, although adopting this now will change prevalence and subsequent recurrent stroke risk in fresh epidemiological cohorts. The established clinical definition is used in this thesis, consistent with that used in the Oxford Vascular Study (OXVASC), a prospective population-based study of all vascular events occurring in a general practice registered population of 91,000 patients (16).

Although there are a number of definitions of minor stroke without consensus in the literature (17), in this thesis, minor stroke is defined as a sudden onset persistent deficit which is mild, has an explanatory cerebral lesion and results in a National Institute of Health Stroke Scale score of less than 5 points (18).

The risk of recurrent stroke is high in the first few days and after TIA or minor stroke. Routinely collected data from the Kaiser Permanente health system in California demonstrated 10% of patients attending an emergency department (ED) after TIA represented with stroke after 90 days, and half of these strokes occurred within the first two days (19). Prospective population-based cohort data of patients presenting to general practices and EDs has similarly shown a high early risk of recurrent stroke after TIA and minor stroke (20;21).

Patients at greatest risk of recurrent stroke after TIA can be identified using clinical prediction rules. A simple rule derived from the Oxford Community Stroke Project (OCSP) data (22) and validated in OXVASC data used age, blood pressure and clinical features (ABCD) to predict stroke risk (23). To improve validation in an international cohort the presence of diabetes was added, forming the ABCD2 score (24). The ABCD and ABCD2 scores have had reports of both high utility (25-29) and low utility (30-34) for predicting recurrent stroke, but nevertheless the ABCD2 score is recommended as a stroke prediction tool in current guidelines internationally (34-37) for determining the urgency of specialist
assessment. Accuracy of predicting recurrent stroke is increased with the inclusion of brain imaging findings (38;39) together with a history of multiple TIAs and carotid imaging (40) although imaging results are not available at the first contact with healthcare services, either in primary care or in most EDs.

The importance of urgent specialist assessment after TIA and minor stroke for reducing recurrent stroke risk has been inferred from observational studies. In the Early Use of Existing Preventive Strategies for Stroke (EXPRESS) study nested within OXVASC, urgent assessment and treatment resulted in an 80% relative risk reduction in recurrent stroke at 90 days, compared with a historical cohort subjected to routine referral practice (41). Similarly, follow up of patients presenting to an urgent TIA clinic on the same day as their symptoms occurred (SOS-TIA study) showed an 80% relative risk reduction compared with predicted stroke risk from baseline ABCD2 scores (42).

Thus stroke risk after TIA (and minor stroke) is high in the first few days after events, is predictable using clinical scores and can be reduced with urgent assessment and treatment. This creates a crucial role for primary care in reducing the burden of cerebrovascular disease, by identifying and appropriately referring patients for secondary prevention, particularly as the majority of these patients initially attend primary care (43). Furthermore given the frequency with which stroke is preceded by TIA, the total burden of disease will be reduced by optimal secondary prevention after TIA - 25% of patients with stroke in treatment trials report a TIA beforehand, with almost half of these occurring in the week before the stroke (44).

1.5 Primary care and secondary prevention – the importance of delay and delayed presentation

In order to reduce recurrent stroke, primary care needs to ensure that the patient with TIA or minor stroke is recognised and assessed as soon as possible after the event. This enables risk stratification for urgency of assessment, and for TIA patients scoring 4 or more on the ABCD2 score, a specialist assessment with access to investigations and treatment should be carried out within 24 hours of symptom onset (36).

Thus primary care needs to understand how it can reduce any delay between symptom onset and specialist assessment for patients who have TIA or minor stroke. Delay from symptom onset to definitive treatment was explored by Andersen in women after a diagnosis of breast or ovarian cancer (45) and a general model of delay deduced (termed “Total Patient Delay”) that involves both patient responses and subsequent healthcare system
responses. Patient responses are characterised by processes that govern the successive time intervals:

1. appraisal delay (time from the first symptom to patient recognition that the somatic experience might be a symptom of an illness)
2. illness delay (time from recognition to the decision to seek medical attention)
3. behavioural delay (time from deciding to seek medical attention to scheduling an appointment with a clinician)

Healthcare system responses are characterised by the following successive time intervals:

1. scheduling delay (time from first call for medical attention to first clinical assessment)
2. treatment delay (time from first medical attention to definitive treatment)

This model is relevant to patients with TIA and minor stroke as the perception that a transient or mild symptom is actually due to an illness may not be straightforward. While there is poor awareness of the specific symptoms of stroke amongst the general public (46-50) this may not actually impact on appraisal delay as patients may still think they are ill, but without necessarily making the correct self-diagnosis (51), particularly if symptoms are more severe (52). Conversely patients could correctly recognise symptoms as due to stroke but not act to access emergency medical attention (53).

Given the short time window for intervention after TIA or minor stroke, it is essential to understand the patient level factors, e.g. highest education attained, deprivation, previous stroke or TIA that could influence the time from symptom onset to the first call for medical attention. This could influence the optimal public health intervention to reduce delay from symptoms to healthcare contacts. Although educational interventions have been tried previously to reduce delay from stroke symptom onset to calling for medical attention (54-56) there has been no consensus on the optimal intervention to reduce this delay.

The choice of attending primary care or ED is also important as there are inherent delays in accessing primary care due to a booked appointment system, compared to an ED which offers a walk-in service. This contributes to scheduling delay and perceived access of care has been shown to influence the choice of first interaction with healthcare (primary care or ED), particularly in the out of hours period, i.e. outside normal office hours of work (57).

Perceptions of availability of primary care may influence delay between symptom onset and first medical attention after TIA or minor stroke. If patients choose to seek healthcare from primary care rather than ED and delay seeking care due to perceived lack of availability,
then this aspect of patient behaviour therefore should be considered as a primary care related factor.

1.6 Primary care and secondary prevention – reducing treatment delay

Treatment delay is the time period from the first medical assessment to receiving definitive treatment. For patients seeking healthcare from general practice, this time period includes the following steps

1. GP recognises that TIA or minor stroke are diagnostic possibilities
2. GP initiates referral to either a TIA /Stroke rapid access clinic or to ED
3. Patient is transferred to ED or clinic makes contact with the patient to arrange a date/time for assessment
4. Patient begins new medication or is referred for carotid surgery (entailing a further delay to surgical assessment and thereafter date of operation)

There are two key roles for primary care in reducing the above treatment delay. Firstly, correct recognition of TIA or minor stroke. The false negative rate for TIA or minor stroke diagnosis in primary care is unknown i.e. patients with TIA or minor stroke who seek care from their GP but the diagnosis is not considered and they are not referred for further care. If patients are not referred at the initial consultation then this results in delay to definitive care.

Secondly, primary care referrals to TIA clinics should have a reasonable prevalence of true TIA and currently clinic audits have estimated this to be around 50% (58;59). Whilst it is neither realistic nor appropriate that GPs should have 100% specificity for any suspected diagnosis in primary care (where all referred patients for a suspected diagnosis actually have that diagnosis), if large numbers of patients without a TIA diagnosis are referred from primary care to secondary care for a TIA/stroke specialist rapid access clinic then this will saturate a fixed capacity out-patient clinic system. This in turn would make it more difficult for patients with true TIA and particularly those at estimated high risk to be seen within the urgent timeframes that are set out in the National Stroke Strategy for England and Wales (60). There are currently no reported studies of interventions to improve the positive predictive value of a GP referral to a TIA clinic, with the specific aim of reducing delay for true TIA cases and in particular high risk cases. The most recent National audit data shows that only 43% of TIA patients at high estimated stroke risk are seen within 24 hours (61). If clinic capacity were to be increased by reducing the numbers of false positive referrals then a greater proportion of high risk TIA patients could be seen within the correct urgent timeframes.
1.7 Need for a literature review

The initial observations from my professional clinical experiences together with the evidence outlined above led me to define the evidence gap by conducting a literature review. The remit of the review was to examine

1. Current knowledge about delays in presentation to healthcare after the onset of symptoms of cerebral ischaemia and determinants of delay, particularly the organisation of primary care.


3. Currently available prediction tools that could be used to aid primary care practitioners after recognising the possibility of cerebral ischaemia as a cause of symptoms i.e. to refine the population of patients that would benefit from referral to specialists.
Chapter 2  

2.1 Introduction

This chapter will review the existing evidence relevant to the following three issues; delay in accessing care after TIA or minor stroke, the ability of GPs to recognise TIA or minor stroke and diagnostic tools that could reduce inappropriate referrals from GPs to cerebrovascular specialists.

The published databases searched were the National Library of Health (Medline) and EMBASE (simultaneously using the OVID interface) and The Cochrane Library via its own website. The primary and secondary search strategies are in Appendix 2.

2.2 Delay in accessing specialist care

Most developed healthcare systems have a portal of entry to specialist care either in the community i.e. primary care, or a triage and diagnostic service in secondary care i.e. the emergency department (ED). Examining what is known about delay in accessing definitive specialist care requires analysis of the determinants of delay along these pathways. Firstly I will address the issue of delayed presentation to healthcare after acute vascular event symptom onset, as well as the determinants of the choice of first route into the healthcare system (which has speed implications in terms of specialist access). Then I will examine what is known about the pre requisites of presentation – knowledge of symptoms and of action prompted by symptoms, together with educational interventions to improve levels of such knowledge in the general population and in at risk populations. Finally I will examine reports of demographic predictors of access to primary care and whether these could explain health inequalities in stroke via the mechanism of delay in receiving specialist care.

2.3 Determinants of delay in presentation to healthcare

Delayed presentation to healthcare after the onset of illness entails a risk of a worse clinical outcome. Improving recognition of illness by reducing appraisal delay (45) and encouraging earlier contact with healthcare by reducing illness and behavioural delays (45) are important public health goals.

Delay in presentation to healthcare after the onset of symptoms has been extensively studied in progressive conditions where early intervention improves long term survival such as cancer. However, identifying demographic factors associated with delay may be specific to the disease in question. For example, demographic factors associated with appraisal delay vary with the type of cancer, although common system factors such as misdiagnosis
and delay in onward referral for definitive investigations have been identified (62). Analyses of health service data agree that demographic influence on delay varies according to diagnosis (63).

Healthcare seeking behaviour in general has been assessed with written vignettes sent to UK patients of descriptions of two scenarios – a lump in the axilla and chest pain. No effect of ethnicity, socioeconomic position or sex was observed in the responses (66% of the total sample) on the reporting of urgent health care seeking behaviour (64). This was interpreted as evidence that demographic associations with delay in providing care must be due to the healthcare system response, and not individual healthcare seeking behaviour.

2.3.1 Determinants of delay after vascular events – myocardial infarction (MI)

Determinants of delay may be also dependent on the type of healthcare system. In the United States (US), delayed presentation after acute MI has been observed in older women, patients with diabetes, African American ethnicity, patients with a lower socioeconomic position, less education and less comprehensive private insurance (65). In a privately insured payment healthcare system, certain risk factors for delayed presentation may also compound the likelihood of poor clinical outcomes – for example lower socioeconomic position (SEP) is associated with a reduced control of diabetes (66). The association between delayed presentation after MI and female sex, lower SEP and diabetes has been observed consistently in the US (67). For patients with a lower SEP, delayed presentation is partly due to concerns over cost (68).

2.3.2 Determinants of delay after vascular events – cerebrovascular disease

At the extremes of delay, we know that patients do not present at all after symptoms compatible with stroke like episodes, and US data shows that lower income, smoking and absence of previous diagnoses of stroke are associated with a complete lack of healthcare seeking (69). Furthermore for those that do present to healthcare after stroke, the size of delay appears to be stable over time. An observational study of arrival times to EDs after stroke onset compared from 1993 to 1999 showed that although more African American patients arrived within three hours, there was no change in the percentage of patients arriving within two hours (70).

In patients that do present to healthcare, observed determinants of delay vary with the healthcare setting. A greater than six hour delay in presentation after stroke in the Catalan healthcare system in Spain was associated with lower educational attainment, diabetes, lower SEP, living alone, and incorrect recognition of the cause of symptoms (71). Lower SEP
is also associated with small increases in pre-hospital assessment and transportation via emergency ambulance services in the US (72). The Copenhagen Stroke Study demonstrated associations with delay after stroke and being alone, being retired, having less severe symptoms and no prior history of TIA but no associations with other demographic or clinical features (73). However, investigating delay outside formal research studies i.e. using routine datasets, can potentially lead to errors due to inaccurate recording in medical records of stroke onset time (74).

Interviews with patients presenting after stroke have identified that the presence of bystanders adds a further influence on delay. Alongside female sex and reduced severity increasing delay, the views of others present at the time of symptoms influences the action taken by patients in the US (75). Furthermore, interviews with bystanders accompanying stroke patients in Wales found that they had variable recognition that a stroke was in progress, a minority thought that the patient’s symptoms were severe whilst some took a ‘wait and see’ approach and only took action when they thought that spontaneous resolution was unlikely (76).

US data from interviews of patients with TIA or stroke who attend hospital have indicated that using emergency services rather than primary care, or private transport direct to hospital, was associated with reduced delay as well as an urgent bystander response (77). This study also demonstrated that female sex was associated with increased delay and perceived severity reduced delay. The Delay in Accessing Stroke Healthcare Study (DASH), also in the US, found that after controlling for all associations with delay in multivariable analysis, using emergency services and bystander recognition of stroke were the only significant predictors of delay in arrival at hospital (78).

Perceived severity reducing delay also explains the findings of an audit of consecutive stroke admissions in South Korea that although a small proportion of patients arrived within three hours, those doing so had greater deficit than those arriving after three hours (79). Choosing a community based service for the first medical attention after stroke together with not perceiving symptoms as serious have also been associated with delayed presentation in Singapore and again this appears to be stable over time (80;81).

Although many interview studies of patients attending ED after a stroke have not identified specific clinical features that prompt more urgent action, other than general perception of severity, a US study specifically designed to investigate the effect of symptoms of TIA or stroke found that delay was greater for posterior circulation presentations of dizziness and difficulty in walking, compared with the largely anterior circulation symptom of speech difficulty where patients acted more quickly (82).
A smaller US study also found that the territory affected within unilateral anterior stroke presentations can influence delay with speech more likely to result in early presentation than unilateral weakness (83). The decision to seek healthcare after TIA is also likely to be influenced by the type of deficit, with a survey of patients without previous cerebrovascular disease showing that one episode of speech deficit would be sufficient to prompt a consultation whereas sensory disturbance alone would not (84). Posterior circulation symptoms may also be associated with delay from assessments of urgent action after symptoms in stroke free cohorts. In the general US population, interviews with people chosen randomly by telephone number selection found that visual disturbance was least likely to prompt an urgent response to seek healthcare (85).

### 2.4 Patient recognition of stroke symptoms and action

Qualitative data suggests that appraisal delay is significant after stroke onset (86) and this may be partly due to denial of deficit (87). This suggests that patient knowledge has a role to play in the development of delay in accessing healthcare although there are conflicting inferences from the literature. Knowledge and correct recognition of stroke have not been associated with quicker arrival times in hospital from one German study (53) but a similarly conducted study in the US did find that recognition that symptoms were due to stroke resulted in earlier presentation (88). Interviews with a small number of Korean patients after stroke have been used to build a decision tree model to explain reasons for delay, and severity rather than recognition was included in a model with modest ability to predict delay in an external validation (89).

Prior knowledge of the key features of stroke and TIA and acquired understanding of the need to take urgent action will clearly influence delay in presentation to healthcare, both by patients and bystanders. Correct recognition of stroke symptoms is low in the general community in the US (90) and in patients admitted to hospital after stroke (91), although this by itself does not imply that patients would not act in an emergency after symptom onset. Echoing the findings from patients after stroke, data from interviews in the general population in the UK suggests that perceived severity would entail urgent action (92).

Stroke knowledge may be dependent on personal exposure to the condition. In a young general population in Switzerland, stroke knowledge was increased with personal experience of a patient with stroke but not with professional healthcare status or university education (93). Samples of stroke free patients in an area of high stroke risk in North Carolina (‘The Stroke Belt’) demonstrated that stroke knowledge was related to age and ethnicity with an inverse relationship in that those at higher risk tended to have less knowledge of stroke (94).
Telephone interviews in the general population, selected from random telephone number generation, in the US have demonstrated very little improvement over time in knowledge of stroke symptoms, risk factors and the potential for urgent treatment (thrombolysis) from 1995 to 2005 (47). However, in patients with TIA or minor stroke from a Dutch study, knowledge of risk factors and their control can be much higher than knowledge of symptoms of stroke itself (95). In a German study, a large representative older general population showed that knowledge of stroke risk factors was good with the media most often cited as the source of knowledge (96).

2.5 Interventions to improve stroke knowledge and action

There have been wide variations in the scale of interventions to improve both stroke knowledge and action, although determining the enduring effect of an intervention via reducing delays in presentation is less often observed. A targeted intervention in a small rural community in the US demonstrated durable improvements in stroke knowledge but only in a subset of patients that were tested after the initial intervention (50% of the sample), but due to the small sample size and short follow up, the effect of education on action after stroke could not be tested (55). Multiple media delivery methods to the general population in North Carolina using interviews on radio and television as well as newspaper articles were tested for their effect on times to presentation for stroke in a before and after analysis (97). More patients, as a percentage of total attendances in each time period, arrived at specialist care units after stroke within 24 hours in the intervention period. The effect of media advertising on presentations of stroke and TIA in Canada suggests that a persisting presence is needed, as ED attendance dropped after an advertising ‘blackout’ (98). This suggests that one-off interventions may not have an enduring effect on the primary outcome of attending healthcare after symptom onset.

The nature of the media used to disseminate educational messages may be relevant in that messages delivered by television rather than by newspaper had greater effect on stroke knowledge, as assessed from randomly selected members of the population exposed to the interventions in the US (99). This has also been suggested by a small trial using a short animation to teach the clinical features of stroke based on weakness of the Face, Arm or deficit of Speech and Time to call for help (FAST) which showed persisting high levels of stroke knowledge but again no ability to test how this would translate into action in response to stroke symptoms (54). The potential for a campaign that solely contains the features of FAST to capture all cases of TIA and stroke was assessed in 12 months of stroke presentations to regional hospitals in Greater Cincinnati and Northern Kentucky (100). FAST
would have detected 90% of TIA and ischaemic stroke patients but fewer patients with haemorrhagic stroke (70%).

Given the US data showing that ethnic differences exist in time to presentation after TIA and Stroke and also that Afro-Caribbean people may have a higher stroke risk from epidemiological data from the South London Stroke Register independent of demographic confounders (101), an examination of ethnic differences in design and delivery of educational interventions is relevant. One US trial examining whether culturally relevant educational material may be more effective in enduring stroke knowledge is yet to report (102). However, a before and after assessment of a community based intervention involving African American beauticians educating their clients during appointments demonstrated that knowledge of stroke symptoms and the need for urgent action increased as a result (103).

Although one large scale (75,000 population base) trial of mass mailing of letters on stroke symptoms and suggested course of action resulted in more patients with stroke arriving within three hours of symptom onset (34% vs 28%, with the difference accounted for by reduced delay in women only) (104), others have found that education appears to influence knowledge with little impact on behaviour (105;106).

2.6 Determinants of delay - healthcare system factors

Delay in accessing specialist care after TIA and stroke has healthcare system factors. After negotiating an initial contact in general practice or a hospital ED, an assessment and then referral process is required to arrange specialist assessment.

2.6.1 ED and Hospital

For patients choosing ED, accessibility of acute hospitals is clearly a determinant of time to admission after stroke. An Australian study reported that location at the time of the stroke was the only predictor of delayed admission, but in rural settings geographical variables and associated travelling times have a wider absolute range (107). For patients accessing care using emergency retrieval systems, recognition of stroke is important but priority for transfer to hospital does not need to rely on this. A study of US emergency 911 call tapes found that although the call handlers did not recognise symptoms as being due to stroke, the nature of the clinical features elicited a high priority response (108).

Changes to the pathway of care in ED for patients with TIA and stroke can reduce delay in access to key investigations such as cranial imaging, a pre-requisite for safe administration of thrombolytic agents in acute ischaemic stroke. Use of ‘acute stroke calls’ to encourage a coordinated emergency response on ED arrival resulted in reduced time to CT brain and
subsequent length of stay at one Australian hospital (109). The implementation of a decision support tool in the form of a flow diagram for investigations of suspected TIA in one US hospital ED found that most (83%) patients were treated according to the pathway and a low 90 day stroke risk of 1.3% was observed on follow up (110). Speeding up access to diagnostics in ED rather than organising them during an acute hospital admission may also be more cost effective in the US context (111). A review of Australian hospitals’ care for patients with stroke and TIA suggests that where stroke services are organised, there is improved care for patients with TIA (112).

For patients with cerebrovascular events occurring during a hospital stay, there is considerable variation in response depending on the clinical specialty environment in which they are nursed. Delay in assessment and treatment was higher on general medical and surgical wards compared with neurology wards in a retrospective analysis of inpatients at two academic secondary care centres in the US (113).

A distinction between routine working hours and the ‘out of hours’ period is evident in the majority of hospitals for most services, apart from ED and intensive care environments. It is therefore plausible that an effect of the time of stroke and the time of presentation to hospital may affect access to specialist care. No effect of ‘in hours’ compared with ‘out of hours’ presentations in thrombolysis rates or clinical outcome was seen in a large prospectively collected stroke register in Germany (114). However, small differences in stroke mortality have been estimated from large national cohorts in the US (115) and Canada (116) with weekends and evening presentations associated with a greater risk of death.

For outpatient management of TIA and minor stroke, clinic frequency is a major determinant of delay in specialist assessment. A UK weekly clinic was associated with a median delay of 16 days from initial event to specialist assessment (117). Outpatient TIA clinics in New Zealand also show delays of greater than one week again related to frequency of provision (118).

2.62 Primary Care

For patients presenting in primary care, access to a doctor is a key step in the onward referral for specialist assessment. GP appointments, particularly for same day requests for care, are arranged by communication with a practice receptionist in the UK. This in itself is not straightforward, as making an appointment in primary care is a complex social process and therefore subject to multiple patient, receptionist and practice factors (119). The ability of practice receptionists to respond to descriptions of symptoms and arrange an appropriate appointment can influence delay to definitive care. A Welsh study of 45 general practices
found that the response of many practice receptionists to telephone descriptions of amaurosis fugax and cortical TIA was a routine or urgent GP appointment with a mean time to being seen of two days (120).

System delays can exist in accessing specialist care depending on the referral routes that GPs choose after suspecting TIA or stroke, which is partly determined by local availability of services. A questionnaire study of German primary care providers (the majority were GPs with others being community based internal medicine specialists) reported that out patient management may be recommended for stroke, depending on their perception of the therapeutic benefit of admission (121). Interestingly, poorer outcomes were seen for patients admitted with stroke in Scotland compared to those managed in the community (122) although examining the effect of place of care on outcomes for acute stroke is likely to be subject to considerable confounding by indication as referral choice will be strongly influenced by severity at first clinical assessment.

2.7 Determinants of access to Primary Care – patient choice and provision from health policy

The determinants of primary care usage for acute onset illness in general are relevant to the factors which influence speed of access of primary care after TIA.

Improving access to primary care was one of the major health policy objectives of the 1997 – 2009 Labour government (123). Specifically practices needed to be able to devote sufficient resource to be able to offer an appointment within 48 hours of a patient request. Given the need for rapid assessment after TIA, the effect of this policy on requests for same day care is important, as practices could achieve this target of 48 hours at the expense of same day appointments. One questionnaire study of practices who did and did not adopt the ‘Advanced Access’ policy found that more same day appointments were available as a result of the policy (124). Although a same day appointment may not be with the doctor that a patient would ordinarily choose to consult, discrete choice experiments of patient opinion about primary care access have shown that waiting time is a priority over choice of GP for new health problems (125), suggesting that for acute presentations after TIA, patients will see whichever GP has the earliest available appointment. In a questionnaire satisfaction study, patients were not more satisfied with the Advanced Access system as there is little flexibility over the day on which patients can be seen (126).

A further Department of Health policy to increase access to primary care was the establishment of ‘NHS walk in centres’ which offered seven days a week care and 8am to 8pm opening hours (127). These were piloted as nurse-led and offering mainly treatment for
minor illness. An analysis of the first wave of such centres using a before and after time series analysis did not show significantly altered ED and GP use within a three km radius of the centre or any change in out of hours (OOH) primary care usage (128). Thus primary care consulting for registered patients is unlikely to have changed significantly as a result of walk in centres.

2.7.1 Primary care in the out of hours (OOH) period

For patients choosing primary care for their first healthcare contact after TIA, the effect of whether events occur during routine hours or in the OOH period is a potentially important determinant of delay. A patient’s own GP is an efficient provider of OOH care - a randomised controlled trial of OOH care found that usual GPs made fewer home visits, spent less time on home visits when they were required and issued fewer prescriptions compared with a deputising service (129). Until the new general medical services (GMS) contract in 2004 patients were calling their own practice for healthcare in the OOH period (130). The likelihood of accessing OOH primary care before the contract change, i.e. one’s own GP, was higher with increasing deprivation (131). However, this is not because more deprived patients seek healthcare in primary care rather than ED in the OOH period, as high OOH primary care usage and high ED usage occur in parallel in more deprived areas (132).

Age and illness severity are important factors in choice of healthcare provider in the OOH period. A study of all emergency contacts within a population in the OOH period shows that patients aged 60 -80 have a higher rate of GP consultation than ED consultation whereas younger adults have similar ED and GP consulting rates (133). Furthermore the range of conditions seen suggests a different population of patients choose ED and primary care in the OOH period in a Dutch study (134). This is echoed by a retrospective time series analysis of the effect of introducing an OOH GP where none existed previously on ED presentations in Australia – a gradual effect was observed with a reduction in non urgent cases and an increase in urgent cases suggesting that patients may self-select a service based on severity of illness (135). Patients who are more anxious about their symptoms choose ED rather than primary care (136) and anxiety has some predictive value for vascular disease having been included in a diagnostic clinical prediction rule for patients presenting to primary care with chest pain (137).

Nevertheless, the decision to use OOH services also depends on the opinion of the service available and the patient’s proximity to ED (138), although qualitative interview data suggests that significant uncertainty remains over whether OOH calls are appropriate (139).
Choosing primary care or ED after a TIA in the out of hours period

A previous report on the distribution of diseases among patients using ED or a GP in the OOH period have used population based time defined cohorts, and TIA as a specific condition was too infrequent to appear as category on its own in either setting (140). Again, healthcare system specific variation may be present as the choice of OOH primary care or ED after TIA was found to be very different in an international comparison with the majority of patients attending an ED in Canada and a GP in the UK (141). The majority of patients with stroke before the change in GMS contract also contacted primary care rather than emergency services in the first instance, in spite of accurate recognition of stroke (142). There are significant consequences of choosing primary care rather than ED after symptom onset for stroke in that the median times to assessment are well beyond the extended window for thrombolysis for stroke (143).

Socio-economic position and stroke outcomes – measurement and presentation delay as a confounder

Stroke mortality has reduced in recent years in most European countries but the reductions in the lowest income groups may not be as great (144). Delay in accessing stroke care has been associated with lower SEP (2.3.1) and lower SEP is also associated with a more severe stroke deficit at initial assessment as well as stroke at a younger age from a contemporary hospital cohort from Scotland (145). US data suggests that lower SEP increases stroke mortality independent of age, sex and stroke severity (146), potentially due to differences in hospital treatment rather than initial stroke deficit (147)- the markers of SEP that predicted stroke outcomes were occupation and household income but not educational level. Income not only predicts mortality but may also be the best predictor for stroke risk. Data from New Zealand found that household income was the strongest socio-economic predictor of stroke risk from case control methodology (148). Given the poorer stroke outcomes after delayed access to care, there is a potential for a confounding relationship to exist between SEP and both delay and poorer outcome from stroke.

Follow up of cohorts who are stroke-free at baseline demonstrates that stroke incidence is higher with lower SEP (149-152) and the degree of inequality in stroke risk among middle aged men varies by country (153). The risk of recurrent stroke is also affected by SEP and this effect may be greater for women than for men (154). Furthermore a review of blood pressure distribution across levels of SEP shows that there is higher blood pressure in patients with lower SEP, greater for women than men, and this is mostly explained by association between body mass index and SEP (155).
There is some variation in observations of the increased stroke risk and stroke mortality with lower SEP. UK data has not demonstrated an association between lower SEP and case fatality after stroke but it is associated with poorer control of vascular risk factors and stroke at younger age (156). In the French population-based Dijon stroke study, higher stroke incidence in a locality was associated with the degree of income inequality in that locality as well as the wealth markers of car ownership and employment. These relationships were stronger in women, and no marker of SEP appeared to predict stroke incidence in men overall (157).

2.8.1 Reversal of health inequality in stroke with ageing

In the very elderly in the US, the association of income and educational level with stroke incidence was reversed over the age of 75 years, with stroke associated with higher income and educational level, and the significance of this association was not reduced after controlling for race, risk factor control, presence of a social network or depression (158). A similar methodology in a larger study of an elderly stroke free cohort at baseline found that strokes occurring after the age of 65 are not associated with lower education and income (159), and has been confirmed in a further multicentre study in France where over the age of 65 years, higher rather than lower income is associated with stroke (160). UK data concurs with the reverse effect of area deprivation on stroke risk in the very elderly (161). In the elderly, marital status and sex are confounded with financial resources in that being widowed or never married confers higher stroke risk but this risk is lessened after adjustment for income (162).

2.8.2 The influence of socioeconomic position – childhood or adulthood?

However, the above studies used markers of current SEP to measure deprivation e.g. current household income or area measures of deprivation. There may be differential effects of deprivation in childhood compared with deprivation experienced as an adult on cardiovascular event risk. Childhood SEP has a stronger association with stroke mortality than adult measures of SEP (163) but this is not the case for coronary heart disease (164;165), suggesting that there is not a generic vascular disease effect.

Childhood SEP effects may explain other observed demographic associations with stroke risk. In US data, childhood SEP predicted 10 year stroke risk in African Americans and Caucasian Americans with a higher stroke incidence seen in African Americans but, after adjustment for both childhood and adult SEP, the racial differential was minimal (166). Father’s occupational class is a strong determinant of stroke risk and there is also a strong
legacy effect in that their upwardly mobile offspring who have non manual jobs have the same stroke risk as men in manual jobs (167).

2.8.3 Socio-economic inequality in stroke outcomes – unequal treatment by the health system?

Hospital care may exacerbate inequality. Higher SEP is associated with a greater chance of post stroke cerebral imaging in a contemporary cohort of 200,000 patients in English hospitals (168), however no consistent pattern of SEP and access to care is evident across all components of stroke care (169). There are higher rates of symptomatic carotid disease in patients with lower SEP but no associated increase in carotid endarterectomies in these patients (170). However, no effect of SEP was seen on delay in referral to a TIA clinic in Scotland (171).

Management in primary care can result in inequalities in stroke outcomes. Post stroke management in primary care in the Netherlands demonstrates an effect of SEP on control of diabetes and BP for women but not for men (172). In the UK prior to the National Stroke Strategy, the major determinant of receiving secondary prevention post stroke was age, with the oldest patients who are at highest risk of events receiving the least amount of secondary prevention (173). After the introduction of the pay for performance contract which offers incentive payments for undertaking annual reviews in primary care for patients with stroke, patients with a lower SEP were found to be less likely to attend and be ‘exception reported’ (174).

2.8.4 Socio-economic position and cohort studies – measurement and representativeness

There are different approaches to forming groups with similar levels of deprivation in published stroke and TIA research. Some use absolute values of income irrespective of the proportion of events contained within those groups, particularly in large community based studies (160) or large hospital cohorts (146), whereas some use fixed categories of educational attainment or occupation (147;150). The effect of deprivation on stroke outcomes was tested by grouping postcode area deprivation into quartiles in the NEMESIS stroke incidence study, although it did not recruit patients with TIA (175).

One hospital based study of stroke risk after TIA identified patients from health provider/insurance records without an assessment of national representativeness (19) although regional representativeness was analysed in the health plan users prior to data collection (176). Among community based TIA studies, The North Dublin TIA study has not
reported national representativeness (177) and neither has the Perth Stroke Study which included patients with TIA (178).

In general community based TIA and stroke studies have not assessed national representativeness. This is relevant in that under-sampling of deprivation in a cohort may result in a type 2 error where lack of influence of deprivation may be concluded even though it may exist within the population at national level.

2.9 Diagnosis of TIA and stroke

Diagnosing cerebrovascular events is a challenge in primary care and ED, particularly for TIA where typically there are no persisting neurological deficits at the time of presentation, and diagnosis relies on history alone. Diagnostic research in the field of TIA suffers from a lack of a gold standard for diagnosis (15). This is less problematic for minor stroke as there must be imaging abnormalities of infarction for a diagnosis along with persisting neurological signs. Although some TIAs are associated with abnormalities on diffusion weighted magnetic resonance imaging (DW-MRI) (179-181), it is not a necessary condition for diagnosis. TIA patients with normal DW-MRI are seen and have a higher risk of recurrent events compared to a background population, indicating that clinical diagnoses without neuroimaging findings are still likely to be correct (182).

2.9.1 Specialist disagreement and TIA diagnosis

The diagnosis of TIA is a complex process, relying on an interpretation of the history given by the patient or a bystander. There are no diagnostic tests from blood or cranial imaging that take the place of a diagnosis from a description of the transient event. Although consensus statements exist on the definition of TIA (15), there does not seem to be the same specialist consensus over who has actually had a TIA. The likelihood of a TIA diagnosis among 55 patients referred to a TIA clinic in Stanford, USA was assessed independently by three vascular neurologists using case records (183). Differently graded rating scales were used to capture likelihood and unlikelihood of TIA, and although the neurologists individually showed some degree of category preference, there was most disagreement over the category of ‘unlikely TIA’ (agreement coefficient of 0.16).

A Dutch study assessed agreement over TIA diagnosis but allowed the neurologists to take histories from the patients and 56 patients were seen by two clinicians independently (184), which showed higher agreement than the US clinic note based study (Cohen’s kappa = 0.65). The Dutch study was later repeated using slightly different methodology with greater
agreement found if symptoms were initially recorded using basic clinical description rather than diagnostic terms (185).

Furthermore, specialist disagreement is not limited to diagnosis in cerebrovascular disease. Specialists can disagree over the severity of neurovascular deficit, particularly for aphasia and facial palsy (186), as well as in determining the affected vascular territory from clinical features of the cerebrovascular event (187).

The generalist and specialist may have different concepts of TIA as a disease entity. Spanish non-neurologists and neurologists include different symptoms that can be caused by TIA and this contributes to the poor predictive value of non-neurologists’ suspected diagnosis of TIA (188).

2.9.2 Symptom questionnaires and the value of self-reporting by patients

Questionnaires have assessed patients’ reporting of previous TIA and stroke as well as symptoms that could be due to TIA and stroke. In the Norwegian Tromso study, patients’ self reported history of stroke had a high predictive value of 75% for an actual recorded stroke event (189) but UK data suggests that there is significant ‘over recall’ i.e. false positives for stroke diagnosis together with under recall i.e. false negative recall (190).

Questionnaires that aim to assess the burden of cerebrovascular disease and unmet need in TIA and stroke care assess the prevalence of symptoms compatible with TIA and stroke in stroke-free cohorts. In the Reasons for Geographical and Racial Differences in Stroke (REGARDS) study, recruited subjects with a mean age of 65 years and no previous presentations to healthcare with TIA or stroke had a 20% prevalence of symptoms of TIA/stroke (191). However, self report of TIA symptoms depends on the prevailing concept of TIA at the time of study design and one UK study included loss of consciousness (192) which is no longer considered to be due to TIA. Nevertheless, using questionnaire based symptom prevalence studies the number of patients with un-assessed TIA in the US has been estimated as greater than five million (193).

2.9.3 Questionnaires and TIA diagnosis

The diagnostic performance (compared with clinician diagnosis) of questionnaires for TIA vary with the population tested. The Questionnaire to Verify Stroke Free Status (QVSFS) has been validated in stroke and general medical out-patient clinics with varying reports of 96% negative predictive value (NPV), 71% positive predictive value (PPV) (194) and 100% NPV and 36% PPV respectively (195). A six item derivative of the QVSFS which only contains questions about clinical symptoms was used in Veterans Affairs clinics in the US
with reported sensitivity of 82% and specificity of 62% in a population of patients with a prevalence of stroke of 50%, which equates to a PPV of 68% and an NPV of 78% (196). These results suggest that a rule out function is more reliable than making positive diagnoses.

In the Asymptomatic Carotid Atherosclerosis Study (ACAS), an algorithm for TIA diagnosis based on questionnaire responses in a selected population of patients presenting with TIA or stroke like symptoms, had an 88% sensitivity and 71% specificity for cerebrovascular diagnoses compared with an expert panel for patients reporting stroke-like episodes within the trial (197).

Reliance on questionnaires may result in a bias towards diagnosing anterior rather than posterior events. A diagnostic algorithm in the Atherosclerosis Risk in Communities (ARIC) study was more likely to diagnose self reported symptoms as TIA if they were anterior rather than posterior circulation (198). Nevertheless, the ARIC algorithm estimated a community prevalence of 5-6% of TIA/Stroke (199) and there is face validity to the algorithm in that those diagnosed with TIA had higher risk of stroke over a subsequent 11 year follow up period (200).

A TIA questionnaire by Wilkinson in 10,000 US care home residents showed that 6.4% had suggestive symptoms of TIA and 15.4% had possible symptoms but in a sample of 1700 respondents, the PPV of the questionnaire for TIA diagnosis by a neurologist was 7% (201). A nurse administered version of this questionnaire in a population of patients referred after suspected TIA to a UK hospital clinic had a positive predictive value of 66% for combined TIA and stroke diagnoses reflecting the much higher prevalence for TIA and stroke compared with the care home population (202).

2.9.4 GPs’ ability to diagnose TIA and stroke – responses to case vignettes

There are very few reports of studies that assess how GPs diagnose cerebrovascular events. Questionnaire studies using case vignettes have been used to test GPs’ recognition of TIA and subsequent management either with free text without cueing TIA, or with a restricted set of diagnoses which automatically cues TIA as a diagnostic possibility. All studies have used artificially constructed cases and are restricted to anterior circulation presentations.

Explicitly varying a parameter in case vignettes has been used to test GPs’ ability to suspect TIA and the influences on decisions to refer for assessment. Quik-van Milligen constructed five matched pairs of cases, varying one aspect of the history in each pair. Anterior
circulation symptoms were used, which was justified by the authors from the low prevalence of posterior circulation presentations (203). The five history parameters varied were; age (dichotomised to <65 or > 65), single vs multiple episodes, history of non-specific symptoms vs no history, brief (few minutes) vs longer (20-24 hrs) duration and cortical vs retinal pattern of symptoms. The cases were mixed with ten distractor cases (five had neurological disease and five had infectious disease) and sent to Dutch GPs, with 376 responding (59% of the sample). The authors state that cases were placed in a random order, although it is not clear if the order was at random for each questionnaire i.e. individual questionnaires had a different presentation order. The GPs were asked “What is your diagnosis?” with free text response and “What are you going to do?” with a restricted choice of suggested options including reassurance, investigations, new prescriptions and referral.

They found that GPs were more likely to diagnose TIA for cortical rather than retinal patterns of symptoms, younger rather than older age (contrary to the prevalence rate (16)), recurrent rather than single attacks, and an absence of non-specific symptoms. Referral to a neurologist was more likely for cortical rather than retinal symptom patterns, although overall there were more referrals to specialists in the retinal vignettes if ophthalmology referrals are included. However this study was reported in 1992 and was therefore carried out before guidelines and national strategies were in place for optimal evidence-based management for TIA.

Permutations in multiple parameters within a single case vignette have been used to assess the features that influence diagnosis and referral for one presentation of anterior circulation TIA - transient monocular visual disturbance (204). In this study of the responses of 866 Dutch GPs (54% of their sample), one TIA case was used from a bank of 16 cases created from two possibilities for each of four characteristics; blurring of vision vs complete visual loss, all of the visual field affected vs part of the visual field, duration of a few minutes vs a few hours and the patient did or did not report covering each eye in turn during the episode. The age was kept at 56 years, a suggested 'neutral' age which would encourage the GPs to focus on clinical presentation as the basis for decision making. Only one of the possible 16 cases was sent to each GP with three distractor cases included (lumbosacral radiculopathy, syncope and polyneuropathy).

The range of GP responses were from fixed lists for diagnosis and for management. The likelihood of each diagnosis of retinal migraine, optic neuritis, glaucoma, retinal detachment, amaurosis fugax, cortical TIA or other was rated visually by placing a mark on a line from ‘very improbable’ to ‘highly probable’. Offered management options included reassurance, medication (free text choice) and referral (specifying routine or urgent and to whom) and the
GPs rated how strongly they agreed with each suggested course of action by marking a line from ‘agree’ to ‘not agree’. Amaurosis was considered more likely for brief, visual loss (rather than blurring) affecting the complete field of vision. Performance of a cover test by the patient did not affect the perceived likelihood of the diagnosis. Of those diagnosing amaurosis or cortical TIA, 72% and 64% respectively recommended specialist referral but again, this study was reported in 1999 before recent advances in evidence based benefit, and timing, of treatment.

A study of Polish GPs' examined whether age or affected territory (within the carotid distribution) influenced correct diagnosis or referral decisions using six carotid territory TIA vignettes (205). The study authors again justified the restriction to the anterior circulation by a low prevalence of posterior circulation symptoms. The three pairs of vignettes consisted of firstly keeping age constant but varying cortical or retinal ischaemia, the second pair consisted of retinal symptoms with age variation (above and below 65 years) and the third pair consisted of cortical symptoms with similar age variation. Each case also contained information about previous episodes and a history of non-specific symptoms, similar to the Quik-van Milligen study (203). Diagnostic responses were free text but management responses were semi-structured with GPs indicating whether they would perform additional tests, initiate medication or refer for specialist assessment.

In 89 respondents (89% of the sample), correct diagnosis of both cortical TIA and amaurosis was less likely with recurrent episodes and a history of non-specific symptoms. For the cases without a history of non-specific symptoms and without previous episodes, younger age reduced the likelihood of a correct diagnosis of cortical TIA but age did not affect the correct diagnosis of amaurosis. In general the likelihood of correct diagnosis was higher for cortical TIA than for amaurosis, a similar finding to the Quik-van Milligen study (203). Although they do not present total referrals for each case, age did not affect referral to neurologists after cortical TIA or amaurosis. In the direct comparison between cortical and retinal territory, more patients with cortical symptoms were referred to a neurologist and more patients with retinal symptoms were referred to an ophthalmologist.

2.9.5 Recognition tools for diagnosis of cerebrovascular disease – easier for stroke than for TIA

Validated tools for diagnosis of cerebrovascular disease exist for stroke, and the original motivation for their derivation was earlier transport of patients to hospital from the community i.e. to aid paramedics in their recognition of stroke. Stroke recognition tools are centred on physical signs with key features from the history which identify stroke mimics. The absence
of physical signs in TIA (by definition) contributes to the difficulty in deriving simple diagnostic tools.

The Cincinnati Prehospital Stroke Scale (CPSS) is a three item scale scoring arm drift, facial droop and speech clarity and in a small sample of paramedics and ED support workers it showed good reproducibility (206). However, a prospective assessment of its use via a one hour training session in a 12 month before and after study showed no effect on recognition of stroke or time spent ‘on scene’ (207).

The Melbourne Ambulance Stroke Screen (MASS) involves initially testing for facial droop, hand grip (rather than the arm drift of CPSS) and speech disturbance (208). If any of these features are present, then the rule out features of age <45 years, history of epilepsy, and blood glucose levels (for hypoglycaemia) are assessed as well as the then thrombolysis ineligibility criterion of being bedridden or chair-bound. The MASS improved paramedic sensitivity in detecting stroke (78% to 94%) and reduced time to medical review in the ED by notification of potential stroke before arrival. This improvement was sustained three years after city wide implementation in terms of sensitivity of diagnosis and specificity (209). In this study, CPSS was calculated retrospectively from case notes and both MASS and CPSS had high NPVs around 95% with PPVs of 56% for CPSS and 65% for MASS.

The Los Angeles Prehospital Stroke Screen (LAPSS) starts with the rule out features of history of epilepsy, age <45, abnormal blood glucose together with the thrombolysis ineligibility features of being bedridden or chair-bound and symptom duration of 24 hours or more. If patients ‘pass’ this test then three motor tests are used – unilateral facial droop, arm weakness and grip strength (i.e. no speech component compared with the CPSS and MASS) (210). The LAPSS was tested retrospectively using clinical data for patients entered into acute stroke trials with a high sensitivity for stroke detection of 93% (210). A prospective validation of LAPSS used by US paramedics on 206 eligible emergency transfers to hospital showed a PPV of 86% and an NPV of 98% for a diagnosis of stroke (211).

The Face, Arm, Speech Test (FAST) uses the presence of unilateral facial weakness or arm weakness or speech disturbance as a stroke recognition tool (212). It differs from the CPSS as the examiner assesses speech by observation rather than asking the patient to repeat a sentence. It was validated by comparing paramedics assessment of the three clinical features with those of an admitting stroke physician in one UK hospital (213). 78% of admitted stroke or TIA patients had FAST signs and the strongest level of agreement between paramedics and stroke physicians was for arm weakness (Cohen’s kappa of 0.8) and slightly weaker agreement for the presence of speech disturbance (kappa = 0.7) (213). Consecutive suspected stroke referrals from paramedics using FAST as a recognition tool
were compared with ED and GP referrals using clinical judgement, and this demonstrated similar PPVs for stroke across the three referral routes (214). The FAST recognition tool has a clear bias towards anterior circulation stroke detection as evidenced by fewer posterior circulation strokes referred from the paramedics compared with ED and GPs (214).

A tool designed to improve the referral of patients with stroke from ED to acute stroke teams rather than from paramedics assessing patients in the community uses a combination of history, clinical features and blood glucose. The Recognition of Stroke in the Emergency Room (ROSIER) scale starts with checking blood glucose and correcting hypoglycaemia if present, then runs through negative predictors for stroke – loss of consciousness/syncope, ictal activity (which score -1 if present, 0 if absent). Positive features of stroke score 1 point each – face, arm or leg weakness, speech deficit and visual field defect (215). A prospective validation phase of the ROSIER scale showed a 93% sensitivity and 83% specificity for stroke diagnosis with PPV of 90% and NPV of 88% (215). These measures of diagnostic performance were higher than for FAST, CPSS and LAPSS which were also calculated in the same cohort of patients. In a high prevalence setting (consecutive admissions to an acute stroke unit), nurses using the ROSIER scale had an equivalent sensitivity and positive predictive value as doctors using clinical acumen for stroke diagnosis (216). A small validation study in an external setting showed a high PPV (94%) for stroke diagnosis (217).

Other studies have examined predictors for stroke or TIA diagnosis rather than deriving a scale. One study of ED presentations with dizziness, vertigo or postural imbalance found that positive predictors of TIA or stroke were higher age and male sex and negative predictors were isolated symptoms and non-hispanic or white ethnicity (218), although the methods were weak as there was no follow up and diagnosis was from record review. Distinguishing TIA from non-TIA in the ED is as difficult as in primary care with a similar 50% prevalence of true TIA in consecutive cases with suspected TIA (219).

However, predictors of true TIA in the ED population may be different to primary care attenders, as a study deriving a logistic regression model for TIA diagnosis from ED referrals found that significant discriminators were gradual onset symptoms, non-specific complaints and a history of prior transient neurological symptoms resulting in an area under the receiver operating characteristic curve of 0.79 (220). The ABCD2 score did not provide any discriminating utility in the study.

Features that predict stroke in a cohort of admitted patients in a UK study were absence of a history of dementia, no signs of systemic illness, any focal neurological sign, an exact time of onset, abnormal vascular findings, clinical features sufficient for an Oxford Community Stroke Project classification (total or partial anterior circulation, lacunar, posterior circulation),
a deficit sufficient for a National Institute of Health Stroke score >1 and neurological signs consistent with a unilateral cerebral lesion (221). Given the degree of expertise required to elicit these features, it would not be possible to use them to derive a scale for a generalist or non-clinician.

The only specific tool for TIA diagnosis in the literature was derived using multivariable logistic regression from routinely recorded secondary care clinical notes (222). Given that by definition a patient with TIA has non-persisting symptoms, the tool is based on the clinical history. A score weighted using the beta coefficients of the regression model consisted of negative points for headache, loss of consciousness/syncope and ictal features, positive points for diplopia, past history of stroke/TIA, unilateral face/limb weakness, speech disturbance and increasing age. Internal validation performed with a 2:1 cost ratio used for misclassifying TIA as non-TIA was used to derive an optimal cut point for clinical usage (222).

The ABCD2 score, although designed for prognostic use (24), may have a role in TIA diagnosis. It has a higher score in patients with an ED physician diagnosis of TIA who are deemed to have had a true TIA by case record review from a neurologist, compared to those patients deemed not to have had a TIA (223). A hospital clinician calculated ABCD2 score has been used to distinguish between true TIA and non-TIA in patients referred by GPs and ED physicians to TIA clinics. One study reported an area under the receiver operator characteristic curve of 0.75 (224) and another of 0.68 (177) for TIA diagnosis using ABCD2.

Other potential tools include blood tests and machine learning algorithms. A search for novel biomarkers using mass spectrometry, by examining patterns caused by combinations of proteins in serum and comparing these in patients with stroke and controls has shown a 77.5% sensitivity and 72% specificity in a small number of matched patients (225). The diagnosis of TIA using neural networks which take an alternative prediction methodology using existing history and examination factors to generate a leaning network of associations has not been validated in external cohorts (226).

2.9.6 Potential impact of GP-specialist difference in ABCD2 score

The England and Wales National Stroke Strategy (60) and NICE guidance (36) recommend that the ABCD2 risk prediction score is used at the first point of contact with healthcare services after TIA or minor stroke in order to predict the risk of recurrent stroke. However the ABCD score in its original form was derived from clinical histories taken by secondary care clinicians(23), and in further iterations as ABCD2(24), ABCD2-I(38), ABCD3 and ABCD3-
(40), only secondary care taken histories or investigations (the “I”) have been used as derivation and validation datasets.

Restricting derivation and validation datasets to secondary care histories is entirely appropriate if that is the context in which the prediction tool is to be used. For ABCD2-I and ABCD3-I, at least in the UK, only secondary care clinicians will have access to the appropriate investigations to use these iterations of the ABCD score. Therefore predicting the risk of recurrent stroke after TIA or minor stroke with ABCD2-I or ABCD3-I will only be undertaken in hospital in which case deriving and validating scores from histories taken by hospital clinicians is appropriate.

However, throughout North America (15), Europe, Australia (35) and New Zealand (37), the ABCD2 score is recommended for use at the first healthcare contact after TIA which typically takes place either in the ED or in primary care. To date no studies have reported on the validation of the ABCD2 score either from ED clinical assessments or from primary care assessments of patients after TIA.

Very little published work has examined the agreement among different clinicians (either intra-specialty or inter-specialty) of the ABCD2 score. One small retrospective study using ABCD referral proformas compared all referrals (50% were primary care, 50% ED) and only examined agreement in ABCD score for true TIA diagnoses, finding significant disagreement between ‘non-stroke specialists’ and ‘vascular neurologists’ (227). Interestingly this study also found disagreement between neurology registrars and consultants, although this was less marked than for the non-stroke specialists.

2.9.6 Arterial territory – a determinant of patient presentation and professional recognition?

Cerebral localisation of function results in different clinical presentations according to affected arterial territory for ischaemic lesions as well as for the direct tissue damage from haemorrhage. The structure of arterial inflow to the brain and brainstem creates the major categories of anterior and posterior lesions, and left and right sided lesions. An analysis of a German stroke registry found marked differences between left and right hemisphere strokes, with larger numbers of TIA and ischaemic stroke patients with symptoms attributable to left cerebral lesions compared with right sided lesions (228). However, there was an equal right–left distribution for haemorrhagic lesions in the registry suggesting that vascular pathology per se should not have any bias to affect one side or another. Given the other features that may make haemorrhage recognisable to patients and lead to presentation to healthcare such as headache, it may be the case that dominant sided ischaemic lesions i.e. left for the
majority, are more noticeable as dominant hand and speech are more likely to be affected. Furthermore, anterior lesions may be more easily diagnosed than posterior lesions. The positive predictive value of referral was much higher in a specific anterior circulation TIA clinic (86%) than the general TIA clinic figure of 50% (229).

2.10 Summary

The determinants of delayed presentation to healthcare after TIA and stroke and the choice of initial healthcare provider may vary with the healthcare system. Although delay is related to lower socioeconomic position and worse clinical outcomes, it is not likely that the associations between deprivation and stroke outcomes are mediated via delay in access to care – the relationship is complex with childhood and adulthood socioeconomic position having varying effects as well as varying associations with age at the time of stroke.

Stroke knowledge and action (by the patient and bystander) are different and can be influenced differentially even if both are the focus of the same educational intervention. The action of accessing healthcare is related to perceived severity and the decision to access care in the out of hours time period may be different from within routine hours.

Healthcare delivery influences the delay in receiving secondary prevention after TIA and stroke, with both primary care and secondary care influencing delay either by patient perception of availability, referral mechanisms or secondary care provision.

Limited data is available to assess how GPs diagnose TIA with methodological flaws in existing studies from use of artificial cases, non-randomisation of case order on questionnaires, cueing as to the potential for a TIA diagnosis and restriction to anterior circulation syndromes. Excluding posterior circulation presentations is not logical as the risk of recurrent stroke is as high as after anterior circulation events (230). Recognition tools are well developed for stroke but not for TIA, and none are based on primary care recorded data and consequently a tool to improve referral decisions does not yet exist.
Chapter 3  The Evidence Gap

3.1  Introduction

The literature review has identified a number of evidence gaps that are relevant to understanding how primary care could reduce delay from symptoms to treatment for patients with cerebrovascular events that confer a risk of major stroke.

3.2  The Oxford Vascular Study - use of cohort studies

The accuracy with which cohort studies in general, and OXVASC in particular, can represent their national population has not been established. This is relevant as I wish to infer generalisable results by examining the patient pathway to definitive healthcare interventions for patients with cerebrovascular disease. The demographic variables of age and deprivation are strongly associated with the incidence of cerebrovascular events and patterns of healthcare system use. If there are differences in both age and deprivation structure between the OXVASC population and the national population then these must be taken into account when assessing the implications of results.

Cerebrovascular disease cohorts have been identified either by presentation to secondary care or by following up a population of patients defined by contiguous geography. Studies which recruit patients at general practice level rather than through secondary care may be subject to clustering effects specific to individual practices. A cohort defined by registration at a general practice may therefore reflect specific local geographical features that may also influence generalisability of results. Failure to present to healthcare after cerebrovascular events has been associated with lower SEP and therefore large differences between practice populations in SEP may result in differences in presentation of TIA or minor stroke. Younger practice populations will have a lower incidence of TIA and minor stroke, thereby reducing clinical exposure to minor cerebrovascular presentations in primary care which could affect recognition by clinicians.

Research Questions

1. Is the OXVASC population representative of the national population in terms of age and deprivation?
2. Do differences in age and deprivation between OXVASC practices affect the presentation of patients with TIA to healthcare?
Research question 1 was answered by comparing the age and deprivation structure of the OXVASC population with that of the population for England, using data available in the public domain.

Research question 2 was answered by firstly establishing the variation in age and deprivation structure of individual practice populations in OXVASC and then examining the extent to which the structure of the registered population influenced presentation of TIA to healthcare. TIA presentation was assessed using the ratio of TIA to major stroke at practice level and was analysed with Poisson regression and graphical display.

3.3 Initial Interface between patient and the healthcare system

3.3.1 Determinants of access to healthcare

The influences on the two core components of healthcare access, the choice of healthcare provider and the time to call for medical attention and have not been determined in a population based prospective study of patients with TIA and minor stroke. This is relevant in understanding how to reduce delays in the care pathway from symptom onset to definitive interventions. Community ascertained cohorts of patients with TIA or minor stroke that have examined delay have used calendar day rather than analysing delay to call as a continuous variable and this is a methodological weakness. The timing of symptom onset of TIA or minor stroke has not been previously analysed as a determinant of healthcare access. Previously identified predictors of calendar day delay (clinical features, deprivation and previous cerebrovascular disease) have not been examined with delay as a continuous variable from symptom onset to time to call for medical attention.

The effect of the organisation of primary care on delay in presentation after non-disabling cerebrovascular events has not been assessed. There is no evidence concerning how changes in primary care delivery influence where patients seek care or how long they take to make initial contact. The change in the general practice contract in 2004 allows for such an analysis, as a very well publicised and closely debated change in out of hours provision occurred. Given that delays in definitive care may vary depending on whether ED or primary care is the first contact after symptoms, the key outcomes to be studied are choice of healthcare provider and the delay in making the initial contact.

Research Questions

3. Is healthcare access after TIA or minor stroke influenced by the following factors; choice of healthcare provider, time of symptom onset, clinical features of the event, deprivation and previous history of cerebrovascular disease?
4. Did the change in general practitioner contract influence healthcare access after TIA or minor stroke?

Research question 3 was answered by firstly establishing the choice of healthcare provider and the delay between symptom onset and calling that provider for medical attention, and then determining differences in these outcomes between groups defined by the identified predictors in the OXVASC cohort.

Research question 4 was answered with a before and after analysis of choice of provider and delay to call during equivalent time periods pre and post the GMS contract change.

3.3.2 Recognition of TIA in primary care

The most important cases of TIA to identify in primary care are those that are high risk for recurrent stroke, and the most difficult cases are those that are missed at initial consultation in primary care. The methodological weaknesses in existing vignette based studies testing TIA recognition in primary care are the restriction to anterior circulation symptoms, artificial case histories, very high prevalence of TIA in vignettes resulting in cueing of the diagnosis and the lack of randomisation of order of presentation. Alteration of single details of vignettes has been used to test hypotheses of the reasons for clinical decision making in general and have been applied to TIA vignettes in recognition studies but with the above methodological weaknesses. It is not clear if failure to recognise TIA is systemic, in that GPs are never trained to the standard of full recognition of the variation in TIA presentations, or whether they receive this knowledge during training but the low prevalence of the condition at individual practitioner level results in an atrophy of this knowledge with time.

As there have been no studies of GP trainees' recognition of TIA from vignettes, there is no evidence to inform strategies to improve recognition of TIA (to alter training or alter post graduate education). This also implies that a pilot study needs to be undertaken to determine the design parameters of recruitment and completion rates to inform an appropriately powered definitive study testing TIA recognition in GP trainees.

Research Questions

5. Do GP trainees fail to recognise high risk TIA cases that have been missed by established GPs?

6. What are the recruitment and completion rates for a web-based vignette study of TIA recognition by GP trainees?

7. Does alteration of one case vignette parameter influence TIA recognition and management decisions?
Research question 5 was answered with an internet-based questionnaire pilot study of GP trainees using case vignettes with a narrative description of results given the pilot status. The cases were constructed from the primary care records of a cohort of patients in OXVASC who presented in general practice after transient neurological symptoms consistent with TIA, but the diagnosis of TIA was not considered by the treating GP, and the patient went on to have a stroke within 30 days of the initial consultation. The missed TIA cases were therefore not artificial, included both anterior and posterior circulation symptoms, and were presented in random order. Distractor (non-TIA) cases were also presented to reduce the prevalence of TIA in the questionnaire to 50%.

Research question 6 was answered from analysing the numbers of trainees starting the questionnaire and completing the cases in full as a proportion of the denominator invited to take part. Research question 7 was answered by including a matched case with each case of missed TIA that varied in one parameter that I judged to have been a potential cause for failure to detect TIA.

3.4 GP and Specialist Disagreement – risk tools and diagnosis

Potential clinical rules for TIA prognosis (ABCD2) and diagnosis (222, hereafter known as the 'Dawson tool'), have been derived using clinical records made by specialists. There is no evidence about whether primary and secondary care records are in agreement over the key clinical features that drive the utility of the rules and whether disagreement is greater for patients who do not have TIA. This has implications for whether currently existing decision tools could be used appropriately in primary care to identify TIA patients at high risk of recurrent stroke for urgent clinic appointments. The only specific TIA recognition tool in the literature, the Dawson tool has not been subject to an external validation outside the clinic site where it was derived. Furthermore the ABCD2 score has been used to discriminate between TIA and non-TIA as causes of transient symptoms and although it was not originally derived as a diagnostic tool it shows reasonable accuracy in specialist recorded histories. Differences in Dawson tool and ABCD2 scores between GPs and specialists will impact on the extent to which a rule used in primary care can increase the positive predictive value of a referral for suspected TIA and crucially not miss patients with TIA.

Research Questions

8. Do GPs’ and specialists’ records of the same patients with suspected TIA agree about the clinical features?

9. Does GP/specialist disagreement vary according to final clinic diagnosis?
10. What impact does GP/specialist disagreement have on the performance of the ABCD2 score for accurate triage of patients with TIA?

11. What impact does GP/specialist disagreement have on the discrimination metrics of existing clinical prediction tools for TIA diagnosis?

Research question 8 was answered by examining all referrals with suspected TIA from primary care to assess the degree of disagreement for individual patients in clinical histories taken by GPs and by specialists. Research question 9 was examined by analysing the degree of disagreement in patient groups defined by final clinic diagnosis. Research question 10 was examined by the proportion of high and low risk TIAs that would have been inaccurately triaged with ABCD2 scores calculated from the primary care clinical histories and the number of strokes that may be expected by inaccurately triaged high risk TIA patients waiting for one week before definitive risk factor reduction. Research question 11 was examined by comparing the calibration and discrimination metrics for TIA diagnosis using the Dawson tool and the ABCD2 score in specialist histories and comparing these metrics with those from Dawson tool and ABCD2 score in primary care histories.

3.5 New Prediction Rules for TIA in Primary Care

There are no prediction rules for TIA diagnosis that have been derived from primary care records. Improving the recognition of TIA in primary care and increasing the predictive value of a primary care referral for TIA diagnosis will require a prediction rule that uses clinical features detected in primary care, and therefore it should be derived from primary care records. If primary care diagnostic models are different from specialist diagnostic models then the clinical predictors used must be different.

There is no evidence to suggest the most appropriate modelling strategy that represents the underlying relationships between predictor variables and outcome. Both logistic regression and classification tree methods have been used for derivation of prediction models in general but comparisons of techniques have not been reported for TIA diagnosis. Prediction using ensemble methods for classification and in particular, the random forest technique, have theoretical advantages over predictions from single models of a given dataset but there are no studies reporting benefits of such modelling methods over standard techniques for TIA diagnosis.

Research Questions

12. Which clinical predictors are included in a prediction rule for TIA diagnosis derived from primary care records in suspected TIA?
13. What are the calibration and discrimination metrics for a model of TIA diagnosis derived from primary care records in suspected TIA?

14. Does choice of statistical model affect discrimination metrics for TIA diagnostic models?

Research question 12 was answered by analysing the clinical records of all referrals with suspected TIA from primary care to a TIA clinic over a four year period. I restricted this to patients presenting with transient symptoms, as for those with persisting symptoms and signs there is less clinical doubt about decision making. General practice records of consultations and referral letters were used to create symptom categories from groups of similar textual descriptions of the presenting problem at first consultation in primary care. Research question 13 was answered by constructing a calibration curve from predicted probabilities from the model output and discrimination metrics calculated using model-based rules both with and without weighting by beta coefficients. Research question 14 was answered by comparing discrimination metrics for three different statistical models of TIA diagnosis.

3.6 Summary

The following reports of detailed methods and results address these evidence gaps, from the security of generalisable inferences from OXVASC, through the initial interface between patient and the healthcare system and the ability of clinicians to identify high risk presentations, as well as the utility of existing and potential future tools to enable more appropriate resource use in secondary care.
SECTION 2  The Oxford Vascular Study
Chapter 4  Oxford Vascular Study – methods and generalisability

4.1  Introduction

Making inferences from cohorts to national populations requires an assessment of their similarity. Given that age and deprivation can affect healthcare seeking behaviour, the research questions answered are

1. Is the OXVASC population representative of the national population in terms of age and deprivation?

2. Do differences in age and deprivation between OXVASC practices affect the presentation of patients with TIA to healthcare?

These questions are answered with comparisons between the OXVASC registered practice population and the population of England in terms of age and deprivation structure. The distribution of age and deprivation between the practices is compared and then potential impact of differences in these parameters on presentation of TIA is assessed.

4.2  The OXVASC methods - population and practices

The Oxford Vascular Study (OXVASC) is a prospective study of all vascular events (TIA/stroke, ischaemic cardiac events and peripheral vascular events) occurring in a population of 91,000 patients registered at nine general practices in urban and rural Oxfordshire (231). Population based studies including all incident cases within stroke medicine are important as the patients that are excluded from trials have a higher mortality but equivalent recurrent vascular event rate (232).

Patients are recruited after presentation to primary care or the ED (‘hot pursuit’) or via regular searches of the general practices’ computer records (‘cold pursuit’) (231). The general practices in OXVASC are those of the Oxford Community Stroke Project carried out 20 years previously (22). At recruitment, a structured proforma is used by either a research nurse or clinical research fellow to collect demographic data and clinical data about the presentation event and then follow up is either active via regular visits to a research clinic run by Professor Peter Rothwell (PMR) or via monitoring of the electronic GP record for vascular events. OXVASC methods were approved by the Oxford Clinical Research Ethics Committee (ref CO.043).

Although the data analysed in this thesis are from patients who were recruited from 1st April 2002 (the start of the study) to 1st April 2007, data on the distribution of age, sex and deprivation of the 91,000 registered patients were not available from the primary care trust
(NHS Oxfordshire) until August 2010 and is constituted by the patients registered at the OXVASC practices on the date 31.8.2010. Although there will have been some registered population turnover from 2007 to 2010, this is a continual process and would have been going on from 2002 to 2007. Given that practices have considerable boundary stability, the postcodes in the registered population are likely to give a stable estimate of the deprivation structure of that practice population even if there is a small percentage turnover of people residing at those postcodes.

4.3 Comparison of the OXVASC population with the total England population and inter-practice comparisons—age and deprivation

Substantial differences between the OXVASC population and the background England population may limit inferences concerning patient behaviour as well as affecting derivation of predictors for TIA diagnosis. Practice populations can also differ substantially in terms of the demographic structure of age and deprivation. Given the very high incidence in elderly populations (16) as well as the effect of deprivation on cerebrovascular disease (149), this may result in different incidence rates for TIA and minor stroke at practice level. This in turn implies that the GPs who refer to OXVASC may have different clinical experience with TIA and stroke with different opportunities for upkeep of diagnostic and management skills.

Socioeconomic position was measured using the Indices of Multiple Deprivation (IMD) 2007. The Office of the Deputy Prime Minister, responsible for policy development in the area of social exclusion, published a map of deprivation using an Index of Deprivation in 2004 (233) and this was updated in 2007 (234). The 2007 IMD consists of a sum of measures along seven dimensions of income deprivation, employment deprivation, health deprivation and disability, education skills and training deprivation, barriers to housing and services, crime and living environment deprivation. The summary IMD statistic in England and Wales is assigned to small regions of roughly 1,500 residents, formed in 2004 from similar contiguous geographical areas. These are termed lower super output areas (LSOAs). All England LSOAs are ranked in terms of their IMD scores and so a national rank is also available for each LSOA.

The age and deprivation structure of the OXVASC population was compared with the total England population data from the Office for National Statistics, which gives an age breakdown (grouped into 0 - 15 years, 16 – 29 years, 30 – 44 years, 45 – 64 years and > 65 years) for each LSOA (235). The total England data was merged with a separate Office for National Statistics file with IMD scores for each LSOA (236) to create a complete age and deprivation structure. The OXVASC population was grouped according to the same age bands for direct comparisons which were made visually from graphical display.
Inter-practice comparisons within the OXVASC population for mean age were carried out with analysis of variance, and for median IMD scores with the Kruskall-Wallis test due to the skewed nature of that distribution.

4.4 Demographic effects and recognition of TIA

Major strokes, by virtue of their absolute requirement for hospital care and certainty of diagnosis from imaging can be viewed as a reference at practice level for cerebrovascular burden. Given the high early risk of stroke after TIA (20), if one particular practice was subject to a bias of patients not presenting early after TIA then one would expect a change in the ratio of TIAs to major strokes occurring in that practice population, as patients would present with stroke (and be ascertained from hospital into OXVASC) during a period of delay after TIA. Also, if there was little exposure to TIA or minor stroke presentations (either due to reduced patient presentation to healthcare or to true low incidence) and such cases were missed due to lack of familiarity then the ratio between ascertained TIA and ascertained major stroke per practice would be affected.

Thus if patient healthcare seeking behaviour or GP accuracy of TIA detection are affected at practice level then the number of ascertained TIAs may differ by practice, potentially varying with age distribution or IMD distribution. However the number of major strokes will not be affected by healthcare seeking behaviour after major stroke or GP accuracy of detection of transient cerebrovascular presentations.

Thus an appropriate estimation of the effect of different age and deprivation demographics on presentation (and referral) would be if a relationship existed between the ratio of TIA to major stroke and age or deprivation structure of the practices.

In order to examine effects of age and deprivation structure, the sampling distribution of the TIA to major stroke ratio should ideally be known. The likely distribution of the two variables used to calculate the ratio, numbers of TIA and numbers of major stroke will be Poisson rather than normal. Furthermore they are in the same underlying population and will be correlated and therefore techniques that assume independence are not appropriate.

In order to investigate the association between practice factors of age distribution and deprivation on the rate of TIA, I modelled the effect of mean IMD and of % ≥ 65 years separately, using a Poisson regression (SPSS version 17). As the underlying distribution of the ratio of TIA to major stroke is unknown I was not able to use this as a dependent variable. Instead, I used the numbers of major strokes as an offset in the modelling equation, to act as a unit of reference. Poisson regression uses maximum likelihood to estimate model
parameters, a process which chooses the model parameters as the ones that are most likely, given the data. This analysis of optimal model parameters is an iterative process which uses a number of starting points and aims to converge on one set of model parameters at the end of each parameter search. Failure of convergence on one set of model parameters is possible with small numbers of practices in which case simple graphical displays can be used to investigate the associations between these factors and the ratio of these two rates.

I am grateful to Dr Richard Germuska at NHS Oxfordshire who provided the postcodes for the 90,235 patients registered at the OXVASC practices.

4.5 England and OXVASC population

Figure 4.1 displays a plot summarising the population structure as available from the UK Office of National Statistics (0 - 15 yrs, 16 - 29 yrs, 30 – 44 yrs, 45 – 64 yrs and > 65 yrs) for all lower super output areas (LSOAs) in England and for the registered population at OXVASC practices. The relative size of the circle at each age band represents the % of the total population contained within the relevant age ranges. From visual inspection, the relative sizes of the age bands are equivalent in OXVASC and in England, but there are differences in mean IMD scores, with the OXVASC population living in more affluent LSOAs than the UK average. The relationship between mean IMD across age bands within each population is similar with the highest mean IMD in the 16 – 29 year band and lowest in the >65 year band.
Figure 4.1   Age and deprivation structure of population of England and population of OXVASC registered patients

4.6   OXVASC Population – total and practice differences

Figure 4.2 shows a more detailed population pyramid by age deciles for the OXVASC registered population with sex differences, demonstrating an excess of females in the upper age deciles.
The overall age structure differs across the nine practices, particularly where there are high numbers of student registrations (Practice 1) as demonstrated by the large population band in the 20-30 year age group shown in figure 4.3.

Table 4.1 shows the mean (SD) for age in each practice and the total population. Analysis of variance demonstrates that the differences in age distribution are statistically significant (F = 228.51, p<0.001). Levene’s test for homogeneity of variances was significant (352.43, df=8, p<0.001) and so the Games-Howell post hoc comparison test was chosen to investigate which practices were significantly different from the rest of the group. Practices 1 and 5 which have lower mean values were significantly different from all other practices at p<0.001 for all comparisons.

However, given that there is variation in age composition of the practices it raises the possibility that there may be practice specific bias in characteristics of the registered population. As such, pooling data on patients across practices may not be appropriate as it will not take clustering of characteristics into account. Such clustering effects may be more likely if small numbers of the at risk population for TIA or minor stroke are registered at a given practice.
Figure 4.3  Age (years) and sex structure of the nine practices forming the OXVASC population.
Table 4.1 Absolute numbers and % of patients aged ≥ 50 and ≥ 65 registered at each OXVASC practice.

<table>
<thead>
<tr>
<th>PRACTICE</th>
<th>TOTAL PATIENTS</th>
<th>Age mean (SD)</th>
<th>≥ 50 yrs</th>
<th>% ≥50 yrs</th>
<th>≥ 65 yrs</th>
<th>% ≥ 65 yrs</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>12846</td>
<td>33.10 (17.60)</td>
<td>2230</td>
<td>17.4</td>
<td>893</td>
<td>7.0</td>
</tr>
<tr>
<td>2</td>
<td>4548</td>
<td>41.13 (22.10)</td>
<td>1678</td>
<td>36.9</td>
<td>644</td>
<td>14.2</td>
</tr>
<tr>
<td>3</td>
<td>5062</td>
<td>41.30 (23.40)</td>
<td>1931</td>
<td>38.1</td>
<td>979</td>
<td>19.3</td>
</tr>
<tr>
<td>4</td>
<td>15709</td>
<td>41.37 (23.76)</td>
<td>5836</td>
<td>37.2</td>
<td>2947</td>
<td>18.8</td>
</tr>
<tr>
<td>5</td>
<td>8686</td>
<td>34.90 (20.64)</td>
<td>1930</td>
<td>22.2</td>
<td>859</td>
<td>9.9</td>
</tr>
<tr>
<td>6</td>
<td>12392</td>
<td>41.91 (23.83)</td>
<td>4760</td>
<td>38.4</td>
<td>2314</td>
<td>18.7</td>
</tr>
<tr>
<td>7</td>
<td>12661</td>
<td>40.66 (23.11)</td>
<td>4633</td>
<td>36.6</td>
<td>2061</td>
<td>16.3</td>
</tr>
<tr>
<td>8</td>
<td>11117</td>
<td>39.92 (22.75)</td>
<td>3766</td>
<td>33.9</td>
<td>1758</td>
<td>15.8</td>
</tr>
<tr>
<td>9</td>
<td>7214</td>
<td>41.38 (22.34)</td>
<td>281</td>
<td>39.1</td>
<td>113</td>
<td>15.7</td>
</tr>
<tr>
<td>Total</td>
<td>90235</td>
<td>39.35 (22.51)</td>
<td>29585</td>
<td>32.8</td>
<td>13589</td>
<td>15.1</td>
</tr>
</tbody>
</table>

4.7 The ‘at risk’ population and the ‘high incidence’ population

95% of patients with TIA and MIS ascertained in OXVASC are aged over 50 and this population i.e those ≥50 yrs registered at OXVASC practices can be described as the ‘at risk’ population. After the age of 65, the incidence of TIA and MIS rises sharply (16) and this sub-population within each practice is therefore likely to yield a large proportion of events.

Given that smaller units of recruitment may lead to bias, the absolute numbers of the at-risk population and high incidence population is required. Figure 4.3 displays the absolute number of patients greater than 50 years old registered at the 9 practices within each ten year band.
Figure 4.4  Numbers of registered patients with age > 50 within each ten year band of the nine general practices forming the OXVASC baseline population.

This shows the large variation in absolute practice numbers of at risk patients although the age band structure within each practice in the registered population over 50 appears stable across practices. Given that the incidence of TIA increases in the elderly and that 75% of all TIA and MIS occur in those greater than 65 years old it is important to confirm that absolute numbers of patients at risk in these older age groups are present in numbers sufficient to minimise bias which for example could manifest as reduced ability to detect TIA or MIS simply due to very low numbers of events per GP occurring in a given practice population.

Table 4.1 shows the absolute numbers of patients that are in the ≥50 age group and ≥65 age group at each practice and as a % of the total at the practice. The practice with high student registrations (practice number 1) has the lowest percentage of the at risk population (≥ 50 yrs) at 13.4% and the lowest in the high incidence population (≥65 yrs) of just 7% (n=893). A large practice with very few high incidence patients suggests that the GPs in the practice will have little exposure (on a per GP basis) to the type of presentations associated with TIA and MIS, compared with a similar sized practice with a much larger group of high incidence
patients. Practice number 6 conversely has a larger fraction of the high incidence age group at 18.7% (n = 2314) with greater opportunity for familiarity with TIA/minor stroke on a per GP basis.

4.8 Distribution of deprivation – IMD 2007 scores

For the total OXVASC population the IMD distribution is skewed with more patients residing in LSOAs with lower IMD scores, i.e. more affluent areas as shown in figure 4.4

![Histogram of IMD 2007 scores for patients registered at all practices.](image)

**Figure 4.5 Distribution of IMD 2007 scores for patients registered at all practices.**

For the OXVASC population as a whole, comparing the ≥65 year old population with the population <65 years shows that there is a statistically significant difference in deprivation distribution with older patients living in more affluent LSOAs (≥65 median IMD =7.44 vs < 65 median IMD = 9.80, Mann Whitney z = -25.9, P<0.001). Practice level populations were also examined for age and deprivation differences.

The mean IMD in each practice may mask differences in the distribution of deprivation within key population sub groups. The average IMD score for each practice is shown in table 4.2. Marked differences are seen in the deprivation scores for the total populations and in the subset populations of interest.
In general, table 4.2 shows that the high incidence groups have a mean IMD lower or similar to the practice mean. Practice numbers 1, 3 and 5 have the highest mean IMD scores in the ≥ 65 age group, yet they have some of the lowest absolute numbers in this group (total for practices 1,3,5 = 1897 patients, % of total ≥ 65 = 19.5%). Thus the majority of the high incidence group of patients in OXVASC are registered at other practices with lower mean IMD scores.

<table>
<thead>
<tr>
<th>PRACTICE</th>
<th>&lt; 50 yrs Mean (SD)</th>
<th>≥ 50 yrs Mean (SD)</th>
<th>&lt; 65 yrs Mean (SD)</th>
<th>≥ 65 yrs Mean (SD)</th>
<th>All patients Mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>16.09 (7.86)</td>
<td>12.43 (7.55)</td>
<td>15.72 (7.90)</td>
<td>11.90 (7.34)</td>
<td>15.45 (7.92)</td>
</tr>
<tr>
<td>2</td>
<td>8.78 (4.30)</td>
<td>8.00 (4.20)</td>
<td>8.61 (4.29)</td>
<td>7.78 (4.15)</td>
<td>8.49 (4.28)</td>
</tr>
<tr>
<td>3</td>
<td>15.60 (6.42)</td>
<td>12.15 (6.82)</td>
<td>14.98 (6.62)</td>
<td>11.36 (6.70)</td>
<td>14.28 (6.78)</td>
</tr>
<tr>
<td>4</td>
<td>9.76 (6.46)</td>
<td>8.81 (5.70)</td>
<td>9.55 (6.33)</td>
<td>8.80 (5.59)</td>
<td>9.41 (6.21)</td>
</tr>
<tr>
<td>5</td>
<td>22.50 (8.37)</td>
<td>22.11 (8.98)</td>
<td>22.38 (8.41)</td>
<td>22.68 (9.36)</td>
<td>22.41 (8.51)</td>
</tr>
<tr>
<td>6</td>
<td>7.22 (2.74)</td>
<td>7.34 (2.68)</td>
<td>7.23 (2.72)</td>
<td>7.42 (2.70)</td>
<td>7.27 (2.72)</td>
</tr>
<tr>
<td>7</td>
<td>8.66 (5.91)</td>
<td>8.32 (5.98)</td>
<td>8.53 (5.72)</td>
<td>8.58 (5.34)</td>
<td>8.53 (5.66)</td>
</tr>
<tr>
<td>8</td>
<td>9.67 (6.23)</td>
<td>8.60 (5.30)</td>
<td>9.41 (6.12)</td>
<td>8.78 (5.20)</td>
<td>9.31 (5.98)</td>
</tr>
<tr>
<td>9</td>
<td>7.87 (3.75)</td>
<td>7.46 (3.74)</td>
<td>7.79 (3.79)</td>
<td>7.26 (3.51)</td>
<td>7.71 (3.75)</td>
</tr>
</tbody>
</table>

Table 4.2 Mean (SD) IMD scores for all registered patients and by grouping into at risk and high incidence age groups.

Given the skewed distribution of IMD score, differences between practices in IMD distribution were tested using a non-parametric test which found statistically significant differences in the high incidence population (Kruskall-Wallis test $\chi^2 = 27796$, df=8, p<0.001)

4.9 Assessing the impact of different age and deprivation structures in practices on recruitment

A priori, the effect of reduced % at risk and reduced % high incidence populations will be to dilute the clinical experience at individual GP level. Thus not only will there be reduced incidence of TIA and MIS at practice level but potentially the accurate detection of symptoms may differ as well. This could also be affected by level of deprivation in the high risk group which can influence incidence and presentation to primary care.
Thus if patient healthcare seeking behaviour or GP accuracy of TIA detection are affected at practice level, then the number of ascertained TIAs may differ by practice, potentially varying with age distribution or IMD distribution. By contrast the number of major strokes (defined as NIHSS score >5) is much less likely to be affected by healthcare seeking behaviour or GP diagnostic accuracy.

The investigation of the effect of age and deprivation distribution at practice level on TIA incidence was attempted with Poisson regression. Using stroke incidence as an offset, the regression models against both % population ≥ 65 years at each practice and mean IMD in the ≥ 65 years population at each practice did not converge. Instead, relationships between the ratio of TIA to major stroke were plotted against practice level % population ≥ 65 years and mean IMD in the ≥ 65 age group and assessed visually.

Figure 4.5 displays the plot of TIA: Major stroke in the practices ordered by increasing % of total population in the high incidence age group. The ratio does not increase with increasing % population ≥ 65 years, suggesting that a low per GP event rate does not systematically result in ‘under finding’ of TIA.

Figure 4.6 displays the plot of TIA: Major stroke in the practices ordered by increasing mean IMD score in the ≥ 65 population. Again there does not appear to be a relationship between this ratio and increasing deprivation in the high incidence population.
4.10 Summary

The OXVASC population is not representative of the national population in terms of deprivation, although the age structure is similar to that of the national population and the relative deprivation between age groups is similar. There is under sampling from areas of high deprivation in the IMD score histogram of the registered population at OXVASC practices. Given the small numbers residing in highest IMD scored postcodes, the use of quartiles creates larger and equal numbers within groups of increasing deprivation but at the price of potentially not detecting effects that are evident at highest levels of deprivation. As the OXVASC population is less deprived than the national population, the strength of associations between deprivation and healthcare seeking behaviour in the cohort may not reflect the strength of association in the national population.

Differences in age and deprivation of the populations at the different practices in OXVASC do not affect the presentation of TIA to healthcare. Although there are significant differences in age and deprivation structure of the practice populations, the ratio of TIA to major stroke does not vary systematically with increasing high risk population at a given practice or with mean IMD in this population. Hence recognition and referral of TIA does not systematically vary with these parameters. As such, it is valid to pool patients from different practices and analyse them as one group.

Figure 4.7 Scattergram of practice TIA: major stroke against increasing mean IMD score in practice registered population ≥ 65 years old
SECTION 3  Initial Interface Between Patient and Healthcare System
Chapter 5  Methods

5.1  Introduction

This chapter outlines the methods used to address the evidence gap concerning access to the first medical assessment after the onset of symptoms of either TIA or MIS and the ability to recognise that TIA or MIS may be the cause of symptoms.

Education over how to access services in a timely manner to facilitate urgent assessment and treatment requires an understanding of potential factors that influence patients’ choices over initial healthcare provider. Previous experience of TIA, minor stroke or major stroke may influence choice of provider and delay to call for medical attention so whether the event that prompted medical attention was incident or recurrent may be associated with provider choice and delay.

The research questions on healthcare access are

1. Is healthcare access after TIA or minor stroke influenced by the following factors; choice of healthcare provider, time of symptom onset, clinical features of the event, deprivation and previous history of cerebrovascular disease?

2. Did the change in general practitioner contract influence healthcare access after TIA or minor stroke?

The key components of health care access, choice of healthcare provider, and delay to calling that provider were measured in the OXVASC cohort 2002 – 2006. Associations between delay to call for help and the choice of provider are examined in the first research question. The effect of patient level factors of demographic and clinical features and the effect of healthcare system factors (routine primary care availability and the change in the GMS contract) on provider choice and delay were tested with strength of statistical associations. The impact of altering primary care availability was estimated from the number of patients who had a TIA in the out of hours period and went on to have a recurrent stroke before their practice re-opened.

TIA or MIS may not be recognised at that first healthcare contact and this is also a cause of delay to definitive treatment, either because underlying pathology is suspected and patients are referred to an intermediary specialist or because no significant cause for the presentation is suspected and patients are sent home. As TIA and MIS are seen infrequently by individual GPs it is not clear if failure to recognise the cause of symptoms is due to low
familiarity from low clinical exposure or because GPs are not trained sufficiently to recognise the various presentations of TIA and MIS. The research questions were

3. Do GP trainees fail to recognise high risk TIA cases that have been missed by established GPs?

4. What are the recruitment and completion rates for a web-based vignette study of TIA recognition by GP trainees?

5. Does alteration of one case vignette parameter influence TIA recognition and management decisions?

These questions were answered with a questionnaire of vignettes of high risk TIA presentations missed in primary care (with matched cases altered by one parameter) which was piloted in GP trainees, testing diagnosis and management.

5.2 Definition of In Hours and Out of Hours (OOH)

A number of general practices offered Saturday morning surgeries under the old GMS contract and in order to compare patients’ healthcare seeking behaviour for events occurring in hours and OOH, it is necessary to define these time periods. All OXVASC practices offered Saturday morning emergency surgeries (personal communication from each practice manager) so for patients who had events in these time periods before the new GMS contract, Saturday mornings between 09:00 and 12:00 were considered to be in hours, as the registered general practice was open. Although patients could contact a known GP from the registered practice throughout the 24 hour period of each day the premises had weekday office hours opening and the time periods of 08:00 to 18:30 were chosen in the old contract time period to reflect this. This is also the time that practices are open under the service of the new GMS contract.

Thus for the old contract in hours were 08:00-18:30 Monday to Friday and 09:00 – 12:00 Saturday, and for the new contract in hours were 08:00 – 18:30 Monday to Friday only. All other times are defined as OOH.

5.3 Methods for research question 1 – determinants of choice of provider and delay to call

Choice of provider was defined as the setting of the first medical assessment (ED including the separate Eye Hospital emergency department, registered general practice, out of hours GP, other hospital clinic or occurring whilst an in-patient). Proportions of patients attending a
given provider in the in hours and OOH periods were compared using z tests assuming
independent proportions (Microsoft Excel software).

Median delay in calling for medical attention after TIA and minor stroke was analysed for the
following comparisons using Mann-Witney U test

1. Incident events vs recurrent events
2. In hours events vs out of hours events
3. Primary care attendance vs secondary care attendance
4. Most deprived quartile vs least deprived quartile
5. Clinical features (weakness, sensory loss, visual change, speech disturbance) - presence vs absence

5.4 Data quality

5.4.1 Quality and Outcomes Framework (QoF)
The QoF incentivises GPs to detect and control chronic disease e.g. diabetes and
hypertension. Given that the QoF was introduced at the mid-point of the time period for data
collection for the above sample it would be of interest to examine whether the prevalences
for chronic diseases and risk factors are different for patients recruited after TIA and minor
stroke in the old and new contract time periods

5.4.2 Events occurring during holidays

In a population based study with active ascertainment some patients may well have events
that occur when they are outside their usual area of residence, which could influence not
only their choice of healthcare provider but also the delay in calling for help. In order to test
the effect of the GMS contract and GP opening hours in general on patients’ choice of
provider and delay in calling them, I excluded 25 patients who had events whilst out of the
UK and were therefore in an unfamiliar healthcare setting at the time of symptom onset.

5.4.3 Measuring delay and missing data

Delay after TIA has been previously measured using calendar days in a combination of
research and routine data in Oxford (43) and similarly in a comparison between UK and
Canada for TIA presentations (141). Delay in this thesis was instead measured in hours for
the interval between symptom onset to the first call for medical assistance - either telephone
to a GP provider, call to emergency services or the time of self-presentation at a GP practice
or ED. For patients who were unable to give timings at their recruitment interview, these were derived from GP consultation notes, GP letters or ambulance transfer sheets. If these were unavailable, then call times were imputed using the modal class of call time from available data.

Given that some patients ascertained in OXVASC sought care many months after their event (which in the majority of these cases was mainly mentioned in passing when attending for other reasons) median times to call for help were used as the most appropriate measure of central tendency of the skewed distributions of delays to call.

5.5 Evidence for combining contract time periods

I hypothesised that the new contract would affect choice of provider (primary care, OOH primary care and ED) but not delays to call that provider i.e. those using primary care would do so with the same delay to call in 2002-2004 as in 2004 – 2006.

This hypothesis was investigated by comparing median call times for all patients in the old contract and all patients in the new contract with the Mann-Witney U test (SPSS 17).

5.6 Research question 2 – the influence of the General Medical Services (GMS) contract change on healthcare access

In April 2004, the new GMS contract for primary care was introduced (130). Alongside a pay for performance structure set out as the Quality and Outcomes Framework (QoF), the mandatory 24 hour responsibility for patient care ended. Practices could opt out of this 24 hour care model, which the vast majority did given that the only penalty was a small reduction in practice income. Local Primary Care Trusts (PCTs) took over the responsibility for ensuring access to primary care services outside office hours which were defined as 18:30 to 08:00 from Monday to Thursday and 18:30 on Friday to 08:00 on Monday i.e. 24 hours on Saturday and Sunday. A proposal to extend general practice opening hours was put forward by the UK government for negotiation with the General Practitioner Committee of the British Medical Association in 2008 (237).

All the OXVASC practices opted out of the 24 hour care model and therefore their practice opening hours were the only times when a patient could access their registered general practitioner. The out of hours contact with primary care is initiated either by ringing the registered practice and then receiving details about contacting the out of hours (OOH) service or having the surgery telephone number redirected to the OOH provider automatically.
As patients were ascertained in OXVASC for two years before the GMS contract change, there is an opportunity to analyse the effect that the change in GMS contract had on decisions to contact primary care in the out of hours time period. At the time of this analysis it had already been established that service level issues may have an influence on the delay to calling for help but no evidence of how the delivery of primary care could influence the delay to specialist assessment (43). A number of potential influences can be hypothesised.

1. Perceived lack of availability would prompt patients to attend ED rather than ringing their practice and being redirected to the OOH primary care service.  
2. Knowing that the practice is closed, patients are aware that they are not disturbing their own doctor and this might prompt more patients to ring the OOH provider  
3. For patients that will only see their own practice, more may wait until the surgery is open again before making any healthcare contact.

In order to assess the impact of the new GMS contract I chose to analyse two equivalent time periods before and after its introduction. As OXVASC began collecting data on the 1st April 2002, two years were chosen before and after the change in contract, therefore all patients with TIA and MIS that were ascertained in the time period 1st April 2002 to 31st March 2006 were selected for analysis.

Proportions of patients attending a given provider in the new and old contract time periods were compared with z tests assuming independent proportions (Microsoft Excel software).

5.7 Early Recurrent Strokes – do GP opening hours contribute?

One of the advantages of the OXVASC population based study with active ascertainment is that patients who present with major stroke are recorded and previous events in these patients that were thought to be due to TIA or MIS are also documented.

If GP opening hours are indeed a barrier to care then, given the high early risk of recurrent stroke, one would expect to find patients who had a TIA or MIS during the OOH period who went on to have a recurrent stroke before their registered practice re-opens. I therefore sought patients who did not seek care after TIA/MIS in the OOH period and were ascertained in OXVASC only after a recurrent stroke before the time that their practice reopens. This assumes that, as they had not gone to the ED, they may have been waiting for their practice to reopen before seeking care.

Alternative primary care centres were proposed in the NHS review led by Lord Darzi, Minister of State for Health (238), where primary care access would be increased to 08:00 to 20:00, seven days a week. If all primary care delivery were to be structured in this way then
there may be benefits in reducing recurrent stroke for patients presenting to services after TIA and MIS, provided that there is capacity in secondary care to respond in a timely manner.

I calculated the number of strokes that may have been prevented with ‘Darzi centre’ opening hours, by examining the number of TIA or MIS cases occurring in the extra period of opening hours that such a centre would offer.

5.8 Distribution of call delay and practice opening hours

The pattern of call delay was displayed visually by plotting call delay against the time of event to demonstrate how the time of event and the opening hours of general practice influence call delay. The frequencies of time bins of calls to general practices were plotted as further evidence of the link between practice opening hours and timing of requests for medical assistance.

5.9 Rationale for testing recognition of TIA

5.9.1 Bias in clinic-case derived tools – ‘missed’ high risk TIA

The classical presentations of TIA are well known to clinicians and the clinic referral cohort of TIA and mimics of TIA have similar characteristics in regard to these features, such as weakness and speech disturbance. However, there may be cases of TIA sharing common features that are less well known to GPs and therefore may be missed after initial presentation in primary care. If these particular TIAs are high risk, then they will present a short time afterwards with stroke.

A major bias of the Dawson diagnostic tool for TIA is that it is derived using patients who have been referred to clinics, and as such GPs or ED clinicians have been concerned enough about the underlying diagnosis in these patients to refer for a specialist opinion.

Patients who may present to both primary care and ED with high risk TIA but are not considered to have TIA by the treating clinician will not be included in the reference pool of clinic referrals. Such cases may share atypical features (as clinicians may not have considered the diagnosis) resulting in a biased tool where not all TIAs have been included at derivation. This is particularly problematic if the missed cases are associated with early recurrent stroke.

As OXVASC recruits patients after all vascular events, all patients that present with stroke rather than TIA are also ascertained. All ascertained patients have primary care records examined for relevant risk factors and their control e.g. office blood pressure level, as well as
previous presentations to healthcare (either records from primary care or letters from ED) with events that could be due to cerebrovascular disease.

This represents an opportunity to examine in more detail patients who were recruited after a major stroke but may have had a TIA beforehand and presented to primary care after transient phenomena but were not diagnosed as having TIA. By virtue of case definition these transient events are very high risk for recurrent stroke.

The patients that are included are those who had a recurrent stroke within 30 days of a TIA but presented to medical attention in the time window after TIA and before the recurrent stroke. A time to recurrent event of $\leq 30$ days after TIA was chosen because it could be argued that some presentations, if atypical, are less likely to represent a high risk TIA if the recurrent stroke occurs beyond this time window.

Although this yet again introduces a bias as patients with high risk events that do not present after TIA are not included, there is no reliable record taken prospectively after the initial TIA in such cases. Furthermore if one makes the assumption that such presentations are associated with a persistent lack of presentation to healthcare then there will not be an opportunity for a clinician to ‘miss’ the diagnosis of TIA prior to presentation with stroke.

The care-seeking behaviour of patients in OXVASC who had a recurrent ischaemic stroke within 30 days of a TIA was used to identify those patients who presented to primary care but the diagnosis of TIA was not considered. Other referral routes from primary care were assessed to define ‘missed events’ in primary care.

Data were available for all patients with ischaemic stroke after TIA, ascertained for the first eight years (2002 to 2010) of OXVASC.

5.9.2 What is a ‘missed’ cerebrovascular event?

A liberal definition of a missed event would include all patients who have had a delay in receiving optimal therapy because the diagnosis was not considered at the initial consultation. This could be for a number of reasons after presenting to primary care -

1. The diagnosis was not considered and the patient had a recurrent stroke which resulted in further presentation to healthcare for investigation and treatment.
2. The diagnosis was not considered until a later date before any recurrent events and then a referral was made for investigation and treatment.
3. The diagnosis was not considered but a referral was made to other specialist teams e.g. ophthalmology who then made the diagnosis and referred onwards for investigation and treatment.

Out of the above patient groups, point 1 defines those patients who are likely to be at highest risk of recurrent stroke and also where primary care may have the greatest difficulty in making a diagnosis. Point 2 defines patients at lower risk as there are no recurrent events after initial presentation during any period of delay. Point 3 defines a group where GPs suspect that there is an underlying condition to diagnose but refer to an intermediary specialist who then either asks the GP to refer on for a TIA assessment or refers the patient directly without further involvement of primary care. This latter group requires an analysis of referral routes to TIA clinic and identification of those patients who arrived via an intermediary specialist.

5.9.3 ‘Hidden TIA’ and prediction rule development

Patients with TIA can be correctly recognised by a referring clinician yet have a recurrent stroke requiring hospitalisation before being assessed in a TIA clinic. In OXVASC, these patients are ascertained as a stroke, rather than a TIA. Therefore by restricting the derivation of decision rules to patients seen in clinic (not only in OXVASC but also in other centres such as Glasgow in the Dawson tool), this will reduce the assessment of true TIA cases referred by GPs as the total set of GP suspected TIA will be missing the patients with highest early risk.

This creates an interesting bias as ideally the highest risk patients should be identified by a diagnostic or referral support tool but they did not appear in the Dawson derivation set by definition (as only clinic-seen patients were included) and have not been used in the derivation of decision rules above, partly to allow for a fair comparison with the Dawson tool.

The incidence of these high risk cases was assessed to determine impact on clinic-based derivation datasets.

5.9.4 Other missed opportunities – intermediary specialist referral

Given that the effects of transient cerebral ischaemia will result in dysfunction of a regional part of the body associated with a particular area of brain or retina, it is perhaps unsurprising that GPs may refer to intermediary specialists (i.e. not TIA specialists). The referral will be based on the presumption that the affected body part is dysfunctional, rather than the controlling homuncular cortical area or retina. This may be more of a problem in patients with
visual symptoms which can be interpreted at first consultation as ocular in origin (due to lenticular, humoural or retinal disease).

Bias could be introduced if such patients had recurrent strokes before being seen in a clinic, or if persisting non-disabling symptoms developed and in OXVASC (and in other clinic-defined cohorts used for TIA diagnosis) were ascertained as a stroke rather than TIA. The extent to which intermediary specialist referral in OXVASC introduces a bias in using clinic-defined TIA populations for diagnostic decision rules will be assessed by estimating the incidence of affected patients.

5.10 Using missed TIA to assess diagnostic decision making - case vignette methods

In order to examine whether the high risk TIA cases that were missed represent cases that GPs in general would find difficulty identifying as TIA, a study is required that tests decision making in clinical situations that, as far as possible, resemble the clinical situation that faced the GPs seeing these identified cases.

Case vignettes are short case descriptions of key clinical features from history, examination and where relevant, investigations. They have been used to test GPs’ diagnostic ability (239-242), the improvement in GPs’ diagnostic ability after training (243;244), GPs’ management of cardiovascular conditions (245-248) and perception of risk of events (249). Case vignettes have been validated as reflecting clinical decision making in clinical practice (250) and have been used across different clinical specialties for assessing the contributions of different variables to making diagnoses and management decisions (251-256). Furthermore, vignettes have been validated as an assessment of quality of care and also for measuring variation in clinical practice (257).

Specifically, vignettes have been used to investigate diagnostic error with the presentation of a single misleading detail (258). Varying the content of a single vignette has been used to understand the factors taken into account by clinicians when they make diagnostic errors in bipolar disorder (259).

The literature review identified the key problems with current vignette evidence on TIA diagnosis in primary care; artificial histories potentially not reflecting real presentation to GPs, cases restricted to anterior circulation, lack of randomised order of questions for each participant and cueing that TIA is a diagnostic possibility which does not reflect the clinical decision making context at initial consultation in primary care. More importantly, it is crucial
that patients at highest risk of recurrence should be detected and then appropriately managed with urgent referral to specialist TIA services.

5.11 TIA decision making – Rationale for GP trainees as participants

I selected GP trainees to be the participants. After leaving GP training, fledgling GPs will no longer be taking part in a systematic educational programme and whilst learning will continue it will be on the basis of individual choice. Thus areas for study and review are chosen by individual GPs from taught courses, online learning modules and journal reading to comply with the requirements of revalidation. If the cases missed by practising GPs are also missed by GP trainees then this suggests that current post graduate education does not equip GPs to detect key TIA presentations. If GP trainees detect the missed TIA cases, then this argues that there may be an element of knowledge atrophy due to the infrequent nature of the condition at individual GP level. The solution to the problem of missed high risk TIA cases could therefore be altering post graduate education or an intervention to keep low prevalence but serious conditions in the minds of GPs.

5.12 Research Questions 3, 4, and 5 - Questionnaire Design

In order to test the diagnostic skill resulting from exposure to hospital-based clinical experience and educational programmes, I chose to base the vignettes on patients who sought care in general practice after TIA, where TIA was not suspected by the GP, yet the patients went on to have a recurrent stroke within 30 days. These are patients who should not be missed in primary care as they are high risk but there may be difficulties in making the diagnosis. This answers the question of whether GP trainees fail to recognise high risk TIA cases that have been missed by established GPs?

However, a pilot trial was needed in order to determine the key parameters of recruitment rate and completion rate of a questionnaire for GP trainees. There are no similar GP trainee surveys in the literature so the design parameters could not be estimated using published reports. These would be needed for a large scale trial to calculate power to detect differences in detection rates within matched pairs. This answers the question of recruitment and completion rates among GP trainees for a web based questionnaire survey.

In order to examine potential reasons for the TIA cases to be missed, I selected a plausible hypothesis for each case and created a matched case where the only materially different factor was related to the hypothesis as to why the case was missed. A number of causes could be relevant but in order to create a matched pair I chose one that appeared to be most
likely. This answers the question of whether alteration of one parameter in a vignette affects recognition and management decisions.

To prevent cueing of TIA, I created a questionnaire of ten vignettes with five TIA cases and five cases due to other pathologies that are prevalent within an elderly population. One of the cases was a non-dominant parietal TIA presenting as ‘confusion’ and the matched case was confusion without focal parietal features, i.e. not TIA. This enabled an equal number of TIA and non-TIA cases in the questionnaire.

A questionnaire was created via the online tool, surveymonkey (www.surveymonkey.com) which allows for secure data collection and 24 hour access for participants to choose when to complete the study. The questions were randomised such that each time the weblink was accessed, a different order of questions was created for the survey.

5.13 Study Design for Recognition of High Risk Missed TIA Using Vignettes

The following design was approved by the University of Oxford Clinical Trials and Research Governance Service and submitted via the Integrated Research Application System (IRAS) for approval from the National Research Ethics Service (NRES) – REC reference 11/EM/0252.

Primary Objective:
To assess the ability of GP trainees to appropriately manage high risk TIA

Secondary Objectives:
To assess the recruitment rate to an email invitation to a clinical vignette study
To assess the completion rate of the full questionnaire of 10 vignettes.
To examine the influence of clinical variables on recognition of TIA.
To examine the perceived educational value of completing the questionnaire.

Summary of Study Design
Cross-sectional questionnaire study of responses to clinical vignettes. Participants will undertake one questionnaire and the responses to 10 vignettes will take 25 minutes to complete.

Primary and Secondary Outcome Measures
Primary outcome measure:
1. % correct management decision to refer for a specialist TIA assessment for each case of TIA. This is the primary outcome as it is the key step in accessing effective
interventions for the patient. Even if the GP does not think that the diagnosis is likely, the decision to refer will enable rapid access to definitive risk reducing treatment.

**Secondary outcome measures:**
1. Recruitment rate (% of GP trainees invited who responded to the survey)
2. Full completion rate (among those who completed the survey, % cases with full responses)
3. % correct TIA diagnosis
4. % cases where parameter change altered decision making in matched pairs of vignettes
5. Self-rating of perceived educational value after being given feedback on performance

**Study Participants**

All GP trainees (ST1-4) in the Oxford deanery

**Study Procedures**

GP trainees will be emailed with an invitation to take part in a study of “common problems that occur in primary care”. They will be informed that this is part of a research study and that they will be given feedback on the group performance after completing the questionnaire. Their views on the educational value of the questionnaire and feedback will be sought.

**Informed Consent**

The email will state that their responses will be recorded anonymously via a web based survey as part of a research study and that they are under no obligation to start or complete the questionnaire. An information sheet giving further details about the study that is not contained within the invitation email will form the first web page of the questionnaire. A formal informed consent procedure will not be undertaken but consent will be implied by completing the questionnaire.

**Study Assessments**

A link to a secure website (hosted by [www.surveymonkey.com](http://www.surveymonkey.com)) will contain 10 clinical case vignettes with free text response to the questions “What would you do next?” and “What do you think is the most likely diagnosis?”. The first question in the survey will simply ask which year of training the respondent is in (ST1-4) and whether they have undertaken any elderly medicine hospital training as a specialist trainee or foundation doctor.
A reminder email will be sent out two weeks after the initial email.

After the responses have been entered and completed, a further email will be sent out with an explanation of the correct management for each case and at that point the trainees will be invited to rate the educational experience of the exercise by another link to a separate web-based survey where they will be asked to score the educational value from 1 (no value) to 5 (highest value) with a free text box to add any comment.

**Definition of End of Study**

The end of study is four weeks after the last email is sent out.

**Number of Participants**

A total of 54 GP trainees are in the Oxford deanery rotation and this is the total pool which will form the group to whom invitations to take part will be sent.

**Analysis of Endpoints**

As this is a pilot study to inform recruitment and completion rates, null hypotheses will not be tested. Descriptive statistics for the primary and secondary endpoints will be summarised by year of training (1 to 4) and by elderly medicine experience as a specialist trainee (dichotomised to yes/no).

**Ethics**

The main ethical consideration of the study concerns the initial invitation to the GP trainees. This study aims to test their ability to recognise TIA but if we were to inform them that this was the motivation for the study then their diagnostic ability would not be truly tested. Also, given that patients do not present with ‘clues’ as to their underlying diagnosis before they are assessed, real life decision making is tested by not alerting trainees to the true nature of the research question. As such I believe that not disclosing the research question is justified so that not only are trainees fairly assessed but also the training that they receive is fairly assessed as well. GP trainees are familiar with clinical vignettes in the exam situation where no hint is given as to the underlying diagnostic possibilities.

**Participant Confidentiality**

GP trainee identity will not be recorded in the clinical cases questionnaire or in the questionnaire of perceived educational value. All email invitations will be sent from a GP training Programme Director in Oxford, who has access to trainees’ email addresses as part of their routine work.
Chapter 6  Choice of Provider after TIA and minor stroke

6.1  Introduction

Clinical and demographic characteristics of the patients recruited into OXVASC after the first TIA or minor stroke occurring during the study period are presented along with the effects of the new GMS contract for primary care and timing of events on the choice of first provider of care.

6.2  Characteristics of the Patients – does QoF influence coding of co-morbidities?

Table 6.1 shows the numbers of patients with TIA and minor stroke in the first four years of OXVASC with basic demographics summarised by diagnostic group.

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>TIA</th>
<th>Minor stroke</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>361</td>
<td>436</td>
</tr>
<tr>
<td>Male n (%)</td>
<td>127 (43.5)</td>
<td>222 (50.9)</td>
</tr>
<tr>
<td>Mean age (SD)</td>
<td>74.2 (12.2)</td>
<td>74.2 (12.0)</td>
</tr>
<tr>
<td>Caucasian n (%)</td>
<td>326 (96.7)</td>
<td>413 (94.7)</td>
</tr>
<tr>
<td>Recurrent event n (%)</td>
<td>95 (26.3)</td>
<td>76 (17.4)</td>
</tr>
<tr>
<td>Previous TIA n (%)</td>
<td>67 (18.6)</td>
<td>58 (13.4)</td>
</tr>
<tr>
<td>Previous Stroke n (%)</td>
<td>48 (13.3)</td>
<td>75 (17.3)</td>
</tr>
<tr>
<td>Previous myocardial infarct n (%)</td>
<td>44 (12.2)</td>
<td>53 (12.2)</td>
</tr>
<tr>
<td>History of angina n (%)</td>
<td>62 (17.2)</td>
<td>63 (14.5)</td>
</tr>
<tr>
<td>History of cardiac failure n (%)</td>
<td>38 (10.5)</td>
<td>43 (9.3)</td>
</tr>
<tr>
<td>Previous peripheral vascular disease n (%)</td>
<td>22 (6.1)</td>
<td>39 (9.0)</td>
</tr>
<tr>
<td>History of hypertension n (%)</td>
<td>186 (51.7)</td>
<td>250 (57.7)</td>
</tr>
<tr>
<td>History of atrial fibrillation n (%)</td>
<td>64 (17.8)</td>
<td>74 (17.1)</td>
</tr>
<tr>
<td>History of diabetes n (%)</td>
<td>38 (10.6)</td>
<td>48 (11.1)</td>
</tr>
<tr>
<td>History of hypercholesterolaemia n (%)</td>
<td>115 (31.9)</td>
<td>99 (22.7)</td>
</tr>
<tr>
<td>Never smoked n (%)</td>
<td>154 (44.0)</td>
<td>176 (40.6)</td>
</tr>
</tbody>
</table>

Table 6.1 TIA and minor stroke patients ascertained in years 1 to 4 of OXVASC. % calculated from denominator of 337 available TIA patients.

The TIA and minor stroke patients are very similar in age, gender and ethnicity. 290 patients (38%) were aged >80. The prevalences of co-morbid vascular disease and traditional vascular risk factors are similar in both groups. However, the presence of such conditions and risk factors are obtained from patient report or GP records and represent baseline morbidity and risk from previous health service encounters either symptomatic or
opportunistic, rather than from the results of blood tests and cardiac or vessel imaging after recruitment to the study. As such there may be patients with co-existing risk factors or organ-specific vascular disease that are not captured in the above table.

Tables 6.2 and 6.3 show baseline characteristics for TIA and minor stroke patients respectively with data summarised for patients recruited in each contract time period.

<table>
<thead>
<tr>
<th>Contract</th>
<th>OLD</th>
<th>NEW</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>192</td>
<td>169</td>
</tr>
<tr>
<td>Male n (%)</td>
<td>82 (42.7)</td>
<td>75 (44.4)</td>
</tr>
<tr>
<td>Age mean(SD)</td>
<td>74.0 (12.5)</td>
<td>74.4 (12.0)</td>
</tr>
<tr>
<td>Caucasian n (%)</td>
<td>177 (97.3)</td>
<td>149 (96.1)</td>
</tr>
<tr>
<td>Recurrent n (%)</td>
<td>60 (31.3)</td>
<td>35 (20.7)</td>
</tr>
<tr>
<td>Previous TIA n (%)</td>
<td>43 (22.4)</td>
<td>24 (14.2)</td>
</tr>
<tr>
<td>Previous Stroke n (%)</td>
<td>29 (15.1)</td>
<td>19 (11.2)</td>
</tr>
<tr>
<td>Previous myocardial infarction n (%)</td>
<td>26 (13.5)</td>
<td>18 (10.7)</td>
</tr>
<tr>
<td>Angina n (%)</td>
<td>33 (17.2)</td>
<td>29 (17.2)</td>
</tr>
<tr>
<td>Cardiac Failure n (%)</td>
<td>21 (10.9)</td>
<td>17 (10.1)</td>
</tr>
<tr>
<td>Peripheral Vascular Disease n (%)</td>
<td>11 (5.7)</td>
<td>11 (6.5)</td>
</tr>
<tr>
<td>Hypertension n (%)</td>
<td>104 (54.2)</td>
<td>82 (48.5)</td>
</tr>
<tr>
<td>Atrial fibrillation n (%)</td>
<td>39 (20.3)</td>
<td>25 (14.8)</td>
</tr>
<tr>
<td>Diabetes n (%)</td>
<td>20 (10.4)</td>
<td>18 (10.7)</td>
</tr>
<tr>
<td>Hypercholesterolaemia n (%)</td>
<td>53 (27.6)</td>
<td>62 (36.7)</td>
</tr>
<tr>
<td>Never Smoked n (%)</td>
<td>83 (43.2)</td>
<td>71 (42.0)</td>
</tr>
</tbody>
</table>

Table 6.2 Prevalence of co-morbid cardiovascular conditions and risk factors among TIA patients across old and new contract time periods. Data available from a total of 337 patients across new and old contract periods.
<table>
<thead>
<tr>
<th>Contract</th>
<th>OLD</th>
<th>NEW</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>227</td>
<td>209</td>
</tr>
<tr>
<td>Male n (%)</td>
<td>112 (49.3)</td>
<td>110 (52.6)</td>
</tr>
<tr>
<td>Age mean(SD)</td>
<td>74.4 (11.6)</td>
<td>74.0 (12.4)</td>
</tr>
<tr>
<td>Caucasian n (%)</td>
<td>213 (93.8)</td>
<td>200 (95.7)</td>
</tr>
<tr>
<td>Recurrent n (%)</td>
<td>51 (22.5)</td>
<td>25 (12.0)</td>
</tr>
<tr>
<td>Previous TIA n (%)</td>
<td>35 (15.4)</td>
<td>23 (11.0)</td>
</tr>
<tr>
<td>Previous Stroke n (%)</td>
<td>51 (22.5)</td>
<td>24 (11.5)</td>
</tr>
<tr>
<td>Previous myocardial infarction n (%)</td>
<td>32 (14.1)</td>
<td>21 (10.0)</td>
</tr>
<tr>
<td>Angina n (%)</td>
<td>39 (17.2)</td>
<td>24 (11.5)</td>
</tr>
<tr>
<td>Cardiac Failure n (%)</td>
<td>28 (12.3)</td>
<td>15 (7.2)</td>
</tr>
<tr>
<td>Peripheral Vascular Disease n (%)</td>
<td>23 (10.1)</td>
<td>16 (7.7)</td>
</tr>
<tr>
<td>Hypertension n (%)</td>
<td>126 (55.5)</td>
<td>124 (59.3)</td>
</tr>
<tr>
<td>Atrial fibrillation n (%)</td>
<td>38 (16.7)</td>
<td>36 (17.2)</td>
</tr>
<tr>
<td>Diabetes n (%)</td>
<td>28 (12.3)</td>
<td>20 (9.6)</td>
</tr>
<tr>
<td>Hypercholesterolaemia n (%)</td>
<td>53 (23.3)</td>
<td>46 (22.0)</td>
</tr>
<tr>
<td>Never smoked n (%)</td>
<td>103 (45.4)</td>
<td>73 (34.9)</td>
</tr>
</tbody>
</table>

Table 6.3 Prevalence of co-morbid cardiovascular conditions and risk factors among minor stroke patients across old and new contract time periods

There is no tendency evident for the prevalence of cardiovascular conditions or risk factors for stroke to increase in the new contract time period, indicating that there is no marked effect of QoF on the recording of risk factors or disease register conditions (although CKD is not included) in this population of patients.

From the perspective of demographics and baseline co-morbidities I conclude that fair comparisons can be made between patients in the old and new contract time periods.

6.3 Missing data and events occurring on holiday

In order to test the effect of the GMS contract and GP opening hours in general on patients’ choice of provider and delay in calling them, I excluded 25 patients who had events whilst out of the UK and where therefore in an unfamiliar healthcare setting at the time of symptom onset. Their characteristics are summarised in table 6.4.

There were no research records to analyse for 4 patients. What information there is on them is included in the demographic table for the total of 797 patients.
Given that these patients had events on holiday they are likely to be more active than the general TIA/minor stroke population and as one would predict they have a lower mean age than the total population. They have very little established cardiovascular disease but risk factors are present.

### 6.4 Provider choice in all TIA and minor stroke patients – GMS contract effects

Table 6.5 shows the numbers (% within each contract time period) of patients whose first clinical assessment was either by primary care (registered GP or OOH GP), ED, mentioned whilst at another outpatient clinic or occurred whilst hospitalised for another reason, and by old and new contract time period. Z tests for the difference between two independent proportions were calculated for GP and ED attendances with two-tailed probabilities of type 2 error in the old and new contract time periods (given the binomial requirement of \( n \, (p) > 5 \) and \( n(1-p) > 5 \) were met, where \( n \) is the denominator for each patient pool in a given contract time period and \( p \) the proportion attending a given healthcare provider). The combined patient group of TIA and minor stroke is presented.
<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>TIA and minor stroke</th>
<th>Z for comparison</th>
<th>P for comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contract</td>
<td>OLD</td>
<td>NEW</td>
<td></td>
</tr>
<tr>
<td>Total n (%) attending provider</td>
<td>GP</td>
<td>306 (75.0)</td>
<td>255 (70.8)</td>
</tr>
<tr>
<td></td>
<td>ED</td>
<td>83 (20.3)</td>
<td>90 (25.0)</td>
</tr>
<tr>
<td></td>
<td>Other Clinic</td>
<td>6 (1.4)</td>
<td>4 (1.1)</td>
</tr>
<tr>
<td></td>
<td>In Patient</td>
<td>13(3.2)</td>
<td>11 (3.5)</td>
</tr>
<tr>
<td>n (%) in-hours</td>
<td>GP</td>
<td>164 (79.6)</td>
<td>137 (75.7)</td>
</tr>
<tr>
<td></td>
<td>ED</td>
<td>31 (15.0)</td>
<td>34 (18.8)</td>
</tr>
<tr>
<td></td>
<td>Other Clinic</td>
<td>3 (1.5)</td>
<td>3(1.7)</td>
</tr>
<tr>
<td></td>
<td>In Patient</td>
<td>8 (3.9)</td>
<td>7 (3.9)</td>
</tr>
<tr>
<td>n (%) out-of-hours</td>
<td>GP</td>
<td>131 (70.8)</td>
<td>113(66.9)</td>
</tr>
<tr>
<td></td>
<td>ED</td>
<td>49 (26.5)</td>
<td>55 (32.5)</td>
</tr>
<tr>
<td></td>
<td>Other Clinic</td>
<td>2 (1.1)</td>
<td>0 (0)</td>
</tr>
<tr>
<td></td>
<td>In Patient</td>
<td>3 (1.6)</td>
<td>1 (0.6)</td>
</tr>
</tbody>
</table>

Table 6.5 Provider choice overall and by whether symptom onset was in hours or OOH and by contract time period

The majority of patients with TIA and minor stroke seek healthcare from primary care. This was found overall within each contract time period and both for events with symptoms starting in the in hours period and in the OOH period. No statistically significant differences were found between the old contract and new contract time periods in the proportions of patients seeking healthcare via primary care or via ED.

6.5 Provider Choice after TIA symptom onset – GMS contract effects

Table 6.6 shows the numbers (% within each contract time period) of patients whose first clinical assessment was by primary care (either registered GP or OOH GP), ED, mentioned whilst at another outpatient clinic or occurred whilst hospitalised for another reason by old and new contract time period. Z tests for the difference between two independent proportions were calculated for GP and ED attendances with two-tailed probabilities of type 2 error in the old and new contract time periods.
### Table 6.6 Provider choice overall and by whether symptom onset was in hours or OOH and by contract time period

<table>
<thead>
<tr>
<th></th>
<th>Diagnosis</th>
<th>TIA</th>
<th>Z for comparison</th>
<th>P for comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Contract</td>
<td>OLD (%)</td>
<td>NEW (%)</td>
<td></td>
</tr>
<tr>
<td>Total n (%)</td>
<td>GP</td>
<td>147 (78.6)</td>
<td>114 (70.8)</td>
<td>1.68</td>
</tr>
<tr>
<td>attending provider</td>
<td>ED</td>
<td>33 (17.6)</td>
<td>42 (26.1)</td>
<td>-1.91</td>
</tr>
<tr>
<td></td>
<td>Other Clinic</td>
<td>4 (2.1)</td>
<td>4 (2.5)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>In Patient</td>
<td>3 (1.6)</td>
<td>1 (0.6)</td>
<td></td>
</tr>
<tr>
<td>n (%) in-hours</td>
<td>GP</td>
<td>78 (84.8)</td>
<td>63 (71.6)</td>
<td>2.15</td>
</tr>
<tr>
<td></td>
<td>ED</td>
<td>11 (12.0)</td>
<td>21 (23.9)</td>
<td>-2.09</td>
</tr>
<tr>
<td></td>
<td>Other Clinic</td>
<td>1 (1.1)</td>
<td>3 (3.4)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>In Patient</td>
<td>2 (2.2)</td>
<td>1 (1.1)</td>
<td></td>
</tr>
<tr>
<td>n (%) out-of-hours</td>
<td>GP</td>
<td>63 (72.4)</td>
<td>49 (70.0)</td>
<td>0.33</td>
</tr>
<tr>
<td></td>
<td>ED</td>
<td>21 (24.1)</td>
<td>21 (30.0)</td>
<td>-0.83</td>
</tr>
<tr>
<td></td>
<td>Other Clinic</td>
<td>2 (2.3)</td>
<td>0 (0)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>In Patient</td>
<td>1 (1.1)</td>
<td>0 (0)</td>
<td></td>
</tr>
</tbody>
</table>

The majority of patients with TIA sought healthcare from primary overall and ED was the first healthcare provider in 26.1% of patients under the new contract and for 17.6% of patients under the old contract. Although this increase was not statistically significant in the combined patient group of both in and out of hours onset of TIA symptoms, the sub group of patients with events occurring in hours did show statistically significant differences between the proportions of patients accessing care from primary care and from ED, with a move away from primary care attendance to ED attendance.

#### 6.6 Provider Choice after minor stroke symptom onset – GMS contract effects

Table 6.7 shows the first healthcare provider assessment by contract time period and whether symptom onset was in hours or OOH for patients with minor stroke. Effect of the change in GMS contract was assessed with z tests and p values are two-tailed.
Table 6.7 Provider choice overall and by whether symptom onset was in hours or OOH and by contract time period

Again the majority of patients with minor stroke seek healthcare from primary care in the first instance overall and in both the in hours and OOH period. There was no effect of the change in GMS contract in proportions of patients within each contract period and time period seeking healthcare from either ED or primary care.

6.7 Use of Primary Care OOH and the GMS Contract

Given that the majority of patients seek healthcare from primary care after TIA or minor stroke and that there is a need to reduce delays to specialist assessment after symptom onset, it is of interest to determine the extent to which primary care is used in the OOH period and if this was affected by the GMS contract.

This analysis was restricted to patients who had events in the OOH period and chose primary care as their first provider for medical assessment. Of the 244 patients (131 under the old contract, 113 under the new contract) a higher proportion of patients used an on call GP after an event OOH under the new contract (20.3% vs 32.4 %, 95% confidence interval of difference = 8.7 – 23.2%, p=0.034).

Thus rather than acting as a barrier to accessing healthcare out of hours, there is suggestive evidence that the provision of a dedicated OOH provider, rather than members of the daytime surgery team working overnight and at weekends, has increased OOH access to primary care, at least after symptoms of TIA and minor stroke.
6.8 NHS Direct

Ten patients called NHS direct over the four year period with variable advice including going directly to ED as well as seeking a routine GP appointment. They are included in the medical provider group in which they first sought care but as a proportion of all cases it is not a significant first contact point for advice after the symptom onset of TIA or minor stroke.

6.9 Demographic Correlates of Provider Choices – Effect of the GMS contract

Table 6.8 summarises numbers and % of patients with that categorical feature present attending either GP or ED for events occurring in hours and OOH, separated by new or old contract time periods. Continuous variables of age and Integrated Measure of Deprivation 2004 (IMD) score are summarised with mean and standard deviation. The majority of patients first seek healthcare from primary care and there is no clear demographic subgroup where this is reversed. Similarly no pattern of age or IMD appears to correlate with provider choice. This general finding of greater primary care use irrespective of age, deprivation, being alone at time of onset, correct recognition of cerebrovascular aetiology of symptoms, place of residence and previous experience of cerebrovascular disease was seen in both new and old contract time periods. Hence the alteration in GP service terms as a result of the new contract has not altered patient choices for provider provision after TIA/minor stroke within the categories assessed.
<table>
<thead>
<tr>
<th>CONTRACT PROVIDER</th>
<th>OLD</th>
<th>NEW</th>
</tr>
</thead>
<tbody>
<tr>
<td>INCIDENT GP</td>
<td>91 (66.4)</td>
<td>85 (58.2)</td>
</tr>
<tr>
<td>INCIDENT ED</td>
<td>42 (30.7)</td>
<td>60 (41.1)</td>
</tr>
<tr>
<td>RECURRENT GP</td>
<td>31 (64.6)</td>
<td>15 (65.2)</td>
</tr>
<tr>
<td>RECURRENT ED</td>
<td>16 (33.3)</td>
<td>7 (30.4)</td>
</tr>
<tr>
<td>ALONE GP</td>
<td>30 (56.6)</td>
<td>23 (54.8)</td>
</tr>
<tr>
<td>ALONE ED</td>
<td>23 (43.4)</td>
<td>19 (45.2)</td>
</tr>
<tr>
<td>NOT ALONE GP</td>
<td>87 (69.0)</td>
<td>72 (61.0)</td>
</tr>
<tr>
<td>NOT ALONE ED</td>
<td>35 (27.8)</td>
<td>45 (38.1)</td>
</tr>
<tr>
<td>TIA suspected GP</td>
<td>41 (66.1)</td>
<td>47 (77.0)</td>
</tr>
<tr>
<td>TIA suspected ED</td>
<td>20 (32.3)</td>
<td>14 (23.0)</td>
</tr>
<tr>
<td>OWN HOME GP</td>
<td>100 (67.1)</td>
<td>86 (61.0)</td>
</tr>
<tr>
<td>OWN HOME ED</td>
<td>46 (30.9)</td>
<td>54 (38.3)</td>
</tr>
<tr>
<td>WARDEN GP</td>
<td>3 (42.9)</td>
<td>2 (40.0)</td>
</tr>
<tr>
<td>WARDEN ED</td>
<td>3 (42.9)</td>
<td>3 (60.0)</td>
</tr>
<tr>
<td>CARE HOME GP</td>
<td>3 (75)</td>
<td>1 (50)</td>
</tr>
<tr>
<td>CARE HOME ED</td>
<td>1 (25)</td>
<td>1 (50)</td>
</tr>
<tr>
<td>AGE</td>
<td>72.2 (12.1)</td>
<td>74.3 (11.4)</td>
</tr>
<tr>
<td>IMD Score</td>
<td>8.83 (6.14)</td>
<td>9.72 (6.66)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8.36 (5.07)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>IN HOURS PROVIDER</th>
<th>OLD</th>
<th>NEW</th>
</tr>
</thead>
<tbody>
<tr>
<td>INCIDENT GP</td>
<td>103 (70.0)</td>
<td>90 (61.6)</td>
</tr>
<tr>
<td>INCIDENT ED</td>
<td>39 (26.5)</td>
<td>52 (35.6)</td>
</tr>
<tr>
<td>RECURRENT GP</td>
<td>36 (62.1)</td>
<td>27 (77.1)</td>
</tr>
<tr>
<td>RECURRENT ED</td>
<td>21 (36.2)</td>
<td>7 (20.0)</td>
</tr>
<tr>
<td>ALONE GP</td>
<td>32 (66.7)</td>
<td>23 (54.8)</td>
</tr>
<tr>
<td>ALONE ED</td>
<td>16 (33.3)</td>
<td>19 (45.2)</td>
</tr>
<tr>
<td>NOT ALONE GP</td>
<td>99 (66.4)</td>
<td>83 (67.5)</td>
</tr>
<tr>
<td>NOT ALONE ED</td>
<td>44 (29.5)</td>
<td>36 (29.3)</td>
</tr>
<tr>
<td>TIA suspected GP</td>
<td>49 (77.8)</td>
<td>40 (67.8)</td>
</tr>
<tr>
<td>TIA suspected ED</td>
<td>14 (22.2)</td>
<td>18 (30.5)</td>
</tr>
<tr>
<td>OWN HOME GP</td>
<td>107 (65.6)</td>
<td>91 (64.1)</td>
</tr>
<tr>
<td>OWN HOME ED</td>
<td>51 (30.7)</td>
<td>47 (33.1)</td>
</tr>
<tr>
<td>WARDEN GP</td>
<td>5 (62.5)</td>
<td>3 (33.3)</td>
</tr>
<tr>
<td>WARDEN ED</td>
<td>3 (37.5)</td>
<td>6 (66.7)</td>
</tr>
<tr>
<td>CARE HOME GP</td>
<td>3 (75.0)</td>
<td>4 (66.7)</td>
</tr>
<tr>
<td>CARE HOME ED</td>
<td>1 (25.0)</td>
<td>1 (16.7)</td>
</tr>
<tr>
<td>AGE</td>
<td>75.7 (10.7)</td>
<td>74.7 (11.7)</td>
</tr>
<tr>
<td>IMD Score</td>
<td>8.83 (5.56)</td>
<td>9.20 (6.26)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>9.42 (4.61)</td>
</tr>
</tbody>
</table>

Table 6.8 n (%) for categorical variables, mean (SD) for continuous variables of age and IMD score for patients choosing GP or ED in hours and OOH, separated by contract time period.

### 6.10 Symptoms and Provider Choice

Table 6.9 summarises the proportion of patients with each symptom who choose either GP or ED for their first healthcare contact after symptom onset, separated by contract time period. Again the majority of patients initially attend primary care.
<table>
<thead>
<tr>
<th>SYMPTOM</th>
<th>Presence</th>
<th>OLD</th>
<th>NEW</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>GP</td>
<td>ED</td>
</tr>
<tr>
<td>Motor</td>
<td>Y</td>
<td>153 (64.6)</td>
<td>46 (31.2)</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>122 (67.0)</td>
<td>74 (25.2)</td>
</tr>
<tr>
<td>Sensory</td>
<td>Y</td>
<td>84 (77.8)</td>
<td>22 (20.3)</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>191 (61.4)</td>
<td>98 (31.5)</td>
</tr>
<tr>
<td>Speech</td>
<td>Y</td>
<td>108 (60.3)</td>
<td>63 (35.2)</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>167 (69.6)</td>
<td>57 (23.8)</td>
</tr>
<tr>
<td>Vision</td>
<td>Y</td>
<td>50 (65.8)</td>
<td>17 (22.4)</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>225 (65.6)</td>
<td>103 (30.0)</td>
</tr>
</tbody>
</table>

Table 6.9 n (%) of patients with symptoms choosing GP or ED in each contract time period
Chapter 7  
Delay in seeking care after TIA and minor stroke

7.1 Time to call for help – Evidence for combining contract time periods

For patients seeking care in primary care after events in hours, median (IQR) time to call under the old contract compared with the new contract was 4.67 hrs (44.8) vs 3.0 hrs (46.0) Mann Witney test z = -0.017 P=0.99. For patients seeking care in primary care after events out of hours, median (IQR) time to call under the old contract compared with the new contract was 13.0 hrs (40.5) vs 10.5 hrs (41.5) Mann Whitney test z = -0.60, P=0.55. Delay to call primary care has not been influenced by the change in GMS contract. The effect of GP opening hours per se will therefore be analysed for the combined old and new contract time periods.

A similar analysis for patients who sought medical attention in secondary care showed no significant difference between the two contract time periods with median (IQR) time to call for help under the old contract of 0.75 hrs (2.92) vs 0.71 hrs (2.46) under the new contract, Mann Whitney test z = -0.38, P = 0.70.

7.2 Time to call for help – Combining Incident Events with Recurrent Events

Median (IQR) time to call for medical attention for patients with incident events was 3.5 hrs (24.6) and for patients with recurrent events was 3.2 hrs (26.3), Mann Whitney test z = -0.16, P = 0.87.

Given that there is no statistically significant difference between these groups of patients, patients with incident and recurrent events will be analysed together.

7.3 Time to call for help - Primary vs Secondary Care, Deprivation and Clinical Features

Delay data were available in 721 patients (94% of the 768 included in the analysis of provider choice). Of these, the imputation methods for calculating delay were required in 63 patients (8.7 % of the delay data). Combining new and old contract time periods, there were remarkable differences in the times to call for medical attention between those seeking care from primary care and those from secondary care. Including all events, the median (IQR) time to call for help from primary care was 11.0 hrs (46.0) and from ED was 1.0 hrs (4.0), Mann Whitney test z = 9.17, P<0.001.

Table 7.1 shows the times to call for medical attention by deprivation quartile and by clinical features.
### Grouping

<table>
<thead>
<tr>
<th>Grouping</th>
<th>Median (IQR) hours</th>
<th>p for comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEPRIVATION</td>
<td>Mean (SD) IMD Score</td>
<td></td>
</tr>
<tr>
<td>Least deprived Q1</td>
<td>3.6 (2.7)</td>
<td>5.1 (28.1)</td>
</tr>
<tr>
<td>Q2</td>
<td>6.2 (0.7)</td>
<td>5.3 (31.3)</td>
</tr>
<tr>
<td>Q3</td>
<td>10.0 (1.5)</td>
<td>6.1 (33.5)</td>
</tr>
<tr>
<td>Most deprived Q4</td>
<td>17.0 (5.1)</td>
<td>7.5 (42.0)</td>
</tr>
</tbody>
</table>

### CLINICAL FEATURES

<table>
<thead>
<tr>
<th>Feature</th>
<th>Present</th>
<th>Absent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Motor symptoms</td>
<td>2.1 (20.1)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Sensory symptoms</td>
<td>7.5 (37.2)</td>
<td>0.12</td>
</tr>
<tr>
<td>Visual disturbance</td>
<td>12.5 (44.0)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Speech disturbance</td>
<td>2.5 (19.9)</td>
<td>0.002</td>
</tr>
</tbody>
</table>

### Table 7.1 Median (IQR) delay to call for medical assistance by deprivation quartile and by presence of clinical features (as recorded in secondary care)

There was no significant effect of deprivation on time to call for help but significantly quicker times to call for medical help were found for the presence of motor symptoms and speech symptoms. Visual symptoms were associated with a significantly longer time to call for help and sensory symptoms were associated with a non-significantly longer time to call.

#### 7.4 Time to call for help – Influence of GP Opening Hours

In order to assess the effect of opening hours on delay to calling for help, events were classified according to whether symptom onset occurred during the GP routine hours of opening as defined by the contract period (i.e. including Saturday morning hours of onset for patients with events during the old contract).

For patients choosing primary care as their provider, median (IQR) time to call for help for events during opening hours was 4.0 hours (44.5) and for events in the OOH period was 12.0 hrs (41.0), Mann Whitney test $z = -2.48$, $P = 0.01$.

Thus patients with events out of hours are waiting significantly longer to contact primary care. However, patients using primary care have a choice of access for out of hours events – either waiting to contact their registered practice or contacting the OOH provider. Those seeking help from primary care during the OOH period (28% of those seeking attention from
primary care after symptom onset out of hours) did so with a median (IQR) delay of 1.0 hrs (2.46) compared with those who waited until their registered practice was open who had a median delay of 24.8 hrs (48.5), Mann Whitney test z = -9.60, P <0.001.

Conversely there was no effect of event timing for delay in patients seeking care from the ED. Median (IQR) delay to call after events in hours was 0.73 hrs (1.63) and out of hours 0.91 hrs (2.35) hours, Mann Whitney test P=0.751.

7.5 Pattern of Calling for Help – Opening Hours Influence

GP opening hours appear to have a greater influence on delay to call rather than the choice of provider, particularly for events occurring when practices are closed. In order to assess the precise influence of opening hours, the pattern of calling for help is required. If opening hours were truly a significant determinant of healthcare seeking behaviour, patients would call at the first available opportunity when the practice reopens. If patients who had waited to call did so without a particular urgency i.e. at random points after practices had re-opened then it is harder to argue that opening hours per se are having a significant influence on the length of time between symptom onset and initiating a healthcare system response.

Figure 7.1 shows the length of time between the onset of symptoms and the time of the call to primary care for the first 72 hours of call delays, in the group of patients who had events in the out of hours period but waited until their practice was open before making contact with primary care.

Patients with events on Monday to Thursday overnight and on Sunday overnight show a delay to call that appears to decrease linearly as time of symptom onset approaches the earliest time that the practice is open i.e. 08:00 in the morning. Similarly patients with events during the day on Sunday wait longer and those on Saturday wait longer still.
In order to investigate why the scattergram in figure 7.1 appears to demonstrate a simple relationship between length of delay and time of symptom onset, the distribution of times that patients called the surgery are presented in figure 7.2. It demonstrates that the vast majority of patients are indeed calling at the very first opportunity in the morning when the practice re-opens.
This suggests that GP opening hours are a very significant determinant of patients’ healthcare seeking behaviour after TIA or minor stroke. Those who choose to seek medical attention from primary care after events in the OOH period and wait to do so at their registered practice, call at the earliest opportunity. Patients may well know that they have a problem that requires medical attention and the fact that the delay to call is dependent on opening hours suggests that this is a barrier to meeting Department of Health Stroke Strategy Targets for specialist assessment.

7.6 Early Recurrent Strokes – Do GP Opening hours contribute?

Of all the patients with major stroke who were ascertained in OXVASC from 2002 to 2006, 37 had a previous TIA or minor stroke for which they did not seek medical attention. Of these, 13 had a TIA or MIS in the OOH period and a recurrent stroke occurring before practice reopening.

Of the 13 patients with TIA and minor stroke occurring OOH, five had a TIA or minor stroke in the extra period of opening hours that a ‘Darzi centre’ would offer. Thus five strokes
occurred in a population base of 91,000 people that may have been prevented by extending opening hours as long as urgent treatment and assessment could be arranged in secondary care.

7.7 Pattern of calling for help – In Hours Events

The distribution of delays in calling for help after events that occur when practices are open is informative. Although the above analysis demonstrates that for patients seeking care from general practice, opening hours clearly influence behaviour for events out of hours, the relationship does not appear to be the same for in hours events. Figure 7.3 demonstrates the distribution for the first 72 hours of delays in seeking medical attention from general practice after in hours events. There is a more diffuse spread of delays with gaps occurring due to practice closure. Not all patients take the earliest opportunity to call for help and some wait until the following day or day after that, even though there is ample opportunity to request care from general practice.

**Figure 7.3 Delay in calling GP after TIA/Minor Stroke occurring in-hours**
Chapter 8  Missed TIA in Primary Care

8.1 Introduction

One outcome of the initial interface with the healthcare system after TIA is that the diagnosis may be missed by the treating clinician. This chapter will present an analysis of high risk TIA cases missed in primary care. Derivation datasets for diagnostic prediction rules based on clinic-referred patients are subject to a bias in that they do not contain missed TIA cases, as they are not referred from primary care. Furthermore, the implementation of diagnostic prediction rules will not automatically improve recognition of missed TIA as the clinician who first interprets the patients' narrative needs to suspect the diagnosis in order to use a diagnostic rule.

The patients that are included in the analysis below are those who had a recurrent stroke within 30 days of a TIA but presented to medical attention in the time window after TIA and before the recurrent stroke. Using clinical features of missed high risk cases i.e. associated with early recurrent stroke, a pilot vignette study in GP trainees was developed and implemented to ascertain the design parameters for a definitive larger study of recognition of high risk TIA cases.

8.2 Recurrent Stroke after TIA in OXVASC

Data were available for all patients with ischaemic stroke after TIA, ascertained for the first eight years (2002 to 2010). A time to recurrent event of ≤ 30 days after TIA was chosen because it could be argued that some presentations, if atypical, are less likely to represent a high risk TIA if the recurrent stroke occurs beyond this time window. There were 159 patients with TIA and a recurrent stroke ≤ 30 days in the first eight years of OXVASC. Their data is summarised in table 8.1.

The two patients who did not have a diagnosis of TIA but were referred to another specialist, were both referred to ophthalmology and had recurrent posterior circulation strokes with ophthalmic symptoms as part of their TIA presentation. Five patients were diagnosed but not referred, either because the GP escalated pre-existing vascular risk reduction medication or because treatment was deemed futile in the face of pre-existing advanced chronic disease (e.g. dementia, terminal cancer).
<table>
<thead>
<tr>
<th>Total Number with stroke at ≤ 30 days</th>
<th>159</th>
</tr>
</thead>
<tbody>
<tr>
<td>No medical attention before recurrent stroke</td>
<td>60</td>
</tr>
<tr>
<td>Medical attention before recurrent stroke</td>
<td></td>
</tr>
<tr>
<td>First presentation to A&amp;E</td>
<td>27</td>
</tr>
<tr>
<td>First presentation to GP</td>
<td>45</td>
</tr>
<tr>
<td>Event whilst in patient</td>
<td>9</td>
</tr>
<tr>
<td>Notes unavailable</td>
<td>17</td>
</tr>
<tr>
<td>Outcomes in those first presenting to GP</td>
<td></td>
</tr>
<tr>
<td>Diagnosis of TIA - Referral to TIA clinic/admit</td>
<td>31</td>
</tr>
<tr>
<td>No diagnosis of TIA - Referral to other specialist</td>
<td>2</td>
</tr>
<tr>
<td>Diagnosis of TIA - No referral</td>
<td>5</td>
</tr>
<tr>
<td>No diagnosis of TIA - No referral</td>
<td>8</td>
</tr>
</tbody>
</table>

Table 8.1 Numbers of patients with stroke ≤ 30 days after TIA accessing healthcare through different routes

Eight patients with early recurrent stroke were seen in primary care after a transient event which was not considered to be TIA by the treating GP, but these events were deemed to be TIA by a senior neurologist (Professor Peter Rothwell, after reviewing the primary care records and the subsequent history taken in secondary care). These eight patients therefore did not have investigations or specialist assessment and went on to have a recurrent stroke within 30 days of TIA. Of these, seven records had sufficient detail to be used to construct a case vignette to examine the decision making of GPs in training when confronted with that particular clinical presentation. 17 records were not available to analyse, and there may be more patients seen in primary care where the presenting high risk TIA was not diagnosed.

8.3 ‘Hidden TIA’ and prediction rule development

High risk patients correctly identified in primary care with TIA who are referred but have a recurrent stroke before being seen in clinic are ascertained as a stroke in OXVASC. They are not missed, but can be considered ‘hidden’ from prediction rule derivation as such cases do not appear in TIA clinic derivation datasets. However, the numbers are small when taken annually (31 over 8 years, approximately 4 per year). The clinical features of these patients with TIA will require further study, to find out if they would be identified by existing decision rules as a TIA, and also identified as high risk.

8.4 Other missed opportunities – intermediary specialist referral

Detailed referral data were collected from the first four years of OXVASC and including all TIA and minor stroke (797 patients in total). There were 15 patients who had an intermediary
specialist referral after presentation to primary care, 14 patients were referred to ophthalmology and 1 referred to cardiology. Of the 14 sent to ophthalmology, 6 had amaurosis fugax, 4 had posterior circulation TIA and 4 had posterior circulation stroke.

Table 8.1 shows intermediary referral for the patients with recurrent stroke and it has a relatively small contribution to delay to specialist assessment - two cases in the eight year dataset that had an early recurrent stroke.

8.5 Development of case vignettes for assessing diagnosis

In seven of the high risk TIAs missed in primary care, sufficient detail was present in the record of the GP consultation to be used to construct a vignette. In each of the cases in Appendix 3 I have suggested a hypothesis concerning one aspect of context or clinical feature which may influence decision making, and hence variation in the identified parameter can be used to construct alternative cases.

The cases appear to be mainly from anterior circulation symptoms with one involving the posterior circulation. The anterior cases are either transient weakness/sensory loss or speech disturbance together with some element in the history that has steered the GP away from a vascular cause, or a higher function TIA reflecting dysfunction in parietal cortex, rather than the M1, S1 or speech cortical area dysfunctions (which produce the ‘classical’ TIA picture).

8.6 Distractor cases and questionnaire content

In order to reduce the chance that participants in the TIA recognition study from case vignettes would be cued to TIA diagnosis, I created distractor cases of commonly presenting conditions in the elderly. One of the distractor cases acts as the paired case for TIA to test whether the presence of a focal feature to a non-focal presentation alters management. The cases are in Appendix 3.

8.7 Questionnaire Content

In order to create a questionnaire that was feasible in length and had a TIA prevalence of not more than 50% of questions to avoid cueing, ten vignettes were selected. Three TIA cases along with their alternative hypothesis driven vignettes were included. Firstly, an anterior circulation TIA during correction of hyperglycaemia (Appendix 3, case 2 matched with similar symptoms during persistent hyperglycaemia), a posterior circulation TIA from the history of a carer but denied by the patient (Appendix 3, case 4 matched with similar symptoms with both carer and patient in agreement) and a further anterior circulation TIA of dressing
apraxia presenting as confusion (Appendix 3, case 3 matched with a non-TIA case of confusion alone). The remaining four distractor cases were added.

In response to reading the vignettes, the key questions to answer are ‘What are you going to do now?’ and ‘What do you think is the most likely diagnosis?’. The former is more important as there are a number of reasons for referral and ‘rule out’ of an important but unlikely condition is part of the function of primary care. Thus a GP could think that TIA is not the most likely diagnosis, yet feel that a specialist opinion from a TIA clinic would be beneficial, given the sequelae of failing to diagnose TIA and control vascular risk. More strokes will be prevented from referral of true TIA that GPs think is unlikely and hence the most important outcome of a clinical encounter after TIA is referral to a TIA clinic, even if the GP does not suspect the condition particularly strongly.

Free text responses were given so that respondents to the questionnaire could write as much as they wished and so that diagnostic choices were not constrained. Two academic GPs in the Department of Primary Care Health Sciences took the questionnaire and a time of 25 minutes for completion was feasible.

The questions were presented in randomised order to each respondent via ‘surveymonkey’ web hosting via the link https://www.surveymonkey.com/s/DNS6K2N. The questionnaire invitation emails to trainee GPs, front sheet for study details and questionnaire content are in the appendix.

8.8 Response and completion rate

The design of a large scale national assessment of GP trainees’ response to vignettes of missed TIAs will require an estimate of response rate to an invitation to participate as well as the completion rate of all questions. Given the randomised order of questions, the fraction of fully completed questionnaires is important as the completion of matched pairs to ascertain if certain parameters are associated with failure to recognise TIA may rely on fully completed questionnaires.

Of 54 GP trainees in the Oxford specialist GP training scheme, a total of 19 started the questionnaire, giving a response rate of 35%. A reminder email was sent out two weeks after the initial invitation to take part in the questionnaire, which resulted in two additional responses. Thus the majority of responses were initiated after the first email invitation with little effect of the reminder email. The inference is that those who did not start the questionnaire did not simply forget to respond after the initial invitation.
Of the 19 trainees responding, 11 completed the questionnaire in full, giving a completion rate of 58% for those starting the questionnaire. Three respondents (16%) clicked on the link but did not respond to any vignettes.

8.9 Demographic of Respondents

Table 8.2 gives the breakdown in the respondents of the year of training and the last recognised training post in elderly medicine. Owing to concerns about confidentiality of responses, sex and years since qualification were not collected.

<table>
<thead>
<tr>
<th>YEAR OF TRAINING</th>
<th>None</th>
<th>F1</th>
<th>F2</th>
<th>ST1</th>
<th>ST2</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>ST1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>ST2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>ST3</td>
<td>4</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td>ST4</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>TOTAL</td>
<td>6</td>
<td>6</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>19</td>
</tr>
</tbody>
</table>

Table 8.2 Respondents by year of training and most recent training post in elderly medicine

8.10 Responses to vignettes – TIA diagnosis

Six paired questions tested the influence of alteration of one clinical parameter on recognition of TIA, as well as if the GP trainees would also recognise TIA in the cases that had been missed in primary care.

<table>
<thead>
<tr>
<th>Case (parameter changed)</th>
<th>All respondents</th>
<th>Respondents answering both cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case 1 (hyperglycaemia correction)</td>
<td>2/13</td>
<td>1/11</td>
</tr>
<tr>
<td>Matched case</td>
<td>7/12</td>
<td>6/11</td>
</tr>
<tr>
<td>Case 2 (patient/carer disagreement)</td>
<td>7/11</td>
<td>8/11</td>
</tr>
<tr>
<td>Matched case</td>
<td>10/12</td>
<td>9/11</td>
</tr>
<tr>
<td>Case 3 (dressing apraxia)</td>
<td>4/15</td>
<td>3/11</td>
</tr>
<tr>
<td>Matched case</td>
<td>1/12</td>
<td>1/11</td>
</tr>
</tbody>
</table>

Table 8.3 TIA diagnosis (n/N responding) in cases (originally missed in primary care) and in matched cases generated by altering one parameter
Table 8.3 shows that TIA was diagnosed by a greater proportion of the trainees when symptoms where associated with persistent hyperglycaemia than corrected hyperglycaemia. There was almost equivalent diagnosis of TIA irrespective of whether the patient and carer agreed on the presence of symptoms and a smaller proportion of trainees diagnosed TIA when there was confusion alone rather than confusion and dressing apraxia.

### 8.11 Response to vignettes – management

<table>
<thead>
<tr>
<th>Case (parameter changed)</th>
<th>All respondents</th>
<th>Respondents answering both cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case 1 (hyperglycaemia correction)</td>
<td>3/13</td>
<td>2/11</td>
</tr>
<tr>
<td>Matched case</td>
<td>6/12</td>
<td>4/11</td>
</tr>
<tr>
<td>Case 2 (patient/carer disagreement)</td>
<td>7/11</td>
<td>8/11</td>
</tr>
<tr>
<td>Matched case</td>
<td>8/12</td>
<td>8/11</td>
</tr>
<tr>
<td>Case 3 (dressing apraxia)</td>
<td>4/15</td>
<td>3/11</td>
</tr>
<tr>
<td>Matched case</td>
<td>1/12</td>
<td>2/11</td>
</tr>
</tbody>
</table>

Table 8.4 Specialist referral (n/N responding) in cases (originally missed in primary care) and in matched cases generated by altering one parameter

Table 8.4 shows the numbers of trainees that would refer either to a TIA clinic or acutely for further management. There are less marked differences in proportions of trainees referring patients as a result of alteration of vignette parameters, compared with Table 8.3.

### 8.12 Vignette responses - diagnosis of TIA and referral for investigation

<table>
<thead>
<tr>
<th>Total number of TIA diagnoses</th>
<th>30</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decisions to refer after TIA diagnosis</td>
<td>25</td>
</tr>
<tr>
<td>Failure to refer after TIA diagnosis</td>
<td>No. of trainees not referring in 2 cases</td>
</tr>
<tr>
<td></td>
<td>No. of trainees not referring in 1 case</td>
</tr>
</tbody>
</table>

Table 8.5 Decisions to refer after TIA diagnosis.

Table 8.5 shows the conversion from TIA diagnosis to referral. Referral was recommended in 25/30 diagnoses (83%). Two trainees did not recommend referral after making a TIA
diagnosis on two occasions, and one trainee did not recommend referral after one TIA diagnosis.

8.13 Feedback from respondents

After a four week period of availability of the internet survey, a summary of the rationale for the study, aggregated group performance and invitation to give feedback about the study utility for personal learning was sent to the whole group of GP trainees. In order to maintain confidentiality of responses, the email addresses of those responding were not recorded and so feedback could only be sent to the group of initial invitees.

2/19 responded to the feedback questionnaire. Both respondents found answering the vignettes useful but found being given the aggregated group performance more useful. Free text comments in general were invited, and one respondent replied saying that the questions appeared very similar and that they were not sure if giving similar answers was right. However, this respondent found the feedback about the cases very useful.
Chapter 9  Summary and Conclusions

9.1  Healthcare-seeking behaviour after TIA and minor stroke – primary care structure, delivery and future research

9.1.1 Influences on healthcare access after TIA or minor stroke

Healthcare access after TIA or minor stroke is influenced by choice of healthcare provider, time of symptom onset and clinical features of the event but not by deprivation or previous history of cerebrovascular disease. The choice of ED or GP was associated with substantial differences in median delay to access medical care. Patients seeking care from general practice have a median delay that is around three hours greater that those seeking care from ED. Furthermore the delay in care seeking for those with events out of hours is substantially greater than for those seeking care after events in hours, but only for patients who first seek care in general practice.

Anterior circulation symptoms of speech and unilateral weakness were associated with shorter delays to call for medical attention than patients with visual deficit or sensory disturbance which has been noted before (82) and interviews with the general public have identified that visual symptoms are least likely to prompt an urgent response (85).

There was no influence of deprivation on median delay as assessed in quartiles of postcode deprivation scores unlike previous stroke studies which found lower social position to increase delay in accessing healthcare (71). This could be due to a low median IMD score in the most deprived group in OXVASC (17 points) compared with the most deprived scores in England of over 40.

The majority of demographic variables were not distributed differently between those patients seeking care from GPs and those seeking care from ED. Deprivation, and degree of dependency in the community as assessed by class of residency did not appear to influence where patients sought care. Fewer patients who were alone at the time of onset sought care from GPs in the out of hours period but the in hours period showed similar proportions of patients seeking care irrespective of the presence of bystanders. This is particularly relevant for delay given the impact that ‘bystanders’ can have on delay to seeking care (76) although no significant differences were found in times to call for medical attention between those who were and were not alone.

Timing of events in the 24 hour cycle with respect to opening hours of routine services and use of different healthcare providers has not previously been examined for patients with TIA and minor stroke. Although calendar day of event and day of seeking care has been
examined in a combined group of routine NHS and OXVASC patients (43), delay as a continuous variable has not been previously reported in OXVASC after TIA and has been mainly reported in cohorts of stroke patients particularly in relation to opportunities for thrombolysis (71;75;78) with few studies including patients with TIA (82).

The importance of practice opening hours in influencing care seeking behaviour is shown most simply by the scattergram of time from symptom onset to calling for medical attention against time of day of onset for those with events out of hours who seek care from general practice. Given that patients seeking care in general practice with events occurring in the out of hours period wait until the earliest opportunity to call for help, the practice opening hours are influencing their care seeking behaviour. Also, some patients seek care in the days after their TIA and the gaps on the scattergram for those with events in hours correspond to times of closure of the practice. There are no comparable studies which have investigated how practice opening hours influence care seeking behaviour for patients with TIA in other healthcare settings.

9.1.2 The influence of the GMS contract on healthcare access after TIA or minor stroke

The change in general practitioner contract has had a small and positive influence on healthcare access after TIA and minor stroke. The new contract has not affected the perception for both in and out of hours events that primary care is the first point of interaction with healthcare services. Although there was much media speculation about how the change in access to registered practices overnight and at weekends would alter availability of primary care, the use of primary care in the out of hours period actually increased after the introduction of the new contract in the OXVASC patients. This suggests that patients have fewer barriers to seeking an opinion from primary care out of hours than before the introduction of the contract.

9.1.3 Limitations of this analysis

Call time data were not available for all patients and where there was sufficient evidence call time data were imputed using the methods in Chapter 5. However, only a small proportion of the data were imputed and their exclusion from the dataset did not affect the magnitude of the difference between groups. Also, it could be argued that there may be potential recall bias and recording bias for digit preferences of times to call and onset times which will create a distribution of call times that are not smoothly continuous. However, patients were seen rapidly after their events when data were recorded and this methodology is similar to those of other reports of delays in seeking care for the stroke cohorts identified from acute hospital admissions.
No multivariable modelling of all potentially relevant factors was carried out on the OXVASC data. Specific hypotheses were tested over deprivation, clinical features and service availability, so the potential interaction between variables has not been assessed. It could be argued that a complete understanding of the causes of delay requires knowledge of how these variables interact. The importance of the route to accessing healthcare in determining delay was demonstrated in one study which showed that on multivariable analysis, the only factors which remained as predictive of delay after stroke were choice of provider and bystander response (78). However, the variables chosen in this analysis have been pre-selected from rational hypotheses of causation of delay rather than from in depth interviews with patients who are able to describe why they may have acted as they did. Even if a complex pattern of interacting variables could be determined it would not be clear how that could be used to inform educational messages for the public or for patients determined to be at high risk of stroke or TIA.

9.1.4 Future research and practice

The importance of delay during the out of hours period was seen in the OXVASC cohort with a number of recurrent strokes observed where patients had events in the OOH period and did not have an opportunity to seek care from registered general practices. A small number of these occurred during times that newer primary care service models would have been open but the protective effect of vascular risk reduction after TIA/minor stroke requires urgent treatment and therefore secondary care capacity would need to be able to meet demand from seven day general practice if extended hours were to contribute to reducing the burden of stroke.

Public education policies have failed in as much as patients who correctly identify symptoms mainly attend primary care, even when they know they may have to wait after events in the OOH period. Similarly prior experience of a cerebrovascular event (patients with recurrent rather than lifetime incident events) does not lead to the majority of patients seeking care from ED, indicating that any previously received education about how to respond to recurrent events during previous episodes of care are not effective.

Although other patients in Oxfordshire identified by referral to NHS TIA clinics also attend initially in primary care (43), this may be a UK phenomenon as an international healthcare seeking behaviour comparison for presentation after TIA found that most patients with TIA in Canada attend ED (141).

A qualitative analysis of delay after TIA or minor stroke would inform why there are the observed associations of delay with clinical factors and service delivery factors.
Understanding delay after the onset of symptoms of myocardial infarction has been explored with qualitative analysis of patients and their partners’ experiences of decision making (260-262) as well as in mixed methods studies categorising patterns of responses and predictors of those responses (263). The qualitative stroke literature is smaller by comparison (264;265) and with no reports of qualitative analyses of delay after TIA. The quantitative findings in this thesis could inform the sampling of a qualitative study to include patients with events out of hours who sought care in primary care using on call services and those waiting until their practice re-opened. The quantitative analysis could be repeated with a later cohort, as the most recently recruited patients were from six years ago, to investigate whether patterns of consulting have changed with time.

Thus, one of the key determinants of healthcare seeking behaviour after TIA or minor stroke is related to time of onset of symptoms and the availability of different parts of the healthcare system, particularly the most familiar setting where patients seek care. These factors have not been explored in detail in existing studies of delay after TIA or stroke and there is little qualitative work to inform education around how to recognise and respond to symptoms of TIA or minor stroke. Nevertheless, this thesis has identified an important aspect of primary care’s contribution to delay in specialist assessment after TIA and stroke – most patients choose to be seen first in primary care and the structure and delivery of primary care results in some patients being subject to long delays in a potentially high risk window for recurrent stroke.

9.2 Hidden and Missed TIA – scale and recognition

9.2.1 GP trainees’ recognition of high risk missed TIA

The majority of GP trainees did not suspect TIA in two of the missed cases and in these cases referral for further investigations and specialist opinion were only recommended by a minority of respondents. Although some trainees did suspect TIA in two of the missed cases, overall the results suggest that education of trainees over broader presentations of TIA may be lacking, but a larger scale study is required to assess this.

The high risk missed TIA cases were identified from an analysis of routes to care for patients with early recurrent stroke after TIA. This showed that patients presenting to primary care may not be included in clinic based prediction rule derivation datasets, either because the diagnosis is not suspected by the treating GP or a referral is made and a recurrent stroke occurs before being seen to confirm the diagnosis. Furthermore, delay to optimal treatment could occur due to intermediary specialist referral (e.g. ophthalmology). Although the absolute numbers affected were small in this dataset, the presenting cases are all
associated with recurrent stroke and are therefore important to recognise. Firm conclusions over whether the key problem is deficiency in training or lack of clinical exposure post training can’t be drawn as this study was performed to pilot the methodology and determine recruitment and completion rates for an appropriately powered test of the question.

9.2.2. Response and completion rates for a web based survey

The response rate was small, and approximately half the trainees completed the survey with paired data, which is essential to answer questions where vignettes are constructed to test the effect of change in one parameter on TIA recognition and appropriate management. The reported vignette diagnosis studies had higher rates of response, and with larger sample sizes (203-205) among qualified GPs. Allowing for 50 responses for each matched case, which was found in one study using permutations to test the effect of cues on diagnosis and management of TIA (204), would require 300 invitees if one third respond and half of those complete the questionnaire as per the pilot.

9.2.3 The influence of altering case vignette parameters on TIA recognition and management decisions

The narrative description of the results suggests that single parameter change may be associated with TIA detection and referral. Higher rates of TIA diagnosis and referral were found in the matched cases where alteration of one parameter was hypothesised to increase TIA detection. One pair was constructed to examine if the presence of focal non dominant parietal symptoms in a presentation of confusion was treated any differently to a presentation of confusion alone – although more trainees suspected TIA in the case with parietal symptoms, the majority of trainees did not treat this case differently from transient confusion alone.

9.2.4 Limitations of this analysis

The main limitation is the size of the sample. A modification to improve the numbers completing the survey would be to reduce the number of questions but allow the full range of TIA vignettes to be tested by presenting only one or two randomly chosen matched pairs with an equivalent number of distractors each time the web questionnaire is accessed. This approach was used to allow for 16 permutations from completing four cases in a single questionnaire with 800 respondents (204).

Increasing the recruitment pool entails contacting multiple deaneries for sufficient numbers of GP trainees to be invited to take part. In order to comply with the Data Protection Act, an extension would require formal collaboration with other deaneries and fresh ethical
applications and owing to time constraints this will be taken forward in the research plans following on from this thesis.

9.2.5 Future research and practice

Derivation datasets comprising clinic diagnosed TIA patients exclude the correctly recognised and very high risk who have a stroke after referral but before assessment, and those that are missed by GPs but are high risk and present with recurrent stroke. Dawson et al did not systematically search for these patients to include them in a derivation dataset. These patients were identified in OXVASC and the research steps following on from this thesis will be to understand in more detail the factors that affect recognition (a prerequisite of using a TIA referral/decision support tool) as well as expanding clinic based derivation datasets.

In order to improve recognition of TIA, particularly high risk TIA, the presentations of these patients should be incorporated into derivation datasets. However, this may require further qualitative analysis of textual information, particularly as some non-dominant parietal features may be coded as confusion even if descriptions identify focal deficit such as dressing apraxia. An alternative strategy would be a prospective systematic data collection of all patients who present with a disturbance in their neurological status, either focal or non-focal such as confusion, in order to get an accurate and systematically derived clinical assessment of all presentations that could be relevant to this question along with follow up for outcome – such presentations have been termed transient neurological attacks (TNAs) (266).
SECTION 4  GP and Specialist Disagreement – risk tools and diagnosis
Chapter 10  Methods

10.1 Introduction

The clinical relevance of discrepancies between GP and specialist histories is that a prediction rule derived from specialist records may perform differently in GP records. However if there is less disagreement over clinical features for patients with TIA compared with those without TIA then accurate identification of true high risk may still be possible even if the rule does not perform well in terms of discrimination i.e. for diagnosis. However, the patients who may be truly high risk i.e. have a an ABCD2 score > 4 from a specialist assessment might not be scored as such at the first healthcare contact and may not receive timely assessment and intervention to reduce early stroke risk. Conversely, patients who are at low risk of recurrent stroke may be scored in the higher risk category at the first healthcare contact and as such be fast-tracked to an urgent clinic slot. This is addressed with the following research questions

1. Do GPs’ and specialists’ records of the same patients with suspected TIA agree about the clinical features?

2. Does GP/specialist disagreement vary according to final clinic diagnosis?

3. What impact does GP/specialist disagreement have on the performance of the ABCD2 score for accurate triage of patients with TIA?

To answer these questions, primary care records (referral letters and consultation notes) among all referred patients in the cohort 2002 – 2006 were analysed for symptom content relevant to the ABCD2 score (motor, speech and duration) and compared with those recorded from specialist assessment.

An ideal test of agreement between GPs and specialists of a risk or diagnostic score would randomise the order of history-taking between GP and specialist. In spite of any systematic differences between the history taking of GPs and specialists it is possible that histories can change in the ‘re-telling’ and so testing agreement from a history derived score by having a non-randomised order of clinician for each patient is a potential criticism. Furthermore the history taking is not contemporaneous in most cases as a referral to a clinic is required resulting in a delay between the GP history taking and the specialist history taking. Although OXVASC patients were seen relatively quickly overall and usually the same day in phase2 of the nested EXPRESS study (41), there was still a delay of some hours if not an overnight period in between history taking.
As there is fixed capacity for urgent slots in hospitals with dedicated TIA clinics (e.g. Oxford Radcliffe Hospitals NHS Trust, University Hospitals Birmingham NHS Trust, Addenbrooke’s Hospital NHS Trust, personal communication) significant resource is spent in getting patients seen urgently. Low risk patients if incorrectly identified could saturate urgent clinic capacity thereby reducing space for true high risk patients to be seen.

Improving the accuracy of referrals from primary care both from a risk perspective and diagnostic perspective (reducing the number of referrals for patients who do not have TIA) will contribute to reducing delays to optimal risk reducing interventions. The potential for existing clinical prediction tools to improve discrimination of TIA from non-TIA in primary care is addressed with the following research question

4. What impact does GP/specialist disagreement have on the discrimination metrics of existing clinical prediction tools for TIA diagnosis?

This was answered by validating the Dawson tool and the ABCD2 score for TIA diagnosis in the cohort 2002 – 2006 and comparing the area under the receiver operator characteristic (ROC) curve with scores based on primary care records and secondary care records.

10.2 Calculation of ABCD2 scores

In OXVASC, GPs are not required to complete a proforma with an ABCD2 score calculated or with the component parts for appointment triage via a calculated score. In the following analysis the ABCD2 score was calculated from a combination of sources, the referral letter from the GP and also the electronic record of the consultation notes. Retrospective calculation of ABCD and ABCD2 scores is an established methodology from secondary care notes and has been performed in order to examine long term stroke risk in a population of patients seen in a TIA clinic, with 14 years of follow up (267).

These patients are a subset of those included in the analysis of delay as that was a population based analysis of all patients with incident TIA, irrespective of their route through the healthcare system. The following analysis is restricted to patients who had been referred to the research clinic and as such had consultation data and referral letter data. Inferences are therefore only generalisable to situations where GPs are referring directly to a TIA clinic, rather than including ED assessments or where GPs are referring directly to acute physicians.

All referred patients from primary care both with and without TIA are included in the analysis, in contrast to previously published work (227). If the degree of disagreement is different
between TIA patients and non-TIA patients then this has implications for urgent clinic capacity, particularly if patients without TIA are classed as high risk.

This approach could be criticised as the GPs were not being asked to calculate an ABCD2 score. GPs are often required to provide an ABCD2 score at referral but there is potential for this to lead to a degree of ‘gaming’ as the score is used to facilitate referral, particularly for the non-TIA patients where urgent reassurance may be required for psychological reasons – a systematic over-scoring of risk in non-TIA patients would be evidence of this. The present analysis is not subject to any degree of score optimisation bias as histories were taken without view to score calculation.

10.3 Analysis of disagreement

Disagreement for the ABCD2 score was assessed by examining the spread of secondary care ABCD2 scores for each GP ABCD2 score, for the whole referred population and for the TIA and non-TIA groups separately. Variation in agreement was also assessed within the TIA patients, grouped by anterior and posterior circulation territory. Altman-Bland plots (268) were constructed by plotting the absolute difference between each patient’s GP and secondary care ABCD2 score against the average of the two scores. This gives a graphical display of bias from the mean level of disagreement as well as how the spread of score differences changes with the average score. The standard deviation (SD) of differences was calculated and lines corresponding to mean, and two SDs either side were plotted on the graphs. Altman-Bland plots use the average of two methods of measurement as a proxy for the likely true underlying value, given that the difference between two scores can be related to either score as a statistical artefact (269).

The agreement between the individual clinical components of the ABCD2 score was assessed using Cohen’s kappa for the presence or absence of symptoms. Age, blood pressure closest to the event and presence of diabetes was constant across GP and specialist records. Cohen’s kappa assesses the agreement that occurs over and above that due to chance alone (0 corresponds to chance agreement, 1 to perfect agreement).

10.4 Impact of GP-specialist disagreement

The major consequence of disagreement occurs when GPs incorrectly label patients as low risk and initiate a slower referral mechanism as set out in the National Stroke Strategy. Given the high early risk of stroke, such inappropriately slow referrals could result in recurrent strokes occurring in patients waiting to be assessed without appropriate full risk factor reduction therapies.
The mathematical expectation of strokes occurring in this way was calculated using the distribution of secondary care ABCD2 scores and their associated seven day stroke risks from validation studies (24).

10.5 **ABCD2 as a diagnostic tool**

If the ABCD2 score has any diagnostic utility it is essential to know firstly in which population of patients it has predictive value i.e. GP referrals to a TIA clinic (a screened population) or patients seeking care in general practice after transient events (a relatively unscreened population) and secondly whether within an identical population of patients the scores are different between groups of GPs and groups of specialists.

One could argue that a diagnostic score has most utility when used by those who need support in diagnostic decision making i.e. generalists. Specialists, as a result of acquired knowledge and skills through training are less likely to need tools to support diagnostic decision making. Therefore if the ABCD2 score is to be fairly assessed as a diagnostic tool one should know how well it validates in terms of discriminating TIA/minor stroke from non-cerebrovascular causes of symptoms in a primary care population.

It is not possible to assess the ABCD2 score for diagnostic discrimination in a primary care population presenting after transient symptoms with patients referred routinely to an NHS TIA service. This is because the primary care practitioner must already have at least a reasonable suspicion for TIA in order to generate the referral. Given that the OXVASC clinic has encouraged GPs to send in all patients that they consider to be even remotely at risk, it is likely that within this clinic seen cohort of patients there may be more non-TIA diagnoses than in a usual NHS service clinic i.e. GPs may be sending more patients than they would do otherwise when they have a low index of suspicion for the diagnosis. As such, the OXVASC clinic cohort offers an opportunity to assess the discriminating ability of the ABCD2 score in a population that may be more representative of a primary care population where patients present after transient symptoms than a usual NHS service clinic population.

10.6 **Metrics of validation – calibration and discrimination**

There are two key parameters that measure how well a prediction rule works. Firstly calibration, which corresponds to how well a predicted outcome matches the observed outcome. In the case of scores, where we do not have predicted probabilities to match with observed outcomes within patient groups based on similar predictions, as scores increase we would expect a greater chance of the diagnosis being present, in other words, the positive predictive value of the score should increase with increasing value of the score.
Secondly, a score should discriminate well i.e. the score can partition those with and those without a disease. The most valid form of validation is *external*, in that the score is derived in one cohort of patients and then the calibration and discrimination are tested in another cohort of patients, preferably not within the same recruitment/data gathering process.

Discrimination between TIA and non-TIA was assessed using receiver operator characteristic (ROC) curves. The plot of the ROC curve demonstrates the trade off between increasing detection of TIA i.e. sensitivity with increasing over –diagnosis i.e. false positives (= 1-specificity). The area under the curve (AUC) corresponds to the probability that a patient with TIA chosen at random will have a higher score than a patient without TIA chosen at random, both within the dataset. An AUC of 1 therefore indicates perfect discrimination as all TIA patients have higher scores than all non-TIA patients. An AUC of 0.5 indicates that the ability of the score to discriminate between TIA and non-TIA is no better than chance. The optimal cut point for a score in terms of maximising sensitivity and minimising false positive results is the point on the curve that is nearest to the coordinates (0,1) i.e the upper left hand corner.

10.7 Calculation of the Dawson et al TIA recognition tool

There are a number of validity issues which may limit generalisability of this tool. Firstly, diagnostic drift over the 13 year period of ascertainment in the derivation cohort, as a number of definitions of TIA have been produced over that time and research published prior to the start of the retrospective cohort drew attention to the low diagnostic concordance among specialists over TIA (270). Secondly there is a problem of mixed TIA and minor stroke in the validation cohort. Persisting signs may increase the discriminating ability of the tool. For example, the presence of persisting upper motor neuron (UMN) facial weakness may be clear on examination but not all patients with transient UMN facial weakness would be able to discriminate and correctly recollect the difference between UMN and LMN weakness they experienced during the period of deficit. Although the prevalence of cerebrovascular disease is broadly similar to clinic audit prevalence (271) the results are only generalisable to a population referred to a TIA clinic, rather than to those attending in primary care after transient symptoms or attending the ED.

The authors of the recognition tool suggest two cut points – the optimal point is 6.1, but in order to reflect the misclassification costs of missing TIA, a 2:1 cost ratio was applied, which reduced the optimal score to 5.4.
Nevertheless, as this is the only specific TIA diagnostic tool in the literature, it was externally validated on OXVASC data for TIA patients, to examine its potential role in supporting referral decisions.

Using vascular risk predictors and age, the features from the history which can increase and decrease the probability of the diagnosis are scored according to the table 10.1.

<table>
<thead>
<tr>
<th>Variable</th>
<th>If Yes</th>
<th>If No</th>
</tr>
</thead>
<tbody>
<tr>
<td>History of TIA/Stroke</td>
<td>0.5</td>
<td>0</td>
</tr>
<tr>
<td>Headache</td>
<td>0</td>
<td>0.5</td>
</tr>
<tr>
<td>Diplopia</td>
<td>1.2</td>
<td>0</td>
</tr>
<tr>
<td>Loss Of Consciousness /Presyncope</td>
<td>0</td>
<td>1.1</td>
</tr>
<tr>
<td>Seizure</td>
<td>0</td>
<td>1.6</td>
</tr>
<tr>
<td>Speech abnormalities</td>
<td>1.3</td>
<td>0</td>
</tr>
<tr>
<td>Unilateral limb weakness</td>
<td>1.7</td>
<td>0</td>
</tr>
<tr>
<td>Upper motor neuron facial weakness</td>
<td>0.6</td>
<td>0</td>
</tr>
<tr>
<td>Age</td>
<td>Multiply by 0.04</td>
<td></td>
</tr>
</tbody>
</table>

Table 10.1 Dawson et al TIA recognition score calculated by summing components.

10.8 Validation of the Dawson score – primary care and secondary care records

A diagnostic tool is not useful in the hands of a specialist in a TIA clinic. By definition they use skills and knowledge from training to derive a diagnosis and decision tools are more likely to be used by generalists as they support a decision making process. This was certainly the aim of tool derivation and is also the driving force behind stroke recognition tools for paramedics and generalist clinicians. Thus a validation from the records of those who are likely to use such a tool will be more helpful to judge the potential effect of implementation.

All cases of TIA ascertained in OXVASC for years 1-4 and all patients with suspected TIA referred to the OXVASC TIA clinic were used as the validation data set to examine the calibration and discriminatory performance of the TIA recognition tool using secondary care
records. This allows an initial comparison with the published data on the tool, which had a validation dataset from a TIA clinic using hospital histories to derive the scores.

Histories from GP consultation records and referral letters were used to populate the Dawson et al TIA recognition score. A history of transient facial weakness was deemed to be UMN, as this would be how the score would be used in routine practice and calibration and discrimination metrics would therefore be generalisable.

The population of patients used for the validation are those seen first in primary care and where the GP has suspected TIA and then later assessed by an OXVASC research fellow with final diagnosis decided after discussion with the senior investigating neurologist (PMR). Calibration as assessed by predictive value for TIA diagnosis with increasing score and discrimination as assessed by ROC curves were calculated for secondary care and for primary care histories.

10.9 Score component analysis

The distributions of the individual components of the score were assessed with univariate analysis by examining the frequency of the presence of predictors in non-TIA and TIA patients. Relative risks (RR) of TIA for a given predictor were calculated and p values calculated using chi-square for RR values greater than 0, Fisher's exact test where RR=0 and t-test for differences in mean age.

10.10 Impact on referral and safety

If decisions to refer within the population of patients attending primary care were based on this particular decision rule then its impact can be assessed by using follow up of false negative patients. If high risk TIA patients are missed then the ascertainment of recurrent events in OXVASC will detect this. In other words, making the assumption that patients with scores below the cut point would not have been referred, the impact on the tool in terms of patient safety can be judged by following up those patients to see if any had a recurrent stroke. This is likely to underestimate the safety aspect of implementing the tool, as referred patients with true TIA have been investigated and treated with a combination of vascular risk reducing interventions. Thus only counting the missed patients with recurrent events is likely to under-estimate the risk of using the tool.
Chapter 11  Difference in histories and ABCD2 score

11.1 Measures of risk score agreement

Table 11.1 shows the numbers of patients (TIA and non-TIA) directly referred to the OXVASC clinic from general practice from 2002 to 2006 with ABCD2 score summary statistics calculated from the GP records and letters as well as ABCD2 scores from the standardised OXVASC proforma completed by secondary care clinicians along with their notes and letters.

<table>
<thead>
<tr>
<th></th>
<th>All Patients</th>
<th>TIA Patients</th>
<th>Non-TIA Patients</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>N</strong></td>
<td>515</td>
<td>212</td>
<td>303</td>
</tr>
<tr>
<td><strong>GP ABCD2</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>3.33 (1.5)</td>
<td>3.85 (1.3)</td>
<td>3.19 (1.5)</td>
</tr>
<tr>
<td>Median (IQR)</td>
<td>3 (2)</td>
<td>4(2)</td>
<td>3 (2)</td>
</tr>
<tr>
<td><strong>Specialist ABCD2</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>3.46 (1.5)</td>
<td>3.94 (1.4)</td>
<td>2.9 (1.4)</td>
</tr>
<tr>
<td>Median (IQR)</td>
<td>4 (3)</td>
<td>4 (2)</td>
<td>3 (2)</td>
</tr>
</tbody>
</table>

Table 11.1 ABCD2 scores from GP notes and secondary care records for patients referred to OXVASC.

Mean scores are marginally higher for TIA patients than non-TIA patients for both GP and specialist scores with this difference being more exaggerated for the hospital based ABCD2 scores. However the agreement between GP and secondary care is not reflected in these summary statistics.

Figure 11.1 shows a plot of the distribution of hospital ABCD2 scores for patients with the same GP ABCD2 score, so for all patients with GP scores of 1, 2, etc, the spread of scores that specialists gave can be viewed. Very few patients were referred with either ABCD2 score of 0 or 7, and for GP ABCD2 scores of 1 to 6, the modal hospital ABCD2 score is the same as the GP ABCD2 score with values in the mid range having the widest spread. Clearly there is a large range of disagreement at the level of individual patients that is not reflected in simple summary statistics.
Figure 11.1 Spread of specialist ABCD2 scores for each GP ABCD2 score – all patients

Figure 11.2 shows an Altman-Bland plot of the average of the GP and hospital ABCD2 scores plotted against the difference between the GP and hospital ABCD2 scores (calculated as specialist score – GP score). This method of displaying the data allows for a visualisation of not only the spread of difference but also any evident bias over certain ranges of scores. Markers are weighted by the number of cases with a given disagreement in score at each level of average ABCD2 score. The mean difference was -0.13 points (thick dashed line) with standard deviation 1.14 points (2 SDs above and below the mean shown as thin dashed lines). The maximum difference between specialist and GP was a score of 4 points. Given that age, BP and diabetes were kept equal across the scores (as the BP chosen was the one soonest after the event taken in primary care), the only variables that contribute to the difference are those that are due to the history i.e. clinical features (speech disturbance or unilateral weakness, and level of duration 0-10 minutes, 10-60 minutes, > 60
minutes). The extreme differences of four points occur in less than 5% of the data, as shown graphically as they lie outside two standard deviations from the mean.

Figure 11.2 Specialist ABCD2 score – GP ABCD2 score plotted against average of specialist and GP ABCD2 scores (All TIA and non-TIA patients)

However, the distribution of the differences shows that for higher average scores of ABCD2 the specialist may be scoring the patients at a higher level, and for lower average ABCD2 scores, the GP may be scoring patients at a higher level. The extreme positive difference values lie at higher average scores and extreme negative values lie at lower average scores.

11.2 Variation in agreement with diagnosis

Differences in the level of agreement in score between patients with and without a diagnosis of TIA have implications for using information from primary care referrals to prioritise urgent clinic slots. Figure 11.3 shows an Altman-Bland plot of specialist ABCD2 score – GP ABCD2 score against average of specialist and ABCD2 score for each patient who does not have a
final diagnosis of TIA, with markers weighted by numbers of cases with a given disagreement at each level of average score. The average specialist – GP ABCD2 score is -0.27 (thick dashed line) and the standard deviation is 1.17 (2 SDs above and below the mean shown as thin dashed lines). The spread of disagreement is much wider than in figure 11.4 which shows the Altman-Bland plot for patients with a TIA diagnosis only. The standard deviation is smaller (0.98) and the mean is positive (0.10).

Figure 11.3 Altman-Bland plot of Specialist ABCD2 – GP ABCD2 score against average of Specialist and GP ABCD2 scores (Non-TIA diagnoses)

Comparing the groups of TIA and non-TIA patients, the distribution of differences between specialist and GP ABCD2 scores is significantly different. Means of differences suggest that from clinical histories, specialists when compared to GPs tend to score true TIA patients with a higher ABCD2 score (mean (specialist – GP) score = 0.10, standard error = 0.07) and tend to score non-TIA patients with a lower ABCD2 score (mean (specialist – GP) score = -0.29, standard error = 0.07). An independent samples t test of the distributions of differences in scores between non-TIA and TIA patients shows that there is evidence to reject the null
hypothesis that the distributions are equal ($t = -4.09, df = 496, p<0.001$) assuming that variances are not equal.

This suggests that although there is measurable disagreement between GPs and specialists in the clinical features elicited and recorded that contribute to ABCD2 score calculation there is less disagreement for patients with a TIA diagnosis than those who do not have a final clinic diagnosis of TIA. Furthermore GPs may be overestimating the ABCD2 score in non-TIA patients, raising the prospect that non-TIA patients may be accorded high risk status and given urgent clinic slots as a result.

![Figure 11.4 Altman-Bland plot of Specialist – GP ABCD2 score against average specialist and GP ABCD2 score (All TIA patients).](image)

11.3 The influence of arterial territory

Measures of disagreement could also vary within the group of TIA patients. Posterior circulation TIAs have the same level of high early risk of recurrent ischaemic stroke as anterior circulation TIAs (230). Thus the ability to correctly attribute high early risk is equally important in posterior circulation TIA patients. Figures 11.5 and 11.6 display the distribution...
of differences between specialist and GP ABCD2 scores across the range of average scores for anterior and posterior TIA patients respectively. Representations of mean, two standard deviations from the mean and the marker size format are as for the preceding plots.

The mean of differences in scores for anterior TIA patients was 0.19 points with SD of 0.90, indicating that specialists tend to higher ABCD2 scores than GPs. For posterior TIA patients, the mean of differences was -0.24 points with SD of 1.02, indicating that specialists tend to lower ABCD2 scores than GPs. Comparing the distributions of the differences between the anterior TIA and posterior TIA patients demonstrates that they are unlikely to come from the same population of differences using an independent samples t test assuming equal variances (t = 2.71 df = 205, p = 0.007).

Figure 11.5 Altman-Bland plot of specialist – GP ABCD2 score against average specialist and GP ABCD2 score (anterior territory TIA)
Given that the four point differences concern three symptom domains, a further analysis is required to determine if there are systematic differences in symptom interpretation across diagnostic groups of patients. The clinical domains are unilateral weakness, speech disturbance and duration of symptoms. The following analysis examines specialist agreement for each domain scored by GPs. 506 patients’ records were available for this analysis.

11.5 Specialist agreement with GP recording of unilateral weakness.

Unilateral weakness was coded as involving the face, upper limb or lower limb. Bilateral weakness is not scored on the ABCD2 score. For the overall group of referred patients, of the 38 patients recorded as having facial weakness in primary care, only 26 were coded as having facial weakness by specialists. Of 468 patients coded as not having facial weakness in primary care, specialists found that facial weakness was in fact present in 30 patients.
Table 11.2 shows agreement over the presence of facial weakness between specialists and GPs assessed with Cohen's kappa (with a value of 0 being at chance, and 1 perfect agreement). Kappa is summarised for the overall group and then within each diagnostic category, for non-TIA patients, anterior TIA patients and posterior TIA patients.

<table>
<thead>
<tr>
<th>TERRITORY</th>
<th>Specialist</th>
<th>GP-Specialist agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Absent</td>
<td>Present</td>
</tr>
<tr>
<td>All Patients</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GP Absent</td>
<td>438</td>
<td>30</td>
</tr>
<tr>
<td>GP Present</td>
<td>12</td>
<td>26</td>
</tr>
<tr>
<td>Total</td>
<td>450</td>
<td>56</td>
</tr>
<tr>
<td>Non-TIA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GP Absent</td>
<td>278</td>
<td>3</td>
</tr>
<tr>
<td>GP Present</td>
<td>5</td>
<td>14</td>
</tr>
<tr>
<td>Total</td>
<td>283</td>
<td>17</td>
</tr>
<tr>
<td>Anterior TIA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GP Absent</td>
<td>121</td>
<td>27</td>
</tr>
<tr>
<td>GP Present</td>
<td>5</td>
<td>12</td>
</tr>
<tr>
<td>Total</td>
<td>126</td>
<td>39</td>
</tr>
<tr>
<td>Posterior TIA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GP Absent</td>
<td>39</td>
<td></td>
</tr>
<tr>
<td>GP Present</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>41</td>
<td></td>
</tr>
</tbody>
</table>

Table 11.2 GP-specialist agreement over presence of facial weakness by diagnostic group

Agreement over presence of this symptom is higher in the non-TIA group but this is partly due to the lower prevalence of this symptom. Table 11.3 shows that for the presence of unilateral limb weakness the agreement for the overall group is similar in the non-TIA group and anterior TIA group. Low symptom prevalences in the specialist group leads to a ‘chance’ value of -0.03 (272), although GPs and specialists agree that the majority of posterior TIA patients do not have unilateral limb weakness.
<table>
<thead>
<tr>
<th>TERRITORY</th>
<th>Specialist</th>
<th>GP-specialist agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Absent</td>
<td>Present</td>
</tr>
<tr>
<td>All Patients</td>
<td>GP</td>
<td></td>
</tr>
<tr>
<td>Absent</td>
<td>344</td>
<td>35</td>
</tr>
<tr>
<td>Present</td>
<td>44</td>
<td>83</td>
</tr>
<tr>
<td>Total</td>
<td>388</td>
<td>118</td>
</tr>
<tr>
<td>Non-TIA</td>
<td>GP</td>
<td></td>
</tr>
<tr>
<td>Absent</td>
<td>219</td>
<td>10</td>
</tr>
<tr>
<td>Present</td>
<td>34</td>
<td>37</td>
</tr>
<tr>
<td>Total</td>
<td>253</td>
<td>47</td>
</tr>
<tr>
<td>Anterior TIA</td>
<td>GP</td>
<td></td>
</tr>
<tr>
<td>Absent</td>
<td>87</td>
<td>24</td>
</tr>
<tr>
<td>Present</td>
<td>8</td>
<td>46</td>
</tr>
<tr>
<td>Total</td>
<td>95</td>
<td>70</td>
</tr>
<tr>
<td>Posterior TIA</td>
<td>GP</td>
<td></td>
</tr>
<tr>
<td>Absent</td>
<td>38</td>
<td>1</td>
</tr>
<tr>
<td>Present</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>40</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 11.3 GP-specialist agreement over presence of unilateral limb weakness by diagnostic group

11.6 Specialist agreement with GP recording of speech disturbance

Table 11.4 displays the same agreement data for the presence of speech disturbance, as recorded by GPs at the initial consultation and by specialists at the second consultation in secondary care. Agreement is generally higher across all groups compared with unilateral facial and limb weakness with the highest agreement in anterior TIA patients. Given that one of the clinical features that leads to a diagnosis of anterior TIA is dysphasia (one of the components of speech disturbance) it is not surprising that the highest agreement is in this diagnostic group.
<table>
<thead>
<tr>
<th>TERRITORY</th>
<th>Specialist Absent</th>
<th>Specialist Present</th>
<th>GP-specialist agreement Kappa</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Patients</td>
<td>GP Absent 337 25</td>
<td>GP Present 39 105</td>
<td>Kappa 0.68</td>
<td>SE 0.04</td>
</tr>
<tr>
<td></td>
<td>Total 376 130</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-TIA GP</td>
<td>GP Absent 230 8</td>
<td>GP Present 26 36</td>
<td>Kappa 0.61</td>
<td>SE 0.06</td>
</tr>
<tr>
<td></td>
<td>Total 256 44</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anterior TIA GP</td>
<td>GP Absent 75 14</td>
<td>GP Present 10 66</td>
<td>Kappa 0.71</td>
<td>SE 0.06</td>
</tr>
<tr>
<td></td>
<td>Total 85 80</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Posterior TIA GP</td>
<td>GP Absent 32 3</td>
<td>GP Present 3 3</td>
<td>Kappa 0.41</td>
<td>SE 0.2</td>
</tr>
<tr>
<td></td>
<td>Total 35 6</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 11.4 GP-specialist agreement over presence of speech disturbance by diagnostic group

It could be argued that GPs may be only eliciting sufficient symptoms to influence decision making e.g. ‘do I have enough information for a referral?’ and as such may not record information beyond a minimum to trigger referral. If unilateral weakness was found this would be enough to refer for suspected TIA without the presence of speech disturbance as well (and although the GPs were not calculating ABCD2 scores, speech adds no extra risk points to unilateral weakness).

Thus GP records may not include certain symptoms if sufficient other symptoms are present to justify referral as further information may not influence decisions. As such, the absence of a symptom may not be evidence of true discordant history taking between GPs and specialists but reflect efficient time management in a consultation.

To investigate this, the 25 patients where GPs did not record speech disturbance but the specialists did record this as being present were analysed for other symptoms. Of the 25, 22 (88%) did not have facial weakness and 18 (72%) did not have limb weakness. This suggests that discordance in history taking is the reason for disagreement, rather than a truncated history taken by the GP once a minimal set of features have been elicited to make a referral decision.
11.7 Specialist agreement with GP recording of duration of symptoms

The ABCD2 score has three levels of duration of symptoms with increasing scores given to longer durations (0-10 minutes = 0, 10–60 minutes = 1, >60 minutes = 2). Duration of symptoms are categorised according to these three levels from GP records and compared with the levels of duration from the specialist notes.

Figure 11.5 demonstrates the variation in duration assessments by GPs and specialists for non-TIA patients, anterior TIA patients and posterior TIA patients. The level of agreement overall is similar within the three diagnostic groups without any evidence of bias in assessment of duration specific to one group. As such it is likely that disagreement in duration of symptoms adds noise rather than poses any risk of either systematically under scoring high risk patients or over scoring low risk patients.

<table>
<thead>
<tr>
<th>TERRITORY</th>
<th>Specialist Duration</th>
<th>GP-Specialist agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>All Patients</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GP Duration</td>
<td>0</td>
<td>54</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>17</td>
</tr>
<tr>
<td>Total</td>
<td>104</td>
<td>129</td>
</tr>
<tr>
<td>Non-TIA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GP Duration</td>
<td>0</td>
<td>31</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>9</td>
</tr>
<tr>
<td>Total</td>
<td>52</td>
<td>75</td>
</tr>
<tr>
<td>Anterior TIA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GP Duration</td>
<td>0</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>Total</td>
<td>37</td>
<td>45</td>
</tr>
<tr>
<td>Posterior TIA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GP Duration</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Total</td>
<td>15</td>
<td>9</td>
</tr>
</tbody>
</table>

Table 11.5 GP-specialist agreement over duration of symptoms (coded as score on ABCD2) by diagnostic group
11.8 Agreement over high and low risk patients

The purpose of the ABCD and ABCD2 scores are to risk-stratify patients at the first point of healthcare contact with a generalist, in order to judge the speed of organising investigations and management from specialty services. Therefore it can be argued that the above analysis of differences in scores between GPs and specialists and potential sources of those differences are only of any real importance if there is significant re-classification of patients from high risk to low risk and vice versa.

Table 11.6 shows the agreement over risk status from ABCD2 scores calculated from GP records and the ABCD2 scores from the specialist records, for all referrals and by diagnostic group. Assuming that the specialist scores are the current gold standard for risk, as these were the derivation and validation sets for the ability of the ABCD2 score to predict early recurrent stroke, the table also allows for calculation of sensitivity and specificity of GPs ability to detect the high risk cases (as defined by specialist taken histories).

GPs show reasonable ability to detect risk status as sensitivity is over 87% for all patients and for anterior TIA patients. The actual number of high risk posterior TIA patients is low so the lower % sensitivity may in this group not be a true reflection of accuracy of detection of risk in posterior TIA.
<table>
<thead>
<tr>
<th>TERRITORY</th>
<th>Specialist risk status</th>
<th>GP Detection Accuracy of high risk features</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LOW</td>
<td>HIGH</td>
<td>Sensitivity</td>
</tr>
<tr>
<td>All Patients</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GP risk status</td>
<td>LOW</td>
<td>212</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td>HIGH</td>
<td>62</td>
<td>207</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>274</td>
<td>236</td>
</tr>
<tr>
<td>Non-TIA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GP risk status</td>
<td>LOW</td>
<td>146</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>HIGH</td>
<td>50</td>
<td>95</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>196</td>
<td>107</td>
</tr>
<tr>
<td>Anterior TIA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GP risk status</td>
<td>LOW</td>
<td>42</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>HIGH</td>
<td>8</td>
<td>104</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>50</td>
<td>116</td>
</tr>
<tr>
<td>Posterior</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GP risk status</td>
<td>LOW</td>
<td>24</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>HIGH</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>28</td>
<td>13</td>
</tr>
</tbody>
</table>

Table 11.6 GP-specialist agreement over risk status according to ABCD2 score (high risk ≥4) by diagnostic group (sensitivity = number of specialist high risk cases identified as high risk by GPs)

11.9 Expected recurrent strokes

Of the 17 high risk patients with TIA that were not identified from GP scores, 6 had an ABCD2 score of 5, with the rest having a score of 4. The recurrent stroke risks at 7 days for ABCD2 scores of 4 and 5 are 5% and 7% respectively (24). Therefore by failing to detect high risk status in these patients one would expect

\[(11 \times 0.05) + (6 \times 0.07) = 0.97\] strokes.

Thus for a 90,000 population over a period of 4 years, there may be one excess stroke due to a failure to detect high risk status in primary care due to arranging time to specialist assessment within 7 days rather than within 24 hours of symptom onset.
Chapter 12  ABCD2 for diagnosis

12.1  Distribution of specialist ABCD2 scores within diagnostic groups

ABCD2 scores were available for 394 patients with TIA and 338 non-TIA patients. Table 12.1 shows the mean specialist ABCD2 scores by diagnosis. Mean ABCD2 scores were higher for TIA than non-TIA patients (4.04 vs 2.86, Mann Whitney test, z = -10.13, p<0.001). Within the TIA group, patients with anterior TIA had significantly higher ABCD2 scores than posterior TIA patients (4.23 vs 3.41, Mann Whitney test, z = -4.5, p < 0.001).

<table>
<thead>
<tr>
<th>Diagnostic Group</th>
<th>N</th>
<th>Mean ABCD2 score</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-TIA</td>
<td>338</td>
<td>2.86</td>
<td>1.39</td>
</tr>
<tr>
<td>TIA</td>
<td>394</td>
<td>4.04</td>
<td>1.48</td>
</tr>
<tr>
<td>Anterior TIA</td>
<td>302</td>
<td>4.23</td>
<td>1.41</td>
</tr>
<tr>
<td>Posterior TIA</td>
<td>92</td>
<td>3.41</td>
<td>1.56</td>
</tr>
</tbody>
</table>

Table 12.1 Mean and SD of specialist ABCD2 score by diagnostic group

The distribution of scores within each diagnostic group is shown in figure 12.1 demonstrating the asymmetric distributions of the anterior and posterior TIA scores. Furthermore the asymmetry is dissimilar as there are opposite skews with anterior TIA patients having a modal score at the upper end of their distribution and the posterior TIA patients having a modal score at the lower end of their distribution.
A formal calibration curve cannot be constructed as the ABCD2 score does not give an expected probability of TIA diagnosis, but the prevalence of TIA can be measured at each value of the score. Increasing scores are associated with higher prevalence of TIA (figure 12.2).
12.3 Discriminating ability of specialist ABCD2 scores for TIA diagnosis

Figures 12.3 to 12.5 display the ROC plots for discriminating ability of specialist ABCD2 scores in separating TIA from non-TIA patients overall and for diagnostic sub groups dependent on arterial territory. The AUC (figure 12.3) for discriminating TIA from non-TIA is 0.71 (S.E. 0.02).

Figure 12.3 ROC curve for discrimination of TIA from non-TIA using specialist ABCD2 scores

The area under the curve (figure 12.4) for discriminating anterior TIA from non-TIA is higher than for the group overall, at 0.75 (S.E. 0.02) and for discriminating posterior TIA from non-TIA (figure 12.5) is 0.59 (S.E. 0.03). Clearly discriminating performance is much worse for posterior TIA and is near chance value.
Figure 12.4 ROC curve for discrimination of anterior TIA from non-TIA using specialist ABCD2 scores

Figure 12.5 ROC curve for discrimination of posterior TIA from non-TIA using specialist ABCD2 scores

12.4 Summary measures of performance at successive cut points

Figure 12.6 shows how the sensitivity and specificity vary with choosing increasing cut points on the ABCD2 score for diagnosis. At the lowest cut point of 0, all patients are labelled as
TIA with the resultant 100% sensitivity as all those who truly have TIA are correctly identified, and specificity is 0% as all those without TIA are falsely identified as having TIA creating a maximum false positive rate. At the highest score of 7, specificity is 100% as patients without TIA all scored less than 7 and so there are no false positives at this cut point. Only a small proportion of patients actually had a score of 7 so choosing this cut point only identifies a small fraction of the total patients with TIA, hence the very low sensitivity. The curves cross at the value of 4 where sensitivity and specificity are approximately equal at 65%.

**Figure 12.6 Sensitivity and specificity for the diagnosis of TIA using specialist ABCD2 scores**

Figure 12.7 shows the change in positive predictive value and negative predictive value with increasing ABCD2 score for TIA diagnosis. At the lowest score of 0, positive predictive value (PPV) is at prevalence of TIA within the sample but as the cut point increases to restrict PPV calculation to patients with higher ABCD2 scores, the value increases but only reaches 100% at the highest score, as clearly there are patients without TIA who are scoring up to 6 points. The negative predictive value is 100% until scores of 2 or more are scored, as no patients with TIA scored 1 or 0.
Figure 12.7 Positive predictive value (PPV) and negative predictive value (NPV) for TIA diagnosis at increasing cut points of specialist ABCD2 score

12.5 ABCD2 scores calculated from GP records within diagnostic groups

Primary care ABCD2 scores were available for 515 patients and mean scores are shown by diagnostic group in table 12.2. Mean GP ABCD2 scores were significantly higher for patients with TIA than patients without TIA (3.85 vs 3.19, Mann Whitney test, $Z = -4.64$, $p < 0.001$). Within the TIA group, patients with anterior TIA had significantly higher ABCD2 scores than patients with posterior TIA (4.02 vs 3.19, Mann Whitney test, $Z = -3.97$, $p<0.001$)

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>N</th>
<th>Mean ABCD2</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-TIA</td>
<td>303</td>
<td>3.19</td>
<td>1.53</td>
</tr>
<tr>
<td>TIA</td>
<td>212</td>
<td>3.85</td>
<td>1.33</td>
</tr>
<tr>
<td>Anterior TIA</td>
<td>169</td>
<td>4.02</td>
<td>1.31</td>
</tr>
<tr>
<td>Posterior TIA</td>
<td>43</td>
<td>3.19</td>
<td>1.24</td>
</tr>
</tbody>
</table>

Table 12.2 Mean (SD) ABCD2 scores from GP records by diagnostic group

The distribution of GP ABCD2 scores within each diagnostic group is shown in figure 12.8. Non-TIA patients have scores throughout the entire range of the ABCD2 scoring system. Within the TIA territory subgroups there are opposite skews for the anterior and posterior TIA patients, with more posterior patients scoring at the lower end of their distribution and more anterior patients scoring at the upper end of their distribution.
Figure 12.8 Distribution of GP ABCD2 scores by diagnostic group

The prevalence of TIA at each GP ABCD2 score point is shown in figure 12.9. This demonstrates that calibration is not as good as with specialist ABCD2 scores, given that there is no linear increase in prevalence with each unit increase in score.

Figure 12.9 Prevalence of TIA at each ABCD2 score – all primary care referrals
12.6 Discriminating ability of GP ABCD2 scores for TIA diagnosis

Figures 12.10 – 12.13 display ROC curves for the discriminating performance of GP ABCD2 scores. For discriminating TIA from non-TIA (figure 12.10), the area under the curve was 0.62 (S.E. 0.03). However, this overall discriminating performance is due to differential performance for the arterial territory sub groups. For anterior TIA discrimination (figure 12.11), the area under the curve was 0.65 (S.E. 0.03) and for posterior TIA discrimination (figure 12.12), the area under the curve was lower than chance at 0.48 (S.E. 0.04).

Figure 12.10 ROC curve for discrimination of TIA from non-TIA using GP ABCD2 scores

Figure 12.11 ROC curve for discrimination of anterior TIA from non-TIA using GP ABCD2 scores
Figure 12.12 ROC curve for discrimination of posterior TIA from non-TIA using GP ABCD2 scores

12.7 Summary Measures of test performance at increasing GP ABCD2 cut points.

Figure 12.13 shows the change in sensitivity and specificity for the diagnosis of TIA using GP ABCD2 scores at increasing cut points. A similar pattern to the specialist scores is demonstrated. Specificity remains low throughout an acceptable range of sensitivity and only increases above 50% when sensitivity has dipped below 60%.

Figure 12.13 Change in sensitivity and specificity with increasing GP ABCD2 cut points for diagnosis.
Figure 12.14 demonstrates the predictive values of the GP ABCD2 score for TIA diagnosis. The negative predictive value remains at 100% until scores of 2 or more are used as cut points. The positive predictive value at this cut point remains low indicating significant numbers of false positives.

Figure 12.14 Positive predictive value (PPV) and negative predictive value (NPV) for TIA diagnosis at increasing cut points of specialist ABCD2 score.

In contrast to the PPVs from the specialist ABCD2 scores, the PPV using GP ABCD2 scores remains below 60% throughout the range of cut points.
Chapter 13 Validation of a TIA Recognition Tool

13.1 Validation using hospital records

All available patients with TIA in OXVASC were used as the hospital validation dataset as well as non-TIA clinic referrals with suspected TIA. 41 patients (5.4% of total patients in this analysis), who were referred but recorded as not having TIA, did not have secondary care histories of the event. Of these 41, eight patients had a primary care record and were used in the primary care validation of the Dawson tool.

13.2 Distribution of scores by diagnostic group – hospital dataset

Scores will range from a minimum of \((\text{age} \times 0.04)\) to a maximum of \((8.5 + (\text{age} \times 0.04))\). In the hospital dataset, the scores for TIA and non-TIA patients are summarised in Table 13.1.

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-TIA</td>
<td>330</td>
<td>6.0</td>
<td>1.12</td>
<td>2.42</td>
<td>9.40</td>
</tr>
<tr>
<td>All TIA</td>
<td>394</td>
<td>7.63</td>
<td>1.37</td>
<td>4.36</td>
<td>10.94</td>
</tr>
<tr>
<td>Anterior TIA</td>
<td>302</td>
<td>7.80</td>
<td>1.31</td>
<td>4.36</td>
<td>10.94</td>
</tr>
<tr>
<td>Posterior TIA</td>
<td>92</td>
<td>6.95</td>
<td>1.34</td>
<td>4.46</td>
<td>10.88</td>
</tr>
</tbody>
</table>

**Table 13.1 Summary statistics for TIA tool scores by diagnosis**

Although the overall mean for TIA patients is higher than for non-TIA patients, anterior TIA patients have a higher mean score than non-TIA patients. Comparing non-TIA, anterior TIA and posterior TIA for TIA tool scores with analysis of variance shows a significant between diagnostic group difference \((F=170, P<0.001)\). Post hoc Bonferroni tests show that all groups are significantly different in tool scores compared with each other at \(<0.001\) significance level.

Figure 13.1 demonstrates the distribution of scores for non-TIA patients, anterior TIA patients and posterior TIA patients. This displays visually that using the prevalence calibration data will not yield useful cut points, as only a minority of patients without TIA have scores below 4 and only a minority of TIA patients have scores above 10.
Figure 13.1 Distribution of TIA recognition scores between non-TIA, anterior TIA and posterior TIA

Figure 13.2 demonstrates the calibration for TIA versus non–TIA, by binning scores and examining the prevalence of TIA within each bin. Prevalence of TIA increases with successive score unit increases. No patients with a score below 4 have TIA and above 10, all patients have TIA.

Figure 13.2 Calibration of TIA recognition score - prevalence of TIA within each successive score bin
13.3 Tests of discrimination – hospital dataset

Figure 13.3 shows the plot of the ROC curve for discrimination between TIA and non-TIA using hospital records of the clinical history. The area under the curve (AUC) is 0.82 (standard error 0.02, 95% confidence interval 0.79 to 0.85). However this overall performance is over two quite distinct clinical populations of anterior and posterior TIA.

![ROC curve for discrimination between TIA and non-TIA](image)

**Figure 13.3 ROC curve for discrimination between TIA and non-TIA**

Figures 13.4 and 13.5 demonstrate ROC curves for the ability of the tool to discriminate between anterior TIA and non-TIA, and posterior TIA and non-TIA respectively. Clearly there is inferior performance for posterior TIA. The AUC for anterior TIA discrimination is 0.85 (S.E. 0.02, 95% CI 0.82 to 0.88) and for posterior TIA is 0.70 (S.E. 0.03, 95% CI 0.64 to 0.76).
Figure 13.4 ROC curve for discrimination between anterior TIA and non-TIA

Figure 13.5 ROC curve for discrimination between posterior TIA and non-TIA

The distribution of the individual components of the score is assessed in table 13.2 with proportions of non-TIA and TIA patients with feature present. Relative risks (RR) >1 indicate that the feature is associated with a diagnosis of TIA and relative risks < 1 indicate that the absence of the feature is associated with a diagnosis of TIA. Analyses are univariate with p
values calculated using chi-square for RR values greater than 0, Fisher’s exact test where RR=0 and t-test for mean age.

<table>
<thead>
<tr>
<th></th>
<th>Non-TIA symptom/patients</th>
<th>TIA symptom/patients</th>
<th>RR</th>
<th>Lower Confidence Interval</th>
<th>Upper Confidence Interval</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>History of TIA/stroke</td>
<td>26/330</td>
<td>119/394</td>
<td>3.83</td>
<td>2.57</td>
<td>5.71</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Headache</td>
<td>52/330</td>
<td>40/394</td>
<td>0.64</td>
<td>0.44</td>
<td>0.95</td>
<td>0.03</td>
</tr>
<tr>
<td>Diplopia</td>
<td>10/330</td>
<td>18/394</td>
<td>1.51</td>
<td>0.71</td>
<td>3.22</td>
<td>0.38</td>
</tr>
<tr>
<td>Fit</td>
<td>9/330</td>
<td>0/394</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Loss Of Consciousness</td>
<td>59/330</td>
<td>0/394</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Speech</td>
<td>51/330</td>
<td>168/394</td>
<td>2.76</td>
<td>2.1</td>
<td>3.64</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Face</td>
<td>19/330</td>
<td>85/394</td>
<td>3.75</td>
<td>2.33</td>
<td>6.03</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Limb</td>
<td>50/330</td>
<td>146/394</td>
<td>2.44</td>
<td>1.84</td>
<td>3.26</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Age mean (SD)</td>
<td>64.72</td>
<td>(16.62)</td>
<td>73.64</td>
<td>(12.33)</td>
<td>-</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Table 13.2 Distribution of sub-components of tool recognition score: non-TIA vs TIA

A forwards conditional stepwise multivariate logistic regression for diagnosis of TIA (with a probability of 0.05 for entry) demonstrated that all the tool recognition variables were kept in the equation apart from the presence of a fit, with overall 72.9% correct classification from the model and a Nagelkerke $R^2$ of 0.40 (this is a measure of how much better the model is at predicting the outcome over assuming that each patient has the same chance of the outcome as the background prevalence in the group - 1 is perfect prediction and 0 is no better than the background).

Seizure activity has a low prevalence overall in this cohort of patients which may reflect a difference in referral practices between OXVASC GPs and GPs in Scotland. This in itself may reflect availability of other services where a rapid access to a physician opinion and cerebral imaging may be lacking, such as first fit clinics.

Thus the recognition tool validates well in this cohort of patients in terms of calibration and discrimination using secondary care histories, with the majority of variables contributing to predictive power. However, there is marked differential performance in discrimination with poorer performance in posterior TIA.
13.4 Cut points and effects on performance

The authors of the recognition tool suggest two cut points – the optimal point is 6.1, but in order to reflect the misclassification costs of missing TIA, a 2:1 cost ratio was applied, which reduced the optimal score to 5.4. Sensitivity, specificity, positive predictive value and negative predictive value at the two cut-points are shown in table 13.3.

<table>
<thead>
<tr>
<th>Cut point</th>
<th>Positive Predictive Value</th>
<th>Negative Predictive Value</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.4</td>
<td>61%</td>
<td>85%</td>
<td>96%</td>
<td>28%</td>
</tr>
<tr>
<td>6.1</td>
<td>69%</td>
<td>76%</td>
<td>86%</td>
<td>54%</td>
</tr>
</tbody>
</table>

Table 13.3 Performance of tool at different cut points

Reducing the cut-point for diagnosis increases the sensitivity but at the expense of increasing false positives, as evidenced by a reduction in positive predictive value and specificity. This is illustrated more clearly in figures 13.6 and 13.7 which plot the effect of changing the diagnostic cut point on sensitivity and specificity as well as positive and negative predictive values.

Figure 13.6 Plot of sensitivity and specificity with increasing diagnostic cut point
13.5 Validation using primary care records

The population of patients used are those seen first in primary care and where the GP has suspected TIA and then later assessed by an OXVASC research fellow with final diagnosis decided after discussion with the senior investigating neurologist (PMR). The records of 513 patients who were referred from primary care were used in the analysis.

13.6 Distribution of scores across diagnostic groups - primary care dataset

Table 13.4 shows summary statistics for the tool scores in primary care. Differences in scores between diagnostic groups were tested with analysis of variance which was statistically significant for a between group effect (F 53.27, P<0.001). Post hoc Bonferroni tests explored which groups accounted for this difference and anterior TIA scores were significantly different from non-TIA and posterior TIA (both at p<0.001). Posterior TIA scores were not significantly different from non-TIA scores.

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-TIA</td>
<td>304</td>
<td>6.33</td>
<td>1.10</td>
<td>3.54</td>
<td>9.04</td>
</tr>
<tr>
<td>All TIA</td>
<td>209</td>
<td>7.23</td>
<td>1.22</td>
<td>3.85</td>
<td>10.17</td>
</tr>
<tr>
<td>Anterior TIA</td>
<td>167</td>
<td>7.44</td>
<td>1.21</td>
<td>3.85</td>
<td>10.17</td>
</tr>
<tr>
<td>Posterior TIA</td>
<td>42</td>
<td>6.41</td>
<td>0.90</td>
<td>4.80</td>
<td>8.34</td>
</tr>
</tbody>
</table>

Table 13.4 Summary statistics for TIA tool scores by diagnosis
The distribution of scores is demonstrated between the diagnostic groups in figure 13.8. This displays more clearly the lack of difference between the posterior TIA and non-TIA score distribution, compared with the anterior TIA distribution.

Figure 13.8 Distribution of scores from GP records by diagnostic group

Figure 13.9 demonstrates the calibration of the tool with predictive value of each score bin. The inter-bin increase is not linear e.g. bins 3 to 4 and 4 to 5 do not show an increase in prevalence of TIA whereas there is a large increase from bin 8 to 9 to bin 9 to 10. All patients with a score of 10 or more have TIA whereas in the lowest bin with scores (bin 3 to 4) there is a small prevalence of TIA.
Figure 13.9 Calibration of % prevalence of TIA within successive GP score bins

13.7 Discrimination between diagnostic groups using primary care records

The ROC plot for discriminating TIA from non-TIA is shown in figure 13.10, with an AUC of 0.70 (S.E. 0.02, 95% CI 0.66 to 0.75). However, this overall performance is from the sum of two distinct clinical populations, anterior and posterior TIA patients. The AUC for discrimination of anterior TIA from non-TIA was 0.75 (S.E. 0.02, 95% CI 0.70 to 0.80). The AUC for posterior TIA was 0.52 (S.E. 0.04, 95% CI 0.43 to 0.60), demonstrating that discrimination from non-TIA was no better than chance. Figures 13.11 and 13.12 display the ROC plots for anterior and posterior discrimination respectively.

Figure 13.10 ROC curve for diagnosis of TIA vs non-TIA using the recognition score with primary care records.
Figure 13.11 ROC curve for diagnosis of anterior TIA vs non-TIA using the recognition score with primary care records.

Figure 13.12 ROC curve for diagnosis of posterior TIA vs non-TIA using the recognition score with primary care records.

In order to investigate which parts of the tool are likely to account for its performance, the distribution of the tool components in patients with and without TIA is shown in table 13.5. No GPs recorded fits or seizures, although nine patients had their histories re-interpreted in secondary care to be related to seizure activity. Therefore fit presence/absence is not shown in table 13.5.
<table>
<thead>
<tr>
<th></th>
<th>Non-TIA Symptom / patients</th>
<th>TIA Symptom / patients</th>
<th>RR</th>
<th>Lower confidence Interval</th>
<th>Upper confidence Interval</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>History of TIA/stroke</strong></td>
<td>22/304</td>
<td>53/209</td>
<td>3.50</td>
<td>2.20</td>
<td>5.58</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td><strong>Headache</strong></td>
<td>32/304</td>
<td>19/209</td>
<td>0.86</td>
<td>0.50</td>
<td>1.48</td>
<td>0.70</td>
</tr>
<tr>
<td><strong>Diplopia</strong></td>
<td>8/304</td>
<td>7/209</td>
<td>1.27</td>
<td>0.47</td>
<td>3.46</td>
<td>0.84</td>
</tr>
<tr>
<td><strong>Reduced conscious level</strong></td>
<td>24/304</td>
<td>1/209</td>
<td>0.06</td>
<td>0.008</td>
<td>0.45</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td><strong>Speech</strong></td>
<td>61/304</td>
<td>81/209</td>
<td>1.93</td>
<td>1.45</td>
<td>2.56</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td><strong>Face</strong></td>
<td>19/304</td>
<td>19/209</td>
<td>1.45</td>
<td>0.80</td>
<td>2.68</td>
<td>0.30</td>
</tr>
<tr>
<td><strong>Limb</strong></td>
<td>70/304</td>
<td>57/209</td>
<td>1.18</td>
<td>0.88</td>
<td>1.60</td>
<td>0.30</td>
</tr>
<tr>
<td><strong>Age mean (SD)</strong></td>
<td>65.20 (16.57)</td>
<td>73.22 (12.76)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Table 13.5 Comparison of sub-components of TIA recognition score between TIA and non-TIA.

Again, p values are calculated using chi-squared for dichotomised variables of presence/absence of symptoms and using an independent samples t test for comparing mean age. Compared with secondary care scores, facial weakness, limb weakness, diplopia and headache are no longer significantly differently distributed between patients with and without TIA. This may reflect the fact that this is a sub-population of patients i.e. table 13.5 represents all those patients seeking care in general practice before secondary care, whereas table 13.3 is for all referrals to clinic including those patients who first presented in ED. However, an alternative explanation is that the histories are differently taken and recorded on the same patients.

To further investigate the sub-components of the score, a forwards stepwise multivariate logistic regression, with entry probability of 0.05 to predict TIA diagnosis, using primary care recorded variables from the recognition score was carried out. At step 4, no further iterations improved model fit and the excluded variables were headache, diplopia, facial weakness and limb weakness, with a Nagelkerke R² of 0.22 and 68% correct classification.

13.8 Cut points and effects on performance – primary care records

Table 13.6 shows positive predictive value, negative predictive value, sensitivity and specificity of the two diagnostic cut points recommended by Dawson – 6.1 and 5.4.
Even at the higher cut point of 6.1, positive predictive value does not improve on published clinic prevalence figures of around 50%. At the recommended cut point of 5.4, to reflect the need for greater sensitivity in TIA detection, PPV remains at around the overall level of 40% in the cohort, due to a higher false positive rate (82%). Figures 13.13 and 13.14 illustrate more clearly the effect of different diagnostic cut points on sensitivity, specificity, PPV and NPV.
Specificity remains low at high levels of sensitivity, but there is a sharp decline in sensitivity as the false positive rate begins to reduce.

13.9 Impact on referral and safety

If decisions to refer within the population of patients attending primary care were based on this decision rule then its impact can be assessed by using follow up of false negative patients. If high risk TIA patients are missed then the ascertainment of recurrent events in OXVASC will detect this. In other words, making the assumption that patients with scores below the cut point would not have been referred, the impact on the tool in terms of patient safety can be judged by following up those patients to see if any had a recurrent stroke. This is likely to underestimate the risk of implementing the tool, as referred patients with true TIA have been investigated and treated with a combination of vascular risk reducing interventions. Thus only counting the missed patients with recurrent events is likely to underestimate the risk of using the tool.

At both cut points of 6.1 and 5.4, two patients were missed who had a recurrent ischaemic stroke within 7 days of the TIA that led them to present to their GP. No additional events were ascertained to 90 days after TIA in the ‘false negative’ cohort as defined by either cut point.

Using the 6.1 cut point would have saved 168 referrals, of whom 42 were diagnosed with TIA and did not have a recurrent stroke by 90 days of follow up. Using the 5.4 cut point would have saved 71 referrals of whom 17 were TIA and didn't have a recurrent stroke on 90 days of follow up.
The presenting ABCD2 scores are available for these patients who were falsely negative at the two cut points and the expected number of strokes can be calculated without treatment using the known associated early stroke risk.

In the 6.1 cut point cohort of 42 false negative recurrence free patients, multiplying the frequency of each patient’s ABCD2 score by its 90 day estimated risk gives a total expectation of 2.08 strokes, for a 91,000 population over 4 years, alongside the additional two cases of stroke.

In the 5.4 cut point cohort of false negative recurrence free patients, the same calculation of mathematical expectation of recurrent strokes is 0.71 for a 91,000 population over 4 years, alongside the additional two cases of stroke.

13.10 Difference in populations or difference in histories?

The diagnostic tool score performs quite differently in secondary care records and in primary care records. Not only are the calibration and discrimination measures different, most notably for the discrimination of posterior TIA patients in primary care recorded clinical features, but the individual components of the score have different levels of statistical significance in their distribution between TIA and non-TIA patients, when one compares across primary and secondary care. Fewer items are distributed differently with an appropriate degree of significance in the primary care records.

These differences could be due to a sub-population effect, whereby the total clinic population is made up of two quite distinct populations of primary care and ED referred patients which in themselves are very different in their distribution of predictors of a TIA diagnosis. Alternatively, there could be marked disagreement between GPs and the TIA specialists in deciding on the presence or absence of clinical features. This latter explanation was seen in chapter 6 where ABCD2 scores were compared between primary and secondary care.

Figure 13.14 demonstrates an Altman-Bland plot of the difference between specialist and GP scores against the average of the specialist and GP scores in patients referred from primary care to OXVASC. For all referred patients the mean difference is -0.09 with a standard deviation of 0.92. Further plots by diagnostic group (figures 13.16 – 13.18) demonstrate a mean difference for non-TIA patients of -0.3 (i.e. GPs score higher than specialists), a mean difference of +0.3 for anterior TIA (specialists score higher than GPs) and a mean difference of 0.02 for posterior TIA.
Figure 13.14 Specialist – GP score against average specialist and GP score for all referred patients from primary care; mean difference (thick line), 2 standard deviations (broken lines).

Figure 13.15 Specialist – GP score against average specialist and GP score for all non-TIA patients; mean difference (thick line), 2 standard deviations (broken lines)
Figure 13.16 Specialist – GP score against average specialist and GP score for anterior TIA patients; mean difference (thick line), 2 standard deviations (broken lines)

Figure 13.17 Specialist – GP score against average specialist and GP score for posterior TIA patients; mean difference (thick line), 2 standard deviations (broken lines)
From mean values, GPs tend to score patients more highly than specialists for non-TIA and lower for anterior TIA. In the overall referred group with suspected TIA, there is no particular bias as the anterior and non-TIA patients effectively cancel each other out. This suggests that the difference in performance metrics for the Dawson tool between primary and secondary care is due to differences in histories and therefore predictor values.

The foregoing analysis rests on the assumption that if a GP hasn’t recorded a symptom then the patient did not have that symptom present. However, patients could have complained of symptoms or GPs elicited symptoms that were not recorded in consultation records or referral letters.

As the GPs were not knowingly taking a history with a teleological function to populate a diagnostic score, there is no risk of score optimisation bias, which is where certain aspects of the history may be amplified with a view to creating a higher score to justify referral.
Chapter 14  Summary and Conclusions

14.1  GP-specialist disagreement and accuracy of triage

14.1.1 Disagreement between GPs’ and specialists’ records of the same patients with suspected TIA

There was substantial disagreement over the presence or absence of key clinical features that determine risk of recurrent stroke from the ABCD2 score – speech disturbance, unilateral weakness and duration of symptoms (24). There was greater agreement over speech disturbance, less so for unilateral limb weakness and less again for facial weakness, with lowest agreement for level of duration of symptoms. This has implications for the appropriate usage of TIA clinic slots as incorrect risk assessment of high risk patients as low risk increases their time to clinic appointment from 24 hours to seven days (60). Disagreement may also reduce the availability of the usually smaller number of urgent clinic slots by incorrectly labelling low risk patients as high risk making it harder for clinics to meet the high risk patient target of 24 hours.

14.1.2 GP-specialist disagreement and final clinic diagnosis

Patients with TIA had a lower mean ABCD2 score from GPs than specialists, but referred patients without TIA had a higher mean score from GPs than specialists. Within the category of patients with TIA, anterior TIA patients had a lower mean score from GPs than specialists and posterior TIA patients had a higher mean score from GPs. The key symptom disagreements underpinning these findings were over facial and limb weakness and speech disturbance with the lowest agreement over duration of symptoms mainly because it is not a dichotomous measure, thereby allowing three values for disagreement rather than two. Differences in the prevalence of symptoms in different diagnostic groups, as assessed either by GPs or specialists makes the kappa coefficient of agreement less straightforward to interpret and identify if there is systematic over or under attribution of symptoms by GPs.

14.1.3 Impact of GP-specialist disagreement on the performance of the ABCD2 score for accurate triage of patients with TIA

From clinical records of patient histories, GPs would underscore patients at higher predicted risk and overscore patients at lower predicted risk. Furthermore in anterior TIA patients GPs would underscore and in posterior TIA patients they would overscore when compared with specialist scores, which act in this case as a gold standard of risk prediction, given the pre-existing validation of the ABCD2 score in specialist derived scores. This results in a good but not perfect sensitivity to detect high risk cases (as defined by specialists) with the risk that if
guidelines for time to assessment are adhered to, a small number of recurrent strokes may occur in a relatively small group of high risk patients who are thought to be lower risk, before they are investigated and have optimal medical management in place.

This analysis suggests that the ABCD2 score may not work as well as a predictive tool in primary care and therefore may not improve the delay that high risk TIA patients currently experience in specialist assessment. A small number of strokes may occur annually as a result of incorrectly triaging high risk TIA as low risk thereby incurring a longer delay to clinic appointment.

14.1.4 Limitations of this analysis

Although this method of retrospective scoring of the ABCD2 score has been used previously with specialist records (25;224) there are no published reports of scoring from the primary care record at the initial consultation in general practice. However, retrospective scoring reduces the chance of score optimisation bias where GPs will know that if a patient has a score of 4 or more then they will be seen urgently and this is a more risk averse strategy than the patient waiting for a week before a definitive assessment.

14.1.5 Future Research and Practice

Even though derivation and validation of the score for early recurrent stroke have been carried out on hospital records (23;24;26-29) it has been widely recommended for use in primary care (35;36;60) where its validation metrics have not been assessed. The phenomenon of derivation and validation in secondary care with implementation recommendations to primary care has been noted as a cause of poor performance in prognostic models (273).

To answer the question of the impact of this disagreement would require prospectively studying the agreement among consecutively referred patients to a TIA clinic where GPs have to use the ABCD2 score as part of the referral process and TIA specialists complete the score as well. The level of agreement could be examined in relation to diagnosis and time since referral from primary care in case disagreement is higher for longer elapsed time periods from the clinical event. Furthermore the key outcome of appropriate urgent clinic slot usage could be determined, as well as disagreement over the individual score, although this is less meaningful than disagreement over dichotomised risk status. The only similar study reported is small, did not examine referring clinician specialty (primary care, ED or general physician), did not examine risk status or appropriate urgent slot usage and would not have been powered to detect influences on the degree of disagreement (227).
14.2 The impact of GP-specialist disagreement on the discrimination metrics of existing clinical prediction tools for TIA diagnosis

14.2.1 ABCD2 score

For both specialist and GP scores, the sensitivity and specificity plots with increasing ABCD2 cut offs demonstrate that there are no acceptable values with sufficiently high sensitivity and specificity. Furthermore, the predictive value plots for the GP scores display low predictive values for increasing cut off points for diagnosis.

For both specialist and GP ABCD2 scores, the mean values in TIA patients were higher than non-TIA patients and anterior TIA scores were higher than posterior TIA scores. For both GP and specialist the distribution of anterior and posterior TIA scores had opposite skews with peak numbers of patients scoring at the upper end for anterior TIA and at the lower end for posterior TIA. Specialist scores appear to calibrate better than GP scores with higher specialist scores associated with an increased chance of TIA diagnosis.

The potential for ABCD2 as a diagnostic rule has been suggested by several authors with similar validation methods for discrimination in a TIA clinic population using ROC curves with AUC values of 0.68 for TIA, 0.70 for combined TIA and minor stroke (177) and 0.75 for combined TIA and minor stroke (224). The specialist record derived ABCD2 score in OXVASC on ROC analysis has an AUC of 0.71 for TIA which is in keeping with these other reports but again there is a marked territory variation in performance with an anterior TIA discrimination AUC of 0.75 and a posterior TIA value of 0.59. The reports cited above did not explore the influence of arterial territory on diagnostic performance.

However given that for the majority of UK patients GPs will be initially suspecting the diagnosis, one should know the diagnostic performance of the ABCD2 score in their hands. The GP derived scores again showed lower AUCs for TIA vs non-TIA of 0.62 and for posterior TIA vs non-TIA of 0.48 i.e. no better than chance discrimination.

Inspite of suggestions that the ABCD2 score could be used for diagnosis, its performance as a diagnostic tool even in specialist hands is limited. The performance assessed by ROC plots demonstrates low AUC values, and no additional value above chance for posterior TIA detection. For GP ABCD2 scores, there is no value that offers appropriate sensitivity and specificity and the positive predictive values across the range of scores do not rise much above general prevalence in the suspected population, indicating that it is not useful for refining referral decisions.
14.2.2 Dawson tool

The AUC for the Dawson score was lower in the GP dataset than the specialist dataset, with the greatest difference for posterior circulation TIA. Dawson et al. did not report an AUC for their model, only the sensitivity/specificity/PPV/NPV for two cut points – an optimal ROC derived score and a further cut point reflecting 2:1 costs of misclassification (222). However, they did suspect that their tool would have different performance metrics in specialist and GP histories, and this is indeed the case from the data presented in this thesis.

Overall prevalence of TIA in the GP referred patients was similar to the complete group of OXVASC clinic referrals (41%) but diagnostic performance in GP derived histories was substantially poorer. GP derived scores had lower AUCs than specialist scores (0.70 (S.E. 0.02) vs 0.82 (S.E. 0.02)) for TIA diagnosis and were at chance on posterior TIA patients (0.52 (S.E. 0.04) vs 0.70 (S.E. 0.03) specialist scores)

At the 5.4 cut point suggested for the recognition tool, the sensitivity (96%) and specificity (28%) of the specialist derived score are comparable to those reported by Dawson et al in their internal validation (97% and 24%). Interestingly the predictive value of specialist score in the OXVASC data although lower than in the Dawson validation sample is in fact a substantial improvement on the OXVASC baseline prevalence. The prevalence of TIA in the Dawson validation sample was 60% which increased at the tool cut point to 68%, whereas using the specialist derived score in OXVASC increased predictive value from 46% baseline prevalence to 61% at the 5.4 cut point.

Although Dawson et al used an outcome of cerebrovascular disease (i.e. TIA and minor stroke), the key diagnostic issue in primary care is for patients with resolved or rapidly resolving deficits. There is no management dilemma for patients with persisting stable deficits at the time of consultation, as there is a clear clinical problem that requires a solution. Transient phenomena are more difficult to assess as there are no robust physical signs and the history of events along with the contextual information of vascular risk factors are the only data available. Thus the outcome of interest is TIA diagnosis and so all patients with transient phenomena i.e. TIA patients and those with transient phenomena in the non-TIA population were used as the validation dataset. The implications of a difference in clinical histories between GP and specialist for the same patient is that a diagnostic rule using variables from the clinical history is likely to have different validation metrics in GP histories and specialist histories.

The analysis of the Dawson tools’ significant predictors in the hospital records dataset shows that diplopia is not distributed differently between patients with and without TIA. Whilst
headache is still a significant negative predictor in OXVASC, the associated p value of 0.03 is much higher than that associated with the other predictors, and higher than the original derivation model in the Glasgow clinic (0.00007 (222) suggesting that it does not have such a difference in prevalence between TIA and non-TIA in OXVASC. This may represent an element of subjectivity of TIA diagnosis, given that all patients with TIA in OXVASC had one senior neurologist ultimately determine diagnosis (PMR), and disagreement over TIA diagnosis between neurologists has been noted previously (183;270).

14.2.3 Limitations of this analysis

The limitations of ABCD2 scoring from primary care records in discussed in 14.1.4. Differences in the reported performance of the Dawson score and performance in OXVASC may represent the difference between routinely collected clinical records and research records. This difference has been noted before particularly for stroke onset times (74). Predictors in derivation and validation should be collected systematically (274) and by using the content of clinical records there is no guarantee of systematic history taking. However, the key associated features of focal deficit such as headache are more likely to be taken systematically by a TIA specialist as part of the exclusion of mimic conditions.

The analysis of individual Dawson tool predictors collected from primary care records showed that a number were no longer significant – headache, diplopia, and unilateral weakness of face or limb. GPs clinical records were taken as part of routine care rather than for research purposes and so the extent of symptom questioning may well be less rigorous than a specialist might undertake, contravening the principle of systematic predictor measurement (274). However, face and limb weakness do not appear to be differently distributed between TIA and non-TIA patients and therefore they drop out as being predictors of diagnosis. This is not because they are not elicited, it is because they are equally prevalent in both groups of patients.

There are no other similar studies to compare the effects of territory on diagnostic tool performance after presentation with TIA. However, the performance of the FAST test of recognition of stroke was compared between anterior and posterior stroke patients with equal sized groups demonstrating differences in sensitivity - 61% of posterior strokes compared with 92% of anterior strokes were detected (275).

14.2.4 Future research and practice

For GP derived scores at the optimal cut point of 5.4, there is not a large difference in predictive values from baseline, with an increase in TIA prevalence from 41% to 44% and
the low specificity of 18% explains why prevalence has only slightly increased – most non-
TIA patients would still be referred at that cut point.

GPs are most likely referring because there is some feature in the clinical history that makes
them suspect the diagnosis i.e. the patient presentation resembles the diagnostic pattern. It
follow that more classical diagnostic features may be present in both TIA and non-TIA
patients as judged by the doctor who is initially suspecting TIA. Thus patient narratives that
are interpreted as weakness will generate the suspicion of TIA and hence membership of the
group of referred patients.

On the other hand, the diagnostician specialist will be seeking to decide who has and who
has not had a TIA from clinical history. They will therefore be reinterpreting the report of
events and those that are not TIA will not have the diagnostic features. This would predict
that specialists would have lower scores than GPs for non-TIA patients as they would be
finding fewer of the diagnostic predictors in these patients, which is the case with a mean
difference of - 0.3 between the specialist score and the GP score. On the other hand, in the
case of patients with TIA, GPs are under finding diagnostic features which the specialists are
recording, with a mean difference of + 0.3 between specialist and GP for patients with
anterior TIA.

Thus a new prediction tool is required that should be derived from primary care taken
histories, as those from specialist records are likely to validate poorly with little gain in terms
of refinement of the referral population.
SECTION 5 Prediction Rules for TIA in Primary Care
Chapter 15  Methods

15.1  Introduction

Primary care records have not been previously explored for their potential to be used in deriving predictors for final clinic diagnosis of TIA. Section 4 demonstrated that existing decision tools derived from secondary care records are unlikely to be useful for primary care to support the diagnosis of TIA. The research questions that address this issue are

1. Which clinical predictors are included in a prediction rule for TIA diagnosis derived from primary care records in suspected TIA?

2. What are the calibration and discrimination metrics for a model of TIA diagnosis derived from primary care records in suspected TIA?

To answer these questions, the primary care records for all referrals to the TIA clinic in the 2002 – 2006 OXVASC cohort were analysed for all symptom content and grouped into categories of neurological dysfunction. These categories were treated as potential predictors and a logistic regression model derived to predict TIA diagnosis. Predicted probabilities from the model were used to construct a calibration curve and scores based on the logistic regression model (both with and without weighting from beta coefficients) were used to assess discrimination from ROC curves.

We do not know what the true underlying relationship is between predictors and the outcome of TIA diagnosis. The different methods of modelling this relationship are just that – different models, without any particular claim on the truth. As such we should admit that there is ‘model uncertainty’ in that there is no clear decision about which model should be used. Although logistic regression is commonly used to derive prediction rules for diagnosis there are alternative methods of statistical modelling which offer potential advantages - classification trees and the random forest method of bootstrapped multiple classification trees. The research question that addresses this is

3. Does choice of statistical model affect discrimination metrics for TIA diagnostic models?

This question was answered by creating a classification tree to discriminate between TIA and non-TIA using the same clinical predictors as the logistic regression model, and by creating a random forest again using the same clinical predictors as the logistic regression model. In order to compare their metrics using the same methodology, the original full cohort dataset was randomly split into training and testing datasets. Then logistic regression,
classification tree and random forest models were derived from the training data and discrimination measured on the testing data using ROC curves.

15.2 Coding GP clinical records

The clinical presentations from histories taken by GPs were coded according to broad symptom categories. These categories were formed after grouping all relevant descriptions of suspected TIAs and forming groups consisting of the smallest number of patients. This maximises the number of symptoms to select whilst allowing pragmatic grouping for smaller units within related themes. For example a specific observation that GPs made was that some patients reported normal speech but transiently, a particular word which they knew eluded them and thus isolated ‘word finding difficulty’ can be identified. However, this could also be grouped within a higher category of ‘all speech disturbance’ and is distinct from nominal aphasia which is a much more severe symptom affecting classes of words e.g. nouns rather than a specific word in the context of otherwise normal speech.

Cases were excluded if they were seen in accident and emergency first before referral (in the UK or abroad) and told that they had a TIA or minor stroke and were asked to see their GP for an onward referral. The history that may be recorded in this situation may not be the one that a primary care practitioner would have taken if seeing the patient first. These patients are attending primary care because the healthcare system demands it having attended another part of the service, rather than reflecting a ‘primary care’ population.

Patients were also excluded if they presented with a non-TIA after recruitment into OXVASC with a TIA or minor stroke at initial ascertainment. This is because the decision making process may be altered by the fact that the GP (and patient) know that they are in a research study with regular follow up where potential recurrent events are assessed. Thus it is not similar to a fresh referral where a GP is trying to decide if a patient’s narrative could be a TIA and therefore a decision rule that is derived using such cases may not validate well outside of the OXVASC research process.

Patients with symptoms persisting for longer than 24 hours were excluded because they pose less of a decision-making problem - a neurological deficit that persists requires investigation in order to ascertain the underlying cause. The key decision that this analysis aims to support is the decision to refer based on history alone with no neurological deficit on clinical examination. Given that only TIA patients ascertained in OXVASC are used in the derivation of this diagnostic decision rule, the non-TIA patients should be those that most closely mimic TIAs, i.e. they have had transient symptoms without persisting deficit.
15.3 Methodologies for diagnostic model derivation – logistic regression

Logistic regression is one of the commonest used techniques for selecting optimal predictors for a dichotomous outcome in clinical prediction research. After selecting plausible predictors either from previous research or clinical consensus, each is tested in turn (univariate analysis) to see if individually they are predictive of a cerebrovascular diagnosis. Those passing a minimum level of statistical significance (probability of the data given that the null hypothesis is true < 0.05) are included in a forward stepwise logistic regression. This method of analysing relationships seeks to select the fewest number of variables that predict a binary outcome (presence or absence of cerebrovascular diagnosis).

Using the logistic transformation, where the probability that cerebrovascular disease is present is transformed to

\[ \log (\text{odds (probability that cerebrovascular disease present)}) \]

a linear predictor of that transformation is estimated. Thus for each variable that has predictive power after the inclusion of other variables (i.e. is significantly associated with cerebrovascular disease presence after adjustment from other included variables), a coefficient is derived. Methods of parameter derivation for logistic regression use maximum likelihood techniques which use iterative methods to refine parameters and estimate them based on convergence to an asymptotic value. The likelihood that is maximised is the probability of finding the observed data assuming that the model is true (276).

In order to calculate the probability that cerebrovascular disease is the cause of symptoms, a reverse transformation of the linear variables is needed, weighted by their associated parameters e.g.

\[ p (\text{cerebrovascular disease present}) = \frac{1}{1 + e^{-lp}} \]

where \( lp = \beta_1x_1 + \beta_2x_2 + \cdots + \beta_ix_i + \text{error} \), and \( x_i \) is the \( i \)th predictor.

Logistic regression in essence will derive the strength of prediction that is due to an individual clinical predictor, given the presence of other significant predictors in the dataset as a whole.

However, there may be situations where a variable such as presence of a certain symptom may only have predictive power in a restricted range of patients, defined by co-existence of other features. By itself that variable may not be significantly associated with the outcome across the dataset as whole. As such it is unlikely to be identified by logistic regression as a
predictor of the outcome, even though it may have utility in terms of discriminating the presence of disease in a defined set of patients.

15.4 Diagnostic rule derivation with regression using GP histories

Using the categories from analysing GP records of clinical events, all were considered as potential predictors alongside age and previous history of TIA/Stroke and were tested to examine univariate associations with diagnosis of TIA and those with significant predictive ability were selected to go into a forward stepwise logistic regression. The model outputs were then used to derive predicted probabilities of TIA and the distributions of predicted probabilities were examined for internal calibration and discrimination.

15.5 Methodologies for diagnostic model derivation – rationale for classification trees

Multivariable logistic regression models use only those variables that still significantly predict the presence of TIA after adjusting for other predictors found on univariate analysis. Thus the model assesses the strength of a predictor allowing for the presence of other variables that also have some predictive value.

However, it may well be the case that a variable has strong predictive power only in a group of patients defined by certain characteristics. In this case that predictor may not be included in a logistic regression model as its statistical effect will be diminished by the presence of other predictors, unless a specific interaction term is inserted into the logistic regression model to explicitly account for the fact that we are interested in the predictive power of this variable given the presence of other features.

With a large number of potential predictors, specifying all the interactions a priori will generate a large and complex logistic regression model, as well as a decision rule that is even more complex as a given predictor may have a different weighted multiplier in a score depending on the co-occurrence of other predictors. Furthermore, even without the complexity of specifying interaction terms, the diagnostic scores derived from logistic regression models are not straightforward to use, particularly if weighted models are used.

Rather than calculating a score which is time consuming and error-prone, an alternative strategy is to derive an algorithm where the response to certain key questions determines the decision to refer for a suspected TIA.

A classification tree model consists of a number of questions about the predictor variables which are asked in a fixed order. This is said to mimic the clinical decision making process
with most important predictive variables asked about first (276). A tree is constructed by producing partitions or splits in a derivation dataset depending on the value of a predictor – either a cut point for continuous predictors such as age or blood pressure, or the presence or absence of a binary categorical predictor such as weakness or confusion.

The splits in the data produce smaller sub groups, and are chosen such that there is maximum difference between the subgroups i.e. an age cut off is chosen which produces two subgroups with the greatest difference in % TIA within the subgroups. The predictor which results in the largest separation in % outcome in the two subgroups is placed at the top of the tree as the most important predictor.

As the tree is ‘grown’ with increasing numbers of predictors used, splitting of the remaining data into further subgroups continues until there are too few cases to produce subgroups (this can be specified in the modelling) or until there is no more improvement in outcome prevalence in subgroups compared with the parent group used to produce the split.

Two advantages of tree models over logistic regression are simplicity of presentation (and usage) and the incorporation of multiple interaction effects, and given the number of branches of trees, these interactions can be of high order (i.e. the effect of variable a, given a value of variable b and c etc. along the branches of a tree). In logistic regression, such interactions would need to be explicitly included and are rarely higher than second order i.e. a term for the interaction of two predictors. The process of tree construction assumes that all predictors interact with each other.

However, although interactions are presumed in classification trees, the higher order interaction is only modelled in a specific branch of a tree and not modelled along other branches. This is because predictors are used to split groups where they produce the greatest subgroup differences, and this may be several nodes away from an initial large split in the dataset. As such, there may be patients channelled along distant parallel branches of a tree that could contribute to an interaction term but are not included in the generation of a subgroup split using that interaction, as they were split off at an earlier stage due to the presence or absence of another predictor. An interaction term in a logistic regression model would be applied across the dataset rather than in a restricted set of patients due to prior branching.

A further disadvantage of classification trees are the categorisation of continuous variables needed to produce a split as this reduces information and would not occur in regression modelling. However, in the dataset under consideration in this thesis only one clinical
predictor is not categorical (age), with the others being the dichotomised presence of predictors from clinical histories and thus this is not a particular disadvantage.

Irrespective of theoretical advantages and disadvantages, a classification tree analysis of the GP dataset is warranted to explore the potential for a clearer set of decisions that could support referral from primary care to secondary care for suspected TIA, without needing to perform potentially complicated calculations in the real time setting of the consultation. If a classification tree has similar discriminating performance to a complicated score then it may have greater utility as it will be easier to follow and incorporate into history taking and clinical assessment.

15.6 ‘Pruning’ a classification tree

The ‘rpart’ package in R software (www.r-project.org) was used to generate a classification tree with TIA diagnosis as the outcome to be predicted, and all GP clinical variables were included alongside age and prior history of cerebrovascular disease as potential predictors. This is similar to the initial analysis for the logistic regression method where all potential predictors are considered in the univariate analyses for statistical significance.

Trees can be ‘pruned’ such that a smaller number of branches are used but ensuring that they still effectively partition the dataset into patients with and without disease. The pruning process is determined by the effectiveness of splits in the data for the distal branches in a tree (277). In the basic tree models, the trees grow until there is no more data or the split doesn’t result in different prevalences of the outcome in the subgroups compared with the parent group at that split. Some splits could produce fairly trivial differences in subgroup prevalences and add complexity without adding much to the discriminating ability of the tree in the overall dataset.

A ‘complexity parameter’ can be defined which will only allow a splitting if there is a big enough improvement in model fit and if not, then the tree will stop growing (rather than being ‘pruned back’, the tree is not allowed to grow, so the gardening metaphor behind the terminology does not reflect the mathematical processes).

The complexity parameter takes into account the number of patients at the ‘node’ or decision point to be split, the number of nodes in the tree as a whole and the change in the predictive ability of the whole tree (the model fit) as a result of the split. The complexity parameter informs the tree growing algorithm that a further split is allowed if it reduces the overall lack of fit (the residual mean square) by a certain factor (the value of the complexity parameter).
Identifying the best value for the complexity parameter comes from a cross-validation exercise. The dataset is randomly split into a number of groups and classification trees are ‘grown’ using the same initial variables but only using the data from outside a given group and tested for how accurate its predictions are on data in the given group, by deriving a misclassification rate. This is done over a range of different complexity parameters, which will therefore result in a range of tree sizes for each attempt to predict the diagnosis in a given group. This whole process is repeated for each of the randomly constructed groups, one at a time, giving a misclassification rate each time a complexity parameter is tested in each of the groups. This results in an estimate across the groups of the misclassification error for each potential value of the complexity parameter.

The ideal complexity parameter (cp) is the one that results in the least error. As tree size increases, associated with lower and lower potential values of the cp, a minimum value can be identified. However, as this could still be associated with a large tree, the ideal value is taken as one that is within one standard error of the minimum value. This allows the choice of a cp which is low enough to reduce misclassification but not so low that it will end up generating a tree that is as complex as the initial one.

15.7 Tree based decision rules for GP diagnosis of TIA

An internal validation was carried out using predicted probabilities from the modelling that created the decision nodes in the tree using R software. Discrimination was assessed using cumulative frequency plots and corresponding ROC curves to demonstrate the difference between unpruned and pruned trees.

The performance measures of sensitivity, specificity and positive and negative predictive values change with progression along the tree. This is equivalent to using only the first one, two or three etc. splits and then examining the results in terms of misclassification summarised via a 2X2 table after each successive split to give the usual four measures of a diagnostic test performance. If the total tree is still too large or complex to use in practice, we can examine how using an initial subset of the tree could help support decision making. We can consider the utility of the predictor choice and degree of importance in terms of how far along the tree they are introduced by examining the cumulative performance at successive splits.

Given the risks of not referring patients with TIA i.e. failing to prevent a recurrent stroke in a proportion of patients, it is clearly more harmful to misdiagnose a patient with TIA as non-TIA, than to misdiagnose a patient without TIA as having had a TIA. Patients without TIA who are diagnosed with the aid of a decision rule in primary care will be referred, having
been given medication (currently aspirin only according to NICE guidelines) and may be investigated. As such the harms of referral for these non-TIA patients are the temporary exposure to aspirin and potential harms from any brain imaging (ionizing radiation exposure via computed tomography, claustrophobia from magnetic resonance imaging) and financial costs to patient and healthcare system of appointments and investigations.

In order to reflect the fact that misdiagnosing a TIA as a non-TIA is more undesirable than misdiagnosing non-TIA as TIA, a cost can be applied to missing a true TIA diagnosis. This cost may result in different predictors or different cut points being chosen such that there are fewer misclassifications of true TIA as predicted non-TIA compared with a tree without such costs. Using the same dataset, a classification tree was derived using a cost matrix specifying that the costs of misclassification of true TIA as predicted non-TIA were twice that of true non-TIA as predicted TIA.

15.8 One tree or multiple trees? Potential for ensemble prediction via Random Forest analysis

Where models are used for forecasting in economics and meteorology, combinations of predictions from different models have been found to reduce errors and provide more accurate estimates than using individual models (278). This approach has not been systematically tested in clinical situations with different clinical prediction models for a given diagnosis.

One further consideration in improving prediction is the generalisability of prediction models generated using the OXVASC dataset. How sensitive are the methods to slight alterations in the prevalence of the clinical features of the patients with and without TIA? If the models are ‘over-fitted’ i.e. they perform well at internal validation but poorly on external validation, then they will have little clinical utility in day to day practice.

Classification trees are data driven in that all predictors are taken into consideration as potentially helpful in splitting a group of data into two sub groups with maximal difference between them in the outcome (276). It could therefore be argued that slight alterations in the data may result in different predictors being chosen at a given split or a different order of similar predictors which would result in different patients being selected for referral or reassurance. Methods that can incorporate the influence of perturbations in the derivation dataset may offer more reliable prediction at external validation.
15.9 Random Forests – theory and potential advantages

A random forest is a group of classification trees which have all been derived from subsets of data, drawn from the same overall dataset (279). A criticism of deriving a single classification tree is that it is susceptible to small changes in the data and particularly the choice of the first node, i.e. the most important predictor, may differ depending on the underlying predictor prevalences within the diagnostic groups (280). In order to have greater confidence in the generalisability of model predictions, the random forest in effect tests susceptibility of predictor choice to random changes in training data.

Each tree is derived from a random two thirds of the original dataset, termed the ‘training set’, which implies that one third of the data is not used to construct the tree (and that third of the data is given the label ‘out of bag’). For successive trees, each one is derived using a different randomly chosen two thirds of the dataset. Thus some data points may be used multiple times or possibly not at all. Each tree selects predictors from a randomly chosen subset of all predictors at each node (the size of this subset can be varied) (277).

The random forest is comprised of all the trees that are derived by the underlying algorithm - the total number of trees can be set for each forest derivation but no more accuracy is gained beyond 500 trees (280).

Each tree can be validated on the data not used in its derivation i.e. the one third of data that is ‘out of bag’, by comparing the predicted diagnosis using the tree with the actual diagnosis for each patient in the out of bag dataset. This can be averaged over all 500 trees in the forest to produce an average ‘out of bag’ error rate. This is calculated as a misclassification rate and is presented as a simple 2X2 table.

However, the actual discriminatory performance of the forest is measured by allowing each tree in the forest to classify a set of data. Thus in this case there will be 500 predictions of the presence of TIA for each patient, which are summarised to give one output. The classification output is by ‘vote counting’ where a simple majority of trees ‘wins’ so if for a given patient in a test dataset 260 trees in the forest predict TIA and 240 trees predict ‘not TIA’ then that patient is predicted to have a TIA. The output of predicted probability of TIA is given as the proportion of trees that classify a patient as TIA, so for the above example the predicted probability of TIA is 260/500 = 0.52.

The importance of variables used by the trees in the forest is determined by two measures. One assesses how well the predictor creates two sub groups that are different with respect to the outcome, in this case TIA prevalence (this is termed a reduction in ‘node impurity’ as
two sub groups should be created with one having a higher TIA prevalence than the parent node and one with a lower TIA prevalence than the parent node). This is known as the Gini index and a variable is more important if it results in a greater decrease in ‘impurity’ after it is used to split a node. The Gini importance is an indication of how often it was used by the different classification trees in the forest and the size of the discriminative value.

The other measure describes the importance of predictors by assessing how much prediction accuracy reduces if they contain random noise rather than the original datapoints. This is calculated by taking all the classification trees using a certain predictor, and then adding ‘noise’ or random error to the values of that predictor in the out of bag data but leaving the other predictors intact in the out of bag data. Then the trees make predictions from the ‘permuted’ out of bag data and the prediction error is compared to that found using the unadulterated out of bag data (279). By surveying all the classification trees in a forest, the importance of different predictors can be found and their relative importance characterised. Of the two measures, the Gini measure is thought to be more robust (281).

Variable selection by a random forest analysis has been used to determine the importance of predictors for clinical conditions such as depression in patients with rheumatoid arthritis (282).

15.10 Derivation of a Random Forest using Primary Care data for diagnosis

The ‘randomForest’ package in R software was used to generate a random forest for TIA diagnosis using GP histories in the full referral dataset. This allows for a comparison of the predictors that the different modelling techniques select.

A random forest will provide perfect discrimination of its training data as the classification trees are grown to maximal depths, even down to single patients in the final nodes and so separate datasets are required to derive sensitivity and specificity. The classification error from the out of bag data does not allow for adjusting the threshold probability for diagnosis as a majority vote is used for classification and as such varying sensitivity and specificity over a range of thresholds is not possible.

Therefore in order to test the predictions from a random forest, I decided to split the primary care patient dataset into two randomly chosen and equal halves. This is so that a forest can be derived from one half, and then the actual prediction probability used from the other half to construct a receiver operator characteristic curve. This allows calculation of sensitivity and specificity for comparable metrics with the outputs from the other prediction methodologies.
15.11 Comparison with other modelling techniques

In order to create a fair comparison, the logistic regression model and the classification tree model were re-derived using the same derivation dataset, and then validation metrics calculated by applying them to the same validation dataset. This is not an external validation as both datasets are from the same underlying population. All three methods were graphically compared on an ROC plot.

15.12 Using diagnosis specific cumulative frequency plots to visually display discrimination

One limitation of reading an ROC curve is that there is no way to determine the actual optimal score, without further calculation. The area under the ROC curve gives the probability that a randomly selected patient with disease will score more highly than a patient without disease in pairwise comparisons for the entire group. This statistic does not have a clinical interpretation, given that the clinician will be faced with a patient with a suspected diagnosis and a score value.

Differences in distributions of scores between affected and unaffected patients will determine the area under the ROC curve. At a given score of a patient with a disease, the probability that a patient without disease will score less is determined by the fraction of disease-free patients with scores less than that of the randomly selected patient with disease. A cumulative frequency plot of the scores of patients with disease alongside a cumulative frequency plot of patients without disease displays these differences in distributions that determine the area under the ROC. The difference in the cumulative frequency curves of the ABCD2 score for patients with and without TIA for example is shown in figure 15.1. As an example, a cut point of 3 is displayed with the probability that a patient with TIA will score more than 3 and the probability that a patient without TIA will score less than 3.
Figure 15.1 Cumulative frequency of specialist ABCD2 scores among TIA and non-TIA patients with cut point at score = 3

The cumulative frequency plot has the advantage of showing how discrimination (as measured by the probability of patients with disease scoring more than patients without disease) varies with the actual score. If the cumulative frequency plot of patients with disease is right shifted from those without disease, then \(1 - \) (cumulative frequency for disease) is greater than (cumulative frequency in non-diseased) and hence the probability of patients with disease having higher scores than those without disease will be greater which will correspond to a higher area under the ROC curve. If the lines cross over, then patients with disease are more likely to score lower than those without disease.

The relationship between disease specific cumulative frequency plots and the ROC curve concerns determination of optimal cut points. Figure 15.2 shows a section of an ROC curve near the upper left point (coordinates 0% (1-specificity), 100% sensitivity), and the optimal cut point on the curve, which should maximise sensitivity and specificity.
Figure 15.2 Optimal cut point on a portion of an ROC curve

The minimum distance between the curve and the coordinates (0, 1) is the hypotenuse of the right angled triangle with sides (1 – sensitivity) and (1 – specificity). This distance is at its minimum when (1 – sensitivity) and (1 – specificity), constrained by the shape of the curve, are at their minimum values (from Pythagoras’ theorem).

However, (1 – sensitivity) at a given cut point is the % of TIA patients who would test negative at that cut point. This equals the cumulative frequency of scores from 0 to that cut point, given that all patients below the cut point are going to be diagnosed as test negative.

(1 – specificity) at a given cut point is the % of non-TIA patients who would test positive, which equals (1 – cumulative frequency) at that cut point i.e non-TIA patients who would score above the cut point. These concepts are illustrated in figure 15.3, with the distance d representing the difference between the two curves.
Figure 15.3 Cumulative frequency plots of specialist ABCD2 scores for TIA and non-TIA patients, illustrating derivation of sensitivity and specificity at optimum cut point.

Given that at any cut point, \((1 - \text{sensitivity}) + (1 - \text{specificity}) + d = 1\), when \((1 - \text{sensitivity})\) and \((1 - \text{specificity})\) are both at a minimum within the confines of their respective curves, \(d\) is at a maximum. Hence at the optimal ROC curve defined cut point, \(d\) on the cumulative frequency plot is at a maximum.

Therefore cumulative frequency plots allow a visual inspection of the optimal cut point defined by the vertical distance between the cumulative frequencies. They also allow an assessment of where scores fail to discriminate. For example an area under the ROC curve of 0.5 could be due to a 0.5 chance of discrimination throughout the range of scores or to strong positive discrimination for one half of the score range and strong negative discrimination for the other half of the score range.
16.1 Coding clinical features from GP records

The clinical records for 496 referred patients with suspected TIA were available for analysis. Table 16.1 lists the symptom categories with included and excluded features.

The initial univariate analyses of the clinical predictors are shown in Table 16.2. For each potential predictor, the prevalence within TIA patients and non-TIA patients are given along with the relative risk (RR) with 95% confidence intervals and p values for comparison using \( \chi^2 \) distributions (Fisher’s exact test for 0 values).
<table>
<thead>
<tr>
<th>SYMPTOM</th>
<th>Included symptoms</th>
<th>Excluded symptoms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Muzzy Head</td>
<td>Muzzy/fuzzy/funny feeling in head</td>
<td></td>
</tr>
<tr>
<td>Headache</td>
<td>Headache/pressure sensation in head</td>
<td></td>
</tr>
<tr>
<td>Dizziness</td>
<td>Vertigo/dizziness/swimming feeling</td>
<td></td>
</tr>
<tr>
<td>Ataxia</td>
<td>Unsteady walking/loss of control of arm without weakness/difficulty walking without leg weakness</td>
<td></td>
</tr>
<tr>
<td>LOC</td>
<td>reduced level of consciousness/reduced responsiveness</td>
<td>Feeling faint</td>
</tr>
<tr>
<td>Facial weakness</td>
<td>unilateral, upper or lower motor neuron not specified</td>
<td>Bilateral weakness</td>
</tr>
<tr>
<td>Limb weakness</td>
<td>unilateral arm/hand/leg, weakness/heaviness</td>
<td>Bilateral weakness</td>
</tr>
<tr>
<td>Facial sensation</td>
<td>unilateral numbness/paraesthesiae</td>
<td>Bilateral sensory change</td>
</tr>
<tr>
<td>Limb sensation</td>
<td>unilateral arm/hand/leg, numbness/paraesthesiae</td>
<td>Bilateral sensory change</td>
</tr>
<tr>
<td>Nausea or vomiting</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Altered visual perception</td>
<td>distortion/flashes/blurring/arc/zig-zag, monocular or binocular</td>
<td></td>
</tr>
<tr>
<td>Absent visual perception</td>
<td>curtain/darkness/dark patch/clouding monocular or binocular</td>
<td></td>
</tr>
<tr>
<td>Diplopia</td>
<td>Double vision, horizontal or vertical</td>
<td></td>
</tr>
<tr>
<td>Speech disturbance</td>
<td>Slurred speech/global word finding difficulties/receptive and expressive dysphasia</td>
<td></td>
</tr>
<tr>
<td>Word finding difficulty</td>
<td>isolated to specific words (rather than classes of words e.g. nouns), with rest of speech normal</td>
<td></td>
</tr>
<tr>
<td>Memory loss</td>
<td>Procedural or declarative</td>
<td></td>
</tr>
<tr>
<td>Confusion</td>
<td>Patient or witness report</td>
<td></td>
</tr>
</tbody>
</table>

Table 16.1 Categories of GP symptoms for suspected TIA referrals
**Table 16.2 Distribution of symptoms between TIA and non-TIA patients**

<table>
<thead>
<tr>
<th>SYMPTOM</th>
<th>No TIA n/N</th>
<th>TIA n/N</th>
<th>RR</th>
<th>Lower 95% CI</th>
<th>Upper 95% CI</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Muzzy Head</td>
<td>6/295</td>
<td>0/201</td>
<td>0</td>
<td>na</td>
<td>na</td>
<td>0.08</td>
</tr>
<tr>
<td>Headache</td>
<td>33/295</td>
<td>19/201</td>
<td>0.88</td>
<td>0.61</td>
<td>1.28</td>
<td>0.59</td>
</tr>
<tr>
<td>Dizziness</td>
<td>31/295</td>
<td>11/201</td>
<td>0.62</td>
<td>0.37</td>
<td>1.04</td>
<td>0.06</td>
</tr>
<tr>
<td>Ataxia</td>
<td>18/295</td>
<td>13/201</td>
<td>1.02</td>
<td>0.67</td>
<td>1.57</td>
<td>0.92</td>
</tr>
<tr>
<td>LOC</td>
<td>26/295</td>
<td>1/201</td>
<td>0.09</td>
<td>0.01</td>
<td>0.59</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Facial weakness</td>
<td>19/295</td>
<td>19/201</td>
<td>1.24</td>
<td>0.89</td>
<td>1.74</td>
<td>0.32</td>
</tr>
<tr>
<td>Limb weakness</td>
<td>67/295</td>
<td>56/201</td>
<td>1.15</td>
<td>0.92</td>
<td>1.45</td>
<td>0.28</td>
</tr>
<tr>
<td>Facial sensation</td>
<td>18/295</td>
<td>15/201</td>
<td>1.12</td>
<td>0.76</td>
<td>1.65</td>
<td>0.72</td>
</tr>
<tr>
<td>Limb sensation</td>
<td>57/295</td>
<td>19/201</td>
<td>0.57</td>
<td>0.38</td>
<td>0.85</td>
<td>0.003</td>
</tr>
<tr>
<td>Nausea or vomiting</td>
<td>14/295</td>
<td>2/201</td>
<td>0.3</td>
<td>0.08</td>
<td>1.09</td>
<td>0.04</td>
</tr>
<tr>
<td>Altered visual perception</td>
<td>37/295</td>
<td>16/201</td>
<td>0.71</td>
<td>0.47</td>
<td>1.09</td>
<td>0.12</td>
</tr>
<tr>
<td>Absent visual perception</td>
<td>20/295</td>
<td>28/201</td>
<td>1.5</td>
<td>1.14</td>
<td>1.94</td>
<td>0.02</td>
</tr>
<tr>
<td>Diplopia</td>
<td>6/295</td>
<td>7/201</td>
<td>1.32</td>
<td>0.79</td>
<td>1.22</td>
<td>0.5</td>
</tr>
<tr>
<td>Global Speech disturbance</td>
<td>55/295</td>
<td>75/201</td>
<td>1.64</td>
<td>1.34</td>
<td>2.01</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Specific word finding difficulty</td>
<td>11/295</td>
<td>5/201</td>
<td>0.76</td>
<td>0.36</td>
<td>1.58</td>
<td>0.58</td>
</tr>
<tr>
<td>Memory loss</td>
<td>27/295</td>
<td>1/201</td>
<td>0.08</td>
<td>0.01</td>
<td>0.57</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Confusion</td>
<td>24/295</td>
<td>4/201</td>
<td>0.33</td>
<td>0.13</td>
<td>0.84</td>
<td>0.006</td>
</tr>
</tbody>
</table>

Significant positive predictors which increase the likelihood of TIA are absence of visual perception, and speech disturbance. Significant negative predictors which reduce the likelihood of TIA are reduced level of consciousness, unilateral limb sensory change, nausea or vomiting, memory loss and confusion.

16.3 Derivation of the logistic regression model

Age and history of previous vascular disease are significantly associated with TIA in the GP referred population (from chapter 12). These variables were added to the significant predictors derived from clinical descriptions found on univariate analysis in table 16.2 in a multivariate forward stepwise logistic regression model.

All included variables survived adjustment and were included in the final model and table 16.3 displays the beta coefficients in the model with 95% confidence intervals of the exponents of the coefficients.
Table 16.3 Predictors of TIA diagnosis from multivariable logistic regression – beta coefficients with standard error and associated Wald tests.

The associated Nagelkerke’s R² for the model is 0.38, with 72.6% correct categorisations of TIA and non-TIA at the final step of addition of predictors.

16.4 Calibration of the logistic regression model

The probability that TIA is present is given from the equation \( \frac{1}{1 + e^{-lp}} \) where

\[ lp = \text{age} \times 0.04 + \text{history of TIA} \times 1.77 + \text{loss of vision} \times 1 + \text{global speech disturbance} \times 0.52 - \text{reduced level of consciousness} \times 3.77 - \text{unilateral limb sensory symptoms} \times 0.78 - \text{nausea/vomiting} \times 1.99 - \text{memory loss} \times 3.3 - \text{confusion} \times 2.2 - 3.3. \]

Predicted probabilities were calculated for each patient and the spread of predicted probabilities for TIA and non-TIA patients is shown in figure 16.1.
Figure 16.1 Spread of predicted probabilities Non-TIA and TIA

The predicted probabilities were grouped into deciles and compared with the observed chance of TIA for patients within the relevant decile and is displayed in figure 16.2. Within the derivation dataset, there is a linear trend for increasing fraction of TIA patients with higher deciles of predicted probability but the observed probability is overestimated by the model. Nevertheless the model fit from Hosmer-Lemeshow test indicated that observed values did not significantly differ from predicted values (p = 0.65)

Figure 16.2 Observed probability within increasing deciles of predicted probability
16.5 Constructing the score

Two scores were constructed from the predictors in the logistic regression model. One was weighted by the predictive strength, i.e. the beta coefficient and the other score was unweighted with a unit change in cumulative score depending on whether it is a positive or negative predictor. Table 16.4 summarises the two scores.

<table>
<thead>
<tr>
<th>Component</th>
<th>Weighted Score</th>
<th>Unweighted score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Present</td>
<td>Absent</td>
</tr>
<tr>
<td>History of TIA/Stroke</td>
<td>1.8</td>
<td>0</td>
</tr>
<tr>
<td>Global speech disturbance</td>
<td>0.5</td>
<td>0</td>
</tr>
<tr>
<td>Visual loss</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Reduced level of consciousness</td>
<td>0</td>
<td>3.8</td>
</tr>
<tr>
<td>Memory loss</td>
<td>0</td>
<td>3.3</td>
</tr>
<tr>
<td>Confusion</td>
<td>0</td>
<td>2.2</td>
</tr>
<tr>
<td>Nausea/vomiting</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Unilateral sensory change</td>
<td>0</td>
<td>0.8</td>
</tr>
<tr>
<td>Age</td>
<td>X by 0.04</td>
<td>X by 0.04</td>
</tr>
</tbody>
</table>

Table 16.4 Weighted and unweighted scores (from beta coefficients) for GP sourced predictors in a TIA diagnostic prediction rule

16.6 Internal validation of the scores – discrimination using cumulative frequency curves and ROC

Table 16.5 shows mean and standard deviation of weighted and unweighted scores by diagnostic group. Comparison with non-TIA scores for each diagnostic category using weighted and unweighted scores were carried out with the Mann-Whitney test, given the non-normal distribution of scores.
<table>
<thead>
<tr>
<th>Diagnostic Group</th>
<th>N</th>
<th>Weighted Mean (SD)</th>
<th>P for comparison with non-TIA</th>
<th>Unweighted Mean (SD)</th>
<th>P for comparison with non-TIA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-TIA</td>
<td>295</td>
<td>13.91 (1.68)</td>
<td>-</td>
<td>7.42 (1.08)</td>
<td>-</td>
</tr>
<tr>
<td>All TIA</td>
<td>201</td>
<td>15.68 (1.13)</td>
<td>&lt;0.001</td>
<td>8.59 (1.0)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Anterior TIA</td>
<td>162</td>
<td>15.75 (1.13)</td>
<td>&lt;0.001</td>
<td>8.67 (1.01)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Posterior TIA</td>
<td>39</td>
<td>15.36 (1.08)</td>
<td>&lt; 0.001</td>
<td>8.28 (0.90)</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Table 16.5 Mean weighted and unweighted scores by diagnostic group.

Mean scores for anterior TIA and posterior TIA are more similar compared with the ABCD2 and Dawson scores, where posterior TIA means were closer to non-TIA mean values. The discriminative ability of the weighted and unweighted scores was visually displayed using cumulative frequency plots.

Figure 16.3 shows the cumulative percentage of scores for each diagnostic group. The vertical distance between the non-TIA line and both the anterior TIA and posterior TIA lines throughout the range of scores demonstrates good discrimination across a range of scores. The anterior and posterior TIA lines are both right-shifted from the non-TIA line without crossover with the non-TIA line. The posterior line closely follows the anterior line indicating similar discriminating ability throughout the score range.

Figure 16.4 shows the same plot for the unweighted score. Comparing this with figure 16.3 demonstrates a difference in the posterior TIA scores. Both scoring systems offer equivalent discrimination for posterior TIA in the first half of the distribution. However at higher values, the unweighted posterior TIA line moves closer to the non-TIA line compared with the weighted posterior TIA line, indicating that at higher scores the discriminative ability of the unweighted score is likely to be weaker compared with the weighted score.
Figure 16.3 Cumulative percentage of weighted scores by diagnostic group

Figure 16.4 Cumulative percentage of unweighted scores by diagnostic group
The change in sensitivity and specificity with increasing weighted score is shown in figure 16.5. Compared with the ABCD2 and Dawson scores using primary care records, sensitivity is relatively well preserved from the lowest score value although there is a steep drop off as specificity rises with the lines meeting at 70%. The variation in sensitivity and specificity for the unweighted score is shown in figure 9.6.

Figure 16.5 Change in sensitivity and specificity for the diagnosis of TIA with increasing weighted score value

Figure 16.6 Change in sensitivity and specificity for the diagnosis of TIA with increasing unweighted score value

Sensitivity with the unweighted score reduces at an earlier point in the score range compared with figure 16.5 and the specificity rises such that the lines meet at 80%. Thus
compared with the weighted score, higher specificity can be achieved without as steep a drop in sensitivity.

Figures 16.7 and 16.8 show the change in PPV and NPV for increasing values of the weighted and unweighted score. At the lowest value, the PPV is at the prevalence of the referred population (40%) as all patients are deemed to have TIA at the lowest cut off.

Figure 16.7 Change in positive predictive value (PPV) and negative predictive value (NPV) with increasing value of weighted score

Figure 16.8 Change in positive predictive value (PPV) and negative predictive value (NPV) with increasing value of unweighted score

The unweighted score has a more restricted range of values as the increments in score are not greater than 1 and fewer patients score at the lower end of the range. This results in an
earlier rise of PPV and fall of NPV as the early component of the distribution is contained within the first score points compared with the weighted score which spreads the same component out over more score points.

Figures 16.9 to 16.11 display the ROC curves for both weighted and unweighted scores for discrimination of TIA, anterior TIA and posterior TIA from non-TIA. Figure 16.9 shows comparable discriminating ability for both scoring systems with AUC values for the weighted score of 0.81 (S.E. 0.02) and unweighted score of 0.79 (S.E. 0.02).
Figure 16.10 ROC curve for discrimination of anterior TIA from non-TIA using weighted and unweighted scores.

Figure 16.10 shows the discrimination of anterior TIA from non-TIA with AUC values for weighted scores of 0.82 (S.E. 0.02) and unweighted scores of 0.80 (S.E. 0.02).

Discrimination of posterior TIA from non-TIA is shown in figure 16.11 with an AUC 0.77 (S.E.0.04) for the weighted score and 0.74 (S.E. 0.04) for the unweighted score.

Figure 16.11 ROC curve for discrimination of posterior TIA from non-TIA
Thus both scores have significant discriminating ability across diagnostic groups with the weighted score having a marginally but consistently higher AUC.
Chapter 17  Classification Tree

17.1 Classification Tree Derivation

The initial tree is shown in figure 17.1. It is clearly a complicated structure with 17 terminal nodes (final sub groups) formed by 16 separate splits and of these, age is used at 7 splits at different cut-off values. This tree classifies 79.6% of patients correctly (compared with 72.6% for the logistic regression model).

Included variables are previous history of cerebrovascular disease, age, reduced level of consciousness, memory disturbance, confusion, visual loss, dizziness, and unilateral limb sensory disturbance. Most of the included predictors in the tree are ‘negative’ in the sense that they are associated with a reduced chance of TIA diagnosis.

Only previous cerebrovascular disease, age and absence of vision are positive predictors of TIA and these are all included with positive beta coefficients in the logistic regression. Interestingly speech disturbance is not included as a predictor in the tree, which is a positive predictor in the logistic regression, and neither is the logistic regression negative predictor of nausea or vomiting. Instead the classification tree identified dizziness as a predictor, which was not included in logistic regression modelling although on univariate analysis in chapter 16 it had a relative risk (RR) of TIA of less than 1 but the confidence interval of this RR included values greater than 1, with an associated p value just outside statistical significance of 0.06.
Figure 17.1 Classification Tree for TIA diagnosis – ‘unpruned’

There is clearly a cost to having such a complex tree as although more patients may be correctly classified it is not likely to be useful in routine practice if there are 16 steps to go through in order to generate a decision about referral to a TIA clinic. Also a large and complex tree is more likely to overfit the training data and therefore be less useful in clinical practice as it will not predict as well.
17.2 ‘Pruning’ the classification tree

A cross-validation exercise was carried out to determine the optimal trade off between reducing the size of a potentially large and complex tree and maintaining reasonable prediction accuracy. Random sub-groups are formed in the dataset and different sized trees are derived from data outside a given sub-group and their prediction accuracy tested on data inside that sub-group. This process determines the optimal complexity parameter (CP) which determines how large a tree will be.

Figure 17.2 shows how the error of a particular tree size changes (as a proportion of the error in a ‘null model’ which assumes that prevalence of TIA in the whole dataset is the chance that an individual has a diagnosis of TIA - the initial position of an infinite CP, which will not generate any decision nodes). The dotted line represents a cut off point that is one standard error above the minimum value of relative error and any CP that is below this line can be considered an appropriate choice.

![Graph showing reduction in misclassification error with reducing values of CP](image)

**Figure 17.2** Reduction in misclassification error (relative to a classification tree without splits) with reducing values of the complexity parameter (CP)
Figure 17.2 shows that a CP value of 0.019 is within one standard error of the minimum value and represents a reasonable balance between reducing misclassification error and keeping tree size and complexity to a minimum.

A 'pruned' classification tree for TIA diagnosis is presented in figure 17.3. It uses eight splits, resulting in nine terminal nodes and overall it classifies 74.8% of patients correctly (again higher than the 72.6% for the logistic regression model).

**Figure 17.3 ‘Pruned’ classification tree for TIA diagnosis**

The included variables are past history of cerebrovascular disease, age (used to split patients without a history of cerebrovascular disease at a cut point of 48 yrs and to split patients without confusion at a cut point of 77 yrs), reduced level of consciousness (LOC), memory disturbance, confusion, visual loss and dizziness.
17.3 Predicted probabilities from the classification trees

Figure 17.4 shows the calibration curve for the probabilities of TIA predicted by the unpruned and pruned trees (patients are grouped into deciles of predicted probability, and the observed probability within each decile is plotted). Both trees estimate the probability of a TIA diagnosis near the ideal calibration line (assuming that the midpoint of the probability decile is the ideal for observed probability).

![Calibration curve for predicted probability deciles of unpruned and pruned classification trees](image)

**Figure 17.4 Calibration curve for predicted probability deciles of unpruned and pruned classification trees**

The cumulative distribution of the probabilities is shown in each diagnostic group in figure 17.5 for the unpruned tree. Posterior and anterior TIA are discriminated equally well from non-TIA with most of the TIA patients having a predicted probability greater than 0.63.
Figure 17.5 Cumulative predicted probability from unpruned classification tree by diagnostic group

Figure 17.6 Cumulative predicted probability from pruned classification tree by diagnostic group
Figure 17.6 demonstrates that the increase in cumulative frequency for TIA patients occurs at an earlier predicted probability of 0.22. Again, there is no difference in the cumulative distribution for anterior and posterior TIA predicted probabilities from the pruned tree.

17.4 Discrimination – ROC curves

Figures 17.7 – 17.9 show the ROC curves using the unpruned and pruned trees to discriminate between TIA and non-TIA, as well as for the diagnostic sub groups. The AUC for discrimination of all TIA from non-TIA was 0.84 (S.E. 0.02) for the unpruned tree and 0.80 (S.E. 0.02) for the pruned tree (figure 17.7). Discrimination of anterior TIA from non-TIA for the unpruned tree showed an AUC of 0.84 (S.E. 0.02) and for the pruned tree was 0.80 (S.E. 0.02). Interestingly the discrimination performance for posterior TIA was very similar with unpruned tree AUC of 0.86 (S.E. 0.02) and a pruned tree AUC of 0.80 (S.E. 0.03).

Figure 17.7 ROC curve showing discrimination of TIA from non-TIA using unpruned and pruned classification trees
Figure 17.8 ROC curve showing discrimination of anterior TIA from non-TIA using unpruned and pruned classification trees

Figure 17.9 ROC curve showing discrimination of posterior TIA from non-TIA using unpruned and pruned classification trees
The discriminating ability for anterior and for posterior TIA, as judged by AUC on an ROC curve, is identical using the pruned classification tree. The logistic regression model of GP histories showed that discriminating anterior TIA from non-TIA was better than discriminating posterior TIA from non-TIA (0.82 vs 0.77) although the differences were not as marked as discrimination using the Dawson score on GP records (0.75 vs 0.52) or ABCD2 on GP records (0.65 vs 0.48).

17.5 Sensitivities, specificities and predictive values

![Figure 17.10 Sensitivity and specificity at successive splits along the pruned classification tree](image)

Figure 17.10 shows the marked variation in sensitivity and specificity which only appear to reach acceptable values together at the end of the tree, from split 6 to split 8. The values do not linearly progress with increasing splits, i.e. including increasing numbers of predictors. Instead there is pronounced variation particularly from split 1 to split 2. This arises because the tree construction is focussed on the end result for the total tree, and so large changes in performance measures for subsets of the tree are not relevant.

Figure 17.11 shows the change in positive and negative predictive values at successive splits along the tree. Again, both values are widely separated until splits 6 to 8, when they converge to jointly acceptable levels. However, the sustained high negative predictive value from split 2 to split 6 suggests a potential role in decision support in primary care.

If a perfectly performing diagnostic rule is not a realistic goal, then perhaps a referral support tool could be an alternative for implementation. Such a tool would need to function as a ‘rule
out’ test in order to reduce the referral of patients without TIA. A high negative predictive value, if it remains high in a significant proportion of the patient group i.e. at a higher split, suggests that the test could reliably rule out enough of the patients without TIA to be worth implementing in practice. High NPVs occurring at the initial part of the tree only may not be as useful as this may only apply to a small fraction of the dataset without impacting significantly on referral numbers.

![Graph showing positive predictive value (PPV) and negative predictive value (NPV) at successive splits along the pruned classification tree.](image)

**Figure 17.11** Positive predictive value (PPV) and negative predictive value (NPV) at successive splits along the pruned classification tree

### 17.6 Decision rule based on the pruned tree – structure and impact on referral

The suggested decision rule for TIA diagnosis is shown in the flow chart in figure 17.12. These are the questions that a GP would follow and depending on the answer, would refer or not to a TIA clinic. Making a diagnosis in primary care would not currently change management for the GP. This is because there is no primary care access to investigations such as brain imaging (CT or MRI) or carotid ultrasound within the timeframes required to reduce recurrent stroke risk and urgent treatment with the full range of vascular risk reducing medications has not been recommended for use until after specialist assessment. Even if a perfectly accurate diagnostic tool, in terms of predictive values, were available, a GP would still need to refer patients to be seen by a specialist under current limitations on access to key timely investigations and National body recommendations for treatment.

Thus the major role for a diagnostic support tool would be in the more accurate identification of patients for referral to a TIA clinic. As such, the most important feature of a diagnostic tool
is a high negative predictive value, so that the diagnosis can be ruled out with sufficient accuracy in primary care.

Figure 17.12 Decision Algorithm from classification tree for decision to refer for specialist assessment, with simpler algorithm stopping at the dashed line

The cumulative performance of this rule was shown in figures 17.10 and 17.11, with very high negative predictive values up to the 5th split (presence or absence of confusion). If this rule were truncated at the fifth split with all patients that do not have confusion referred to specialists for assessment then this would simplify the rule considerably (up to the dashed line in figure 17.12). It would in essence act as a ‘rule out’ checklist by identifying features that would preclude referral, apart from the first split where all those with a history of previous cerebrovascular disease would be referred.

In this dataset, using the rule up to and including the 5th split would have prevented 117 referrals of patients without TIA (39.7% of all those without TIA), at a cost of missing 5 patients with TIA (2.5% of all those with TIA). This would have resulted in a 25% reduction in referrals overall.
However, the key reason to refer a patient with suspected TIA is to reduce the chance of a recurrent stroke. The risks of recurrent stroke after TIA are high, and if a referral rule does miss cases, then ideally it should not miss patients that are high risk for TIA.

Theoretically if a patient presents to primary care with classical TIA but also with other features that appear in the rule, for example confusion, then it could be argued that they would not be referred when there would be compelling clinical grounds to do so.

17.7 Who does the rule miss?

Understanding which patients are missed by a decision rule provides potential users of that rule with valuable information. Also, where a rule has very different external validation performance metrics to its internal validation ones, an awareness of why a rule misses patients could help in any modification to improve generalisability.

In the internal validation of the decision tree, five patients with TIA are missed. Given the early age cut off, there are four that are missed at the first split, as they are <48 years old. One of these had an inherited disorder of connective tissue supporting vascular structures with a visual presentation. The other patients <48 years old had anterior territory presentations with one having facial numbness, one with facial weakness and one with slurred speech.

The 5th missed patient was 92 years old without a history of cerebrovascular disease and had slurred speech as well as what was described as confusion by the GP. It is likely that the confusion may have been a direct result of speech disturbance with a receptive as well as the recorded expressive component.

Each of the five patients was followed up for eight years and no recurrent vascular events were recorded over that time period. Thus none of these patients were high risk as measured by the occurrence of a recurrent stroke on treatment of their TIA.

However, it could be argued that given the reduction in recurrent stroke risk due to treatment of vascular risk factors post TIA these patients may have been stroke free on follow up, i.e. low risk, because of treatment - but they may have had a stroke had they not have been referred. An assessment of this can be gauged in the short term at least by examining their ABCD2 scores for short term stroke risk.

Of the five missed TIA patients, three had an ABCD2 score of 4, one had a score of 2 and one had a score of 1 (all scored by secondary care). Given the estimated stroke risk at 90 days with these scores, the mathematical expectation of missed strokes at 90 days is
The OXVASC population is 91,000 and this can be estimated as 0.0015 of the UK population. Thus an approximate estimation of the number of missed strokes at 90 days post TIA if this decision tool were to be used in the UK to govern primary care referrals to TIA clinics would be 181 over four years, or 45 strokes per year at 90 days post TIA.

There is no equivalent validated stroke prediction tool for long term stroke risk but it will be higher than the 90 day risk. Therefore the cumulative number of missed strokes will increase with time as patients with TIA missed by the rule will have recurrent vascular events outside the short term 90 day risk window.

17.8 The effect of weighting misclassification with ‘costs’

Applying a misclassification cost resulted in a complex tree with 14 splits and 15 final nodes with 77.4% correct classification. Eight predictors were used to generate the splits (age, reduced level of consciousness, loss of memory, previous history of cerebrovascular disease, confusion, dizziness, altered visual perception and unilateral limb sensory change), with age used seven times with seven different cut points.

A pruned tree was generated after cross-validation showed that the optimal complexity parameter was 0.045, and the resulting classification tree had three splits (age, reduced level of consciousness and memory loss) and four terminal nodes. This classified 61.5% of patients correctly, with 110 true non-TIA patients predicted as non-TIA (37.3% specificity) and 6 true TIA patients predicted as non-TIA (97% sensitivity).

These figures do not improve on the first 5 splits of the unweighted tree used above and although the weighted pruned tree is simpler it in fact misses more patients with TIA. The cost matrix reduces the number of missed patients overall compared with the total unweighted tree but as the initial section of the unweighted tree was used to generate a decision rule there has been no improvement in the performance as a rule out algorithm.
Chapter 18  Random Forest

18.1  Choice of Predictors from a random forest - comparison with other techniques

The contribution made by different variables to TIA diagnosis in the whole GP dataset is shown with a list in descending order of importance in figure 18.1. Two methods to determine variable importance are displayed; accuracy reduction, which tests the strength of a predictor by replacing the predictor with random values and assessing how much worse the predictions are, and the ‘Gini’ assessment, which assigns importance in terms of how often trees in the forest use a predictor and the improvement in classification that it gives.

The Gini assessment of importance shows that the nine most important predictors are age, previous history of vascular disease, speech disturbance, sex, reduced level of consciousness, limb weakness memory disturbance and absence of vision. The Dawson tool's nine predictors do not include sex, memory disturbance or absence of visual symptoms, but this was not based on GP recorded clinical features.

The logistic regression model (Chapter 16) has nine predictors and it shares seven with the random forest - age, history of vascular disease, speech disturbance, reduced level of consciousness, absence of vision, memory loss and unilateral sensory disturbance in a limb. The pruned classification tree has seven predictors and of the most important seven in the random forest it shares four - age, history of vascular disease, reduced level of consciousness and memory disturbance.
Figure 18.1 Variable importance measures in random forest prediction of TIA from training dataset

18.2 Misclassification rate for full dataset of referred patients

Using the total GP dataset, the random forest reported a misclassification rate of 27.42% from the out of bag data, i.e. a correct classification rate of 72.58%. This compares with the logistic regression internal validation of 72.6% and the classification tree internal validation of 79.6% correct classification. The random forest may give a better indication of performance at external validation by incorporating variation in predictor prevalences (279).
18.3 Derivation from the training dataset – comparing probabilities in the testing dataset

To compare the model methodologies fairly, the techniques of logistic regression, classification tree derivation and random forest derivation were repeated on the training dataset. Predicted probabilities for TIA in the testing dataset were calculated for the three models and the calibration of observed probability within predicted probability deciles is shown in figure 18.2. The calibration to the testing dataset demonstrates over and under prediction across increasing deciles of predicted probability with no model clearly calibrating better than another. The classification tree does not provide the same spread of predicted probability as the random forest or logistic regression.

Figure 18.2 Observed probability within each decile of predicted probability for TIA diagnosis – logistic regression, classification tree and random forest

Figure 18.3 demonstrates the cumulative distribution of predicted probabilities from the random forest by diagnostic group. Patients with posterior TIA are not disadvantaged by using this modelling technique as there is no reduction in the ability to discriminate between posterior TIA and non-TIA compared with anterior TIA and non-TIA.
18.3 Cumulative frequency of predicted probability in of TIA by diagnostic group using a random forest model

18.4 Discrimination with different modelling techniques – ROC curves

Figure 18.4 shows the ROC curve for prediction of TIA diagnosis in the testing dataset, with predictions from random forest alongside logistic regression and classification tree models from the same derivation dataset. The AUCs are similar with the highest from the logistic regression model with 0.80 (S.E. 0.03), classification tree 0.79 (S.E. 0.03) and random forest 0.72 (S.E. 0.03).

Figure 18.5 demonstrates similar discriminating ability for anterior TIA from non-TIA in the testing dataset, with similar AUCs with logistic regression 0.80 (S.E. 0.03), classification tree 0.79 (S.E. 0.03) and random forest 0.72 (S.E. 0.03).

Figure 18.6 demonstrates a similar pattern for discriminating posterior TIA from non-TIA with AUCs for logistic regression of 0.81 (S.E. 0.05), classification tree 0.78 (S.E. 0.05) and random forest 0.75 (S.E. 0.06).

Although the random forest predictions show good discrimination for both anterior and posterior TIA, they are not as good as predictions from logistic regression or a single classification tree in terms of AUC values.
Figure 18.4 Discrimination of all TIA from non-TIA – comparison across modelling techniques

Figure 18.5 Discrimination of anterior TIA – comparison across modelling techniques
Figure 18.6 Discrimination of posterior TIA – comparison across modelling techniques

18.5 Predictive values, sensitivity and specificity with random forest analysis

As random forests give a predicted probability in terms of vote counting (i.e. the number of trees where a particular set of parameter values give a positive diagnosis / the number of trees in the forest), rather than relying on the simple majority vote which determines the misclassification rate, the predicted probability itself could be used to determine a cut point for diagnosis.

The potential high rule out function is shown more easily in the sensitivity and specificity plot in figure 18.7, where a very high sensitivity of 97% is associated with the lowest decile cut point of 0.1.

Figure 18.8 shows the changes in positive and negative predictive value using the random forest predictions in the testing dataset for different cut points of predicted probability.
Figure 18.7 Sensitivity and specificity of random forest derived predicted probabilities of TIA with increasing decile diagnostic cut points

At the lowest probability decile cut point of 0.1, there is a very high negative predictive value so the forest predictions have a good rule out function at this level. Using this cut point as a referral tool would reduce non-TIA referrals by 29% whilst missing 3% of patients with TIA. The patients missed by the rule did not have recurrent cerebrovascular events at six years of follow up.

Figure 18.8 Positive predictive value and negative predictive value with increasing deciles of predicted probability of TIA from a random forest voting
Chapter 19  Summary and Discussion

19.1  Clinical predictors included in a prediction rule for TIA diagnosis derived from primary care records in suspected TIA

A qualitative analysis of routine clinical records and referral letters from GPs identified groupings of symptoms which were then tested as potential predictors alongside age and past history of vascular disease which are routinely available in primary care. Using forward stepwise multivariable logistic regression, a similar process to the approach of Dawson (222), a model for prediction of TIA diagnosis was constructed with positive predictors of age, previous TIA/CVA, loss of vision and speech disturbance and negative predictors of reduced level of consciousness, limb sensory disturbance, nausea or vomiting, memory disturbance and confusion.

Diagnostic models have been viewed as a special case of prediction models (283), where data is collected at a certain time and a future state is to be predicted, in this case the knowledge of diagnosis after a test to be carried out in the near future. The development process for stroke and TIA recognition tools have used predictors from specialist assessment of patients. A diagnostic prediction rule for primary care, based on factors elicited in primary care was derived and then internally validated.

Comparing the predictors found by Dawson, the GP model shares only one focal neurological predictor that increases the probability of TIA - speech disturbance (excluding isolated specific word finding difficulty). The GP model also includes a new focal symptom predictor - reduced visual perception (rather than present but altered visual perception). Dawson included hemianopic visual disturbance as a potential predictor but did not specify whether this was positive phenomena (altered perception) or negative phenomena (reduced or absent visual perception) and this did not add predictive value in their logistic regression after adjustment by other predictors (222). The only visual symptom in the Dawson model is diplopia, which was not a significant predictor in the GP data and so was not included in model building.

Unilateral weakness was not a predictor in the GP model, but both facial and limb weakness were independently predictive of TIA in the Dawson model. Sufficient numbers of non-TIA cases are felt by GPs to have weakness so that the discriminating ability of this symptom for TIA diagnosis disappears.

The GP model includes a focal neurological symptom as a negative predictor. This is prime facie a counter-intuitive finding as focal symptoms are more likely to have focal pathology i.e.
regional cerebral dysfunction. However there are a number of peripheral neurological aetiologies for localised sensory symptoms such as nerve entrapment or plexus pathological changes. A recognised functional presentation (or medically unexplained symptom) in neurological clinics is hemisensory syndrome (284), which may also explain why unilateral sensory symptoms are negative predictors in GP recorded symptoms (the recording of these symptoms in secondary care may be affected by clinicians’ diagnoses).

There is one shared negative predictor in both models – altered level of consciousness. This suggests global rather than regional cerebral dysfunction with a potential cardiac cause due to transient reduced effective circulating blood volume.

GPs did not record data on seizure activity, which may reflect the fact that locally a pre-defined seizure clinic is in operation and so any patients with suspected ictal activity will be referred there or that when complex seizures are present e.g. as a differential for altered responsiveness/level of consciousness, GPs do not suspect seizure activity as a cause and so it is not recorded in their notes. The prevalence of headache between the TIA and non-TIA patients was not significantly different, and so was not included as a potential predictor in the GP model.

Nausea or vomiting, memory loss and confusion were additional negative predictors in the GP model, and these were not even considered as potential negative predictors by Dawson. Their inclusion is a direct effect of the strategy of analysing spontaneous recordings by GPs and then assessing predictive strength individually, rather than relying on pre-existing secondary care derived scoring systems to identify potential predictors. Although nausea and vomiting may be due to posterior circulation TIA and have been identified as appearing more commonly in posterior stroke (275) they are a strong negative predictor among GP records, possibly as a better identifier of migraine in the GP referred population than headache, which was not a significant positive or negative predictor.

This model shares speech disturbance and reduced consciousness with the Dawson tool but importantly only speech disturbance and loss of vision are identified as positive predictors from GP histories. Visual symptoms improve the sensitivity of the FAST test to detect posterior circulation stroke (275) so their inclusion in a recognition tool may improve identification of posterior circulation TIA. However, I separated visual symptoms into altered perception and absence of perception i.e. positive and negative phenomena as a natural grouping of the symptoms. This grouping resulted in different predictive abilities in this dataset.
19.2 Calibration and discrimination metrics for a model of TIA diagnosis derived from primary care records in suspected TIA

The logistic regression model calibrated well. The major finding from ROC analysis was that posterior TIA patients had similar discrimination compared with anterior TIA. AUCs for anterior and posterior TIA discrimination from non-TIA were 0.81 and 0.77 for the weighted score and 0.79 and 0.74 for the unweighted score respectively. The NPV is very important in primary care as we wish to safely reduce the TIA clinic burden by identifying TIA patients while reducing the non-TIA referrals. The weighted score at a cut point of 14 has a sensitivity of 98.5% and a specificity of 52% with an NPV of 97% and a PPV of 40%. These metrics are an improvement on that reported for the internal validation of the Dawson tool. For an increase in 10% of predictive value over baseline prevalence, a smaller fraction of TIA patients are missed. For a similar increase in predictive value of 10% over baseline, the unweighted score metrics are a sensitivity of 95% and a specificity of 65% which misses more patients with TIA but reduces more non-TIA referrals.

19.3 Choice of statistical model and discrimination metrics for TIA diagnostic models

Although ‘uncertainty’ in modelling is usually restricted to predictor choice and uncertainty over parameter values (285), there is a further dimension of uncertainty over the choice of statistical model to represent the relationships between predictors and outcome. There are few reports in the literature where authors have compared different modelling methodologies to reflect such model uncertainty, and they tend to be restricted to situations where there are large numbers of physiological variables in complex settings such as intensive care (286;287). The approaches of classification trees and ensemble prediction i.e. combining multiple model outputs were tested in the GP referral dataset.

A classification tree was constructed to explore if an alternative prediction methodology to logistic regression could be used to derive a decision support tool for TIA referrals from primary care. Even without adjusting for the fact that missing a patient with TIA is worse than referring a patient without TIA to clinic, a simplified decision algorithm could deliver a referral support tool that is very sensitive to TIA diagnosis and does not vary with arterial territory. However, a small number of patients with TIA are still missed which on a large enough scale would result in patients potentially suffering from preventable recurrent stroke. Adding in a weighting factor to try to minimise missing patients with TIA did not improve this.

Two tree based models were derived, one subject to a complexity parameter for simplification. Cumulative predicted probability distributions showed that both anterior and
posterior TIA patients were discriminated similarly from non-TIA patients across scores for both unpruned and pruned trees. The majority of predictors used in the pruned tree were ‘rule out’ i.e negative predictors after the first node split from history of TIA/stroke.

Internal validation showed similar AUC values on ROC curves compared with logistic regression and the pruned tree was equal in discriminating anterior and posterior TIA from non-TIA (AUCs of 0.80). To further reduce complexity, a simple first section of the tree was analysed for its ability to refine a referral population showing a PPV of 53% (an increase of 12% over baseline prevalence) and an NPV of 96% with a sensitivity of 98%, missing 5 true TIA patients in the dataset. Using the methods of Lavallee et al in predicting counterfactual stroke risk in treated patients (42), the internal validation metrics suggest that a very small number of strokes may occur if this rule is implemented nationally whilst reducing referrals by 40%.

The most important predictors selected by the random forest, taking the first nine to compare with the logistic regression models, showed closer similarity to those of the Dawson tool, with limb weakness included. Although the classification rate for the random forest derived from the whole GP dataset was similar to the internal validation classification rate for logistic regression, an experimental comparison by splitting the dataset into testing and training data to examine predicted probabilities showed that logistic regression outperformed both random forest and single classification tree in terms of discrimination. Nevertheless, as a rule out for referral the random forest performed well by reducing referrals by 29% and missing 3% of patients with TIA, although cut points of predicted probability were used to decide on an optimal point of high negative predictive value, rather than relying on the dichotomous random forest output i.e. classified correctly or incorrectly, from majority voting. However, there is no clear benefit to the use of such complex models for clinical diagnosis.

Selecting variables by their performance in one particular dataset may result in a phenomenon termed ‘testimation bias’ (285), which is a similar concept to regression to the mean. If, in a certain dataset we detect a significant predictor, this could be due to a type 1 error i.e. in two groups defined by disease presence or absence, by chance there is an unequal distribution of a non-causative variable which we would then select for use in a prediction model. This would lead to poorer external validation as the predictor is not a true predictor of the outcome of having the disease and this is usually a larger problem for weaker associations (285). By creating multiple bootstrapped derivation and validation datasets, such randomly occurring and weak associations may be less likely to be selected from strong performance across the variations in a forest but only for predictors with equal categories for responses, such as presence or absence of clinical features (288). Similarly
not including one particular predictor because it was not useful in a given dataset also creates a bias and Moons et al argue that this may also result in a poorer model (273).

The use of random forests to improve diagnosis, a problem of classification, has not been explored in clinical medicine. The statistical field is young and the methodology is still emerging (289). Nevertheless the three different modelling approaches were given identical derivation and testing datasets by splitting the whole set of GP referred data so that predicted probabilities from each model could be compared on ROC curves. Despite the theoretical advantages of the random forest in terms of variable selection and prediction accuracy it had poorer AUC values for all discriminations than logistic regression and classification trees.

19.4 Limitations of this analysis

No other clinician examined the categorisation of symptoms or suggested alternative categorisations. Furthermore although the total derivation dataset was a total pool of 515 patients, there were a number of referred patients’ records that were not available for qualitative analysis (11 TIA patients and 8 non-TIA patients) but this represents a data loss of only 4%.

Another limitation is that the GP variables were not collected systematically by every GP for every patient as the GPs were acting as part of their routine practice. A limitation of the models developed in this thesis therefore relates to generalisability, given the non-systematic way in which the clinical variables were recorded. For prognostic models and by logical extension, diagnostic models, clearly defined and systematically collected predictors are ideal (274). However, in other settings spontaneous report of new symptoms has been used to support clinical recognition particularly for the early features of meningitis with parental report rather than clinical notes being used to identify features early in the time course of the illness (290).

19.5 Future Research and Practice

A GP history derived regression model offers promise at referral refinement whilst maintaining the identification of patients with TIA, particularly posterior TIA. However, the regression weighted scoring system is complex to use, as it requires not only multiplication but addition as well of different numbers either because of the presence or because of the absence of a symptom. This may limit the generalisability of logistic regression-derived models in general. Removing the weights to create a simpler score may improve usage as
the reduction in overall discriminating ability is modest. Nevertheless a simpler rule based tool may offer greater ease of use.

The classification tree output is the simplest to implement without any need for calculation or indeed for going through an entire set of variables. The decision tree could identify patients that should be referred from very few variables and is therefore simpler to implement. The next research steps will address the key limitation of this thesis – non-systematically collected predictors without an external validating dataset.

The different modelling approaches of logistic regression and classification trees have produced different diagnostic tools which have very different implementations in terms of variable choice and what is done with those variables. Both have the potential to improve clinic usage but miss patients with TIA, although an external validation will be required to determine likely impact of using outputs from these models.

It is not known to what extent GPs would use a decision support tool for TIA clinic referrals. No tools have yet been established for TIA diagnosis and GPs are therefore used to referring a low prevalence condition from instinct and acumen. Qualitative work will therefore be required to understand if GPs would be willing to use a tool as a rule out mechanism so that a TIA clinic and confirmatory opinion from a specialist would not be needed for reassurance, either for themselves or the patient.

Secondly, in order to ensure as far as possible that predictors are collected in a systematic way, the symptom categories with included and excluded descriptions used for derivation of tools should be made available to GPs, possibly via clinic referral forms, to create a prospectively collected set of clinic referral data at multiple clinic sites for derivation and validation datasets. This will also ensure that there is a variation in specialist decisions over diagnosis, as relying on one specialist for outcome definition may not capture all patients with TIA given the known disagreement between specialists.

Thirdly, an implementation trial would be needed where GPs are given information about likelihood of TIA from their referral which could take a number of forms, such as a flowchart with the decision tree, or an online referral tool with live feedback about TIA likelihood from inputted data. Outcomes of such a trial would be positive predictive value of GP referral, appropriate clinic usage for urgent slots, follow up of patients who are deemed not to have TIA by the decision rule to assess rate of stroke above background for a non-TIA population, as well as ease of use and how often the rule was countermanded by strong clinical suspicion.
Chapter 20 Final Conclusions

The OXVASC registered population does not represent the national population in terms of deprivation and this limits generalisation of statistical associations with deprivation or lack thereof, in the patient cohort in OXVASC. Inspite of variations between practices in age and deprivation structure, presentation of TIA to healthcare has not been affected which allows pooling across practice populations to answer research questions over healthcare access using observational data.

Healthcare access after TIA and minor stroke is influenced by choice of provider, timing of symptoms and clinical features of the event. The majority of patients after TIA and minor stroke seek care in general practice and will wait to be seen at their registered practice if their symptoms start in the OOH period. Delay in calling for medical attention is strongly influenced by timing of symptom onset and availability of routine primary care at that time, unless patients chose ED for their initial healthcare assessment or request an assessment from OOH primary care services. Therefore, the organisation and delivery of primary care influences healthcare seeking behaviour after TIA and minor stroke and currently is a barrier to achieving the National Stroke Strategy targets, due to delays in access for events occurring at weekends.

Since the introduction of the new GMS contract, the majority of patients with TIA and minor stroke continue to present initially to primary care during, and outside of, normal working hours. The GMS contract has, however, increased the use of OOH primary care services. Therefore the contract has improved utilisation of OOH primary care, at least for patients with TIA and minor stroke.

A pilot study of GP trainees showed a failure to recognise high risk TIA cases from vignettes that had been constructed from TIA cases missed in routine clinical practice and associated with early recurrent stroke. In order to construct the vignettes, patients with high risk TIA were identified by examining primary care records before an acute admission for completed stroke. This demonstrated a cohort of patients who, by definition, will not be included in outpatient clinic databases. This is because such patients are seen in general practice after TIA but have a completed stroke and admission before the opportunity to attend a specialist outpatient clinic or they have consulted in general practice after a TIA but the diagnosis was not considered.

The recruitment and completion rates of the pilot study were low indicating a much larger recruitment pool will be needed to answer the question of whether the high risk TIA phenotype is initially well understood during training but there is reduced recognition of TIA.
at later stages in the professional career due to low clinical exposure, or whether such clinical presentations are not well recognised throughout training and independent clinical practice.

The alteration of single cues in the vignettes affected recognition and management decisions in the pilot sample of trainees. Whilst an appropriately powered study is required to test the influence of single parameter change, this suggests that hypotheses for why certain cases are missed can be tested using vignettes which can then lead on to an appropriate educational intervention.

Disagreement exists between GP and specialist over the clinical features of TIA, with greatest disagreement over duration of symptoms, less for weakness and least for speech disturbance. This will affect the performance of prognostic and diagnostic tools that have been derived from specialist accounts of events if they are to be used in primary care.

The disagreement in clinical features is not random in that there is a tendency for specialists to document fewer features than described by the GP in patients who do not have TIA and vice versa for patients with TIA. This explains the variation in mean differences in ABCD2 scores between GP and specialists for patients with and without TIA.

The difference in ABCD2 scores between GP and specialists implies inaccurate risk stratification in primary care. This will result in a missed opportunity to prevent a small number of early recurrent strokes as a small number of patients with high risk TIA will be inaccurately classed as low risk. It will also result in inappropriate use of urgent clinic slots as low risk patients with be inaccurately classed as high risk, thereby contributing to difficulties in meeting National Stroke Strategy targets for urgent access out-patient clinic assessment for patients with high risk TIA.

The existing tools that have been proposed to aid diagnosis of TIA perform reasonably well in specialist records of patients with suspected TIA but perform poorly in validations using primary care clinical records, particularly for patients with posterior circulation symptoms. Therefore they are not appropriate to be used in primary care.

The clinical predictors included in a diagnostic model derived from routinely recorded clinical notes made by GPs of patients referred with suspected TIA do not match the classical TIA phenotype. Furthermore, weakness, an important prognostic marker of recurrent stroke risk, was not included as GPs found weakness to be equally prevalent in patients with and without TIA. This limits the usefulness of this method of diagnostic model derivation.
Nevertheless, the primary care based prediction model calibrated well and showed good discrimination for TIA. It did not demonstrate a differential accuracy for anterior and posterior TIA symptoms on internal validation. However, as a large number of predictors were included, and this complexity is increased by weighting the subsequent diagnostic score using the beta coefficients from the regression model, it is likely to be too cumbersome to be used in routine practice.

The choice of statistical technique for derivation of a diagnostic model affects the metrics of discrimination. The theoretically advantageous and more complex modelling technique using ensemble prediction had inferior performance compared with logistic regression and single classification tree methods. Although classification tree models have similar performance in terms of accurate identification of patients with TIA, they are likely to be simpler to implement and can have high negative predictive values. This is important for primary care as we wish to safely rule out the diagnosis, thereby increasing the positive predictive value of referral and appropriate usage of TIA clinic slots.

The limitations of this thesis include the method of imputation for missing data for the analysis of call delay, the sample size of the pilot study of GP trainees, the retrospective scoring of the ABCD2 score from primary care records, the observational nature of the primary care clinical data in that predictors were not systematically collected using standardised definitions, I alone coded the primary care clinical data into predictor categories of nervous system dysfunction and the derivation dataset for the diagnostic model did not include all high risk TIA patients in the OXVASC cohort.

In order to address some of these limitations and to derive a diagnostic tool that is likely to validate well in primary care, a prospective study of all patients presenting to primary care with transient symptoms of neurological dysfunction is required. This will include patients where the diagnosis of TIA is not recognised, as the cohort will be defined by transient alteration of nervous system function irrespective of the underlying aetiology. It will also include patients who are recognised as having TIA but go on to have a recurrent stroke before being seen in an out-patient clinic (those with hyperacute stroke risk). Furthermore, because clinic-based derivation datasets are ‘pre-selected’ by referrers in primary care, the data is subject to selection bias. This is likely to explain why weakness is not included in the TIA diagnostic model from clinic datasets – patients without TIA were also found to be weak and their similarity to the TIA phenotype is why GPs referred them to the TIA clinic. Given that a TIA diagnostic rule should be applied to all patients with transient symptoms in order to be effective, it should not be derived from patients where GPs have already considered the diagnosis and referred.
Nevertheless such a study is challenging. Firstly because of the time required to collect data at the initial consultation in primary care and this will reduce recruitment and therefore population capture of all the phenomena that could predict TIA diagnosis. Secondly patients in primary care commonly present with multiple problems and the transient event may be at the end of a list of issues that patients wish to discuss, and GPs may not think of recruitment given the complexity of such consultations. Thirdly patients access a primary care opinion other than in their registered surgeries. For example, home visits are not usually the ideal situation for recruitment partly due to lack of supporting research materials either electronic or paper-based and presentation to out of hours primary care is unlikely to result in recruitment and data capture given that the clinician will for the most part not be from the patients’ own practice.
Reference List


Ref Type: Online Source


(37) Stroke Foundation of New Zealand. For the assessment and management of people with recent transient ischaemic attack (TIA). Wellington: Stroke Foundation of New Zealand Inc.; 2008.


(43) Giles MF, Flossman E, Rothwell PM. Patient behavior immediately after transient ischemic attack according to clinical characteristics, perception of the event, and predicted risk of stroke. Stroke 2006 May;37(5):1254-60.


Cragg DK, McKinley RK, Roland MO, Campbell SM, Van F, Hastings AM, et al. Comparison of out of hours care provided by patients’ own general practitioners and commercial


(142) Carroll C, Hobart J, Fox C, Teare L, Gibson J. Stroke in Devon: knowledge was good, but action was poor. J Neurol Neurosurg Psychiatry 2004 Apr;75(4):567-71.


(186) Josephson SA, Hills NK, Johnston SC. NIH Stroke Scale reliability in ratings from a large sample of clinicians. Cerebrovasc Dis 2006;22(5-6):389-95.


(223) Josephson SA, Sidney S, Pham TN, Bernstein AL, Johnston SC. Higher ABCD2 score predicts patients most likely to have true transient ischemic attack. Stroke 2008 Nov;39(11):3096-8.


(227) Kinsella JA, Tobin WO, Cogan N, McCabe DJ. Interobserver agreement in ABCD scoring between non-stroke specialists and vascular neurologists following suspected TIA is only fair. J Neurol 2010 Dec 25.


Appendix 1

Information for MD Examiners

Where the work was carried out

Department of Primary Care Health Sciences, University of Oxford, 2nd Floor 23-38 Hythe Bridge Street, Oxford OX1 2ET

Stroke Prevention Unit, Department of Clinical Neurology, University of Oxford, John Radcliffe Hospital, Headley Way, Oxford OX3 9DU

Statement of Work Done

I conceived and designed the comparison of the OXVASC registered population with the national population and was advised by Dr Rafael Perera, (Statistician, Department of Primary Care Health Sciences) on the use of Poisson regression.

I conceived, designed and analysed the study of the influence of primary care delivery and the new GMS contract on healthcare seeking behaviour. I collected the data on routes to care and delay for patients with TIA, and Dr Arvind Chandratheva (Research Fellow, Stroke Prevention Unit) collected the care route and delay data for the minor stroke patients.

I conceived and designed the study of TIA recognition and collected the data for questionnaire development. I designed the questionnaire and collected the respondents’ data and produced a narrative summary.

I conceived and designed the study for the comparison of primary and secondary care histories and analysed the validation of the ABCD2 and Dawson recognition tools for diagnostic accuracy in referred populations.

I conceived and designed the study for derivation of novel prediction rules using routine observational data and collected the predictor variables from the GP records. I analysed the data and was advised by Dr Patrick McSharry (Mathematician at the Smith School for Enterprise and the Environment, University of Oxford) to use random forests for classification but undertook the training in R software and analysis personally.

Supervision

Professor David Mant supervised the development of the research questions and methodologies as well as the writing of the thesis.

Professor Peter Rothwell supervised the study of provider choice and delay in presentation to healthcare.

Publications arising from this work

Ethics Committee Applications

The TIA recognition study was given favourable opinion by the East Midlands Research Ethics Committee via the Integrated Research Application System (ref:11/EM/0252).

The Oxford Vascular Study ethical approval is from the Oxfordshire Research Ethics Committee (ref: CO.043)
Appendix 2

Search Strategy

Primary searches

OVIDSP was used as the interface for MEDLINE and EMBASE, incorporating subject headings and ‘*’ for wildcard search. Age limits to include all adult patients were specified separately for EMBASE and MEDLINE.

Delay in seeking healthcare after TIA or stroke

1. transient isch*emic attack
2. TIA
3. transient cerebral isch*emia
4. mini stroke
5. 1 OR 2 OR 3 OR 4
6. stroke
7. Cerebrovascular disease*
8. Cerebrovascular accident
9. brain infarct*
10. cerebral infarct*
11. cerebellar infarct*
12. CVA
13. cerebrovascular disease/di, dm, dt, ep, et, pc, th [Diagnosis, Disease Management, Drug Therapy, Epidemiology, Etiology, Prevention, Therapy]
14. 6 OR 7 OR 8 OR 9 OR 10 OR 11 OR 12 OR 13
15. 5 OR 14
16. delay*
17. time
18. timing
19. timelin*
20. access
21. health care access /
22. 16 OR 17 OR 18 OR 19 OR 20 OR 21
23. 15 AND 22
24. patient* perception
25. patient* recognition
26. 24 OR 25
27. 15 AND 26
28. primary medical care /
29. primary health care /
30. 28 OR 29
31. 15 AND 30
32. 23 OR 27 OR 31
33. limit 32 to humans
34. limit 33 to " all adult (19 plus years)"
35. limit 33 to (adult < 18 to 64 > years or aged< 65+ years)
36. 34 OR 35
Relationship between deprivation and stroke

1. transient isch*emic attack
2. TIA
3. transient cerebral isch*emia
4. mini stroke
5. 1 OR 2 OR 3 OR 4
6. stroke
7. Cerebrovascular disease*
8. Cerebrovascular accident
9. brain infarct*
10. cerebral infarct*
11. CVA
12. cerebrovascular disease/di, dm, dt, ep, et, pc, th [Diagnosis, Disease Management, Drug Therapy, Epidemiology, Etiology, Prevention, Therapy]
13. 6 OR 7 OR 8 OR 9 OR 10 OR 11 OR 12
14. 5 OR 13
15. economic status
16. socio*economic status
17. socio*economic position
18. depriv*
19. health inequality
20. 15 OR 16 OR 17 OR 18 OR 19
21. 14 AND 20
22. limit 21 to humans
23. limit 22 to " all adult (19 plus years)"
24. limit 22 to (adult < 18 to 64 > years or aged< 65+ years)
25. 23 OR 24

Diagnosis of TIA and stroke

1. transient isch*emic attack
2. TIA
3. transient cerebral isch*emia
4. mini stroke
5. 1 OR 2 OR 3 OR 4
6. stroke
7. Cerebrovascular disease*
8. Cerebrovascular accident
9. brain infarct*
10. cerebral infarct*
11. CVA
12. cerebrovascular disease/di, dm, dt, ep, et, pc, th [Diagnosis, Disease Management, Drug Therapy, Epidemiology, Etiology, Prevention, Therapy]
13. 6 OR 7 OR 8 OR 9 OR 10 OR 11 OR 12
14. 5 OR 13
15. diagnosis/
16. computer assisted diagnosis/
17. delayed diagnosis/
18. diagnostic accuracy/
19. diagnostic error/
20. diagnostic reasoning/
21. diagnostic test/
22. diagnostic test accuracy study/
23. diagnostic value/
24. differential diagnosis/
25. early diagnosis/
26. qualitative diagnosis/
27. quantitative diagnosis/
28. 15 OR 16 OR 17 OR 18 OR 19 OR 20 OR 21 OR 22 OR 23 OR 24 OR 25 OR 26 OR 27
29. 14 AND 28
30. limit 29 to humans
31. limit 30 to "all adult (19 plus years)"
32. limit 30 to (adult < 18 to 64 > years or aged< 65+ years)
33. 31 OR 32

**The Cochrane Library search strategies are listed below**

**Delay in seeking healthcare after TIA or stroke**

1. transient isch*emic attack
2. TIA
3. transient cerebral isch*emia
4. mini stroke
5. #1 OR #2 OR #3 OR #4
6. stroke
7. Cerebrovascular disease*
8. Cerebrovascular accident
9. brain infarct*
10. cerebral infarct*
11. cerebellar infarct*
12. CVA
13. MeSH descriptor Cerebrovascular Disorders explode all trees
14. #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13
15. #5 OR #14
16. delay*
17. time
18. timing
19. timelin*
20. access
21. MeSH descriptor Health Services Accessibility explode all trees
22. #16 OR #17 OR #18 OR #19 OR #20 OR #21
23. #15 AND #22
24. patient* perception
25. patient* recognition
26. #24 OR #25
27. #15 AND #26
28. MeSH descriptor Primary Health Care explode all trees
29. #15 AND #28
30. #23 OR #27 OR #29

**Relationship between deprivation and stroke**

1. transient isch*emic attack  
2. TIA  
3. transient cerebral isch*emia  
4. mini stroke  
5. #1 OR #2 OR #3 OR #4  
6. stroke  
7. Cerebrovascular disease*  
8. Cerebrovascular accident  
9. brain infarct*  
10. cerebral infarct*  
11. cerebellar infarct*  
12. CVA  
13. MeSH descriptor Cerebrovascular Disorders explode all trees  
14. #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13  
15. #5 OR #14  
16. economic status  
17. socio*economic status  
18. socio*economic position  
19. depriv*  
20. health inequality  
21. #16 OR #17 OR #18 OR #19 OR #20  
22. #15 AND #21

**Diagnosis of TIA and stroke**

1. transient isch*emic attack  
2. TIA  
3. transient cerebral isch*emia  
4. mini stroke  
5. #1 OR #2 OR #3 OR #4  
6. stroke  
7. Cerebrovascular disease*  
8. Cerebrovascular accident  
9. brain infarct*  
10. cerebral infarct*  
11. cerebellar infarct*  
12. CVA  
13. MeSH descriptor Cerebrovascular Disorders explode all trees  
14. #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13
15. #5 OR #14
16. MeSH descriptor Delayed Diagnosis explode all trees
17. MeSH descriptor Diagnosis, Computer-Assisted explode all trees
18. MeSH descriptor Early Diagnosis explode all trees
19. MeSH descriptor Diagnostic Errors explode all trees
20. MeSH descriptor Diagnosis, Differential explode all trees
21. MeSH descriptor Diagnostic Techniques, Neurological explode all trees
22. MeSH descriptor Diagnostic Self Evaluation explode all trees
23. #16 OR #17 OR #18 OR #19 OR #20 OR #21 OR #22
24. #15 AND #23

Secondary Searches
Reference lists of articles retrieved from the primary searches were examined for additional articles relevant to delay in presenting to healthcare, the link between deprivation and stroke and methods to diagnose TIA and stroke.
Appendix 3

TIA case vignettes

1. A 59 year old man attends surgery. He is concerned by two recent episodes – during the first one, he suddenly dropped his keys and noticed that fingers in the right hand felt weak. This lasted five minutes and then resolved spontaneously. He had a further episode a week later, which came on suddenly whilst writing, and this too resolved completely after a few minutes. He has had no previous medical problems and he is a smoker. He has a normal motor and sensory examination with BP 158/72 and a regular pulse of 62.

Alternative – Weakness is not associated with specific tasks (holding keys, writing) and therefore comes on ‘suddenly’ (Hypothesis - Does onset unrelated to action make GPs more likely to suspect TIA?)

2. A 44 year old man with a 24 year history of type 1 diabetes attends the surgery for an urgent appointment. That morning he woke with general malaise but no fever and had a BM of 19.2 and no ketones on his home urine dipstick. He gave himself an extra dose of actrapid and went to work where one hour later his colleagues noticed that his speech had started to sound slurred, he re-checked his BM and it had come down to 4.1. When you see him in the surgery in the afternoon he looked well, was afebrile, BP 138/62 P 72 and regular, he had a clear chest, normal heart sounds, normal speech and normal motor and sensory exam in the limbs. His BM was 5.1, and there was a trace of protein on urine dip.

Alternative – BM does not fluctuate but remains elevated throughout day, even after dose of short-acting insulin (Hypothesis – Does a reducing blood sugar make GPs less likely to suspect TIA if focal symptoms are present?)

3. You are asked to visit an 82 year old man at home. After waking, his wife went downstairs to make a cup of tea and on returning upstairs found him sitting with his shirt on back to front and a ‘confused look on his face’. After this he got himself dressed and went downstairs. He has a history of controlled hypertension and had a hip replacement 2 years ago. He is usually independently mobile and he and his wife manage with some support from their son who lives nearby. On examination he is orientated in time, person and place and remarks that he “can’t see what all the fuss is about”. BP is 151/72, P 68 and regular, he has normal speech, no motor or sensory deficit and he is walking normally. Temperature 37.2, chest clear, heart sounds normal, normal abdomen and urine dip NAD.

Alternative – no dressing apraxia present (Hypothesis – Is confusion treated the same as confusion with dressing apraxia?)

4. You are asked by a carer to visit an 81 year old woman whom you know well. The carer noticed that the patient’s speech was a little slurred and she was more unsteady than usual yesterday and was concerned. On arrival, the patient has no complaints whatsoever and feels well. The patient did not notice any problems yesterday. You examine her and find no focal deficit of speech, cranial nerves or upper or lower limbs. She is afebrile with a clear chest and nothing on her urine dip. She is furniture walking but you remember this as being her usual state from previous visits.

Alternative – the patient also complains of the symptoms in agreement with the carer (Hypothesis – Do disconcordant patient/witness histories reduce the likelihood that GPs will suspect TIA?)
5. A 67 year old woman comes to the surgery and describes a two week history of intermittent numbness and weakness on the right hand side, affecting both arm and leg, although not always at the same time. The episodes have no obvious precipitating cause and seem to resolve spontaneously after a few minutes. She feels lethargic and generally unwell. On examination she is afebrile with BP 132/82. Not disorientated, with normal power and sensation in upper and lower limbs.

**Alternative – absence of lethargy and general malaise (Hypothesis – does the presence of symptoms of systemic upset reduce the likelihood that GPs will suspect TIA as the cause of focal symptoms?)**

6. You are taking triage telephone calls during an Out of Hours shift on a Sunday morning. A 48 year old man rings you saying that he has woken up feeling dizzy with the room spinning and had to go back to bed. He has no weakness or sensory loss and no headache. He has not been vomiting. He has no history of hypertension, angina or diabetes and takes no regular medication. He is a smoker

**Alternative – presence of nausea and vomiting (Hypothesis – is vertigo with vomiting treated the same as isolated vertigo?)**

7. A 75 year old woman comes to see you and complains of increasing episodes of left arm weakness over the last two weeks. A few months before she injured her elbow in a car accident but did not sustain a fracture. Her elbow is normal, there is no sensory loss and you detect slight thenar eminence wasting in the left hand. Power is symmetrical and reflexes are symmetrical. She has a history of asthma, glaucoma, sciatica and is a smoker. She takes regular inhalers and no other medication.

**Alternative – no injury to the elbow (Hypothesis – does an alternative but less plausible explanation for symptoms reduce the likelihood of a TIA diagnosis?)**

**Distractor (non TIA) cases**

1. A 71 year old man presents for the third time with backache that has not resolved with anti-inflammatory medication. He has had several courses of antibiotics for recurrent chest infections over the past four months and says he now feels breathless walking up stairs. He has hypertension and gout, and takes bendrofluazide, amlodipine and allopurinol. He is afebrile BP 152/68, clear chest, normal heart sounds with a regular pulse of 78 and no ankle oedema. Urinalysis shows ++ proteinuria. **Diagnosis - Myeloma**

2. An 81 year old woman who is usually very active and lives alone without support comes to her annual review for CKD and ischaemic heart disease. She complains of increasing breathlessness initially walking up hill but now she notices it walking up stairs. She has had no chest pain. Medications include ramipril 5mg, bisoprolol 2.5mg, aspirin and simvastatin. Routine bloods taken a year ago were normal with eGFR 51. Examination shows a regular pulse of 74, BP 118/72, soft first heart sound with apical systolic murmur, bibasal chest crackles and pitting ankle oedema. **Diagnosis - Heart failure**

3. A 68 year old woman attends morning surgery having experienced visual disturbance the day before, where the left side of her vision suddenly became dark for five minutes. She covered her left and right eye alternately and the problem was confined to the left eye. She has a headache which has been present for two weeks, and is throbbing in nature without any nausea. Over the last few months she has been tired and generally unwell. She has no past medical history and is on no regular medications. Visual fields, fundoscopy and cranial nerve exam are normal. She is
afebrile, BP 162/88, regular pulse of 68 with a clear chest, and you hear an early ejection systolic murmur. **Diagnosis – Temporal arteritis**

4. A 78 year old man attends for the second time in a month with upper abdominal discomfort. He has had no change in bowel habit, no rectal bleeding, vomiting or haematemesis. However he reports that he is finding it increasingly difficult to swallow food and feels that it gets stuck pointing to his mid-chest, although he has no problem with liquids. He has nodal osteo-arthritis and hypertension, and takes regular co-codamol, lisinopril and prn NSAIDS. He is an ex-smoker and drinks little alcohol. On examination there is no abdominal tenderness or organomegaly. **Diagnosis - Oesophageal neoplasm**

5. The husband of an 85 year old woman requests a home visit. On arrival he describes that after waking that morning she was confused and wasn’t sure where she was, but her speech sounded normal. Afterwards she seemed much better and got herself dressed and came downstairs. She takes bendrofluazide and amlodipine for hypertension. On examination she has normal visual fields and cranial nerves, normal speech and is orientated in time, person and place. MTS is 10/10. Limb examination is normal. She has a clear chest, normal heart sounds with BP 138/71, pulse 88 and regular and is afebrile. Urine dip shows a trace of protein only. **Diagnosis – Transient confusion, to act as the paired case for TIA case**
Appendix 4

Materials for GP Trainee High Risk TIA Recognition Study

1. **Email to GP Trainees**

SUBJECT: GP Trainee Questionnaire Study

Dear Trainee

The Department of Primary Care Health Sciences, University of Oxford is conducting a research study of how GP trainees manage common problems that occur in primary care (The Chief Investigator is Dr Daniel Lasserson, Clinical Lecturer). This consists of a questionnaire with ten short written cases describing a clinical presentation to a GP. After each case there are two questions “What would you do now?” and “What do you think is the likely diagnosis?”. You are under no obligation to start or complete the questionnaire. No personally identifiable data will be recorded in the questionnaire although it will ask for your year of training (ST1 to ST4) and whether you have had training in elderly medicine. The questionnaire should take no more than 25 minutes to complete.

There is a link to the questionnaire in this email below. If you decide to answer the questions, I would ask that you find time outside clinical working hours to complete it.

In four weeks time, the link will close and then the performance of the group as a whole will be emailed to all the trainees that were invited to take part. Detailed descriptions of the cases will be sent then as well, including suggested optimal answers to the questions. Your views will then be sought (anonymously) about the educational value of this exercise.

Thank you for considering completing the questionnaire

2. **Front sheet of Questionnaire**

**GP Trainee Questionnaire Study**

Thank you for considering taking part in this questionnaire research study. Common conditions can have a wide range of presentations in general practice and there are some diseases where prompt recognition and management in primary care is important for improving long term health outcomes. This study has been designed to assess how well current training prepares future GPs for the more challenging clinical cases that they might encounter in routine practice.

You have been invited to take part in this study because you are currently training to be a GP on the Oxford rotation, either undertaking a hospital based or community based post from ST1 to ST4.

The questionnaire consists of ten case vignettes, which are short descriptions of a presentation to a GP either in surgery or at home, with key features from history, examination or simple tests that are widely available in primary care. After each description there are two questions “What would you do now?” and “What do you think is the likely diagnosis?”. The responses are in free text and you can write as much or as little as you wish. It should take no longer than 25 minutes to complete and we request that it is undertaken outside of clinical working hours.

You will also be asked which year of training you are in and whether you have undertaken elderly medicine as a GP ST in your training so far. No personally identifiable information will be recorded and all responses are anonymous.
After the questionnaire study period has closed, you will be sent more details on the clinical cases and the aggregated performance of all those who undertook the questionnaire. You will then be invited to answer a question about how valuable you felt this exercise was for your training, which will also be anonymous.

All research in the NHS is looked at by an independent group of people, called a Research Ethics Committee, to protect your interests. This study has been reviewed and given favourable opinion by East Midlands Research Ethics Committee.

Given the nature of this study, it is highly unlikely that you will suffer harm by taking part. However, the University has arrangements in place to provide for harm arising from participation in the study for which the University is the Research Sponsor.

If you wish to complain about any aspect of the way in which you have been approached or treated during the course of this study, you should contact Dr Daniel Lasserson, (email Daniel.lasserson@phc.ox.ac.uk or telephone 01865 289357) or you may contact the University of Oxford Clinical Trials and Research Governance (CTRG) office on 01865 572224 or the head of CTRG, email heather.house@admin.ox.ac.uk

Thank you for considering taking part

Dr Daniel Lasserson
Clinical Lecturer
Department of Primary Care Health Sciences
University of Oxford.

3. Vignettes used in the questionnaire

The husband of an 85 year old woman requests a home visit. On arrival he describes that after waking that morning she was confused and wasn't sure where she was, but her speech sounded normal. Afterwards she seemed much better and got herself dressed and came downstairs. She takes bendrofluazide and amlodipine for hypertension. On examination she has normal visual fields and cranial nerves, normal speech and is orientated in time, person and place. MTS is 10/10. Limb examination is normal. She has a clear chest, normal heart sounds with BP 138/71, pulse 88 and regular and is afebrile. Urine dip shows a trace of protein only.

A 71 year old man presents for the third time with backache that has not resolved with anti-inflammatory medication. He has had several courses of antibiotics for recurrent chest infections over the past four months and says he now feels breathless walking up stairs. He has hypertension and gout, and takes bendrofluazide, amlodipine and allopurinol. He is afebrile BP 152/68, clear chest, normal heart sounds with a regular pulse of 78 and no ankle oedema. Urinalysis shows ++ proteinuria.

An 81 year old woman who is usually very active and lives alone without support comes to her annual review for CKD and ischaemic heart disease. She complains of increasing breathlessness initially walking up hill but now she notices it walking up stairs. She has had no chest pain. Medications include ramipril 5mg, bisoprolol 2.5mg, aspirin and simvastatin. Routine bloods taken a year ago were normal with eGFR 51. Examination shows a regular
pulse of 74, BP 118/72, soft first heart sound with apical systolic murmur, bibasal chest crackles and pitting ankle oedema

You are asked to visit an 82 year old man at home. After waking, his wife went downstairs to make a cup of tea and on returning upstairs found him sitting with his shirt on back to front and a confused look on his face. After this he got himself dressed and went downstairs. He has a history of hypertension and a left hip replacement due to osteo-arthritis. He and his wife manage with some support from their son who lives nearby. On examination he is orientated in time, person and place and remarks that he “can’t see what all the fuss is about”. BP is 151/72, P 68 and regular, he has normal speech, no motor or sensory deficit and he is walking normally. Temperature 37.2, chest clear, heart sounds normal, normal abdomen and urine dip NAD.

A 68 year old woman attends morning surgery having experienced visual disturbance the day before, where the left side of her vision suddenly became dark for five minutes. She covered her left and right eye alternately and the problem was confined to the left eye. She has a headache which has been present for two weeks, and is throbbing in nature without any nausea. Over the last few months she has been tired and generally unwell. She has no past medical history and is on no regular medications. Visual fields, fundoscopy and cranial nerve exam are normal. She is afebrile, BP 162/88, regular pulse of 68 with a clear chest, and you hear an early ejection systolic murmur.

A 44 year old man with type 1 diabetes attends afternoon surgery for an urgent appointment. That morning he woke with general malaise but no fever and had a BM of 16.2 and no ketones on his home urine dipstick. He gave himself a shot of short acting insulin aspart and went to work where one hour later his colleagues noticed that his speech had started to sound slurred, he re-checked his BM and it had come down to 4.1. He takes aspart three times daily, long acting insulin glargine once daily, simvastatin and ramipril. On examination he looks well, is afebrile, BP 138/62 P 72 and regular, he has a clear chest, normal heart sounds, normal speech and normal motor and sensory exam in the limbs. His BM is 5.1, and there is a trace of protein on urine dip.

You visit an 84 year old man who lives alone but has twice daily visits from a carer. He has been more withdrawn recently due to depression after the death of his wife and has a history of hypertension. Earlier that morning he suddenly felt very unsteady and called out to his carer but felt that his speech was slurred. His carer said that she found him holding on to the back of the sofa and when she asked him what was wrong he sounded as if he was drunk. On examination he has normal speech and is fully orientated. He has normal upper and lower limb sensory examination and is walking slowly and cautiously but does not feel unsteady.

A 78 year old man attends for the second time in a month with upper abdominal discomfort. He has had no change in bowel habit, no rectal bleeding, vomiting or haematemesis. However he reports that he is finding it increasingly difficult to swallow food and feels that it gets stuck pointing to his mid-chest, although he has no problem with liquids. He has nodal osteo-arthritis and hypertension, and takes regular co-codamol, lisinopril and prn NSAIDS. He is an ex-smoker and drinks little alcohol. On examination there is no abdominal tenderness or organomegaly.
You are asked by a carer to visit an 81 year old woman whom you know well, with a past history of hypertension and depression, taking amlodipine and venlafaxine. The carer noticed that the patient’s speech was a little slurred and she was more unsteady than usual yesterday and was concerned. On arrival, the patient does not complain of any symptoms and feels well. She did not notice any problems yesterday. You examine her and find no focal deficit of speech, cranial nerves or upper or lower limbs. She is afebrile with a clear chest, BP 163/88, a regular pulse of 72, urine dip NAD. She is furniture walking but you remember this as being her usual state from previous visits.

You see a 49 year old man with type 1 diabetes in an Out of Hours clinic on a Saturday morning. The day before he felt unwell on waking and his post breakfast BM was 18.2. He didn’t have any ketones using his own urine dipstick, and gave himself an extra shot of actrapid. His BMs were still high before lunch and before his evening meal (13.3, 15.8). His partner noticed that his speech was slurred later that evening and his BM at the time was 9.1. She thought he should see a doctor, so he contacted the OOH service this morning but now feels better with BM 6.2. He takes actrapid three times daily and once daily insulatard. On examination he has normal speech, normal motor and sensory exam in the limbs and normal gait. BP is 118/64 and pulse 58 regular. Urinalysis NAD.