doi:10.1093/fampra/cmn071 Family Practice Advance Access published on 7 October 2008

Diagnostic difficulty and error in primary care—a systematic review

Olga Kostopoulou, Brendan C Delaney and Craig W Munro

Kostopoulou O, Delaney BC and Munro CW. Diagnostic difficulty and error in primary care—a systematic review. *Family Practice* 2008; **25**: 400–413.

Background. Diagnostic error in primary care can have serious implications for the patient, the clinician and the health-care system, possibly more so than other types of error.

Objective. To identify common characteristics of diseases that GPs may misdiagnose.

Methods. Systematic search of the MEDLINE and EMBASE databases for primary research on diagnostic error/delay in primary care. Papers on system errors, patient delay, case reports, reviews, opinion pieces, studies not based on actual cases and studies not using a systematic sample were excluded from the review. Twenty-one papers were included. All papers were assessed for quality using the GRADE system. Common features were identified across diseases and presentations that made diagnosis difficult and led to error/delay.

Results. Most studies were retrospective cohorts of patients recruited in hospital and collected data from patient interviews and/or hospital records, resulting in incomplete and potentially biased information. It was usually not possible to determine preventability of the delay. Some conditions were extremely rare, suggesting a specialist research interest rather than an increased rate of misdiagnosis. Conditions investigated were malignancies, myocardial infarction, meningitis, dementia, iron deficiency anaemia, asthma, tremor in the elderly and HIV. Common features of difficulty were atypical presentations, non-specific presentations, very low prevalence, the presence of co-morbidity and perceptual features and could be missed.

Conclusions. Misdiagnosis in primary care covers a wide range of conditions that may be related in the manner in which they present. The challenge is to identify ways of supporting the diagnostic process in potentially difficult presentations.

Keywords. Atypical presentation, diagnosis, diagnostic delay, diagnostic difficulty, diagnostic error, diagnostic overshadowing, disease presentation, non-specific presentation.

Introduction

In recent years, health-care systems worldwide have become aware of the problems posed by patient safety events and the importance of learning from them.[1,2] Given that the majority of clinical encounters occur in primary care, there is potential for a significant volume of such events to occur. The role of GPs as gatekeepers to specialist care in the UK and other countries, e.g. The Netherlands, Canada and Australia, makes diagnosis crucial for the prompt identification and treatment of serious disease. Though diagnostic errors are rarely reported to incident monitoring systems,[3] there is some evidence that they can result in the most serious consequences for patients.[4] Diagnostic error accounts for the greatest proportion of claims against GPs in the UK (63%)[5] and for a third of negligent adverse events in the US primary care (34%).[6] Furthermore, diagnostic errors are those that GPs remember the most and consider the most serious in their career.[7,8]

Diagnostic error can seriously affect patients, health professionals and the health-care system. Malignant or rapidly evolving conditions can cause great harm if undiagnosed. Doctors can suffer litigation and loss of trust, while the cost to the health service is primarily financial. Diagnostic error is due to multiple causes—with cognitive factors being the most prevalent[9,10] and has been associated with different medical conditions. For example, a qualitative study looking at US family

Received 7 July 2008; Revised 4 September 2008; Accepted 9 September 2008.
School of Health and Population Sciences, College of Medicine and Dental Sciences, University of Birmingham, Birmingham, UK. Correspondence to Olga Kostopoulou, Primary Care Clinical Sciences building, University of Birmingham, Birmingham, B15 2TT, UK; Email: o.kostopoulou@bham.ac.uk

physicians' most memorable errors found that the 30 diagnostic errors reported were related to cancers, myocardial infarction (MI), trauma, bowel obstruction and meningitis.[8] An analysis of 1000 UK general practice negligence claims found that delayed or erroneous diagnoses was associated with malignant neoplasms, trauma and infections.[5] An analysis of 5921 negligent adverse events in the US primary care found that the commonest conditions were acute MI and cancers, though 'no single condition accounted for more than 5% of all negligent claims'.[6] It is likely that diagnostic error is not condition specific but that the misdiagnosis of certain conditions is more memorable to GPs and leads to litigation more often due to the severity of consequences.

We conducted a systematic review to determine (i) the conditions where GP diagnostic error or delay have been investigated and (ii) the features that make these conditions difficult to diagnose, increasing the likelihood of error.

Methods

We searched the MEDLINE (1950–2007) and EMBASE (1980–2007) electronic databases for articles about diagnostic error or delay in primary care using the terms below. MeSH terms are signified by a '/', while other terms are text words that were searched for in titles and abstracts.

PRIMARY CARE: (Primary Health Care/OR Primary Medical Care/
OR Family Practice/OR Physicians, Family/
OR General Practice/OR General Practitioner/)
AND
DIAGNOSIS: Exp Diagnosis/
AND
ERROR: (Skill$ OR Delay$ OR Error$ OR Expert$
OR Bias$ OR Fault$ OR Mis$
OR Difficult$ OR Judgment
OR Judgement OR Decision
OR Problem Solving OR Reasoning
OR Information Search OR Information Seeking).

This strategy resulted in 3584 hits. The search strategy was very non-specific and resulted in a large number of papers describing the evaluation of diagnostic tests or diagnostic strategies where one of the 'error' terms appeared in the abstract. In order not to reduce search sensitivity, we did not amend the search strategy but reviewed the results of titles and abstracts for relevance to the review. The total number of potentially relevant hits was 96. We subsequently applied a number of inclusion/exclusion criteria. Inclusion criteria were original research describing diagnostic error or delay in primary care. Exclusion criteria were articles describing only systems-related errors or patient delay in seeking treatment, case reports, reviews, opinion pieces, qualitative studies reporting only GP's opinion and not based on any actual cases and studies using *ad hoc* case series rather than a systematic sample.

After applying these criteria, 21 papers were included in the review. We also reviewed bibliographies of the papers selected, but did not find additional studies. All authors read the papers. They independently assessed the quality of the studies using the GRADE system[11] and agreed on the grading following discussion. The GRADE system allows comparative scores to be made across differing methodologies. Papers are scored for design and quality. Specifically, a design score is given on the basis of the suitability of the study's design to answer the research question. The quality score reflects the appropriateness of the methodology and data analyses. Points are added for each flaw. For example, a point is added if a study estimates diagnostic delay on the basis of patient recall. The evidence provided by each paper was graded on the basis of the combined design and quality scores as of high (1), intermediate (2–3) or low (>3) quality.

In each paper, the authors identified the reasons for the diagnostic error or delay, focusing on features of the diseases or the presentations that made diagnosis difficult and led to the error. We concentrated on the GP diagnosis or referral decision and did not systematically examine delay by other health professionals or patients. Each author listed the diseases and associated features identified. These were then discussed among all authors and an agreement was reached.

Results

Using the GRADE system, only one of the studies included in the review was of high quality, 13 were of intermediate quality and seven were of low quality (Table 1). Both prospective and retrospective studies were identified, as well as an observational study with standardized patients (Table 2). In retrospective cohort studies, a cohort of patients with a particular condition was recruited in hospital or identified from a database. Data were then collected from a range of sources, for example, GP referral letters, interviews with patients and record review. The aim was to determine when and how the patients presented to the GP and identify reasons for diagnostic error or delay. The GPs who had made the referrals were not interviewed. In one study, patients diagnosed with a specific disease in primary care were recruited and assessed using gold-standard diagnostic criteria to measure the extent

TABLE 1 *GRADE scoring of the papers included in the review*

Study	Type of study	Design score	Quality score	Evidence grade	Quality score explained
Bleeker *et al.*[19], NL	Retrospective cohort	2	1	I	Symptoms at presentation based on patient recall only.
Bouma *et al.*[20], NL[0]	Retrospective cohort	2	1	I	Delay estimate and symptoms at presentation depended on patient recall only.
Burgess *et al.*[17], UK	Retrospective cohort	2	2	L	Small sample size. Time-to-referral estimates based on patient recall.
Charlton *et al.*[30], UK	Retrospective cohort	2	2	L	No objective diagnosis of asthma. No data on reasons for diagnostic delay.
Dixon-Woods *et al.*[32], UK	Retrospective, qualitative, purposive sample	3	1	L	Symptoms at presentation were not objectively recorded and relied on parental recall. Therefore, reasons for delay cannot be ascertained.
Goddard *et al.*[29], UK	Retrospective cohort	2	1	I	It is not entirely clear what data were confirmed by record review.
Goyal *et al.*[21], UK	Retrospective cohort	2	1	I	Data based only on hospital records—information about symptoms at GP presentation may be incomplete.
Jiwa *et al.*[23], UK	Retrospective case review	3	3	L	Single practice. Very small number of cases reviewed. Data on presentation not systematically collected.
Jones and Sykes[31], UK	Retrospective cohort	2	3	L	Single practice. No objective diagnosis of asthma. Delay in diagnosis relied on parental recall. No data on the reasons for delay. Data not adequately analysed and presented.
Kantola *et al.*[18], Finland	Retrospective cohort	2	0	I	
Levack *et al.*[22], UK	Retrospective cohort	2	0	I	
Lopponen *et al.*[14], Finland	Cross-sectional prospective cohort study with retrospective review of records	1	0	H	
Meara *et al.*[12], UK	Retrospective cohort	2	1	I	Reasons for overdiagnosis of Parkinson's disease not explored.
Mikulin and Hardcastle[26], UK	Retrospective cohort study	2	1	I	Symptoms and estimates of delay based only on patient recall.
Meineche-Schmidt and Jorgensen[16], Denmark	Prospective cohort study	1	1	I	No information about reasons for not referring patients with alarm symptoms.
Paauw *et al.*[28], US	Observational study with three standardized patients	2	1	I	Small number of cases.
Schnetler[25], UK	Retrospective cohort	2	2	L	Measurements of delay in diagnosis made from referral letter (date of first presentation not checked in primary care records). No reasons for delay identified.
Sorensen *et al.*[27], Denmark	Retrospective cohort	2	0	I	
Van Schayck *et al.*[15], NL	Prospective cohort	1	1	I	No reasons were provided for the missed diagnoses.
Vasudev *et al.*[24], UK	Retrospective cohort	2	3	L	Low response rate (54%). Data based on patient recall only. No data about reasons for delay.
Yates *et al.*[13], UK	Retrospective cohort	2	0	I	

The quality score ranges from 0 (highest) to 3 (lowest). The evidence grading is based on the total score, which is the sum of the design and quality scores. Evidence is graded as low (L), intermediate (I) or high (H). NL, The Netherlands.

of overdiagnosis; however, underdiagnosis could not thus be measured.[12] In another study, patients found to have anaemia 4 years previously were identified from laboratory records, and GP and hospital records were reviewed to examine what had happened since (how the patients had been managed and whether disease had developed).[13]

Few prospective studies were identified. Two prospective cohort studies screened a population for a particular disease and then reviewed the primary care records to check whether the patients identified with the disease had already been diagnosed by the GP.[14,15] The advantage of these studies is that they can directly compare patients who have been diagnosed and patients who have not yet been diagnosed with the disease and identify predictors of the diagnosis. In another type of prospective study, patients with a particular condition were recruited prospectively, and the GPs were asked to record their presenting symptoms and signs.[16] Patients were followed up and their final diagnoses were determined. By linking symptoms at presentation to the final diagnosis, the researchers identified predictors of error or delay. The advantage of this type of study is that the data recorded at presentation are usually more complete and less prone to bias than in retrospective studies where

TABLE 2 *Details of the papers included in the review (condition and population studied, methods, outcome measures and findings)*

Study	Condition	Population studied	Methods	Outcome measures	Findings
Bleeker *et al.*[19], NL	MI	300 patients admitted with MI at one of three hospitals.	Retrospective cohort. Data collected from hospital, ambulance records and patients shortly after diagnosis.	Length of delay from symptoms to admission, broken down by patient delay, GP delay, and ambulance delay. Symptoms at presentation based on patient recall.	Median GP delay was 84 minutes. Longer GP delay associated with atypical symptoms and a smaller overall number of symptoms.
Bouma *et al.*[20], NL	MI	400 patients aged <75 years admitted with MI in three hospitals.	Retrospective cohort. Questionnaire and interview study.	Length of delay from symptoms to admission, broken down by patient delay, GP delay, and ambulance delay. Symptoms at presentation based on patient recall. Patient explanations.	Female sex associated with longer delay. Effect of sex disappeared when 'displacement' was taken into account (patient attributing symptoms to non-cardiac causes).
Burgess *et al.*[17], UK	Breast cancer	185 women with breast cancer recruited at a breast cancer clinic.	1. Retrospective cohort, $n = 132$, <60 years of age.2. Case control study, $n = 53$, ⩾60 years of age. Women interviewed 2 months after diagnosis. Tumour characteristics from clinic records.	Time from symptoms to presentation to referral on the basis of patient recall.	32 (17%) cases were not referred by the GP at initial presentation. The only factor associated with delay was the absence of a palpable mass (OR 3.6, 95% CI 1.6–8.1).
Charlton *et al.*[30], UK	Asthma	212 children with a diagnosis of asthma in two practices, identified from practice records.	Retrospective cohort. Record review.	Delay in asthma diagnosis: age at first consultation for respiratory symptoms minus age at first asthma diagnosis.	Median delay of 3 years or seven consultations. No data available on the actual reasons for delayed diagnosis.
Dixon-Woods *et al.*[32], UK	Childhood cancers	20 parents of children with a range of cancers being treated at a paediatric oncology unit.	Retrospective interview study with parents, focusing on the period before the child's diagnosis. Hospital record review to confirm dates of consultations and investigations.	Parents' experience with the diagnostic process and potential impact on early diagnosis.	10 parents had had disputes with health professionals regarding diagnosis and seven of these had lengthy disputes (2.5–8 months). Parental concerns were not taken seriously by health professionals.
Goddard *et al.*[29], UK	Retinoblastoma	Parents of 100 children with retinoblastoma identified at a UK supra-regional centre.	Retrospective interview study with parents and record review.	Measured delays at several points in diagnosis, including from presentation to GP to referral. Presenting symptoms according to parental recall, corroborated by medical record.	23% had diagnosis delayed more than 8 weeks. Long delay associated with younger child ($P < 0.01$), squint ($P < 0.05$) and initial presentation to health visitor rather than GP or optician ($P < 0.001$).
Goyal *et al.*[21], UK	Bone cancer	103 patients diagnosed with osteosarcoma or Ewing's sarcoma at two hospitals, during 1990–2002.	Retrospective cohort. Hospital record review.	Measured delays at several stages of diagnosis, including from presentation to GP to referral. Delay adjusted for tumour type and professional seen by regression.	GPs delayed longer than Accident and Emergency doctors (1.8 versus 1.3 months). Longer doctor delay (not just GP) associated with age ⩽12 and axial (rather than limb) tumours (4.7 versus 1.4 months).

TABLE 2 *Continued*

Study	Condition	Population studied	Methods	Outcome measures	Findings
Jiwa *et al.*[23], UK	Cancer	56 patients with cancer identified from record review at a single practice.	Retrospective case series.	Reported interval from presentation to diagnosis. Cases with longest delay were discussed among staff to identify patient, professional, task and system factors.	Eight cases were subject to the longest delays and were discussed. A multitude of diverse, contributing factors were identified.
Jones and Sykes[31], UK	Asthma	187 children with asthma (receiving anti-asthma therapy) at a single practice.	Retrospective cohort. Data collected from parents via a questionnaire.	Symptoms (frequency and precipitating factors of wheeze and cough). Family history of atopy. Length of diagnostic delay.	Delays of 'approximately 40% of the age of the child at diagnosis'. Reasons for delay not investigated. 45% 'may' have had bronchitis diagnosed rather than asthma.
Kantola *et al.*[18], Finland	Tongue cancer	75 cases of tongue cancer—all cases in a district in 20 years.	Retrospective cohort. Record review. Data from hospital and primary care linked and anonymized. A stepwise Cox regression model was built to relate delay to survival.	Delay from presentation to final histologically confirmed diagnosis.	65% (49) referred at first GP visit. 48 patients presented with a lesion or pain in tongue and 37 of these were referred promptly. 25 presented with no tongue pain or lesion and only 10 of these were referred promptly.
Levack *et al.*[22], UK	MSCC	319 patients with MSCC diagnosed consecutively at three hospitals.	Retrospective cohort. Review of primary care and hospital records and interview with patients.	Signs, symptoms and delays in diagnostic process.	GP delay 3–66 days (median 18 days). Early symptoms of back and root pain are non-specific for MSCC.
Lopponen *et al.*[14], Finland	Dementia	All 1260 patients of >65 years of age in a small town in Finland.	Prospective cohort study with review of GP records and screening of patients for dementia (using the Mini-Mental State Examination, the Hachinsky ischaemic scale and the Clinical Dementia Rating Scale—patients categorized according to DSM IV criteria).	Prevalence of dementia and rate of detection by GPs.	9% of sample had dementia; only 4% of these had been diagnosed by GPs. Patients with undocumented dementia were more likely to be male, diagnosed with depression, have milder dementia, live at home and be better at activities of daily living.
Meara *et al.*[12], UK	Parkinson's disease (PD)	402 cases of 'presumed' PD (on anti-parkinsonian drugs but not for neuroleptic side effects) from 74 practices.	Retrospective cohort study. Patients were assessed using recommended clinical diagnostic criteria for Parkinson's disease, multiple system atrophy, progressive supra-nuclear palsy and other non-parkinsonian diagnoses (essential tremor, Alzheimer's).	Actual diagnoses	Only 53% (213) had definite diagnosis of Parkinson's disease.
Mikulin and Hardcastle[26], UK	Gastric cancer	83 patients with newly diagnosed gastric cancer during 1981–1982, at a single hospital.	Retrospective cohort. Patients were interviewed upon recruitment and followed up for 2 years to investigate survival.	Patient delay, GP delay and hospital delay.	GP median delay: 7 weeks (3–14).

TABLE 2 *Continued*

Study	Condition	Population studied	Methods	Outcome measures	Findings
Meineche-Schmidt and Jorgensen[16], Denmark	GI cancer	7274 patients presenting with dyspepsia during 1991–1993.	Prospective cohort study. Data on symptoms collected by GP at presentation in structured way. Final diagnoses collected in 1994 by postal questionnaire to GPs.	Mortality and gastrointestinal morbidity per 1000 person-years in patients with and without alarm symptoms. Predictive value of alarm symptoms for cancer and ulcer.	32% of patients without 'alarm symptoms' were referred compared to 48% with alarm symptoms. However, the predictive value of individual alarm symptoms was very low. The presence of an alarm symptom only increased the incidence of cancer from 4.1 to 12.5 per 1000 person-years. Individually, anaemia was the most predictive sign, but of colorectal rather than upper GI cancer (13% of subjects with anaemia).
Paauw *et al.*[28], US	HIV	Stratified, random sample of 134 primary care physicians, selected from a larger sample of 599 physicians recruited by mail.	Observational study. Physicians saw one to three standardized patients (real HIV patients).	Detection and diagnosis of Kaposi's sarcoma, oral hairy leukoplakia and lymphadenopathy.	Kaposi's sarcoma was noticed by 47% (42/89) and misdiagnosed by 45% of these (19/42). Oral hairy leukoplakia was noticed by 56% (54/97) and misdiagnosed by 59% (32/54). Lymphadenopathy was noticed by 17% (23/133). Experience with HIV patients increased correct identification of oral hairy leukoplakia only.
Schnetler[25], UK	Oral cancer	96 cases of oral cancer—all cases seen in three oral surgery departments during 1986–1991.	Retrospective cohort. Data based on analysis of referral letters.	GPs' differential diagnoses, symptoms at presentation, time from presentation to referral, all based on referral letters.	GPs correctly diagnosed tumour in 52% of referrals. Lymphadenopathy and larger tumour present in half of the correct diagnoses. Delay in referral (>2 days) in 36% of cases. No reasons identified.
Sorensen *et al.*[27], Denmark	Meningitis	177 patients with meningococcal disease, identified from a database of public health notifications in a single county.	Retrospective cohort. Analysis of referral letters and hospital records.	Number of GP contacts before admission. GP 'referral diagnosis'. Symptoms and signs on arrival. Actual diagnosis based on microbiological confirmation (n = 166) or typical clinical picture (n = 11).	Disease prevalence was 3.63/100 000 persons/year. 92 patients (52%) had more than one GP contact before admission. GP referral diagnosis more likely to be correct if neck stiffness or petechiae present. Non-petechial rash made correct diagnosis less likely (OR 0.3).
Van Schayck *et al.*[15], NL	Asthma	1155 subjects aged 25–70, randomly selected from 10 urban and rural practices.	Prospective population cohort screened for asthma with FEV_1 and reversibility. Questionnaire and record review.	Disease prevalence, rate of consultation with respiratory symptoms, diagnostic accuracy of GP.	7.5% (86) had objective evidence of asthma. Only 29 of these had ever presented to the GP with symptoms; six of these (21%) had been missed by GP. No reasons are provided.

TABLE 2 *Continued*

Study	Condition	Population studied	Methods	Outcome measures	Findings
Vasudev *et al.*[24], UK	Testicular cancer	180 men newly diagnosed with testicular cancer during 1998–2002.	Retrospective cohort. Patient questionnaire.	Time from GP presentation to being seen in hospital and management decision by GP based on patient recall.	40% referred promptly (within 2 weeks), 51% were referred for scans and/or prescribed antibiotics and 6% inappropriately reassured.
Yates *et al.*[13], UK	IDA	431 patients identified retrospectively from laboratory records as having been found to have microcytic anaemia between June 1997 and May 1998.	Data collected from GP records up to 12 months after the finding of anaemia and from hospital records and death notifications up to 2003 (4 years).	Management and diagnoses. Survival up to 4 years.	43% had had GI investigations. 48 GI cancers and 150 patients with other GI causes were found. 23 had non-GI cancer and 43 other non-GI causes. Patients less likely to be investigated for cause of IDA if female, aged <65, haemoglobin≥9,0 and previous anaemia. The standardized mortality rate was increased three-fold with anaemia and of the 53 who died, 41 had had no investigations and 12 had advanced cancer.

NL, The Netherlands.

presentation data are extracted from the record or based on patient recall.

By definition, diagnostic error implies some avoidable delay in the final diagnosis. The studies used the term 'delay' in different ways. Some studies used the term simply to indicate that referral did not take place at first presentation to the GP.[17,18] Other studies used the term to estimate the time interval between initial presentation to the GP and either referral or eventual diagnosis, implying that this interval exceeded some unspecified threshold of clinical acceptability.[19–23] However, these studies consider cancers and acute MI where urgent referral is an expectation. A small number of studies defined this time interval by guidelines for the specific condition.[24–26] Some studies measured diagnostic accuracy rather than delay, as the extent to which a diagnosis confirmed in hospital was mentioned in the GP's referral letter,[25,27] or by screening a population of patients with a specific diagnosis using gold-standard criteria[12] or by asking clinicians to diagnose cases with diagnoses known to the investigators.[28] One study investigated appropriateness of management for patients with anaemia rather than diagnostic accuracy.[13] With the exception of this study that showed inappropriate lack of investigation of anaemia, the other studies could not determine preventability of delay or diagnostic inaccuracy, since they did not have detailed information about the specific GP–patient encounters. However, certain studies provided evidence of patient characteristics that were associated with increased delay or misdiagnosis, which was the focus of this review.[13,14,19–21,27,29]

A number of conditions were associated with diagnostic error or delay in primary care, cancer being the commonest. We identified five features of potential diagnostic difficulty: atypical presentations, non-specific presentations, very low prevalence, co-morbidity and perceptual features that could be missed. Misdiagnosed diseases often had multiple features, for example, a rare disease presenting with non-specific or atypical symptoms, making its diagnosis even more difficult (Table 3).

Atypical presentations

Atypical presentations have a shortage of prototypical features. These can be defined as features that are most frequently encountered in patients with the disease, features encountered in advanced presentations of the disease or simply features of the disease commonly listed in medical textbooks. Atypical presentations may also have features with unexpected values, for example, right-arm radiation of chest pain in a patient with ischaemic heart disease (left-arm radiation considered more typical of cardiac chest pain).

Breast cancer. One hundred and eighty-five women were interviewed at a breast cancer clinic, 2 months

TABLE 3 *Conditions and associated features of difficulty*

Conditions	Features of difficulty				
	1. Atypical presentation	2. Non-specific presentation	3. Very low prevalence	4. Co-morbidity	5. Perceptual features
Breast cancer[17]	X				
Testicular cancer[24]	X				
Oral cancer[25]	X				
MI[19,20]	X				
Meningococcal disease[27]	X	X			
Dementia (and depression)[14]	X			X	
Asthma[15,30,31]		X			
Childhood cancers[21,32]		X	X		
Upper GI cancer[26]		X			
Tremor in the elderly[12]		X			
MSCC[22]		X	X		
IDA[13]		X			
Tongue cancer[18]	X		X		
Retinoblastoma[29]	X	X	X		
Cancers (various)[23]				X	
HIV[28]			X		X

Conditions in bold suggest misattribution of symptoms to an obvious aetiology or readily available explanation.

after their diagnosis. Seventeen per cent were not referred by the GP at their first visit. This was more likely to happen to patients presenting without a breast lump (a typical symptom of breast cancer) [odds ratio (OR) 3.6, $P < 0.01$].[17]

Testicular cancer. A survey of 180 men diagnosed with testicular cancer found that the mean referral time for the period 1998–2002 (3.55 weeks) was similar to that in 1985 (4.8 weeks).[24] Of the 171 patients who saw their GP at first presentation, 40% were referred promptly, i.e. within the recommended 2-week time frame; 51% were sent for an ultrasound scan and/or prescribed antibiotics at first visit, which delayed referral; 11 men (6%) were reassured and not given a further appointment. Their misdiagnoses ranged from 'nothing wrong' to hydrocoele, strain or 'one testicle bigger than the other'. Although the study does not determine directly how these patients presented, the authors hypothesize that misdiagnoses were due to the GPs not finding a 'typical' testicular mass suggestive of cancer.

Oral cancer. A retrospective cohort study examined the referral letters of all patients seen in three oral surgery departments with a diagnosis of oral cancer.[25] Ninety-six cases were identified, 50 of whom were referred by GPs. A correct diagnosis (mention of tumour or sinister lesion in the referral letter) was made in 52% of GP referrals (26/50). In half of the correct diagnoses, the patient had 'typical features' of lymphadenopathy and a larger size of tumour. The most common misdiagnosis was infection (11/24), suggesting that these cases lacked a typical lesion. However, a lack of information on the presenting symptoms hampers the interpretation of the study. There was a delay in the referral of 18 cases (36%) of more than the recommended 2 days from presentation (median 4 weeks, range 1 week to 8 months).

Myocardial infarction. A retrospective cohort study of 300 patients, admitted to three hospitals in The Netherlands with MI, collected data on pre-hospital delay from a range of sources including interviews with patients about their presenting symptoms.[19] Atypicality of presentation was found to increase delay from presentation to referral. Data regarding the length of time taken by the GP to make a diagnosis were collected, dividing the group into <95 minutes to admission (175 patients) and ⩾95 minutes (125 patients). Patients with either atypical symptoms, such as 'stabbing' chest pain, dizziness, epigastric pain, or a smaller number of typical symptoms (mid-chest pain, left arm pain, sweating, chest tightness) were more likely to be referred to hospital with a delay, with potential consequences for their survival.

A similar study of 400 patients in The Netherlands admitted to hospital with MI found that GPs delayed longer before referring female patients to hospital (52 versus 36 minutes).[20] Female sex is considered to be less typical because of the lower incidence of MI in women. The only factor that accounted for the sex difference in doctor delay was 'displacement', i.e. patients attributing their symptoms to non-cardiac causes. This suggests that females are more likely to misattribute their cardiac problems, contributing to a delay in diagnosis.

Meningococcal disease. Meningitis is one of the most serious infections, where prompt antibiotic use is life

saving. A case of missed meningitis in a child can be a serious issue for a GP, yet the condition is difficult to diagnose for a number of reasons. One retrospective cohort study of 177 cases admitted to hospital divided patients into those where the GP suspected meningococcal disease and those where they did not.[27] The GP diagnosis at referral was more likely to be correct in the presence of neck stiffness [OR 4.5, 95% confidence interval (CI) 2.2–9.5] and petechiae (OR 2.1, 95% CI 1.0–4.6). Both features are typical of meningitis. In contrast, a correct diagnosis was less likely in the presence of a non-petechial rash (OR 0.3, 95% CI 0.2–0.7). A petechial rash was present in 66% of cases of meningococcal disease in the study (116/177), whereas a non-petechial rash was present in 47% of cases (83/177). The two types of rash may coexist. Therefore, a non-petechial rash is less common, i.e. less typical, than a petechial rash in presentations of meningococcal disease but should not be used to exclude the disease, as it is clearly a possible presenting feature. It is also a non-specific feature, common to a number of less serious childhood conditions.

Dementia. A cross-sectional prospective cohort study screened all 1260 residents over the age of 65 in a small town in Finland for dementia and then examined the GP records to determine if this had been recognized by the GP.[14] Fifty-four (4%) had already been diagnosed with dementia or 'memory impairment', while 58 (5%) had not been diagnosed. Patients with undocumented dementia were more likely to be male, depressed, have milder dementia, live at home and be better at activities of daily living. There were no significant differences in GP contact rates between those with documented and undocumented dementia. Underdiagnosis could be attributed to a failure to recognize mild, and therefore possibly atypical, symptoms of dementia that did not interfere with a person's daily life. It is also possible that dementia symptoms were attributed to an existing diagnosis (depression).

Non-specific presentations

Non-specific presentations contain clinical features with poor diagnostic value that do not distinguish well between the differential diagnoses.

Asthma. Cough in a child is typical of asthma but much more frequently associated with self-limiting illness of an infectious aetiology. A study based on retrospective record review of children diagnosed with asthma in two practices found a median delay of 3 years or seven consultations.[30] A study that estimated the length of diagnostic delay in a single practice on the basis of parental recall found delays of 'approximately 40% of the age of the child at diagnosis'.[31] Given that the data in these two studies were collected almost 20 years ago, they are likely to give a biased picture of the extent of diagnostic delay of asthma. A more recent prospective cohort study of 1155 randomly selected subjects used more objective evidence for the diagnosis of asthma (FEV_1 and reversibility) and found little evidence of delay in its diagnosis. Evidence of asthma was found in 86 subjects, only 29 of whom had ever presented to the GP with asthma-related symptoms. Of these, 23 had been diagnosed. No information is provided about why the six cases had been missed.[15] All three papers speculate that asthma symptoms, such as recurrent cough and wheeze, may be attributed to respiratory infections, the airflow obstruction going undiagnosed.

Childhood cancers. Although detecting childhood cancers is problematic because of their rarity, non-specific presentations add an extra layer of diagnostic difficulty. A retrospective cohort study of 103 children with primary bone tumours found that GPs delayed referral longer than Accident and Emergency doctors (mean 1.8 versus 1.3 months). Longer delays overall were associated with age $\leqslant$12 and axial rather than limb tumours (mean 4.7 versus 1.4 months). Axial tumours presented with non-specific symptoms such as back and chest pain, which were misdiagnosed as 'muscle strain' or 'growing pains'.[21] A particular problem was that many patients were fit, young people active in sport, thus their symptoms were often misdiagnosed as injuries. A qualitative study selected a purposive sample of 20 parents with children diagnosed with cancer and interviewed them about the diagnostic process.[32] Ten parents had disputes with the GP or hospital doctor about the diagnosis and management of their child. Lengthy disputes were about delays in the investigation mostly of children who were eventually found to have solid tumours, e.g. Hodgkin's disease, Ewing's sarcoma, osteosarcoma. Fewer disputes had occurred about the diagnosis of leukaemia, which can be diagnosed with a simple blood test. The authors suggest that the rarity and non-specific presentation of childhood cancers make diagnosis difficult but that investigations are not always appropriately ordered and parental concerns are not taken seriously by health professionals.

Upper gastrointestinal cancer. A retrospective cohort study of 83 patients with newly diagnosed gastric cancer found that most had presented with dyspeptic symptoms, and in some cases, these symptoms had been present for over 2 years before diagnosis.[26] The median delay in diagnosis that was attributed to the GP (from presentation to referral) was 7 weeks (3–14), based on patient recall. The most common first symptom was simple epigastric pain, a common and non-specific symptom occurring in up to 40% of the population per year.[33] The study included patients diagnosed in 1981–1982, and it is possible that delay has

now decreased in response to the UK NICE guidelines on the prompt investigation of suspected cancer. These state that patients over 55 years of age with unexplained dyspeptic symptoms should have a gastroscopy within 2 weeks.[34] Nevertheless, suspecting and deciding to investigate a rare disease such as cancer in patients presenting with non-specific, common symptoms remain a diagnostic challenge in primary care.

A more robust and recent prospective cohort study of 7274 patients presenting with dyspepsia was conducted in Denmark.[16] GPs referred more patients with 'alarm symptoms' than simple dyspepsia (48% versus 32%). However, over 3 years of follow-up, the positive predictive value of typical 'alarm' symptoms such as weight loss, dysphagia, jaundice and anaemia was only 3%, while half of the diagnosed cancers occurred in patients without alarm symptoms. This further emphasizes the inherent difficulty in diagnosing gastrointestinal cancers in primary care. Most guidelines around the world simply rely on the increasing prevalence of cancer above age 55 to select patients with dyspepsia for endoscopic investigation.

Tremor in the elderly. A retrospective cohort study of 402 cases of presumed Parkinson's disease diagnosed in general practice (identified as taking anti-parkinsonian drugs but not for neuroleptic side effects) found that only 213 cases (53%) had definite diagnosis of Parkinson's disease based on application of gold-standard diagnostic criteria. Twenty-six per cent had alternative diagnoses including drug-induced parkinsonism, vascular pseudo-parkinsonism, dementia and essential tremor and the remaining 21% 'possible' Parkinson's.[12] The study did not address possible reasons for overdiagnosing Parkinson's disease; however, tremor is a feature of all these neurological conditions and typical of Parkinson's disease.

Malignant spinal cord compression. Malignant spinal cord compression (MSCC) has a substantial adverse effect on quality of life and survival, especially as studies indicate that most patients are diagnosed only after they develop a sensory level by which time the majority will be unable to walk. Between 2.5% and 5% of patients with terminal cancer have MSCC within the last 2 years of illness, and for some patients MSCC is their presenting symptom.[35] A retrospective study of 319 patients diagnosed with MSCC in Scotland showed that the mean delay from onset of symptoms to diagnosis was 3 months.[22] Two hundred and sixty-one patients agreed to be interviewed and 94% described the onset of pain as either localized back pain or nerve root pain, either thoracic or upper lumbar. It is unclear from the study whether the GPs had considered MSCC or whether they attributed the nerve root symptoms to a prolapsed intervertebral disc, for which urgent MRI is not usually available. Patients who were already known to have cancer were diagnosed more quickly (median 49 days versus 90 days, $P < 0.001$).

Iron deficiency anaemia (IDA). Iron deficiency anaemia (IDA) is not a disease entity in itself but an indicator of underlying pathology. It is fairly non-specific, potentially resulting from a broad range of conditions, including cancers. Four hundred and thirty-one patients with IDA were identified retrospectively from laboratory records and followed up 4 years later to determine eventual diagnosis and potential diagnostic delays. The study found that GPs were less likely to investigate female patients (OR 0.40, 95% CI 0.23–0.69), those under 65 (OR 0.48, 95% CI 0.27–0.85), those with a haemoglobin of more than 9 g/dl (OR 0.37, 95% CI 0.22–0.62) and those with a history of anaemia (OR 0.43, 95% CI 0.25–0.73).[13] At 1-year follow-up, 43% had had GI investigations, with serious pathology found in 30% of those investigated. Given that IDA is a common presenting symptom of upper and lower GI cancers, coeliac disease and inflammatory bowel disease, there is potential to miss serious disease. Indeed, having anaemia increased the risk of death three-fold and 12 patients who were not investigated initially died of advanced cancer. IDA has a higher incidence in women than men at all ages[36] but is commonest in menstruating women, where it may mask other important causes. The non-investigation of anaemia in these groups of patients could be attributed to not considering the possibility of serious pathology in younger patients; explaining away anaemia in women as a result of menstruation and not attaching significance to mild anaemia.

Very rare conditions

Disease prevalence is known to influence diagnostic decision making, with rare diseases less likely to be considered. Even when symptoms and signs are quite specific to a rare condition, there will be a high rate of false positives, and the positive predictive value of symptoms (the ratio of true positives to all those with the symptom) will be low. Only very diagnostic (pathognomonic) features can make a rare disease more likely. The review identified two rare conditions where the diagnosis can be missed: tongue cancer and retinoblastoma. In the case of tongue cancer, it was not clear how patients in the study presented and whether features of the presentation compounded the rarity of the disease and led to misdiagnosis. In the case of retinoblastoma, it was evident that the presentation also played an important role.

Tongue cancer. A Finnish study estimated the consultation prevalence of oral symptoms in primary care to be 0.55%. Approximately 1:2000 patients with oral symptoms have tongue cancer; therefore, the average GP could expect to see one case every 130 years.[18]

Despite collecting data on cases going back 20 years, only 75 cases were identified; twenty-six of these (35%) were not referred for specialist opinion at initial presentation and 14 of these were not followed up. The median professional delay (from initial consultation to final diagnosis on the basis of histology) was 0.7 months (0.1–18.2 months), though it is unclear to what extent this was due to GP delay and to what extent it was due to patients waiting for investigations. The longest and most fatal delays were experienced by patients who had not been followed up (median delay of 5.2 months, adjusted relative hazard of death 6.3, 95% CI 1.7–22.9). The grade of malignancy was significantly higher in those patients whose diagnoses were subject to delay. Most of the patients who consulted with a specific sign or pain in the tongue were referred (37/48, 77%). In contrast, only 40% (10/25) of the patients who consulted with an unrelated symptom or a non-specified oral symptom were referred.

Retinoblastoma. The parents of all patients with retinoblastoma, treated at a London hospital during 1993–1996 (100 patients), were interviewed about their child's diagnosis in primary care (100% response rate).[29] Fifty-four parents had consulted the GP first, the others consulting a health visitor or optician. Patient records were used in 90% of the cases to verify the date of the diagnosis, tumour laterality and treatment. The median time interval from GP presentation to first consultation with the ophthalmologist was 1 week (range 1–56). A quarter of patients (23/100) had experienced a delay in referral of more than 8 weeks. Children presenting with squint had significantly longer delays (median 7 weeks) than those presenting with other more typical symptoms and signs, such as leucocoria and change in eye appearance (median 1 week). This would suggest that the rarity of the condition was compounded by its presentation (atypical/non-specific), which increased diagnostic difficulty.

Co-morbidity

Co-morbidity is the presence of one or more disorders in addition to the disorder under investigation. In this case, one condition may mask another, i.e. it may either alter the presentation of another, or it may provide an obvious explanation for the presenting symptoms, thus, a second morbidity is not considered.

Cancers. A retrospective study of 56 cases of cancer diagnosed at a single practice asked the practice team to review and discuss eight cases with the longest diagnostic delay (from initial presentation to final diagnosis). The team discussed cases of one gastric, three colorectal, two ovarian, one prostate and one lung cancers, where delay ranged between 233 and 1283 days (though it is not clear to what extent this was due to patients waiting for investigations). Included among the patient, professional and task factors identified were the presence of complex multiple pathologies and failure to consider the diagnosis.[23] The study, however, did not systematically collect information about how the patients presented. It is only mentioned that 'in some cases, patients were relatively asymptomatic' (p. 301), which could suggest atypical and/or non-specific presentations.

Dementia and depression. The study of dementia mentioned under 'atypical presentations' above also found that individuals with unrecognized dementia were more likely to suffer from co-morbid conditions such as depression.[14] The study suggested that elderly patients who are diagnosed with depression are less likely to be assessed for dementia, because their symptoms are attributed to the depression. Detailed psychiatric assessment may be needed to make the distinction.

Perceptual features

Misdiagnosis can also result from a failure to detect and/or recognize visual or auditory symptoms and signs. The difficulty in noticing fairly typical visual features was investigated by Brooks *et al.*[37] They gave both medical students and clinicians pictures of patients copied mostly from textbook illustrations (e.g. jaundice for pancreatitis, a butterfly rash on the face for lupus erythematosus, swollen parotid glands for mumps). They found that clinicians were more accurate in their diagnoses than the students but still made mistakes. Importantly, the accuracy of both groups increased by similar amounts (around 20%) when a verbal description of the visual feature was also provided, i.e. when the feature was pointed out to them. Finally, clinicians were more likely to notice visual features when the correct diagnosis was suggested to them. The authors suggest that visual features that may seem quite obvious in hindsight, i.e. after the diagnosis is known, may be seen as normal variation and go unnoticed. Furthermore, features may not be noticed and recognized if the correct diagnosis is not being considered.

Human Immunodeficiency virus. A US study that used real HIV patients as standardized patients found that only a quarter of the 134 participating primary care physicians both detected and correctly diagnosed prominent lesions suggestive of HIV infection (Kaposi's sarcoma and oral hairy leukoplakia).[28] Approximately half of the participants did not detect the lesions, even though the standardized patients drew their attention to the anatomic areas where the lesions were located. Even fewer participants (17%) detected lymphadenopathy in a patient with prominent diffuse lymphadenopathy and a history of persistent fatigue, fever and arthralgias. Correct diagnosis (including

detection) of cardiac murmurs related to aortic insufficiency in two of the standardized patients was more accurate (56–58% depending on patient, 38% for doctors who saw both patients). Experience with treating HIV patients increased identification only of oral hairy leukoplakia.

Discussion

This systematic review identified a range of conditions in primary care where diagnostic error or delay has been investigated. In addition to cancers, ischaemic heart disease and meningitis that have also been identified by other methods (GP self-reports of memorable errors, litigation databases), there is evidence for non-investigation of IDA and underdiagnosis of dementia in certain groups of patients. Some conditions, especially cancers, were extremely rare, suggesting a specialist research interest rather than an inappropriate rate of misdiagnosis. The studies reviewed are best considered as a series of '*ad hoc*' investigations that do not chart the profile of diagnostic error in primary care. Furthermore, there is no systematic comparison between different clinical domains in the literature.

Cancers were the most frequently identified conditions, being discussed in 11 papers. Cancers are missed because they are rare (e.g. tongue cancer, retinoblastoma, childhood cancers, MSCC) and/or present atypically (e.g. breast cancer presenting without a breast lump) or with features that are common to other, less serious conditions (e.g. upper GI cancer presenting with dyspeptic symptoms). Other conditions, such as MI and meningitis, may also get missed for similar reasons, i.e. the way that they present.

A psychological process that could explain many of the misdiagnoses is the misattribution of the presenting symptoms to an obvious explanation or a readily available aetiology, especially in the absence of typical features. This is known as 'diagnostic overshadowing' in the mental health and intellectual disability literatures, where an existing mental illness 'overshadows' and leads to the underdiagnosis of other co-morbidities.[38] An existing diagnostic label can hamper the clinician's ability to restructure the diagnostic problem and look for alternative explanations. The presence of a co-morbidity is not necessary for misattribution. In fact, co-morbidity appeared only in two papers as a potential factor in the diagnostic error.[14,23] Symptoms were misattributed to an obvious aetiology or a readily available explanation in at least three other studies: delays in diagnosing MI in female patients (influenced by the patients' non-cardiac explanations),[20] missing bone cancers in children by attributing fairly non-specific symptoms, like back or chest pains, to growth or sports injuries[21] and not investigating anaemia in young women by explaining it away as a result of menstruation.[13]

Using an accepted quality assessment tool,[11] only one of the 21 studies reviewed was assessed as being of high quality. Most studies were retrospective and suffered from potentially biased and incomplete information. For example, symptoms at first presentation to GP and estimates of delay were based on patient recall or hospital records. The reasons for the delay were often not investigated and could not be extrapolated from the reported results. Due to insufficient detail about the specifics of individual consultations, the studies cannot be used to make inferences about preventability, i.e. whether the delay could have been avoided and do not constitute evidence that the GPs' diagnostic skills were inadequate. Furthermore, studies originating in small cohorts of patients with a rare diagnosis recruited in specialist centres shed little light on how to diagnose rare conditions in primary care, as they do not include patients with false-positive diagnoses who may have unnecessary referrals and investigations.

The studies can, however, be used to hypothesize about difficulties in the nature of diagnosis in primary care. In the range of conditions described, presentation was clearly influential on the accuracy of the GP's differential diagnosis, and consequently, the promptness of referral. Difficult presentations are likely to get missed, so the question is how to support performance in those cases, without increasing investigation and referral rates unnecessarily. One of the reviewed studies suggested the need to take parental concerns seriously,[32] another indicated the need for follow-up.[18] A US study of closed malpractice claims (patients alleging missed or delayed diagnosis) in the ambulatory setting found that failure to gather sufficient and appropriate information was responsible for most errors.[9] Appropriate information gathering was found to be associated with diagnostic accuracy in a recent study by our group. We found that GPs were more likely to diagnose difficult cases correctly, if they gathered more 'critical' information.[39] This was defined as information with diagnostic value for any differential diagnosis that was relevant to the problem at hand. Therefore, what seems important is formulating an appropriate set of diagnostic hypotheses and selecting the right questions to test them. Ongoing, in-depth analyses of misdiagnosed cases suggest that hypothesis generation is the key. In the majority of misdiagnosed cases, the correct hypothesis had not been considered. In the absence of the correct hypothesis, appropriate information was either not gathered or was dismissed.[40]

Clinicians can quickly recognize what is wrong with a patient from just a few features, very early on in the consultation or entertain a small number of hypotheses (2–4).[41] With experience, they solve familiar problems with speed and accuracy, through instant

recognition of patterns, matching of prototypes or recall of similar patients.[42,43] This is a kind of 'satisficing' behaviour that does not involve meticulous consideration of alternative hypotheses and exhaustive information gathering. It thus reduces time, risk and cost from unnecessary investigations and referrals, while leading to the correct diagnosis and management most of the time. However, difficult presentations and rare diseases such as those identified in this review are likely to be misdiagnosed with the usual satisficing approach. The challenge for research is therefore to understand how to support diagnosis of difficult cases, without seriously obstructing, continuously interrupting or impractically lengthening the natural reasoning process.

Declaration

Funding: The Patient Safety Research Programme (Department of Health) (PS/027). The views and opinions expressed herein do not necessarily reflect those of the Secretary of State for Health.
Ethical approval: None.
Conflicts of interest: None.

References

[1] Department of Health. *An Organisation with a Memory.* London: The Stationery Office, 2000.
[2] Shaw R, Drever F, Hughes H, Osborn S, Williams S. Adverse events and near miss reporting in the NHS. *Qual Saf Health Care* 2005; **14:** 279–283.
[3] Kostopoulou O. Do GPs report diagnostic errors? *Fam Pract* 2008; **25:** 1–2.
[4] Bhasale A. The wrong diagnosis: identifying causes of potentially adverse events in general practice using incident monitoring. *Fam Pract* 1998; **15:** 308–318.
[5] Silk N. What went wrong in 1000 negligence claims. *Health Care Risk Rep* 2000; 13–16.
[6] Phillips RL Jr, Bartholomew LA, Dovey SM, Fryer GE Jr, Miyoshi TJ, Green LA. Learning from malpractice claims about negligent, adverse events in primary care in the United States. *Qual Saf Health Care* 2004; **13:** 121–126.
[7] Fisseni G, Pentzek M, Abholz HH. Responding to serious medical error in general practice—consequences for the GPs involved: analysis of 75 cases from Germany. *Fam Pract* 2008; **25:** 9–13.
[8] Ely JW, Levinson W, Elder NC, Mainous AG, Vinson DC. Perceived causes of family physicians' errors. *J Fam Pract* 1995; **40:** 337–344.
[9] Gandhi TK, Kachalia A, Thomas EJ *et al*. Missed and delayed diagnoses in the ambulatory setting: a study of closed malpractice claims. *Ann Intern Med* 2006; **145:** 488–496.
[10] Graber ML, Franklin N, Gordon R. Diagnostic errors in internal medicine. *Arch Intern Med* 2005; **165:** 1493–1499.
[11] GRADE Working Group. Grading quality of evidence and strength of recommendations. *BMJ* 2004; **328:** 1490.
[12] Meara J, Bhowmick BK, Hobson P. Accuracy of diagnosis in patients with presumed Parkinson's disease. *Age Ageing* 1999; **28:** 99–102.
[13] Yates JM, Logan EC, Stewart RM. Iron deficiency anaemia in general practice: clinical outcomes over three years and factors influencing diagnostic investigations. *Postgrad Med J* 2004; **80:** 405–410.
[14] Lopponen M, Raiha I, Isoaho R, Vahlberg T, Kivela SL. Diagnosing cognitive impairment and dementia in primary health care—a more active approach is needed. *Age Ageing* 2003; **32:** 606–612.
[15] van Schayck CP, Der Heijden FM, Den Boom G, Tirimanna PR, van Herwaarden CL. Underdiagnosis of asthma: is the doctor or the patient to blame? The DIMCA project. *Thorax* 2000; **55:** 562–565.
[16] Meineche-Schmidt V, Jorgensen T. Alarm symptoms' in patients with dyspepsia: a three-year prospective study from general practice. *Scand J Gastroenterol* 2002; **37:** 999–1007.
[17] Burgess CC, Ramirez AJ, Richards MA, Love SB. Who and what influences delayed presentation in breast cancer? *Br J Cancer* 1998; **77:** 1343–1348.
[18] Kantola S, Jokinen K, Hyrynkangas K, Mantyselka P, Alho OP. Detection of tongue cancer in primary care. *Br J Gen Pract* 2001; **51:** 106–111.
[19] Bleeker JK, Simoons ML, Erdman RA *et al*. Patient and doctor delay in acute myocardial infarction: a study in Rotterdam, The Netherlands. *Br J Gen Pract* 1995; **45:** 181–184.
[20] Bouma J, Broer J, Bleeker J, van Sonderen E, Meyboom-de Jong B, DeJongste MJ. Longer pre-hospital delay in acute myocardial infarction in women because of longer doctor decision time. *J Epidemiol Community Health* 1999; **53:** 459–464.
[21] Goyal S, Roscoe J, Ryder WD, Gattamaneni HR, Eden TO. Symptom interval in young people with bone cancer. *Eur J Cancer* 2004; **40:** 2280–2286.
[22] Levack P, Graham J, Collie D *et al*. Don't wait for a sensory level—listen to the symptoms: a prospective audit of the delays in diagnosis of malignant cord compression. *Clin Oncol (R Coll Radiol)* 2002; **14:** 472–480.
[23] Jiwa M, Reid J, Handley C *et al*. Less haste more speed: factors that prolong the interval from presentation to diagnosis in some cancers. *Fam Pract* 2004; **21:** 299–303.
[24] Vasudev NS, Joffe JK, Cooke C, Richards F, Jones WG. Delay in the diagnosis of testicular tumours—changes over the past 18 years. *Br J Gen Pract* 2004; **54:** 595–597.
[25] Schnetler JFC. Oral cancer diagnosis and delays in referral. *Br J Oral Maxillofac Surg* 1992; **30:** 210–213.
[26] Mikulin T, Hardcastle JD. Gastric cancer—delay in diagnosis and its causes. *Eur J Cancer Clin Oncol* 1987; **23:** 1683–1690.
[27] Sorensen HT, Moller-Petersen J, Krarup HB, Pedersen H, Hansen H, Hamburger H. Diagnostic problems with meningococcal disease in general practice. *J Clin Epidemiol* 1992; **45:** 1289–1293.
[28] Paauw DS, Wenrich MD, Curtis JR, Carline JD, Ramsey PG. Ability of primary care physicians to recognize physical findings associated with HIV infection. *JAMA* 1995; **274:** 1380–1382.
[29] Goddard AG, Kingston JE, Hungerford JL. Delay in diagnosis of retinoblastoma: risk factors and treatment outcome. *Br J Ophthalmol* 1999; **83:** 1320–1323.
[30] Charlton I, Jones K, Bain J. Delay in diagnosis of childhood asthma and its influence on respiratory consultation rates. *Arch Dis Child* 1991; **66:** 633–635.
[31] Jones A, Sykes A. The effect of symptom presentation on delay in asthma diagnosis in children in a general practice. *Respir Med* 1990; **84:** 139–142.
[32] Dixon-Woods M, Findlay M, Young B, Cox H, Heney D. Parents' accounts of obtaining a diagnosis of childhood cancer. *Lancet* 2001; **357:** 670–674.
[33] Jones RH, Lydeard SE, Hobbs FD *et al*. Dyspepsia in England and Scotland. *Gut* 1990; **31:** 401–405.
[34] NICE. *The Management of Dyspepsia in Adult Patients in Primary Care.* London: National Institute of Clinical Excellence, 2004.
[35] Prasad D, Schiff D. Malignant spinal-cord compression. *Lancet Oncol* 2005; **6:** 15–24.
[36] McCormick A, Fleming D, Charlton J. *Morbidity Statistics from General Practice: Fourth National Study 1991–1992.* London: HMSO, 1995.
[37] Brooks LR, LeBlanc VR, Norman GR. On the difficulty of noticing obvious features in patient appearance. *Psychol Sci* 2000; **11:** 112–117.

[38] Jopp DA, Keys CB. Diagnostic overshadowing reviewed and reconsidered. *Am J Ment Retard* 2001; **106:** 416–433.
[39] Kostopoulou O, Oudhoff J, Nath R *et al.* Predictors of diagnostic accuracy and safe management in difficult diagnostic problems in family medicine. *Med Decis Making* 2008; **28:** 668–680.
[40] Kostopoulou O, Devereaux-Walsh C, Delaney BC. Missing celiac disease in family medicine: the importance of hypothesis generation. *Med Decis Making* (in press).
[41] Elstein AS, Shulman LS, Sprafka SA. *Medical Problem Solving: An Analysis of Clinical Reasoning.* Cambridge, MA: Harvard University Press, 1978.
[42] Grant J, Marsden P. The structure of memorized knowledge in students and clinicians: an explanation for diagnostic expertise. *Med Educ* 1987; **21:** 92–98.
[43] Schmidt HG, Norman GR, Boshuizen HPA. A cognitive perspective on medical expertise: theory and implications. *Acad Med* 1990; **65:** 611–621.