John Busby, Knut Schroeder, Wolfram Woltersdorf, Jonathan AC Sterne, Yoav Ben-Shlomo, Alastair Hay and William Hollingworth

Temporal growth and geographic variation in the use of laboratory tests by NHS general practices:

using routine data to identify research priorities

Abstract

Background
Laboratory tests are extensively used for diagnosis and monitoring in UK primary care. Test usage by GPs, and associated costs, have grown substantially in recent years.

Aim
This study aimed to quantify temporal growth and geographic variation in utilisation of laboratory tests.

Design and setting
Retrospective cohort study using data from general practices in the UK.

Method
Data from the General Practice Research Database, including patient demographics, clinical details, and laboratory test results, were used to estimate rates of change in utilisation between 2005 and 2009, and identify tests with greatest inter-regional variation, by fitting random-effects Poisson regression models. The study also investigated indications for test requests, using diagnoses and symptoms recorded in the 2 weeks before each test.

Results
Around 660 000 tests were recorded in 230 000 person-years of follow-up. Test use increased by 24.2%, from 23 872 to 29 644 tests per 10 000 person-years, between 2005 and 2009. Tests with the largest increases were faecal occult blood (121%) and C-reactive protein (86%). There was substantial geographic variation in test utilisation; GPs in some regions requested tests such as plasma viscosity and cardiac enzymes at a rate more than three times the national average.

Conclusion
Increases in the use of laboratory tests have substantial resource implications. Rapid increases in particular tests may be supported by evidence-based guidelines, but these are often vague about who should be tested, how often, and for how long. Substantial regional variation in test use may reflect uncertainty about diagnostic accuracy and appropriate indications for the laboratory test. There is a need for further research on the diagnostic accuracy, therapeutic impact, and effect on patient health outcomes of the most rapidly increasing and geographically variable tests.

Keywords
clinical laboratory techniques; economics; evidence-based medicine; physician practice patterns; primary health care; trends

INTRODUCTION

Laboratory tests are extensively used to make or support diagnoses and for chronic disease monitoring and management. In the UK, NHS laboratory services process in excess of 700 million requests annually, costing around £2.5 billion, and influence over 70% of clinical decisions.[1] More than one-third of requests for laboratory tests are initiated in primary care.[1] Testing is expensive for laboratories and practices, as it incurs substantial primary care staff and patient costs per specimen due to sampling, transport, result interpretation, and result communication.

Test use by GPs has seen particular growth recently.[2] In part, this is due to the increasing use of point-of-care testing and the introduction of the Quality and Outcomes Framework (QOF), which offers financial rewards to GPs for more intensive monitoring of patients.[3] Requests by GPs for biochemistry tests increased, on average, by more than 20% between 2002/2003 and 2005, with growth in demand predicted to continue.[1] Concerns that many laboratory tests requested by GPs are unnecessary[4,5] may be supported by considerable regional variation, both in the UK and worldwide.[6–11] Wennberg's 'professional uncertainty hypothesis' suggests that some geographic variation occurs because of regional differences in physicians', or a regional opinion leader's, beliefs about the value of the test or treatment, rather than geographic variations in patients' need.[12] Inappropriate test use has also been attributed to factors such as patient pressure and the use of 'defensive' diagnostic testing to reduce malpractice liability.[13,14] As budgets tighten and UK GPs become responsible for commissioning decisions, and ultimately laboratory budgets, there will be increased incentives for more cost-effective use of laboratory tests.

The aim of this study was to identify the laboratory tests with the largest temporal growth and geographic variation in utilisation, using NHS GP records of use of routine laboratory tests requests between 2005 and 2009. The evidence supporting rapidly increasing use of specific tests is discussed, and possible causes of geographic variation are explored.

METHOD

Data from the General Practice Research Database (GPRD) were used; the GPRD holds approximately 5 million active patient records from 636 general practices across the UK. Information collected includes practice characteristics, patient demographics, records of primary care visits including symptoms and diagnoses, and data on laboratory test requests and results. The GPRD has an in-house data

J Busby, MSc, research assistant in health economics; **K Schroeder**, MSc, PhD, MD, DRCOG, DCH, MRCGP, MRCP, GP and honorary senior clinical lecturer; **JAC Sterne**, BA, MSc, PhD, professor of medical statistics and epidemiology and head of school; **Y Ben-Shlomo**, MSc, PhD, MRCP, professor of clinical epidemiology; **A Hay**, MD, MRCP, DCH, MRCGP, DFFP, FHEA; GP and reader in primary health care; **W Hollingworth**, BSocSc, MSc, PhD, professor of health economics; School of Social and Community Medicine, University of Bristol, Bristol. **W Woltersdorf**, MD, MRCP, FRCPath, consultant chemical pathologist, Department of Laboratory Medicine, University Hospitals Bristol NHS Foundation Trust, Bristol.

Address for correspondence
Will Hollingworth, School of Social and Community Medicine, Canynge Hall, 39 Whatley Road, Bristol, BS8 2PS.
E-mail: William.Hollingworth@bristol.ac.uk

Submitted: 5 January 2012; **Editor's response:** 28 May 2012; **final acceptance:** 13 September 2012
This is the full-length article (published online 25 Mar 2013) of an abridged version published in print. Cite this article as: **Br J Gen Pract 2013; DOI: 10.3399/bjgp13X665224**

How this fits in

Test utilisation by UK GPs is thought to have grown rapidly in recent years with concerns that some tests are unnecessary. This study confirms these beliefs, with an overall increase in the utilisation of laboratory tests of 24.2% between 2005 and 2009. It also strengthens previous evidence indicating substantial geographic variation in laboratory test use.

quality-assurance mechanism to ensure practices are 'up to research standard'.

Patients can have several registration periods if they move between GPRD practices. When this occurred, only data from the last registration period were used. A random sample was drawn of 100 000 patients registered with a GPRD general practice at any time from 1 January 1999 to 31 December 2008. Patients were excluded from the initial cohort if they left the practice or died before 1 January 2005 (n = 25 233), or left the practice or died before the practice was deemed 'up to research standard' (n = 924) from the initial cohort, leaving 73 843 patients for analysis.

Table 1. Cohort characteristics

Characteristic	2005	2006	2007	2008	2009
Number of practices	444	451	455	461	467
Region, *n* (%)					
North West	66 (14.9)	67 (14.9)	67 (14.7)	67 (14.5)	68 (14.6)
London	57 (12.8)	59 (13.1)	60 (13.2)	61 (13.2)	63 (13.5)
South West	42 (9.5)	43 (9.5)	43 (9.5)	43 (9.3)	43 (9.2)
West Midlands	41 (9.2)	41 (9.1)	43 (9.5)	43 (9.3)	44 (9.4)
South Central	41 (9.2)	41 (9.1)	41 (9.0)	43 (9.3)	43 (9.2)
South East Coast	41 (9.2)	41 (9.1)	41 (9.0)	41 (8.9)	41 (8.8)
Scotland	36 (8.1)	39 (8.6)	40 (8.8)	40 (8.7)	42 (9.0)
East of England	34 (7.7)	34 (7.5)	34 (7.5)	35 (7.6)	35 (7.5)
Wales	30 (6.8)	30 (6.7)	30 (6.6)	31 (6.7)	31 (6.6)
Northern Ireland	17 (3.8)	17 (3.8)	17 (3.7)	17 (3.7)	17 (3.6)
Yorkshire and Humber	16 (3.6)	16 (3.5)	16 (3.5)	16 (3.5)	16 (3.4)
East Midlands	15 (3.4)	15 (3.3)	15 (3.3)	15 (3.3)	15 (3.2)
North East	8 (1.8)	8 (1.8)	8 (1.8)	9 (2.0)	9 (1.9)
Deprivation quintile, *n* (%)					
0 (least deprived)	82 (18.5)	82 (18.2)	83 (18.2)	84 (18.2)	84 (18.0)
1	87 (19.6)	91 (20.2)	91 (20.0)	92 (20.0)	94 (20.1)
2	92 (20.7)	94 (20.8)	95 (20.9)	95 (20.6)	96 (20.6)
3	87 (19.6)	87 (19.3)	89 (19.6)	91 (19.7)	91 (19.5)
4 (most deprived)	96 (21.6)	97 (21.5)	97 (21.3)	99 (21.5)	102 (21.8)
Number of patients	38 579	41 648	44 696	47 860	50 679
Mean age (SD), years	35.8 (22.5)	35.8 (22.5)	35.8 (22.5)	35.9 (22.6)	35.9 (22.6)
Males, *n* (%)	18 994 (49.2)	20 612 (49.5)	22 167 (49.6)	23 767 (49.7)	25 135 (49.6)
Number of tests	99 569	116 466	133 309	149 446	161 838

Percentages do not add to 100 due to rounding.

This was an 'open' cohort of variable size during the study period, because of births, deaths, and transfers.

Test results were recorded separately in the GPRD, even if requested as part of a panel, leading to multiple entries for the same test request. For example, when a full blood count was requested, red blood cell, white blood cell, and platelet counts were recorded separately. Results corresponding to the same test panel were regrouped and recorded as one test request (Appendix 1). Although the GPRD collects information solely from primary care, the results of tests carried out in secondary care and returned to the GP might be recorded within the GPRD. Therefore, tests that are unlikely to be requested by a GP, and some miscellaneous test groups, were excluded from the analyses (Appendix 2).

Test-utilisation rates were calculated by dividing the number of test results by the total person-years of follow-up. For each test, age–sex-standardised annual utilisation rates were calculated using direct standardisation with the population of England as the standard population. Standardisation ensures that observed differences in test rates are not simply due to differences in the age or sex of the population over time or between practices.[15] Rates were expressed as tests per 10 000 person-years. To focus on the most common and economically important tests, the analysis was limited to laboratory tests with an average annual utilisation of more than 50 per 10 000 person-years. Tests with the largest increase in utilisation, in both relative and absolute terms, were identified.

Temporal changes and geographic differences in the demographics and frequency of patients undergoing testing were explored, and reasons for test use were identified, using information on symptoms and diagnoses, which are recorded in the GPRD using Read Codes. Because the Read Code system is complex, with numerous ways to code the same event,[16] Read Code groupings (RCGs) supplied by the Information Services Division, NHS National Services Scotland were used.[17] RCGs use headings based on the chapters of the International Classification of Diseases (ICD-10), to categorise around 80 000 Read Codes into around 260 more homogeneous groups. Probable reasons for test use were evaluated by examining the frequency of RCGs (including both new and chronic symptoms and diagnoses) recorded in the 14 days before the laboratory test. For each test, a rate ratio was calculated, comparing the relative frequency of RCGs recorded in

patients who had that test (cases) compared with a cohort of patients matched on sex, age, and geographic region (controls), who underwent a different laboratory test in the same quarter of the year as the case. When possible, 10 controls were sampled for each case. Symptoms and diagnoses most strongly associated with the ordering of the test should lead to the largest rate ratios. Hand filtering was used to remove RCGs relating to administrative events (for example, patient reviewed) or related to the test itself (for example, urine sample taken). To focus on the most relevant symptoms and diagnoses, rate ratios were only calculated for those that were recorded in more than 50 patients in the 2 weeks preceding the test.

Geographic variation in the average rate of test utilisation from 2005 to 2009 among 13 UK regions was explored using Poisson regression models, allowing for between-region variation in test utilisation. Calculating variance estimates from random effects Poisson regressions rather than directly from the observed rates is a more statistically rigorous approach, which appropriately adjusts for chance variability. The most variable tests are those with the largest between-region standard deviation. Analyses were conducted using Stata (version 11).

Table 2. Temporal and geographic variation in laboratory tests, tests per 10 000 person-years (age–sex standardised)

Test	2005	2006	2007	2008	2009	Relative % increase (rank)	Absolute increase (rank)	Between-region SD (rank)
Faecal occult blood	42	40	64	91	93	121.3 (1)	51.1 (16)	0.38 (13)
C-reactive protein	440	503	589	708	817	85.8 (2)	377.2 (5)	0.34 (14)
Haematinics	326	387	445	473	571	75.0 (3)	244.9 (11)	0.22 (18)
Immunoglobulins	61	77	95	94	106	73.4 (4)	45.1 (17)	0.31 (15)
Serum iron studies	42	51	59	74	72	72.2 (5)	30.3 (21)	0.87 (6)
Urine biochemistry	336	378	406	445	560	66.9 (6)	224.5 (12)	0.19 (22)
Blood trace elements/vitamins	70	75	81	86	112	60.4 (7)	42.0 (18)	1.25 (3)
Bone profile	756	898	1019	1092	1150	52.0 (8)	393.3 (4)	0.59 (7)
Clotting tests	766	963	1062	1126	1112	45.1 (9)	345.3 (7)	0.56 (8)
Prolactin level	41	50	52	51	55	36.0 (10)	14.6 (22)	0.44 (10)
Prostate-specific antigen	220	224	253	261	295	34.2 (11)	75.3 (15)	0.25 (16)
Plasma viscosity	129	132	134	170	170	31.4 (12)	40.5 (19)	3.14 (1)
Full blood count	2898	3106	3306	3545	3642	25.7 (13)	743.7 (2)	0.13 (27)
Liver function tests	2717	3030	3165	3293	3405	25.3 (14)	688.1 (3)	0.14 (25)
Erythrocyte sedimentation rate	779	815	896	930	968	24.2 (15)	188.7 (13)	0.40 (11)
Urea and electrolytes	3212	3525	3750	3862	3987	24.1 (16)	774.5 (1)	0.11 (28)
Urine MCS	1631	1741	1886	1921	2000	22.6 (17)	368.3 (6)	0.19 (20)
Sex hormones	140	157	164	163	171	21.9 (18)	30.7 (20)	0.19 (21)
Thyroid function tests	1843	1917	1939	2103	2180	18.3 (19)	337.0 (8)	0.10 (29)
Glucose	2069	2167	2184	2263	2356	13.9 (20)	287.2 (9)	0.16 (23)
Rheumatoid factor	89	88	93	111	100	12.4 (21)	11.1 (24)	0.21 (19)
Lipid profile	2082	2251	2260	2286	2341	12.4 (22)	259.1 (10)	0.14 (26)
HbA1c — diabetes control	702	722	739	753	787	12.0 (23)	84.5 (14)	0.15 (24)
Creatine phosphokinase level	173	188	222	197	185	6.8 (24)	11.8 (23)	0.93 (4)
Film report	63	64	66	62	65	3.3 (25)	2.1 (25)	0.89 (5)
Vaginal swab	106	107	95	98	106	–0.4 (26)	–0.4 (26)	0.38 (12)
Urine dipstick	1021	1100	1063	1091	995	–2.6 (27)	–26.8 (29)	0.23 (17)
Cardiac enzymes	82	78	91	82	76	–8.0 (28)	–6.6 (27)	2.01 (2)
Pregnancy test	74	69	67	62	57	–23.7 (29)	–17.6 (28)	0.48 (9)
Average of all tests[a]	23 872	25 889	27 231	28 578	29 644	24.2	5772	—

HbA1c = glycosylated haemoglobin. MCS = microscopy, culture and sensitivities. SD = standard deviation. [a]Average of all tests, including those that are utilised less than 50 times per 100 000 person-years.

RESULTS

Table 1 shows that cohort demographics (age, sex, and practice location) were stable during the study period, although fewer practices in the North East, East Midlands and Yorkshire and Humber regions contributed data to the GPRD than in other areas. Around 660 000 tests were recorded in 230 000 person-years. The number of practices and patients, and tests ordered, increased substantially throughout the study period.

Table 2 shows the annual utilisation rate for each test, together with the relative and absolute utilisation increases, and the between-region standard deviation. Test use increased by an average of 24.2% during the study period, from 23 872 to

Table 3. Details of test use for relatively increasing tests

Name of test	Mean age (SD), years		Male, *n* (%)		Frequent diagnoses/symptoms	Incidence rate ratio (95% CI)[a]	Tests per patient[b] (SD)	
	2005	2009	2005	2009			2005	2009
Faecal occult blood	62.67 (17.37)	63.52 (13.17)	60 (41.38)	170 (43.04)	Number (%) with any diagnosis/symptom recorded in the 2 weeks pre-test = 671 (51.42)		1.56 (0.84)	1.14 (0.46)[c]
					Iron deficiency anaemia (*n* = 52)	23.54 (14.30 to 38.75)		
					Activities related to digestive/abdominal S&S (*n* = 144)	14.06 (10.91 to 18.12)		
					Digestive/abdominal S&S (*n* = 266)	3.46 (3.01 to 3.98)		
C-reactive protein	52.44 (19.96)	53.70 (19.72)[c]	566 (35.22)	1300 (34.23)	Number (%) with any diagnosis/symptom recorded in the 2 weeks pre-test = 6342 (48.77)		1.31 (1.09)	1.43 (1.41)[d]
					Rheumatoid arthritis and systemic connective tissue disorders (*n* = 320)	4.35 (3.82 to 4.96)		
					Activities related to neurological/musculoskeletal S&S (*n* = 131)	3.95 (3.23 to 4.84)		
					Joint disorders w/e (*n* = 476)	3.12 (2.82 to 3.46)		
					Soft tissue disorders (*n* = 663)	2.28 (2.09 to 2.48)		
					Diseases of intestines and peritoneum w/e (*n* = 193)	2.25 (1.92 to 2.63)		
Haematinics	53.70 (21.53)	53.08 (21.12)	325 (27.11)	836 (30.85)[c]	Number (%) with any diagnosis/symptom recorded in the 2 weeks pre-test = 5062 (53.18)		1.21 (0.54)	1.23 (0.56)
					Anaemias excluding iron deficiency (*n* = 117)	7.79 (6.12 to 9.92)		
					Iron-deficiency anaemia (*n* = 125)	6.05 (4.84 to 7.56)		
					Diseases of the oral cavity, salivary glands and jaws w/e (*n* = 71)	5.03 (3.78 to 6.69)		
Immunoglobulins	40.07 (19.06)	41.98 (18.97)	48 (18.18)	131 (23.48)	Number (%) with any diagnosis/symptom recorded in the 2 weeks pre-test (%) = 1359 (64.47)	1.55 (1.18 to 2.04)	1.05 (0.29)	1.08 (0.37)
					Skin S&S (*n* = 59)			
Serum iron studies	54.22 (23.60)	53.13 (21.67)	42 (27.45)	94 (27.57)	Number (%) with any diagnosis/symptom recorded in the 2 weeks pre-test (%) = 771 (59.81)		1.17 (0.50)	1.12 (0.39)
					Fatigue, tiredness, malaise and dizziness (*n* = 82)	1.78 (1.41 to 2.24)		
					Diseases of the skin and subcutaneous tissue w/e (*n* = 62)	1.74 (1.33 to 2.26)		

n = number of times the symptoms was recorded in the 2 weeks prior to the test in the patients receiving the test. SD = standard deviation. S&S = signs and symptoms. w/e = with specific exclusions. [a]Comparing the relative frequency of Read Code groups recorded in patients who had that test (cases) compared with a cohort of patients matched on sex, age, and geographic region (controls) who underwent a different laboratory test in the same quarter of the year as the case. [b]Average number of tests per patient per year, given that they had at least one test in that year. [c] Significant at the 0.1% level when fitting a Poisson model (tests per patient), linear model (age); logistic model (% male) with year as covariate. [d] Significant at the 1% level.

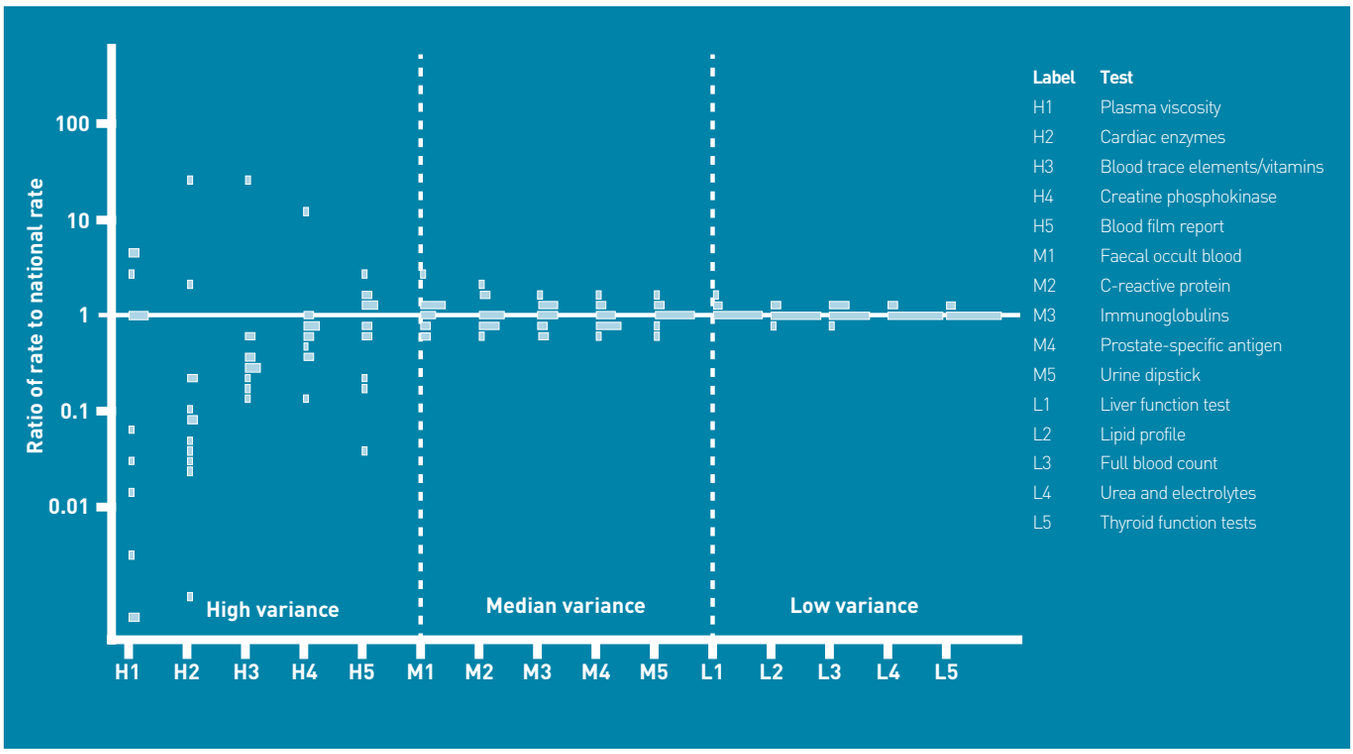

Figure 1. Regional variation in test-ordering behaviour, stratified by high, median, and low variance. For H1, H2, no procedures were observed in some regions. The lowest bar for this procedure refers to a notional rate of 0.1 procedures per 10 000 person-years.

29 644 per 10 000 person-years. Between 2005 and 2009, several tests, including faecal occult blood, C-reactive protein, and haematinics, had relative increases of more than 75%. The largest absolute increases, in excess of 680 tests per 10 000 person-years, were observed for creatinine, urea and electrolytes, full blood count, and liver function test. Few tests decreased in utilisation, and the magnitude of decreases was modest in both absolute and relative terms.

Table 3 displays a more detailed analysis of testing patterns and reasons for testing, for the five tests with the greatest relative increases in utilisation. Increased utilisation appeared to result mainly from testing more patients, rather than from testing the same patients more frequently. C-reactive protein was the only test for which there was a clear increase in tests per patient tested between 2005 and 2009 (1.31 versus 1.43; $P = 0.002$). Symptoms and diagnoses recorded relatively more frequently in the 14 days before the test request corresponded to the expected indications for these tests, but in a high proportion of cases no diagnoses or symptoms were recorded.

There was substantial geographic variation in the use of some tests (Figure 1). For example, the average rate of plasma viscosity testing varied widely between regions; no testing was recorded in two regions (East of England, South East Coast), while 770 tests per 10 000 person-years were recorded in the South West region. In some regions, rates of plasma viscosity, cardiac enzymes, and blood trace elements/vitamins testing were more than three times the national average. The test with least geographic variability, thyroid function, had a utilisation rate varying from 1935 tests per 10 000 person-years in the East Midlands to 2845 per 10 000 person-years in Northern Ireland.

Geographic variation may be partly due to more frequent testing of the same patients (Table 4). Patients in regions of high use were tested more frequently for plasma viscosity (1.60 versus 1.25 tests per person per year; $P<0.001$) and cardiac enzymes (1.48 versus 1.04 tests per person per year; $P = 0.01$) than patients in areas of low use. Other geographically variable tests showed trends in the same direction. There were minor differences in the age profile of patients tested in regions of high use; for example, patients for whom blood trace elements/vitamins and clotting tests were ordered were older in high utilisation regions.

DISCUSSION

Summary

Between 2005 and 2009, utilisation of laboratory tests in primary care increased by 24.2%, from 23 872 to 29 644 tests per

Table 4. Details of test use for geographically variable tests

Name of test	Mean age (SD), years		Male, *n* (%)		Frequent diagnoses/symptoms	Incidence rate ratio (95% CI)[a]	Tests per patient[b] (SD)	
	High use	Low use	High use	Low use			2005	2009
Plasma viscosity	55.31 (18.49)	53.65 (17.38)	1018 (36.40)	103 (40.87)	Number with any diagnosis/symptom recorded in the 2 weeks pre-test = 1323 (43.42)		1.60 (2.01)	1.25 (0.83)[c]
					Rheumatoid arthritis, systemic connective tissue disorders (*n* = 62)	3.28 (2.46 to 4.37)		
					Back and neck disorders (*n* = 60)	2.16 (1.63 to 2.86)		
					Joint disorders w/e (*n* = 72)	1.91 (1.49 to 2.46)		
					Neurological/musculoskeletal S&S (*n* = 177)	1.80 (1.53 to 2.11)		
Cardiac enzymes	57.35 (19.33)	54.28 (16.40)	653 (40.48)	26 (53.06)	Number with any diagnosis/symptom recorded in the 2 weeks pre-test = 600 (36.12)		1.48 (1.07)	1.02 (0.14)[d]
Blood trace elements/vitamins	57.03 (20.17)	52.04 (21.98)[c]	506 (34.99)	112 (35.90)	Number with any diagnosis/symptom recorded in the 2 weeks pre-test = 731 (41.58)		1.49 (1.03)	1.15 (0.41)[c]
					Neurological/musculoskeletal S&S (*n* = 101)	1.55 (1.25 to 1.91)		
Creatine phosphokinase	59.58 (17.54)	61.51 (15.44)[b]	1298 (45.16)	491 (51.58)[b]	Number with any diagnosis/symptom recorded in the 2 weeks pre-test = 1668 (43.60)	2.52 (2.18 to 2.93)	1.32 (0.83)	1.19 (0.51)[e]
					Soft tissue disorders (*n* = 218)			
Blood film report	51.89 (23.74)	49.28 (21.48)	427 (40.02)	90 (33.46)[e]	Number with any diagnosis/symptom recorded in the 2 weeks pre-test = 805 (60.25)	1.83 (1.46 to 2.28)	1.19 (0.54)	1.22 (0.58)
					Fatigue, tiredness, malaise and dizziness (*n* = 92)			

n = number of times the symptoms was recorded in the 2 weeks prior to the test in the patients receiving the test. SD = standard deviation. S&S = signs and symptoms. W/e = with specific exclusions. [a]Comparing the relative frequency of Read Code Groups recorded in patients who had that test (cases) compared with a cohort of patients matched on sex, age, and geographic region (controls) who underwent a different laboratory test in the same quarter of the year as the case. [b]The total number of tests they received during the year, given that they had at least one test. [c] Significant at the 0.1% level when fitting a Poisson model (tests per patient), linear model (age); logistic model (% male) with year as covariate. [b]Significant at the 1% level. [e]Significant at the 5% level.

10 000 person-years, after accounting for changes in the age and sex composition of the cohort. The largest relative increases, more than 75% over the 5-year study period, were observed in faecal occult blood, C-reactive protein, and haematinic tests. Most of the largest relative temporal increases reflect the testing of more patients rather than testing the same patients more frequently. In absolute terms, substantial increases of more than 650 tests were found in the urea and electrolytes, full blood count, and liver function tests. There was substantial geographic variation; in some regions, rates of plasma viscosity, cardiac enzymes, and blood trace elements/vitamins testing were more than three times the national average.

Strengths and limitations

A large sample of patients afforded substantial statistical power to investigate temporal trends and geographic variations. Several studies have found GPRD data to be of high accuracy and completeness when compared to other databases.[18–20] The GPRD is a voluntary scheme, but participating general practices are quality assured and broadly representative of the UK population,[21] suggesting the present results are reliable and generalisable. GPRD guidelines ask that GPs record all clinically 'significant' test results.[22] Therefore, it is possible that tests within normal range have not been recorded in the database, and improvements in recording, specifically movement to electronic test recording, might be responsible for the temporal trends observed. However, electronic reporting of laboratory tests was used in 90% of GPRD practices by 2005.[23] Furthermore, during preliminary analysis of selected tests (C-reactive protein, folate level, glycosylated haemoglobin (HbA1c), thyroid stimulating hormone, serum creatinine, blood glucose, calcium, alkaline phosphatase), it was found that the proportion of tests with numerical results recorded in the GPRD (85.9%, 93.6%, and 94.6% in 1995–1999, 2000–2004, and 2005–2009 respectively), and the proportion of

results within the normal range (54.6%, 71.3%, and 73.0 respectively) had reached a plateau by 2005. However, given the small number of GPRD practices in some regions (for example, North East), geographic variations may be influenced by differences in recording completeness.

The GPRD does not require GPs to record a reason or 'problem heading' for ordering a test. Some insight into reasons for test utilisation was provided in this study by presenting symptoms and diagnoses that were relatively more frequently recorded in the 2 weeks before test ordering. However, these results must be interpreted with caution, owing to the large proportion of tests where no reason for testing could be ascertained. Furthermore, this approach may be better able to detect acute indications, rather than chronic-disease-monitoring activity, because the GPRD guidelines require that only new occurrences of symptoms or episodes of illness are recorded.[22]

Comparison with existing literature

There are no official statistics on laboratory test usage in the UK, and to the authors' knowledge this is the first study of its kind in the UK primary care sector. Data from the Bettering the Evaluation and Care of Health study, conducted in Australia, reported an annual increase in utilisation of pathology tests of 7% between 2004 and 2008.[24] Several UK-based and international studies have concluded that regional variation exists in primary care test-ordering behaviour.[6–11] One of these studies observed large differences, probably driven by clinical practice variation,[10] in the patterns of pathology test requests of 22 UK general practices. Some studies have investigated variations at the clinician[8] or practice[9,10] level; however, individual practices or practitioners could not be identified within the present study.

Implications for practice and research

This study has demonstrated widespread increases in test utilisation, some of which will be appropriately generated through adherence to guidelines for chronic disease management, but others (for example, regional variations in plasma viscosity use) may reflect a growth in inappropriate testing. For example, the introduction of clinical guidelines for rheumatoid arthritis, based on evidence from randomised controlled trials, recommending 'regular' use of C-reactive protein testing for monitoring and optimisation of therapy, particularly in patients with recent onset of disease, might be one factor underlying the large increases in the use of this test during the study period. Increases in the use of faecal occult blood tests might be directly or indirectly related to the introduction of the national bowel cancer screening programme in 2006. Results from the screening programme are not typically recorded in GP records, but may directly contribute to some of the increase observed in the present study. Furthermore, the screening programme has probably led to greater patient pressure for testing and increased awareness among clinicians around the appropriate indications for testing. Conversely, the decrease in GP-initiated pregnancy testing that was observed in this study might reflect an increased reliance on results from home pregnancy testing kits.[25] The QOF,[3] which directly encourages the use of some tests in the present study (for example, HbA1c and thyroid function tests), is likely to be responsible for some of the increases seen in this study. However, the widespread increases in test use, transcending both QOF and non-QOF tests, suggest that other mechanisms play an important role in driving utilisation.

For some tests, the observed extensive regional variation might indicate that GPs are uncertain about the appropriate rate of usage. Alternatively, variation in local laboratory policies rather than GP preferences, for example differences in the numerical criteria used by laboratories when deciding whether a blood film is necessary,[26] may drive geographic variation. It is unclear whether this variation reflects inappropriate testing in regions of high use, or, conversely, failure to provide effective care in regions of low use; however, exploration of geographic variations in test use may help identify priorities for future research. For example, plasma viscosity tests in patients with suspected rheumatoid arthritis and other musculoskeletal disorders varied markedly by region. While plasma viscosity testing might have some prognostic value in arthritis,[27] it is not recommended by the National Institute for Health and Clinical Excellence's (NICE's) guidelines for rheumatoid arthritis.[28]

Guidelines are available to disseminate knowledge about appropriate test use to practitioners.[29–31] However, research suggests that they generally lag behind best practice, are often supported by poor evidence, and GP adherence may be low.[32–34] Numerous other, more intensive, interventions, such as tailored education/feedback,[35–38] decision support,[39,40] redesign of test-order forms,[41] and physician quality

Funding

This project was funded by the MRC Methodology Research Panel Grant 'Making diagnostic systematic reviews fit for purpose: improving decision making in the NHS' (Grant Number: G0801405).

Ethical approval

None required.

Provenance

Freely submitted; externally peer reviewed.

Competing interests

The authors have declared no competing interests.

Acknowledgements

The authors would like to thank Chris Salisbury and Tjeerd Van Staa for their comments an earlier draft of this paper and Tarita Murray-Thomas for advice on the GPRD dataset.

Discuss this article

Contribute and read comments about this article on the Discussion Forum: **http://www.rcgp.org.uk/bjgp-discuss**

improvement,[42] have been shown to effectively reduce demand. However, such interventions can be costly, and in some cases the savings made may be less than the programme cost.[43]

The present findings highlight the need for research on the appropriate use of laboratory tests for diagnosis, treatment monitoring, and prognosis in primary care, at the level of both practices and clinicians. Where there is relatively strong evidence that the test has good accuracy (for example, C-reactive protein for monitoring disease activity in rheumatoid arthritis), the key questions for clinicians and policy makers relate to better definition of the patient subgroups that should be tested, the optimal interval between screening or monitoring tests, and the optimal duration of monitoring. In the absence of such information, vague guideline recommendations that patients should be monitored 'regularly', or until 'treatment has controlled the disease', may continue to foster the large increases in test use observed since 2005.[44] For other tests, such as plasma viscosity, where regional variation in test use is highest and evidence may be poor quality or sparse, additional primary diagnostic/prognostic accuracy studies and systematic reviews are needed to inform appropriate policy.

REFERENCES

1. Lord Carter of Coles. *Report of the review of NHS pathology services in England.* London: DoH Publications, 2006.
2. Commission for Healthcare Audit and Inspection. *Getting results: pathology services in acute and specialist trusts.* London: Commission for Healthcare Audit and Inspection, 2007.
3. NHS Employers and General Practitioners Committee. *Quality and Outcomes Framework guidance for GMS contract 2009/10. Delivering investment in general practice.* London: The NHS Confederation (Employers) Company Ltd, 2009.
4. Bayram C. *Evidence–practice gap in GP pathology test ordering: a comparison of BEACH pathology data and recommended testing.* Sydney: Family Medicine Research Centre, University of Sydney, 2009.
5. Rao GG, Crook M, Tillyer ML. Pathology tests: is the time for demand management ripe at last? *J Clin Pathol* 2003; **56(4):** 243–248.
6. Mindemark M, Wernroth L, Larsson A. Costly regional variations in primary health care test utilization in Sweden. *Scand J Clin Lab Invest* 2010; **70(3):** 164–170.
7. Song Y, Skinner J, Bynum J, *et al.* Regional variations in diagnostic practices. *N Engl J Med* 2010; **363(1):** 45–53.
8. Verstappen WH, ter Riet G, Dubois WI, *et al.* Variation in test ordering behaviour of GPs: professional or context-related factors? *Fam Pract* 2004; **21(4):** 387–395.
9. McNulty CA, Freeman E, Bowen J, *et al.* Diagnosis of genital chlamydia in primary care: an explanation of reasons for variation in chlamydia testing. *Sex Transm Infect* 2004; **80(3):** 207–211.
10. Smellie WS, Galloway MJ, Chinn D, Gedling P. Is clinical practice variability the major reason for differences in pathology requesting patterns in general practice? *J Clin Pathol* 2002; **55(4):** 312–314.
11. Smellie WS, Clark G, McNulty CA. Inequalities of primary care microbiology testing between hospital catchment areas. *J Clin Pathol* 2003; **56(12):** 933–936.
12. Wennberg JE, Barnes BA, Zubkoff M. Professional uncertainty and the problem of supplier-induced demand. *Soc Sci Med* 1982; **16(7):** 811–824.
13. Studdert DM, Mello MM, Sage WM, *et al.* Defensive medicine among high-risk specialist physicians in a volatile malpractice environment. *JAMA* 2005; **293(21):** 2609–2617.
14. van Bokhoven MA, Pleunis-van Empel MC, Koch H, *et al.* Why do patients want to have their blood tested? A qualitative study of patient expectations in general practice. *BMC Fam Pract* 2006; **7:** 75.
15. Pan American Health Organization. Standardization: a classic epidemiological method for the comparison of rates. *Epidemiol Bull* 2002; **23(3):** 9–12.
16. Gray J, Orr D, Majeed A. Use of Read codes in diabetes management in a south London primary care group: implications for establishing disease registers. *BMJ* 2003; **326(7399):** 1130.
17. Information Services Division Scotland. *Grouping of codes for conditions (RCGs).* http://www.isdscotland.org/Health-Topics/General-Practice/PTI-Statistics/Grouping-clinical-codes.asp (accessed 12 Mar 2013).
18. Khan NF, Harrison SE, Rose PW. Validity of diagnostic coding within the General Practice Research Database: a systematic review. *Br J Gen Pract* 2010; DOI: 10.3399/bjgp10X483562.
19. Herrett E, Thomas SL, Schoonen WM, *et al.* Validation and validity of diagnoses in the General Practice Research Database: a systematic review. *Br J Clin Pharmacol* 2010; **69(1):** 4–14.
20. Thomas SL, Edwards CJ, Smeeth L, *et al.* How accurate are diagnoses for rheumatoid arthritis and juvenile idiopathic arthritis in the general practice research database? *Arthritis Rheum* 2008; **59(9):** 1314–1321.
21. Lawrenson R, Williams T, Farmer R. Clinical information for research; the use of general practice databases. *J Public Health Med* 1999; **21(3):** 299–304.
22. Medicines and Healthcare Products Regulatory Agency. *GPRD recording guidelines for vision users.* London: Crown Copyright, 2004.
23. Anonymous. GPRD Users' Group meeting. *Scribe* 2005; **8(4):** 20.
24. Australian Association of Pathology Practices Inc. *An analysis of pathology test use in Australia.* Sydney: Family Medicine Research Centre, University of Sydney, 2011.
25. The King's Fund. *The role of GPs in maternity care — what does the future hold?* London: The King's Fund, 2010.
26. Galloway MJ, Osgerby JC. An audit of the indications for the reporting of blood films: results from the National Pathology Benchmarking Study. *J Clin Pathol* 2006; **59(5):** 479–481.
27. Vercoutere W, Thevissen K, Bombardier C, Landewe RB. Diagnostic and predictive value of acute-phase reactants in adult undifferentiated peripheral inflammatory arthritis: a systematic review. *J Rheumatol Suppl* 2011; **87:** 15–19.
28. National Collaborating Centre for Chronic Conditions (UK). *Rheumatoid arthritis: national clinical guideline for management and treatment in adults.* London: Royal College of Physicians, 2009.
29. Smellie WS, McNulty CA, Collinson PO, *et al.* Best practice in primary care pathology: review 12. *J Clin Pathol* 2010; **63(4):** 330–336.
30. Smellie WS, Vanderpump MP, Fraser WD, *et al.* Best practice in primary care pathology: review 11. *J Clin Pathol* 2008; **61(4):** 410–418.
31. Smellie WS, Shaw N, Bowley R, *et al.* Best practice in primary care pathology: review 10. *J Clin Pathol* 2007; **60(11):** 1195–1204.
32. Wathen B, Dean T. An evaluation of the impact of NICE guidance on GP prescribing. *Br J Gen Pract* 2004; **54(499):** 103–107.
33. Gerrits M, Avery T, Lagro-Janssen A. Urinary incontinence management in women: audit in general practice. *J Eval Clin Pract* 2008; **14(5):** 836–838.
34. van den Brandhof WE, Bartelds AI, Koopmans MP, van Duynhoven YT. General practitioner practices in requesting laboratory tests for patients with gastroenteritis in the Netherlands, 2001–2002. *BMC Fam Pract* 2006; **7:** 56.
35. Bunting PS, Van WC. Effect of a controlled feedback intervention on laboratory test ordering by community physicians. *Clin Chem* 2004; **50(2):** 321–326.
36. Mindemark M, Larsson A. Long-term effects of an education programme on the optimal use of clinical chemistry testing in primary health care. *Scand J Clin Lab Invest* 2009; **69(4):** 481–486.
37. Sood R, Sood A, Ghosh AK. Non-evidence-based variables affecting physicians' test-ordering tendencies: a systematic review. *Neth J Med* 2007; **65(5):** 167–177.
38. Thomas RE, Croal BL, Ramsay C, *et al.* Effect of enhanced feedback and brief educational reminder messages on laboratory test requesting in primary care: a cluster randomised trial. *Lancet* 2006; **367(9527):** 1990–1996.
39. Poley MJ, Edelenbos KI, Mosseveld M, *et al.* Cost consequences of implementing an electronic decision support system for ordering laboratory tests in primary care: evidence from a controlled prospective study in the Netherlands. *Clin Chem* 2007; **53(2):** 213–219.
40. van Wijk MA, van der Lei J, Mosseveld M, *et al.* Assessment of decision support for blood test ordering in primary care: a randomized trial. *Ann Intern Med* 2001; **134(4):** 274–281.
41. Shalev V, Chodick G, Heymann AD. Format change of a laboratory test order form affects physician behavior. *Int J Med Inform* 2009; **78(10):** 639–644.
42. Verstappen WH, van der Weijden T, Dubois WI, *et al.* Improving test ordering in primary care: the added value of a small-group quality improvement strategy compared with classic feedback only. *Ann Fam Med* 2004; **2(6):** 569–575.
43. Schroeder SA, Myers LP, McPhee SJ, *et al.* The failure of physician education as a cost containment strategy. Report of a prospective controlled trial at a university hospital. *JAMA* 1984; **252(2):** 225–230.
44. National Institute for Health and Clincial Excellence. *Rheumatoid arthritis: the management of rheumatoid arthritis in adults.* (CG79). London: National Institute for Health and Clinical Excellence, 2009.

Appendix 1. Groupings of tests

Battery	Test unit
Full blood count	Eosinophil count Haemoglobin Mean corpuscular haemoglobin Mean corpuscular haemoglobin concentration Mean corpuscular volume Monocyte count Neutrophil count All platelet tests Red blood cell count Total white blood cell count Lymphocyte count Full blood count Packed cell volume Basophil count Differential white cell count Red blood cell size
Creatinine, urea and electrolytes	Serum creatinine Potassium Sodium Urea blood Urea and electrolytes Renal function tests GFR (calculated based on MDRD) Serum electrolytes
Liver function test	Albumin Alkaline phosphatase Alanine aminotransferase Aspartate aminotransferase Bilirubin Gamma glutamyl transpeptidase Total protein Liver function tests Serum globulin (calculated) Liver enzymes
Lipid profile	Serum cholesterol High-density lipoprotein (HDL) Low-density lipoprotein (LDL) Triglycerides Blood lipids HDL:LDL ratio
Thyroid function tests	Thyroxine Thyroid stimulating hormone Thyroid function Free tri-iodothyronine (T3) Free thyroxine Serum T3 level
Urine MCS	Urine test Midstream specimen of urine Microscopy culture and sensitivities
Urine dipstick	Urine ketones Urinalysis glucose Urinalysis protein Urine dipstick for glucose Urine dipstick for protein Urine dipstick for ketones Urine dipstick for blood Urine dipstick for pH
Serum troponin[a]	Serum troponin T Serum troponin I

... continued

Appendix 1 continued. Groupings of tests

Glucose	Blood glucose Glucose tolerance test Fasting glucose
Bone profile	Serum inorganic phosphate Calcium Calcium adjusted Bone studies
Haematinics	B12 levels Serum ferritin Folate B12 and folate RBC folate
Iron studies	Serum iron studies Total iron-binding capacity Iron studies
Clotting tests	Clotting tests Prothrombin time Partial thromboplastin time Thrombin time INR
Sex hormones	Female sex hormones Follicle stimulating hormone Luteinising hormone Male sex hormones Oestradiol level Progesterone Sex hormone binding globulin Testosterone
Urine biochemistry	Urine biochemistry Urine microalbumin Drug levels Other drug levels
Vaginal swab	Vaginal swab High vaginal swab

GFR = glomerular filtration rate. INR = international normalised ratio. MCS = microscopy, culture and sensitivities. MDRD = modification of diet in renal disease. RBC = red blood cell.[a]Utilisation less than 50 tests per 10 000 person-years and so not included in Table 2.

Appendix 2. Tests that are excluded from the study tables/figures

Test	Reason
Other laboratory tests	Miscellaneous
Other bacteriology tests	Miscellaneous
Large unstained cells	Miscellaneous
Infection titres	Miscellaneous
Other autoantibodies	Miscellaneous
Histology	Miscellaneous
Other biochemistry test	Miscellaneous
Other immunology	Miscellaneous
Other drug levels	Miscellaneous
Other swab	Miscellaneous
Uric acid blood level	Secondary care
Serum chloride	Secondary care
Serum bicarbonate	Secondary care
Anion gap	Secondary care
Blood gases	Secondary care
Ascitic fluid examination	Secondary care
Schilling test	Secondary care