

BMJ Open Primary care use of laboratory tests in Northern Ireland's Western Health and Social Care Trust: a cross-sectional study

Magda Bucholc,¹ Maurice O'Kane,² Ciaran Mullan,³ Siobhan Ashe,² Liam Maguire¹

To cite: Bucholc M, O'Kane M, Mullan C, *et al.* Primary care use of laboratory tests in Northern Ireland's Western Health and Social Care Trust: a cross-sectional study. *BMJ Open* 2019;9:e026647. doi:10.1136/bmjopen-2018-026647

► Prepublication history and additional material for this paper are available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2018-026647>).

Received 12 September 2018
Revised 27 March 2019
Accepted 13 May 2019



© Author(s) (or their employer(s)) 2019. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

¹School of Computing, Engineering and Intelligent Systems, University of Ulster - Magee Campus, Londonderry, UK

²Clinical Chemistry, Altnagelvin Area Hospital, Londonderry, UK

³Western Local Commissioning Group, Health and Social Care Board, Londonderry, UK

Correspondence to

Magda Bucholc;
m.bucholc@ulster.ac.uk

ABSTRACT

Objectives To describe the laboratory test ordering patterns by general practitioners (GPs) in Northern Ireland Western Health and Social Care Trust (WHSC) and explore demographic and socioeconomic associations with test requesting.

Design Cross-sectional study.

Setting WHSC, Northern Ireland.

Participants 55 WHSC primary care medical practices that remained open throughout the study period 1 April 2011–31 March 2016.

Outcomes To identify the temporal patterns of laboratory test ordering behaviour for eight commonly requested clinical biochemistry tests/test groups in WHSC. To analyse the extent of variations in laboratory test requests by GPs and to explore whether these variations can be accounted for by clinical outcomes or geographical, demographic and socioeconomic characteristics.

Results The median number of adjusted test request rates over 5 consecutive years of the study period decreased by 45.7% for urine albumin/creatinine ratio ($p < 0.000001$) and 19.4% for lipid profiles ($p < 0.000001$) while a 60.6%, 36.6% and 29.5% increase was observed for HbA_{1c} ($p < 0.000001$), immunoglobulins ($p = 0.000007$) and prostate-specific antigen (PSA) ($p = 0.0003$), respectively. The between-practice variation in test ordering rates increased by 272% for immunoglobulins ($p = 0.008$) and 500% for HbA_{1c} ($p = 0.0001$). No statistically significant relationship between ordering activity and either demographic (age and gender) and socioeconomic factors (deprivation) or Quality and Outcome Framework scores was observed. We found the rural–urban differences in between-practice variability in ordering rates for lipid profiles, thyroid profiles, PSA and immunoglobulins to be statistically significant at the Bonferroni-adjusted significance level $p < 0.01$.

Conclusions We explored potential factors of the interpractice variability in the use of laboratory tests and found that differences in requesting activity appear unrelated to either demographic and socioeconomic characteristics of GP practices or clinical outcome indicators.

INTRODUCTION

Despite the important role of laboratory testing in the diagnosis and monitoring of disease, there is concern about the increasing number of requested tests and in particular,

Strengths and limitations of this study

- The study provides a comprehensive analysis of temporal changes in laboratory test utilisation patterns and establishes the extent of variability in test requesting activity across general practices in Northern Ireland's Western Health and Social Care Trust.
- The variation in test ordering, not related to demographic and socioeconomic characteristics of practices, practice location or clinical outcome indicators, may reflect inappropriate laboratory test utilisation and hence, suggest a potential for more efficient demand management of laboratory services.
- Given a cohort of general practices within one catchment area, our results provide evidence of differences in behaviour of individual general practitioner (GP) when managing patients with similar clinical symptoms.
- Failure to collect and cross-tabulate data on characteristics of GPs, such as GP's age, years of experience, medical training was a study limitation and a missed opportunity in assessing the influence of practitioner factors on the variation in test ordering behaviour.

large differences in laboratory utilisation between clinical teams.¹ In the UK, laboratory test orders grew by approximately 5% per year in recent years and inappropriate test requests are considered to be an important cause for this increase.^{2,3} Although pathology expenditures account for only 5%–6% of the UK total health budget, they are viewed as a potential source of savings, most likely because the costs can be easily identified and measured.⁴ According to the Department of Health, the rationalisation of pathology services, including demand management of laboratory tests and elimination of unnecessary requesting, could produce savings of at least £500 million.⁵

Unnecessary testing is not only wasteful of resources but impacts on patients directly through the requirement for venepuncture and the follow-up of minor (and possibly

insignificant) abnormalities which may cause patient anxiety. On the other hand, inappropriate under requesting may cause harm through failure to diagnose or manage disease optimally. Several studies suggested that unnecessary and inappropriate utilisation of laboratory services is closely linked to interpractitioner variability in test requesting.^{6–8} Despite the increased availability of clinical management guidelines promoting harmonisation of the use of laboratory tests, there is still substantial variation in test utilisation among general practitioners (GPs).^{7,9} These differences appear to be unrelated to demographic characteristics of patient populations, socioeconomic status of GP practices, disease prevalence or clinical outcome indicators.^{6,7,10,11} Even if some of these variables have been shown to have an effect on test ordering patterns, the variation in requesting rates is so large that it can only be explained by differences in attitudes towards the use of laboratory tests of individual practitioners.¹⁰ Accordingly, factors such as confidence in clinical judgement, clinical experience, an attitude about clinical practice guidelines, a lack of knowledge regarding the correct use of tests and fear of litigation have been identified as potential sources of practice variation.^{12–15}

Since unwarranted variation can lead to suboptimal clinical outcomes,^{16,17} identification of factors contributing to differences in test requesting can provide useful information for optimising utilisation of laboratory services. The aim of our study was to establish the extent of variability in test requesting and characterise temporal changes in test ordering patterns across general practices within the catchment area of the Northern Ireland (NI) Western Health and Social Care Trust (WHSC) for a range of most commonly requested clinical biochemistry tests/test groups. In addition, we investigated potential factors associated with interdoctor variability in the use of laboratory tests including geographical, demographic and socioeconomic factors as well as Quality and Outcome Framework (QOF) scores.

MATERIALS AND METHODS

Study design and data sources

We conducted a cross-sectional study of laboratory test ordering activity across general practices in the WHSC in the period from 1 April 2011 to 31 March 2016. The data on the use of laboratory tests were obtained from a HSC Business Object XI clinical information system. We investigated requesting rates for eight clinical biochemistry tests/test groups including electrolyte profile, thyroid profile (FT4 and TSH), liver profile, lipid profile, urine albumin/creatinine ratio (ACR), glycosylated haemoglobin (HbA_{1c}), prostate-specific antigen (PSA) and immunoglobulins. The profile tests, that is, electrolyte, lipid, liver and thyroid profiles contained a number of different related analytes. Individual elements of the profile tests could not be requested separately. Other tests, such as HbA_{1c}, comprised one single analyte. The standardised number of laboratory tests requested

in each general practice was determined by dividing the total number of requested tests by the number of registered patients and expressed as requests per 1000 patients, with the exception of HbA_{1c} and ACR for which ordering rates were expressed as tests per patient with diabetes and PSA rates, calculated per 1000 male patients. GP practice list size data (including the number of patients with diabetes) and patient demographic data were extracted from the Business Services Organisation Family Practitioner Service Information and Registration Unit system. The rural–urban distribution of GP practices was based on the data from the Census Office of the Northern Ireland Statistics and Research Agency (NISRA).¹⁸ Socioeconomic characteristics of GP practices were determined using the NISRA Neighbourhood Information System.¹⁹ QOF scores for individual practices were extracted from the website of the Northern Ireland Department of Health.²⁰

Participants and setting

Data on laboratory tests requested from 55 general practices within the WHSC were collected from the laboratory databases of Clinical Chemistry departments of the Altnagelvin Area Hospital, Tyrone County Hospital and the Erne Hospital (subsequently the South West Acute Hospital). WHSC provides health and social care services throughout the west of NI, across the council areas of Derry City and Strabane District Council, Limavady in the Causeway Coast and Glens Borough Council, and Fermanagh and Omagh District Council.

Inclusion criteria

We examined laboratory test requests from 55 separate primary care medical practices within the catchment area of WHSC that remained open throughout the study period, that is, during 5 consecutive years: April 2011–March 2012, April 2012–March 2013, April 2013–March 2014, April 2014–March 2015 and April 2015–March 2016. To ensure the completeness and consistency of the data, test orders from WHSC GP practices that were closed (two GP practices) or taken over by other practices (one GP practice) during the period of investigation were not taken into account.

Variables and characteristics

Our analysis was limited to eight most frequently requested clinical biochemistry tests/test groups including electrolyte profile, thyroid profile, liver profile, lipid profile, ACR, HbA_{1c}, PSA and immunoglobulins. All considered tests/profiles (with the exception of immunoglobulins and ACR) were listed on a paper pathology request form and ordered by ticking a box adjacent to the test name; requests for ACR and immunoglobulins were made by writing the test name in an open text field.

Rates of test requests were analysed in the context of geographical and socioeconomic characteristics of practices, patient demographics (age and gender) and clinical outcome indicators. Note that we were not able to

obtain a consolidated data on both gender and age of patients (ie, the number of patients split into age bands and broken down by gender) registered in individual GP practices.

The practice setting (rural vs urban area) was determined using the urban–rural classification of the Department for Environment, Food and Rural Affairs²¹ while the size of settlements was obtained from the NISRA Census Office.¹⁸ A GP practice was defined as rural if its physical address was situated in a settlement of less than 10 000. Accordingly, we classified 24 practices as rural and 31 as urban. Furthermore, GP practices were categorised based on the NI Multiple Deprivation Measure (MDM) identified for individual Super Output Areas (SOAs).¹⁹ The MDM comprises a weighted combination of seven component measures (Income Deprivation; Employment Deprivation; Health Deprivation and Disability; Education, Skills and Training Deprivation; Proximity to Services; Living Environment and Crime and Disorder) and ranges from 1 (most deprived SOAs) to 890 (least deprived SOAs). The Health Deprivation and Disability rank (one of the MDM domains)¹⁹ was analysed separately as a potential factor contributing to interdoctor variability in the use of laboratory tests.

Given a detailed breakdown of the age of patients on the individual practice lists, we examined the relationship between the percentage of patients over age 65 and ordering rates for electrolyte, thyroid, liver and lipid profiles. In addition, we investigated the link between gender and requesting activity for PSA (for males) and thyroid profiles (for females).

To evaluate the impact of test requesting patterns for diabetes mellitus on the QOF clinical indicator scores in diabetes, we used the overall QOF points under clinical indicator DM007 ('The percentage of patients with diabetes in whom the last HbA_{1c} is 59 mmol/mol or less in the preceding 15 months'), DM008 ('The percentage of patients with diabetes in whom the last HbA_{1c} is 64 mmol/mol or less in the preceding 15 months') and DM009 ('The percentage of patients with diabetes in whom the last HbA_{1c} is 75 mmol/mol or less in the preceding 15 months'). Specifically, we combined the number of points for each of these three categories, calculated the overall achievement rate and investigated its relationship with interpractitioner variability in test requesting of HbA_{1c}.

Outcome measures

Our outcome of interest was to identify the presence and extent of variations in primary care laboratory tests ordering and to evaluate temporal changes in both the standardised number of test requests and between-practice variability in requesting. In addition, we studied demographic, socioeconomic, geographical and clinical factors that may explain this variation.

Statistical analysis

Interpractice variability in test requests was assessed by calculating the variance (σ^2). Furthermore, we computed

'variability index' (Var_i) defined as the top decile divided by the bottom decile of standardised test request rates.¹⁰ Var_i is a dimensionless measure of dispersion allowing us to compare the amount of variation in request rates of individual tests despite their differences in scale. The Shapiro-Wilk test of normality was used to determine if the distribution of test ordering data deviated from a normal distribution.²² Since the distribution of laboratory test request rates was found to be non-normal, the non-parametric statistics were implemented to perform further analysis. Mann Whitney U (MWU) test was employed to compare distributions of laboratory test rates between GP practices located in rural and urban areas.²³ The homogeneity of variances for requesting activity in rural and urban GP practices was assessed with the Fligner-Killeen (FK) test.²⁴ Significance of temporal changes in median and variation in test requesting rates was examined with the Mann–Kendall (MK) test.²⁵ To test the relationship between adjusted requesting rates and (1) MDM; (2) Health Deprivation and Disability Deprivation rank; (3) proportions of patients over age 65; (4) distribution of patient's gender and (5) QOF clinical indicator scores, we calculated the Kendall rank correlation coefficient (τ) with the corresponding p-value.²⁶ The Kendall coefficient ($-1 \leq \tau \leq 1$) measures the strength of a correlation between two variables and assesses the degree of overall correspondence of variables' ranking.²⁶

Since the test statistics were used to simultaneously evaluate the significance of observations for different case scenarios (ie, differences over time or differences between clinical biochemistry tests/test groups), we took a conservative approach in selecting the significance level α by applying the Bonferroni correction.²³ This multiple comparison correction technique minimises the risk of obtaining false positive results by using an adjusted α for each single test. Single threshold values α were calculated in such a way that the family wise error probability p_{fwe} (here adopted $p_{\text{fwe}} < 0.05$) was retained at the global level.²³ Accordingly, the significance level α for a single test was approximated by dividing the global error probability p_{fwe} by the total number of independent tests.²³

Patient and public involvement

No patients or GPs were involved in defining the research question or outcome measures nor were they engaged in the design or implementation this study. The study was approved by the WHSCT. All identifiable information used for the purpose of the study was anonymised and not traceable to individual patients or GPs.

RESULTS

Test ordering patterns

We analysed the laboratory test request rates of 55 general practices within the catchment area of the WHSCT comprising a total of 316 382 (2011–2012), 316 688 (2012–2013), 318 057 (2013–2014), 319 383 (2014–2015) and 326 429 (2015–2016) patients. [Figure 1](#) shows

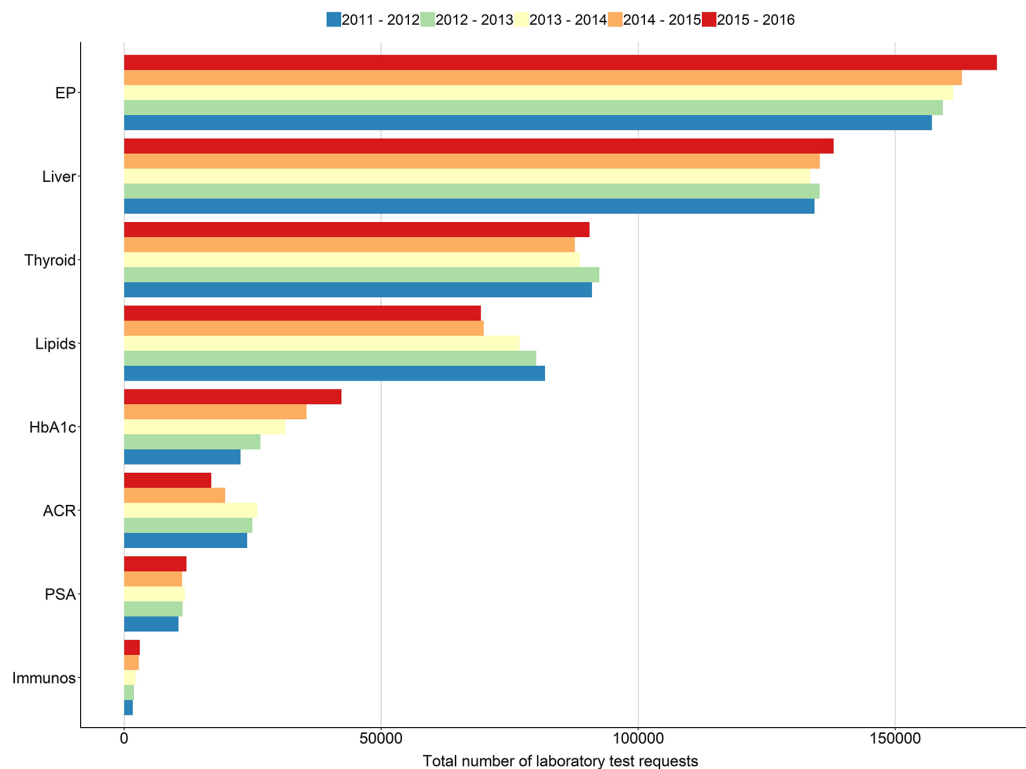


Figure 1 The total number of eight considered clinical biochemistry tests/test groups ordered by 55 general practices during the period 1 April 2011–31 March 2016. ACR, albumin/creatinine ratio; EP, electrolyte profiles; PSA, prostate-specific antigen.

the total number of eight considered clinical biochemistry tests/test groups ordered during the study period of 5 consecutive years. The total number of ordered tests was 523 111 (2011–2012), 531 849 (2012–2013), 531 583 (2013–2014), 525 146 (2014–2015) and 542 118 (2015–2016) with electrolyte profiles found to be the most frequently ordered tests, making up approximately 30% of all requests in each year.

The median number of adjusted test request rates between the 2011–2012 and 2015–2016 decreased by 45.7% for urine ACR ($p < 0.000001$) and 19.4% for lipid profiles ($p < 0.000001$). We observed a 60.6%, 36.6% and 29.5% increase for HbA_{1c} ($p < 0.000001$), immunoglobulins ($p = 0.000007$) and PSA ($p = 0.0003$), respectively. The Bonferroni-adjusted probability $p < 0.01$ was used in testing for statistical significance (table 1).

The between-practice variation in the number of laboratory test requests is shown in table 1. In the considered period, the Var_i was lowest for electrolyte profiles (2.0–2.3), liver profiles (2.2–2.5) and HbA_{1c} (2.2–2.9) and highest for immunoglobulins (19.2–69.8). The number of outlier GP practices, that is, practices with outlier ordering rates of laboratory tests contributed to the overall variability (figure 2). The number of requested tests for thyroid profiles, HbA_{1c} and immunoglobulins per patient was over three times higher in the outlier GP practices (table 2). We however acknowledge that a statistical outlier in terms of test utilisation is not necessarily an example of inappropriate practice. When compared with 2011–2012, the variance in test utilisation was 16.8%

lower for electrolyte profiles, 15.3% for lipid profiles, 28.3% for thyroid profiles and 23.5% for liver profiles in 2015–2016. For ACR, immunoglobulins and HbA_{1c}, the variance was 40%, 272% and 500% higher, respectively.

Trends in the number and variation of test requests over time

We observed no statistically significant linear trend in the median standardised number of laboratory tests over the period of investigation as indicated by the MK test p-values (figure 3). Furthermore, no linear trend was found in the annual changes in variation in test requesting rates (figure 3). Note that to control the false discovery rate in multiple comparisons (for eight laboratory tests), the Bonferroni-adjusted probability $p < 0.00625$ was used in testing for statistical significance.

Test ordering variation between general practices and explanatory factors

The relationship between test requesting rates and potential explanatory factors was established based on the information on the number of ordered tests, patient demographics and Quality and Outcomes Framework indicator scores obtained for the period 1 April 2015–31 March 2016.

Demographic characteristics of patient population

Proportion of the oldest age category of patients (>75) constituted a relatively small group in each GP practice (mean=6.1%, 95% CI = 5.7% to 6.5%). Hence, we combined the 65–74 and >75 age categories to create a

Table 1 Overview of general statistics

	EP	Lipids	PSA	Thyroid	HbA1c	Liver	ACR	Immunos
2011–2012								
Median, CI	520.5 (465.5 to 545.0)	268.9 (242.2 to 300.5)	53.2 (48.3 to 79.8)	296.7 (266.2 to 339.3)	1.8 (1.7 to 2.0)	435.7 (394.3 to 473.9)	1.9 (1.6 to 2.1)	4.1 (3.1 to 6.0)
Range	267–1002	101.8–954.5	19.6–279.3	121–969.4	1.1–3.1	254–989.1	0.6–3.8	0.3–16.8
Variance	15660.2	14519.1	2193.8	16320.0	0.2	16187.4	0.5	15.8
Var _i	2.3	3.4	7.0	3.3	2.2	2.5	3.4	18.3
2012–2013								
Median, CI	491.7 (474.6 to 542.5)	257.4 (236.1 to 288.5)	59.2 (51.5 to 88.3)	291.7 (266.5 to 342.1)	2.0 (1.8 to 2.2)	432.3 (399.0 to 473.4)	1.7 (1.6 to 2.0)	4.5 (3.6 to 7.9)
Range	283.6–1041.8	105.8–986.3	19.9–396.1	115.5–995.2	1.1–3.4	258.7–1016.4	0.5–4.8	0–24
Variance	14299.4	14494.8	3896.4	17230.2	0.3	14997.6	0.5	28.0
Var _i	2.2	3.4	7.6	3.3	2.4	2.4	3.7	69.8
2013–2014								
Median, CI	508.8 (469.2 to 534.2)	239.3 (221.3 to 271.2)	59.5 (54.9 to 83.5)	264.3 (246.4 to 318.0)	2.1 (1.9 to 2.5)	408.3 (381.6 to 453.3)	1.8 (1.6 to 2.1)	4.4 (3.7 to 8.4)
Range	325.1–1014.9	111–950.7	23.1–527.6	141.5–962.2	1.3–4.6	267.1–987.8	0.7–4.7	0.5–24.8
Variance	12831.5	13134.3	5153.6	15099.2	0.6	13739.4	0.5	36.2
Var _i	2.0	3.1	7.2	3.0	2.8	2.3	3.3	19.2
2014–2015								
Median, CI	508.7 (476.5 to 541.4)	224.0 (198.6 to 248.4)	63.9 (55.6 to 81.0)	260.4 (244.9 to 311.8)	2.3 (2.2 to 2.8)	422.4 (389.5 to 458.4)	1.3 (1.2 to 1.7)	5.9 (4.7 to 9.9)
Range	274.9–976.9	82.8–939.5	17.1–274	152.7–951.7	1.5–5.9	260.7–968	0.2–4	0–46.6
Variance	12696.9	13814.0	2192.6	14324.6	0.8	13251.9	0.5	67.3
Var _i	2.1	3.5	6.3	3.0	2.9	2.3	4.7	33.6
2015–2016								
Median, CI	527.5 (482.9 to 556.9)	216.6 (193.2 to 244.6)	68.9 (60.2 to 86.7)	268.4 (251.6 to 315.1)	2.8 (2.6 to 3.3)	426.3 (391.7 to 457.7)	1.0 (0.9 to 1.3)	5.6 (5.0 to 10.4)
Range	316–925.3	96.9–846.5	26.4–296.9	142.5–857.2	1.7–8.1	274.6–898.3	0.1–5	0–33.9
Variance	13048.3	12303.9	2961.0	11708.1	1.2	12378.9	0.7	58.9
Var _i	2.0	3.7	6.5	2.8	2.9	2.2	9.4	24.5

ACR, albumin/creatinine ratio; EP, electrolyte profiles; PSA, prostate-specific antigen.

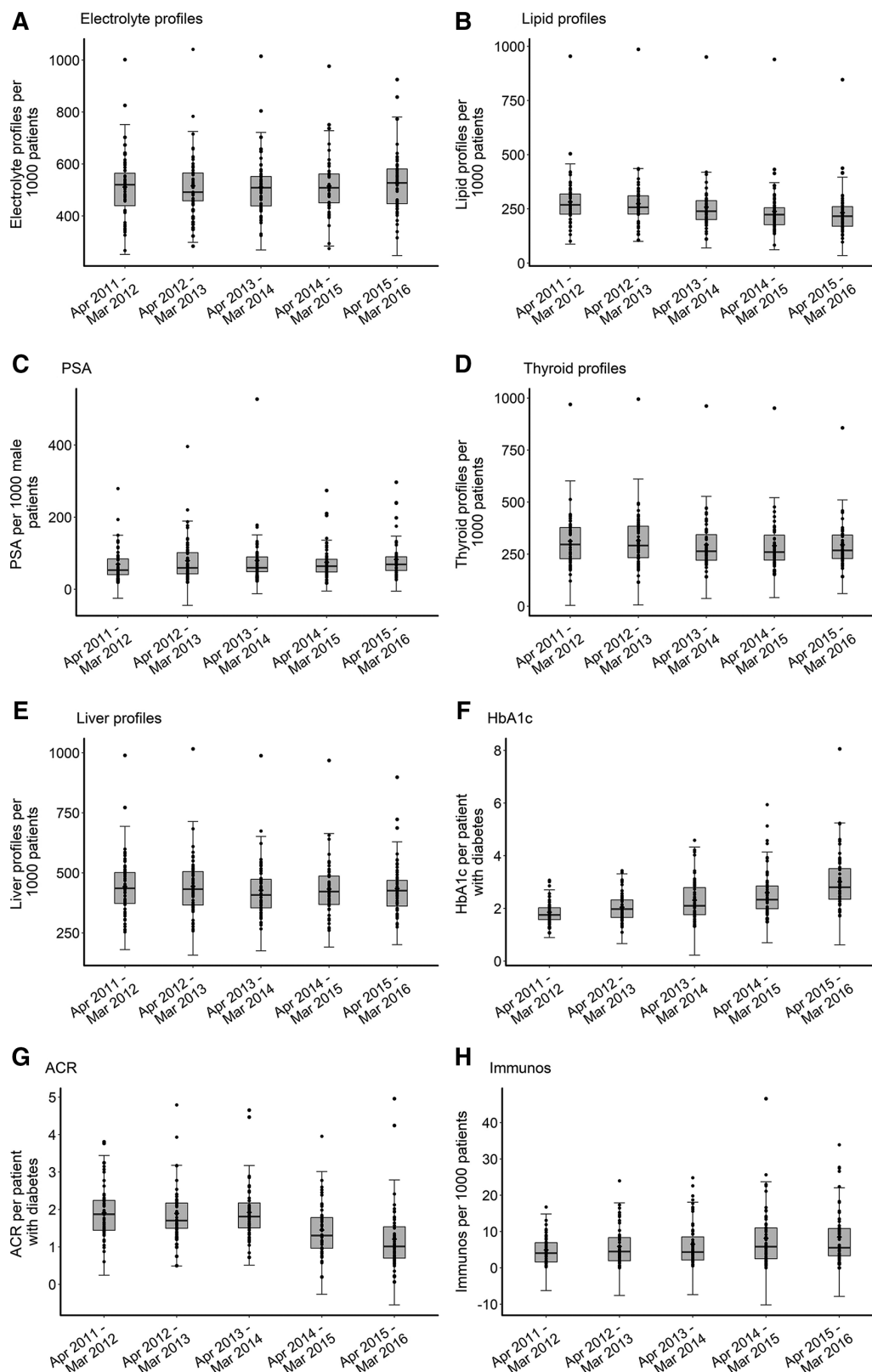


Figure 2 Temporal variability of the standardised laboratory test requests for (A) electrolyte profiles, (B) lipid profile, (C) prostate-specific antigen (PSA), (D) thyroid profiles, (E) liver profiles, (F) HbA_{1c}, (G) urine albumin/creatinine ratio (ACR) and (H) immunoglobulins (immunoglobulins) for 55 considered general practices. Each data point (dot): a single practice. Solid, horizontal line inside the box: median. Lower and upper 'hinges' of the boxplots: first and third quartiles, respectively. Lower and upper extremes of whiskers: interval boundaries of the non-outliers (black dots). Data outside interval: outliers.

Table 2 Number of requested tests per patients in outlier (O) and non-outlier (NO) GP practices

Period	Electrolyte profiles		Lipid profiles		PSA		Thyroid profiles		HbA _{1c}		Liver profiles		ACR		Immunos	
	O	NO	O	NO	O	NO	O	NO	O	NO	O	NO	O	NO	O	NO
2011–2012	0.88	0.49	0.65	0.25	0.22	0.06	0.97	0.28	3.04	1.78	0.84	0.42	3.80	1.83	0.017	0.005
2012–2013	0.58	0.50	0.99	0.25	0.28	0.07	1.00	0.29	3.32	1.99	1.02	0.42	4.61	1.80	0.024	0.006
2013–2014	0.85	0.50	0.95	0.24	0.22	0.07	0.96	0.28	4.59	2.32	0.74	0.41	4.50	1.84	0.024	0.006
2014–2015	0.78	0.50	0.52	0.21	0.22	0.07	0.95	0.27	5.00	2.46	0.97	0.42	3.96	1.43	0.034	0.008
2015–2016	0.87	0.51	0.51	0.21	0.23	0.07	0.86	0.27	8.06	2.93	0.74	0.41	4.55	1.13	0.027	0.008

The requesting rates for (1) electrolyte profiles, liver profiles, thyroid profiles, lipid profiles and immunoglobulins are expressed as the number of tests per patient; (2) PSA—as the number of tests per male patient and (3) ACR and HbA_{1c}—as the number of tests per patient with diabetes.

ACR, albumin/creatinine ratio; GP, general practitioner; PSA, prostate-specific antigen.

more meaningful group of patients in the older age band. We examined four tests for which differences in proportions of older patients were expected to be reflected in test requesting, that is, electrolyte profiles, liver profiles, lipid profiles and thyroid profiles. Since we were unable to obtain a consolidated data on both gender and age of patients, we did not assess the relationship between the PSA requesting rates and the category of males aged 65 and over. We found no statistically significant relationship between the adjusted request rates of selected laboratory tests and the percentage distribution of patients of age over 65 (online supplementary table S1).

Since previous studies reported on a higher rate of testing of thyroid hormones in females,²⁷ we looked at the strength of a relationship between the percentage of females in individual GP practices and requesting rates for thyroid profiles. We found no statistically significant association between these two characteristics ($\tau=-0.19$, $p=0.05$) (online supplementary figure S1).

In addition, we examined the effect of gender distribution on the standardised number of PSA tests, with the PSA ordering rates expected to be higher for GP practices with higher percentage of males (online supplementary figure S2). No statistically significant correlation between these two variables was found ($\tau=0.15$ and $p=0.1$). Note that we acknowledge the fact that a combined effect of sex and age distributions might have had a more significant effect on PSA requesting activity. However, the restricted access to such data limited the opportunity of more insightful analysis.

Practice setting

Figure 4 shows a temporal median–variance relationship of the standardised number of laboratory test requests for rural and urban areas. The median number of tests ordered annually by practices located in rural areas was higher by approximately 27%–37% for PSA, 14%–30% for lipid profiles, 14%–38% for thyroid profiles and 8%–23% for liver profiles. For ACR and immunoglobulins, the median requesting rates were lower in rural areas by 1%–27% and 18%–57%, respectively.

Across five consecutive time periods, the MWU test showed significant differences between ‘rural’ and ‘urban’ distributions of laboratory test rates for lipid profiles ($p_{\text{MWU}}=0.0004$ in 2011–2012, 0.001 in 2012–2013 and 0.0097 in 2013–2014), thyroid profiles ($p_{\text{MWU}}=0.009$ in 2011–2012), PSA ($p_{\text{MWU}}=0.003$ in 2015–2016) and immunoglobulins ($p_{\text{MWU}}=0.0096$ in 2015–2016) while no statistically significant rural–urban difference in between-practice variation in ordering rates was found for any of the studied* tests (table 3). Note that p_{MWU} and p_{FK} lower than Bonferroni-corrected significance cut-off $\alpha < 0.01$ (corrected for five different measurements over time) were considered statistically significant.

Socioeconomic factors

We observed no statistically significant relationship between the MDM and test requesting activity (online supplementary table S2). Similarly, the relationship between the Health Deprivation and Disability rank and the standardised number of test requests was found insignificant for any of the considered tests (online supplementary table S2). Note that to adjust for multiple comparisons at a family-wise simultaneous error rate of $p_{\text{fwe}} < 0.05$, the Bonferroni-adjusted $\alpha < 0.003125$ (corrected for 16 different comparisons) was used to test for statistical significance.

Quality and outcomes framework indicators

To evaluate the relationship between standardised number of laboratory tests and QOF indicator scores, we looked at the management of diabetes. There are two main reasons for that. First, the guidelines for the management of diabetes are widely available to practitioners to improve clinical practice and care of diabetic patients. Second, HbA_{1c} is a specific test in diabetes care and does not play an important role in the monitoring or diagnosis of any other condition.

All practices achieved the maximum 17 points available under QOF clinical indicator DM007. All but nine practices attained the maximum eight points available under DM008 (range: 4.1–8.0) and 31 practices attained the

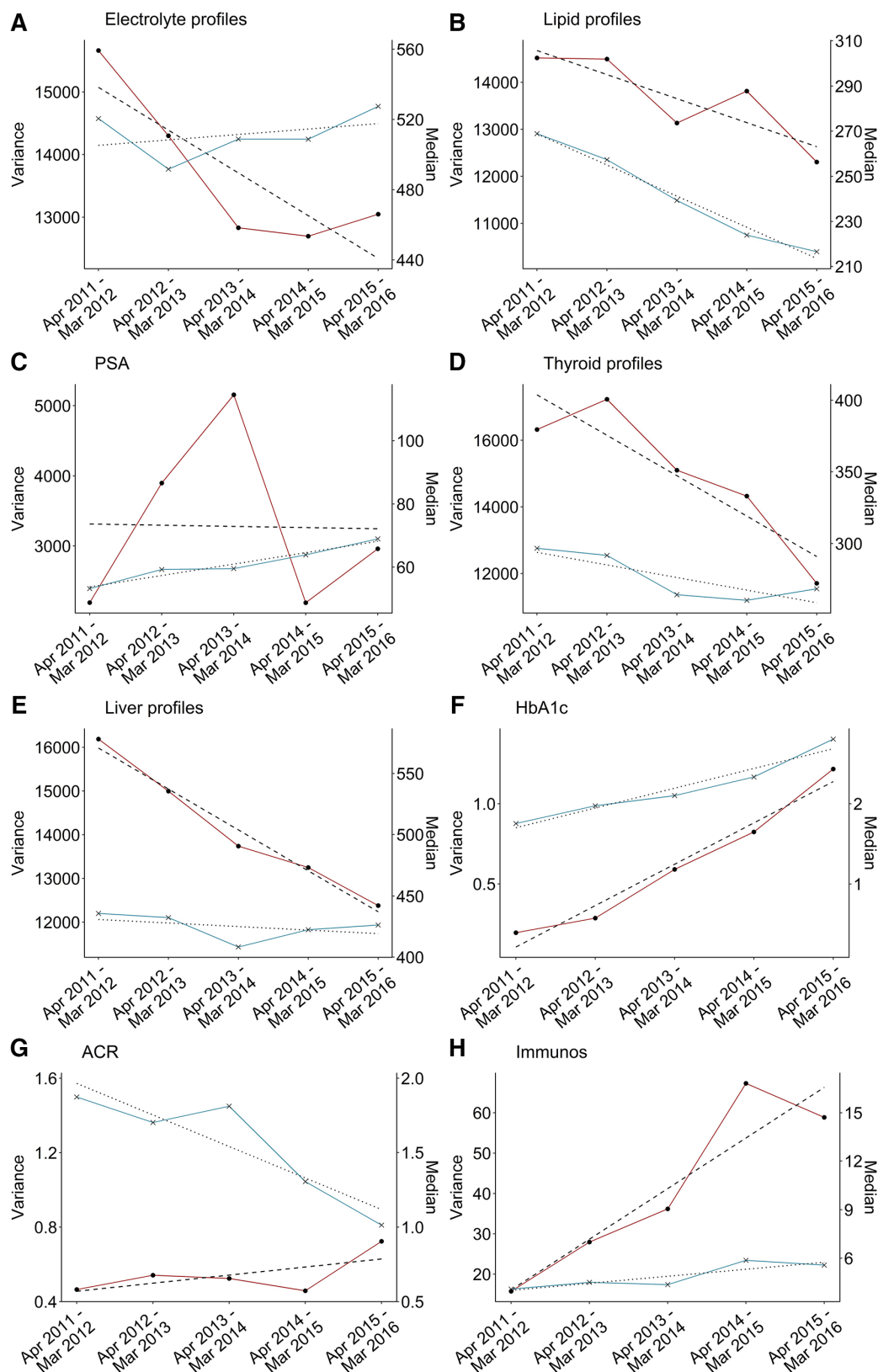


Figure 3 Trend lines for median (blue) and variance (red) of the standardised number of (A) electrolyte profiles, (B) lipid profile, (C) prostate-specific antigen (PSA), (D) thyroid profiles, (E) liver profiles, (F) HbA_{1c}, (G) urine albumin/creatinine ratio (ACR) and (H) immunoglobulins (immunos).

maximum 10 points available under DM009 (range: 6.36–10.0). The relationship between the number of HbA_{1c} tests performed and the GP practice effectiveness, as measured

by the QOF overall achievement rate (combined DM007, DM008 and DM009), was found statistically insignificant ($\tau=0.12$, $p=0.2$).

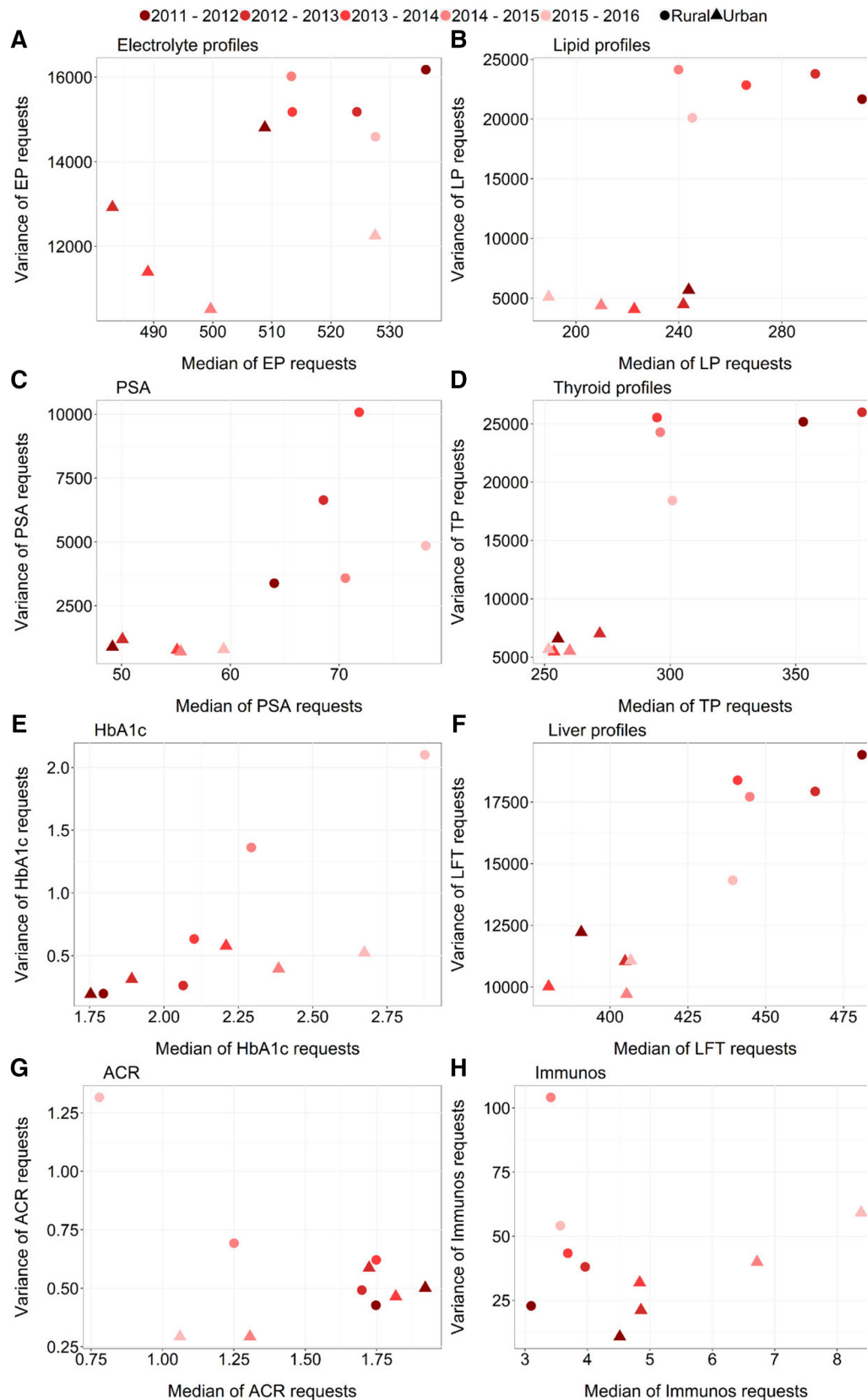


Figure 4 The temporal median–variance relationship of the standardised number of laboratory test requests across the years for (A) electrolyte profiles (EP), (B) lipid profile (LP), (C) prostate-specific antigen (PSA), (D) thyroid profiles (TP), (E) liver profiles (LFT), (F) HbA_{1c}, (G) urine albumin/creatinine ratio (ACR) and (H) immunoglobulins (immunoglobulins). Circle/triangle: rural/urban general practices.

Table 3 The significance of differences in the distribution and variability in test request rates between GP practices located in rural and urban areas

	Electrolyte profiles	Lipid profiles	PSA	Thyroid profiles	HbA1c	Liver profiles	ACR	Immunoglobulins
p_{MWU}								
2011–2012	0.1	0.0004*	0.01	0.009*	0.3	0.03	0.4	0.2
2012–2013	0.08	0.001*	0.03	0.01	0.7	0.02	0.5	0.4
2013–2014	0.8	0.0097*	0.02	0.07	1.0	0.3	0.5	0.4
2014–2015	0.6	0.01	0.02	0.1	0.9	0.3	0.5	0.03
2015–2016	0.7	0.03	0.003*	0.07	0.7	0.3	0.08	0.0096*
p_{FK}								
2011–2012	0.4	0.5	0.03	0.2	0.6	1	0.2	0.8
2012–2013	0.3	0.6	0.01	0.3	0.9	0.4	0.7	0.3
2013–2014	0.2	0.6	0.06	0.03	0.5	0.7	1	0.5
2014–2015	0.9	0.6	0.1	0.02	0.5	0.9	0.3	0.6
2015–2016	0.7	0.7	0.2	0.2	0.2	0.7	0.8	0.1

ACR, albumin/creatinine ratio; GP, general practitioner; PSA, prostate-specific antigen.

p_{MWU} , the Mann Whitney U p-value assessing differences in distributions of laboratory test rates between rural and urban practices.

p_{FK} , the Fligner-Killeen p-value referring to the significance level of differences in variances. In both cases, the Bonferroni-corrected $p < 0.01$ was considered statistically significant (*). Bonferroni-corrected significance cut-off of $\alpha < 0.01$ was established by dividing the family wise error probability $p_{fwe} < 0.05$ by the number of tested associations that is, five measurements over time.

DISCUSSION

We evaluated temporal changes in variability and number of laboratory tests ordered by individual GP practices in the WHSCT. We also investigated a range of key demographic, socioeconomic, geographical and clinical factors to assess whether any of these factors are likely to explain a part of the observed variation in requesting activity.

Interpractice variability in test requests expressed by the index of variability was the lowest for profile tests and the highest for HbA_{1c} and immunoglobulins. This observation was consistent with previous studies.^{10 28 29} In addition, we observed a 23.3% decrease in profile tests (electrolyte, lipid, liver and thyroid profiles combined) between 2011–2012 and 2015–2016 ($p=0.009$, the Bonferroni-adjusted $\alpha < 0.01$ used). This might have been related to the implementation of the demand optimisation intervention in the WHSCT over the 3-year period from April 2012 to March 2015. The intervention was developed to support appropriate use of laboratory testing in the WHSCT and consisted of several elements including educational materials on the benefits of the optimal tests utilisation, information on minimal retest intervals, the review of the test requesting processes and financial incentive. Educational material covered the major clinical indications for a range of most commonly requested tests (including profile tests) and was circulated electronically to all GPs. The educational material was supplemented by face-to-face educational sessions with primary care teams and presentation of data showing the local variability on test requesting rates. Furthermore, all primary care teams were asked to engage in the process of reviewing test requesting procedures within their practice (ie, what staff are allowed to request tests and what is the process for test

requesting), and to reflect on the information provided on their practice test requesting rates and ranking in comparison to other practices. GPs also received a very small incentive to participate in the scrutiny of their requesting processes and activities. The 500% increase in variance of HbA_{1c} ordering rates (statistically significant at the Bonferroni-adjusted $\alpha < 0.01$), on the other hand, may have been caused by between-practice differences in the adoption of HbA_{1c} as a diagnostic test (from 2012) and inconsistent implementation of new guidelines on appropriate rate of diabetes monitoring. This may indicate that recommended guidelines did not predispose GPs to change their perceptions on the value and role of HbA_{1c} test in patient assessment.

The interpractitioner variations in test ordering were found unrelated to demographic and socioeconomic factors. We showed that distributions of patients' age and sex had little impact on ordering rates for a predetermined set of pathology tests. The socioeconomic status of GP practice also did not appear to identify low or high requestors of laboratory tests. No significant relationship between test ordering and age, gender or deprivation measures was reported by other studies.^{6 30–32}

The practice location was found to be a significant factor associated with variability in use of lipid profiles, thyroid profiles, PSA and immunoglobulins. This may reflect the differences in clinical practice associated with the specific aspects of practice organisation and workflow in rural and urban practices as well as personal characteristics of GPs. The practice location was previously related to the between-practice variation in prescribing.^{33–35} Further studies are needed to explore in more detail the potential reasons for the rural–urban discrepancy in test

utilisation, especially that these differences appear to be significant only for some clinical biochemistry tests/test groups.

Finally, we found no evidence of a significant correlation between test requesting rates and clinical outcomes. This observation is consistent with other studies.¹⁰ The fact that the majority of practices attained maximum points for the three target levels for HbA_{1c} (59, 64 and 75 mmol/mol) implies that the glycaemic control was generally good and could not explain the variability in HbA_{1c} testing.

Note that we took a conservative approach in determining the significance cut-off α when evaluating the significance of observations for multiple comparisons (eg, differences in laboratory test requests over time or differences in request rates between considered clinical biochemistry tests/test groups) by applying the Bonferroni correction. We are aware that while the Bonferroni adjustment decreases the risk of type I errors, that is, a probability of a false significant result, this happens at cost of inflating type II errors, that is, the probability of accepting the null hypothesis when the alternative is true. However, by applying this correction technique, we believe we reinforced confidence in the significance of our results.

Implications for practice and research

The findings of our study suggest that there may be considerable potential for the rationalisation of test ordering through minimising the between-practice variability in test utilisation. Yet, it is important to establish whether the observed variation is, and to what extent, associated with over-requesting (unnecessary repeat requesting of tests) in GP practices with high ordering rates or on the contrary, it reflects a failure to prescribe clinically indicated tests by GP practices characterised by low use of tests. Further analysis on the degree and detection of inappropriate use of laboratory resources in primary care could contribute to improving the consistency, efficiency and cost-effectiveness of patient diagnostics, monitoring and treatment and hence, reduce unnecessary costs to patient care.

Further exploration of variations in requesting activity in the context of factors not considered in this study (eg, practitioner-specific characteristics) may help identify and implement appropriate optimisation strategies to manage demand for laboratory tests. Previous studies reported on the number of successful approaches in fostering best practice and reducing variation in test utilisation including implementing locally agreed clinical guidelines, changing test order forms, incorporating clinical decision support tools with embedded retest interval rules, conducting audits of GPs on their request rates and providing financial incentives.^{36–42} In fact, several studies showed that multifaceted interventions were most successful in optimising laboratory demand.^{43 44}

Finally, it is essential to better understand implications of variability in laboratory test requesting for the cost and

quality of care. In particular, an important question to be answered is whether or not the growing costs associated with increased use of laboratory services has led to commensurate benefits to the patient.

Strengths and limitations

Test requesting data were directly extracted from the HSC Business Object XI clinical information system that captures information on tests' use from three clinical chemistry departments of Altnagelvin Area Hospital, Tyrone County Hospital, and South West Acute Hospital. Accordingly, our analysis was based on all available data regarding requesting activity in the Northern Ireland WHSCT in the period of 1 April 2011–31 March 2016 and hence, was not subject to selection bias. Furthermore, our study provides a first comprehensive insight into the use of laboratory tests and factors accounting for the variation in between-practice test utilisation in the WHSCT primary care system.

Due to data unavailability, we were unable to investigate the relationship between laboratory use patterns and practitioner-specific characteristics including GP's age, education and medical training. Such analysis could help identify potential reasons behind variation in clinical practice. Besides the lack of information on the practitioners, the present study was limited by a paucity of research evidence in this area. We were also unable to retrieve consolidated data *on both gender* and age of patients registered in individual GP practices and therefore, assess the combined effect of sex and age distributions on test requesting activity. Furthermore, despite our attempt to maintain consistency by analysing laboratory test requests only from primary care medical practices that remained open throughout the study period, we acknowledge that some GPs in those practices could have been replaced, moved to alternative locations or quit direct patient care during the period of investigation. This could have an impact on the centre-associated requesting rates of laboratory tests. Finally, although the rationale of the study, research objectives, variables, inclusion criteria, outcome measures and statistical methods to be used in our study were identified in advance to guide the analysis of the primary care use of laboratory tests in the WHSCT, the a-priori defined analysis plan was not published. We acknowledge that this could be perceived as a potential limitation of our study.

CONCLUSION

This study investigated the patterns and temporal changes in request rates across a range of frequently ordered laboratory tests. In addition, it explored potential factors of the interpractice variability in the use of laboratory tests and found that differences in requesting activity appear unrelated to either demographic and socioeconomic characteristics of GP practices or clinical outcome indicators. Our results highlight the need for further investigations

to identify other potential factors that may account for the differences in test utilisation between practitioners.

Contributors MB and MO had the original idea for this study. SA led the data collection. MB designed the methodology, performed the analysis and drafted the manuscript. MO, CM and LM contributed to the drafting and critical revision of the manuscript.

Funding This project was supported by the EU's INTERREG VA Programme, managed by the Special EU Programmes Body (SEUPB). The views and opinions expressed in this paper do not necessarily reflect those of the European Commission or the Special EU Programmes Body (SEUPB).

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement No additional data are available.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

REFERENCES

- Newman-Toker DE, McDonald KM, Meltzer DO, *et al*. How much diagnostic safety can we afford, and how should we decide? A health economics perspective. *BMJ Qual Saf* 2013;22:ii11–20.
- Karakusevic S, Edwards N, Lewis R, *et al*. *The future of pathology services*: Nuffield Trust, 2016.
- Smellie WS. Demand management and test request rationalization. *Ann Clin Biochem* 2012;49:323–36.
- Bareford D, Hayling A. Inappropriate use of laboratory services: long term combined approach to modify request patterns. *BMJ* 1990;301:1305–7.
- Chaired by Lord Carter of Coles. An Independent Review for the Department of Health. Report of the Review of NHS Pathology Services in England. 2006.
- Smellie WS, Galloway MJ, Chinn D, *et al*. Is clinical practice variability the major reason for differences in pathology requesting patterns in general practice? *J Clin Pathol* 2002;55:312–4.
- Sá L, Teixeira ASC, Tavares F, *et al*. Diagnostic and laboratory test ordering in Northern Portuguese Primary Health Care: a cross-sectional study. *BMJ Open* 2017;7:e018509.
- Kwok J, Jones B. Unnecessary repeat requesting of tests: an audit in a government hospital immunology laboratory. *J Clin Pathol* 2005;58:457–62.
- Vaidya B, Ukoumunne OC, Shuttleworth J, *et al*. Variability in thyroid function test requests across general practices in south-west England. *Qual Prim Care* 2013;21:143–8.
- O'Kane MJ, Casey L, Lynch PL, *et al*. Clinical outcome indicators, disease prevalence and test request variability in primary care. *Ann Clin Biochem* 2011;48:155–8.
- Smellie WS, Galloway MJ, Chinn D. Benchmarking general practice use of pathology services: a model for monitoring change. *J Clin Pathol* 2000;53:476–80.
- Weydert JA, Nobbs ND, Feld R, *et al*. A simple, focused, computerized query to detect overutilization of laboratory tests. *Arch Pathol Lab Med* 2005;129:1141–3.
- Wong ET, McCarron MM, Shaw ST. Ordering of laboratory tests in a teaching hospital. Can it be improved? *JAMA* 1983;249:3076–80.
- Young DW. Improving laboratory usage: a review. *Postgrad Med J* 1988;64:283–9.
- Epstein AM, McNeil BJ. Relationship of beliefs and behavior in test ordering. *Am J Med* 1986;80:865–70.
- van Diepen S, Bakal JA, Lin M, *et al*. Variation in critical care unit admission rates and outcomes for patients with acute coronary syndromes or heart failure among high- and low-volume cardiac hospitals. *J Am Heart Assoc* 2015;4:e001708.
- Adam MA, Goffredo P, Youngworth L, *et al*. Same thyroid cancer, different national practice guidelines: When discordant American Thyroid Association and National Comprehensive Cancer Network surgery recommendations are associated with compromised patient outcome. *Surgery* 2016;159:41–51.
- Northern Ireland Statistics and Research Agency. Review of the Statistical Classification and Delineation of Settlements. 2015 <http://www.nisra.gov.uk/archive/geography/review-of-the-statistical-classification-and-delineation-of-settlements-march-2015.pdf> (Accessed Aug 2018).
- Northern Ireland Statistics and Research Agency. Northern Ireland Multiple Deprivation Measures 2017. 2017 <https://www.nisra.gov.uk/publications/nimdm17-results> (Accessed Aug 2018).
- Department of Health. Quality and outcomes framework (QOF) achievement data 2015/16. 2016 <https://www.health-ni.gov.uk/publications/quality-and-outcomes-framework-qof-achievement-data-201516> (Accessed Aug 2018).
- Department for Environment, Food & Rural Affairs. Rural Urban Classification. 2016 <https://www.gov.uk/government/collections/rural-urban-classification> (Accessed Aug 2018).
- Sen A, Srivastava M. *Regression analysis: theory, methods, and applications*: Springer Science & Business Media, 2012.
- Jekel JF, Katz DL, Elmore JG, *et al*. *Epidemiology, biostatistics and preventive medicine*: Elsevier Health Sciences, 2007.
- Conover WJ, Johnson ME, Johnson MM. A Comparative study of tests for homogeneity of variances, with applications to the outer continental shelf bidding data. *Technometrics* 1981;23:351–61.
- Wk L, Stanford DA, Yu H. *Advances in Time Series Methods and Applications*: Springer, 2016.
- Gibbons JD, Fielden JD. *Nonparametric measures of association*: Sage, 1993.
- Gupta S, Verma M, Gupta AK, *et al*. Are we using Thyroid Function Tests Appropriately? *Indian J Clin Biochem* 2011;26:178–81.
- Mogyorósy Z, Mogyorósy G. [Practice pattern and geographic variation in test ordering. A literature review]. *Orv Hetil* 2006;147:25–31.
- Larsson A, Palmer M, Hultén G, *et al*. Large differences in laboratory utilisation between hospitals in Sweden. *Clin Chem Lab Med* 2000;38:383–9.
- Selvin E, Steffes MW, Zhu H, *et al*. Glycated hemoglobin, diabetes, and cardiovascular risk in nondiabetic adults. *N Engl J Med* 2010;362:800–11.
- Miyakis S, Karamanof G, Liontos M, *et al*. Factors contributing to inappropriate ordering of tests in an academic medical department and the effect of an educational feedback strategy. *Postgrad Med J* 2006;82:823–9.
- Beckett GJ, Toft AD. First-line thyroid function tests - TSH alone is not enough. *Clin Endocrinol* 2003;58:20–1.
- Bucholc M, O'Kane M, Ashe S, *et al*. Prescriptive variability of drugs by general practitioners. *PLoS One* 2018;13:e0189599.
- Morrison J, Anderson MJ, Sutton M, *et al*. Factors influencing variation in prescribing of antidepressants by general practices in Scotland. *Br J Gen Pract* 2009;59:e25–e31.
- Holm M, Olesen F. Factors affecting prescription of psychotropic drugs in general practice. *Scand J Prim Health Care* 1988;6:169–73.
- Bugter-Maessen AM, Winkens RA, Grol RP, *et al*. Factors predicting differences among general practitioners in test ordering behaviour and in the response to feedback on test requests. *Fam Pract* 1996;13:254–8.
- Baker R, Falconer Smith J, Lambert PC. Randomised controlled trial of the effectiveness of feedback in improving test ordering in general practice. *Scand J Prim Health Care* 2003;21:219–23.
- Buntinx F, Knottnerus JA, Crebolder HF, *et al*. Reactions of doctors to various forms of feedback designed to improve the sampling quality of cervical smears. *Qual Assur Health Care* 1992;4:161–6.
- Bunting PS, Van Walraven C. Effect of a controlled feedback intervention on laboratory test ordering by community physicians. *Clin Chem* 2004;50:321–6.
- Showstack JA, Schroeder SA, Matsumoto MF. Changes in the use of medical technologies, 1972–1977: a study of 10 inpatient diagnoses. *N Engl J Med* 1982;306:706–12.
- Pysher TJ, Bach PR, Lowichik A, *et al*. Chemistry test ordering patterns after elimination of predefined multitest chemistry panels in a children's hospital. *Pediatr Dev Pathol* 1999;2:446–53.
- Martins CM, da Costa Teixeira AS, de Azevedo LF, *et al*. The effect of a test ordering software intervention on the prescription of unnecessary laboratory tests - a randomized controlled trial. *BMC Med Inform Decis Mak* 2017;17:20.
- Oxman AD, Thomson MA, Davis DA, *et al*. No magic bullets: a systematic review of 102 trials of interventions to improve professional practice. *CMAJ* 1995;153:1423–31.
- Grimshaw JM, Russell IT. Effect of clinical guidelines on medical practice: a systematic review of rigorous evaluations. *Lancet* 1993;342:1317–22.