‘Diagnostic downshift’: clinical and system consequences of extrapolating secondary care testing tactics to primary care

Imran Mohammed Sajid ,[1,2] Kathleen Frost,[3] Ash K Paul[4]

10.1136/bmjebm-2020-111629

[1]NHS West London Clinical Commissioning Group, London, UK
[2]University of Global Health Equity, Kigali, Rwanda
[3]NHS Central London Clinical Commissioning Group, London, UK
[4]NHS South West London Health and Care Partnership STP, London, UK

Correspondence to:
Dr Imran Mohammed Sajid,
NHS West London Clinical Commissioning Group, London, London, UK;
imransajid@nhs.net

To cite: Sajid IM, Frost K, Paul AK. *BMJ Evidence-Based Medicine* Epub ahead of print: [please include Day Month Year]. doi:10.1136/bmjebm-2020-111629

ABSTRACT

Numerous drivers push specialist diagnostic approaches down to primary care (‘diagnostic downshift’), intuitively welcomed by clinicians and patients. However, primary care’s different population and processes result in under-recognised, unintended consequences. Testing performs poorer in primary care, with indication creep due to earlier, more undifferentiated presentation and reduced accuracy due to spectrum bias and the ‘false-positive paradox’. In low-prevalence settings, tests without near-100% specificity have their useful yield eclipsed by greater incidental or false-positive findings. Ensuing cascades and multiplier effects can generate clinician workload, patient anxiety, further low-value tests, referrals, treatments and a potentially nocebic population ‘disease’ burden of unclear benefit. Increased diagnostics earlier in pathways can burden patients and stretch general practice (GP) workloads, inducing downstream service utilisation and unintended ‘market failure’ effects. Evidence is tenuous for reducing secondary care referrals, providing patient reassurance or meaningfully improving clinical outcomes. Subsequently, inflated investment in per capita testing, at a lower level in a healthcare system, may deliver diminishing or even negative economic returns. Test cost poorly represents ‘value’, neglecting under-recognised downstream consequences, which must be balanced against therapeutic yield. With lower positive predictive values, more tests are required per true diagnosis and cost-effectiveness is rarely robust. With fixed secondary care capacity, novel primary care testing is an added cost pressure, rarely reducing hospital activity. GP testing strategies require real-world evaluation, in primary care populations, of all downstream consequences. Test formularies should be scrutinised in view of the setting of care, with interventions to focus rational testing towards those with higher pretest probabilities, while improving interpretation and communication of results.

Introduction

Are more tests, earlier in pathways, within primary care helpful? While welcomed by clinicians, patients and policymakers, we explore under-recognised consequences.

Several drivers push specialist diagnostic approaches down to the broader primary care layer of a health system, which we describe as ‘diagnostic downshift’. Aspirations for earlier disease detection or capacity pressures in specialist and cancer pathways underlie shifting of tests from high-cost hospital settings to primary care. Such assumptions have led to procurement and growth of GP diagnostics, with unfettered direct access to physiology tests, endoscopy, ultrasound, MRI, CT, and biochemical and immunological tests, often only evaluated in secondary care. It is increasingly expected, sometimes mandated, for GPs to perform secondary care-based testing strategies prior to referral. Earlier diagnostics are presumed to accelerate patient journeys (eg, decision-making at first outpatient appointment) or reduce referrals by empowering GPs. However, diagnostic growth, heavily cited for low-value overuse,[1] has unintended consequences. In the COVID-19 pandemic context, reduced hospital access and increasing virtual consultations may proliferate community testing.

Clinicians intuitively welcome tests, bolstering autonomy and professional confidence, while playing into patients’ biases. Medical risks should be discussed, yet test inaccuracy and cascades go unrecognised,[2] while clinicians poorly interpret results.[3 4] Tests can guide management, but also generate anxiety, low-value disease labels, fear avoidance behaviours, further investigations, referrals and treatment cascades of little benefit.[5–7] Consider vitamin D, recommended only in select patients, now ubiquitously screened, with almost 100-fold increases, reflecting massive costs, time and prescribing of spurious value.[8]

Testing in primary care is different from secondary care

Tests are integral to primary care, with overlapping symptoms between benign and serious conditions and uncertainty in up to 40% of consultations.[9] However, secondary care has different populations, workflows and expertise; thus, diagnostic tactics should not be blindly extrapolated. General practice performs a technical but also wider psychosocial role for non-specific presentations, with a more person-orientated, than pathophysiology-orientated focus.[10] While uncertainty management depends on psychological factors,[11 12] there tends to be diminishing decision-making value from additional tests.[13]

Diagnostic downshift's pretest considerations (table 1) include inflated (inappropriate) tests per capita, test indication creep and inadvertent screening. Suspicion to trigger a test will typically be lower than that for a referral. Thus, more patients are tested than otherwise referred. Post-test dynamics (table 2) include altered performance (false-positive paradox and spectrum bias) with lower positive predictive values, greater false-positive rates, a burden of incidental findings, misinterpretation problems (particularly for serial testing), limited reassurance and 'multiplier effects' of low-value cascades.

Diagnostic sensitivity trades off against specificity. Compared with secondary care encounters, GPs do not require immediate high-sensitivity testing tactics, as patients can be referred onwards for evaluation, as well as readily reattend for persisting or worsening symptoms. Test specificity is more critical to manage referral appropriateness. In low-prevalence settings, without near-100% specificity, benefit (diagnostic yield) is eclipsed by greater false-positives or incidental findings (see table 2). With pretest probability <10% (common for primary care), even with 90% specificity, Bayesian analysis shows greater false-positives than true-positives, with positive predictive value no better than a coin toss. For example, carotid artery ultrasound screening, with 92% specificity, across 100 000 patients, generates 7920 false-positives versus only 940 true-positives.[14]

High-quality studies rarely demonstrate benefit from advanced testing in primary care

Despite disseminated use in different populations, for wider indications, traditionally specialist tests are rarely robustly evaluated in primary care, reliant on haphazard postmarket surveillance, such as audits. Referral reduction is often based on self-report without capturing downstream utilisation and typically lacks usual care comparator analysis. The few randomised controlled trials, such as knee MRI, hysterosalpingography for fertility or low-dose CT for lung cancer, fail to demonstrate meaningful impact.[15–17] MRI access does not reduce orthopaedic referrals nor clinically benefit patients.[17–19] Systematic review shows little to no high-quality evidence of clinical or cost benefits to support increasing tests in primary or community settings, with only low-quality evidence of reducing referrals, suggesting such diagnostic strategies may be more politically motivated.[20]

Earlier testing does not necessarily improve cancer outcomes

Rhetoric around cancer detection system delays often drives diagnostic expansion.[21 22] However, there is a paucity of evidence that advanced GP testing improves outcomes (survival rate statistics are misleading due to lead time, length bias or overdiagnosis of indolent disease). Impact of 'delayed diagnosis' is mixed, including the so-called 'waiting-time-paradox' ('delay' associated with improved outcome for some cancers).[23–25] Diagnostic strategy, particularly for low-but-not-no-risk presentations, is complex. Systematic review of GP direct access testing suggests, although time-to-test may improve, there is no change in time-to-diagnosis or outcomes.[26] Novel pathway triage may enable prehospital diagnostics, although it has its own drawbacks; in 'straight-to-test' pathways, alternate diagnostics may have been preferred by specialists.[27] Furthermore, pre-referral laboratory cancer tests are broadly unreliable, including many biomarkers.[28] Without evidence, novel technologies should be cautioned, considering already pressured workloads.

Outside of low-income countries, there is little evidence to suggest increased diagnostic direct access resolves the problem of misdiagnosis in primary care, which is due to a myriad of factors, including cognitive reasoning errors.[29 30]

Expansive testing in primary care creates a population 'disease' burden of unclear benefit (overdiagnosis)

Many expanding disease definitions, increasingly determined by test results rather than symptoms, have questionable impact on outcomes, for example, polycystic ovarian syndrome, pre-diabetes

Table 1 Pretest considerations relevant in primary care

Description	Evidenced examples
Inappropriate/unnecessary tests per capita	
Representing 90% of health service encounters,[83] increased GP testing inflates (inappropriate) tests per capita. Majority of test volumes are now primarily ordered by GPs, not specialists. Guidelines, often opinion-based,[84] are not always the best indicator of 'appropriateness', offering testing at such low prevalence that impact is rare, for example, echocardiography.[85]	Time or patient pressures, limited specialist capacity, uncertainty intolerance, defensive practice, political factors and cognitive biases all drive low-value testing. A third of GP consultations result in laboratory testing. A third of laboratory tests,[64] 98% of knee[86] and half of hip radiographs,[87] half of endoscopies,[83] 66% of internal auditory MRIs,[88] upto 95% of musculoskeletal MRIs,[18] and 94% of spinal CTs in primary care appear inappropriate.[89]
Rate of indication creep	
Earlier, undifferentiated GP presentations, with siloed practice, engender variation and 'indication creep' (testing outside recommendations). Test growth creates choice overload and decisional fatigue for generalists, while diagnostic training is largely absent from curricula. Risk is not normally distributed; most have below-average risk.[90] Yet clinicians gravitate towards overtesting due to time pressures and biases,[91] such as action bias. Persuading patients to accept evidence-based testing recommendations is challenging.[92 93]	There is up to 100% variation in non-guideline GP testing.[83] Primary care presentations require fewer tests than secondary care, for example, only four tests for recent fatigue.[94] GP faecal calprotectin use as a screening, rather than a rule-out test, has grown.[95] GP CT diagnostic yield for renal calculi is 7% vs >44% in secondary care.[96] GP endoscopies similarly have lower diagnostic yield.[97]
Inadvertent screening	
Demand can result in inadvertent screening,[93] not meeting the Wilson criteria, without adequate consent. Inequalities can widen from unapproved screening of low-risk individuals.	A full blood count is commonly requested without indication, with numerous indices where statistical but clinically irrelevant abnormalities are common.[98] High-cost case-finding (screening) GP schemes, such as UK's health check, lack supporting evidence.[99]

GP, general practice.

Table 2 Post-test considerations relevant in primary care

Description	Evidenced examples
Higher false-positive rates (false-positive paradox) and reduced test performance (spectrum bias)	
Lower pretest probabilities in primary care generate lower positive predictive values.[100] Tests without near-100% specificity, in low-prevalence environments, generate innumerable false-positives, known as the *'false-positive-paradox'* or *base-rate fallacy*. Sensitivity and specificity are not constant across settings, varying from meta-analytical averages.[101] Comorbidities, prevalence and severity affect performance, known as, *'spectrum bias'*.[102] Secondary care test accuracy is therefore not applicable to primary care.	58% of abnormal GP laboratory results may be false-positives.[103] Prevalence changes performance up to 40%.[100] Sensitivity and specificity of faecal calprotectin in secondary care were 93% and 94%, yet as low as 80% and 67% in primary care.[104 105] The false-positive paradox limits accuracy of the same diagnostic tool across settings in a system, such as qSOFA for sepsis or NEWS for clinical deterioration.[106]
Interpretation challenges	
Few results are dichotomous and clinicians overestimate positive predictive values.[3 4] Even common tests such as lipids or HbA1c require tools (for example, QRISK® and QDiabetes®), incorporating pretest factors, to translate risk, which are usually neglected.[107]	Recommended immunoglobulin tests cause confusion for almost all GPs.[108] GPs may inadequately interpret upto 90% of lipids, causing overtreatment and undertreatment.[107 109] GPs correctly interpret only 17% and mismanage 65% of Musculoskeletal MRI results,[18] associated with worse outcomes.[110 111]
Incidental findings	
Diagnostic downshift generates a burden of, mostly benign,[112] incidental findings, causing GP anxiety and workload.[113] Management may be more challenging and inconsistent for non-specialists,[114] requiring further input.	19% of chest radiographs,[115] 22% of brain MRIs,[112] 37% of renal CTs,[116] 26% of emergency abdominal ultrasounds,[117] 87% of musculoskeletal MRIs[18] and 67% of neck ultrasounds show incidental findings.[118] Studies on follow-up burden are limited.[119 120] While 200 MRIs are required to identify one acoustic neuroma, one in six suffer overdiagnosis cascades.[88]
Interval testing and reference change values	
Serial testing is common due to greater patient access and (largely unevidenced-based)[84] disease monitoring. Under-recognised test 'reference change values' (RCVs) describe normal biological and analytical variation.[121] Many RCVs (eg, liver enzymes, cholesterol, free thyroxine, etc) are ~20%.[121 122] Normal test-to-test variation may be inappropriately acted upon, for example, with dose changes.	Testing occurs more frequently than recommended, for example, 22% of HbA1c tests are repeated prematurely,[65] and 60% of cholesterol tests are repeats, 70% of which unnecessary.[123] Bone density scans are ordered frequently, despite annual change being lower than the scan's analytical error.[124] Analytical variation is even greater for unstable samples transported from primary care, for example, spurious hyperkalaemia.[125 126]
Lack of real-time feedback to correct illusory correlations	
As signal-to-noise declines with low prevalence, individuals can increasingly overidentify false targets.[127] Illusory correlation/causation between false-positive or incidental findings and patient symptoms may cause belief reinforcement of testing behaviours. Without real-time feedback, GPs are unable to refine heuristics.	Respiratory auscultation's poor accuracy inappropriately influences prescribing.[128] Most musculoskeletal MRI patients are surgically referred for clinically irrelevant findings.[18] Vitamin D is erroneously associated with non-specific complaints.[129–131]
Follow-up discontinuity	
While some GPs hold expertise for certain tests, colleagues, including administrative staff, may relay results to patients, out of context to clinical history and expectations.	Roughly half of musculoskeletal MRI results are conveyed by staff other than the requesting clinician familiar with the presentation.[18]
Cascade and multiplier effects	
Low-value test cascades (further tests, referrals, overtreatment) are highly prevalent.[5 6] Primary care's system value can be undermined by new early pathway activity generating larger downstream costs, known as, 'multiplier effects'. [132] Such cascades are poorly recognised.	GP inflammatory markers, or musculoskeletal and auditory MRIs generate expansive cascade costs from low-value findings, often greater than test costs. [18 88 111 133] Cross-sectional imaging in particular has cost consequences.[134 135] Cascades include complications; for example, in lung screening, 23% of false-positive investigated patients suffered complications.[136 137]
Spurious reassurance	
Up to 40% of GP patients have medically unexplained symptoms. Incidental findings and high false-positive rates undermine testing for reassurance.[38] Test overuse may shift focus away from unmet psychosocial needs, without resolving ongoing symptoms, which likely require additional support.	Systematic reviews show tests alone contribute little to reassurance[138 139]; for example, neuroimaging provides no sustained reassurance for headache.[140]

GP, general practice; HbA1c, haemoglobin A1c; NEWS, national early warning score; qSOFA, quick sequential organ failure assessment.

or dementia, . Not all diagnoses provide resultant 'therapeutic yield' (positive change in management). Increased per capita provision from diagnostic downshift, combined with expanding disease definitions, yields a greater population 'disease' burden (figure 1) of unclear utility.

For example, diagnostic access has resulted in a 'cancer' epidemic in high-income nations, without concurrent cancer mortality change.[31] Liver 'disease' shows similar growth, without mortality increase.[32] Tests are often 'abnormal' (whereby normal typically refers to 'bio-statistically average'), regularly identifying indolent, incidental 'disease', particularly with an ageing population. Examples include early-stage hypertension, lumbar disc protrusions, joint changes, small-vessel disease, hypogonadism, hypercholesterolaemia, osteopaenia, raised liver enzymes, renal cysts, vitamin D insufficiency and so on. Such endless asymptomatic 'disease' often has minimal clinical relevance. Laudable aspirations for earlier treatment or lifestyle change are often unevidenced. For example, diabetes or hypertension

Figure 1 Diagnostic market growth and disease burden.

labels deliver marginal-to-zero behaviour change,[33–35] while increasing anxiety.[36] Diagnostic unintended harms include physical, psychological, social, financial and treatment burden and healthcare dissatisfaction.[37] As a doctor–patient 'gesture', appeasing emotional needs,[38] tests' unintended psychological consequences can endure for years.[39 40] Medicalising labels can be nocebic, broadly reducing self-reported health,[41] raising anxiety, perceived severity and preference towards more invasive management.[42] Expansive GP testing runs counter to the World Organization of Family Doctors' strategic priority to address low-value overdiagnosis.[43]

Diagnostic downshift is rarely cost-effective

Diagnostic downshift is seen as a way to reduce costs, yet the opposite is often true. While 'value' lies in outcomes, policymakers focus on test unit costs and productivity. Recommended large supplier economies of scale,[44] based around 'technical efficiency', inflate volume, with paradoxical effects on 'population value'. The Donabedian curve[45] highlights earlier diminishing returns from productivity increases in medical technologies with a given rate of harm (figure 2). Medical technologies confer effectiveness for specific patients or settings (for diagnostics, often those with higher pretest probability of disease) and become decreasingly helpful when applied wider. As test accuracy decreases in lower-acuity settings (see table 2), harm rate increases. While diagnostic downshift may benefit some, this is eclipsed by greater false-positives or overdiagnosis, causing psychological harm, further healthcare utilisation and low-value overtreatment. Downstream congestion of services can even negatively impact access for those beneficially diagnosed earlier.

Diagnostic downshift incurs costs as commissioners procure supply beyond current capacity, potentially fragmenting care across additional providers, often without clinical or information technology integration, which can contribute to repeat testing. Supply-induced demand must be considered. Minimal barriers to access can induce utilisation, making timely diagnostics harder for sicker patients.

Traditional cost-per-QALY (quality-adjusted life years) analysis may not reflect affordability for high-volume tests, nor account for indication creep and inaccuracy (see tables 1 and 2). With lower prevalence, more tests are required per diagnosis in primary care and lower 'diagnostic yield' is less likely to meet cost-effectiveness thresholds. Economic models rarely include outcome data, societal costs or false-positive effects,[46 47] relying on opinion with considerable uncertainty.[48] Indolent diagnoses and cascades generate low-value utilisation across services. In a 'rival market', with limited capacity, this reduces resources for others. Such negative externalities (costs imposed on third parties, eg, secondary care) are intangible in clinical encounters, as clinicians gravitate towards overtesting, nor to diagnostic providers, unreflected in their (increasingly cheaper) test costs. Diagnostic economies of scale, expanding per capita testing, thereby risk market failure (inefficient distribution of limited resources for a socially optimal outcome) across a healthcare economy.

Particularly with diagnostic accuracy publication biases,[49 50] test evaluation should adopt frameworks beyond just accuracy, including practical considerations, diagnostic and therapeutic yields, psychological factors and outcomes relevant to patients.[51 52] Assessing accuracy in low-prevalence settings presents challenges[53]; however, real-world 'technology management' cannot be neglected.[54] Cost consequence analysis, incorporating indication creep and all disaggregated cascades, may help policymakers appraise diagnostic downshift.[55] Providers can assist in capturing outcomes to evidence their 'value'.

While diagnostics may reduce some 'inappropriate' referrals, hospital capacity is a fixed cost, where freed capacity is consumed by other demand. Without concurrently reducing hospital capacity, the additional cost burden of primary care diagnostics rarely releases savings. Furthermore, if GP testing allows greater first-outpatient decision-making (largely presumptive), this increases hospital throughput, potentially increasing spending.

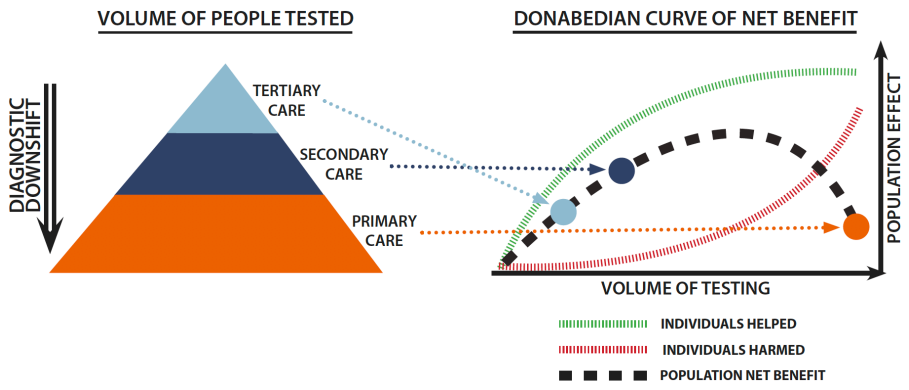

Figure 2 Diminishing 'population value' with diagnostic downshift.

Consider liver guidelines. Specialist investigations for abnormal enzymes are hugely expensive for GPs, with lower yield of rare conditions, highlighting differences between generalist and specialist testing agendas.[56] Antinuclear antibodies, neither sensitive nor specific, generate 99.9% false-positive rates, prone to misinterpretation.[57 58] Novel GP ELF™ and FibroScan® testing inflates spending by tens of millions with no evidenced outcome benefit.[59] While perceived referral 'appropriateness' may improve, there is no significant reduction in referral volume, risking increased referrals.[60]

Diagnostic cost growth must be evaluated against competing priorities. For conditions related to wider determinants, such as obesity, hypertension, diabetes, liver disease or even musculoskeletal conditions, public health investment may be more impactful.[61] For example, haemoglobin A1c-based diabetes prevention programmes are a questionable use of scarce GP resource, for mass detection of predisease, compounded by insufficient accuracy.[62]

Diagnostics represent profitability in healthcare, particularly for readily expandable tests. Activity-based payments provide no volume control incentive for providers and the market is unlikely to self-regulate supply. With consistent 5%–10% annual growth,[63] without improved outcomes, payment structures encouraging demand management (eg, capitated budgets, or inclusion within fixed appointment tariffs) should be considered. Pathway position can affect cost; in the UK hospital pathology testing is often included within outpatient tariffs (discouraging overtesting). While local arrangements vary, GP pathology can be separately billed, inflating costs, undermining primary care's historically low-cost, high-value.

How do we optimise primary care testing?

Interventions often fail to sustain long-term improved requesting.[64 65] 'Nudging',[66] through choice architecture or default bundles in electronic ordering, shows promise.[67–74] Decision aid impact is still unclear.[75]

Dichotomous cut-offs (eg, for C-reactive protein, faecal calprotectin, prostate-specific antigen, etc) fail to consider uncertainty. Segmenting results into post-test risk categories,[76] as well as further advice in reports,[77 78] may help. Probability tools, incorporating pretest factors, may support uncertainty communication to improve shared decision-making.[79]

Testing policies should comment on the setting of care. For example, the National Institute for Health and Care Excellence recommended lumbar MRI only occur in specialist settings.[80] Joint prostheses blood metal testing has complex interpretation; thus guidelines resisted shifting responsibility to GPs.[81]

There are of course scenarios of undertesting, where diagnostic downshift represents value. High prevalence (higher pretest probability) maintains performance; for example, malaria testing is invaluable in certain regions. For housebound patients, point-of-care testing may provide benefit in trained hands. Natriuretic peptide testing for heart failure, with primary care-evaluated cut-offs, potentially reduces delayed diagnosis and hospital admissions (although interestingly, not mortality).[82]

Unintended consequences of earlier test access also apply to direct-to-consumer tests and wearable technologies, which should undergo stringent scrutiny.

Summary

Problems

- Diagnostic downshift reduces test performance. In low-prevalence environments, greater equivocal, false-positive, incidental or difficult-to-interpret results generate anxiety, further investigations, referrals, nocebic disease labels and low-value treatments, with associated costs and workload. Test cascades can be magnified earlier in pathways for larger populations.
- Moving more tests down to primary care, task shifting out of hospital for efficiency, risks overmedicalising a broader population across a health system's lower level. While some benefit, more may be harmed.
- GP diagnostics, an additional cost pressure, rarely release savings. Occasionally, earlier testing represents 'value'; however, this tends to be the exception. High-quality evidence broadly does not support diagnostic downshift to improve outcomes or cost-effectively reduce referrals. This has implications for GP education, often based on specialist approaches.

Solutions

- Pretest interventions (eg, choice architecture, rationalised default test bundles and a scrutinised GP diagnostic catalogue) focusing appropriate testing to those with high pretest probability, as well as post-test interventions to improve interpretation, may minimise harm and improve access for those most likely to benefit.
- Research priority areas include uncertainty communication and risk-based decision-making. Diagnostic suppliers can support real-world outcome measurement for provision of optimal therapeutic yield for a test in a given setting, with minimal harm, to justify costs.
- Diagnostic economics require better understanding, beyond parochial focus on test price, which poorly reflects system costs. Supply-side volume controls and payment structures can avoid per capita overprovision, which generates low-value population 'disease' burden and negative externalities, that is, 'market failure'.

Twitter Imran Mohammed Sajid @imransajid

Contributors Authorship includes expertise in primary care, public health commissioning, diagnostic pathway development and contract management. IMS and KF have led on NHS sector-wide diagnostic optimisation, procurement and quality improvement work for the past 4 years, which underpinned the learning for this article. Stakeholder engagement has included commissioners, GPs, specialists, laboratory and radiology staff, as well as patient representatives via several forums, echoing commonly under-recognised recurrent themes. IMS, general practitioner, adjunct lecturer in health policy and clinical commissioner, drafted the manuscript, which was reviewed and edited by coauthors KF and AKP. KF is a commissioning programme lead, previously a physiotherapist and pathology manager. AKP is a consultant in public health also working in commissioning of services, as well as the NHS national evidence-based interventions programme, who provided further evidence review for the article. IMS is responsible for the overall content and is the primary correspondent.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

ORCID iD

Imran Mohammed Sajid http://orcid.org/0000-0002-0332-9704

References

1 **Morgan DJ**, Dhruva SS, Coon ER. Update on medical overuse. *JAMA Intern Med* 2019.
2 **van Bokhoven MA**, Pleunis-van Empel MCH, 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.
3 **Whiting PF**, Davenport C, Jameson C, *et al*. How well do health professionals interpret diagnostic information? A systematic review. *BMJ Open* 2015;**5**:e008155.
4 **Morgan DJ**, Pineles L, Owczarzak J. Accuracy of practitioner estimates of probability of diagnosis before and after testing. *JAMA Intern Med* 2021.
5 **Deyo RA**. Cascade effects of medical technology. *Annu Rev Public Health* 2002;**23**:23–44.
6 **Ganguli I**, Simpkin AL, Lupo C, *et al*. Cascades of care after incidental findings in a US national survey of physicians. *JAMA Netw Open* 2019;**2**:e1913325.
7 **Ganguli I**, Lupo C, Mainor AJ, *et al*. Assessment of prevalence and cost of care cascades after routine testing during the Medicare annual wellness visit. *JAMA Netw Open* 2020;**3**:e2029891.
8 **Rockwell M**, Kraak V, Hulver M, *et al*. Clinical Management of Low Vitamin D: A Scoping Review of Physicians' Practices. *Nutrients*2018;**10**:493.
9 **Koch H**, van Bokhoven MA, ter Riet G, *et al*. What makes general practitioners order blood tests for patients with unexplained complaints? A cross-sectional study. *European Journal of General Practice* 2009;**15**:22–8.
10 **Rosser WW**. Approach to diagnosis by primary care clinicians and specialists: is there a difference? *J Fam Pract* 1996;**42**:139–44 http://www.ncbi.nlm.nih.gov/pubmed/8606303
11 **Meyer AND**, Giardina TD, Khanna A. Pediatric clinician perspectives on communicating diagnostic uncertainty. *Int J Qual Heal Care* 2020.
12 **Begin AS**, Hidrue M, Lehrhoff S. Factors associated with physician tolerance of uncertainty: an observational study. *J Gen Intern Med* 2021:1–7.
13 **Johnson HA**. Diminishing returns on the road to diagnostic certainty. *JAMA* 1991;**265**:2229.
14 **Jonas DE**, Feltner C, Amick HR, *et al*. Screening for asymptomatic carotid artery stenosis: a systematic review and meta-analysis for the U.S. preventive services Task force. *Ann Intern Med* 2014;**161**:336–46.
15 **Wilkes S**, Murdoch A, Steen N, *et al*. Open access tubal aSsessment for the initial management of infertility in general practice (the oats trial): a pragmatic cluster randomised controlled trial. *Br J Gen Pract* 2009;**59**:329–35.
16 **Guldbrandt LM**, Fenger-Grøn M, Rasmussen TR, *et al*. The effect of direct access to CT scan in early lung cancer detection: an unblinded, cluster-randomised trial. *BMC Cancer* 2015;**15**:934.
17 **Andronis L**, Atwell C, Brealey S. Effectiveness of GP access to magnetic resonance imaging of the knee: a randomised trial. *Br J Gen Pract* 2008;**58**:767–74.
18 **Sajid I**, Parkunan A, Frost K. Unintended consequences: an evaluation quantifying benefits, iatrogenic harms and downstream cascade costs of musculoskeletal magnetic resonance imaging (MRI) in UK primary care. SSRN, 2020. Available: https://papers.ssrn.com/abstract=3735015 [Accessed 21 Nov 2020].
19 **van Oudenaarde K**, Swart NM, Bloem JL, *et al*. General practitioners referring adults to MR imaging for knee pain: a randomized controlled trial to assess cost-effectiveness. *Radiology* 2018;**288**:170–6.
20 **Chambers D**, Booth A, Baxter SK, *et al*. Evidence for models of diagnostic service provision in the community: literature mapping exercise and focused rapid reviews. *Health Serv Deliv Res* 2016;**4**:1–362.
21 **Harris M**, Thulesius H, Neves AL, *et al*. How European primary care practitioners think the timeliness of cancer diagnosis can be improved: a thematic analysis. *BMJ Open* 2019;**9**:e030169.
22 **Rose PW**, Rubin G, Perera-Salazar R, *et al*. Explaining variation in cancer survival between 11 jurisdictions in the International cancer benchmarking partnership: a primary care vignette survey. *BMJ Open* 2015;**5**:e007212.
23 **Neal RD**, Tharmanathan P, France B, *et al*. Is increased time to diagnosis and treatment in symptomatic cancer associated with poorer outcomes? systematic review. *Br J Cancer* 2015;**112**:S92–107.
24 **Lopez-Cedrún JL**, Otero-Rico A, Vázquez-Mahía I, *et al*. Association between hospital interval and survival in patients with oral cancer: a waiting time paradox. *PLoS One* 2019;**14**:e0224067.
25 **Kruger S**, Schirle K, Haas M, *et al*. Prolonged time to treatment initiation in advanced pancreatic cancer patients has no major effect on treatment outcome: a retrospective cohort study controlled for lead time bias and waiting time paradox. *J Cancer Res Clin Oncol* 2020;**146**:391–9.
26 **Smith CF**, Tompson AC, Jones N. Direct access cancer testing in primary care: a systematic review of use and clinical outcomes. *Br J Gen Pract* 2018:1–10.
27 **Aljarabah MM**, Borley NR, Goodman AJ, *et al*. Referral letters for 2-week wait suspected colorectal cancer do not allow a 'straight-to-test' pathway. *Ann R Coll Surg Engl* 2009;**91**:106–9.
28 **Watson J**, Mounce L, Bailey SER, *et al*. Blood markers for cancer. *BMJ* 2019;**361**:l5774.
29 **Kostopoulou O**, Delaney BC, Munro CW. Diagnostic difficulty and error in primary care--a systematic review. *Fam Pract* 2008;**25**:400–13.
30 **Singh H**, Schiff GD, Graber ML, *et al*. The global burden of diagnostic errors in primary care. *BMJ Qual Saf* 2017;**26**:484–94.
31 **Welch HG**, Fisher ES, Income FES. Income and cancer overdiagnosis — when too much care is harmful. *N Engl J Med Overseas Ed* 2017;**376**:2208–9.
32 **Rowe IA**. Too much medicine: overdiagnosis and overtreatment of non-alcoholic fatty liver disease. *Lancet Gastroenterol Hepatol* 2018;**3**:66–72.
33 **Chong S**, Ding D, Byun R, *et al*. Lifestyle changes after a diagnosis of type 2 diabetes. *Diabetes Spectr* 2017;**30**:43–50.
34 **Hackett RA**, Moore C, Steptoe A, *et al*. Health behaviour changes after type 2 diabetes diagnosis: findings from the English longitudinal study of ageing. *Sci Rep* 2018;**8**:16938.
35 **Ineke Neutel C**, Campbell NRC. Changes in lifestyle after hypertension diagnosis in Canada. *Can J Cardiol* 2008;**24**:199–204.
36 **Hamer M**, Batty GD, Stamatakis E, *et al*. Hypertension awareness and psychological distress. *Hypertension* 2010;**56**:547–50.
37 **Korenstein D**, Chimonas S, Barrow B, *et al*. Development of a conceptual map of negative consequences for patients of overuse of medical tests and treatments. *JAMA Intern Med* 2018;**178**:1401–7.
38 **Watson J**, de Salis I, Banks J, *et al*. What do tests do for doctors? A qualitative study of blood testing in UK primary care. *Fam Pract* 2017;**34**:735–9.
39 **Brodersen J**, Siersma VD. Long-Term psychosocial consequences of false-positive screening mammography. *The Annals of Family Medicine* 2013;**11**:106–15.
40 **Darlow B**, Dowell A, Baxter GD, *et al*. The enduring impact of what clinicians say to people with low back pain. *Ann Fam Med* 2013;**11**:527–34.
41 **Jørgensen P**, Langhammer A, Krokstad S, *et al*. Diagnostic labelling influences self-rated health. A prospective cohort study: the HUNT study, Norway. *Fam Pract* 2015;**32**:492–9.
42 **Nickel B**, Barratt A, Copp T, *et al*. Words do matter: a systematic review on how different terminology for the same condition influences management preferences. *BMJ Open* 2017;**7**:e014129.
43 **WONCA**. Overdiagnosis and action to be taken – position paper 2018 | WONCA Europe, 2018. Available: https://www.woncaeurope.org/kb/overdiagnosis-and-action-to-be-taken---position-paper-2018 [Accessed 5 May 2020].
44 **Lord Carter of Coles**. Report of the review of NHS pathology services in England, 2006. Available: https://www.networks.nhs.uk/nhs-networks/peninsula-pathology-network/documents/CarterReviewPathologyReport.pdf [Accessed 22 Sep 2020].
45 **Watson J**, Salisbury C, Jani A, *et al*. Better value primary care is needed now more than ever. *BMJ* 2017;**359**:j4944.
46 **Sutton AJ**, Breheny K, Deeks J, *et al*. Methods used in economic evaluations of chronic kidney disease testing - A systematic review. *PLoS One* 2015;**10**:e0140063.

47 **Verbakel JY**, Turner PJ, Thompson MJ, *et al*. Common evidence gaps in point-of-care diagnostic test evaluation: a review of horizon scan reports. *BMJ Open* 2017;**7**:e015760.
48 **Mistry H**, Mason J. Diagnostic assessment reviews: is cost-effectiveness analysis helpful or necessary? *J Health Serv Res Policy* 2018;**23**:222–42.
49 **Treanor L**, Frank RA, Cherpak LA, *et al*. Publication bias in diagnostic imaging: conference Abstracts with positive conclusions are more likely to be published. *Eur Radiol* 2020;**30**:2964–72.
50 **Treanor LM**, Frank RA, Atyani A, *et al*. Reporting bias in imaging diagnostic test accuracy studies: are studies with positive conclusions or titles submitted and published faster? *American Journal of Roentgenology* 2021;**216**:225–32.
51 **Ferrante di Ruffano L**, Hyde CJ, McCaffery KJ, *et al*. Assessing the value of diagnostic tests: a framework for designing and evaluating trials. *BMJ* 2012;**344**:e686.
52 **Horvath AR**, Lord SJ, StJohn A, *et al*. From biomarkers to medical tests: the changing landscape of test evaluation. *Clin Chim Acta* 2014;**427**:49–57.
53 **Holtman GA**, Berger MY, Burger H, *et al*. Development of practical recommendations for diagnostic accuracy studies in low-prevalence situations. *J Clin Epidemiol* 2019;**114**:38–48.
54 **Bryan S**, Mitton C, Donaldson C. Breaking the addiction to technology adoption. *Health Econ* 2014;**23**:379–83.
55 **Drummond M**, Sculpher M, Claxton K. Methods for the economic evaluation of health care programmes. fourth edition. Oxford university press, 2015. Available: https://global.oup.com/academic/product/methods-for-the-economic-evaluation-of-health-care-programmes-9780199665884?cc=gb&lang=en [Accessed 11 May 2020].
56 **Huddy J**. Writing liver function test guidelines: how hard can it be? *Br J Gen Pract* 2016;**66**:426.
57 **Avery TY**, van de Cruys M, Austen J, *et al*. Anti-Nuclear antibodies in daily clinical practice: prevalence in primary, secondary, and tertiary care. *J Immunol Res* 2014;**2014**:401739.
58 **Zeman MV**, Hirschfield GM. Autoantibodies and liver disease: uses and abuses. *Can J Gastroenterol* 2010;**24**:225–31 https://pubmed.ncbi.nlm.nih.gov/20431809/
59 **Jarvis H**, Hanratty B. Detecting liver disease in primary care: are we ready for change? *Br J Gen Pract* 2017;**67**:202–3.
60 **Srivastava A**, Gailer R, Tanwar S, *et al*. Prospective evaluation of a primary care referral pathway for patients with non-alcoholic fatty liver disease. *J Hepatol* 2019;**71**:371–8.
61 **Martin S**, Lomas JRS. Is an ounce of prevention worth a pound of cure? : Estimates of the impact of English public health grant on mortality and morbidity. Centre for Health Economics, University of York, 2019. Available: http://eprints.whiterose.ac.uk/148342/ [Accessed 13 Jul 2019].
62 **Barry E**, Roberts S, Oke J, *et al*. Efficacy and effectiveness of screen and treat policies in prevention of type 2 diabetes: systematic review and meta-analysis of screening tests and interventions. *BMJ* 2017;**356**:i6538.
63 **O'Sullivan JW**, Stevens S, Hobbs FDR, *et al*. Temporal trends in use of tests in UK primary care, 2000-15: retrospective analysis of 250 million tests. *BMJ* 2018;**363**:k4666.
64 **Cadogan SL**, Browne JP, Bradley CP, *et al*. The effectiveness of interventions to improve laboratory requesting patterns among primary care physicians: a systematic review. *Implement Sci* 2015;**10**:167.
65 **Fryer AA**, Smellie WSA. Managing demand for laboratory tests: a laboratory toolkit. *J Clin Pathol* 2013;**66**:62–72.
66 **Yoong SL**, Hall A, Stacey F, *et al*. Nudge strategies to improve healthcare providers' implementation of evidence-based guidelines, policies and practices: a systematic review of trials included within Cochrane systematic reviews. *Implement Sci* 2020;**15**:50.
67 **Sajid I**. Commissioning diagnostics: system biases and how to influence. in: preventing overdiagnosis. Quebec, 2017. Available: https://www.researchgate.net/publication/327231406_PODC_2017_Quebec_Commissioning_Diagnostics_System_Biases_How_to_Influence
68 **Whiting D**, Croker R, Watson J, *et al*. Optimising laboratory monitoring of chronic conditions in primary care: a quality improvement framework. *BMJ Open Qual* 2019;**8**:e000349–8.
69 **Main C**, Moxham T, Wyatt JC, *et al*. Computerised decision support systems in order communication for diagnostic, screening or monitoring test ordering: systematic reviews of the effects and cost-effectiveness of systems. *Health Technol Assess* 2010;**14**:1–227.
70 **Goldzweig CL**, Orshansky G, Paige NM, *et al*. Electronic health record-based interventions for improving appropriate diagnostic imaging: a systematic review and meta-analysis. *Ann Intern Med* 2015;**162**:557–65.
71 **Roshanov PS**, You JJ, Dhaliwal J, *et al*. Can computerized clinical decision support systems improve practitioners' diagnostic test ordering behavior? A decision-maker-researcher partnership systematic review. *Implementation Sci* 2011;**6**.
72 **Lapić I**, Juroš GF, Rako I, *et al*. Changing the electronic Request form proves to be an effective tool for optimizing laboratory test utilization in the emergency department. *Int J Med Inform* 2017;**102**:29–34.
73 **Seppänen K**, Kauppila T, Pitkälä K, *et al*. Altering a computerized laboratory test order form rationalizes ordering of laboratory tests in primary care physicians. *Int J Med Inform* 2016;**86**:49–53.
74 **O'Keeffe M**, Traeger AC, Hoffmann T, *et al*. Can nudge-interventions address health service overuse and underuse? protocol for a systematic review. *BMJ Open* 2019;**9**:e029540.
75 **Riikonen JM**, Guyatt GH, Kilpeläinen TP, *et al*. Decision AIDS for prostate cancer screening choice: a systematic review and meta-analysis. *JAMA Intern Med* 2019;**179**:1072–82.
76 **Shinkins B**, Perera R. Diagnostic uncertainty: dichotomies are not the answer. *Br J Gen Pract* 2013;**63**:122–3.
77 **Ridley L**. The incidental finding and the false-positive paradox. *J Med Imaging Radiat Oncol* 2020;**64**:75–7.
78 **Farmer CI**, Bourne AM, O'Connor D, *et al*. Enhancing clinician and patient understanding of radiology reports: a scoping review of international guidelines. *Insights Imaging* 2020;**11**:62.
79 **Power M**, Fell G, Wright M. Principles for high-quality, high-value testing. *Evid Based Med* 2013;**18**:5–10.
80 **National Institute for Health and Care Excellence**. Low back pain and sciatica in over 16S: assessment and management, 2016. Available: https://www.nice.org.uk/guidance/ng59 [Accessed 28 Jan 2018].
81 **Matharu VK**, Matharu GS. Metal-On-Metal hip replacements: implications for general practice. *Br J Gen Pract* 2017;**67**:544–5.
82 **Bottle A**, Kim D, Aylin PP, *et al*. Real-World presentation with heart failure in primary care: do patients selected to follow diagnostic and management guidelines have better outcomes? *Open Heart* 2018;**5**:1–8.
83 **O'Sullivan JW**, Albasri A, Nicholson BD, *et al*. Overtesting and undertesting in primary care: a systematic review and meta-analysis. *BMJ Open* 2018;**8**:e018557.
84 **Elwenspoek MMC**, Patel R, Watson JC, *et al*. Are guidelines for monitoring chronic disease in primary care evidence based? *BMJ* 2019;**365**:l2319.
85 **Bethge A**, Penciu O, Baksh S, *et al*. Appropriateness vs value: echocardiography in primary care. *Clin Cardiol* 2017;**40**:1212–7.
86 **Chen A**, Balogun-Lynch J, Aggarwal K, *et al*. Should all elective knee radiographs requested by general practitioners be performed weight-bearing? *Springerplus* 2014;**3**:707–7.
87 **Bhatt R**, Rajesh A, Morgan B, *et al*. An audit of hip radiographs performed for general practitioners. *Clin Radiol* 2001;**56**:970–2.
88 **Sajid I**, Frost K. Hear me out: rethinking internal auditory meatus (IAM) magnetic resonance imaging (MRI) in primary care. A cohort evaluation. SSRN, 2020. Available: https://papers.ssrn.com/abstract=3735206 [Accessed 22 Nov 2020].
89 **Logan GS**, Dawe RE, Aubrey-Bassler K, *et al*. Are general practitioners referring patients with low back pain for CTS appropriately according to the guidelines: a retrospective review of 3609 medical records in Newfoundland using routinely collected data. *BMC Fam Pract* 2020;**21**:236.
90 **Vickers AJ**, Kent DM. The lake Wobegon effect: why most patients are at Below-Average risk. *Ann Intern Med* 2015;**162**:866–7.
91 **Duddy C**, Wong G. Efficiency over thoroughness in laboratory testing decision making in primary care: findings from a realist review. *BJGP Open* 2020:bjgpopen20X101146.
92 **Zikmund-Fisher BJ**, Kullgren JT, Fagerlin A, *et al*. Perceived barriers to implementing individual choosing Wisely® recommendations in two national surveys of primary care providers. *J Gen Intern Med* 2017;**32**:210–7.

93 **Grover M**, McLemore R, Tilburt J. Clinicians report difficulty limiting low-value services in daily practice. *J Prim Care Community Health* 2016;**7**:135–8.
94 **Koch H**, van Bokhoven MA, ter Riet G, *et al*. Ordering blood tests for patients with unexplained fatigue in general practice: what does it yield? results of the vampire trial. *Br J Gen Pract* 2009;**59**:e93–100.
95 **Ng S**, Khan K, Box B. Faecal calprolectin - a good test used badly (Poster Abstracts). *Color Dis* 2017;**19**:14–64.
96 **Arshad K**, Stephenson JA, Mulcahy K, *et al*. Renal colic investigation in patients referred from general practice (GP). *Clin Radiol* 2015;**70**:S14–15.
97 **Aljebreen AM**, Alswat K, Almadi MA. Appropriateness and diagnostic yield of upper gastrointestinal endoscopy in an open-access endoscopy system. *Saudi J Gastroenterol* 2013;**19**:219–22.
98 **Allan GM**, Young J. Complete blood count for screening? *Can Fam Physician* 2017;**63**:772 http://www.ncbi.nlm.nih.gov/pubmed/29025804
99 **McCartney M**. Where's the evidence for NHS health checks? *BMJ* 2013;**347**:f5834.
100 **Leeflang MMG**, Rutjes AWS, Reitsma JB, *et al*. Variation of a test's sensitivity and specificity with disease prevalence. *Can Med Assoc J* 2013;**185**:E537–44.
101 **Riley RD**, Ahmed I, Debray TPA, *et al*. Summarising and validating test accuracy results across multiple studies for use in clinical practice. *Stat Med* 2015;**34**:2081–103.
102 **Willis BH**. Spectrum bias--why clinicians need to be cautious when applying diagnostic test studies. *Fam Pract* 2008;**25**:390–6.
103 **Naugler C**. More than half of abnormal results from laboratory tests ordered by family physicians could be false-positive. *Can. Fam. Physician* 2018;**64**:202–3.
104 **Freeman K**, Willis BH, Fraser H, *et al*. Faecal calprotectin to detect inflammatory bowel disease: a systematic review and exploratory meta-analysis of test accuracy. *BMJ Open* 2019;**9**:e027428.
105 **Conroy S**, Hale MF, Cross SS, *et al*. Unrestricted faecal calprotectin testing performs poorly in the diagnosis of inflammatory bowel disease in patients in primary care. *J Clin Pathol* 2018;**71**:316–22.
106 **Tusgul S**, Carron P-N, Yersin B, *et al*. Low sensitivity of qSOFA, SIRS criteria and sepsis definition to identify infected patients at risk of complication in the prehospital setting and at the emergency department triage. *Scand J Trauma Resusc Emerg Med* 2017;**25**:108.
107 **Finnikin S**, Willis BH, Ryan R. Factors predicting statin prescribing for primary prevention: a historical cohort study. *Br J Gen Pract* 2020;**71**:bjgp20X714065.
108 **Cadogan SL**, McHugh SM, Bradley CP. General practitioner views on the determinants of test ordering: a theory-based qualitative approach to the development of an intervention to improve immunoglobulin requests in primary care. *Implement Sci* 2016;**11**:1–12.
109 **Finnikin S**, Ryan R, Marshall T. Statin initiations and QRISK2 scoring in UK general practice: a thin database study. *Br J Gen Pract* 2017;**67**:e881–7.
110 **Webster BS**, Bauer AZ, Choi Y, *et al*. Iatrogenic consequences of early magnetic resonance imaging in acute, work-related, disabling low back pain. *Spine* 2013;**38**:1939–46.
111 **Webster BS**, Choi Y, Bauer AZ, *et al*. The cascade of medical services and associated longitudinal costs due to nonadherent magnetic resonance imaging for low back pain. *Spine* 2014;**39**:1433–40.
112 **O'Sullivan JW**, Muntinga T, Grigg S, *et al*. Prevalence and outcomes of incidental imaging findings: umbrella review. *BMJ* 2018:k2387.
113 **Najim R**, Booth TC, Petkova H. Incidental findings (IFs) discovered during imaging: the impact on primary care. *Br J Gen Pract* 2018;**68**:bjgp18X696725.
114 **Zafar HM**, Bugos EK, Langlotz CP, *et al*. "Chasing a Ghost": Factors that Influence Primary Care Physicians to Follow Up on Incidental Imaging Findings. *Radiology* 2016;**281**:567–73.
115 **van Vugt S**, Broekhuizen L, Zuithoff N, *et al*. Incidental chest radiographic findings in adult patients with acute cough. *The Annals of Family Medicine* 2012;**10**:510–5.
116 **Meyer HJ**, Pfeil A, Schramm D. Renal incidental findings on computed tomography: frequency and distribution in a large non selected cohort. *Medicine* 2017;**96**:e7039.
117 **Tewari A**, Shuaib W, Maddu KK, *et al*. Incidental findings on bedside ultrasonography: detection rate and accuracy of resident-performed examinations in the acute setting. *Can Assoc Radiol J* 2015;**66**:153–7.
118 **Hoang JK**, Langer JE, Middleton WD, *et al*. Managing incidental thyroid nodules detected on imaging: white paper of the ACR incidental thyroid findings Committee. *Journal of the American College of Radiology* 2015;**12**:143–50.
119 **Lumbreras B**, Donat L, Hernández-Aguado I. Incidental findings in imaging diagnostic tests: a systematic review. *Br J Radiol* 2010;**83**:276–89.
120 **Booth TC**, Boyd-Ellison JM. The current impact of incidental findings found during neuroimaging on neurologists' workloads. *PLoS One* 2015;**10**:e0118155.
121 **Ricós C**, Cava F, García-Lario JV, *et al*. The reference change value: a proposal to interpret laboratory reports in serial testing based on biological variation. *Scand J Clin Lab Invest* 2004;**64**:175–84.
122 **Bugdayci G**, Oguzman H, Arattan HY, *et al*. The use of reference change values in clinical laboratories. *Clin Lab* 2015;**61**:251–7.
123 **Doll H**, Shine B, Kay J, *et al*. The rise of cholesterol testing: how much is unnecessary. *Br J Gen Pract* 2011;**61**:e81–8.
124 **McCormack JP**, Holmes DT. Your results may vary: the imprecision of medical measurements. *BMJ* 2020;**368**:m149.
125 **Smellie WSA**. Spurious hyperkalaemia: pitfalls of testing and summary of guidance. *Br Med J* 2007;**334**:693–5.
126 **Sinclair D**, Briston P, Young R. Seasonal pseudohyperkalaemia. *J Clin Pathol* 2003;**56**:385–7.
127 **Levari DE**, Gilbert DT, Wilson TD, *et al*. Prevalence-induced concept change in human judgment. *Science* 2018;**360**:1465–7.
128 **Hopstaken RM**, Butler CC, Muris JWM, *et al*. Do clinical findings in lower respiratory tract infection help general practitioners prescribe antibiotics appropriately? an observational cohort study in general practice. *Fam Pract* 2006;**23**:180–7.
129 **Havdahl A**, Mitchell R, Paternoster L, *et al*. Investigating causality in the association between vitamin D status and self-reported tiredness. *Sci Rep* 2019;**9**:1–8.
130 **Martin KR**, Reid DM. Is there a role for vitamin D in the treatment of chronic pain? *Ther Adv Musculoskelet Dis* 2017;**9**:131–5.
131 **Straube S**, Derry S, Straube C. Vitamin D for the treatment of chronic painful conditions in adults. *Cochrane Database Syst. Rev* 2015;**2015**.
132 **Schneeweiss R**, Ellsbury K, Hart LG. The economic impact and multiplier effect of a family practice clinic on an academic medical center. *JAMA* 1989;**262**:370–5 http://www.ncbi.nlm.nih.gov/pubmed/2739039
133 **Watson J**, Salisbury C, Whiting P, *et al*. Added value and cascade effects of inflammatory marker tests in UK primary care: a cohort study from the clinical practice research Datalink. *Br J Gen Pract* 2019;**69**:e470–8.
134 **Adams SJ**, Rakheja R, Bryce R, *et al*. Incidence and Economic Impact of Incidental Findings on ^{18}F-FDG PET/CT Imaging. *Can Assoc Radiol J* 2018;**69**:63–70.
135 **Morgan AE**, Berland LL, Ananyev SS, *et al*. Extraurinary Incidental Findings on CT for Hematuria: The Radiologist's Role and Downstream Cost Analysis. *American Journal of Roentgenology* 2015;**204**:1160–7.
136 **Huo J**, Xu Y, Sheu T, *et al*. Complication rates and downstream medical costs associated with invasive diagnostic procedures for lung abnormalities in the community setting. *JAMA Intern Med* 2019;**179**:324.
137 **Lafata JE**, Simpkins J, Lamerato L. The economic impact of false-positive cancer screens. *Cancer Epidemiol Biomarkers Prev* 2004;**13**:2126–32 http://www.ncbi.nlm.nih.gov/pubmed/15598770
138 **Rolfe A**, Burton C. Reassurance after diagnostic testing with a low pretest probability of serious disease. *JAMA Intern Med* 2013;**173**:407.
139 **van Ravesteijn H**, van Dijk I, Darmon D, *et al*. The reassuring value of diagnostic tests: a systematic review. *Patient Educ Couns* 2012;**86**:3–8.
140 **Howard L**, Wessely S, Leese M, *et al*. Are investigations anxiolytic or anxiogenic? a randomised controlled trial of neuroimaging to provide reassurance in chronic daily headache. *J Neurol Neurosurg Psychiatry* 2005;**76**:1558–64.