education

FROM THE JOURNALS Edited highlights of weekly research reviews on https://bit.ly/2PLtil8

White coat hypertension

This isn't any old high blood pressure. This is blood pressure that is only high in the doctor's office. It can be one of two varieties, treated or untreated. The former is called white coat effect and the latter white coat hypertension. The difference is important because many believe the latter not to be clinically significant, thus steering away from antihypertensive medication. Cohen et al show us otherwise by looking at the risk of death and cardiovascular events in these two groups compared with people with normal blood pressure. They carefully meta-analysed 27 observational studies, in which people had blood pressure monitoring in and out of the doctor's office. They found no difference in risk between people with white coat effect and those with normal blood pressure, but found increased risk in those with white coat hypertension. The caveat is that confounding is still likely to be have been present no matter the quality of this analysis. Hence the results are unlikely to directly affect patient care.

Ann Intern Med doi:10.7326/M19-0223

Blood donor sex

This study finds that people receiving transfusions from female donors who were previously pregnant or from sex discordant donors do not have higher mortality. It's good news because blood for transfusion is a limited resource, and it's hard enough to match for blood group, let alone sex considerations. Edgren et al conducted an analysis of two US cohorts and a Scandanavian cohort of blood transfusions. Less than a quarter of the transfusions were from parous or previously pregnant donors. In-hospital mortality was no different whether the donor was female, previously pregnant, or sex-discordant. The authors also looked at longer term mortality, and again there was no difference based on these features. Unfortunately, pregnancy status was not known for many of the donors, so imputation had to be used, but this may be the best data we can get on this subject.

JAMA doi:10.1001/jama.2019.7084

New predictor of acute kidney injury

This cohort study found that elevated concentrations of dickkopf-3 in the urine (a marker for renal tubular stress) before cardiac surgery were associated with increased risk of acute kidney injury. This study may hail a new biomarker for predicting acute kidney injury. A cynic might ask what good is predicting risk if we can't modify it. I'd reply that, at least, it can be used to help plan postoperative care and, at most, it is a step towards finding an intervention that modifies risk because it better identifies patients in whom to target the intervention.

Lancet doi:10.1016/S0140-6736(19)30769-X

Prevention of type 1 diabetes

Everyone would be astounded if a drug could prevent type 1 diabetes. Is it going to be teplizumab, an anti-CD3 antibody? Herold et al tested a two week course of this intravenous drug in a double blind, randomised controlled trial of 76 people who had a relative with type 1 diabetes and were at high risk of developing the disease. To be included in the trial, they had to be over the age of 8 years and have diabetes related autoantibodies and impaired glucose tolerance. The result was a clear reduction in progression to diabetes in the teplizumab group compared with placebo. Immune therapy might well modify the early course of type 1 diabetes, but more work is needed to assess safety and the optimal duration and frequency of treatments.

N Engl J Med doi:10.1056/NEJMoa1902226

Light and obesity?

Presence of artificial light at night while sleeping has been linked with obesity before. Park et al analysed data from over 40 000 women in the US and Puerto Rico, performed adjustments for confounders, and found the same link. The risk increase was statistically significant, but relatively small. This is fascinating and finding the mechanism even more so. Obesity is more than just overconsumption of calories, there is a complex interaction between neurological and metabolic processes. In terms of interventions to combat obesity, it's easy to, say, avoid television and switch the lights off, but, on the basis of these data, I wouldn't expect that to have much impact as it doesn't get to the core of what is making the light lead to weight gain. Are there features of personality, mental state, or economic status associated with use of artificial light at night that could be contributing? These factors can't be fully accounted for in an observational study.

JAMA Intern Med doi:10.1001/jamainternmed.2019.0571

Alex Nowbar is a clinical research fellow at Imperial College London.

Evaluating the impact of healthcare interventions using routine data

CPD READING
0.5 HOURS

Geraldine M Clarke,[1] Stefano Conti,[2] Arne T Wolters,[1] Adam Steventon[1]

[1]The Health Foundation, London
[2]NHS England and NHS Improvement, London
Correspondence to: G Clarke Geraldine.clarke@health.org.uk

Interventions to transform the delivery of health and social care are being implemented widely. Assessing the impact of these interventions enables healthcare teams to learn and to improve services, and can inform future policy.[3] However, some healthcare interventions are implemented without high quality evaluation, in ways that require onerous data collection, or may not be evaluated at all.[4]

A range of routinely collected healthcare data could be used to evaluate these interventions. However, there is a lack of guidance as to where relevant routine data can be found or accessed and how they can be linked to other data. A diverse array of methodological literature can also make it hard to understand which methods to apply to analyse the data. This article provides an introduction to help clinicians, commissioners, and other healthcare professionals wishing to commission, interpret, or perform an impact evaluation of a health intervention.

Impact evaluations

Formative	Summative	Examples
Conducted during the development or implementation of an intervention	Conducted after the intervention's completion, or at the end of a programme cycle	A formative evaluation of the Whole Systems Integrated Care (WSIC) programme, aimed at integrating health and social care in London, found that difficulties in establishing data sharing and information governance, and differences in professional culture were hampering efforts to implement change[10]
Aims to fine tune or reorient the intervention	Aims to render judgment, or make decisions about the future of the intervention	A summative impact evaluation of an NHS new care model vanguard initiative found that care home residents in Nottinghamshire who received enhanced support had substantially fewer attendances at emergency departments and fewer emergency admissions than a matched control group.[13] This evidence supported the decision by the NHS to roll out the Enhanced Health in Care Homes Model across the country.[2]

HOW PATIENTS WERE INVOLVED IN THE CREATION OF THIS ARTICLE

No patients were involved in the creation of this article.

WHAT YOU NEED TO KNOW

- Assessing the impact of healthcare interventions is critical to inform future decisions
- Compare observed outcomes with what you would have expected if the intervention had not been implemented
- A wide range of routinely collected data is available for the evaluation of healthcare interventions

What are interventions, impacts, and impact evaluations?

A health intervention is a combination of activities or strategies designed to assess, improve, maintain, promote, or modify health among individuals or an entire population. Interventions can include educational or care programmes, policy changes, environmental improvements, or health promotion campaigns. Interventions that include multiple independent or interacting components are referred to as complex.[5] The impact of any intervention is likely to be shaped as much by the context (eg, communities, work places, homes, schools, or hospitals) in which it is delivered, as the details of the intervention itself.[6-9]

An impact is a positive or negative, direct or indirect, intended or unintended change produced by an intervention. An impact evaluation is a systematic and empirical investigation of the effects of an intervention; it assesses to what extent the outcomes experienced by affected individuals were caused by the intervention in question, and what can be attributed to other factors such as other interventions, socioeconomic trends, and political or environmental conditions. Evaluations can be categorised as formative or summative (table).

With either type of evaluation, it is important to be realistic about how long it will take to see the intended effects. Assessment that takes place too soon risks incorrectly concluding that there was no impact. This might lead stakeholders to question the value of the intervention, when later assessment might have shown a different picture. For example, in a small case study of cost savings from proactively managing high risk patients, the costs of healthcare for the eligible intervention population initially increased compared with the comparison population, but after six months were consistently lower.[14]

This article focuses on impact evaluation, but this can only ever address a fraction of questions.[15] Much more can be accomplished if it is supplemented with other qualitative and quantitative methods, including process evaluation. This provides context, assesses how the intervention was implemented, identifies any emerging unintended pathways, and is important for understanding what happened in practice and for identifying areas for improvement.[16] The economic evaluation of healthcare interventions is also important for healthcare decision making, especially with ongoing financial pressures on health services (fig 1).[17]

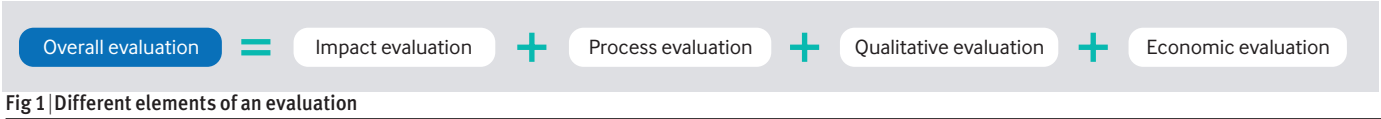


Fig 1 | Different elements of an evaluation

What are the right evaluation questions?

An effective impact evaluation begins with the formulation of one or more clear questions driven by the purpose of the evaluation and what you and your stakeholders want to learn. For example, *"What is the impact of case management on patients' experience of care?"*

Formulate your evaluation questions using your understanding of the idea behind your intervention, the implementation challenges, and your knowledge of what data are available to measure outcomes (fig 2). Review your theory of change or logic model[21 22] to understand what inputs and activities were planned, and what outcomes were expected and when. Once you have understood the intended causal pathway, consider the practical aspects of implementation, which include the barriers to change, unexpected changes by recipients or providers, and other influences not previously accounted for. Patient and public involvement in setting the right question is strongly recommended for additional insights and meaningful results.

What methods can be used to perform an impact evaluation?

Randomised control designs, where individuals are randomly selected to receive either an intervention or a control treatment, are often referred to as the "gold standard" of causal impact evaluation.[24] In large enough samples, the process of randomisation ensures a balance in observed and unobserved characteristics between treatment and control groups. However, while often suitable for assessing, for example, the safety and efficacy of medicines, these designs may be impractical, unethical, or irrelevant when assessing the impact of complex changes to health service delivery.

Observational studies are an alternative approach to estimate causal effects. They use the natural, or unplanned, variation in a population in relation to the exposure to an intervention, or the factors that affect its outcomes, to remove the consequences of a non-randomised selection process.[25] The idea is to mimic a randomised control design by ensuring treated and control groups are equivalent—at least in terms of *observed* characteristics. This can be achieved using a variety of well documented methods, including regression control and matching,[26] eg, propensity scoring[27] or genetic matching.[28] If the matching is successful at producing such groups, and there are also no differences in *unobserved* characteristics, then it can be assumed that the control group outcomes are representative of those that the treated group would have experienced if nothing had changed, ie, the counterfactual. For example, an evaluation of alternative elective surgical interventions for primary total hip replacement on osteoarthritis patients in England and Wales used genetic matching to compare patients across three different prosthesis groups, and reported that the most prevalent type of hip replacement was the least cost effective.[29]

Assessing similarity is only possible in relation to *observed* characteristics, and matching can result in biased estimates if the groups differ in relation to *unobserved* variables that are predictive of the outcome (confounders). It is rarely possible to eliminate this possibility of bias when conducting observational studies, meaning that the interpretation of the findings must always be sensitive to the possibility that the differences in outcomes were caused by a factor other than the intervention. Methods that can help when selection is on unobserved characteristics include difference-in-difference,[30] regression discontinuity,[31] instrumental variables,[18] or synthetic controls.[32].

Observational studies are often referred to as natural (for natural or unplanned interventions), or quasi (for planned or intentional interventions) experiments. Natural experiments are discussed to evaluate population health interventions.[41]

Patient and public involvement (PPI) in setting the right question is strongly recommended

What's wrong with a simple before-and-after study?

Before-and-after studies compare changes in outcomes for the *same* group of patients at a single time point before and after receiving an intervention without reference to a control group. These differ from interrupted time series studies, which compare changes in outcomes for *successive* groups of patients before and after receiving an intervention (the interruption).

Before-and-after studies are useful when it is not possible to include an unexposed control group, or for hypothesis generation. However, they are inherently susceptible to bias since changes observed may simply reflect regression to the mean (any changes in outcomes that might occur naturally in the absence of the intervention), or influences or secular trends unrelated to the intervention, eg, changes in the economic or political environment, or a heightened public awareness of issues.

For example, a before-and-after study of the impact of a care coordination service for older people tracked the hospital utilisation of the same patients before and

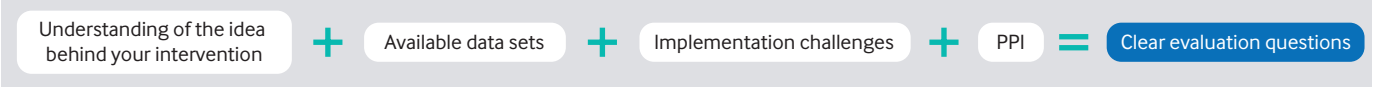


Fig 2 | Developing clear evaluation questions

after they were accepted into the service. They found that the service resulted in savings in hospital bed days and attendances at the emergency department.[42] Reduced hospital utilisation could have reflected regression to the mean here rather than the effects of the intervention; for example, a patient could have had a specific health crisis before being invited to join the service and then reverted back to their previous state of health and hospital utilisation for reasons unconnected with the care coordination service.

Various tools are available to evaluate the risk of bias in non-randomised designs due to confounding and other potential biases.[43 44]

Where can I find suitable routine data?

Healthcare systems generate vast amounts of data as part of their routine operation. These datasets are often designed to support direct care, and for administrative purposes, rather than for research, and use of routinely collected data for evaluating changes in health service delivery is not without pitfalls. For example, any variation observed between geographical regions, providers, and sometimes individual clinicians may reflect real and important variations in the actual healthcare quality provided, but can also result from differences in measurement.[45] However, routine data can be a rich source of information on a large group of patients with different conditions across different geographical regions. Often, data have been collected for many years, enabling construction of individual patient histories describing healthcare utilisation, diagnoses, comorbidities, prescription of medication, and other treatments.

Some of these data are collected centrally, across a wider system, and routinely shared for research and evaluation purposes, eg, secondary care data in England (Hospital Episode Statistics), or Medicare Claims data in the US. Other sources, such as primary care data, are often collected at a more local level, but can be accessed through, or on behalf of, healthcare commissioners, provided the right information governance arrangements are in place. Pseudonymised records, where any identifying information is removed or replaced by an artificial identifier, are often used to support evaluation while maintaining patient confidentiality.

Healthcare records can often be linked across different sources as a single patient identifier is commonly used across a healthcare system, eg, the use of an NHS number in the UK. Using a common pseudonym across different data sources can support linkage of pseudonymised records. Linking into publicly available sources of administrative data and surveys can further enrich healthcare records. Commonly used administrative data available for UK populations include measures of GP practice quality and outcomes from the Quality and Outcomes Framework (QOF),[52] deprivation, rurality, and demographics from the 2011 Census,[53] and patient experience from the GP Patient Survey.[54]

There is no standard, "one size fits all" recipe for a good evaluation

Are there any additional considerations?

It is essential to consider threats to validity when designing and evaluating an impact evaluation; validity relates to whether an evaluation is measuring what it is claiming to measure. See Rothman and colleagues[55] for further discussion.

Internal validity refers to whether the effects observed are due to the intervention and not some other confounding factor. Selection bias, which results from the way in which subjects are recruited, or from differing rates of participation due, for example, to age, gender, cultural or socioeconomic factors, is often a problem in non-randomised designs. Care must be taken to account for such biases when interpreting the results of an impact evaluation. Sensitivity analyses should be performed to provide reassurance regarding the plausibility of causal inferences.

External validity refers to the extent to which the results of a study can be generalised to other settings. Understanding the societal, economic, health system, and environmental context in which an intervention is delivered, and which makes its impact unique, is critical when interpreting the results of evaluations, and considering whether they apply to your setting.[56] Descriptions of context should be as rich as possible.

Often, the impact of an intervention is likely to vary depending on the characteristics of patients. These can be usefully explored in subgroup analyses.[57]

Clear and transparent reporting using established guidelines (eg, STROBE[58] or TREND[59]) to describe the intervention, study population, assignment of treatment, and control groups, and methods used to estimate impact should be followed. Limitations arising as a result of inherent biases, or validity, should be clearly acknowledged.

Around the world, many interventions designed to improve health and healthcare are under way. An evaluation is an essential part of understanding what impact these changes are having, for whom and in what circumstances, and help inform future decisions about improvement and further roll out. There is no standard, "one size fits all" recipe for a good evaluation: it must be tailored to the project at hand. Understanding the overarching principles and standards is the first step towards a good evaluation.

Competing interests: We have read and understood BMJ policy on declaration of interests. All authors work in the Improvements Analytics Unit, a joint project between NHS England and the Health Foundation, which provided support for work reported in references of this report.[13-60]

Supported by The Health Foundation.

Cite this as: *BMJ* 2019;365:l2239

Find the full version with references at http://dx.doi.org/10.1136/bmj.l2239

EDUCATION INTO PRACTICE

- What interventions have you designed or experienced aimed at transforming your service? Have they been evaluated?
- What types of routine data are collected about the care you deliver? Do you know how to access them and use them to evaluate care delivery?
- What resources are available to you to support impact evaluations for interventions?

Are guidelines for monitoring chronic disease in primary care evidence based?

Martha M C Elwenspoek,[1 2] Rita Patel,[1 2] Jessica C Watson,[1 2] Penny Whiting[1 2]

[1]National Institute for Health Research Collaboration for Leadership in Applied Health Research and Care West (NIHR CLAHRC West), University Hospitals Bristol NHS Foundation Trust

[2]Population Health Sciences, Bristol Medical School, University of Bristol

Correspondence to: M M C Elwenspoek martha.elwenspoek@bristol.ac.uk

This is one of a series of occasional articles that highlight areas of practice where management lacks convincing supporting evidence. The series advisers are Sera Tort, clinical editor, Nai Ming Lai, clinical editor, and David Tovey, editor in chief, the Cochrane Library. To suggest a topic for this series, please email us at uncertainties@bmj.com

CPD READING
0.5 HOURS

Primary care clinicians rely on guidelines for common chronic diseases such as type 2 diabetes, chronic kidney disease, and hypertension to inform them which tests they should recommend to their patients and how frequently these should be done. With rates of pathology tests rising—at an estimated annual cost of £1.8bn to primary care in the UK[1]—and the potential for harm from over-testing, it is important to consider the evidence base for these recommendations.

In this article, we review monitoring strategies in current UK guidelines for patients with type 2 diabetes, chronic kidney disease, and hypertension (box 1), highlighting the uncertainties in these guidelines and the need for further research.

Box 1 | Search strategy and guideline selection

We searched for published UK guidelines for the management of patients with type 2 diabetes, chronic kidney disease stages 1-3*, or hypertension using the following sources:

- National Institute for Health and Care Excellence (NICE)
- Scottish Intercollegiate Guidelines Network (SIGN)
- Royal Colleges of Pathologists (RCPath), Physicians, and General Practitioners
- Quality Outcomes Framework (QOF)

The following guidelines are included in this review:

- SIGN 116 Management of diabetes (2017)[2]
- NICE CG127 Hypertension, the clinical management of primary hypertension in adults (2011)[3]
- NICE CG182 Chronic kidney disease (partial update) (2014)[4]
- NICE NG28 Type 2 diabetes in adults (2015)[5]
- NICE PH38 Evidence reviews (Type 2 diabetes: prevention in people at high risk) (2017)[6]
- RCPath: National minimum retesting intervals in pathology (2015)[7]

We extracted any guidance on the use of laboratory tests for disease monitoring, the recommended frequency of testing, and the level of evidence on which the guidance was based. We did not search the primary literature itself. As a consequence, we may have missed evidence that is not picked up by the guidelines or was published after the guideline was written.

*Chronic kidney disease stages 4 and 5 are generally monitored in secondary care and are therefore not included in our analysis.

HOW PATIENTS WERE INVOLVED IN THE CREATION OF THIS ARTICLE

We held a discussion workshop with members of the CLAHRC West Health Systems Panel. Participants had chronic conditions requiring blood test monitoring or had family members with chronic conditions. Participants told us that which tests are done is rather "hidden" and they felt that their GPs did not always explain what the tests are. They were surprised that guidelines about monitoring are largely based on expert opinion. There was a general expectation that tests results are always 100% accurate and that over-testing could not cause any harms.

WHAT YOU NEED TO KNOW

- Current UK guidelines for monitoring type 2 diabetes, chronic kidney disease, and hypertension are largely based on expert opinion; robust evidence for optimal monitoring strategies and testing intervals is lacking
- Unnecessary testing in primary care can lead to false positive and false negative results, increased workload for clinicians, and increased costs for the health service
- Patients and healthcare professionals should be aware of these uncertainties when making shared decisions about chronic disease monitoring

What is the evidence of uncertainty?

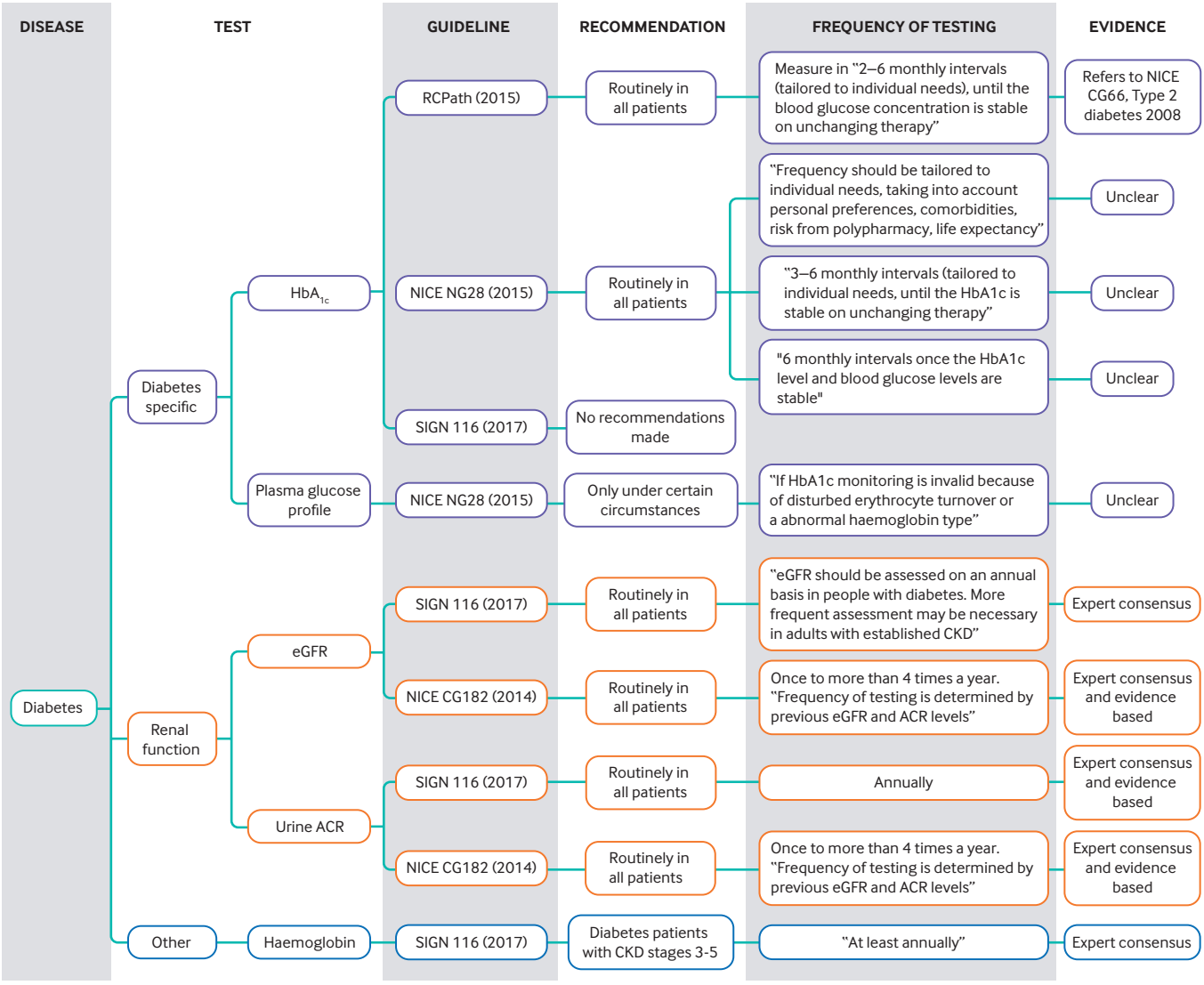
Tests recommended by guidelines

For the chronic diseases reviewed, the recommended tests are similar across guidelines. In the case of type 2 diabetes the monitoring tests recommended across guidelines are glycated haemoglobin (HbA_{1c}), plasma glucose profile, and renal function tests such as estimated glomerular filtration rate (eGFR) and urine albumin:creatinine ratio (ACR) (fig 1). Surprisingly, there is no clear recommendation in the SIGN 116 diabetes guideline to measure HbA_{1c} routinely.[2]

For chronic kidney disease (stages 1-3), guidelines recommend measuring eGFR and ACR routinely, but not serum calcium, phosphate, parathyroid hormone, or vitamin D (fig 2, see bmj.com).

For hypertension, recommended monitoring tests are urine ACR, haematuria, electrolytes and creatinine, total and high density lipoprotein (HDL) cholesterol, renal profile, HbA_{1c}, lipid profile, blood glucose, and eGFR (fig 3).

Testing recommendations are scattered across most guidelines with no specific sections on monitoring. Consequently, clinicians need to read an entire guideline to get an overview of all recommended tests. An overview of monitoring recommendations



ACR = albumin:creatinine ratio; eGFR = estimated glomerular filtration rate; CKD = chronic kidney disease; RCPath = Royal Colleges of Pathologists; NICE = National Institute for Health and Care Excellence; SIGN = Scottish Intercollegiate Guidelines Network

Fig 1 | Guidelines and evidence for tests to monitor type 2 diabetes. Tests are referred to by the same names as in each relevant guideline.

from several guidelines is provided by the RCPath national minimal retesting intervals report, but this document refers to outdated guidelines and awaits updating.[7]

Retesting intervals in guidelines are often missing or unclear

Recommended frequency of testing varies between guidelines or is sometimes not specified at all. For example, SIGN recommends annual testing of renal function in patients with diabetes,[2] whereas NICE suggests that test intervals should be determined by previous renal function results.[4] NICE recommends that individual needs are taken into account when determining the frequency of monitoring, although it is not specified how testing intervals should be adjusted.[5]

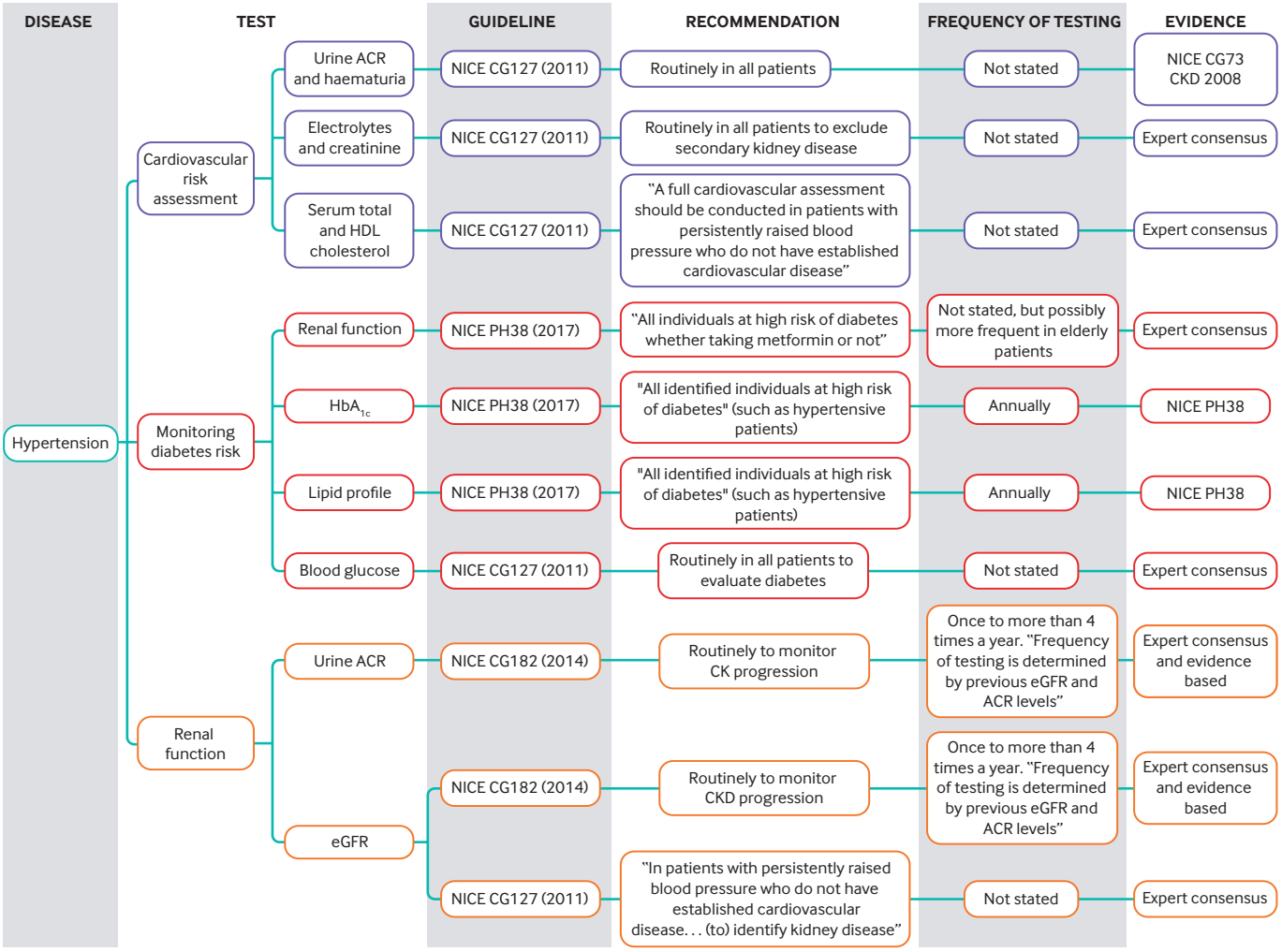
Robust evidence for optimal monitoring strategies and testing intervals is lacking

Most of these recommendations are based on expert opinion, provided by the respective guideline development groups. Where evidence is cited it does not address the fundamental question of whether the test in question is necessary or beneficial. For instance, in support of ACR monitoring, the SIGN diabetes guideline cites a meta-analysis of 10 diagnostic cohort studies in patients with diabetes.[8] However, these studies investigate test performance of ACR, not whether ACR monitoring has an impact on disease progression or mortality in patients with diabetes.

There is no evidence to support frequency of testing of any test in any of the guidelines. Recommendations regarding frequency of testing are entirely based on expert opinion.

Will ongoing research provide relevant evidence?

We searched ClinicalTrials.gov, a clinical trial registry, for studies addressing optimising chronic disease monitoring, using combinations of the following search terms: "type 2 diabetes," "chronic kidney disease," "hypertension," "primary care," "general practice," "laboratory test," and "monitoring." No relevant ongoing studies were identified. Current studies focus of the diagnostic or prognostic accuracy of certain tests, but there is significant uncertainty about how to determine optimal testing frequency or how to evaluate whether monitoring is appropriate.



ACR = albumin:creatinine ratio; HDL = high density lipoprotein; eGFR = estimated glomerular filtration rate; CKD = chronic kidney disease; NICE = National Institute for Health and Care Excellence

Fig 3 | Guidelines and evidence for tests to monitor hypertension. Tests are referred to by the same names as in each relevant guideline

What should we do in light of the uncertainty?

We recommend using the current guidelines where clear testing recommendations are given, as they are based on the best available evidence. These guideline recommendations should feed into, rather than over-ride, discussions with patients that incorporate their values and preferences. In the absence of clear evidence, it is all the more important that clinicians consider with their patients which tests are likely to influence disease management. GPs should ensure that there is a clear clinical rationale for each test that they perform. As chronic disease monitoring is often delegated to nursing staff or healthcare assistants, GPs should consider offering training about these uncertainties and the potential harms of over-testing to the wider primary care team.

Shared decision making

Patients' values and preferences about monitoring should always be taken into account. Some patients may prefer more frequent testing, others will opt for less. Information about testing, including the uncertainties raised here, should be discussed with patients to promote shared decision making.

Avoiding unnecessary testing

Unnecessary testing in a low prevalence setting such as primary care is more likely to lead to false positives, which in turn can lead to cascades of follow-up testing.[10 11] This can generate anxiety for patients, increased workload for doctors, and increased costs for the health service.[12 13] False negative results, on the other hand, may lead to false reassurances and delayed diagnosis.[11]

A substantial proportion of pathology testing may be unnecessary, or even inappropriate. In one study of cholesterol testing rates in Oxfordshire, 42-79% of cholesterol tests were estimated as potentially unnecessary.[14] However, there is no consensus on what an inappropriate test is, and estimates of inappropriate test ordering vary substantially (0.2%–100%).[15] Most studies examining inappropriate testing compare testing rates to guideline recommendations rather than to robust evidence on what constitutes an appropriate or inappropriate test.

Competing interests: None declared.

Cite this as: *BMJ* 2019;365:l2319

Find the full version with references at http://dx.doi.org/10.1136/bmj.l2319

EDUCATION INTO PRACTICE

- When you order blood tests for your patients, is there always a clear rationale for each test?
- How do you explain to patients which tests they are having and why? How do you discuss the limitations of blood tests with patients?
- Do you use local practice protocols for blood tests in patients with chronic diseases? Do these contain any unnecessary extra tests in addition to those recommended by current guidelines?

Patient centred care for multimorbidity improves patient experience

The study

A patient-centred intervention to improve the management of multimorbidity in general practice: the 3D RCT

Salisbury C, Man M-S, Chaplin K, and colleagues

Health Serv Deliv Res 2019;7(5). This study was funded by the National Institute for Health Research Health Services and Delivery Programme (project number 12/130/15)

NIHR | National Institute for Health Research

NIHR Signals provide decision makers in the NHS, public health, and social care with the latest important research from the NIHR and other health research organisations.

To read the full NIHR Signal, go to: http://bit.ly/2J9GaBi

CPD READING

0.5 HOURS

Why was this study needed?

The UK has an ageing population. A 2018 study of multimorbidity in England found that 54% of people over 65 had multimorbidity in 2015. This is expected to rise to 68% by 2035, with 17% of people over 65 expected to have four or more conditions.

National guidelines tend to focus on single conditions. People with multimorbidity may have their conditions managed individually without due consideration of the overall burden of their diseases and treatments. Recognition of this prompted the National Institute for Health and Care Excellence (NICE) and other international organisations to develop guidelines on managing multimorbidity.

This trial assessed a care model that incorporated all strategies recommended by guidelines.

What did this study do?

This cluster randomised controlled trial allocated 33 general practices in England and Scotland to provide comprehensive, three dimensional reviews of multimorbidity or to continue with usual care. Medical records were used to identify adults with at least three of 17 chronic conditions listed in the Quality and Outcomes Framework (QOF). A total of 1546 participants were included (average age 71 years).

Three dimensional reviews were conducted six monthly. They included a nurse appointment to discuss the patient's key health concerns and their effects on daily life, and to screen for depression and dementia. A pharmacist reviewed the patient's medications. The patient met with their doctor, who considered the reviews from the nurse and pharmacist and agreed with the patient a health plan, including realistic goals.

What did it find?

- The multimorbidity review had no effect on quality of life at 15 months. There was 0.00 difference in EQ-5D-5L quality of life score between groups (95% confidence interval −0.02 to +0.02) following adjustment for baseline variables, practice location, and list size.
- There was no difference in any measure of illness burden, which included self rated health, anxiety, and depression scores, or in measures of treatment burden, which included number of drugs prescribed and medication adherence.
- Patients having the multimorbidity review did, however, have better patient centred care than those receiving usual care. This included higher Patient Assessment of Care for Chronic Conditions scores, which ranged from 1 to 5 (adjusted mean difference 0.29, 96% confidence interval 0.16 to 0.41), and greater proportions of participants reporting being very satisfied with care (56% compared with 39%) and having the opportunity to discuss problems of greatest importance to them (42% compared with 26%).

What does current guidance say on this issue?

The NICE 2016 guideline on clinical assessment and management of multimorbidity provides guidance on identifying patients with multimorbidity. It recommends establishing the patient's burden of disease and treatment and their values and priorities, reviewing medications, and agreeing an individualised management plan. The central aim of this approach is to enable patients to actively participate in their care, ensure services meet their needs, improve continuity of care and relationships, and ultimately improve the patient's quality of life.

What are the implications?

Comprehensive care reviews for multimorbidity appear to improve patients' experience of care but have no effect on quality of life. It may be because the intervention was not delivered at full dose/intensity (only 49% received the full two sessions) or monitored for long enough. However, the results are in line with other large trials included in a recently updated Cochrane review.

The principles of care in the control group are generally consistent with NICE guidelines. Yet there is likely to be variation across trusts and practices in the specifics of how reviews are conducted and how frequently. This is the largest trial to date of approaches to improve management of multiple conditions in primary care. Cost effectiveness analysis was uncertain with only small differences in costs and outcomes.

Competing interests: *The BMJ* has judged that there are no disqualifying financial ties to commercial companies. Further details of other interests, disclaimers, and permissions can be found on bmj.com

Cite this as: *BMJ* 2019;364:k4439

CASE REVIEW An unstable lump in the groin

A 49 year old man with type 2 diabetes presented with a three day history of severe right sided groin swelling and pain, feeling generally unwell, fatigue, and fever. His medical history included two myocardial infarctions (11 and 16 years earlier) and a previous right sided inguinal abscess (two years earlier) that healed after incision and drainage. A 20×10 cm lump and skin crepitations were felt in the right groin. Overlying skin appeared healthy with no evidence of erythema, necrosis, or pallor. Glasgow coma scale (GCS) was 15, blood pressure 85/64 mm Hg, heart rate 114 beats/min, temperature 38.4°C, oxygen saturations 98% on room air, and respiratory rate 28 breaths per minute. Blood test abnormalities are shown in the table.

The patient required immediate incision and drainage (fig 1). There was a foul smell after the initial incision but no frank pus.

1 What are the differential diagnoses of groin lumps?

2 What is the most likely diagnosis in this patient?

3 Does this patient meet the qSOFA definition of sepsis?

Submitted by Zain Sheikh and Augustine Akali

Patient consent obtained.

Cite this as: *BMJ* 2019;365:l2214

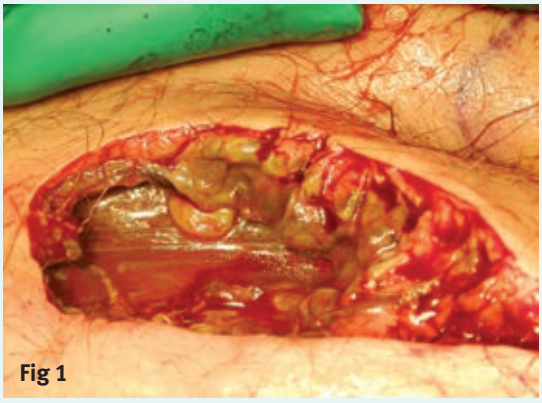


Fig 1

Blood test abnormalities

	Result	Normal range
White cell count	18.1×10^9/L	$4\text{-}11\times10^9$/L
C reactive protein	311 mg/L	0-10 mg/L

answers CASE REVIEW An unstable lump in the groin

1 What are the differential diagnoses of groin lumps?

- Inguinal abscess: painful fluctuant mass of pus which can recur
- Necrotising fasciitis: necrotic tissue following untreated/insufficiently treated abscess, trauma, insect bite, or skin break
- Incarcerated inguinal hernia: protrusion through inguinal canal
- Inguinal lymphadenopathy: reaction to lower limb pathology—eg, severe cellulitis, malignancy
- Unstable pseudoaneurysm: pulsatile mass.

2 What is the most likely diagnosis in this patient?

Necrotising fasciitis caused by a new or recurrent abscess.

The condition can develop over hours or days. Hallmarks include

- Discharge that is foul smelling or with a dishwater appearance (fig 2)
- Necrotic fascia (fig 2) or lack of bleeding
- No muscular fascia resistance to blunt dissection
- Haemodynamic instability
- Skin crepitations.

Risk factors include previous or recurrent abscesses, hyperglycaemia, and impaired immune function.

3 Does this patient meet the qSOFA definition of sepsis?

Yes, based on his respiratory rate and blood pressure.

Sepsis is suggested by the presence of two or more of the following parameters from the modified/quick Sequential Organ Dysfunction Assessment (qSOFA):

- GCS <15
- Systolic blood pressure <100 mm Hg
- Respiratory rate >22 breaths per minute.

Of patients with necrotising fasciitis, 16.3% develop sepsis.

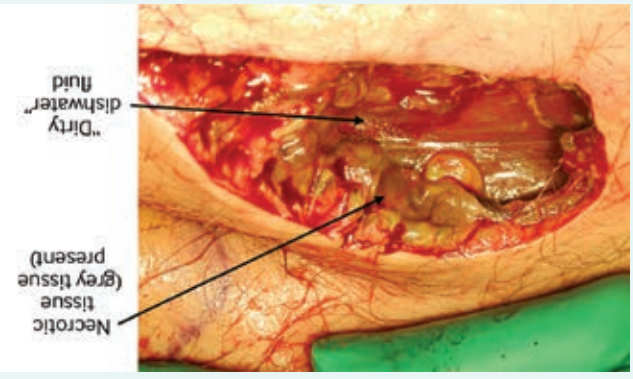


Fig 2 | After the initial incision, showing hallmarks of necrotising fasciitis

LEARNING POINT

- qSOFA scores and the TIME mnemonic are useful for suspected sepsis:
- Temperature: high or low.
- Infection: look for signs.
- Mental decline: confused, sleepy, difficult to rouse.
- Extremely ill: severe discomfort/pain, "Feels like I'm dying."

For extra material, including patient outcome, go to bmj.com/endgames

CPD READING 0.5 HOURS

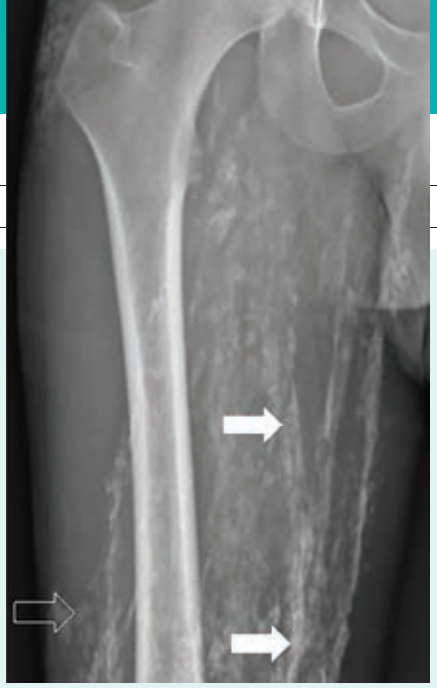
You can record CPD points for reading any article. We suggest half an hour to read and reflect on each.

LEARNING MODULE

Articles with a "learning module" logo have a linked BMJ Learning module at http://learning.bmj.com.

MINERVA

Hard lumps under the skin



A 16 year old boy had a two year history of skin lumps on his legs and progressive muscle weakness. Examination revealed a purplish rash on his eyelids, reduced power in the proximal upper and lower limb muscles, and disseminated, non-tender, subcutaneous, hard, nodular lumps appearing in clumps with no overlying skin abnormalities on his upper and lower legs. Radiography showed radio-opaque subcutaneous (clear arrow) and intra-muscular (white arrows) sheet-like calcific deposits. Inflammatory markers and muscle enzymes were raised.

Juvenile dermatomyositis with calcinosis was diagnosed. Although it is a clinical diagnosis, imaging assesses the extent of calcification and excludes calcinosis mimickers (eg, xanthomas, gouty tophi, subcutaneous cholesterol crystals, osteoma cutis, mycetoma).

Coziana Ciurtin (c.ciurtin@nhs.net); FangEn Sin; Margaret Hall-Craggs; Debajit Sen, University College London Hospitals NHS Trust, London, UK

Patient consent obtained.

Cite this as: *BMJ* 2019;365:l2291

If you would like to write a Minerva picture case, please see our author guidelines at http://bit.ly/29HCBAL and submit online at http://bit.ly/29yyGSx

Legislation to improve vaccination rates

In 2016 California tightened its immunisation requirements for schoolchildren. Parental requests for exemption because of personal or religious beliefs were no longer granted. This had an immediate impact and the proportion of children starting school without up-to-date vaccination decreased from 7% to 4% (*Paediatrics*). A year later however, the number of unvaccinated schoolchildren given an exemption on medical grounds increased. The likely explanation is that parents who don't want their children to be vaccinated find new ways to avoid it.

Dopaminergic therapy for stroke

Dopamine is a central neurotransmitter with an important role in movement control so it's not too far fetched an idea that dopamine agonists might facilitate recovery after stroke. Several small studies have reported benefit but, as so often happens, the hypothesis doesn't survive testing in a properly powered randomised trial. Among nearly 600 patients randomly assigned either to co-careldopa an hour before physiotherapy sessions or to physiotherapy alone there were no signs that the drug improved the ability to walk independently, which was the primary outcome, or reduced mortality at 12 months (*Lancet Neurol*).

Misreporting of trial results

A survey of the abstracts of recently published papers in the field of emergency medicine identified 114 randomised controlled trials in which the primary endpoint hadn't been statistically significant. These trials should have been reported as negative or inconclusive. However, in nearly half some sort of positive interpretation had been put on the results either by focusing on secondary endpoints and the results of subgroup analyses or by claims to have achieved an objective that had never been a prespecified outcome (*J Ann Emerg Med*).

Diagnostic test for chronic fatigue syndrome

The pattern of electrical impedance shown by peripheral blood mononuclear cells in response to hyperosmotic stress differs between healthy controls and patients with chronic fatigue syndrome, according to a study from Stanford, California. The investigators hope that their nano-electronic assay will provide a cheap and non-invasive diagnostic test for the condition (*PNAS*). Minerva thinks that they used the wrong comparison group. What's important is not whether the test can tell healthy and ill people apart but if it can distinguish between chronic fatigue syndrome and other conditions in which symptoms of fatigue and tiredness are prominent.

Unexpected antibacterial activity of ticagrelor

Ticagrelor is an antiplatelet drug widely used in the secondary prevention of cardiovascular disease. It may have other benefits as well. Retrospective analysis of trial data found that patients treated with ticagrelor had a lower risk of infection related death (*JAMA Cardiol*). In vitro assays and experiments in mice have now shown that ticagrelor has bactericidal activity against multi-resistant staphylococci and enterococci. The pharmacokinetics of ticagrelor in the mouse differ from those in humans but nonetheless it's an observation that's worth taking further.

Cite this as: *BMJ* 2019;365:l2427