



This is a repository copy of *A systematic review of international performance indicators and metrics relevant to UK general practice*.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/232124/>

Version: Published Version

Article:

Chambers, D. orcid.org/0000-0002-0154-0469, Mawson, R., Mettle-Nunoo, J. et al. (2 more authors) (2025) A systematic review of international performance indicators and metrics relevant to UK general practice. *BMJ Open Quality*, 14 (4). e003477. ISSN: 2399-6641

<https://doi.org/10.1136/bmjoq-2025-003477>

Reuse

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:

<https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>

A systematic review of international performance indicators and metrics relevant to UK general practice

Duncan Chambers ¹, Rebecca Mawson,² Justina Mettle-Nunoo,¹ Anthea Sutton,¹ Andrew Booth ¹

To cite: Chambers D, Mawson R, Mettle-Nunoo J, et al. A systematic review of international performance indicators and metrics relevant to UK general practice. *BMJ Open Quality* 2025;14:e003477. doi:10.1136/bmjog-2025-003477

► Additional supplemental material is published online only. To view, please visit the journal online (<https://doi.org/10.1136/bmjog-2025-003477>).

Received 31 March 2025
Accepted 21 September 2025



© Author(s) (or their employer(s)) 2025. Re-use permitted under CC BY. Published by BMJ Group.

¹Sheffield Centre for Health and Related Research (SCHARR), The University of Sheffield, Sheffield, England, UK

²School of Medicine and Population Health, The University of Sheffield, Sheffield, UK

Correspondence to

Dr Duncan Chambers;
d.chambers@sheffield.ac.uk

ABSTRACT

Background A wide variety of performance indicators/metrics are used to measure the performance of healthcare systems and to promote quality improvement (QI). We sought to identify indicators relevant to QI and organisational development (OD) within primary care/general practices and to evaluate the evidence for their use in QI and OD interventions in UK general practice and primary care.

Methods We used a framework based on UK National Health Service primary care documents to structure the review. Separate literature searches were performed in four databases to identify relevant reviews and primary studies. Studies were included if (1) the main focus was a metric or indicator that fell within the review framework or (2) they reported an OD or QI initiative or intervention in UK primary care that used one or more of the previously identified metrics or indicators. We mapped studies in group 1 against our framework domains. We performed a narrative synthesis of studies in group 2, again organised by the overall framework.

Results We included 28 studies, 24 (11 reviews and 13 international primary studies) for metrics or indicators and 4 for initiatives or interventions. The number of individual indicators or groups of indicators in group 1 studies ranged from 1 to 773. Three of the four UK QI/OD studies focused on initial access to general practice services; the other dealt with a programme to encourage self-care for long-term conditions. Mapping of the group 1 studies identified potentially relevant indicators across all domains but the process was methodologically challenging.

Conclusions Although numerous potential indicators exist, they tend to be poorly defined and lack examples of their use in practice. Further work is needed to identify and evaluate candidate indicators.

BACKGROUND

In the UK National Health Service (NHS), general practitioners (GPs) and general practice staff provide primary medical care for acute and long-term conditions and refer patients for further investigation/treatment as required. A large percentage of the population is registered with a general practice and GPs treat patients throughout their lifespan. Primary care also covers community pharmacy, dental and optometry services. There are currently more than 6000 GP practices

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Performance indicators and metrics are widely used to support quality improvement (QI) in general practice and primary care, mainly focusing on delivery of clinical services.

WHAT THIS STUDY ADDS

⇒ Our systematic review identified many studies of any indicators and groups of indicators that could be used to support internal (non-clinical) QI initiatives. However, we found limited evidence of such indicators being used in UK general practice and primary care.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ Our findings should prompt a discussion about the measures and metrics that really matter to different stakeholders, for example, policymakers, frontline primary care staff, patients, clinicians and researchers, and how to optimise their use in practice.

of varying sizes in England, and collectively, these now deliver over 360 million patient appointments per year (<https://www.rcgp.org.uk/representing-you/key-statistics-insights#appointments>).

Given the central role of general practice in the NHS, and increasing demand pressures on the workforce, it is important to collect and analyse data on the performance of the system and identify areas for improvement. Performance indicators have become increasingly prominent since the 1990s¹ and the Quality and Outcomes Framework (QOF), introduced in 2004, linked general practice funding to the achievement of targets against a range of predominately clinical indicators. A 2015 review by the Health Foundation of general practice quality indicators in England² argued that the publication of indicators provides transparency; supports informed choice and helps to empower patients and service users; supports accountability and performance management; and helps researchers to investigate workings of

the healthcare system. A review of QOF was conducted in 2018 and recommended the inclusion of a quality improvement (QI) domain.³ A consultation on primary care incentives was launched in March 2024,⁴ although the results of this are not yet available, possibly because of the change of government in July 2024. Therefore, how to incentivise and measure quality improvement and performance in primary care remains a policy priority.

However, primary care performance is multidimensional, encompassing aspects such as clinical quality, access, patient experience, staff satisfaction, continuity of care, efficiency and equity. Despite the existence of various descriptive and analytical frameworks,^{5–7} capturing this complexity through a comprehensive set of indicators is challenging. While quantitative indicators like QOF targets and appointment availability provide valuable information, incorporating patient-reported outcome measures (PROMs) and patient-reported experience measures (PREMs) can offer a more holistic understanding of the quality of care from the patient's perspective. While national indicators can facilitate benchmarking and comparisons between health providers, and used to target improvement efforts and interventions, it is equally important to allow for local adaptation and flexibility to address the specific population health needs, priorities of individual practices and requirements of local health systems and regions, including tackling health inequalities, as highlighted by an evaluation of the introduction of QI modules into the QOF from 2019.⁸ Involving different stakeholders, including healthcare professionals, patients and policymakers, in the development and implementation of performance indicators can further enhance their relevance, acceptance and usefulness.

While performance indicators can drive improvement, there is a risk of unintended consequences, such as narrowing the focus to only measurable aspects of care or encouraging gaming behaviours, especially when indicators are linked to financial payment and national targets. The availability of high-quality data and up-to-date information from across different parts of the health system is also key to successful implementation and use.

Difficulties experienced by patients attempting to access general practice services have been increasingly highlighted in recent years. Access can be measured quantitatively in terms of indicators like availability of same-day appointments or time to answer telephone calls. However, some researchers have argued that this approach fails to capture the real-world complexity of patients' interactions with the healthcare system.⁹ Others have pointed out that general practices and primary care networks vary markedly in terms of contextual factors such as practice size, staffing levels, urban or rural location and social and economic deprivation.^{8 10 11} The presence of deprivation may lead to a practice's performance appearing worse based on indicators such as the QOF or the measures used by the Care Quality Commission (CQC).¹⁰ Furthermore, appointment activity data does not fully capture the operational and business activity associated with running and

managing a general practice, such as administrative tasks, digital transformation and handling ad hoc requests.

The current policy context in the UK NHS implies a need for improved measurement to support quality improvement initiatives in general practice. For example, NHS England's General Practice Improvement Programme (GPIP), which started in 2023, offers tailored support to general practices and Primary Care Networks (PCNs, groups of linked general practices) to encourage improvement and change ways of working to help manage demand and align to a 'modern general practice access' model (<https://www.england.nhs.uk/gp/national-general-practice-improvement-programme>). We sought to identify and characterise relevant indicators from the international literature and to evaluate studies from the UK that involved QI/organisational development (OD) interventions or initiatives and measured impact on 'real world' service quality. An important aim of the review was to identify the types of measures available for improvements in non-clinical areas given a preponderance of measures traditionally focused on clinical outcomes in patients rather than practice processes.

METHODS

Protocol and registration

The protocol was registered with PROSPERO (registration number CRD42023481064).

The full protocol is available via the funder's website (<https://fundingawards.nihr.ac.uk/award/NIHR164763>). Variations from the protocol involved modifying the framework used for synthesis and using the study rather than the indicator as unit of analysis.

Review questions

The review questions were as follows:

1. Which non-clinical metrics and indicators have been used to quantify service improvements in general practice and primary care internationally and which aspects of services do they address?
2. Which quality improvement or organisational development interventions/initiatives have been delivered in general practice and primary care in the UK to address these improvement domains?

Eligibility criteria

Studies were included if (1) the main focus of the study was a metric or indicator of primary healthcare activity that corresponds to one or more of the domains of the framework described below (see 'Synthesis methods') and/or (2) they report an organisational development or quality improvement initiative or intervention that seeks to achieve change against one or more of the previously identified metrics or indicators. Detailed inclusion and exclusion criteria are presented in [table 1](#).

Information sources

Two searches were conducted: (1) to identify systematic reviews and (2) to identify primary research. Searches for

Table 1 Inclusion and exclusion criteria

	Inclusion	Exclusion
Participants/population	General practice or primary healthcare staff and service users	Dental and community pharmacy staff and patients
Intervention/exposure	1. Metrics and indicators (qualitative or quantitative) of non-clinical quality improvement covered by the review framework (indicator groups containing clinical and non-clinical indicators were included) 2. Quality improvement and organisational developments in UK general practice and other primary care settings addressing any of the above domains	Metrics and indicators specific to a specific disease or condition; metrics of capability building (ie, data-driven improvement, culture, leadership, governance, digital capabilities)
Comparator(s)	1. Metrics and indicators from high-income countries with comparable primary care/general practice systems to the UK (Australia, Canada, Ireland, New Zealand, Scandinavia and Netherlands). 2. Comparative and non-comparative studies	Individual metrics/indicators developed in the USA
Outcomes 1.	Feasibility and usefulness of the metric or indicator; external validity; face validity; ease of interpretation; limitations of the metric/indicator; staff satisfaction with metrics and indicators; patient/carer acceptance of the metric or indicator	
Outcomes 2.	Facility to address specific metrics or indicators Effect size or proportion of change achieved Patient acceptability of the initiative/intervention Staff acceptance of the initiative/intervention Mitigation of inequalities through the intervention/initiative Feasibility of providing data collected systematically and regularly and used to inform improvements Additional outcomes: details of development of the metric or indicator; process data on implementation; resource costs/time/time savings (if present)	
Study type	1. Any study type 2. Studies must report research (quantitative or qualitative; from research, evaluation or audit)	
Other	Peer-reviewed published research or substantive 'grey literature' reports	Studies focusing on low-income and middle-income countries Conference abstracts, theses or dissertations

systematic reviews were conducted on Ovid MEDLINE, Embase via Ovid, HMIC (Health Management Information Consortium database) via Ovid, Scopus, Science and Social Sciences Citation Indexes via Web of Science. Due to the rapid nature of the review, a condensed number of information sources were prioritised to identify primary research; Ovid MEDLINE, Embase via Ovid, Science and Social Sciences Citation Indexes via Web of Science. Grey literature was also searched, specifically for primary research.

Search strategy

Systematic reviews search (Search 1)

The search for systematic reviews aimed to map the metrics and indicators used to quantify service improvements in general practice and primary care. It was conducted between 27 October and 28 November 2023. Search terms (subject headings and free-text) were grouped according to three concepts: general practice/

primary care setting, metrics and indicators, and NHS England's GPIP domains. A methodological search filter to identify systematic reviews was combined with the topic search terms. No search limits were applied.

Primary research studies search (Search 2)

The second search aimed to identify primary research on quality improvement or organisational development interventions/initiatives delivered in general practice and primary care in the UK to address the GPIP improvement domains. It was conducted between 28 November and 18 December 2023. Search terms (subject headings and free-text) were grouped according to three concepts: general practice/primary care setting, metrics and indicators, and NHS England's GPIP domains. A geographical search filter to identify studies conducted in the UK was combined with the topic search terms. Searches were limited to 2013 onwards. The following websites were searched for grey literature: The Health

Foundation, NHS England and The Strategy Unit. The TRIP (Turning Research into Practice) database was searched, along with the preprint server medRxiv. The search engines Google and Google Scholar were also searched using terms relating to metrics, indicators and general practice.

Sample search strategies are presented in Appendix 1 and the full set of search strategies is available as online supplemental file 1.

Study selection process

Study selection at the title and abstract stage involved dual independent selection by two reviewers from 10% of the records. Following confirmation of an acceptable level of consistency, remaining records were screened by a single reviewer. At full-text screening, studies were assessed against the inclusion criteria by two reviewers acting independently; disagreements were resolved by consensus among the review team. Included studies were filtered into those reporting the indicator/metric and those that reported use of a metric/indicator within a primary care quality improvement or organisational development initiative/intervention.

Data collection process

For the review of metrics/indicators, our initial aim was to extract data using the metric/indicator as the unit of analysis, supported by all identified publications relevant to that metric or indicator. However, a higher than anticipated number of indicators meant that it was only feasible to extract data using the study as the primary unit of analysis. Available data on individual indicators were extracted from primary research studies providing relevant information. We also extracted data on relevant indicators (ie, excluding condition-specific indicators) from the supplementary file published with the umbrella review by Ramalho *et al.*¹² Interventions/initiatives were analysed by each study describing that individual initiative/intervention.

Data were extracted by one reviewer and checked by a second reviewer using prespecified forms in Microsoft Excel or Word tailored to the amount and type of data to be extracted.

Data items

For review question 1, we extracted data on study design and methodology, metric/indicator name, characteristics, organisational context (eg, urban/rural, practice size (patient list), workforce, practice-level Index of Multiple Deprivation (IMD)), validation, together with any measures of effect (where applicable).

Data items for review question 2 were study characteristics, setting, intervention components, metrics used and effect size or change achieved.

We also extracted data where possible on the prespecified outcomes for each group of studies (see table 1 for list of outcomes).

Risk of bias assessment

There is no formal method for assessing studies reporting metrics and indicators. Instead, we focused on the individual metrics or indicators themselves, assessing information from any studies that contributed to an understanding of that metric or indicator. Specifically, we based our assessments on the Agency for Healthcare Research and Quality (AHRQ) quality indicator properties (<https://www.ahrq.gov/talkingquality/measures/measure-questions.html>). The AHRQ criteria measure standardisation, comparability across organisations, data availability, timeliness, relevance to target audience, validity, experience, stability, evaluability, distinguishability and credibility.

Studies of QI/OD interventions were not formally assessed for risk of bias because our focus was on the role of indicators/metrics. Where data allowed, organisational development and quality improvement interventions and initiatives were evaluated according to the following predefined measures of intervention effectiveness:

- ▶ Facility to address specific metrics or indicators.
- ▶ Effect size or proportion of change achieved.
- ▶ Patient acceptability of the initiative/intervention.
- ▶ Staff acceptance of the initiative/intervention.
- ▶ Mitigation of inequalities through the intervention/initiative.
- ▶ Feasibility of providing data collected systematically and regularly and used to inform improvements (ie, quality control and assurance, eg, quality scorecards, dashboards, clinical audit, patient surveys).

Synthesis methods

We conducted a narrative synthesis grouping studies by the two review questions. Studies relevant to question 1 were used to develop a 'map' of available indicators. Studies relevant to question 2 were used to evaluate the quantity and quality of UK-based QI/OD interventions. For both syntheses, we grouped studies using a framework based on the GPIP practice level support framework (<https://www.england.nhs.uk/gp/national-general-practice-improvement-programme/support-level-framework>) validated by reference to key NHS primary care documents. The framework had six domains: access to primary care systems, access to primary care clinical services, care navigation and triage, managing demand and capacity, managing the practice workload, and an additional category covering key objectives including time and cost savings and improvements in patient experience.

We anticipated that features of the metrics and indicators, and associated interventions, would be heterogeneous. We planned to investigate this heterogeneity using (1) the outcomes presented in table 1 and (2) the metric/indicator properties used to assess metric/indicator quality, and to present synthesised data in the form of tables, narrative summaries and diagrams as appropriate.

Public involvement

We discussed the project with members of the standing public advisory group for the Sheffield HS&DR Evidence Synthesis Centre. We used the domains listed above as a focus for the meeting to identify the aspects of most importance to patients. We also asked the group for suggestions in relation to the evaluation of QI/OD programmes in general practice.

Three members of the public advisory group of diverse age, sex/gender and ethnicity provided feedback. Members identified access to primary care systems as being particularly important, with continued availability of face-to-face appointments seen as a key metric. Participants stressed the importance of requesting feedback from patients by regular surveys as a means of evaluating initiatives to improve quality of general practice and other primary care services. Participants also emphasised the dramatic changes that have taken place in general practice following the COVID-19 pandemic and suggested that a distinction between research from before and after the pandemic would be helpful in evaluating the relevance of the indicators used.

RESULTS

Results of literature search

The results of the literature search and screening process are summarised in the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 flow diagram (figure 1).

We included a total of 28 studies (30 publications: 28 peer-reviewed and 2 grey literature reports), of which 24 addressed review question 1 (mapping of available indicators) and 4 (6 publications) addressed question 2 (use of indicators in UK QI/OD interventions and initiatives).

Mapping of relevant metrics/indicators

Map of studies

The map of available indicators drew on data from 11 systematic and non-systematic reviews (including one umbrella review¹² and 13 primary research studies). The number of individual indicators or groups of indicators covered ranged from 1 to 773. Studies varied widely in scope and methodology and covered diverse topics related to measuring and improving quality in primary care^{13–15 16 17–20} (see online supplemental file 2).

Map of indicators

Indicators included in the umbrella review by Ramalho *et al*¹² are summarised in online supplemental file 3. Of 727 indicators included in the umbrella review, 15 mapped to our framework domains 1a and 1b, 7 to domain 2, 6 to domain 3, 1 to domain 4 and 8 to domain 5. A considerable majority of indicators measured clinical processes such as prescribing and monitoring of long-term conditions.

Summary description of indicators

Most included studies provided descriptions of indicators in terms of how they are calculated in the text (eg, Alsabbagh 2020²¹), in supplementary files or by reference to earlier papers. However, the predefined 'indicator outcomes' for this review were less commonly reported. It should also be borne in mind that most of the indicators and metrics included in reviews that addressed our review question 1 covered clinical aspects of general practice and primary care outside the scope of the review. For these reasons, we extracted 'indicator outcome' data from four studies^{22–25} (table 2).

Quality of indicators

As noted above, the study (covering one or multiple indicators) rather than the indicator (using information from multiple studies where available) was designated as the preferred unit of analysis. Ten primary studies reporting on indicators or groups of indicators were assessed against the AHRQ criteria (online supplemental file 4). While data were lacking for some items, most indicators were assessed positively for relevance, validity, evaluability and credibility.

Use of relevant metrics/indicators in UK QI/OD interventions or initiatives

Study characteristics

The four included UK quality improvement studies (table 3) evaluated four diverse interventions aimed at improving general practice processes: online consultations,²⁶ GP-led or nurse-led telephone triage,²⁷ promotion of self-care²⁸ and use of shorter prebookable review appointments for patients needing follow-up after an initial appointment.²⁹ Three of the studies had a control group and one was a cluster randomised controlled trial.

Indicators used in the included studies (table 3) mainly focused on measures of access (domains 1a and 1b, for example, numbers of face-to-face, online and telephone appointments; consultation length; number of same-day appointments) and practice workload (domain 4, primarily administrative staff workload^{26 29}). The study of self-care for people with long-term conditions²⁸ differed from the others in presenting a mainly qualitative summary of findings (supported by relevant quotations), although it was clear from the description of the programme that quantitative data were also collected.

Additional QI work to support implementation of the study interventions was reported in three studies.^{27–29}

Intervention effectiveness

Interventions (including both the primary intervention and supporting QI work) showed no clear pattern of effects (in terms of magnitude or direction) on the reported indicators (table 3). The indicators used were generally appropriate for the intervention being evaluated and quantitative measures of change were captured in three studies.^{26 27 29} Patient and staff acceptability was addressed in three studies,^{26 27 29} with additional measures

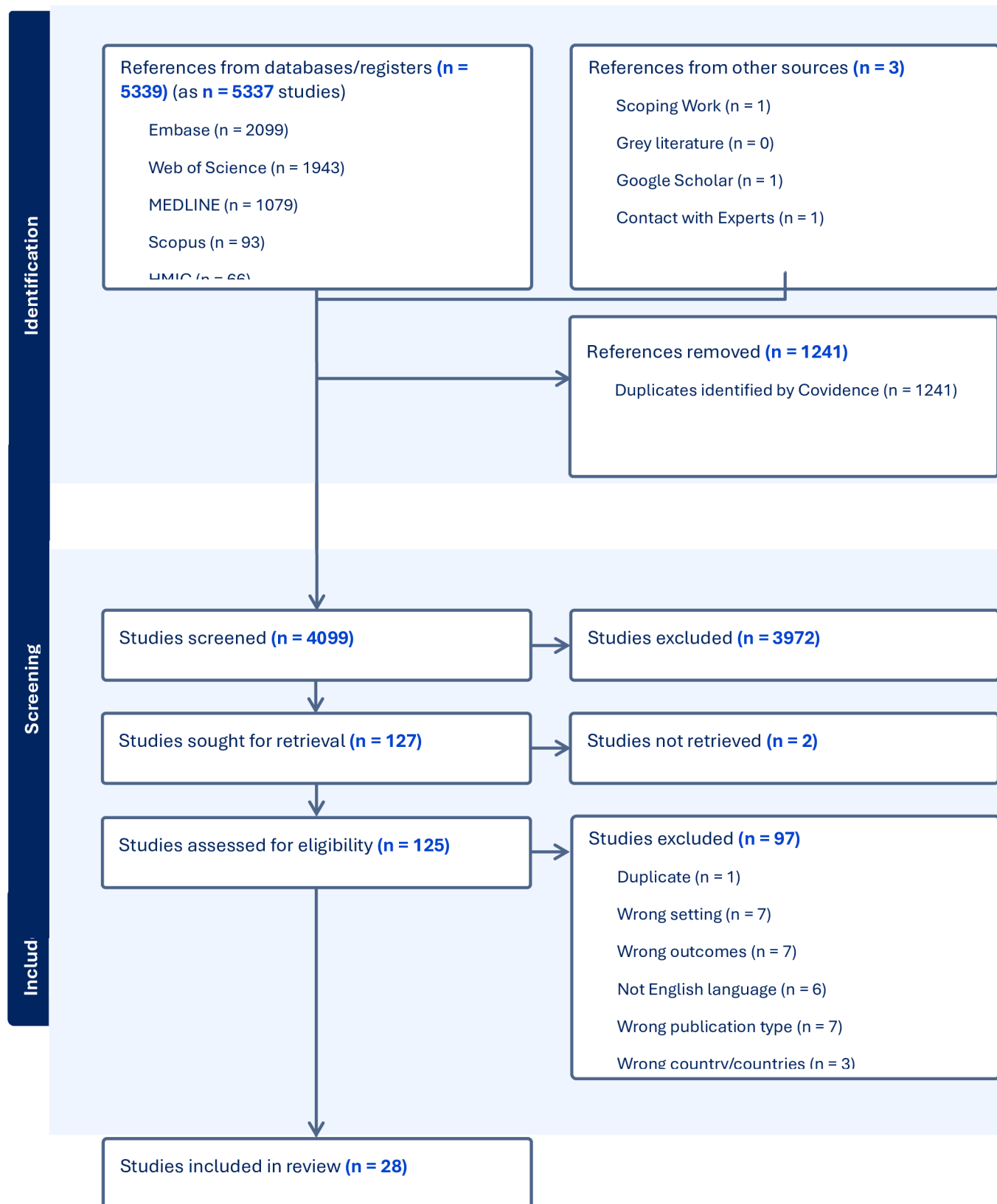


Figure 1 PRISMA 2020 flow diagram. HMIC, Health Management Information Consortium.

of workload in two studies.^{26 29} None of the included studies reported explicitly on mitigation of inequalities by the intervention. However, introduction of electronic consultations²⁶ may risk increasing inequality in access

at the expense of patients without access to digital technology. The other reported interventions could potentially reduce inequality by matching appointments to need^{27 29} and by prioritising a vulnerable population

Table 2 Summary of 'indicator outcomes'

	Indicator and framework domain(s)* covered	Feasibility and usefulness	External validity	Face validity	Ease of interpretation	Limitations	Staff satisfaction	Patient/carer acceptance
Crossland 2014 ²²	Primary Care Practice Improvement Tool (PC-PIT) Domain 2	19/28 participants rated PC-PIT as useful	Participants emphasised the relevance of the PC-PIT to everyday practice work and planning	19/28 said PC-PIT captured key elements of practice organisation and function	Required a high reading age; many administrative and reception staff found it difficult to understand	Some simplification required	Paper reports pilot study Administrative staff less satisfied than GPs	NR
Howie 2000 ²³	Consultation Quality Index (CQI) Domain 5	Considered useful for assessing quality of care by doctors and practices; requires patients to complete a questionnaire	Results were independent of practice case mix and level of deprivation	Authors state results had strong face validity	Divided into bands 1 (least good) to 6 (best) so relatively easy to interpret	Seen as one of a range of quality measures	NR	NR but definition of need and assessment of outcome were both determined exclusively by patient responses.
Levesque 2012 ²⁴	Group of 17 validated tools Domains 1b; 2; 5	Provides data to assess potential usefulness	Tools had undergone different levels of quantitative and qualitative validation		NR	Important attributes of primary care were not covered by any of the validated tools	NR	NR but all instruments aimed to evaluate quality from the patient perspective
Sidaway-Lee 2019 ²⁵	St Leonard's Index of Continuity of Care (SLICC) Domain 5	Described as simple to calculate and useful because continuity of care has benefits for patients	NR but includes all appointments	NR but comparison with UPC (Usual Provider of Care) index provided	High (percentage of appointments with patient's personal doctor)	Study involved one practice where all GPs were part-time	NR but authors note that GPs have incentive to record appointments accurately for workload monitoring	NR
*1a, access to primary care systems; 1b, access to primary care clinical services; 2, care navigation and triage; 3, managing demand and capacity; 4, managing the practice workload; 5, key outcome objectives. GPs, general practitioners; NR, not reported.								

Table 3 Summary of UK quality improvement intervention studies

Study reference	Study design	Setting	Details of QI intervention or initiative	Metric(s) or indicator(s) used	Results (effect size or change achieved, relevant to GPIIP domains)
Carter 2018 ²⁶	Mixed method study	England, UK	Intervention: webGP (online consultation) Comparator: face-to-face consultations	Volume of consultations Patient surveys Staff surveys Administrative staff workload	No impact on practice workload 72% of webGP requests required face-to-face or telephone consultation Good survey responses from patients regarding timeliness and quality of care (low response rate—35.1%) Partial acceptance of intervention by staff with concerns of negative interactions with pre-existing practice systems.
Murdoch 2015 ²⁷	Cluster randomised controlled trial	England, UK	Intervention: GP-/nurse-led telephone triage Comparator: usual care	Patient safety Patient satisfaction Healthcare costs Patient-provider communication Staff surveys	28% reduction in patient-GP contact time 31% reduction in GP face-to-face contacts, and a relative reduction of 1.4 min in overall GP contact time Nurse triage costs equivalent to both GP triage and usual care
Robertson 2013 ²⁸	Mixed method study	England, UK	Intervention: Supporting Self-Care in General Practice programme	GP appointment time duration Telephone consultation Online consultations Patient-centred care	NR (mainly qualitative data reported)
Slater 2021 ²⁹	Mixed method study	Scotland, UK	Intervention: shorter prebookable review appointments Comparator: 'On the day' review appointments	Patient accessibility to care Number of GP appointments per day Efficiency of appointment system Number of calls per day Continuity of care Administrative staff workload	43% increase in available appointments (554 to 792 appointments per 2-week period) Increased number of patients seen at the GP Median 93% same day appointments Administrative staff workload remained the same in terms of number of phone calls received per day Administrative staff workload reduced in terms of dealing with frustrated patients Improved staff survey results Increased rate of DNAs (did not attend) with appointments prebooked

GP, general practitioner; GPIIP, General Practice Improvement Programme; QI, quality improvement.

of patients with long-term conditions.²⁸ All four of the included studies included data that could be collected systematically and used to inform further improvements.

Synthesis across evidence sources

Online supplemental file 5 summarises the relationship between indicators mapped from the literature and those used in UK studies. For domains 1a and 1b, focusing on access, the UK QI studies used a limited number of measures, primarily related to number, type (face-to-face vs online) and duration of consultations. The Ramalho umbrella review identified other potentially useful indicators (eg, use of urgent appointments; practice opening hours; out-of-hours services; and home visits). There was little overlap between the umbrella review and UK studies. Metrics and indicators related to telephone access were not present in the umbrella review and were only present in one of the UK studies.²⁹

With respect to care navigation and triage (domain 2), the UK studies gathered data from patient surveys. Ramalho *et al* also included indicators of patient compliance with advice received which could reflect quality of triage by general practice staff. Other indicators mentioned in the umbrella review were less clearly defined, for example, 'Differentiates appropriately between important and minor issues' and 'Does not raise patient expectations that recommendations will be implemented'.

The QI study by Slater *et al*²⁹ corresponded with managing demand and capacity (domain 3) as it aimed to implement and evaluate a more efficient appointment system. Ramalho *et al* mentioned some other indicators around this domain, but again, limited detail was available.

In terms of managing the whole practice workload (domain 4), two UK QI studies used administrative staff workload as an indicator.^{28 29} A similar indicator ('Development of the primary care workforce') was identified by Ramalho *et al*. The four questionnaires described by Schafer *et al*³⁰ included questions potentially relevant to this domain, although these were split between different questionnaires.

The final framework domain (domain 5) covers diverse key QI outcomes, such as improved patient experience. In the UK QI studies, this domain was represented by staff surveys in three studies.^{31–33} Ramalho *et al* identified other indicators in this domain, for example, around record keeping (online supplemental file 5) and our mapping review included several relevant studies, particularly the recent work of Benson *et al*^{34 35} and the older European Practice Assessment instrument.^{31 36}

DISCUSSION

Main findings

We identified large numbers of indicators/indicator groups but few were relevant to our review focus on indicators of improvement in practice processes; most

measured clinical aspects of primary care such as treatment of specific conditions, monitoring of long-term conditions or prescribing. Indicators often appeared to be designed for academic research rather than for pragmatic general practice, that is, to measure effects of interventions in studies rather than supporting front line improvement.

The UK QI/OD literature was very limited, in contrast to plentiful examples of initiatives and interventions focusing on clinical indicators. The four included studies covered a diverse group of interventions and supporting QI work. Three of the studies focused on initial access to GPs' services,^{26 27 29} reflecting the distribution of indicators seen in the wider literature (online supplemental file 5) and the views expressed by our patient/public involvement (PPI) group.

The balance of the literature was skewed towards reviews and descriptions of groups of indicators, with relatively few studies focusing on a single metric or indicator. This meant that reporting of indicator properties was varied and made it difficult for us to use the indicator as the unit of analysis as originally planned. The process of identifying indicators at different levels of literature (umbrella reviews, reviews and primary studies) was challenging. The umbrella review by Ramalho and colleagues was helpful but most of the included reviews were not specific to non-clinical QI in primary care or general practice.

The indicators and metrics identified could be broadly grouped into those intended for internal use within practices (eg, to support ongoing quality improvement work) and those intended for higher level decision-makers and/or patients and the public. Single measures tended to fall into the first category^{23 25} while decision-makers and the public were offered diverse groups of measures, sometimes identified as 'monitors' or 'dashboards'.^{37 38} None of the included groups of indicators corresponded closely with domains of the framework used for this review or focused mainly on non-clinical aspects of general practice. This suggests that evaluation of large-scale QI initiatives would require purposive selection among existing indicators and/or development of new ones.

Strengths and limitations

Strengths

The review has a number of strengths. We used a dual approach to map the available indicators against those that have been used in practice in the UK. Identification and re-analysis of a subset of relevant indicators among the large number included in an umbrella review increased the yield and added value by increasing the relevance of the review to QI in UK general practice. We built on standard systematic review processes by use of novel methods to characterise indicators/metrics per se and in relation to their role in QI interventions/initiatives.

Limitations

We acknowledge the large and diverse literature covering QI in primary care (and healthcare generally) which

touches on different dimensions of improvement. It is likely that a large number of indicators used in the literature are clinical and have been driven by policy levers (eg, QOF and CQUIN (Commissioning for Quality and Innovation)) and clinical academic research. Indicators related to team climate and culture³⁹ were excluded at the protocol stage to manage the size of the review and could be an area for future review work. Use of the study rather than the indicator as the unit of analysis limited our ability to assess indicator quality and should be taken into account in interpreting the findings. The limited coverage of this review reflects the prespecified review questions and the limited number of QI/OD interventions/initiatives identified from the UK. Finally, although we endeavoured to exclude individual indicators from the USA, the international literature was inevitably dominated by evidence from primary care systems very different from those of the UK nations. This should be borne in mind when evaluating indicators and associated incentives derived from social insurance-based (like most of Europe) or fee for service-based (common in the USA) primary care models.

Implications for decision makers

For policymakers, commissioners and practitioners, further work is needed to cover operational and other improvement domains in primary care and associated metrics. Our findings suggest that there may be a need for a discussion about the measures and metrics that really matter to different stakeholders, for example, policymakers, frontline primary care staff, patients, clinicians and researchers. It appears that there could be a disconnect between the academic literature and QI metrics reported and what is actually needed to track quality improvement and transformation/change over time in primary care practice, in a way that is meaningful to patients and staff. The increasing emphasis in UK practice on delivery of QI through PCNs, and increasing co-operation and co-location of GPs with other health professionals, also support the need to rethink how QI is delivered, measured and reported. Our findings also support the need for caution when using indicators to compare practices serving different population groups, which can penalise practices in more deprived areas (by around 7% when weighted for need in one recent study⁴⁰).

Situations also arise where indicators may conflict, for example, staff satisfaction and patient satisfaction. Happy staff in a practice with clear boundaries around access might mean unhappy patients who see themselves as denied the access they want. The challenge for decision-makers is to find a set of measures to drive improvement efforts (as opposed to performance management) which give an accurate picture of what modern day practice life and activity looks like, from the view of patients and staff; for example, unlimited patient access to a GP is not an achievable aim. A report from the National Institute for Health Research Collaboration for Leadership in Applied Health Research and Care (NIHR CLAHRC) South West

Peninsula (excluded from the review because of lack of empirical data) examined the SERVQUAL approach, which measures service quality using self-report.⁴¹ This approach views quality as a subjective concept, making it amenable to use with quality impact assessment.

The National Voices patient organisation published 'A vision for the future of primary care' in June 2023, identifying nine priority areas for improvement.⁴² We initially considered a broad agenda such as this as the framework for our review but subsequently agreed to focus on a more limited group of domains. Our approach to mapping indicators and metrics against studies could potentially be adapted and extended to cover additional aspects of primary care and general practice.

CONCLUSIONS

We identified large numbers of indicators/indicator groups but few were relevant to our review questions. The UK QI/OD literature was very limited, although the four included studies covered a diverse group of interventions and supporting QI work. Findings of our mapping review suggest that a large number of indicators used in the literature are clinical and have been driven by policy levers (eg, QOF and CQUIN (Commissioning for Quality and Innovation)) and clinical academic research. Further work is needed to identify and evaluate candidate indicators from the dispersed literature and to optimise their value for patients, practitioners and decision-makers concerned with measuring quality in general practice and primary care.

Acknowledgements We thank the members of the public advisory group for their insights and suggestions, and Dr Jean Ledger for comments on an earlier draft of this paper.

Contributors DC: project co-ordination; study selection; data extraction and analysis; report writing; wrote first draft of the paper. RLM: study selection; data extraction and analysis; report writing; general practice expertise. JMN: study selection; data extraction and analysis; report writing. AS: literature searching; report writing. AB: report writing; methodological expertise. AB and DC developed the review protocol with input from AS. All authors contributed to writing the paper and have approved the version to be submitted for publication. DC is the guarantor.

Funding This study was funded by Health Services and Delivery Research Programme (NIHR164763).

Competing interests None declared.

Patient and public involvement Patients and/or the public were involved in the design, or conduct, or reporting, or dissemination plans of this research. Refer to the Methods section for further details.

Patient consent for publication Not applicable.

Ethics approval Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement All data relevant to the study are included in the article or uploaded as supplementary information.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines,

terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution 4.0 Unported (CC BY 4.0) license, which permits others to copy, redistribute, remix, transform and build upon this work for any purpose, provided the original work is properly cited, a link to the licence is given, and indication of whether changes were made. See: <https://creativecommons.org/licenses/by/4.0/>.

ORCID iDs

Duncan Chambers <http://orcid.org/0000-0002-0154-0469>

Andrew Booth <http://orcid.org/0000-0003-4808-3880>

REFERENCES

- 1 McColl A, Roderick P, Gabbay J, *et al*. Performance indicators for primary care groups: an evidence based approach. *BMJ* 1998;317:1354–60.
- 2 Dixon J, Spenceley H, Howells A, *et al*. Indicators of Quality of Care in General Practices in England. London, England, 2015.
- 3 NHS England. NHS England: report of the review of the quality and outcomes framework in England. 2018.
- 4 NHS England. Department of health and social care: role of incentive schemes in general practice. 2024.
- 5 Bodenheimer T, Ghorob A, Willard-Grace R, *et al*. The 10 building blocks of high-performing primary care. *Ann Fam Med* 2014;12:166–71.
- 6 Senn N, Breton M, Ebert ST, *et al*. Assessing primary care organization and performance: Literature synthesis and proposition of a consolidated framework. *Health Policy* 2021;125:160–7.
- 7 Starfield B: New paradigms for quality in primary care. *Br J Gen Pract* 2001;51:303–9.
- 8 Bramwell D, Hotham S, Peckham S, *et al*. Evaluation of the introduction of QOF quality improvement modules in English general practice: early findings from a rapid, qualitative exploration of implementation. *BMJ Open Qual* 2022;11:e001960.
- 9 Voorhees J, Bailey S, Waterman H, *et al*. Accessing primary care and the importance of “human fit”: a qualitative participatory case study. *Br J Gen Pract* 2022;72:e342–50.
- 10 Fisher R, Dunn P, Asaria M, *et al*. Level or Not? Comparing General Practice in Areas of High and Low Socioeconomic Deprivation in England. In: London: The Health Foundation, 2020.
- 11 Morciano M, Checkland K, Hammond J, *et al*. Variability in size and characteristics of primary care networks in England: observational study. *Br J Gen Pract* 2020;70:e899–905.
- 12 Ramalho A, Castro P, Gonçalves-Pinho M, *et al*. Primary health care quality indicators: An umbrella review. *PLoS One* 2019;14:e0220888.
- 13 Brennan SE, Bosch M, Buchan H, *et al*. Measuring team factors thought to influence the success of quality improvement in primary care: a systematic review of instruments. *Implement Sci* 2013;8:20.
- 14 Derriennic J, Nabbe P, Barais M, *et al*. A systematic literature review of patient self-assessment instruments concerning quality of primary care in multiprofessional clinics. *Fam Pract* 2022;39:951–63.
- 15 Rendell N, Rosewell A, Lokuge K, *et al*. Common Features of Selection Processes of Health System Performance Indicators in Primary Healthcare: A Systematic Review. *Int J Health Policy Manag* 2022;11:2805–15.
- 16 Rhydderch M, Edwards A, Elwyn G, *et al*. Organizational assessment in general practice: a systematic review and implications for quality improvement. *Evaluation Clinical Practice* 2005;11:366–78.
- 17 Simou E, Pliatsika P, Koutsogeorgou E, *et al*. Quality Indicators for Primary Health Care: A Systematic Literature Review. *J Public Health Manag Pract* 2015;21:E8–16.
- 18 Stange KC, Etz RS, Gullett H, *et al*. Metrics for assessing improvements in primary health care. *Annu Rev Public Health* 2014;35:423–42.
- 19 Yapi SM, Poitras M-E, Donnelly C, *et al*. Identify and classify interprofessional primary care performance indicators: a scoping review protocol. *BMJ Open* 2023;13:e072186.
- 20 International Alliance of Patients’ Organizations. Patient-Centred Healthcare Indicators Review. London: International Alliance of Patients, 2012.
- 21 Alsabbagh MW, Kueper JK, Wong ST, *et al*. Development of comparable algorithms to measure primary care indicators using administrative health data across three Canadian provinces. *Int J Popul Data Sci* 2020;5:1340.
- 22 Crossland L, Janamian T, Sheehan M, *et al*. Development and pilot study of the Primary Care Practice Improvement Tool (PC-PIT): an innovative approach. *Med J Aust* 2014;201:S52–5.
- 23 Howie JG, Heaney DJ, Maxwell M, *et al*. Developing a “consultation quality index” (CQI) for use in general practice. *Fam Pract* 2000;17:455–61.
- 24 Levesque JF, Haggerty J, Beninguisse G, *et al*. Mapping the coverage of attributes in validated instruments that evaluate primary healthcare from the patient perspective. *BMC Fam Pract* 2012.
- 25 Sidaway-Lee K, Gray DP, Evans P. A method for measuring continuity of care in day-to-day general practice: a quantitative analysis of appointment data. *Br J Gen Pract* 2019;69:e356–62.
- 26 Carter M, Fletcher E, Sansom A, *et al*. Feasibility, acceptability and effectiveness of an online alternative to face-to-face consultation in general practice: a mixed-methods study of webGP in six Devon practices. *BMJ Open* 2018;8:e018688.
- 27 Murdoch J, Varley A, Fletcher E, *et al*. Implementing telephone triage in general practice: a process evaluation of a cluster randomised controlled trial. *BMC Fam Pract* 2015;16:47.
- 28 Robertson S, Witty K, Braybrook D, *et al*. “It’s coming at things from a very different standpoint”: evaluating the “Supporting Self-Care in General Practice Programme” in NHS East of England. *Prim Health Care Res Dev* 2013;14:113–25.
- 29 Slater J, Malik S, Davey P, *et al*. Improving access to primary care: a mixed-methods approach studying a new review appointment system in a Scottish GP practice. *BMJ Open Qual* 2021;10:e001279.
- 30 Schäfer WLA, Boerma WGW, Kringos DS, *et al*. Measures of quality, costs and equity in primary health care instruments developed to analyse and compare primary care in 35 countries. *Qual Prim Care* 2013;21:67–79.
- 31 Engels Y, Dautzenberg M, Campbell S, *et al*. Testing a European set of indicators for the evaluation of the management of primary care practices. *Fam Pract* 2006;23:137–47.
- 32 Haj-Ali W, Hutchison B, Primary Care Performance Measurement Steering Committee. Establishing a Primary Care Performance Measurement Framework for Ontario. *Healthc Policy* 2017;12:66–79.
- 33 Kringos D, Nuti S, Anastasy C, *et al*. Re-thinking performance assessment for primary care: Opinion of the expert panel on effective ways of investing in health. *Eur J Gen Pract* 2019;25:55–61.
- 34 Brennan SE, Bosch M, Buchan H, *et al*. Measuring organizational and individual factors thought to influence the success of quality improvement in primary care: a systematic review of instruments. *Implement Sci* 2012;7:121.
- 35 Benson T, Benson A. Routine measurement of patient experience. *BMJ Open Qual* 2023;12:e002073.
- 36 Obucina M, Harris N, Fitzgerald JA, *et al*. The application of triple aim framework in the context of primary healthcare: A systematic literature review. *Health Policy* 2018;122:900–7.
- 37 Breton M, Gaboury I, Bordeleau F, *et al*. Use of Electronic Medical Record Data to Create a Dashboard on Access to Primary Care. *Healthc Policy* 2023;18:72–88.
- 38 Kringos DS, Boerma WGW, Bourgueil Y, *et al*. The European primary care monitor: structure, process and outcome indicators. *BMC Fam Pract* 2010;11:81.
- 39 Beaulieu M-D, Haggerty J, Tousignant P, *et al*. Characteristics of primary care practices associated with high quality of care. *CMAJ* 2013;185:E590–6.
- 40 Fisher R. Levelling up’ General Practice in England: What Should Government Prioritise? London: The Health Foundation, 2021.
- 41 Manzi S. Examining Quality of Care and the Impact of Change on Care Quality in the Context of Healthcare. South West Peninsula: NIHR CLAHRC, 2021.
- 42 National Voices. A Vision for the Future of Primary Care. London: National Voices, 2023.