HealthPros
International Training Network for Healthcare Performance Intelligence Professionals

A PRACTICAL GUIDE TOWARDS ACTIONABLE HEALTHCARE PERFORMANCE INDICATORS:
Selecting healthcare performance indicators that are fit for purpose and use for various stakeholders

Healthcare Performance Intelligence Series No. 1.3 2022

HealthPros Fellows
Erica Barbazza
Pinar Kara
Nicolas Larrain
Mekha Matthew
Bernardo Meza-Torres
Mircha Poldrugovac
Claire Willmington
Óscar Brito Fernandes
Damir Ivanković

Principal Investigators
Dionne Kringos
Niek Klazinga
Simon de Lusignan
Fabrizio Carinci

Series editors

HealthPros Fellows

Erica Barbazza, Amsterdam UMC, University of Amsterdam

Pinar Kara, Aalborg University Hospital

Mekha Mathew, University of Surrey, University of Oxford

Nicolas Larrain, Optimedis AG

Kendall Gilmore, Sant'Anna School of Advanced Studies

Damir Ivanković, Amsterdam UMC, University of Amsterdam

Óscar Brito Fernandes, Corvinus University of Budapest

Sophie Wang, Optimedis AG

Bernardo Meza, University of Surrey, University of Oxford

Claire Willmington, Sant'Anna School of Advanced Studies

Mircha Poldrugovac, Amsterdam UMC, University of Amsterdam

Véronique Bos, Amsterdam UMC, University of Amsterdam

Principal Investigators

Niek Klazinga, Amsterdam UMC, University of Amsterdam

Dionne Kringos, Amsterdam UMC, University of Amsterdam

Jan Mainz, Aalborg University Hospital

Søren Paaske Johnsen, Aalborg University Hospital

Simon de Lusignan, University of Oxford

Fabrizio Carinci, University of Bologna

Oliver Gröne, Optimedis AG

Sabina Nuti, Sant'Anna School of Advanced Studies

Anna Maria Murante, Sant'Anna School of Advanced Studies

László Gulácsi, Corvinus University of Budapest

Petra Baji, Corvinus University of Budapest

Support staff

Frank Groen, Amsterdam UMC, University of Amsterdam

Laurian Jongejan, Amsterdam UMC, University of Amsterdam

PREFACE

HealthPros is a H2020 Marie Sklodowska-Curie Innovative Training Network for Healthcare Performance Intelligence Professionals under grant agreement No 765141, running from January 2018–April 2022. Healthcare performance intelligence can be defined as a structured approach to acting on health policies, using knowledge and information generated through scientific methods and health data to systematically measure indicators of health system performance. The network set out with the aim to train a first generation of Healthcare Performance Intelligence Professionals (HealthPros Fellows) that can make effective use of available healthcare performance data in countries to improve integrated services delivery, patient engagement, equality in access to healthcare, health outcomes and reduce waste in healthcare.

Since 2018, HealthPros Fellows have completed innovative research and multidisciplinary training in Canada, Denmark, Germany, Hungary, Italy, the Netherlands and the United Kingdom. As part of their training, Fellows also completed secondments at partner organizations as an opportunity to obtain local guidance and conduct applied research.

Throughout the programme, HealthPros Fellows have worked to develop tools and implement methods to streamline healthcare performance measurement, develop and apply performance-based governance mechanisms and optimize the use of healthcare performance intelligence by different end-users. Topics explored through a healthcare performance intelligence lens in their work include: actionability of performance indicators; composite measures; integrated care; corporate governance tools; patient and citizen engagement; nudging; use of routine databases for performance improvement; and, long-term care. As the COVID-19 pandemic paralleled the HealthPros programme, many Fellows and the network at-large, sought opportunities to conduct a number of COVID-19-related studies at pace with the pandemic's changing context.

Outputs of the HealthPros programme have continuously been published as open access studies in international, peer-reviewed journals. Additionally, Fellows have actively contributed to webinars, conferences, the delivery of courses, policy dialogues, direct country support, and media engagements, among other types of dissemination to continuously share new findings throughout the programme.

This **Healthcare Performance Intelligence Series** represents the culmination of key research findings by the network into a collection of reports providing methodological, practical, and policy guidance. Reports in the series are tailored to different audiences, ranging from policy-makers, hospital

A practical guide towards actionable healthcare performance indicators: Selecting healthcare performance indicators that are fit for purpose and use for various stakeholders; Healthcare Performance Intelligence Series No. 1.3 2022

EU H2020-Agreement No. 765141

HealthPros

managers, clinicians, and the general public. The development of each report in the series has relied on close collaboration across the HealthPros network. The range of topics and resources making up this series includes the following:

- Practical experience with implementing disparity and composite measures in large-scale routine quality improvement work to support transferability to other HC systems (No. 1.2 2022)
- A practical guide towards actionable healthcare performance indicators: Selecting healthcare performance indicators that are fit for purpose and use for various stakeholders (current)
- Policy guidance on advancing the performance assessment of integrated healthcare systems (No. 1.4 2022)
- Policy guidance on the use of PREMs to improve health system performance (No. 2.2 2022)
- Policy summary report on the value of results-based tools in health care management-Lessons learned from COVID-19 dashboards (No. 2.3 2022)
- Business model for effectively involving patients in the financial decision-making of health insurance funds- A guide to health care insurers on fostering the engagement of citizens based on recent experiences in the Netherlands. (No. 2.4 2022)
- Policy summary report on best practices for linking financial incentives to health care performance at individual health care provider, institutional and regional level- A business case for value-based health care systems based on performance intelligence (No. 2.5 2022)
- Policy recommendations on the role of nudging for health care performance assessment agencies (No. 3.2 2022)

The full series of reports can be found online (https://www.healthpros-h2020.eu/). For questions related to the series or HealthPros network please contact Dionne Kringos, PhD (d.s.kringos@amsterdamumc.nl).

A practical guide towards actionable healthcare performance indicators: Selecting healthcare performance indicators that are fit for purpose and use for various stakeholders; Healthcare Performance Intelligence Series No. 1.3 2022

EU H2020-Agreement No. 765141

HealthPros

CONTENTS

Key messages

- Healthcare performance measurement plays an important role in guiding the decisions of healthcare system actors with respect to quality of care.
- The validity and reliability of an indicator does not guarantee that it is useful for decision-making.
- To be *actionable* indicators should be both fit for purpose and fit for use.
- An indicator's fitness for purpose reflects its ability to address a specific information need.
- To gauge an indicator's fitness for purpose, key questions to consider include: what is the intended use of the indicator? Who are the intended users (decision-makers) of the indicator?
- An indicator's fitness for use relates to its methodological qualities, the intended context of use and its handling in practice.
- The accompanied practical guide to assessing fitness for purpose and use should be applied by stakeholders working to develop a set of healthcare performance indicators.

A practical guide towards actionable healthcare performance indicators: Selecting healthcare performance indicators that are fit for purpose and use for various stakeholders; Healthcare Performance Intelligence Series No. 1.3 2022

EU H2020-Agreement No. 765141

1. Introduction

1.1.   Why does healthcare performance measurement matter?

Healthcare performance measurement is important to guide the decisions of healthcare system actors, be it national or regional policy-makers, healthcare managers, clinicians, patients or the public. Since the early 2000s, the importance of prioritizing performance measurement in healthcare has received widespread attention, to the point that it is now considered standard practice across healthcare systems. **Healthcare performance indicators** are a tool to help healthcare system actors understand how their services are doing and where there are opportunities for improvement (**Box 1**).

Knowing what indicator to measure can be a challenge, especially considering there is no universally agreed upon criteria for selecting indicators. Fortunately, the attention that healthcare performance measurement has received has increased our understanding of how best to select valid and reliable indicators. There are a number of resources available to guide this process. For example, the Appraisal of Indicators through Research and Evaluation (AIRE Instrument) was developed to assess an indicator's quality and gauge their suitability for use in practice [1]. The RAND/UCLA Appropriateness Method is another available tool that details a consensus based method to selecting indicators using the best-available scientific evidence and judgement of experts [2]. Other authors have offered steps on planning, developing and testing indicators (e.g. [3]) and ways to classify indicators for quality improvement (e.g. [4]).

While following commonly used criteria and methods can help to ensure a strong indicator, this does not guarantee that the information it provides is actually useful. For that the information needs to meet the specific needs of the intended user. Depending who and what the information is for, different indicators may be required and in effect, different data sources, levels of precision, timeliness and comparisons, may also be needed.

For example, a policy to reduce the rate of antimicrobial resistance may invest in the measurement of primary care antibiotic prescribing. Which indicators to select for measuring this will vary depending on how the information will be used and by whom. A primary care clinician, trying to understand their prescribing rate, might need an indicator that assesses new and re-prescribing of antibiotics in their practice on a regular basis. An insurer, issuing incentives to affiliated practices, is more likely to measure the adherence of physicians to prescribing guidelines yearly. And a policy-maker, trying to compare prescribing across the country and monitor long-term trends, is most likely to measure the total volume of antibiotics prescribed per 100,000 people by region, nationally or in comparison with

other countries. This example demonstrates that a 'good' indicator is not only one that is scientifically sound but also requires that the information it provides is **actionable** (i.e. an 'actionable indicator')**.**

1.2. What are actionable healthcare performance indicators?

To be actionable, it is generally agreed an indicator should be two things. The first is **fit for purpose**, meaning it serves an intended decision-making function, that is, a task or specific use. It should also be **fit for use**, meaning it is possible to get the right information, into the right hands, at the right time. While there is agreement on the importance of an indicator's actionability, it still remains a challenge to define, assess and operationalize the assessment of actionable indicators. And, without a clear understanding of what it means for an indicator to be actionable, the tendency is to select indicators on their *potential* to be useful. Importantly, when indicators fail to add useful information, that information may produce more noise than signals. In effect, indicators without a cause risk to create confusion and may even lead to bad decisions. What is worse, when misused, performance measurement can contribute to unintended consequences such as gaming and manipulation.

It is important to keep in mind an indicator's fitness for purpose and fitness for use should be taken together to appraise actionability. For example, in the scope of measuring pressure ulcer rates for international comparisons, an analysis across four measurement systems was conducted (**Box 2**). While rates of pressure ulcers are a commonly used performance indicator in long-term care facilities, the use of this indicator for international comparisons is less developed. For the purpose of international comparison, the fitness for use of the indicator requires further development, specifically around the better alignment of case definitions.

Box 1. Glossary

Indicators refer to a quantifiable variable measured to provide simplified information about a larger area of interest, typically measured over time.

Healthcare performance indicators refer to indicators for quality-driven decision-making to improve performance on one or more of the six dimensions of quality: safe, effective, patient-centred, timely, efficient and equitable care.

Healthcare performance measurement seeks to monitor, evaluate and communicate the extent to which various aspects of the health system meet key objectives [5].

Healthcare performance intelligence is a structured approach to act on healthcare priority improvement areas by using knowledge and information generated through scientific methods using comparable healthcare data to systematically measure indicators of healthcare performance [6].

Fitness for purpose refers to the extent to which an indicator serves an intended decision-making function, that is, a task or specific use [7].

Fitness for use refers to the potential for an indicator to get the right information into the right hands at the right time [7].

Box 2. Aligning methods for measuring pressure ulcer rates internationally would likely impact resources required for its measurement and use.

Pressure ulcer rates are some of the most common performance indicators used in long-term care facilities. Considering the broad use of pressure ulcer measures, it is important to understand to what extent different measurement systems are comparable and therefore, whether different information sources could be used for international comparisons of pressure ulcer rates. To explore existing measurement systems, four measurement systems of pressure ulcers, all based on point prevalence approaches, were assessed by Poldrugovac et al. [8].

The study found the existence of a regularly updated international guideline on the prevention and treatment of pressure ulcers [9] contributed significantly to aligning definitions of pressure ulcers internationally. Nonetheless, some differences were found in definitions used by the four measurement systems considered. Another important aspect of the measurement systems is the use of either a head-to-toe skin assessment supported by ad hoc training or of a validated pre-existing documentation to collect data on pressure ulcers of long-term care facility residents. Such approaches are known to increase the reliability of the measurement. Furthermore, if measures are to be used for international comparisons, it is also essential to consider the representativeness of the sample of long-term care facilities and residents involved.

To improve the comparability of pressure ulcer data internationally, some improvements could be achieved by better aligning case definitions. However other changes, such as the introduction of a head-to-toe skin assessment based on ad hoc training if not already employed, can be very resource intensive. The requirement of considerable resources may reduce the commitment of some countries or some long-term care facilities within a country to engage in such a measurement system. This may in turn reduce the representativeness of the results. Hence decision-makers at micro, meso and macro-level, including in international institutions, have to find a balance between resource intensity and reliability of the measurement systems, if international comparisons of this kind of indicator is pursued.

Source: [8]. Poldrugovac, M et al. International comparison of pressure ulcer measures in long-term care facilities: Assessing the methodological robustness of 4 approaches to point prevalence measurement. Journal of Tissue Viability, 2021, 30(4): 517–526.

1.3. How to use this guide

This guide is designed to support healthcare system actors, be it clinicians, facility managers, professional networks, policy-makers, among others, to select and use healthcare performance indicators that work. That is, the guide aims to provide a barometer for gauging the potential actionability of healthcare performance indicators. Importantly, the considerations listed can be applied to different settings of care (e.g., primary care, specialist care, long-term care, etc.) and healthcare system types, though may be limited to developed country contexts.

The guide draws on findings of a literature review and interviews with experts and real-world users of performance indicators [7]. To illustrate the meaning of fitness for purpose and fitness for use in practice, related studies by HealthPros Fellows are described throughout.

Section two explores the meaning of fit for purpose indicators by the micro (clinical), meso (organizational) and macro (policy) context of healthcare systems. Typical uses and users of healthcare performance indicators and their unique information needs are described. In section three, the meaning of fit for use indicators is elaborated by three main clusters: methodological (or technical) considerations, contextual considerations related to where (in which system) the indicator will be used, and managerial considerations relating to the indicator's use in practice. In section four, a self-guided tool to gauge the actionability of indicators is available for users working on a specific project or framework to determine the potential fitness for purpose and use of indicators for measurement.

2. Fitness for purpose

Decision-making in healthcare systems can be differentiated by three main contexts that reflect three main types of uses of healthcare performance indicators. One is improving *processes of care* at the **micro-level** or clinical setting. Another is improving the performance of *organisations and networks* at the **meso-level**. And lastly, is the use of indicators for improving *policy processes* at the **macro-level** (Figure 1). Gauging an indicator's fitness for purpose requires in a first instance, to be sure of the intended context in which an indicator will be used.

Table 1 (page 12-13) lists common uses of healthcare performance indicators and in effect, the different information they can provide. The list is not exhaustive, but rather demonstrates the importance of being precise in clarifying what information is in fact needed from a specific indicator.

Figure 1. Overview of different contexts of the healthcare system
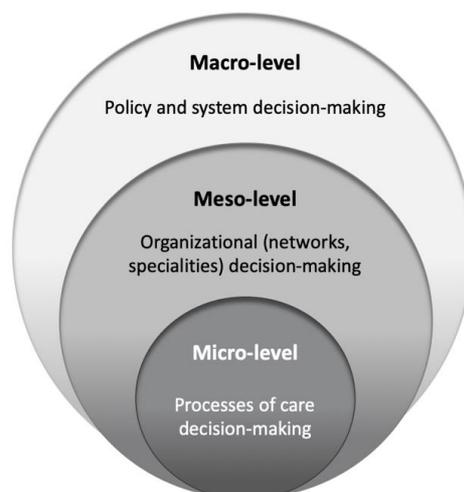
Source: [7] Barbazza E, Klazinga NS, Kringos DS. Exploring the actionability of healthcare performance indicators for quality of care: a qualitative analysis of the literature, expert opinion and user experience. BMJ Quality & Safety. 2021;30(12):1010.

2.1.   Micro-level: processes of care decision-making

Indicators targeted to improve processes of care share a common focus on small units of analysis (like individual physicians, teams, practices and departments) and frequent reporting cycles (from quarterly, to weekly, and in some instances, in real-time). Specific uses may include: to inform patients on their choice of healthcare professionals, treatments or care plans; to improve the performance of healthcare professionals by providing insights into their individual panel of patients (practice) and/or services provided like in real-time dashboards; and to improve the performance of teams, when the performance healthcare professionals is benchmarked against their peers like practice report cards.

To illustrate the uses of healthcare performance indicators in practice at the micro-level, **Box 3** describes the development of an indicator to measure practice variation in diabetes care across primary care practices in England. In **Box 4**, the use of indicators in the context of Tuscany, Italy is described, specifically related to capture differences in the prescription of antibiotics prescribing between individual physicians and across practices.

Box 3. Exploring primary care practice variation for the management of type 2 diabetes during COVID-19 in England

England's National Institute for Health and Care Excellence (NICE) uses healthcare data to assess which type of care will give the best possible outcome. In the scope of the *Quality and Outcomes Framework*, a "bundled indicator of eight checks" for people with diabetes to best manage their condition and reduce the risk of complications has been proposed. The aim is to ensure the diabetes population is consistently being offered the complete set of checks across England. The eight checks include: HbA1c, blood pressure, cholesterol, serum creatinine, urine albumin, foot surveillance, body mass index and smoking.

To quantify practice variation, the Orchid–Royal College of General Practitioners surveillance system dataset and diabetes audit dataset will be relied on. Practice variations will be compared by variables such as deprivation quintile of the practice, full-time equivalent staff, and NHS region between the clinics that provided the whole eight checks and those that did not. For this, the NICE process of care will be divided into three bands 0-3, 4-7 and 8 and analysed using mixed effects ordinal model.

Besides measuring the practice variations, the study will also focus on the difference in monitoring during the COVID-19 pandemic compared to the previous year. Two cohorts—2019–2020 (pre-COVID-19) and 2020–2021 (COVID-19) will be separately analysed. The study also aims to support the diabetes population to be engaged with routine check-ups, requiring proper public health communication.

Source: Matthew M, et al. (in-progress). Management and monitoring of people with type 2 diabetes during COVID-19 pandemic: A retrospective cohort study.

HealthPros

> **Box 4.** Measuring the prescribing of antibiotics by general practitioners and group practices in Tuscany, Italy
>
> Given the global rise of antimicrobial resistance, unwarranted variation is of particular concern when it comes to the prescription of antibiotics. To explore this, variation in antibiotic use in Tuscany's primary care was studied by using seven performance indicators reported at the general practitioner (GP) and group practice-level. In Tuscany, all GPs are affiliated to one of the region's 116 group practices, distributed across 26 local health districts, which are further regrouped in one of three local health authorities (LHAs).
>
> The indicators of interest ranged from the general consumption of antibiotics, to the prescription of certain types of antibiotics, including fluoroquinolones, amoxicillin-based antibiotics, macrolides and 3rd generation antibiotics. Trends in antibiotic prescribing are recorded at the district, group practice as well at the GP level. Depending on the user, performance data will have different purposes. Reporting of individual as well as group performance to GPs allows them to identify potential gaps and improve prescribing behavior individually and as a group. Furthermore, performance in antibiotic prescribing may contribute to the supplementary income provided to GPs through a pay for performance scheme. Performance data of Tuscany's districts and group practices is also publicly available via an online platform and reports that are updated regularly. With the use of effective data visualization, different stakeholders, ranging from patients to policy-makers, are able to identify strengths and weaknesses of their local health system, and subsequently take appropriate action. For instance, heads of LHAs, who receive financial rewards for good performance, may use performance data to set directives that address inappropriate prescribing of antibiotics.
>
> The results from this study suggested that the majority of the variation was due to differences between GPs themselves (75% to 98%) as opposed to the influences exerted by their peers or institutional mechanisms. This means that despite the availability of performance data to GPs, the variation in antibiotic prescribing not only persists but also reflects a wide variety in individual practice styles among GPs. Based on these findings, it is recommended that representatives of primary care practices place greater emphasis on group performance related to antibiotic prescribing and lead their peers to become more aware of their own performance and harmonize clinical behavior in line with best practices.
>
> *Source: [10]. Willmington C, Vainieri M, Seghieri, C. Estimating variations in the use of antibiotics in primary care: Insights from the Tuscany region, Italy. Int J Health Plann Manage. 2021 doi: 10.1002/hpm.3388.*

2.2. Meso-level: Organizational (networks, specialists) decision-making

The use of indicators at the meso-level goes beyond an individual physician or team and assesses trends to alert organizations and networks of care when measures related to quality fall outside a normal range. The use of healthcare performance indicators at this level may focus on improving performance across networks and areas of specialization, measuring the adherence to guidance in order to issue incentives, or on professional development and regulatory uses for the issuing of accreditations, certificates or licences. In Germany, indicators for measuring and monitoring physician cooperation have been explored using insurance claims data (**Box 5**). Despite innovative methods to measure performance and the growing use of measurement data at the meso-level, a study of hospital

managers in the European context found there is need to further expand use of performance measurement for strategic decision-making and learning opportunities (**Box 6**).

Box 5. Cooperation improvement in an integrated healthcare network: A social network analysis

Cooperation is a core feature of integrated healthcare systems. The premise is that providers who cooperate can achieve more efficient use of healthcare services while improving health outcomes. However, indicators for measuring and monitoring physician cooperation are not considered in performance assessment frameworks of integrated healthcare systems. Moreover, traditional cooperation assessment methods use surveys—an expensive and ineffective approach for systematic monitoring.

Several authors have defined and validated the use of shared patients to identify information-sharing relationships among physicians, hence, patient sharing networks can be used to construct cooperation networks. Using claims data and social network analysis, we constructed the physician cooperation network of an integrated healthcare initiative in southern Germany and measured its evolution over its 14 years of existence. Cooperation was studied by analyzing network properties at two levels. At network-level, we focused on network density and network mean distance. At physician-level, we focused on three measures of centrality; degree, betweenness, and eigenvector. Furthermore, using a dynamic panel analysis with fixed effects we were able to understand if the evolution of cooperation was more favorable for physicians participating in the integrated initiative.

Our findings show an increasingly cooperative physician network, led by physicians participating actively in the integrated initiative. Thereby we also provide evidence for cooperation being one of the mechanisms that have been driving the success of integrated healthcare. Moreover, we provide a tool for monitoring cooperation among a network's members for healthcare providers and/or insurers by using claims data. As exposed in this paper, integrated health systems can use said indicators to assess the system's performance in improving professional cooperation, a key concept in the integrated approach value-creating mechanism.

Source: Larrain, Wang, Stargardt, Groene. Physician cooperation improvement in an integrated health care network: A social network analysis. Under review.

Box 6. Despite the substantial and increasing use of performance data for evidence-based management in healthcare organisations around Europe, there is room and need for improvement

Managing hospitals, and other healthcare organisations, requires a delicate combination of strategic and operational management of clinical and all other processes that provide support for clinical work. Performance intelligence, in the form of indicators, provide the evidence necessary to carry out all three basic managerial functions: planning, decision-making and controlling. Management of clinical and support processes is often the domain of middle management, linking the worlds of evidence-based clinical medicine to evidence-based management of healthcare delivery. Mid-level managers supervise frontline clinical workers and are themselves being supervised by an organization's senior managers. Research on middle managers' commitment to the implementation of innovations shows that it is, in large part, influenced by personal perception of the potential benefit of the innovation for patients and the ease with which an innovation can be implemented.

In 2019, we set out to explore the actual use of performance data in hospitals and other healthcare organisations in Europe, and opportunities to enhance its use. We aimed at understanding why performance data are collected, reported and used, what data are collected, reported and used for performance management, and how are performance data used for decision-making in healthcare organisations. We did so through a descriptive cross-sectional study based on a survey, delivered through an online self-reported questionnaire, and a follow-up interactive workshop.

We surveyed 125 healthcare managers, mostly working in publicly-owned hospitals in 20 different European countries, with an even distribution of managerial responsibilities between strategic, clinical and support-process management. We found that although a substantial amount of performance data is being regularly collected, its potential is still somewhat underused for decision-making purposes. A very similar issue is recognised in benchmarking: while being recognised as valuable, benchmarking between and within the organisations is still underused. Additionally, in collecting and reporting the performance data, motivation is found both internally and externally, and is aimed at both improvement as well as accountability purposes. Furthermore, even though a wide range of data sources is used, more should be done on conceptualising, collecting, reporting and using patient-reported data. When it comes to organisations' ownership, managers working for privately-owned organisations reported a greater use of performance data compared to the ones working in the public organisations. Moreover, the strategic levels of management are reported to mostly use performance data to justify their decisions, while the managers working on the operational and clinical level predominantly use it for day-to-day operational decision-making.

Our study showed that, despite the substantial and increasing use of performance data for evidence-based management, there is room and need to further explore and expand its role in strategic decision-making and to support a shift in healthcare from organisational accountability towards the model of learning organisations.

2.3. Macro-level: Policy and system decision-making

At the macro-level, healthcare performance indicators are focused on outcomes, to understand the burden of illness and quality of life of the population, and to manage and evaluate (the contribution of) health system interventions. Performance indicators at this level are also an input for accountability.

While uses of health care performance indicators in this context aim overall to inform policy decisions, distinctions between uses include: system performance monitoring—signalling to system stakeholders, often including the public, the performance of the system as a whole, answering "How is my health care system doing?"; strategy development—signaling to ministries, departments of health or similar with the aim of identifying priority areas, monitor trends and ultimately answering "Have I chosen the right areas to prioritize?"; or system quality assurance—informing decisions of

A practical guide towards actionable healthcare performance indicators: Selecting healthcare performance indicators that are fit for purpose and use for various stakeholders; Healthcare Performance Intelligence Series No. 1.3 2022

EU H2020-Agreement No. 765141

HealthPros

health service executives, quality inspectors or quality observatories for an overview of care processes and signalling of incidents, answering "Is care being delivered as intended?"

Differentiating between system uses of healthcare performance indicators was a key component in the development of Ireland's health system performance assessment framework and its accompanied suite of indicators (**Box 7**). Three different purposes of use and target users were defined in the scope of this work.

Box 7. Three different uses of Ireland's first Health System Performance Assessment (HSPA) framework

In Ireland, improving the governance, accountability, and performance of the health system were set out as key priorities in the 10-year reform programme *Sláintecare* 2019–2028. At the outset of the programme, key stakeholders recognized that a comprehensive performance measurement framework and management system was needed to foster accountability and capture achievements against the objectives of *Sláintecare*. To enable the evaluation of priority areas of the reform and to ensure that the health system is more responsive to the needs of the population, the development of a Health System Performance Assessment (HSPA) framework was launched.

Three functions for the HSPA framework were defined. Each function reflects a different intended use and user (decision-maker). The uses, aim and target users included the following:

1. **Measuring system performance.** To signal the performance of the system as a whole to foster accountability to the public.
2. **Monitoring system reforms.** To signal the performance related to priority areas such as the integration of services, public-private partnerships and regionalization measures to inform the decision-making by the Department of Health.
3. **Improving the delivering of services.** To signal the performance of services delivered/y for short-term planning and priority setting across delivery platforms by the Health Service Executive.

The three uses of the HSPA framework informed the selection of indicators, together with an indicator's measurability and methodological robustness.

Source: [12] (Ivankovic D, Jansen T, Barbazza E, Brito Fernandes Ó, Klazinga N, Kringos D. Health information system in Ireland and its fitness to support health system performance assessment: A multimethod qualitative assessment. (under review) and [13] (Brito Fernandes Ó, Barbazza E, Ivankovic D, Jansen T, Klazinga N, Kringos D. Engaging citizens in development of a health system performance assessment framework: a case study in Ireland. Health Res Policy Sys 19, 148 (2021). https://doi.org/10.1186/s12961-021-00798-8) for detailed assessments of the framework's development.

Table 1. Differentiating uses of health care performance indicators across healthcare systems

Purpose of use	Illustrative uses	Illustrative users	Illustrative information need
Macro: policy and system decision-making			
System performance monitoring	Signaling the performance of the system as a whole; comparing performance internationally; publicly reporting system performance	Public; ministry of health; regional (provincial, state) authorities; health service executive (authority)	How is my health care system doing? How does it compare with others?
Strategy development	Setting health policy priorities; identifying emerging health priority areas; and monitoring trends in current priority areas	Government and ministries; regional (provincial, state) authorities; accountable care organizations; health maintenance organizations	Have I chosen the right areas to prioritize? What is the impact of strategies that are in place?
System quality assurance	Measuring care processes; reporting of incidents and never events	Quality inspectorate; national quality observatory; health and safety executive	Is care being delivered as intended? Where do problems in the delivery of care lie?
Meso-level: Organizational (networks, specialists) decision-making			
Regulation (professional, facility, pharmaceuticals)	Informing accreditation, certification and/or licensing processes	Medical councils, chambers, college of physicians; medicines and health care products regulatory agency	Does the performance of organizations, facilities, medicines, etc., meet established standards?
Professional development	Reporting internally and benchmarking within profession or specialty	Societies of medical professionals; professional associations; training institutions	How do health care professionals of a specific specialty perform?
Quality-based financing	Issuing performance-based payment (pay-for-performance);	Health care insurers; health care providers	Are existing guidelines or standards being adhered to?

10

	value-based contracting		Does this merit the issuing of incentives?
Organization/ network performance improvement	Improving performance of hospitals, networks, care groups; assessing local needs and geographic differences	Hospital management; integrated care; networks/groups; local collaboratives of care	Are affiliated practices/facilities performing optimally?

Micro-level: processes of care decision-making			
Practice or team performance improvement	Convening audit and feedback, plan-do-study-act, and/or collaborative, team-based improvement cycles; comparing across practices	Primary care practices; specialist departments or units; pathways of care	How is my team performing? How can we improve our performance? How do I perform relative to my team members?
Individual performance improvement	Identifying trends in the management of patients; tailoring services to target groups	Individual physicians; nurse/practitioners; other health care professionals	How am I managing my practice panel? How can I improve my performance?
Informed choice	Selecting a health care provider; participating in care decision-making; self-managing care needs	Patients; family members and carers; public	What treatment options or providers are best for me?

Source: [7] Barbazza E, Klazinga NS, Kringos DS. Exploring the actionability of healthcare performance indicators for quality of care: a qualitative analysis of the literature, expert opinion and user experience. BMJ Quality & Safety. 2021;30(12):1010.

3. Fitness for use

The second component of actionability—fitness for use—can be assessed by three main types of considerations: methodological, contextual and managerial. These relate to an indicator's technical qualities, its intended context of use and its handling across what can be characterised as a use cycle. It means that to gauge an indicator's fitness for use, a range of aspects should be assessed that span, for example, 'Does the indicator signal a clear direction?' to 'Can needed data be accessed?' and 'What is the relevant reporting cycle?'. The different types of considerations are described to follow.

3.1. Methodological considerations

Beyond reliability, validity and other generally agreed upon criteria for the selection of indicators, a range of methodological considerations can be said to influence what is a 'good' indicator from a *use* perspective. First, an indicator should measure what matters. That is, the indicator should be of importance to the target audience. Second, the extent to which an indicator resonates with a range of stakeholders is a key gauge of its ability to facilitate a 'what can **we** do' approach, rather than limiting action to an individual user. Third, the easier an indicator can be interpreted, the higher the end user's confidence is in understanding and using the information it provides. Fourth, the extent to which an indicator is clearly defined is a key contributor to trust in what it signals. Fifth, an indicator should be able to be broken down into its related parts to make change points clear. When changes are too remote or disconnected it can be difficult for the measure to be acted upon. This consideration was explored with regard to the use of composite indicators applied to quality-of-care measures in Denmark (**Box 8**). Sixth, an indicator should measure a phenomenon as true to lived experience as possible. The tendency to focus on specific areas of care can reduce performance to overly narrow aspects of care, missing the 'system-ness' of quality. Lastly, the ability of an indicator to be sufficiently sensitive to change based on its intended use is intuitive, yet often a challenge for an indicator to meet.

Box 8. Constructing actionable composite indicators in Denmark

We constructed composite indicators—multiple individual indicators compiled into a single index—for six Danish national clinical registry databases (four cardiovascular and two mental care) (1) to investigate the overall quality of care provided to patients and (2) to facilitate comparisons between regions and healthcare providers in Denmark. We investigated two of the most used approaches: opportunity scoring and all-or-none scoring. These approaches emphasize different aspects regarding quality of care. While opportunity scoring rewards partial performance, all-or-none scoring only rewards complete care and promotes excellence. We obtained composite quality scores based on process indicators for multiple levels: national level, regional level and healthcare provider level.

In our study, it was concluded that, firstly, composite indicators can be useful and actionable tools to quantify quality of care especially when we have many healthcare providers and individual indicators in the study. For example, for the schizophrenia database 12 process indicators were included in the report and there were over 40 providers, resulting in more than 480 numbers to consider in order to make comparisons between healthcare providers. Composite indicators can be very valuable in such circumstances, providing an overall picture of quality and summarizing the quality of care with a single number for each region or provider.

Second, a potential limitation of composite indicators is that there is not a standard approach to construct them and using different methods can give in different results. Therefore, to obtain reliable, robust and actionable composite indicators it is very important to construct composite indicators step by step in a methodologically sound way and to be transparent regarding the construction process to avoid possible misuse and misinterpretation of the results.

3.2. Contextual considerations

Contextual considerations refer to critical factors pertaining to the setting in which an indicator is used Firstly, the information infrastructure can affect the ability to collect, store and extract information. Relevant considerations included the interoperability of information systems (i.e., linkages, output format) and overall data quality (i.e., consistency in field, codes, maintenance). **Box 9** describes the potential, but also challenge, to use routine databases as a source of data for the purposes of performance measurement.

Secondly, characteristics of governance, related to political will and vision, regulatory arrangements for data exchanges, as well as cross-sector partnerships and financing structures, can also influence the use of indicators in practice. Thirdly, workforce capacity, specifically the data literacy skills of people across the healthcare system and the availability of dedicated time for the healthcare workforce to use data can influence the uptake of information. Lastly, professional norms and culture of using performance indicators, be it in clinical practice, health care organizations, professional networks or government agencies, are a key predictor of the importance placed on measurement and ultimately, the use of an indicator.

Box 9. Using automated algorithms to extract cohorts from routine databases

Routine databases and disease registries derived from electronic medical records can be a useful source of patient data. It can be used for predictive statistical modelling, useful to monitor and evaluate the efficacy of health policy interventions on specific populations and specific (disease) outcomes, such as Type 2 diabetes.

Previous work on the EUBIROD network has demonstrated that national registries can be homogenized through a common data dictionary or ontology, such as the Data Collection Reference Guide for Type 1 and Type 2 Diabetes from the International Consortium of Health Outcome Measurements (ICHOM).

Once database columns have been mapped to their relevant name and structure, automated data extraction algorithms can work on predefined columns based on a common benchmarking target through a traditional PICO approach (population, intervention, control, and outcomes) (e.g. effect of organizational arrangements on preventing lower extremity amputations in people with type 2 diabetes). Similarly, statistical models may work on predefined inputs.

The main challenges of working with national registries include i) the acquisition of the relevant permissions and credentials, and ii) physically accessing the databases. Sensitive healthcare data originating in one country would not generally be allowed to be extracted remotely, hence the need for an in-person access or on-site partnership for every registry. However, this challenge is partly tackled by the use of automated algorithms, which require only one on-site visit for the installation of the process. The algorithm subsequently generates aggregated extracts which can be shared across countries, under the current regulations.

We are investigating the mechanisms through which feedback loops can turn data intelligence into actionable policy. For example, through actionable dashboards, integrated care policy feedback, and monitoring unwarranted variation, all under the principle of timeliness and targeted to the intended audience's needs.

Overall, the main recommendations include the following; The use of routine databases and nationally based registries for the monitoring of the quality of care provided to people with type 2 diabetes is recommended. This is a subgroup with high risk for cardiovascular complications as well as COVID-19 related complications, which deserves special attention. In addition, type 2 diabetes is used as an example to transfer this approach to the monitoring of other chronic conditions. Routine registries provide data with adequate granularity, follow-up periods, and sample sizes to conduct robust observational studies on low incidence complications (e.g. lower extremity amputation), controlling for multiple confounders. This provides a novel and low-cost approach for the monitoring and research of chronic conditions and their complications.

How to transfer these findings to other registries, identifying optimal data granularity to extend modelling across Europe in a privacy protected mode, in collaboration with the EUBIROD network will be further explored.

Source: [15] Meza-Torres B, Cunningham S, Heiss C, Joy M, Leese GP, De Lusignan S, et al. High Quality and Timely Foot Care Predict Amputation-Free Survival Among People With Type 2 Diabetes: Results From a Large Population-Based Longitudinal Cohort in Scotland (under review) and [16] Meza-Torres B., Cunningham S., Heiss C., Joy M., Feher M., Leese G., Carinci F., de Lusignan S. Predictors of major lower extremity amputations in people with type 2 diabetes and foot ulcers. International Diabetes Federation 2021. Available at: https://doi.org/10.26226/morressier.617c37317c09fc044a975261.

3.3. Managerial considerations

Managerial considerations relate to an indicator's use across what can be characterized as an indicator's use cycle, as visualized in Figure 2. Relevant considerations include managing the selection of an indicator including gaining clarity around its intended use, construction, data needs and measurement; accessing data to ensure data is available, of quality or can feasibly be collected; applying methods of analysis for the calculation of values that correspond to the intended purpose (see for example **Box 10** testing the use of entropy balancing to evaluate system integration); displaying findings, including decisions around how data is visualized and the degree of story-telling

to describe and interpret results to support understanding of what is meant and any caveats; and actually reaching decision-makers, with decisions needed as to the frequency of dissemination, channel used for delivering information and guidance (if any) to facilitate the use of information provided. **Box 11** describes results exploring the use of dashboards as a reporting modality for disseminating COVID-19 data to the public. While dashboards have been used widely during the pandemic, their actionability is not guaranteed.

Figure 2. Use cycle for managing health care performance indicators
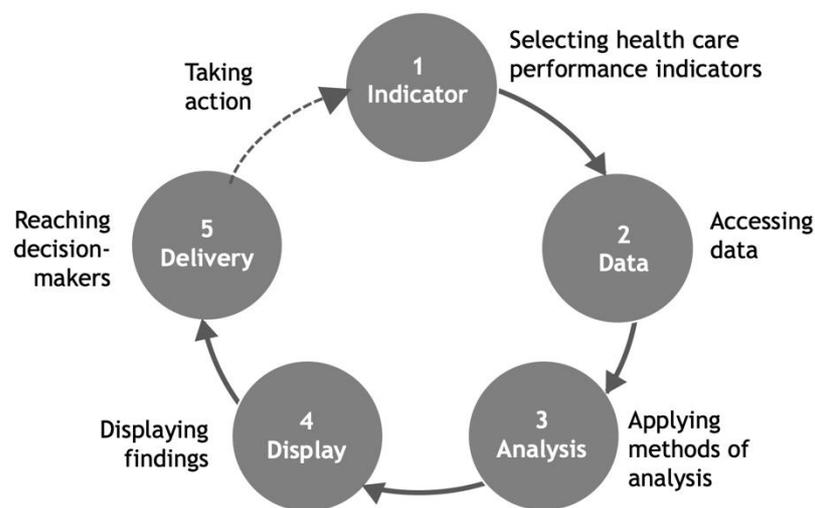
Source: [7] *Barbazza E, Klazinga NS, Kringos DS. Exploring the actionability of healthcare performance indicators for quality of care: a qualitative analysis of the literature, expert opinion and user experience. BMJ Quality & Safety. 2021;30(12):1010.*

Box 10. Measuring the effectiveness of integrated healthcare systems to improve population health outcomes

Evidence linking the effectiveness of integrated healthcare systems improvements in population health is scarce. Previous literature uses quasi-experimental designs based on a combination of exact and propensity score matching with this objective. However, when evaluating an integrated healthcare initiative with the proposed design, an important percentage of the available sample is lost due to a lack of equivalent data. Moreover, the lost sample was significantly associated with high healthcare needs. Because of integrated healthcare initiative's whole system approach, interventions are introduced over the whole spectrum of care services. In this context, the non-random exclusion of a portion of the sample can heavily bias the results of the evaluation in an unknown direction. We provide an updated design to evaluate the effectiveness of integrated care on population health outcomes that overcomes these challenges by using entropy balancing (a multivariate reweighting method to produce balanced samples in observational studies).

Using claims data from 2004 to 2018, we compared participants of an integrated initiative to a control group created with entropy balancing and follow them for 5 years. Population health indicators of survival, mortality ratio, mean age at the time of death, and years of life lost or gained were measured. As comparison, a secondary evaluation was made following the propensity score matching design outlined in previous literature.

Besides measuring the positive effect of integrated care over population health, our findings show that previous approaches for evaluation overestimate said effect by excluding patients with high healthcare needs. Consequently, our results suggest that health gains resulting from the integrated care approach diminish for patients with high healthcare needs. Our design was able to deal with the shortcomings of propensity score matching-based designs by not eliminating any available sample in the treatment group, while achieving better balanced samples at base line.

Source: [17] Larrain N, Groene O. 'Quasi experimental evaluation of an integrated healthcare system using Entropy Balancing'. Working paper. Jun 2021.

Box 11. Dashboards are a tool to visually deliver data to users though require thoughtful consideration of key features to safeguard their actionability

Dashboards are a powerful vehicle for communication, providing a dynamic means to visually display information at-a-glance. In the health sector, dashboards have been relied on for delivering results of health system performance assessments and internal management. Public, web-based dashboards have also been widely adopted for reporting on the COVID-19 pandemic. However, the mere accessibility of COVID-19 dashboards does not guarantee data-informed decision-making.

To explore the state of the art of publicly available web-based COVID-19 dashboards, Ivankovic et al. [14] described and assessed 158 COVID-19 dashboards from 53 countries worldwide. The study reports a snapshot of this landscape in the early stages of the pandemic (July 2020), describing their purpose and users ("why"), content and data ("what") and analyses and displays ("how" they communicate COVID-19 data). In total, 20/158 dashboards (12.7%) were appraised as highly actionable and seven common features were identified between them. Actionable COVID-19 dashboards (1) know their audience and information needs; (2) manage the type, volume, and flow of displayed information; (3) report data sources and methods clearly; (4) link time trends to policy decisions; (5) provide data that are "close to home"; (6) break down the population into relevant subgroups; and (7) use storytelling and visual cues.

While there is no one-size-fits-all template or model to deliver performance data using dashboards, the identified features should be adopted to enhance their actionability.

Source: [18] Ivanković D, Barbazza E, et al. Features Constituting Actionable COVID-19 Dashboards: Descriptive Assessment and Expert Appraisal of 158 Public Web-Based COVID-19 Dashboards. J Med Internet Res 2021;23(2):e25682 doi: 10.2196/25682PMID: 33577467PMCID: 7906125.

4. Practical guide to assessing fitness for purpose and use

Purpose of use

> For a specific measurement area or aim, consider: what is the intended use of the indicator? Which context of decision-making will it inform? Who are the intended users (decision-makers) of the analyzed data? Specify these details in the table below. One or more cell may apply.

Illustrative purpose of use	Illustrative information need	Specific purpose	Specific user
Macro: policy and system decision-making			
System performance monitoring	How is my health care system doing? How does it compare with others?		
Strategy development	Have I chosen the right areas to prioritize? What is the impact of strategies that are in place?		
System quality assurance	Is care being delivered as intended? Where do problems in the delivery of care lie?		
Meso-level: Organizational (networks, specialists) decision-making			
Regulation (professional, facility, pharmaceuticals)	Does the performance of organizations, facilities, medicines, etc., meet established standards?		
Professional development	How do health care professionals of a specific specialty perform?		
Quality-based financing	Are existing guidelines or standards being adhered to? Does this merit the issuing of incentives?		

HealthPros

Organization/ network performance improvement	*Are affiliated practices/facilities performing optimally?*		
Micro-level: processes of care decision-making			
Practice or team performance improvement	*How is my team performing? How can we improve our performance? How do I perform relative to my team members?*		
Individual performance improvement	*How am I managing my practice panel? How can I improve my performance?*		
Informed choice	*What treatment options or providers are best for me?*		

Fitness for use

For each consideration, reflect on the guiding question. Consider each of the intended uses of indicators based on the previous table as different responses may apply to varied intended uses.

Clusters	Considerations	Guiding questions for considering an indicator's use
Methodological		
	Measures what matters	• Does anybody care?
	Wide engagement	• What can *we* do?
	Easily interpreted	• Does the indicator signal a clear direction?
	Clear standardization	• Is the indicator clearly defined and replicable?
	Alignment of accountability	• Are entry-points for taking action feasible?
	Measurement matches delivery	• Is the indicator a reflection of the system?
	Sensitive to meaningful change	• Is the indicator sufficiently sensitive to change?
Contextual		
Information infrastructure	Interoperability	• Can needed data be accessed?
	Data quality	• Is the data of quality?
Governance	Political will and vision	• Is there high-level commitment and direction for use?
	Regulation for data protection	• Does existing legislation facilitate use?
	Cross-sector partnerships	• Are cross-sector partnerships in place?
	Aligned financing structures	• Do financing structures encourage the intended use?
Workforce capacity	Data and quality expertise	• Are the competencies to interpret and use data in place?
	Time dedicated to improvement	• Is time allocated to encourage use?
Culture	Learning-orientation	• Is an environment for learning cultivated?

	Shared responsibility for health	• Do users feel accountable for improvement?
Managerial		
Selecting health care performance indicators	Clear purpose of use	• What is the purpose of use? (e.g., strategy development)
	Target end-user is known	• Is the target audience known? (e.g., clinicians, public)
	Conceptual framework	• Is the dimension of quality pursued clear?
	Indicator quality	• Is the indicator scientifically sound?
	Source, type and availability of data	• What data is needed and is it available? (e.g., administrative, clinical, survey data, wearables)
	Standards for appraisal	• How will improvements in performance be assessed?
	Degree of public disclosure	• Is the indicator for internal or external (public) use?
	Accompanying indicators	• Are there relevant accompanied indicators?
	Previous use	• Has the indicator been used previously?
Accessing data	Representativeness of data	• Is the data complete?
	Data linkages	• Can relevant data sources be linked?
	Data collection tools	• How will data be collected? (e.g., paper-based, automated electronically, manual electronic entry)
	Unity of language/coding	• Is there consistency in coding across data to be used?
Applying methods of analysis	Type of analysis	• How will the data be analyzed? (e.g., benchmarking, time trend, case mix correction)
	Aggregation of indicators	• How can composites/indices be used to simplify data?
	Reference group	• Who is the reference group?
	Breakdowns/cohorts	• How will the data be disaggregated? (e.g., age, sex, ethnicity, geographically)

HealthPros

	Calculation of values	• How will values be calculated? (e.g., means, median, standard deviation, top 10% mean)
	Time interval	• Should a time trend be reported and at what interval?
	Application of risk-adjustments	• How will risk adjustments be applied? (e.g., variable specification, source, weighting scheme)
	Managing missing data	• How will missed data points be handled?
	Contextualizing data	• What other data is needed to give the indicator meaning?
Displaying findings	Chart options	• How will the data be visualized? (e.g., chart, map, table)
	Simplification techniques	• What techniques to simplify the meaning can be applied? (e.g., colour, size variation, icons)
	Customization of display	• How can users customize the data? (e.g., change of display, change of information)
	Narrated interpretation	• How can the quality and the meaning of data be narrated?
	Format of reporting	• How will it be reported? (e.g., print, mobile, web-based)
Reaching decision-makers	Frequency of reporting	• What is the relevant reporting cycle (e.g., real-time, quarterly, annually, biennially)
	Dissemination channels	• How will users be reached? (e.g., mail, email, champions)
	Guidance on use	• How can users be supported to make use of findings?

References

1.   de Koning J, Burgers J, Klazinga N. Appraisal of indicators through research and evaluation (AIRE) Amsterdam University of Amsterdam 2008.

2.   Fitch K, Bernstein SJ, Aguilar MD, Burnand B, LaCalle JR, Lazaro P, et al. The RAND/UCLA Appropriateness Method User's Manual. Santa Monica, CA RAND; 2001.

3.   Mainz J. Developing evidence-based clinical indicators: a state of the art methods primer. International Journal for Quality in Health Care. 2003;15(suppl_1):i5-i11.

4.   Mainz J. Defining and classifying clinical indicators for quality improvement. International Journal for Quality in Health Care. 2003;15(6):523-30.

5.   Smith P, Mossialos E, Papanicolas I, Leatherman S. Part 1: Principles of performance measurement. In: Smith P, Mossialos E, Papanicolas I, Leatherman S, editors. Performance measurement for health system improvement: experiences, challenges and prospects Copenahgen WHO Regional Office for Europe 2008.

6.   20/20 HS. The health system assessment approach: A how-to manual 2012.

7.   Barbazza E, Klazinga NS, Kringos DS. Exploring the actionability of healthcare performance indicators for quality of care: a qualitative analysis of the literature, expert opinion and user experience. BMJ Quality & Safety. 2021;30(12):1010.

8.   Poldrugovac M, Padget M, Schoonhoven L, Thompson ND, Klazinga NS, Kringos DS. International comparison of pressure ulcer measures in long-term care facilities: Assessing the methodological robustness of 4 approaches to point prevalence measurement. J Tissue Viability. 2021;30(4):517-26.

9.   European pressure ulcer advisory Panel, National pressure injury advisory Panel, Pan pacific pressure injury alliance. Prevention and treatment of pressure ulcers/injuries: clinical practice guideline. EPUAP/NPIAP/PPPIA; 2019.

10.   Willmington C, Vainieri M, Seghieri C. Estimating variations in the use of antibiotics in primary care: Insights from the Tuscany region, Italy. Int J Health Plann Manage. 2021.

11.   Ivankovic D, Poldrugovac M, Garel P, Klazinga NS, Kringos DS. Why, what and how do European healthcare managers use performance data? Results of a survey and workshop among members of the European Hospital and Healthcare Federation. PloS one. 2020;15(4):e0231345.

12.   Ivankovic D, Jansen T, Barbazza E, Brito Fernandes Ó, Klazinga N, Kringos D. Health information system in Ireland and its fitness to support health system performance assessment: A multimethod qualitative assessment. under review.

13.   Brito Fernandes Ó, Barbazza E, Ivankovic D, Jansen T, Klazinga N, Kringos D. Engaging citizens in development of a health system performance assessment framework: a case study in Ireland. *Health Res Policy Sys* 19, 148 (2021). https://doi.org/10.1186/s12961-021-00798-8.

14.   Kara, P., Johnsen, SP. Construction and Use of Composite Indicators for Rkkp Databases. https://www.rkkp.dk/siteassets/om-rkkp/rapporter/rkkp_report__081120kompositte_.pdf.

15.   Meza-Torres B, Cunningham S, Heiss C, Joy M, Leese GP, De Lusignan S, et al. High Quality and Timely Foot Care Predict Amputation-Free Survival Among People With Type 2 Diabetes: Results From a Large Population-Based Longitudinal Cohort in Scotland (under review).

16.   Meza-Torres B., Cunningham S., Heiss C., Joy M., Feher M., Leese G., Carinci F., de Lusignan S. Predictors of major lower extremity amputations in people with type 2 diabetes and foot ulcers. International Diabetes Federation 2021. Available at: https://doi.org/10.26226/morressier.617c37317c09fc044a975261.

17.   Larrain N, Groene O. 'Quasi experimental evaluation of an integrated healthcare system using Entropy Balancing'. Working paper. Jun 2021.

18.   Ivanković D, Barbazza E, Bos V, Brito Fernandes Ó, Jamieson Gilmore K, Jansen T, et al. Features Constituting Actionable COVID-19 Dashboards: Descriptive Assessment and Expert Appraisal of 158 Public Web-Based COVID-19 Dashboards. J Med Internet Res. 2021;23(2):e25682.