OECD Health Working Papers No. 176

The economics of diagnostic safety

Luke Slawomirski,
David Kelly,
Katherine de Bienassis,
Kadri-Ann Kallas,
Niek Klazinga

https://dx.doi.org/10.1787/fc61057a-en

OECD

OECD Health Working Papers No. 176

The Economics of Diagnostic Safety

JEL classification: H51, I11, I18, L86

Authorised for publication by Stefano Scarpetta, Director, Directorate for Employment, Labour and Social Affairs.

Luke Slawomirski, David Kelly, Katherine de Bienassis, Kadri-Ann Kallas and Niek Klazinga

OECD Health Working Papers

Reports and research papers | OECD

OECD Working Papers should not be reported as representing the official views of the OECD or of its member countries. The opinions expressed and arguments employed are those of the authors.

Working Papers describe preliminary results or research in progress by the author(s) and are published to stimulate discussion on a broad range of issues on which the OECD works. Comments on Working Papers are welcomed, and may be sent to health.contact@oecd.org.

This series is designed to make available to a wider readership selected health studies prepared for use within the OECD. Authorship is usually collective, but principal writers are named. The papers are generally available only in their original language – English or French – with a summary in the other.

This document and any map included herein are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

The statistical data for Israel are supplied by and under the responsibility of the relevant Israeli authorities. The use of such data by the OECD is without prejudice to the status of the Golan Heights, East Jerusalem and Israeli settlements in the West Bank under the terms of international law.

Abstract

Diagnosis is complex and iterative, therefore liable to error in accurately and timely identifying underlying health problems, and communicating these to patients. Up to 15% of diagnoses are estimated to be inaccurate, delayed or wrong. Diagnostic errors negatively impact patient outcomes and increase use of healthcare resources. The direct financial burden of misdiagnosis, underdiagnosis and overdiagnosis combined is estimated to be 17.5% of total healthcare expenditure, or 1.8% of GDP in a typical OECD country where one tenth of GDP is spent on health care. Reducing diagnostic error has the potential for large cost savings through improvements in healthcare efficiency and reductions in patient harm. Halving rates of diagnostic error could lead to savings of 8% of healthcare expenditure. This report 1) defines the scope of diagnostic error, 2) illustrates the burden of diagnostic error in commonly diagnosed conditions, 3) estimates the direct costs of diagnostic error 4) provides policy options to improve diagnostic safety.

Zusammenfassung

Diagnostik ist komplex und iterativ und daher fehleranfällig, wenn es darum geht, die zugrunde liegenden Gesundheitsprobleme richtig und rechtzeitig zu erkennen und den Patient*innen mitzuteilen. Schätzungen zufolge sind bis zu 15 % aller Diagnosen ungenau, verzögert oder falsch. Diagnosefehler wirken sich negativ auf die Patientenergebnisse aus und erhöhen den Einsatz von Ressourcen in der Gesundheitsversorgung. Die direkte finanzielle Belastung durch Fehldiagnosen, Unterdiagnosen und Überdiagnosen wird auf 17,5 % der gesamten Gesundheitsausgaben bzw. 1,8 % des BIP in einem typischen OECD-Land beziffert, in dem ein Zehntel des BIP für die Gesundheitsversorgung aufgewendet wird. Eine Reduzierung von Diagnosefehlern birgt großes Einsparpotenzial durch eine Verbesserung der Effizienz im Gesundheitswesen und eine Verringerung der Patientenschäden. So könnte eine Halbierung der Fehlerquote zu Einsparungen in Höhe von 8 % der Gesundheitsausgaben führen. Dieser Bericht analysiert das Ausmaß von Diagnosefehlern, veranschaulicht deren Bedeutung bei häufig diagnostizierten Erkrankungen, schätzt die daraus resultierenden direkten Kosten und unterbreitet Politikoptionen zur Verbesserung der Diagnosesicherheit.

Acknowledgements

The work was enabled by a voluntary contribution from the Ministry of Health of Germany (BMG). The authors would especially like to acknowledge Ingo Härtel for his input and support.

Austria, Australia, Belgium, Czech Republic, Colombia, Finland, France, Germany, Iceland, Japan, Latvia, Luxembourg, Norway Sweden, Slovenia, and Switzerland participated in the snapshot survey in November and December 2024.

This report benefited from the advice and comments of Hardeep Singh, José Artur Paiva, Ayda Taha, Patrick Romano, Wendy Levinson, Jean-Luc Tilly, Ines Vukic, Ingo Haertel, Le Duong and Jutta Järvelin. Wendy Levinson provided the examples in Box 4.1.

The authors are grateful to Mark Putland, Eric Sutherland, Emily Hewlett, and Francesca Colombo for their valuable comments and Sabina Kovacic for her assistance and technical support.

This work was coordinated and authored by the OECD, including Luke Slawomirski, David Kelly, Katherine de Bienassis, Kadri-Ann Kallas and Niek Klazinga.

Table of contents

FIGURES

TABLES

BOXES

Acronyms

AI	Artificial intelligence
ADHD	Attention-Deficit/Hyperactivity Disorder
ASD	Autistic Spectrum Disorders
CIRS	Critical Incident Reporting System
CT	Computed tomography
DSM-V	Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition
ED	Emergency department
EHR	Electronic Health Record
FOPH	Federal Office of Pubic Health (Switzerland)
FQC	Federal Quality Commission (Switzerland)
HSPSC	Hospital Survey of Patient Safety Culture
IVD	In vitro diagnostic medical devices
LLM	Large language models
MRI	Magnetic resonance imaging
NASEM	National Academy for Science and Engineering and Medicine
PET	Positron Emission Tomography
POCT	Point of care testing
PSA	Prostate-specific antigen
OECD	Organisation for Economic Cooperation and Development
SPADE	Symptom-Disease Pair Analysis of Diagnostic Error
TPAD	Test results pending at discharge
WHO	World Health Organization

Executive summary

1. Diagnosis is a fundamental part of health care. The correct and timely identification of a health condition is a first step in ensuring that it is properly treated or managed. **Diagnostic error is the failure to deliver an accurate or timely diagnosis, or failure to communicate this to the patient**. It includes, *misdiagnosis*, *delayed diagnosis* and *underdiagnosis*. This report also considers *overdiagnosis* in the scope of diagnostic error. **Depending on the healthcare setting, up to 15% of diagnoses are estimated to represent a diagnostic error**.

2. Tests, tools, diagnostic procedures and information systems are proliferating across healthcare settings to help patients and providers identify the exact nature of health problems. **Despite these technological advances, health systems may still fail to identify and communicate health conditions correctly or in a timely way**. Challenges may arise due to suboptimal clinical skills, problems in decision making, work-environment and organisational factors, fragmented care delivery and limited data and information. This can lead to duplication, unnecessary care and wasteful use of healthcare resources.

3. Most people will experience at least one diagnostic error in their lifetime, sometimes resulting in severe patient harm, as it is estimated that **80% of all harm caused by delayed or misdiagnosis may be preventable**. Most diagnostic errors involve missed, delayed or inaccurate diagnosis of common health conditions, and can be caused by many contributory factors. For instance, variation in application of diagnostic criteria, atypical clinical presentation and lack of specific diagnostic tests are drivers of misdiagnosis of mental health disorders such as ADHD (estimated 1 in 9 adults misdiagnosed in the UK) and of sepsis (an estimated 10% diagnostic error in the United States).

4. **Diagnostic errors negatively impact patient outcomes and increase the use of healthcare services, with associated increased costs**. An estimated 2.6 million diagnostic errors occur in the United States each year, resulting in approximately 371 000 deaths and 424 000 permanent disabilities due to misdiagnosis. **This report estimates that the direct consequences of diagnostic error on healthcare budgets account for 17.5% of total healthcare expenditure.** In the United States this would amount to USD 870 billion each year.

5. **Diagnosis is not a one-off activity, but an iterative and complex ongoing process of information gathering and evaluation**. Risk and uncertainty are intrinsic to the process of diagnosis, and harm sometimes occurs despite high quality care. **Deficits in health system design and governance, clinical environments, and individual provider competencies can drive poor diagnostic outcomes**. Data on diagnostic performance and quality assurance indicators are not routinely collected, analysed or reported by the majority of surveyed OECD countries.

6. **Internationally, guidelines and standards on accurate and timely diagnosis for health conditions can be lacking and not systematically adopted**. Diagnostic stewardship involves ordering the right tests for the right patient at the right time and encompasses the reporting and interpretation of results for optimal clinical management. The Choosing Wisely Initiatives are a practical means of assessing diagnostic performance and instilling a culture of diagnostic safety and excellence at an institutional or

national level. International or national consensus led by professional medical specialty associations on accepted diagnostic criteria for common diagnoses, and best practice diagnostic testing is needed to harmonise clinical practice and reduce diagnostic error and costs.

7. National policies for quality improvements can influence better diagnostic performance. **Policy levers for improving diagnostic safety and reducing diagnostic error include behavioural, technological, and work environment interventions**. Policies promoting a work culture of peer consultation and discussion for diagnostic review, rationalisation and prudent use of diagnostic tests, implementation of electronic health records to track and alert to diagnostic results, and increased patient involvement in the diagnostic process are means of optimising diagnostic practice. Rationalisation of diagnostic radiology, laboratory testing, and complex genomic testing are challenges and opportunities facing healthcare systems, which will require careful governance to avoid overdiagnosis and associated costs and harms to patients and healthcare systems. Advances in, and access to, novel diagnostic technologies and testing risk further inflating the economic costs and health consequences from diagnostic error if left unaddressed.

8. Investment in key policy areas—inappropriate diagnostic testing and treatment, adherence to diagnostic standards and guidelines, instituting a clinical culture of diagnostic review and discussion, and harnessing health data infrastructure for measurement and improvement—can deliver a healthy return. Even a relatively modest target of **halving diagnostic error rates would not only reduce considerable patient suffering and distress but could free up as much as 8% of healthcare expenditure. Across OECD countries, this would equate to USD 676 billion a year.**

1 Understanding diagnosis and diagnostic safety

9. Diagnosis is foundational to the practice of medicine. A safe diagnosis constitutes the correct and timely identification of a health problem or condition. It is a first, critical step in safe, high-quality care that ensures a patient's condition is properly treated or managed. Diagnosis has evolved from history taking and physical examination to include a growing number of tests, tools, and applications. However, failure to provide correct and timely diagnosis is common. Most people experience it at least once in their lifetime, sometimes resulting in severe patient harm (Balogh et al., 2015[1]; Newman-Toker et al., 2023[2])). Depending on the healthcare setting, up to 15% of diagnoses are estimated to represent a diagnostic error (Graber, 2013[3]) (Cheraghi-Sohi et al., 2021[4]) (Singh, Meyer and Thomas, 2014[5]) (Gunderson et al., 2020[6]). But, as with other healthcare interventions, diagnostics, can not only be 'overdone' but also hold an ever-present risk of iatrogenic harm (i.e. harm caused by medical intervention).

10. Demand for healthcare is potentially limitless. Technological advances in all aspects of care, including diagnostics, increase prices and healthcare costs are putting nations' ability to pay under pressure (OECD, 2017[7]; Marino and Lorenzoni, 2019[8])). The result is a growing gap between what can, and what should, be done – a gap that needs to be actively managed if countries are to ensure the best possible patient outcomes without letting healthcare crowd out other valuable areas of spending. With as much as 80% of harm caused by poor diagnostic practice deemed preventable (OECD, 2017[9])), improving diagnostic safety represents an avenue to improve patient outcomes while reducing waste and unnecessary costs.

Diagnosis in the context of the Economics of Patient Safety series

11. In recent years, the OECD has advanced on work to explore the role of policy levers for improving patient safety and expanded in scope of policy analysis (see Box 1.1). A major aspect of this work has been quantifying the costs of patient safety lapses in health systems, to motivate investment in safer health systems. Work began primarily looking at hospital care but has been expanded into assessments of other settings (primary and long-term care) and other aspects of safe care (healthcare worker safety, medication safety and patient engagement).

Box 1.1. OECD Economics of Patient Safety Reports

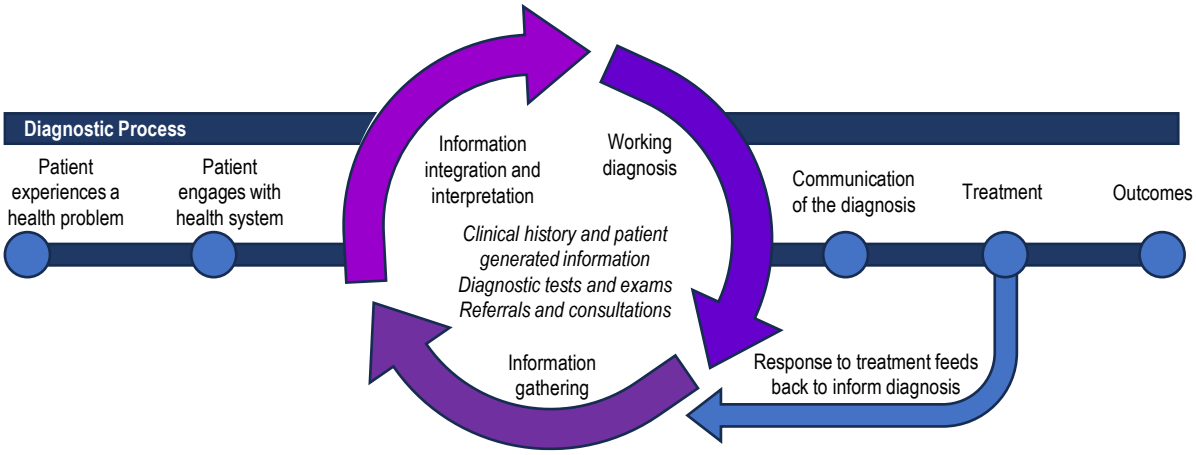
- The Economics of Patient Safety: Patient engagement for patient safety - The why, what, and how of patient engagement for improving patient safety (2023)
- The Economics of Patient Safety: From analysis to action (2022)
- The Economics of Patient Safety Part IV: Safety in the workplace - Occupational safety as the bedrock of resilient health systems (2021)
- The Economics of Patient Safety Part III: Long-term care - Valuing safety for the long haul (2020)
- The Economics of Patient Safety in Primary and Ambulatory Care: Flying Blind (2018)
- Measuring Patient Safety: Opening the Black Box (2018)
- The Economics of Patient Safety: Strengthening a value-based approach to reducing patient harm at national level (2017)

12. This report explores the economics of diagnostic safety. Chapter one describes the key concepts in diagnosis and diagnostic safety, illustrating the unique import of diagnosis on patient safety with case studies. Chapter two examines the scope and health impact of common diagnostic errors in case studies. Chapter three covers the burden of diagnostic error. Chapter four investigates the levers for improving diagnostic safety and chapter five provides a brief conclusion. The content of this report is based on desk research, a 2024 *Survey on the Assessment of the Adoption of Systems* and *Interventions to Improve Diagnostic Safety* of OECD member countries, and discussions with experts in diagnostic safety and quality.

Diagnosis is complex, iterative and error-prone

13. The complexity of diagnosis an important factor in the quality of care. The diagnostic process (see Figure 1.1) is in often iterative (increasingly so given rising chronic disease multimorbidity) meaning initial assessments may not yield a definitive answer but instead guide further investigation. Not arriving at the correct diagnosis immediately isn't a mistake, but rather a step in the process of refining and narrowing down possible causes. This complexity, fragmentation, and iterative nature of diagnosis inflates the risk of errors, delays, and associated harm (Ben-Assuli et al., 2020[10]).

Figure 1.1. The diagnostic process is iterative



Source: Adapted from (Balogh et al., 2015[1]) "Improving Diagnosis in Health Care", https://doi.org/10.17226/21794.

14. **Diagnostic safety** is optimised just before the point where the incremental cost of additional diagnostic testing begins to exceed its incremental benefit (Newman-Toker, McDonald and Meltzer, 2013[11]). The incremental value of each additional assessment is not constant and may rise or diminish based on many factors. Beyond this point, more *diagnostics* does not generate better *diagnosis* in terms of accuracy or timeliness (nor treatment or outcomes).

Diagnosis and diagnostics

15. Diagnosis and diagnostics, while related, refer to distinct concepts. **Diagnosis** is the process of identifying the presence of a disease or condition in a patient. Typically, the diagnostic process is informed by a variety of information sources, including the evaluation of symptoms by their care provider, patient and family histories, and the results of laboratory or other tests. Historically, it has been thought of as an explanation by healthcare providers or as a result after consideration of various investigative processes.

16. The Bayesian nature of diagnosis, where the probability of a diagnosis is continually re-evaluated based on new evidence, means that every clinical interaction is a diagnostic opportunity. How a patient responds to a specific treatment, for example, can provide useful clues for correct diagnosis. A skilled practitioner will note information provided by the patient and/or their family/carers, even if provided incidentally, and use it to develop, refine or confirm a diagnosis.

17. The dynamic and iterative nature of diagnosis and the central role of uncertainly is detailed by Burstin and Cosby (2022): *"During the diagnostic process, it is not unusual, or incorrect, for working diagnostic labels to change as new information is acquired and as the patient's condition evolves both naturally and in response to interventions. The language used to communicate risk of disease and uncertainty about diagnosis is not uniform and may be overly ambiguous (e.g., "cannot rule out," "consider the possibility")* (Burstin and Cosby, 2022[12]).

18. **Diagnostics**, on the other hand, refer to the tools, methods, and procedures used to determine a diagnosis. Many of the methods are based on technologies, including blood tests, imaging studies (e.g. MRI, CT, or ultrasound for example), biopsies, and other medical tests. In areas where there are limited biological markers for disease, such as mental healthcare, patient-reported assessment tools are used to determine if symptoms fulfil the criteria and definitions of mental health disorders (Churruca et al., 2021[13]).

Accuracy of diagnostic tests is a key consideration in the diagnostic process

19. The results of many diagnostic analyses are not necessarily straightforward to report nor interpret, and are open to error. Factors that need to be considered include the likelihood of a patient having a disease along with the reliability and validity of the test results. Diagnostic accuracy relates the ability of a diagnostic test or procedure to properly identify who does or does not have a particular diagnosis. It is impacted by the **reliability** (consistency of the test results over repeated administrations) and **validity** (the accuracy of the test in measuring what it is intended to measure) of diagnostic tools or processes. A reliable test is reproducible under similar conditions, with limited variation. Validity is typically assessed using the concepts of sensitivity and specificity (Shreffler and Huecker, 2023[14]).

20. The **sensitivity** of a diagnostic tool describes if the results properly identify people with the diagnosis of interest (true positives, i.e. a positive diagnosis where the disease is truly present), while **specificity** captures the ability of a test to correctly identify those who do not have the disease or condition (true negatives, i.e. a negative diagnosis where the disease is truly absent). A diagnostic test with high validity would be both sensitive and specific, minimising both false positive and false negative test results. Low accuracy of diagnostic tests can lead to both unnecessary care and stress (in the case of false positive diagnosis) or untimely diagnosis, potential disease progression, and delayed care (in the case of false negative diagnosis). Both sources of error impact health outcomes and can lead to avoidable burdens on patients and health systems. Another important checkpoint in the diagnostic process – and a potential source of diagnostic error – is the interpretation of diagnostic results by clinicians, and how they communicate the result and adapt treatment appropriately.

Diagnostic error comprises three main categories: misdiagnosis, underdiagnosis and overdiagnosis

21. Failure of diagnostic safety results in **diagnostic error**, which is defined here as a failure to deliver an accurate and timely explanation of the patient's health problem or to communicate that explanation to the patient (see Box 1.4 and Diagnostic error rarely stems from a single incident). Within the definition of diagnostic error are the concepts of wrong, delayed and missed diagnosis. In this report, overdiagnosis and underdiagnosis are included in the scope of diagnostic error.

22. **Misdiagnosis**—comprising wrong, delayed and missed diagnosis is the most common and burdensome type of diagnostic error. Most people can expect to experience misdiagnosis during their lifetime. It can result in severe harm (Balogh et al., 2015[1]). Misdiagnosis can be highly visible to patients and communities, and there have been well publicised cases of diagnostic harms affecting public figures. Table 1.1 describes the themes of recent high-profile cases of misdiagnosis, delayed diagnosis, or diagnostic related harm that have been covered in media or become matters of public debate in 2023 and 2024 in OECD countries.

Table 1.1. Examples of publicised cases of diagnostic errors and harms

Australia	Delayed ADHD assessment, colonoscopy recalls, missed sepsis diagnosis
Japan	Cases of overlooking suspected cancer findings in computed tomography (CT) scans1
France	Delayed administration of diagnostic tests in the ER.2
Norway	Cancer, neurodevelopmental disorders and women's health
Slovenia	Cancer misdiagnosis due to histology error
Sweden	Gynecological cancer misdiagnoses

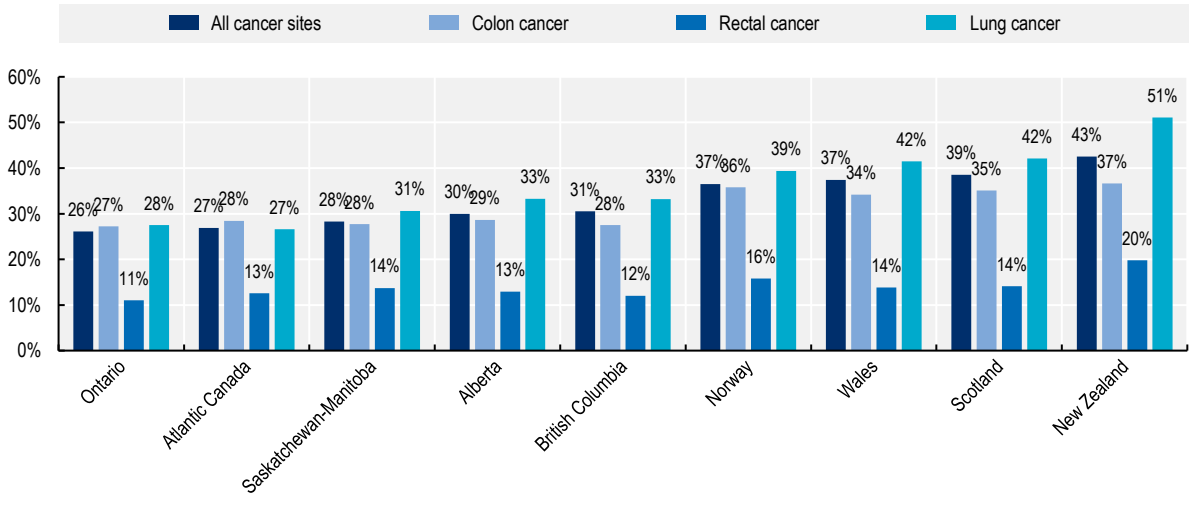
Source: OECD Survey on the Assessment of the Adoption of Systems and Interventions to Improve Diagnostic Safety, 2024

Box 1.2. Diagnostic safety and cancer

Indicators assessing cancer diagnosis resulting from emergency presentations can be used to identify gaps in timely diagnosis by capturing late-stage cancer diagnoses, which often result in poorer patient outcomes ((Zhou et al., 2016[15]). Cancer diagnoses following emergency presentation varies between countries and regions by proportion (Figure 1.2), and has been associated with lower survival and worse patient-outcomes as compared to patients with non-emergency diagnoses, even after adjusting for cancer stage at diagnosis (McPhail et al., 2022[16]). In the United States, 21% of lung cancer and 22% of colorectal cancers were diagnosed following emergency presentation. These delayed or late lung and colorectal cancer diagnoses were associated with 80-90% and 60-90% higher mortality respectively (Kapadia et al., 2024[17]).

In some cases, emergency presentation is unavoidable and is the result of rapidly advancing disease. However, in other cases, emergency presentation in cancers such as lung may reflect disease progression when people are not seeking care through more appropriate channels (e.g. primary care) or for a prolonged period following initial symptoms and diagnosis. For other cancers, diagnosis should preferably occur as part of organised screening efforts rather than via emergency symptoms. Some emergency presentations may therefore be preventable through improved screening and access to diagnostic care, particularly for colorectal and lung cancer (Askari et al., 2017[18]; te Marvelde et al., 2019[19]; Pettit, Al-Hader and Thompson, 2021[20]).

Figure 1.2. Percentage of patients diagnosed through emergency presentation (defined as diagnosis of cancer within 30 days of an emergency hospital admission) by cancer site



Source: International Cancer Benchmarking Partnership (ICBP) population-based study (McPhail et al., 2022[16])

23. **Underdiagnosis** is distinct from misdiagnosis and diagnostic delay in that it encompasses systematic and structural tendencies to neglect appropriate and requisite diagnostics for certain diseases and population groups (Newman-Toker, 2014[21]; Newman-Toker, 2009[22]). Examples include underdiagnosis of stroke in younger patients, ischaemic heart disease in middle-aged females, and depression in older patients. Underdiagnosis is not only inequitable but is burdensome for the patient as well as on the healthcare system as the chronically underdiagnosed will continue to seek increasingly costly medical help as their condition deteriorates.

24. **Overdiagnosis** relates to the diagnosis of a health condition that won't result in negative symptoms or problems for the patient. Often over-diagnosed health conditions result from diagnostic tests and screening methods that detect abnormalities which may never progress or impact the individual's health. Systematic reviews suggest that there is substantial overuse of diagnostic testing present across healthcare settings, with substantial variation in use of similar diagnostic services (Müskens et al., 2021[23]). Overdiagnosis can lead to unnecessary treatments, anxiety, and healthcare costs, without providing any real benefit to the patient but exposing them to iatrogenic harm. The proliferation of increasingly complicated diagnostic methods and techniques has contributed to this risk (Balogh et al., 2015[1]).

Box 1.3. Common instances of processes leading to overdiagnosis

- Imaging for low back pain.
- Imaging for headaches.
- Dual energy X-ray absorptiometry (used to measure bone mineral density).
- Preoperative checks in low-risk patients (electrocardiography, stress electrocardiography, chest radiography).
- Cardiac imaging in low-risk patients.
- Some opportunistic cancer screening activities in low-risk populations (cervical smear test, CA-125 antigen for ovarian cancer, prostate-specific antigen screening, mammography, CT-scan for lung cancer).

Source: (OECD, 2017[24])

Diagnostic error rarely stems from a single incident

25. Patient harm in relation to diagnostic error can result from a single event but given the iterative nature of diagnosis, diagnostic error more often develops over time. Historically, diagnostic errors have been thought of as individual failures—resulting from the shortcomings of individual doctors who misjudge or lack the needed knowledge. While this is still the sometimes the case, diagnostic errors are also the result of systemic shortcomings, resulting from unfavorable conditions to the delivery of accurate or timely diagnosis.

26. At the clinical level, diagnostic error can result in physical and psychological harm to patients. It also impacts care delivery at the organisational level, typically in the form of duplication, revisiting and lengthier care trajectories. At the system level, diagnostic error aggregates to suboptimal safety and quality, manifesting in wasted resources, poorer health outcomes and healthcare efficiency (Auraaen, Slawomirski and Klazinga, 2018[25]).

27. Diagnosis of many common conditions—including sepsis (Angus et al., 2016[26]), irritable bowel syndrome (Enck et al., 2016[27])—still lack consensus on agreed upon classifications and the non-obvious nature of their clinical presentation. Mental health disorder such as Attention-Deficit/Hyperactivity Disorder (ADHD) have seen changes to their diagnostic criteria and application by diagnosing clinicians. Other conditions—such as tuberculosis, cysts and tumours—are frequently misdiagnosed for a wide variety of other (incorrect) conditions (Li et al., 2020[28]). However, diagnostic error isn't limited to rare, complex disorders. It occurs just as frequently in conditions such as asthma, diabetes, cardiovascular disease, cancer and infection, all commonly encountered in primary and emergency care settings (Laposata, 2022[29]; Balogh et al., 2015[1]) (Singh et al., 2013[30]). Key diagnostic concepts are summarised in Box 1.4.

28. Diagnostic error may begin with well-intentioned efforts to offer preventive care. General health checks are offered in many health systems with the objectives of detecting disease and managing risk factors, however the evidence suggesting their benefit in improving cardiovascular or cancer health outcomes is mixed (Krogsbøll, Jørgensen and Gøtzsche, 2019[31]) (McCracken et al., 2024[32]). While the UK study demonstrated reduced long-term mortality in those who underwent NHS Health Checks, healthy participant bias may explain some of the observed results, and may give rise to health inequalities.

29. Uncoordinated population screening activities may also lead to adverse outcomes and higher costs resulting from false positives or overdiagnosis (Kherad and Carneiro, 2023[33]), depending on the reliability of the tests used. For example, asymptomatic individuals are offered individual health assessments that include CT scans, which not only deliver high doses of radiation but are more likely to produce incidental findings ('incidentalomas') than actual pathology (Malone et al., 2016[34]; WHO, 2017[35]). This not only entails potentially negative consequences for the individual but represents questionable use of healthcare resources. Moreover, the majority of the commonly used individual screening tests offered have been incompletely validated and can lead to increased use of diagnostic and therapeutic interventions (Krogsbøll, Jørgensen and Gøtzsche, 2019[36]). For example, when used in asymptomatic people, COVID-19 antigen tests have been found to correctly identify the absence of infection only 55% of the time (Cochrane COVID-19 Diagnostic Test Accuracy Group, 2022[37]). The importance of correct interpretation of diagnostic test results is paramount to avoid misdiagnosis in the process (Kulasekere et al., 2024[38]).

Box 1.4. Key diagnostic concepts

- **Diagnosis:** the iterative and dynamic process of detecting and correctly identifying a disease in a patient.
  A complex, patient-centered, collaborative activity that involves information gathering and clinical reasoning with the goal of determining a patient's health problem. This process occurs over time, within the context of a larger healthcare work system that influences the diagnostic process (National Academies of Sciences, 2015[39]).
- **Diagnostic error:** the failure to provide an accurate and timely explanation of the patient's health problems or communicate that explanation to the patient (National Academies of Sciences, 2015[39]). Missed opportunities to make a correct or timely diagnosis based on the available evidence, regardless of patient harm (Singh and Sittig, 2015[40]).
- **Safety II**: the ability to make things go right and note merely the absence of failure or adverse outcomes. In the context of diagnosis, learning from instances where diagnoses are correct and the diagnostic process functions (discussed section N.N).

Source: (Shreffler and Huecker, 2023[14])

Diagnostic activities account for more than 10% of healthcare spending

30. What do health systems spend on diagnosis and diagnostics? The global diagnostics market was valued at USD 158.6 billion in 2021 and expected to reach almost USD 350 billion in 2030 (PRNewswire, 2022[41]). The latest available estimates suggest that the United States devotes approximately 9% of its healthcare expenditure to diagnostic imaging, and a further 2.3% to in vitro laboratory tests (Horný et al., 2024[42]; Ducatman et al., 2020[43]). These two diagnostic activities therefore account for 11.3% of the United States' healthcare spend – or just over two percent of GDP, a figure that is likely to rise as molecular and genetic testing for precision medicine become mainstream (Schulman and Tunis, 2010[44]) (Douglas and Kumar, 2022[45]) In this new medical frontier, boundaries between diagnosis and treatment are increasingly overlapping. For example, the implementation of targeted cancer therapies based on a person's genetic background can only be achieved if paired with bespoke in vitro testing (OECD, 2019[46]).

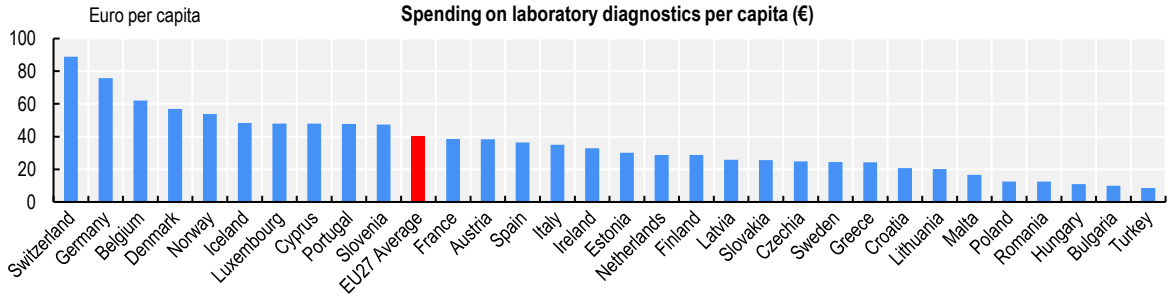
31. The actual expenditure and time invested in diagnosis is probably underestimated. As outlined above, diagnosis entails more than just diagnostic tests and investigations. Every feature of the subjective (i.e. history taking) and objective clinical examination (i.e. physical examination) is de facto a diagnostic 'test'. These embedded diagnostic practices generate cost in the form of practitioner and patient time.

32. Given the fundamental importance of a diagnosis in high-quality care, it is quite reasonable to expect a considerable proportion of resources to be devoted to it. The question is if implemented diagnostic practices represent value. In other words, **do the incremental health and societal benefit that diagnostic activities produce enable exceed their costs** and, **what can be done to (a) identify high- and low-value practices, (b) promote the former and (c) eradicate the latter**.

Regulation and access to diagnostics are evolving

33. The increased variety, availability, and use of diagnostic tests and techniques offers significant opportunities to improve patient outcomes. However, this growth also introduces potential challenges, such as overtesting, overdiagnosis, misinterpretation of results, and requires robust systems to ensure patient safety and quality care. Market research suggests that while the results of in vitro diagnostic medical devices (IVDs) influence as many as 70% of clinical decisions, while IVDs account for just 1% of total healthcare expenditure in the EU European IVD Market Statistics Report (MedTech Europe, 2022[47]). On average across European countries, spending on laboratory diagnostics is just over 40 euros per capita—and over 70 in Switzerland and Germany (see Figure 1.3).

Figure 1.3. European countries spend an average of 40 euros per capita on diagnostic laboratory tests



Source: (MedTech Europe, 2022[47])

34. Recent regulatory changes in the EU have been implemented with a view of improving safety, and the implementation of Regulation (EU) 2017/746 on IVDs, which increases the requirements for clinical evidence, market surveillance and conformity assessments. In particular this legislation mandates that manufacturers gather, record and analyse metrics of IVD quality, performance and safety more extensively post-market authorisation (Danish Medicines Agency, n.d.[48]).

35. Point-of-care testing (POCT) and distributed access to diagnostics can enable improved access and accessibility of diagnostics—facilitating faster decision-making and earlier interventions. However, they also raise patient safety concerns, including the potential for inaccurate results or misinterpretation. Rapid tests for COVID-19 increased the acceptability and support for use of POCT, and POCTs area also widely used in glucose monitoring, coagulation testing, and cardiac monitoring (Clearstate, 2023[49]). A review of published studies found that home-based testing is often preferred by patients as compared to clinic-based testing, and that—importantly—follow-up treatment after a positive home-based test is generally high and, in some cases, even higher than tests done in a clinic setting (Versluis et al., 2022[50]).

Measuring diagnostic error is difficult

36. Estimating the incidence and sequelae of diagnostic error differs to harm from curative care. With the latter, it is a case of calculating the additional clinical management and care resulting from the adverse event, which typically requires additional tests and treatment, an extended hospital stay or a hospital admission to treat the harm. Failure in diagnostic safety, meanwhile, is more remote to the outcomes desired by patients, providers and payers. It can precipitate a variety of consequences depending on the type of failure.

37. Defining and measuring diagnostic error to report diagnostic error rates is challenging. Depending on the disease and the setting, determining how and when a disease is presenting or progressing to warrant being diagnosed at a given time, while balancing the relative risks of under- and overdiagnosis, is subject to debate (Zwaan and Singh, 2015[51]). Owing to the iterative nature of diagnosis, subjectively identifying the optimal or reasonable window period which represents a missed opportunity for diagnosis is equally challenging, due to lack of information on the clinical context. Furthermore, different studies employ different definitions for diagnostic error, preventing direct comparison of estimates.

38. Studies to accurately detect diagnostic error are resource-intensive, often requiring detailed analysis of medical records' free-text, given the absence of a 'misdiagnosis' field plus the reluctance of healthcare professionals in many countries to write these terms in the records (Kerber et al., 2011[52]). (Kulasekere et al., 2024[38]). They also rely on assumptions about the counterfactual because, inter alia, a comparison cohort of similar patients who did not experience that specific diagnostic error with the same pathological and clinical consequences is not easily identified.

39. Harms from curative care such as nosocomial infections or inpatient falls are typically documented in the medical record, then coded and placed in administrative, morbidity or registry data. While the causes of these harms care can certainly be complex, the sequelae are – in most cases – circumscribed quite clearly and, with some exceptions like venous thromboembolism, become evident more or less immediately. Diagnostic error, on the other hand can be more occult and even silent. The sequelae may be latent for months even years (Balogh et al., 2015[1]). Even if clearly documented in a subsequent medical record entry, it most often isn't picked up by monitoring systems due to the passage of time, fragmentation of medical records, or the absence of clinical codes to capture diagnostic problems.

Diagnostic error may be greatly underestimated

40. The handful of well-designed studies examining diagnostic error suggest that it occurs more often than commonly thought. For example, a prospective United States study using unannounced standardised patient visits[7] found that the true costs of diagnostic error during internal medicine consultations were 20 times higher than would be discerned by retrospective record review (Schwartz et al., 2012[53]). A study of using electronic health records to identify diagnostic error in emergency paediatric admissions found that existing reporting systems identified less than 10% of diagnostic error (Lam et al., 2022[54]). The actual burden of diagnostic error therefore also likely to be greater than currently thought. Chapter 2 examines these issues in specific case studies and Chapter 3 explores the scope of the burden in more detail.

Safety II: complexity, resilience and learning from what goes right

41. Diagnostic safety, as defined in this report, is closely linked to diagnostic excellence, which "involves making a correct and timely diagnosis using the fewest resources while maximising patient experience and managing uncertainty" (Meyer & Singh 2019). Avoiding mistakes is of course fundamental, but a good diagnostic process also provides an agreeable experience for the patient while using healthcare resources prudently. In practice, this includes learning from things go well as well as when things go wrong.

42. The traditional approach based on learning from what goes wrong in healthcare (Safety I) must be complemented by a proactive approach that tries to understand and optimize success. 'Safety II' can be described as "the ability to make things go right and note merely the absence of failure or adverse outcomes" (Braithwaite, Wears and Hollnagel, 2015[55]; Hollnagel, Wears and Braithwaite, 2015[56]). It promotes learning from what goes right, from examples of resilience, and from positive deviance or innovative safety-enhancing practices. In the context of diagnosis, this means gathering data and learning from areas, settings, organisations and diseases/conditions where diagnostic safety is consistently high.

43. Safety II accepts uncertainty as an innate feature of the business of healthcare. Pathology manifests differently between individuals, who also respond differently to investigations and treatments. Safety II is therefore uniquely suited to diagnostic practice. All investigations are neither 100% accurate nor 100% safe. All entail risk harm and a degree of uncertainty, which can never be fully eradicated with additional investigations. Good diagnostic practice means balancing the risks and benefits of testing in a dynamic way because, not only does uncertainly vary between patients with similar characteristics or presenting symptoms, but uncertainty is also non-linear (it changes with additional diagnostics *and* as disease progresses). Then there are the financial risks. The cost of the many tests and scans can be substantial. So can the costs of interventions that flow from any incorrect diagnosis. Importantly, Safety II should be seen as complementary to Safety I, with both playing an important role in the continuing objective of minimising iatrogenic harm (Verhagen et al., 2022[57]).

44. Safety II has for some time been embraced other high-risk industries (air travel, automotive, oil and gas) as the preferred way to manage risk. The underpinning micro-economic principle is relatively simple: weighing the costs of preventing errors against the costs incurred by error. Investing in the prevention of harm (prevention costs) thus creates long term value through the reduction of the costs incurred by ameliorating adverse events (failure costs). Despite fundamental differences between health care and these other industries, the enablers and barriers of such an industry-wide approach translate relatively well. These range from the cultural (e.g. vertically aligned commitment to prevent harm) to the socio-technical (e.g. using data for continuous improvement) and the financial and reputational (failure costs are felt by throughout the entire organisation) (Slawomirski and Klazinga, 2022[58]; Auraaen, Saar and Klazinga, 2020[59]).

2 Case studies illustrating diagnostic error

45. The majority of diagnostic errors are accounted for by commonly encountered conditions in medical practice, which increases their burden in terms of health outcomes and costs from diagnostic harm. Together, cardiovascular events, infections, and cancer account for 76% of diagnostic harm reported by malpractice claims in the United States (Newman-Toker et al., 2024[60]). Diagnostic error stemming from individual providers or the healthcare system may originate from the variable clinical presentation, lack of consensus or implementation of diagnostic criteria, uncertainties over interpretation of diagnostic results, or communication and follow-up care of these. This section examines the considerable risk of diagnostic error and overdiagnosis in the process of diagnosing the following common conditions: mental health disorders, sepsis, cancer screening, and Long COVID, through a series of case studies, as well as their related health consequences and costs. Included are findings from a survey of OECD countries on monitoring and reporting of indicators for diagnostic error and diagnostic review for these conditions.

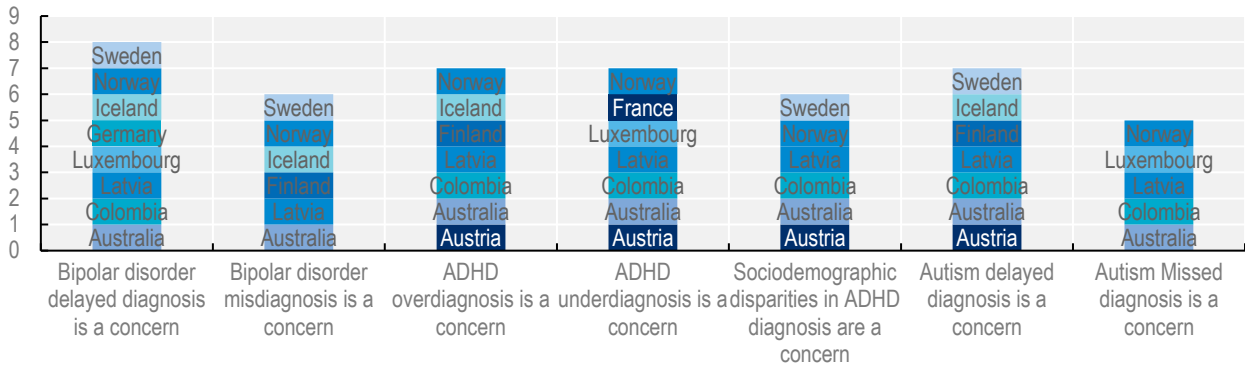
Case study: Mental Health disorders

46. The spectrum of clinical presentations, cognitive biases of clinicians, and the overlap between aspects of different conditions mean that diagnosis of mental health disorder is sometimes challenging, resulting in delayed, misdiagnosis or overdiagnosis (Bradford et al., 2024[61]). For example, diagnosis of anxiety disorder made in primary care may be incorrect in 60% of cases (Fletcher et al., 2020[62]).

47. Diagnostic biomarkers for mental health disorders are limited. Instead, standard diagnostic criteria such as Diagnostic and Statistical Manual of Mental Disorders (DSM-V) and structured interviews are routinely used as psychological assessment and testing. Errors in diagnosis can also driven by changing perceptions and demand for intervention for mental health conditions such as Attention Deficit Hyperactive Disorder (ADHD) and autistic spectrum disorder (ASD) (Davidovitch et al., 2017[63]). Cognitive biases can con contribute to underdiagnosis in certain populations, such as among adults or children.

48. The costs of diagnostic error for patients and healthcare systems include prolonged use of psychotropic medication without review of the original diagnosis (McCool et al., 2022[64]). In the 2024 OECD *Survey on the Assessment of the Adoption of Systems and Interventions to Improve Diagnostic Safety*, countries commonly reported concerns around misdiagnosis of bipolar disorder, overdiagnosis of ADHD, and delayed diagnosis of ASD (see Figure 2.1). Mental health disorders are among the most commonly *underdiagnosed* conditions, at considerable societal cost. A cross-sectional study of Indiana residents found that almost 430 000 had untreated mental illness. The total economic cost associated with this was estimated at $4.2 billion per annum – just over 1% of Indiana's GDP – most of which ($3.3 billion) were indirect costs from unemployment and lower productivity (Taylor et al., 2023[65]).

Figure 2.1. Delayed diagnosis is a concern for bipolar disorder and autism, whereas over- and underdiagnosis is an issue for ADHD



Source: OECD Survey on the Assessment of the Adoption of Systems and Interventions to Improve Diagnostic Safety, 2024

Misdiagnosis of bipolar disorder is associated with increased emergency presentations

49. Bipolar disorder affects approximately one in 150 adults or 40 million people worldwide (Collaborators, 2022[66]) and is associated with annual healthcare costs of $21 000 USD per patient related to hospitalisations and medications (Dembek et al., 2023[67]). Misdiagnosis of bipolar disorder as (unipolar) major depressive disorder results in diagnostic delay before initiation of appropriate treatment. Regardless of whether first consultation for major depressive disorder was in the primary care or psychiatric care sector, the delay to a correct diagnosis of bipolar disorder exceeded 300 days (mean of 405 days versus 396 days) (McIntyre et al., 2022[68]). At one year post first major depressive disorder diagnosis in the United States, 42% of patients with bipolar disorder still had not been correctly diagnosed. This diagnostic delay results in higher rates of emergency presentation and hospitalisation for misdiagnosed patients with bipolar disorder, and 45% higher costs compared to those correctly diagnosed (McIntyre et al., 2022[68]).

50. Eight of the OECD countries surveyed reported concerns about delayed diagnosis, and six reported concerns around misdiagnosis of bipolar disorder (see Figure 2.1). None of the surveyed OECD countries reported the mean delay from initial presentation to diagnosis of bipolar disorder.

Increasing trends in ADHD diagnoses in children indicates overdiagnosis, while many adults remain undiagnosed

51. ADHD is a neurodevelopmental disorder characterised by persistent inattention, hyperactivity and impulsive behaviour, that has a negative effect on social functioning. ADHD is estimated to affect 2.8% of adults worldwide (Fayyad et al., 2017[69]). A UK study reported that 0.3% of adults in primary care had a diagnosis of ADHD, which suggests that only 1 in 9 adults are accurately diagnosed (O'Nions et al., 2025[70]). The finding that an ADHD diagnosis is associated with a life expectancy reduction of 6-9 of PYLL, means the consequences of undiagnosed ADHD in adulthood may include premature mortality.

52. The diagnosis of ADHD in children has increased over a 10 year period in the United States (from 8% to 11.0%) and doubled in Israel (7% to 14%) (Davidovitch et al., 2017[63]) (Visser et al., 2014[71]) (Davidovitch et al., 2017[63]). This trend has been accompanied by increased prescribing of ADHD medication (3.6% to 8.5%) in Israel, where changing attitudes and expectations towards a diagnosis of ADHD suggested as the likely explanation. Some of this trend likely represents overdiagnosis, stemming from increased recognition of the syndrome, changes in the guidelines for diagnostic criteria, or inappropriate application of the latter by healthcare professionals (Manos, Giuliano and Geyer, 2017[72]).

The United States reported increased misuse of stimulants by 5 million persons aged over 12 years, with concerns this may stem from overdiagnosis of ADHD resulting from changes to DSM-V diagnostic criteria or variations in prescribing patterns, in addition to inappropriate procurement of ADHD medication (Moustafa, Chauhan and Rummans, 2022[73]).

53. European countries including Germany, Denmark, United Kingdom, Netherlands reported increases in ADHD medication prescribing in children of more than 50% from 2006 to 2012, indicative of increasing ADHD diagnosis (Bachmann et al., 2017[74]). More recent trends in ADHD diagnosis in Germany up to 2018 suggest that diagnoses among children peaked in 2012 and have subsequently declined, while remaining stable among adults (Grimmsmann and Himmel, 2021[75]).

54. Among OECD countries surveyed, **Austria, Australia, Colombia, Latvia, Finland, Iceland and Norway** reported concerns about ADHD overdiagnosis (see Figure 2.1). Just as many countries reported concerns about ADHD underdiagnosis (**Austria, Australia, Colombia, France, Latvia and Luxembourg and Norway**), indicating the considerable potential diagnostic error for this mental health condition among child and adult populations. **Finland** and **Norway** reported concerns over geographical disparities in diagnosis rates of ADHD, while **Australia** and **Sweden** reported socioeconomic disparities, indicative of over- or under diagnosis.

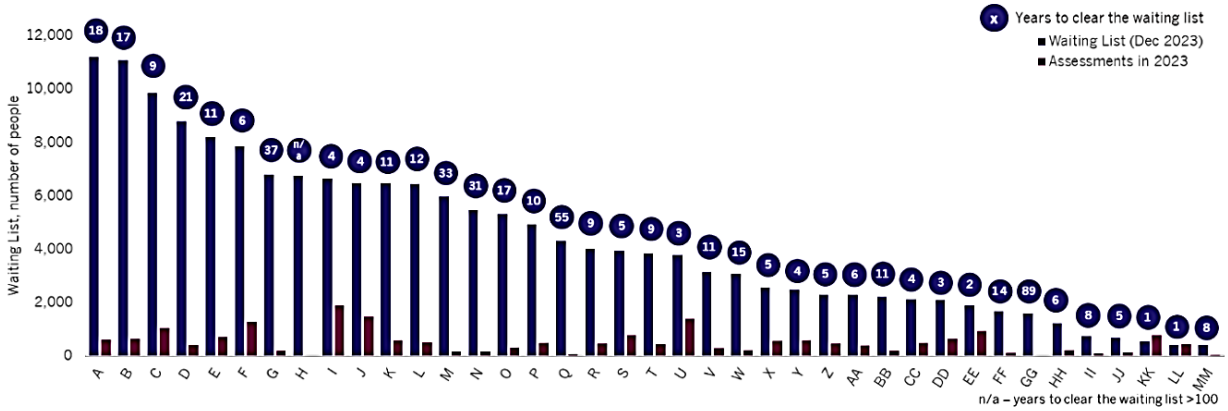
55. While evidence suggests ADHD overdiagnosis in children, the opposite may be the case in adults (Ginsberg et al., 2014[76]; Oliva et al., 2020[77]). In the United States the cost of undiagnosed and untreated ADHD in adults is estimated to be as high as USD 190 billion. Over 80% of these costs are – again – indirect, driven by unemployment and loss of productivity (Doshi et al., 2012[78]; Schein et al., 2022[79]). The reasons cited include adults' internalisation of hyperactivity symptoms (Ginsberg et al., 2014[76]; Oliva et al., 2020[77]; Rao and Place, 2011[80]), compensation strategies to mask symptoms (Weibel et al., 2020[81]), and misdiagnosis as anxiety or personality disorders (Schein et al., 2022[79]).

Overdiagnosis of ADHD is associated with increased medication costs and burden on mental healthcare services

56. The increased demand for ADHD assessments risks overwhelming available healthcare resources in the UK, which at current rates would take on average 8 years to clear the backlog in the demand for adult ADHD assessments (see Figure 2.2). There is no consensus on the explanation behind the dramatic increase in demand for assessment for ADHD, whether due to unmet need in demand for assessment as awareness of ADHD has increased, or the result of self-diagnosis stemming from by misinformation (Darzi, 2024[82]).

Figure 2.2. Increased demand for ADHD assessment suggests overdiagnosis and strains healthcare resources

Implied clearance time in years for adult ADHD assessments based on activity and waiting lists (44 providers, UK)



Source: (Darzi, 2024[83])

57. The patient follow-up and medication prescribing related to overdiagnosis of ADHD has considerable economic costs, given that two thirds of those diagnosed with ADHD in the United States are on ADHD medication. An extra 2 million children were diagnosed with ADHD in the United States in 2011, with associated healthcare costs estimated at USD 143-266 billion, which represents a considerable burden to the US healthcare system (Visser et al., 2014[71]), in addition to patient harm caused by overdiagnosis and overtreatment (Doshi et al., 2012[78]).

Delayed or missed diagnosis of autism reflects lack of access to services and awareness of ASD

58. Autism (autistic spectrum disorders) are neurodevelopmental disorders affecting approximately 1 in 100 children or 28 million people worldwide (Zeidan et al., 2022[84]) (Collaborators, 2022[66]), which typically manifest and are diagnosed in early childhood. The mean age at diagnosis was 55 months (range 7 to 223 months) in both a UK study (Brett et al., 2016[85]), and in a United States study, although this age has not decreased with successive birth cohorts in the UK (Brett et al., 2016[85]), unlike in the United States (Hanley et al., 2021[86]). Missing this early diagnosis can lead to children entering adolescence and even adulthood without a diagnosis of ASD, with considerable adverse impact on their social and occupational functioning. Of the OECD countries surveyed, seven reported concerns around delayed diagnosis of ASD (see Figure 2.1). None of the surveyed OECD countries reported the mean age at ASD diagnosis.

Review of mental health diagnosis and medications to reduce diagnostic error harms

59. The need for review of initial diagnosis and of prescribed psychotropic medication is apparent in mental health, where symptoms may improve or fluctuate over time. This is one means of reducing the consequences of diagnostic error, and of particularly importance given the adverse drug reactions and health risks of long-term psychotropic medication use (Correll et al., 2017[87]). While the initial mental health diagnosis may be made on clinical assessment and fulfilment of diagnostic criteria, improvement or deterioration in the patient's mental state or development should warrant a diagnostic review to ensure misdiagnosis has not occurred.

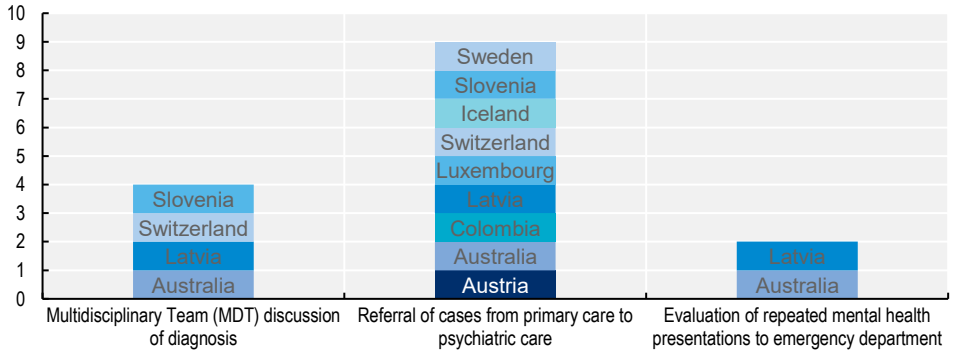
60. A study of primary care patients in Ireland reported an steady increase in the proportion of renewed antidepressant prescriptions from 2016 to 2020, indicating increased trend for long-term prescribing for

depression (McCool et al., 2022[64]) (McCool et al., 2022[64]). Primary care physicians in the UK surveyed in 2019 reported major uncertainty around diagnostic review and cessation of antidepressant medication, with only a minority (17%) confident in their able to difficulty distinguish antidepressant withdrawal from a relapse in depression symptoms (Read et al., 2020[88]) (Read et al., 2020[88]). A UK survey of primary care patients diagnosed with depression and prescribed antidepressants reported 65% had never had a discussion with their provider about cessation of medication, and 48% did not have a frequent medication review by their provider (Read et al., 2019[89]) (Read et al., 2019[89]).

61. Retrospective review of initial mental health diagnosis and associated prescribing can help avoid inadvertent consequences related to inappropriate pharmacological treatment, as well as reducing the financial burden for patients and healthcare systems. The Safer Dx instrument applied to the review of unspecified anxiety disorder diagnoses made in primary care identified diagnostic error in 19% of cases (Fletcher et al., 2020[62]). Routine diagnostic review or consensus on patient diagnosis via multidisciplinary teams or secondary referral services for mental health, are policy options to improve patient safety here.

62. The WHO technical series on *Safer Primary care: Diagnostic errors*, proposes solutions to improve reduce diagnostic error and improve safety, such as improving clinical education and training and improving health systems and information technology for reviewing patient data (World Health Organization, 2016[90]). The theme of World Patient Safety Day 2024 focuses on improving diagnosis, specifically diagnostic review andrelated to diagnosis. Only four countries surveyed (**Australia, Latvia, Slovenia** and **Switzerland**) reported routine multidisciplinary discussion of mental health diagnosis, whereas a majority reported referral of cases from primary to specialist psychiatric care (see Figure 2.3).

Figure 2.3. A majority of surveyed OECD countries report review of primary care diagnosis by psychiatric care, but few report a multidisciplinary team approach



Source: OECD Survey on the Assessment of the Adoption of Systems and Interventions to Improve Diagnostic Safety, 2024

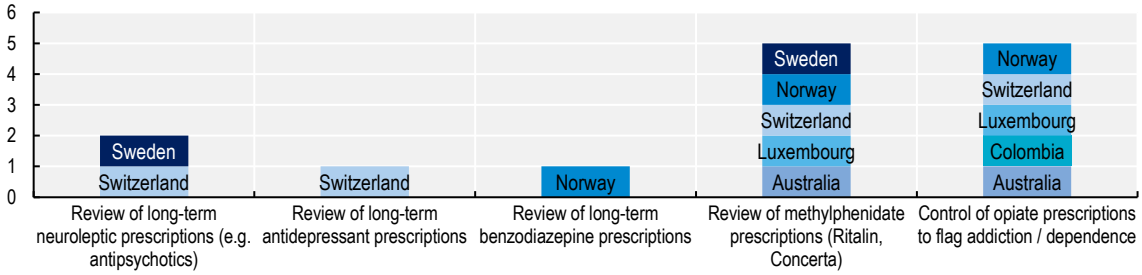
Diagnostic error in mental health often leads to long-term prescribing of psychotropic medication

63. Antipsychotic medication use in people without a diagnosis of psychosis is widely regarded as inappropriate in the care of older adults in long-term care (LTC). Sometimes dementia-related behaviours like aggression and responsive behaviours are overtreated with antipsychotics in LTC settings. Inappropriate antipsychotic medication use poses significant risks, including strokes, falls, fractures, and even death. This practice highlights a critical safety and quality of care issue, as it fails to address the underlying diagnosis.

64. Financially, Canadians spent over $123 million CAD on potentially inappropriately prescribed antipsychotic medications (Huon et al., 2024[91]). With recent data from the Canadian Institute for Health Information revealing rising rates of potentially inappropriate use of antipsychotics in LTC to 24.5%, there is a significant need to establish national targets to guide their appropriate use. The National Appropriate Use Coalition is working to establish these national targets as part of a coordinated effort to reduce inappropriate prescribing and ensure consistent, high-quality care across LTC homes.

65. Long-term prescribing of methylphenidate, antidepressant, benzodiazepine, or antipsychotic medications for misdiagnosed mental health disorders are associated with a range of adverse events that can endanger patient safety. Additionally, their associated costs are considerable as well as the potential cost savings from medication review. Five OECD countries reported reviews of long-term prescribing for ADHD medication (**Australia, Luxembourg, Norway, Sweden** and **Switzerland**) and control of opiate prescribing (**Australia, Colombia, Luxembourg, Norway** and **Switzerland**), but not for other mental health disorders where overdiagnosis or misdiagnosis may be a concern (see Figure 2.4).

Figure 2.4. Few OECD countries implement review of long-term medication prescribing for mental health disorders



Source: OECD Survey on the Assessment of the Adoption of Systems and Interventions to Improve Diagnostic Safety, 2024

Case study: Sepsis

66. Sepsis, defined as a 'life-threatening organ dysfunction caused by a dysregulated host response to infection' (Singer et al., 2016[92]), is among the most common causes of in-hospital death and one of the most expensive conditions to treat, with an estimated 2.7% of national healthcare budgets dedicated to care for sepsis (Van den Berg et al., 2022[93]). The fast progression of sepsis to life-threatening state of multiple organ failure makes timely and accurate diagnosis imperative. Globally, there are nearly 50 million cases and 11 million deaths each year (Rudd et al., 2020[94]) with higher risk in children, the elderly, and immunocompromised.

Lack of a reliable diagnostic test contributes to misdiagnosis and delayed diagnosis for sepsis

67. A fundamental challenge is differentiating sepsis from non-infectious inflammation which can present with similar clinical symptoms. Identification of the presence of infection remains they key to diagnosis in most cases. While various diagnostic tests and tools exist, none provide unambiguous results quickly enough to identify all sepsis cases (Duncan et al., 2021[95]). Initial evaluation of patients with suspected sepsis should include basic laboratory tests, blood cultures, imaging studies, and sepsis biomarkers (Gauer, Forbes and Boyer, 2020[96]). Nevertheless, an estimated one third of confirmed sepsis cases do not identify a conclusive cause of the infection (Meyer and Prescott, 2024[97]).

68. Particularly in cases where the infection source is unknown, diagnosis relies heavily on clinical symptoms, which presents several difficulties. The condition can manifest differently depending on infection site, pathogen, affected organs and patient's baseline health status. While advised against as a standalone screening tool due to low sensitivity, a large change in the Sequential Organ Failure Assessment (qSOFA) score over a short time in the presence of a suspected infection can alert clinicians to a diagnosis of sepsis (Evans et al., 2021[98]). While emerging technologies such machine learning tools using multiple inputs and electronic health records show potential in improving diagnostic accuracy, these have not yet been validated with sufficient accuracy to be widely adopted in practice (Duncan et al., 2021[95]).

Misdiagnosis of sepsis impedes timely interventions, with drastic impacts on patient outcomes

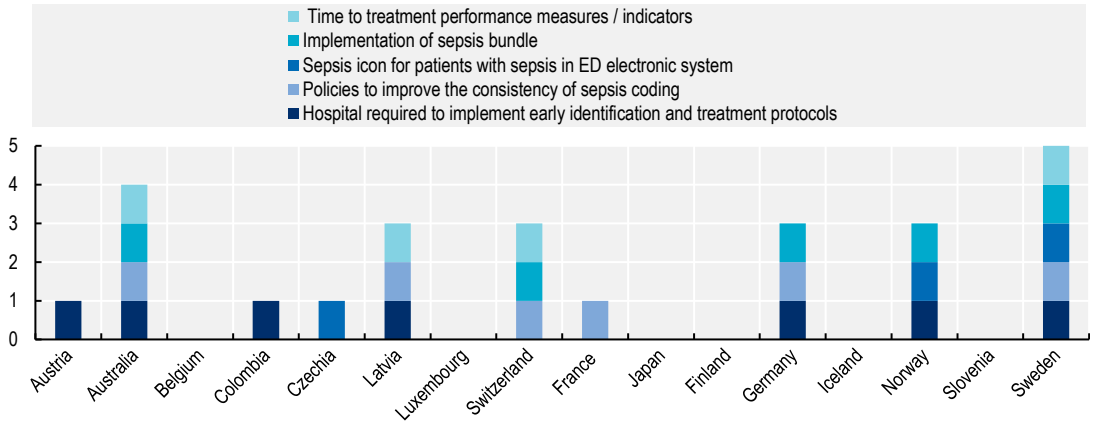
69. Accurate identification and management of sepsis without unnecessary delays drastically improves outcomes (Evans et al., 2021[98]). Sepsis if left untreated and the infection not controlled in a timely manner, can escalate to septic shock – associated with worse outcomes including multiple organ failure and death. Delays in diagnosis of sepsis delay interventions such as fluid resuscitation and antimicrobial treatment, with guidelines advising starting antimicrobials within three hours of suspicion of sepsis. Some guidelines recommend shortening this to one hour (Evans et al., 2021[98]). These diagnostic delays can have serious consequences, potentially affecting not only survival rates but also long-term outcomes including hospital readmission, physical disability, cognitive impairment, and quality of life.

70. However, misdiagnosis of sepsis in patients without infection leads to unnecessary interventions, and delays in appropriate treatment for their true diagnosis. The diagnostic error rate is estimated at 9.9% of sepsis cases by a review of clinical studies in the United States (Newman-Toker et al., 2024[60]). Other estimates have found higher rates, with up to one third of patients treated for bacterial sepsis not having had an infection in retrospect (Meyer and Prescott, 2024[97]). Diagnostic error can also lead to inappropriate use of broad-spectrum antimicrobials – which can be appropriate for infections of unknown origin – but which fuel antimicrobial resistance. With one in five infections across OECD countries caused by resistant bacteria, this presents an important risk for effective management of sepsis (OECD, 2023[99]). For patients with suspected sepsis but unconfirmed infection, it is recommended to continuously review the diagnosis, and discontinue antimicrobials if an alternative illness is strongly suspected (Evans et al., 2021[98]).

A minority of OECD countries report procedures to improve diagnosis and monitoring of sepsis

71. Numerous clinical and coding interventions such as sepsis protocols and care pathways have been developed to improve the early detection, diagnosis and survival from sepsis (Burke et al., 2019[100]). Among the OECD countries surveyed as part of the 2024 Diagnostic Safety survey, **Sweden** and **Australia** lead the way in terms of the number of intervention types implemented, closely followed by **Latvia, Switzerland, Germany** and **Norway** (see Figure 2.5).

Figure 2.5. Interventions to improve the recognition and diagnosis of sepsis



Source: OECD Survey on the Assessment of the Adoption of Systems and Interventions to Improve Diagnostic Safety, 2024

72. The implementation of systematic sepsis screening protocols for high-risk patients has been recommended since 2016 in patients presenting with severe infection or acute organ disfunction that is not attributable to a non-infectious cause (Singer et al., 2016[92]). While seven surveyed countries reported implementation of a sepsis diagnosis and treatment protocol, six of the countries surveyed reported none of the interventions in place. Five countries implement sepsis bundles with key care elements which promote timely management of sepsis.

73. A sepsis code protocol implemented in Spain is a strategy aimed to reduce variability of care and mobilise trained personnel with a focus on early diagnosis and rapid treatment of sepsis. Emerging results suggest a reduction sepsis-related mortality and increased early administration of treatments. However, variability in activation criteria and inconsistent application are ongoing challenges. Similarly in Portugal, emergency department *Via Verde* sepsis screening and fast-track protocols have been implemented in hospitals, ensuring rapid assessment on presentation and availability of trained personnel to manage suspected sepsis cases, although national monitoring of performance is not ongoing.

74. Inconsistencies in recognising and recording sepsis in administrative data make it challenging to obtain reliable metrics on sepsis care. Inconsistent application of diagnostic criteria for sepsis, and sepsis documentation outside the hospital settings (where the majority of sepsis first presents and is mis- or underdiagnosed), mean that measuring the incidence and timing of sepsis, and related diagnostic error, is difficult to capture from medical records (Angus and Bindman, 2022[101]). Twelve of the surveyed countries reported an absence of any indictors for pre-hospital diagnosis of sepsis, while four reported measuring quality measures such as time to administration of antibiotics or consultation. Six countries reported policies to improve the consistency of sepsis coding. Only four OECD countries conduct audits of sepsis diagnosis and delays, with **Germany** and **Norway** doing so on a national level, and **Sweden** and **Australia** on a regional level.

Case study: Cancer

75. Cancer is a condition which may be difficult to detect and diagnose in its early stages, due to the non-specific and often mild symptoms, or its atypical occurrence in younger age groups. Symptoms of cancer may be confused with other common causes of illness, and the diagnosis missed due to failure to investigate with relevant diagnostic testing.

76. The reported rates of diagnostic error depend on the cancer type and the definition used to determine missed opportunities to diagnose cancer symptoms at an earlier time, and the acceptable time

delays relative to diagnosing each cancer type. Estimated diagnostic error rates for colorectal cancer in the United States range from 31% where a missed opportunity was retrospectively identified (Singh et al., 2009[102]), to 24% where diagnostic delay was greater than 8 months (Pruitt et al., 2013[103]). Estimated diagnostic error rates for lung cancer in the United States range from 38% where a missed opportunity for diagnosis was retrospectively identified (Singh et al., 2010[104]), to 23% for delayed care (Nadpara, Madhavan and Tworek, 2015[105]).

77. Cancer screening is not practical or effective for all cancer types, but is common practice for breast, lung, and cervical cancer, and is organised via population level screening programmes in the majority of OECD countries. Screening for common cancers has different consequences and costs of diagnostic error, which vary depending on the cancer types and rate of progression (Esserman et al., 2014[106]).

Overdiagnosis is a costly consequence of cancer screening

78. Overdiagnosis in cancer screening involves detection and diagnosis of cancers that would otherwise have remained asymptomatic and not impacted an individual's health in their lifetime had they remained undetected. This is an unintended consequence of cancer screening programmes, which aim to detect cancers at an earlier often asymptomatic phase to improve survival. However, the consequence of erroneously detecting, diagnosing and treating these asymptomatic or slow growing (indolent) cancers are harmful, due to the psychosocial impact of a cancer diagnosis and the side effects of cancer treatment.

79. Overdiagnosis of breast cancer from mammography screening is estimated at less than 10% once corrected for lead time bias and baseline breast cancer risk (Puliti et al., 2012[107]) (Bulliard et al., 2021[108]). A United States study estimated 15% or 1 in 7 breast cancers represented overdiagnosis (Ryser et al., 2022[109]). Similarly, screening for lung cancer using low dose CT scanning is estimated to result in an overdiagnosis rate ranging from 4% to 19% (Esserman et al., 2014[106]) (Callister, Sasieni and Robbins, 2021[110]). The costs of overdiagnosis to the patient are the psychosocial stress of a cancer diagnosis and the healthcare costs of consultations for invasive biopsies, scans and cancer treatment. Overtreatment is an inevitable consequence of cancers over-diagnosed by screening, as it is virtually impossible to predict which cancers will evolve into invasive disease and death in a patient's lifetime.

80. Overdiagnosis represents an even larger problem for prostate cancer screening using prostate specific antigen (PSA) measurement alone, estimated at 42% to 57% of detected prostate cancers (Heijnsdijk et al., 2009[111]), and contributes to considerable differences in reported prostate cancer incidence between countries. In the late 1990s and early 2000s, the incidence of prostate cancer increased considerably due to increased practice of PSA screening, however prostate cancer mortality rates remained stable or even declined (Vaccarella et al., 2024[112]). This decoupling of screening, diagnosis and mortality demonstrated limited benefit of opportunistic prostate cancer screening using PSA alone and revealed significant overdiagnosis.

81. The consequence of prostate cancer overdiagnosis in younger men is harmful overtreatment of asymptomatic or indolent cancers with intensive cancer therapy. Overtreatment has considerable harmful effects on sexual and genitourinary function (Esserman et al., 2014[106]), for asymptomatic prostate cancer that would otherwise not have impacted the patient's risk of death. The costs of overdiagnosis from PSA screening are estimated to account for 40% of total healthcare costs associated with prostate screening (EUR 24 million to screen 100 000 men) in one modelling study (Heijnsdijk et al., 2009[111]). An estimated 76% of thyroid cancers diagnosed are due to overdiagnosis, which represented approximately 1.7 million cases diagnosed worldwide from 2013-2017 (Li et al., 2024[113]). This overdiagnosis results in costly overtreatment including thyroidectomy surgery and thyroid hormone replacement therapy (Novelli G, 2025[114]). In France, the direct healthcare costs driven by overdiagnosis of thyroid cancer from 2011 to 2015 were estimated at EUR 60-160 million (Li et al., 2023[115]).

Delayed cancer diagnosis confers poorer prognosis and higher treatment costs

82. While organised screening programmes exist for breast, cervical and colorectal cancer, other cancer types rely on early clinical or radiological diagnosis and prompt referral for treatment to improve survival. When this does not occur as a result of diagnostic error or delays in referral for medical investigations, cancer diagnosis can be delayed and the cancer allowed to progress to more advanced stage (Camidge et al., 2022[116]) (van der Veer et al., 2023[117]). This progression increases the risk of emergency presentation, which results in poorer outcomes for patients, including reduced cancer survival.

83. Cancer screening errors and delays in results account for a minority of the diagnostic error in cancer diagnosis. The majority of diagnostic error occurs in the initial presentation of cancer symptoms and referral for investigation. Delays in cancer diagnosis may result from diagnostic error at the initial clinical assessment, in the ordering and interpretation of diagnostic tests, or in the communication and referral of results. Diagnostic delay occurs frequently in primary care, where patients may present multiple times in the six month interval prior to being subsequently diagnosed with cancer (Christensen et al., 2012[118]), although not all of these primary care presentations may represent a missed opportunity for diagnosis (Lyratzopoulos, Vedsted and Singh, 2015[119]).

84. Delayed cancer diagnosis is harmful, as diagnosis at a more advanced stage of cancer results is associated with reduced cancer survival. Lung cancer is an example of where symptoms or radiological anomalies on imaging, if not detected and lung cancer not correctly diagnosed at an earlier stage (Ciello et al., 2017[120]), can considerably reduce survival (Groome et al., 2007[121]). Missed diagnosis of early lung cancer presentation often results in later presentation to the emergency department at a more advanced stage of cancer, and is associated with 80% higher mortality (Kapadia et al., 2024[17]). The healthcare utilisation and treatment costs were estimated to be three times lower for lung cancer diagnosed at stage I compared to stage IV, among insured patients in the United States in 2011 (Gildea et al., 2017[122]).

85. Delayed diagnosis can also be a frequent feature of colorectal cancer, despite the existence of cancer screening programmes in most OECD countries. Poor uptake of colorectal cancer screening, patients ignoring subtle symptoms, or doctors missing non-specific symptoms such as abdominal pain in younger patients, can result in later presentation to the emergency department with advanced cancer stage. At this point, the cancer may require emergency surgery and more extensive follow-up treatment with chemo- and/or radiotherapy, which are highly costly to the healthcare system. The delayed diagnosis at a later cancer stage confers a poorer prognosis and reduced cancer survival for the patient.

Delayed or miscommunicated cancer diagnosis is harmful

86. Diagnosis, besides being timely and accurate, requires the appropriate communication of positive results to inform timely intervention. Considerable negative consequences arise when the results of a cancer diagnosis are either delayed or miscommunicated to a patient or their provider. Patients may be “lost to follow-up” if their result is not communicated to them and assumed to be normal, or if a positive result is miscommunicated as a negative. A review of follow-up rates of abnormal laboratory and radiology test results in the outpatient setting reported failure to follow-up 7% of abnormal laboratory and 1% of radiology results where cancer was suspected (Callen et al., 2012[123]). The report highlighted the favourable impact of electronic health records in ensuring a safer and systematic diagnostic process.

87. Consequences of failure to follow up on abnormal test results include missed cancer diagnoses. Delayed communication of results can also delay timely referral of patients to diagnostic services or cancer care and reduce their survival. The impact of a patient learning a miscommunicated or delayed cancer diagnosis is difficult to quantify, but often results in considerable stress for those involved, as it may equate to a loss of survival chance. The healthcare costs of diagnostic error resulting from delayed or miscommunicated diagnosis are likely to be considerable given the progression to later cancer stage during the time elapsed, in addition to medicolegal costs (Gildea et al., 2017[122]).

88. Where cancer screening is not part of an organised programme, the communication and follow-up of positive screening results with confirmatory diagnostic investigations may not routinely occur. Delay or failure to follow-up cancer screening results reduces cost-effectiveness of cancer screening. A metanalysis estimated that between 2 to 11% of patients with positive colorectal cancer screening results refuse or neglect to engage with follow-up (Dalton, 2018[124]).

Better monitoring and tracking systems can improve timely diagnosis rates

89. Several interventions exist to improve the timeliness of cancer diagnosis (Graber et al., 2024[125]). These underscore the need for referral and tracking systems to improve the quality of diagnostic safety in cancer screening and are particularly relevant for opportunistic cancer screening through private providers outside of organised screening programmes. Primary care practice in the United States demonstrated higher rates of patient follow-up care of their abnormal cancer screening results among those who received electronic health record reminders and outreach communication (31%), compared to standard care (23%) (Atlas et al., 2023[126]).

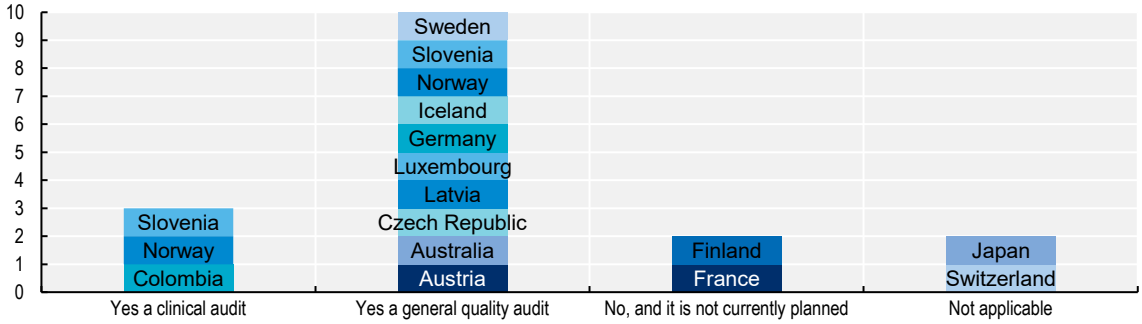
90. Ensuring that confirmatory testing of screening results and specialist consultations take place can reduce delays and complete the diagnostic process for cancer (Graber et al., 2024[125]). Improving patient communication and engagement and reducing the proportion of abnormal cancer screening results lost to follow-up should be prioritised to reduce diagnostic error from cancer screening.

Monitoring quality indicators for cancer screening in OECD countries

91. The EU Cancer Control Joint Action (CanCon) recommends that organised cancer screening programmes are monitored for quality assurance, to ensure their effectiveness at population level (Tit Albreht, 2017[127]). Part of this requires an audit of the screening activity, to benchmark the quality of the cancer screening against international or recognised standards (The International Agency for Research on Cancer, 2023[128]). This audit involves routine monitoring and reporting of quality indicators such as rates of screening coverage and participation rates, false negative results, recall rates for positive screens, diagnostic referral rates and interval cancer diagnoses (Csanádi et al., 2019[129]). Auditing can help to identify and monitor diagnostic error related to cancer screening, which may be minimised through the quality improvement cycle. Opportunistic cancer screening activity outside of organised programmes not subject to rigorous standards of auditing and quality assurance has a less favourable risk-benefit balance compared to organised screening programmes (Tit Albreht, 2017[127]).

92. Of the 16 OECD countries surveyed to inform this report, 11 reported having either a general or a clinical audit of a national cancer screening programme in place (**Austria, Australia, Colombia, Czechia, Latvia, Luxembourg, Germany, Iceland, Norway, Slovenia** and **Sweden**), though not necessarily for all programmes. Countries that lacked an audit of cancer screening reported either regional or opportunistic screening activity, rather than a national organised programme. Two countries surveyed (**France** and **Finland**) reported not having a formal audit process in place for cancer screening activity (see Figure 2.6).

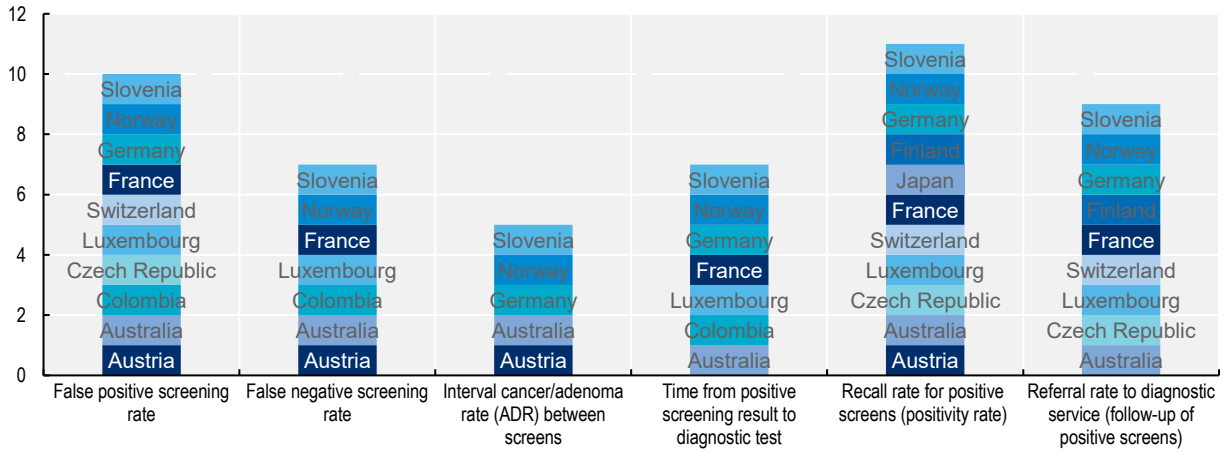
Figure 2.6. A majority of surveyed OECD countries audit National cancer screening programme(s)



Source: OECD Survey on the Assessment of the Adoption of Systems and Interventions to Improve Diagnostic Safety, 2024

93. Most of the 16 surveyed OECD countries monitor diagnostic accuracy in cancer screening through quality assurance indicators for false positive (overdiagnosis) results, but only 7 report indicators on false negative results (**Austria, Australia, Colombia, France, Luxembourg, Norway** and **Slovenia**) (see Figure 2.7). To ensure timely follow-up of screening results, the recall rate for positive screens is monitored and reported in 11 OECD countries, and the referral of patients to diagnostic services to confirm diagnosis is reported by nine surveyed countries. Timely diagnosis can be inferred by the rate of interval cancer development between subsequent cancer screens, which is reported by only five countries (**Austria, Australia, Germany, Norway** and **Slovenia**).

Figure 2.7. A majority of surveyed OECD countries monitor diagnostic error in cancer screening through quality assurance indicators



Source: OECD Survey on the Assessment of the Adoption of Systems and Interventions to Improve Diagnostic Safety, 2024

Case study: Long COVID

94. As a new and evolving disease entity, Long COVID represents a major diagnostic challenge for patients, healthcare professionals and healthcare systems, with considerable potential for diagnostic error. Lack of application of a standardised case definition, its multisystemic clinical presentation, and lack of specific biomarkers for Long COVID are challenges to detection and accurate diagnosis of this condition.

Misdiagnosis of Long COVID increases the medico-administrative burden for patients

95. A prerequisite for developing Long COVID is a diagnosis of prior COVID-19 infection, the confirmation and timing of which may be complicated to ascertain. Furthermore, asymptomatic or mildly symptomatic infection may not be diagnosed, delaying recognition of subsequent Long COVID symptoms. The estimated 409 million people living with Long COVID worldwide is therefore an underestimate based on symptomatic infections (Al-Aly et al., 2024[130]).

96. In the era of the COVID-19 pandemic, the scale of increased use of laboratory test in home settings was unprecedented. The sizable self-testing market share has been supported by the scaled-up use of rapid diagnostic tests and self-administered sampling for COVID-19 outside of the clinical setting. Although consumer testing kits have numerous benefits such as increased access, patient empowerment and lower costs, the documentation, communication and correct interpretation of their results for use in the formal diagnostic process can be lacking. Retrospective misdiagnosis of COVID-19 infection due to diagnostic error related to inappropriate diagnostics thus contributes to delayed diagnosis of Long COVID.

97. Comparing prevalence of Long COVID based on diagnosis by physicians to self-reported symptoms suggests delayed or missed diagnosis of this complex condition in health systems. In Ireland, self-reported prevalence is 7%, whereas physician diagnosis is 3% among the same adult population surveyed in 2024 (Government of Ireland, 2024[131]). In the UK, 3.3% of people self-reported Long COVID in 2023/2024 (Office for National Statistics (ONS), 2024[132]), whereas a retrospective analysis of 19 million primary care patient records from 2020 – 2023 identified only 20,000 patients with a diagnostic code related to Long COVID (Henderson et al., 2024[133]). Prevalence estimates based on patient-reported symptoms range from 6 to 7% of the general population (Al-Aly et al., 2024[130]), depending on the country surveyed and case definition used.

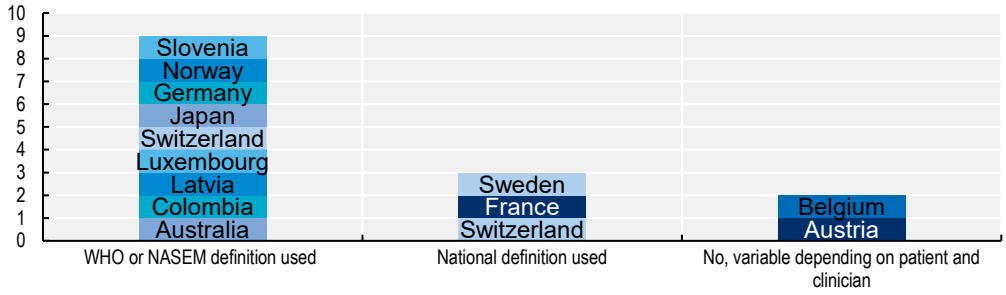
Variation in Long COVID symptoms is a challenge for consensus on case definition

98. No single standard case definition exists for diagnosing Long COVID internationally. Indeed, there appear to be several subtypes of Long COVID affecting different organ systems, rather than a single defined clinical syndrome (Gentilotti et al., 2023[134]) (Wang et al., 2025[135]). Nevertheless, the World Health Organization and the National Academy for Science and Engineering and Medicine have published two similar working definitions (World Health Organization, 2022[136]) (National Academies of Sciences and Medicine, 2024[137]). Due to variation in clinical presentation, severity and duration of symptoms, and lack of standardised criteria and diagnostic tools, Long COVID is especially challenging to diagnose (Espinosa Gonzalez and Suzuki, 2024[10]). The consequence for patients can be a delayed diagnosis of Long COVID or a misdiagnosis of another condition, with a negative impact on quality of life.

Lack of standardised diagnostic criteria and care pathways are a challenge for healthcare professionals

99. From a health systems perspective, clinical understanding and detection of Long COVID syndrome is not standardised among healthcare professionals, and its variable clinical presentation does not fit neatly into a single medical specialty. As a result, there is a lack of consensus on the case definition and standard diagnostic criteria that can be applied and universally adopted by medical professional associations to raise knowledge and awareness of Long COVID in the medical profession. Of the surveyed OECD countries, nine reported use either the WHO or NASEM case definitions (**Australia, Colombia, Latvia, Luxembourg, Switzerland, Japan, Germany, Norway** and **Slovenia**) while two (**Austria** and **Belgium**) reported lack of standardised use of these case definitions to diagnose Long COVID (see Figure 2.8**).**

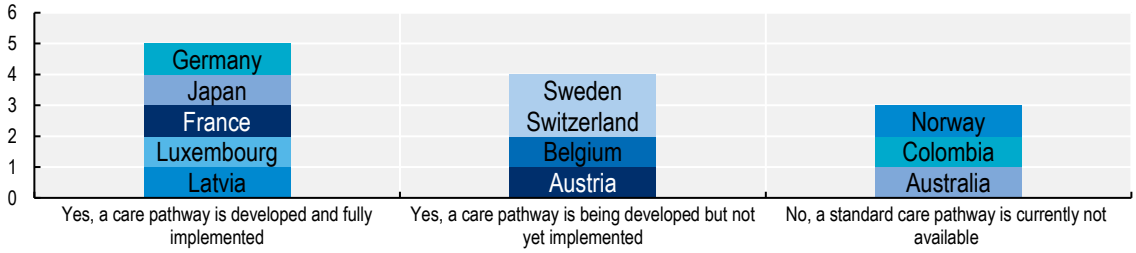
Figure 2.8. Half of OECD countries surveyed use WHO or NASEM case definition to diagnose Long COVID



Source: OECD Survey on the Assessment of the Adoption of Systems and Interventions to Improve Diagnostic Safety, 2024

100. Primary care physicians are often the first consulted by patents with Long COVID, may diagnose Long COVID differently accordingly to training and availability of national guidelines, and may lack clarity in referral options for more complex or severe cases. OECD countries report lacking a standardised care pathway for patients with suspected Long COVID, leading to further delay in diagnosis and management. Only five of the surveyed countries (**France, Germany, Japan, Latvia** and **Luxembourg**) reported having developed and fully implemented a standardised care pathway for Long COVID (see Figure 2.9).

Figure 2.9. A standard care pathway influences timely detection and diagnosis of Long COVID, and exists in four OECD countries surveyed



Source: OECD Survey on the Assessment of the Adoption of Systems and Interventions to Improve Diagnostic Safety, 2024

101. The prevalence of Long COVID among the primary care population is estimated at 7% based on the PaRIS survey of 17 OECD countries in 2023, and is set to remain high due to the cumulative effect of COVID-19 infections in the population. Conservative estimates forecast economic costs of Long COVID in the region of EUR 2 billion for the UK economy (Cambridge Econometrics, 2024[138]). Additionally, primary care patients with Long COVID report having to repeat information that should be in their medical records at a higher rate than patients without Long COVID (33% versus 25%).[1] An OECD and WHO Europe joint initiative supported by the European Commission aims to build consensus among countries on the key needs and priorities of patients, health systems, and societies for Long COVID and to develop good practices, policies and systems to effectively address these. A specific action is the adoption of a consensus case definition and updating clinician knowledge and training to improve detection and diagnosis of Long COVID.

[1] OECD, The impact of Long COVID in the primary care population of OECD countries. Findings from the OECD PaRIS survey, Forthcoming

3 The burden of diagnostic error

102. Estimating the burden of diagnostic error differs to estimating the burden of iatrogenic harm from curative care. The latter is a case of calculating the additional clinical management and care resulting from the adverse event, which typically requires additional tests and treatment. Diagnostic error is more remote to the outcomes desired by patients, providers and payers. It can precipitate a variety of consequences depending on the type of failure. The consequences can, in many cases, be both broader and deeper. This chapter seeks to estimate the burden of misdiagnosis, underdiagnosis and overdiagnosis. Estimates are based on the literature examining incidence and costs of diagnostic error in the context of previous OECD estimates on the burden of patient harm, which focused on harm during curative care.

The burden of misdiagnosis is underestimated

103. Misdiagnosis (wrong, delayed and missed diagnosis) incurs costs by taking the patient down an incorrect treatment path, administering therapies that are unnecessary and thus wasteful, that may in themselves predispose the patient to iatrogenic harm, and that will need to be eventually rectified once the correct diagnosis is (hopefully) made. In the likely event that the disease has progressed, the eventual correct treatment is likely to be of higher intensity (and therefore cost) than had the correct diagnosis been made in a timely way. When a correct diagnosis is delayed, simpler treatment – in many cases perhaps limited to secondary prevention – is supplanted with therapies that are more onerous, invasive and costly.

104. Misdiagnosis will affect most people in their lifetime. A commonly cited estimate is that approximately 10–15% of all rendered diagnoses may be incorrect (Graber, 2013[139]). Population based estimates suggest that diagnostic errors affect at least one in 20 US adults in the general population each year (Singh, Meyer and Thomas, 2014[140]). A study focusing on misdiagnosis of major diseases across clinical settings that included ambulatory clinics, emergency department and inpatients estimated that 2.59 million diagnostic errors occur in the United States each year, resulting in approximately 371 000 deaths and 424 000 permanently disabled due to misdiagnosis, making it the single largest source of serious harms from safety failures (Newman-Toker et al., 2024[141]). It should be noted that these estimates rely on (1) the definition and rates of diagnostic error used in this study, (2) the healthcare settings in which they occur, and (3) how they are detected and reported.

105. Misdiagnoses can be lifelong. For example, up to 4 in 10 people living with hypertension —a common condition that is relatively 'easy' to detect—in the United States are unaware of their diagnosis (Fryar et al., 2024[142]; Kocher and Emanuel, 2022[143]). Research from Sweden found that more than 30% of examined autopsies revealed clinically significant undiagnosed diseases (Friberg et al., 2019[144]). A systematic review found that 28% of autopsies reported at least one misdiagnosis (Winters et al., 2012[145]).

106. The annual cost of type of diagnostic error in the United States has been estimated at USD 100 Billion, a figure that includes malpractice litigations costs.[2] A study of over 76,000 patients also in the United States found that misdiagnosis and delayed diagnosis of appendicitis adds USD $2,712 (23%) to

[2] https://www.ahrq.gov/sites/default/files/wysiwyg/patient-safety/reports/issue-briefs/dx-leadership.pdf

the cost of the index hospital admission, incurring additional costs of USD 21 Billion a year before the costs of follow-up treatment and management are considered (Kulasekere et al., 2024[38]). Delayed diagnosis occurs in only 2.7% of appendicitis cases, but one may assume that other common presentations exhibit similar rates of mis- or delayed diagnosis. While perhaps modest individually (if USD 21 Billion a year can be considered modest), these common conditions, when combined, constitute substantial unnecessary cost borne by health systems. For example, an estimated USD 1 billion could be saved in the United States through better diagnosis of patients presenting to emergency departments with dizziness (Newman-Toker, McDonald and Meltzer, 2013[11]).

Misdiagnosis can lead to unnecessary healthcare as well as negative public health consequences

107. The burden of misdiagnosis includes the overtreatment caused by a lack of information on a patient's exact pathology. This can have individual as well as public health repercussions. For example, a significant cause of inappropriate antimicrobial prescribing is inability to discriminate between bacterial and viral infection in the absence of a relatively simple and cheap microbiological analysis. Apart from the ever-present risk of patient harms such as adverse drug reaction, the broader problem with unnecessary use of antibiotics is antimicrobial resistance – a significant threat to global public health (Pew[146])

108. Up to 40% of antibiotics are prescribed unnecessarily for acute respiratory tract infection, are due to misdiagnosis of viral as bacterial infection, contributing to the development of antibiotic-resistant bacteria (Barlam et al., 2016[147]). The OECD estimates that by 2035, resistance to third-line antimicrobials could more than double in member countries compared to 2005, and highlighted the significant health and economic burdens posed by AMR, emphasising the need for testing and prudent antibiotic use (OECD, 2017[24]; OECD, 2023[99]). Antimicrobial-resistant microorganisms will lead to higher incidence of pneumonia and urinary tract infections, leading to lead to increased healthcare costs and mortality rates (Cecchini, Langer and Slawomirski, 2015[148]). Rapid diagnostic tools can help distinguish between bacterial and viral infections, ensuring that antibiotics are prescribed only when necessary. This benefits individual patients and helps curb the societal impact of antimicrobial resistance (OECD, 2017[24]; O'Neill, 2015[149]).

Overdiagnosis is common and burdensome

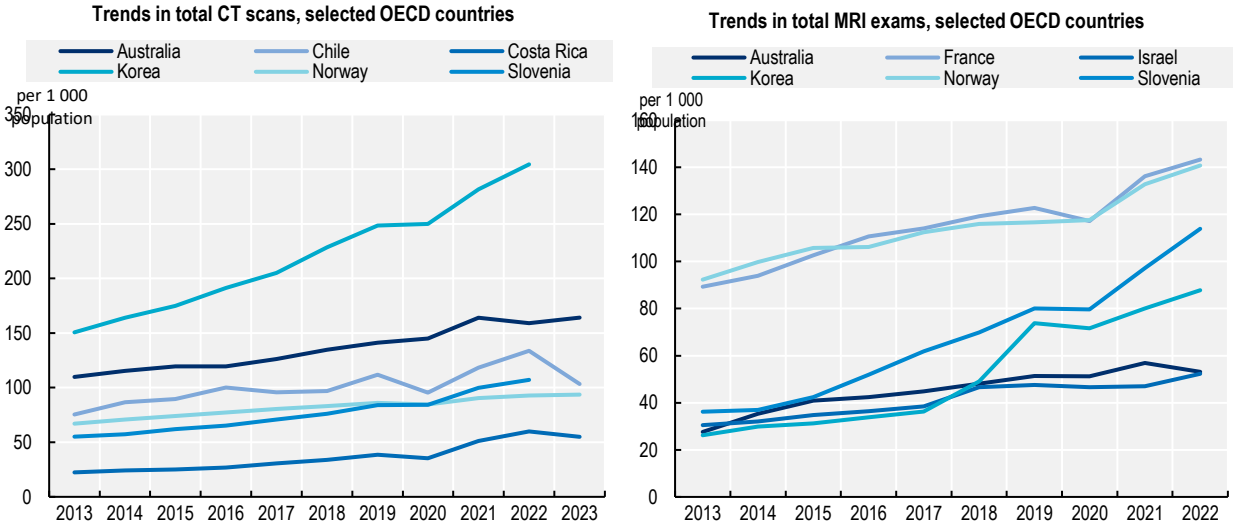
109. Overdiagnosis is costly and potentially harmful (most investigations are invasive). It can also set in train unnecessary medical interventions, that are costly and potentially also harmful. For example, imaging of uncomplicated low back pain not only uses scarce healthcare resources, it can generate an incidental diagnosis that is unrelated to the patient's symptoms yet create a cascade of referral and inappropriate (sometimes invasive) interventions that come with a high cost but little benefit (Sajid, Parkunan and Frost, 2021[150]).

110. Over-testing and overdiagnosis are significant concerns. A 2020 systematic review found "substantial overuse of diagnostic testing is present with wide variation in overuse", with a median proportion of over-testing of 30%. Preoperative testing and imaging for non-specific low back pain were the most frequently identified low-value diagnostics (Müskens et al., 2022[151]).

111. The number of diagnostic tests and screening is growing. For example, large increases in CT and MRI exams per 1 000 population are evident in several countries from 2013 to 2023 (see Figure 3.1). The number of CT exams more than doubled in Costa Rica, Korea and Slovenia, and the number of MRI exams increased by more than 50% in Australia, France, Israel, Korea, Norway, and Slovenia. The use of these diagnostic technologies dropped across many OECD countries in 2020 due to delayed or cancelled

diagnostic exams early in the COVID-19 pandemic. From 2021-2023 however, diagnostic exams increased and were typically above 2019 levels (OECD, 2023[152]).

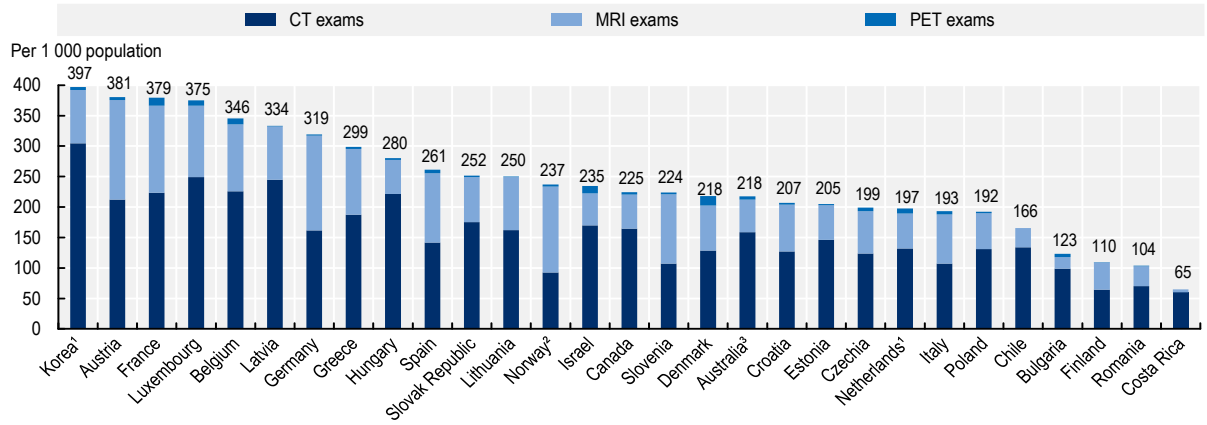
Figure 3.1. Trends in CT and MRI Scans, selected countries, 2011-23



Source: OECD Health Statistics, 2023

112. Substantial variation in the rate of investigations is observed. Across the OECD, the combined use of CT, MRI and positron emission tomography (PET) diagnostic scanners is highest in Korea, Austria, France and Luxembourg (see Figure 3.2). Variations persist also at the sub-national level. In Belgium, for example, recent analysis showed a 50% variation in use of diagnostic exams of the spine across provinces in 2017, and this variation was even larger across smaller areas (Devos et al., 2019[153]). In Australia, the age and sex-adjusted rate of heart perfusion scans varies 50--fold across geographic regions (ACSQHC, 2018[154]).

Figure 3.2. CT, MRI and PET exams, 2022 (or nearest year)



Notes. 1. Data excludes privately funded exams. 2. Data includes only exams outside hospital. 3. Data excludes exams on public patients (services that do not attract a Medicare benefit,).
Source: OECD Health Statistics 2024, https://www.oecd.org/en/data/datasets/oecd-health-statistics.html.

113. In addition to imaging, laboratory diagnostics are often also overused. A Netherlands study showed that hospitalised patients had, on average, 5.7 laboratory orders done during the first week of admission[3] and repeat testing of normal test results occurred in up to 85% of patients (Vrijsen et al., 2020[155]). While the costs of laboratory diagnostics are relatively small (less than 5% of hospital spending), they can have significant impacts, as laboratory results influence the majority (between 60 and 70%) of downstream medical decisions (Shaik et al., 2024[156]). Furthermore, excessive use of laboratory tests can lead to poor outcomes, such as resulting in hospital-induced anaemia, low patient satisfaction (excessive needle pricks, anxiety), and incidental and over diagnosis. Moreover, the proliferation of self-testing technologies (e.g. wearable devices) and greater access to tests will likely increase overdiagnosis and overtreatment. The likelihood of diagnosing a condition, however minor or incidental, rises as more tests are ordered (McCoy et al., 2015[157]). Many diagnostic tests are invasive or expose patients to potential harm such as radiation, allergic reactions. Adverse events for more invasive diagnostic procedures, such as liver biopsies, though infrequent, include infection, bleeding, hospitalisation, and even death (Boyum et al., 2016[158]; Thomaides-Brears et al., 2022[159]).

114. Like misdiagnosis, it is common, well-known conditions that are prone to overdiagnosis and consequent overtreatment. The National Institute for Health and Clinical Excellence (NICE) found that up to a third of asthma sufferers in England (4.1M adults) may have been wrongly diagnosed with the condition, proceeding to issue a guideline to improve diagnosing, monitoring and managing asthma in adults and children two years later.[4] Over 70 000 women a year may be over diagnosed with breast cancer in the United States (4000 a year in Australia) (Bleyer and Welch, 2012[160]) and 8000 Australian men a year are over-diagnosed with prostate cancer (Newman-Toker, 2014[21]). False positive mammograms and breast cancer overdiagnoses, for example, have been estimated to exceed USD four billion annually in the United States (Ong and Mandl, 2015[161]).

115. Globally, 500 000 people a year are over-diagnosed with thyroid cancer (Vaccarella et al., 2016[162]; Lortet-Tieulent et al., 2018[163]). Between 1999 and 2008, Korea saw an explosion in thyroid cancer diagnoses after testing was added to the national fee-for-service payment schedule. The associated with a rapid rise in surgical procedures. Yet over 90% of these cancers were small and they were detected mainly through increased screening rates (Ahn, Kim and Welch, 2014[164]).

116. Other examples of systematic overdiagnosis include:

- ***Melanoma***: an Australian study found that over 70% of melanoma in situ is over diagnosed, with direct annual healthcare costs of approximately AUD18 million (Lindsay et al., 2024[165]).
- ***Prostate cancer***: prostate specific antigen (PSA) testing carries a 60% risk of overdiagnosis (Welch and Black, 2010[166])
- ***Pulmonary embolism***: CT angiography has been shown to have an aggregate yield rate of 3% or less, and can lead to detection of small emboli that may not require treatment (Kline et al., 2020[167]; Prasad, Rho and Cifu, 2012[168])
- ***Attention deficit hyperactivity disorder in children***: Evidence suggests that boys born at the end of the school year (i.e. younger boys) are 30% more likely to be diagnosed and 40% more likely to be medicated than those born at the beginning of the year (Morrow et al., 2012[169])
- ***Polycystic ovary syndrome***: expanded definition may be wrongly labelling women as having the disease (Copp et al., 2017[170])

[3] Guidelines advise performing laboratory testing no more than twice per week.
[4] https://www.wired-gov.net/wg/news.nsf/articles/NICE+consults+on+draft+guideline+to+improve+asthma+diagnosis+28012015152033?open

Healthcare practitioners are aware of overdiagnosis

117. A 2017 survey of over 2,000 physicians in the United States on their perspectives on unnecessary medical care reported that more than 20% of overall medical care was not needed. Moreover, respondents said that a quarter of diagnostic tests were unnecessary (Lyu et al., 2017[171]). If this figure is accurate, over 2.5% of the United States' healthcare expenditure can be said to be consumed by unnecessary diagnostic activity (if, as outlined previously, 11.3% is spent on diagnostics). Using current expenditure figures (USD 4.8 Trillion per annum) would put the total amount wasted in this manner at USD 135 Billion a year. This is before the costs of unnecessary treatment stemming from these tests are considered.

Not all overdiagnosis leads to overtreatment (and vice versa)

118. The 2017 OECD report 'Tackling Wasteful Spending on Health' estimated that 20% of healthcare resources are misallocated on activities and practices that do not contribute to health outcomes, and that a considerable portion of this waste stems from inappropriate and unnecessary diagnostics and therapies (OECD, 2017[24]). In the United States, where most of the literature on inefficient resource allocation is from, estimates of the cost incurred by unnecessary clinical activity exceed USD 200 Billion (Moynihan, Doust and Henry, 2012[172]; Hackbarth, 2012[173]; National Academies, 2013[174]), which amounts to approximately 8% of health expenditure.

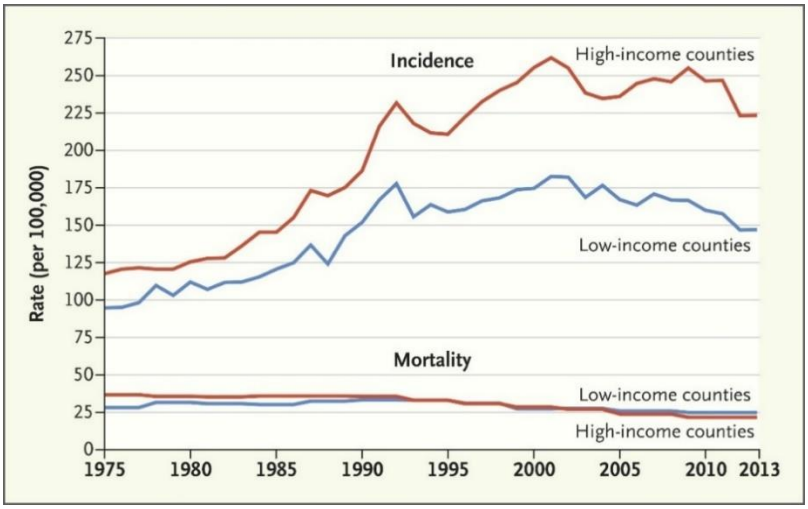
119. However, not all overtreatment is the result of overdiagnosis, and can occur despite a correct, timely diagnosis made with parsimonious use of diagnostic tests. A United States study of over 4,000 patients with osteoarthritis who underwent a total knee replacement found that clinically significant improvement in function and quality of life was limited to those with more severe symptoms, with most of the study participants reporting minimal or no improvement. This suggests that the procedure was performed on many patients inappropriately despite a correct diagnosis (Ferket et al., 2017[175]).

The inverse care law, in the case of overdiagnosis, can benefit the less affluent

120. Overdiagnosis is not limited to high-income countries. A 2022 scoping review of overdiagnosis and overuse of diagnostics and screening in low- and middle-income countries, found widespread overdiagnosis and overuse of tests, generating significant harm and waste. For example, direct costs associated with overdiagnosis of malaria was estimated at US$86 million in one country (Sudan) alone. Imaging, laboratory tests and procedures such as colonoscopy were the most common investigations. Reported drivers included the expanding disease definitions and lower diagnostic thresholds, echoing those of high-income countries (Albarqouni et al., 2022[176]). Overdiagnosis of thyroid cancer is likewise not limited to high-income countries (Lortet-Tieulent et al., 2018[163]).

121. Nevertheless, in many cases affluence is correlated with overdiagnosis. Welch and Fisher (2017) found that between 1975 and 2013 the incidence of four cancers (breast, prostate, thyroid, and melanoma) grew in high-income counties of the United States, compared with low-income counties, mortality rates have remained virtually unchanged (see Figure 3.3). This is likely to be predominantly due to overdiagnosis among the affluent (Welch and Fisher, 2017[177]).

Figure 3.3. Incidence and Mortality Trends for Breast Cancer, Prostate Cancer, Thyroid Cancer, and Melanoma in High- and Low-Income Counties in the United States, 1975–2013



Source: (Welch and Fisher, 2017[177])

122. As already discussed, incidence of several other cancers has also been rising in high-income countries but even a slight corresponding increase in cancer mortality has not been observed (Moynihan, Doust and Henry, 2012[172]). Modern cancer therapies, while expensive, are getting more effective at treating cancer, but most extend patients' lives and not cure them (Bach, 2015[178]). While this is slowly changing with the advent of new-generation interventions such as CAR T-cell[5] or genetic therapy these are relatively recent additions to the therapeutic arsenal, making it unlikely that the unchanging mortality rates in these studies can be put down to better medical management and intervention (Liu et al., 2024[179]).

Underdiagnosis generates unnecessary burden

123. Underdiagnosis occurs when a diagnosis is ignored and therefore missed. It is distinct from misdiagnosis and diagnostic delay in that it encompasses systematic and structural tendencies to neglect appropriate and requisite diagnostics for certain diseases and population groups (Newman-Toker, 2014[21]; Newman-Toker, 2009[22]). Underdiagnosis is not only inherently unfair but is burdensome for the patient as well as on the healthcare system as the chronically underdiagnosed will continue to seek increasingly costly medical help as their condition deteriorates. Underdiagnosis incurs unnecessary costs in the long term because an individual's worsening condition will cause them to seek care to the point where they require hospital admission that is more costly than timely treatment would have been. Underdiagnosis and undertreatment also causes harm by omission and unmet medical for individuals and across populations. The worst-case scenario is when undiagnosed conditions progress to a point at which they are incurable (Camillo, 2023[180]).

124. The prevalence and costs of underdiagnosis are difficult to estimate (Graber, 2013[139]). Globally up to 70% of persons with chronic obstructive pulmonary disease (COPD) or asthma do not receive a formal diagnosis of the condition, potentially impacting outcomes and quality of life, and leading to greater healthcare utilisation and poorer work productivity (Aaron et al., 2024[181]). Asthma is a particular challenge as it is both frequently over and under-diagnosed. Findings from the United Kingdom show that asthma overdiagnosis and underdiagnosis among children were potentially as high as 15% and 40% respectively

[5] https://www.cancer.gov/about-cancer/treatment/research/car-t-cells

(Lo et al., 2018[182]).Women with cardiovascular diseases tend to be underdiagnosed (Keteepe-Arachi, 2017[183]; Lebrun and Bond, 2018[184]).

125. Underdiagnosis occurs in all healthcare settings.[6] For example, general practitioners may underdiagnose chronic obstructive pulmonary disease (Ancochea et al., 2013[185]). Stigmatized diseases such as sexually transmitted infections and substance abuse disorders are underdiagnosed (Aho et al., 2022[186]). Minority population can be overrepresented in underdiagnosis, such as First Nations children in Canada and Indigenous children Australia (Lindblom, 2014[187]; Coleman et al., 2018[188]).

Underdiagnosis can reflect overdiagnosis along a socioeconomic gradient

126. Under- and overdiagnosis can often represent two sides of the same coin. An Australian study found that coronary angiography rates were most strongly associated with private medical insurance status and not, as one might expect, with underlying rates of cardiovascular disease (Chew et al., 2016[189]). This suggests some populations are more likely to unnecessarily receive this invasive test (which should, in most cases result in a percutaneous coronary intervention (PCI)) while their counterparts without private insurance may be less likely to receive care they would benefit from, potentially resulting in lower quality of life and premature death. Shifting this medical activity to those populations with unmet need may or may not deliver savings, but it is highly likely to improve outcomes and therefore value.

127. For example, evidence exists on the effect of undiagnosed COVID-19 cases contributing to the transmission of the SARS-CoV-2 virus, exacerbating the pandemic's already substantial health and economic burdens. A study using Indian data found that undetected cases contributed to the transmission risk of the disease with a reproduction number for undetected cases of (R3) of 0.323, suggesting that better identification could have mitigate the negative health and economic impact of epidemic in that country (Saha and Saha, 2021[190]). A Singaporean study evaluating expanded screening criteria found that higher rates of case identification would produce savings of USD 2.34 million during the initial stages of the pandemic (Lim et al., 2020[191]).

128. In addition, minorities were more likely to be undiagnosed with the virus, once again underscoring the racial and socio-economic gradient in diagnostic error. For example, during the first six months of the pandemic in the United States, approximately 16.8 million SARS-CoV-2 infections went undiagnosed, compared to 3.5 million diagnosed infections – a 4.8-fold difference. The highest estimates were in African American Hispanic participants, and residents of urban centres (Kalish et al., 2021[192]).

The direct costs of diagnostic error may approach a fifth of healthcare expenditure

129. The assertion that “diagnostic errors represent the ’bottom of the iceberg’ of patient safety – a hidden, yet large, source of morbidity and mortality” rings true (Liberman and Newman-Toker, 2018[193]). In terms of **misdiagnosis**, the most recent and convincing assessment of its health burden suggests unequivocally that “diagnostic error is probably the single largest source of deaths across all care settings linked to medical error [and] my exceed estimated deaths from all other patient safety concerns combined, regardless of which prior estimate of total deaths due to medical error is considered” (Newman-Toker et al., 2023[2]). (Noting the caveats listed in a previous section of this report).

130. Previous OECD analyses, these have estimated these direct costs of iatrogenic harm across the three main healthcare settings – acute, primary/ambulatory and long-term care – to be over 12% of healthcare expenditure (Slawomirski and Klazinga, 2022[58]). However, the previous estimate needs to be

[6] https://www.longdom.org/open-access/health-care-systems-primary-secondary-tertiary-and-quaternary-care-97476.html

revised down because the primary care component included misdiagnosis. To avoid double counting the costs of misdiagnosis, the revised estimate for primary care should be 1.7% -- down from 3.3% (the cost of harm in this setting is easier to decompose than the costs of wrong/delayed diagnosis into its setting-based components). The revised total for the direct costs of patient harm during curative care combined is therefore 11% of health expenditure (acute care: 5.4%; primary care: 1.7%; long-term care: 3.9%). The direct costs of **misdiagnosis** therefore amount to approximately 11% of healthcare expenditure.

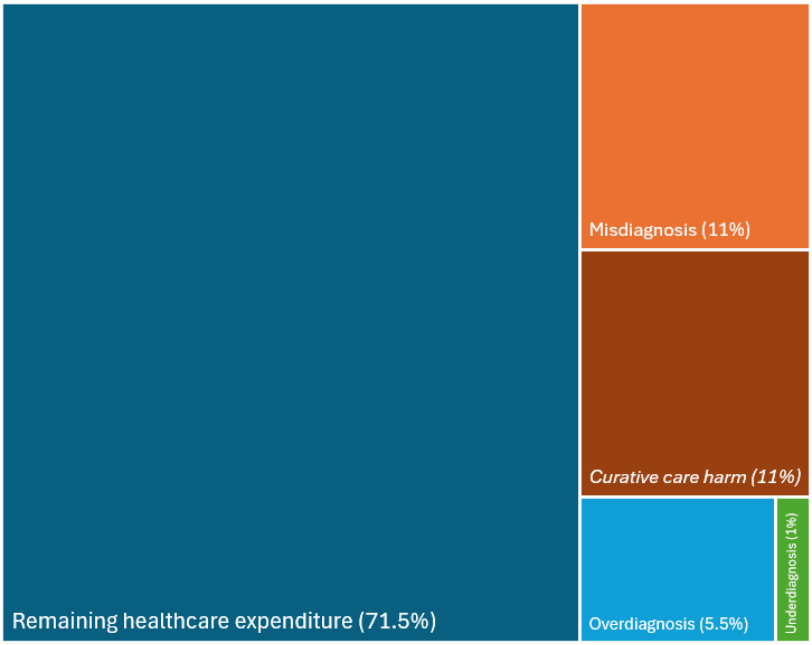
131. Estimating the costs of **overdiagnosis** must first consider the overuse of diagnostics, and then the cost of any ensuing unnecessary care. The latter is attenuated by three factors: that not all overdiagnosis results in overtreatment, not all overtreatment is caused by overdiagnosis, and likely variation between countries due to differences in how health services are structured, funded and regulated. In a typical OECD country. The evidence outlined above suggests that 25-30% of investigations are unnecessary, which would put the direct costs at approximately 2.5% of health expenditure, while overtreatment resulting from overdiagnosis can be estimated at 3% of expenditure. The combined direct costs of overdiagnosis and consequent overtreatment therefore to amount to approximately 5.5% of healthcare expenditure – a figure likely to be greater in the United States.

132. The direct costs of **underdiagnosis** are the most difficult to estimate. A conservative estimate, based on the evidence of its prevalence examined above, would be 1% of health expenditure. As with overdiagnosis, the factors driving underdiagnosis are linked to the way the healthcare services are structures, delivered and funded. Most of the economic costs of underdiagnosis may be indirect. The previous chapter outlined that the costs of underdiagnosing mental health disorders in Indiana incurred an estimated burden of $4.2 billion per annum (just over 1% of Indiana's GDP), comprising primarily of $3.3 billion in indirect costs from unemployment and lower productivity (Taylor et al., 2023[65]).

133. This brings the estimated total direct cost of diagnostic error to **17.5% of health expenditure.** Close to a fifth of what OECD countries spend of healthcare is consumed by harm stemming from diagnostic error, most of which – it must be again noted – is preventable. In the United States this would currently amount to about US$870 Billion a year. In a typical OECD country that allocates 9-10% of its GDP to healthcare, it would account for about 1.8% of GDP.

134. The combined figure for the direct finical burden of safety failures in diagnostic and curative care is therefore 28.5% of health expenditure (Figure 3.4), which amounts to trillions globally each year. In terms of (in)efficiency, the situation can be compared to an engine running only on three of its four cylinders—and it is difficult to imagine any other industry continuing to do business despite causing this amount of harm and waste. Unaddressed, the burden of diagnostic error is likely to grow principally due to the arrival of evermore sophisticated diagnostic technologies as well as the scaling up of personal digital diagnostic tools (wearables) that will become cheaper and more accessible with time.

Figure 3.4. The proportion of healthcare expenditure consumed by managing patient harm from diagnostic error and harm during curative care



Source. Authors

Indirect costs encompass health and environmental sustainability

135. Like harm stemming from curative care, harm from diagnostic error incurs costs beyond the health system. Firstly, the excess mortality and morbidity generated by diagnostic error carry a negative impact on the broader economy, the latter due to reduced productivity of those affected (principally patients and their informal carers), who may need to rely on social supports such as unemployment benefits and disability support payments. This is compounded by the reduced taxation take and other, less tangible contributions, to society. Given the lack of solid evidence, an estimate of the indirect costs of diagnostic error is difficult. However, these costs are likely comparable to the 'brake' on economic growth caused by harm during curative care, estimated at 0.7% of GDP per year (Slawomirski and Klazinga, 2022[58]).

136. The costs of diagnostic error extend to an additional dimension: public health. In addition to unnecessary prescribing of antibiotics contributing to antimicrobial resistance (discussed above), diagnostic error concerning infectious disease will not only affect the individual but also others whom the individual may consequently infect because they fail to take precautionary action due to the false negative diagnosis. COVID-19 serves as an example. While there is plenty of emerging literature on the health and economic burden of the virus, little is known about the effect of missed or delayed COVID-19 diagnosis particularly examining the infection of other individuals.

137. With healthcare contributing about 4.6% of global greenhouse gas emissions (Romanello et al., 2023[194]), all types of iatrogenic harm, including diagnostic error, impact the environmental sustainability of healthcare. Most obviously, overdiagnosis, which sets in train a cascade of further tests and treatment of questionable necessity and value. Overdiagnosis generates carbon emissions without improving health, both alone and when followed by overtreatment – an unnecessary cost on the environment and humanity (Barratt and McGain, 2021[195]). Missed or delayed diagnosis, meanwhile, generates more healthcare activity than would have been needed if the condition was identified earlier, driving up net emissions.

4 Addressing the causes of diagnostic error

138. Despite the significant scope and impacts of diagnostic safety events, they have not often been a focus of organizations responsible for the quality and safety of care and have received less attention than other medical errors (HAS, 2024[196]). There are numerous levers that can improve diagnostic safety. Deficits in these areas can drive poor diagnostic outcomes, while improvements can influence better diagnostic performance. Of particular interest are potential trade-offs within health systems related to accessibility and accuracy. Mechanisms include those applicable at the macro level, i.e. health system design and governance—through to the clinical environment, and down to individual provider competencies. Each level plays a role in improving diagnostic safety.

139. This chapter focuses on the causes of diagnostic error, and then the strategies and interventions to address these and improve diagnostic safety – including measures being taken by countries to achieve this. The causes encompass factors ranging from lack of measurement and information to cognitive bias, to system fragmentation, misaligned incentives, medical culture lacking in the necessary attributes to ensure safe diagnosis every time, and a lack of communication and patient engagement. Interventions, meanwhile, can be boiled down to three priorities: better measurement and feedback of diagnostic performance, creating a policy and regulatory environment that promotes diagnostic safety, and reorienting medical practice and behaviours towards better diagnostic outcomes. All must be underpinned by continuous learning and improvement throughout entire healthcare systems.

140. The challenge springs from the unique nature of diagnosis itself, which – as defined in this report – requires equipoise across several considerations and careful calculation and trade-off between potential risks and benefits. Safe diagnosis can mean both more and less – more for patients lacking access and for neglected conditions, and less for those with no real objective need. To avoid excess testing and overdiagnosis while promoting necessary testing and reducing underdiagnosis, similar nuance will need to be applied to measurement. The parameters for diagnostic safety and its measurement will evolve with our understanding of pathology and with the introduction of new diagnostic technologies and therapies. Current advances in genetic testing of cancers, and machine learning in diagnostics means that this evolution will need to be agile and nuanced.

Diagnostic error has a range of causes

141. As with other types of harm, diagnostic error has many root causes that can be found at the micro, meso and macro level of health systems. This includes the education and preparation of providers for clinical practice. The conditions for the three main types of diagnostic error to occur are discussed below. The common theme across all three are cognitive biases, lack of information, misaligned professional and financial incentives, and the values and culture of medical practice.

Diagnostic error entails a behavioural dimension and shares underlying factors with other types of harm

142. Just as the devices and procedures themselves, the practitioners that administer and oversee them can be sources of diagnostic error. For one, healthcare professionals can be affected by cognitive biases and personality traits when making diagnoses. Factors such as overconfidence or low risk tolerance, as well as specific cognitive errors including framing, anchoring, availability, search satisficing and premature closure can impact the accuracy of diagnostic outcomes – with such biases associated with inaccuracies in diagnosis for up to 75% of studied scenarios (Saposnik et al., 2016[197]; Lee et al., 2013[198]). Research from the Netherlands concluded that over 80% of the identified diagnostic adverse events found using patient record review were preventable, with the main causes of diagnostic adverse events being human errors related to knowledge-based mistakes and problems with information transfer (Zwaan et al., 2010[199]).

143. Historically, diagnostic errors have been thought of as individual failures – resulting from the shortcomings of a practitioner who misjudge or lack the needed knowledge. While this is sometimes the case, diagnostic error are also the result of systemic shortcomings in the organisation and incentives for care delivery, resulting in unfavourable conditions for accurate or timely diagnostic care. In a survey of primary care doctors in England found that system-related factors, such as poor communication between primary and secondary care, was the most cited reason for the occurrence of diagnostic delays (Car et al., 2016[200]).

144. Deficient and poorly designed information technology can interfere with diagnostic safety. For example, the software configuration and interface of electronic health records can create conditions for important information to be missed (Singh et al., 2009[201]). However, such problems are often part of broader organisational factors related to workflow and the availability of computer terminals (Singh and Sittig, 2015[202]). The good news is that these causes of missed or delayed diagnosis can be tackled through better measurement, information sharing, training and education.

Underdiagnosis is principally about access to services but also implicit bias

145. The dominant factor in underdiagnosis is a lack of access to healthcare. Studies of the Oregon Medicaid lottery found that gaining access to medical coverage was associated with higher use of services and improved outcomes, especially for mental health conditions such as anxiety and depression (Oregon Health Study Group, 2012[203]) Participants who ‘won’ coverage tended to have better self-reported mental health, reduced prevalence of undiagnosed depression by almost eight percent and reduced untreated depression, as well as substantial improvement in the symptoms of depression (Baicker et al., 2018[204]).

146. Like missed or delayed diagnosis, underdiagnosis can also be the result of behavioural and cognitive factors, such as implicit bias on the part of the diagnostician (Gopal et al., 2021[205]). For example, a practitioner may disregard symptoms reported based on the patient’s ethnicity, or a lack of research on how a condition manifests in minority demographic group (Camillo, 2023[180]). For this reason, interventions to improve or recalibrate practitioners’ diagnostic cognition must also address these implicit biases, which may be uncomfortable to acknowledge let alone discuss.

Overdiagnosis is rooted in the culture of medicine but influenced by policy and regulation

147. The key drivers of overdiagnosis’ include advancing technology, which enables the detection of ever-smaller abnormalities that increases disease prevalence (Moynihan, Doust and Henry, 2012[172]). In turn, successful treatment of milder diseases, in turn, creates a “false feedback loop” that fuels further testing and treatment of increasingly questionable disease states (Black, 1998[206]).

148. Most overdiagnosis isn't caused by overzealous practitioners, demanding patients or incorrect interpretation of results, but by continually expanding definitions and parameters of what constitutes disease (and a commensurately narrowed definition of normal). Few pathologies are binary. Most are defined as existing past a certain point on a continuum. Where this threshold is placed exactly can be an arbitrary, subjective decision. In most cases it is shifted so that growing numbers of people fall into the disease category (Schwartz and Woloshin, 1999[207]). Other examples of expanded disease parameters include gestational diabetes (Cundy, 2012[208]), osteoporosis (Herndon et al., 2007[209]), cardiovascular disease (Kaplan and Ong, 2007[210]), chronic kidney disease (Winearls and Glassock, 2011[211]), and many mental health disorders (Stein et al., 2010[212]).

149. Changes to diagnostic criteria are made by professional societies and expert panels comprising practitioners in the relevant medical specialty. It is sometimes unclear whether these panels have financial ties to biomedical technology companies that benefit from a growing number of people diagnosed with the disease (Moynihan, 2011[213]; Lichtenfeld, 2011[214]). Further complicating matters are the powerful incentives to do more than what may be necessary on providers, on developers of diagnostic technologies, and on patients, who may not be made aware that the incremental benefit of diagnostic testing can diminish rapidly at the economic margin especially when weighed against the risks. Additionally, clinicians face pressures on ordering diagnostic testing from the work environment they practice in, depending on the prevailing organisational and broader culture. Clinicians may over-test and over-diagnose out of undue caution owing to a prevailing peer culture. A more evidence based and data-driven work environment may rationalise diagnostic testing to manage diagnostic error.

150. Any discussion of excessive intervention – be it diagnostic or curative – must mention the role of healthcare funding models that pay providers for single outputs. It is by now well established that this exacerbates the tendency to over test and overtreat, especially (as is the case in most OECD countries) a third party pays for all the tests and treatments ordered by the practitioner. Practitioners, in most cases, are paid more for intervening than conservatively managing a problem identified by various diagnostic procedures, with little benefit to patients, all and at (tax)payers' expense. For example, the 15-fold growth in thyroid cancer diagnosis from 1993 to 2011 in Korea (see Box 4.3) started when thyroid screening with ultrasonography was added to the fee-for-service schedule for paid for by the government (Ahn, Kim and Welch, 2014[164]; Ahn and Welch, 2015[215]).

151. Medical culture and remuneration can be intertwined. Since the changes in the fee schedule for thyroid cancer screening in South Korea, thyroid cancer screening and treatment became big business with hospitals expanding thyroid clinics, hiring surgeons and creating an industry of robot-assisted thyroid surgery. Even a Thyroid Association comprising of Korean endocrinologists and thyroid surgeons was established. Concerns expressed drew a strong negative reaction from this Association, which claimed that screening and treatment are a "basic human right" (Ahn and Welch, 2015[215]).

Digital technology can fuel diagnostic error

152. The proliferation of wearable diagnostic devices like blood pressure, heart rate and glucose monitors, or the Apple Watch Electrocardiogram are touted as intrinsically positive for human health (Ibrahim T, 2023[216]). However, these devices are sold directly to the public and marketed to the healthy with limited evidence-based oversight. By definition, they medicalise the healthy and promote overuse with limited benefits for most users (Shih et al., 2022[217]).

153. More fundamentally, their growing introduction and use is likely, over time, to contribute towards the growing medicalisation of everyday life, promote ever-expanding disease definitions and create a perception of diagnosis as something that is done routinely instead of when pathology is suspected. All of this is not only likely to fuel overdiagnosis and overtreatment in real clinical settings simply through the sheer volume of potential pathologies detected. More concerning is that general practitioners have been

found to be ill-prepared for handling the data produced by wearables (Haase and Brodersen, 2023[218]). A horde of "worried well" may thus muddy the waters and may end up contributing to misdiagnosis.

Genomic tests can improve precision of diagnosis and treatment, but their misinterpretation represents a challenge

154. Genomic tests enhance the accuracy of diagnosis by identifying genetic markers associated with specific diseases, potentially enabling earlier and more accurate identification. The expanded scope of use of genetic tests in health systems has the potential to be cost-saving by establishing more efficient and accurate diagnosis, fewer irrelevant investigations, and faster turnover times. In addition, genetic testing data can be used to optimise treatment and reduce complications, particularly for metabolic diseases (Ministère de la Santé, n.d.[219]). Despite potential gains, genetic tests can be costly—with average costs of genetic tests used for precision medicine totalling $2,291 USD in Western Europe and $1,471 USD in North America in 2022—though costs are predicted to decline in the coming years (Statista, 2023[220]).

155. In addition, the rapid evolution of genomic technologies and complexity in the interpretation of genetic information may create an environment where clinicians may not be adequately equipped to appropriately interpret diagnostic data (Helm, Ayers and Kean, 2018[221]). For example, the provision of genetic testing by non-genetic healthcare providers can lead to incorrect tests being ordered or the misinterpretation of genetic test results (Shaw et al., 2023[222]). Despite the potential for increased accuracy and precision for diagnosis and subsequent treatment, the complexity of genomic testing and data requires specific knowledge and intelligence to synthesise and meaningfully interpret, which may exceed current human clinician capacity, and is also subject to poor diagnostic outcomes.

156. Furthermore, communication of the implications of genetic test results to patients carries a risk of increased anxiety and failure to make the lifestyle modifications to minimise the disease risks compounded by genomic factors. Involvement of dedicated genetic counsellors can enhance accurate comprehension, manage patient expectations, and offer psychosocial support in the context of genetic testing.

157. Over 90% of available genetic tests are used for diagnosis (Halbisen and Lu, 2023[223]). OECD countries who indicated they are actively tracking the use of genetic tests include **Czechia**, **Colombia**, **Latvia**, **Switzerland**, **France**, **Japan** and **Germany**. In **Switzerland**, for example, Genetic laboratories must be authorised by the Federal Office of Public Health (FOPH) in accordance with the Swiss law on human genetic testing (SR 810.12), and authorised laboratories must report annually to the FOPH the number of tests carried out in their laboratory[7]. In **Czechia**, **Japan**, and **Germany** this data is available via administrative health insurance datasets.

Defensive medicine and patient pressure may have less influence on diagnosis than commonly thought

158. When asked why unnecessary diagnostics are ordered, 85% United States physicians surveyed in 2017 stated fear of being sued for malpractice as the reason being sued for malpractice (Lyu et al., 2017[171]). However, evidence gathered when tort reform was being considered in the United States suggests that the extent and costs of "defensive medicine" may be overstated (Mello et al., 2010[224]). Moreover, comparisons of geographies with varying risk of litigation have failed to detect variation in medical practice (Thomas, Ziller and Thayer, 2010[225]). While a significant proportion of ordered tests may have a defensive component, few of these are entirely driven by it. Low-value care ordered only because of fear of lawsuits is estimated to comprise less than 3% of overall costs (Rothberg et al., 2014[226]).

[7] see also https://www.bag.admin.ch/bag/fr/home/zahlen-und-statistiken/kennzahlen-genetische-tests.html

159. Patient pressure is often invoked as a reason for overdiagnosis. However, literature confirming this phenomenon is scarce and the most cited drivers of overdiagnosis are firmly on the supply side of healthcare (Müskens et al., 2021[23]). In fact, doubt exists over the veracity of the patient pressure argument. For example, evidence suggest that doctors can overestimate how much intervention their patients really want, and that doctors' belief regarding patient expectations may be inaccurate (Karras et al., 2003[227]; Mangione-Smith et al., 1999[228]).

Diagnostic performance must be measured more accurately and routinely

160. The relative lack of attention on diagnostic safety may be partly down to the fact that diagnostic error and its have, to date, been underestimated. Measurement is therefore among the first things that must be addressed to improve diagnostic safety. However, the complexity of diagnosis and diagnostic error —in addition to their and their cognitive and systemic origins—make them difficult to identify and measure (HAS, 2024[196]). Despite some existing estimates from the literature there is relatively little measurement of the scope and occurrence of diagnostic safety lapses at the national level in most countries. No countries indicated that they had conducted a study or clinical audit estimated the national incidence of diagnostic error for sepsis, stroke, heart attack, cancer screening, bipolar disorder, and autism, for example, except for **Germany** (for sepsis and mammography screening).

161. Voluntary reporting by practitioners has been tried with varying success (Singh and Sittig, 2015[202]). While self-reporting should be encouraged, the structural and cultural challenges associated with it make it unsuitable as a core component of routine measurement. For one, it uses the most precious resource of all – clinicians' time – and is challenging to sustain (Graber et al., 2014[229]). More importantly, practitioners are frequently unaware of their diagnostic errors (Schiff, 2008[230]). Nevertheless, practitioner reporting can add clinical context to diagnostic errors identified by other means.

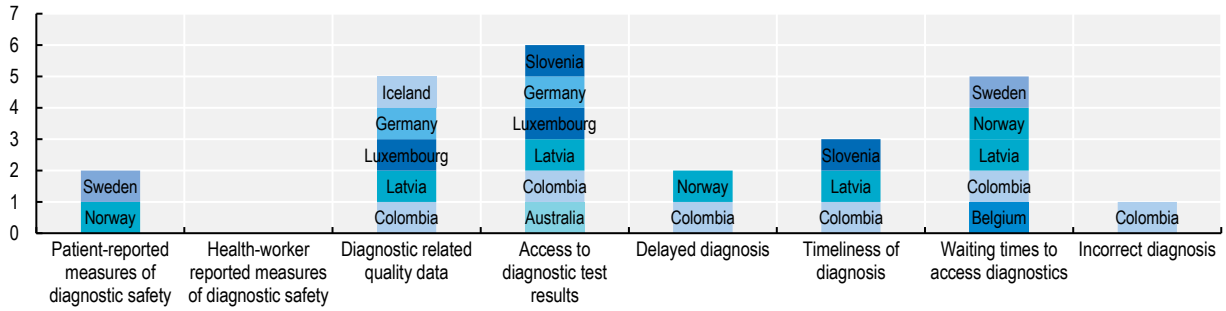
162. Other measurement approaches also have limitations. Autopsies, while highly detailed, are slow, costly and likely unrepresentative of the problem (only a small number of diagnostic errors lead to death). It is unreasonable to expect them to be performed routinely, or to rely on their results to be extrapolated. Similarly, analyses of malpractice claims would likely fail to capture the 'tail' of the distribution that may have a low impact individually, incur considerable aggregate costs due to higher (if under-detected) incidence. Root-cause analysis is labour-intensive and impossible to be conducted at scale without incurring large (opportunity) costs (Graber, 1999[231]). Unannounced standardised patients – while useful in that they have shone a light on the true extent of diagnostic error (Schwartz et al., 2012[53])—are costly and unlikely to ever form part of routine diagnostic safety surveillance for similar reasons.

163. Despite the lack of clinical audits and the assessment of national estimates, several countries have put in place specific programmes or reviews to monitor and reduce diagnostic error (see Figure 4.1). Measures for generating and reviewing diagnostic related quality data have been adopted nationally in five reporting countries (**Colombia, Iceland, Germany, Luxembourg,** and **Latvia**). In Colombia, this relates to reports on quality of diagnosis for rheumatoid arthritis, haemophilia, chronic kidney disease, HIV, Cancer and Hepatitis C. In **Latvia**, a number of policies are in place for mammography services—including requirements for double-blind reading mammography results—with systems that automatically recognizes and passes on results to a third radiologist in cases of discrepancies. In **Australia**, the Australian Dementia Network Registry evaluates the proportion of participants who had comprehensive assessments completed as part of the diagnostic process for dementia and mild cognitive impairment (MCI), including cognitive and functional assessments, core blood tests and structural neuroimaging[8]. In **Finland** and **Germany** the

[8] https://www.australiandementianetwork.org.au/initiatives/clinical-quality-registry/#anchor-3

implemented quality measures relate primarily to laboratory diagnostics (and also radiology diagnostics in **Finland**).

Figure 4.1. The scope of specific programmes or reviews to monitor and reduce diagnostic error at the national level



Source: OECD Survey on the Assessment of the Adoption of Systems and Interventions to Improve Diagnostic Safety, 2024

164. Measurement of diagnostic error and assessment of diagnostic performance should begin with defining a set of indicators that focus on structures, processes and outcomes of diagnosis (Singh, Graber and Hofer, 2019[232]). Implementation of any proposed diagnostic indicators on accuracy, timeliness and communication should be carefully balanced against overburdening pressured healthcare workers and services with more performance management.

Systematic uptake of policies to collect data from patients and health care workers is still lacking

165. Given the importance of monitoring and evaluating patient-reported experiences of safety to improve the quality of healthcare and promote people-centred care delivery, in recent years, an increasing number of OECD countries have developed and validated (population-based) surveys that include questions about experiences of safety (Kendir et al., 2023[233]). However, experiences of diagnostic safety are not always routinely measured. **Sweden** and **Norway** have developed programs at the national level capturing patient-reported measures of diagnostic safety—the Swedish one is assessed by Löf, the Swedish national patient injury Insurance company. Collection of patient-reported diagnostic safety is also collected via the *beschwerdemanagement* project in **Germany** and the reporting system for safety incidents in health care provider organizations and patient injury claims data in **Finland**. In **Latvia**, specific health care institutions have reporting/learning systems where patient safety incidents can be submitted. In specific clinical areas, **Austria** is developing patient-reported diagnostic safety measures related to breast cancer screening[9]. Finally, in **Switzerland**, the FQC, is developing a pilot project for cross-sector implementation of PROMs[10].

166. Data on patient experiences of safety should be monitored as part of broader data collections on patient safety including professionals' experiences with data from health-workers and patient safety events measured through administrative data sources (Kendir et al., 2023[233]). Healthcare worker-reported measures of diagnostic safety were not reported to have been implemented nationally in any of the countries responding to the 2024 survey, however, regional approaches were reported by **Finland** related to reporting systems for safety incidents in health care provider organizations. In **Switzerland**, health-

[9] https://www.sozialministerium.at/Themen/Gesundheit/Nicht-uebertragbare-Krankheiten/Krebs/Brustkrebs-Fr%C3%BCherkennungsprogramm.html

[10] https://www.bag.admin.ch/dam/bag/de/dokumente/kuv-leistungen/eqk/jahresbericht_2023.pdf.download.pdf/EQK_Jahresbericht_2

workers in hospitals and acute care settings typically report critical incidents via critical incident reporting systems (CIRS), however these systems are not widely established in other healthcare facilities like birthing centres and nursing homes[11]. In **Germany**, CIRS are mandatory but not specific to diagnostic safety.

Electronic health records are the backbone of measuring diagnostic performance

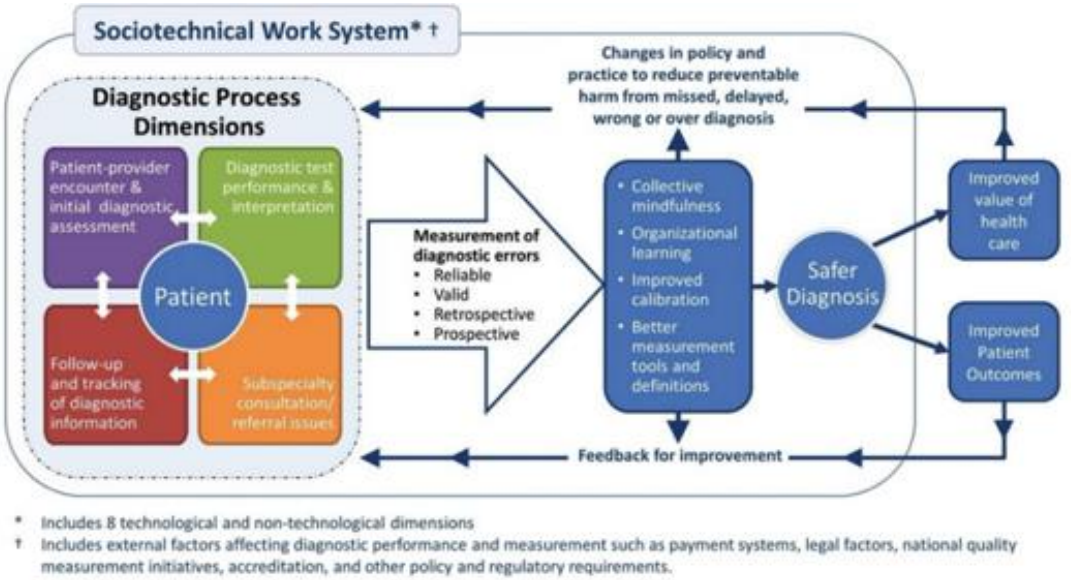
167. The widespread use of EHRs in medical practices is an important step towards new ways of supporting patient diagnosis and treatment as well as population health monitoring and research. To harness this potential, interoperability – the ability to link and share data across different systems – is critical to enable the sharing of information captured in single offices across providers. Identifying relevant data sources and developing a measurement infrastructure is a key step. If the goal is systematic and continuous surveillance of diagnostic safety, relevant data must be accessible. This can be challenging due to limitations of current data infrastructure and governance in many countries. These include the inability to track patients over time regardless of where they seek care, failure to recognise 'red flags' in patient presentation despite most care settings now using electronic medical records. Accurate assessment of diagnostic safety (and failure) must trawl through medical records, which – even with digital technology such as free text analysis and machine learning – can be expensive and prone to bias and error, further hampered by poor record documentation, poor reliability, hindsight bias and costly labour.

168. H0wever, resources to use medical records in this way are emerging. E-triggers alert personnel to potential failures enabling more targeted medical record review can be adapted to diagnostic safety (Murphy et al., 2013[234]; Singh et al., 2011[235]; Schiff, 2013[236]). Machine Learning shows promising results in its enhancing electronic detection of diagnostic errors in emergency department patients, while reducing the administrative burden of manual medical record review (Zimolzak et al., 2024[237]).

169. The Safer Dx Trigger Tools Framework can help identify and measure potential diagnostic errors, study contributing factors, and define goals to improve diagnostic safety using electronic health records (Murphy et al., 2018[238]). The Safer Dx Instrument (see Figure 4.2) is a screening tool that can help in detecting the presence or absence of a diagnostic error through review of medical records. The instruments developed for primary have been successfully adapted to acute care at an organisational level but require further validation on larger datasets. For example, detecting differences between the admission diagnosis and the discharge diagnosis in emergency departments (Malik et al., 2022[239]; Lam et al., 2022[54]).

[11] https://www.gdk-cds.ch/fileadmin/docs/public/gdk/themen/qualitaet/BT_Positionspapier_Q

Figure 4.2. Safer Dx Framework



Source: (Singh and Sittig, 2015[40])

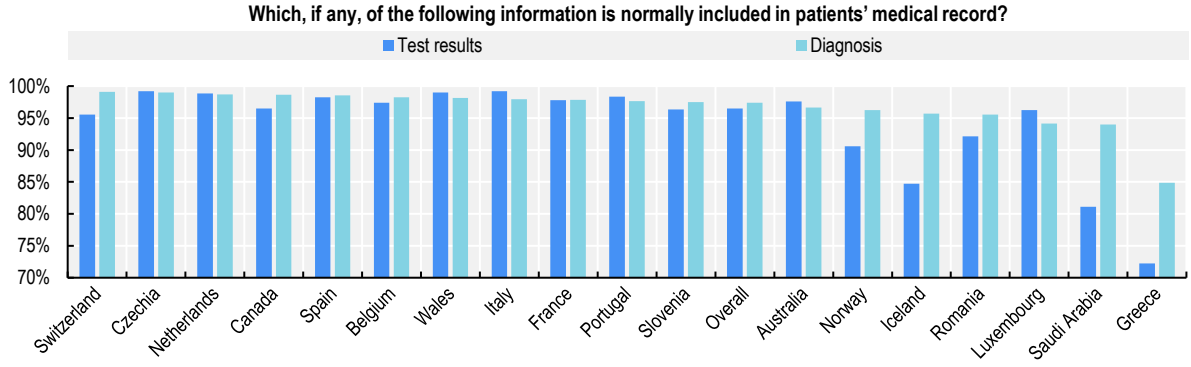
170. Expanding the current list of data elements and fields in electronic health records could enable better detection of diagnostic error and the granularity of information stemming from it. This would enable measuring the association of clinical history, diagnostic information and patient outcomes to inform better diagnosis in real time as well as contributing towards building the evidence base (Burstin and Cosby, 2022[12]).

171. Another concern with electronic medical records is their fragmentation across providers, settings and sectors (Slawomirski et al., 2023[240]). The diagnostic process is now rarely confined a single episode, practitioner and healthcare facility and setting. To accurately measure diagnostic safety, one must be able to follow the patient and their data along their entire healthcare journey. As it stands, this is difficult if not impossible in many OECD health systems. Diagnostic safety, perhaps more than any other aspect of quality, serves to illustrate the current shortcomings in countries health data infrastructure and the need to promote data integration and interoperability (OECD, 2022[241]; Oderkirk, 2021[242]).

172. In recent years, many countries have made considerable progress in adopting structured data elements to register key medical information—including patient diagnosis, medications, laboratory test results, medical imaging results and surgical procedures. While in 2016, only four out of 28 countries (**Denmark, Estonia, Japan** and the **United Kingdom (England))** reported using structured elements to capture and share data for each of these categories, in 2021 this number increased to 13 out of 28 countries (**Australia, Belgium, Costa Rica, Denmark, Estonia, Iceland, Israel, Italy, Japan, Mexico, Norway, Türkiye** and the **United States**). (Slawomirski et al., 2023[240]). In a study of primary care providers participating in the PaRIS survey, the documentation of diagnosis and test results in patient's medical records was found to be generally high—with an average of 96% participating practices in each country routinely coding this for test results and 97% coding this for diagnosis (see Figure 4.3).

Figure 4.3. Documentation of diagnosis and test results is routinely recorded in primary care

% of PaRIS survey primary care providers indicating the following



Note: *Data for Italy refer to patients enrolled in outpatient settings for specialist visits in selected regions.
Source: OECD PaRIS 2024 Database.

Routinely collected data must be harnessed

173. Detecting diagnostic error in routinely collected medical administrative coding data is currently difficult. While datasets typically contain information on principal diagnoses, comorbidities and interventions, no overt signals on whether the diagnoses are correct and timely, or whether the interventions are appropriate, exist within the classification and coding algorithms that underpin them. The timing and sequence of diagnoses and treatments can provide clues on their accuracy and appropriateness but, in many cases, exact times and sequences are not coded for an admission, while assessing several episodes typically requires data linkage (see below). In addition, administrative data only capture what is done and are therefore silent on missed diagnoses and underdiagnosis. While there have been studies using administrative data to identify diagnostic error in specific conditions, researchers in the field of diagnostic safety have generally dismissed their potential utility (Burstin and Cosby, 2022[12]; Singh and Sittig, 2015[202]; Burstin and Schneider, 2022[243]).

174. However, a few changes and additions could potentially make administrative data more useful in helping measure diagnostic safety, in the same manner that previous changes have made the detection of adverse events possible. For example, a field that flags if a diagnosis was incorrect or delayed. Given that these data are abstracted from medical records, this is not an unreasonable proposition. Algorithms could be developed to identify interventions that occurred due to misdiagnosis but would not have occurred had a correct diagnosis been made. These algorithms could also identify when a correct treatment did not occur due to the wrong or delayed diagnosis.

175. Disease registries can also be enlisted to measure and improve diagnostic safety. This may seem odd given that every record in a registry is there based on a specific (presumably correct) diagnosis. However, registries can contain granular information on severity and type of the illness (many conditions contain categories, subtypes and variations), symptoms, interventions and – most importantly – outcomes. These can be useful in measuring diagnostic quality within specific conditions of interest such as cancer or diabetes. Moreover, their granularity and detail could prove to be an excellent way to systematically examine overdiagnosis, overtreatment and optimal treatment pathways.

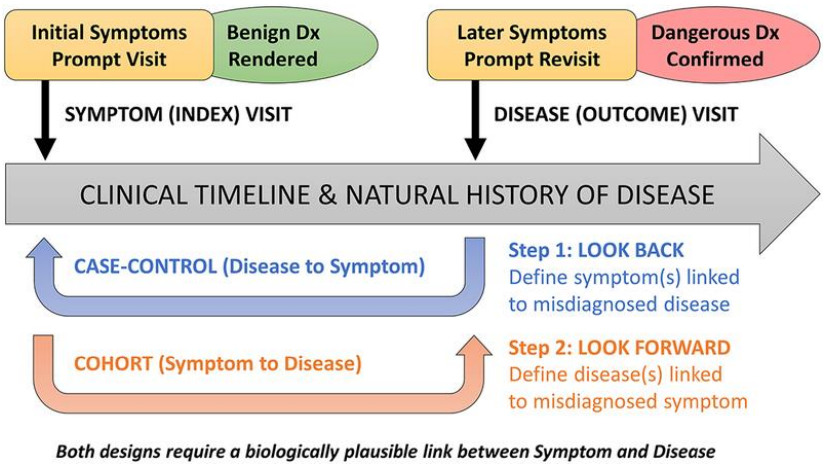
Routine data linkage can be a game changer

176. The data sources outlined above can each contribute useful information for measuring diagnostic safety. Combining them with each other and with other data sources (e.g. mortality data) through data

linkage increases the utility of information. Linking data at the individual record level is the only way, in most data landscapes, to gather a comprehensive picture people's health and healthcare journey -- theoretically from the 'cradle to the grave' – and therefore the most tractable way to systematically measure diagnostic safety. For example, linking people's hospital admission records, with their primary care data, any relevant registry data, as well as a regional or national death database that contains date and cause of death will provide a very comprehensive, chronological record of their diagnoses, treatments and outcomes over time. This can be interrogated for of diagnostic error.

177. For example, the Symptom-Disease Pair Analysis of Diagnostic Error (SPADE) model developed by Liberman et al (2018) measures misdiagnosis-related harm on the simple notion of change in diagnosis over time (see Figure 4.4). It combines "what is known about a condition's history and pathophysiology to develop an inferential model for identifying misdiagnosis-related harms based on time-linked markers of diagnostic delay that are clinically sensible, biologically plausible and specific to symptom-disease pairs" (Liberman and Newman-Toker, 2018[193]). SPADE requires a longitudinal information about a patient, comprising the first 'index' contact with the healthcare system and subsequent contacts, which is predicated on linking medical records and claims/administrative datasets. It enables both look-back (disease to symptom) and look-forward (symptom to disease) analyses (Liberman and Newman-Toker, 2018[193]; Liberman et al., 2023[244]).

Figure 4.4. Conceptual model for Symptom-Disease Pair Analysis of Diagnostic Error (SPADE)



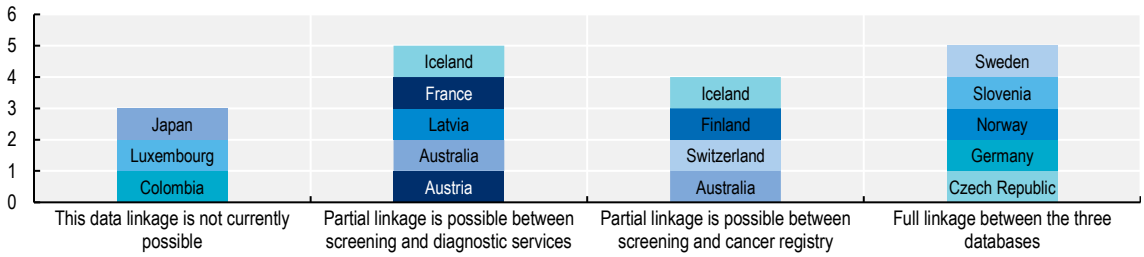
Source: (Liberman et al., 2023[244])

178. Studies using SPADE have examined data sets containing 20,000 service episodes to identify misdiagnosis-related harm (Liberman and Newman-Toker, 2018[193]) point to several advantages including potential automation (making it comparatively cheap to some of the methods outlined above) and its utility in providing an organisation- and system-level picture of misdiagnosis-related harm. These make SPADE a good candidate for routine measurement and surveillance of misdiagnosis, plus a proof of concept for detecting other types of diagnostic error. The underlying idea – linking datasets to provide a longitudinal picture of diagnosis, treatment and outcomes – can be adapted to target overdiagnosis and underdiagnosis. In fact, Australian researchers developed a model that uses clinical data to detect cases of sepsis overdiagnosis by predictive algorithms, that estimated a 4.3% rate of overdiagnosis (Fedyukova, Pires and Capurro, 2021[245]).

179. Effective cancer screening requires appropriate referral and follow-up of positive screening results to confirm diagnosis. Linkages between datasets for different stages of the diagnostic process such as screening records, diagnostic service results, and cancer registries allow for more comprehensive

checking of completeness, to ensure diagnoses are not missed by different providers in the healthcare system. This requires investment and interoperability of information systems. Five countries surveyed for this report—**Czechia, Germany, Norway, Slovenia,** and **Sweden**—indicate that linkage is possible between screening and diagnostic services or screening and cancer registries (see Figure 4.5).

Figure 4.5. Data linkage allows for better follow-up of cancer screening results to reduce error



Source: OECD Survey on the Assessment of the Adoption of Systems and Interventions to Improve Diagnostic Safety, 2024

Measurement framework should encompass all types of diagnostic error

180. The size and complexity of the task of measuring and improving diagnostic safety will require dedicated investment of financial and human resources, as well as political capital. Researchers suggest that initial targets should be conditions that are already known to be susceptible to the most harm from diagnostic error such as cardiovascular events, infectious disease and cancer, where the diagnostic process often breaks down, for example, transition across settings, pathology follow-up (Newman-Toker et al., 2019[246]; Burstin and Cosby, 2022[12]). Overdiagnosis must be part of the measurement framework. Choosing what diagnostics to target can be guided by evidence of variation in medical practice. Underdiagnosis and diagnostic equity should also feature (Burstin and Cosby, 2022[12]). The measurement approach can consequently be adjusted to target not only areas of greatest need but also those with a greater capacity to benefit.

Improving information infrastructure is a priority

181. An information infrastructure that enables measurement of diagnostic performance *and* the use of these data for improvement are key foundational requirements. The OECD has for more than a decade advocated for data governance frameworks that enable health data to be put to work while ensuring privacy through strict security technologies and requirements (OECD, 2016[247]). A good data infrastructure will require accurate recording of not only what is done but the outcomes of interest, including function and quality of life. Routine collection of PREMs and PROMs should become mandatory (OECD, 2025[248]). Such initiatives can have corollary effect such as promoting people-centeredness in provision and policy.

182. The OECD is currently piloting several indicators for potential international comparison on diagnostic safety in cancer related to stage at diagnosis and diagnosis following emergency presentation that are intended to capture the quality of diagnostic screening practices. International benchmarks that capture the quality and safety of diagnosis in other settings and conditions—for example priority areas such as mental health, sepsis, long-term care could provide a springboard to drive improvement in diagnostic outcomes.

The policy environment can set the scene for better diagnostic practice

183. Change is difficult without foundations at the system level. Legislation, regulations, governance and other policy settings set the parameters and context for behaviour, activities and culture across a health system. The impact of unsuitable policy settings can be overcome at the organisational and even clinical level, but this requires exceptional leadership and commitment. Better to set up the policy levers in a way that promote and incentivise practice to achieve overall objectives. The need for an overarching information infrastructure was discussed above. This section discusses other important policy levers to promote better diagnosis across healthcare settings.

Policies promoting timelier diagnosis are being introduced

184. Specific programmes or reviews to monitor and reduce diagnostic error related to delayed diagnosis at the national level are only implemented nationally in two countries: **Colombia** and **Norway. P**rograms or reviews related to timeliness of diagnosis are only implemented nationally in three: **Colombia, Latvia,** and **Slovenia**. In **Slovenia**, waiting times for all healthcare services, including diagnostic procedures, are monitored nationally. In **Colombia** reports have been published on the average diagnostic delays for pediatric leukemia and breast cancer. Work is currently being proposed on the 30-day colonoscopy following abnormal cervical smear. The **Australian** Institute for Health and Welfare tracks annual indicators on timeliness of dementia diagnosis (median wait time) as part of an annual dashboard aimed at ensuring Australia's 10-year national policy framework (the National Dementia Action Plan). **Latvia** introduced fast-track access for cancer patients (called the Green Corridor) in 2016, fully paid by state budgets, to streamline diagnosis and treatment decisions for suspected cancer cases. This requires specialist consultation and diagnostic examination within 10 working days of the date of referral. Fast-track access for recurrent cancer patients (called the Yellow Corridor) was also established to ensure timely access to care. In **Norway**, the national efforts on diagnostic delays focus on cancer and mental health. In **Finland** there are concerted efforts related to timeless for stroke care and for cancer diagnosis in **Sweden**.

185. Six of sixteen responding countries have independent standards for timeliness for certain diagnosis (**Australia, Colombia, France, Iceland, Slovenia,** and **Sweden**) and an another six note there are standards for timeliness are that are integrated into clinical guidelines (**Australia, Czech Republic, Latvia, Germany, Norway,** and **Sweden**). In **Slovenia**, there are "Regulations on ordering and managing waiting lists and maximum permissible waiting times" sets maximum permissible waiting times by level of urgency. In **Australia**, standards pertain to acute coronary syndrome, sepsis, colorectal cancer and dementia and in France for stroke, myocardial infarction, and other defined chronic illnesses[12]. In **Sweden** the standards refer to cancer pathways.

186. Data on waiting times for diagnostic services can provide insights on the timeliness of diagnosis and the following care. Almost all responding countries indicated the presence of specific programs or reviews to monitor and reduce diagnostic error related to waiting times to access diagnostics at the national (**Belgium, Colombia, Latvia, Norway, Sweden**) or regional level (**Finland, Norway**), or in selected health care settings (**Austria, Australia, Czech Republic, Luxembourg, Iceland, Norway**). In **Austria**, for example, there are quality measures related to the time between hospital admission and first imaging reported by the Stroke-unit registry. **Australia** reports the proportion of participants in the National Dementia Registry who had their initial appointment with a specialist diagnostic service within three months of referral and **Belgium** evaluates the waiting times for MRI tests. In **Finland**, data is also assessed related to delays in obtaining radiology and laboratory results. In Latvia, information is made available to patients

[12] https://www.has-sante.fr/upload/docs/application/pdf/2009-07/avc_prise_en_charge_precoce_-_recommandations.pdf
https://www.ameli.fr/seine-saint-denis/assure/sante/urgence/pathologies/avc

showing waiting times for medical institutions provide state-funded healthcare services—however it is self-reported by the institutions and only updated monthly[13].

Box 4.1. The Choosing Wisely Canada initiative to reduce unnecessary laboratory testing

The burden and scope of ordering unnecessary laboratory testing

In Canada, laboratory testing is the single highest volume medical activity costing the system approximately $5.9 billion annually[14]. It is estimated that 16-56% of lab tests ordered in Canada provide little to no clinical value. There are also risks to low-value lab testing, including false positive results, unnecessary follow-up referrals, overdiagnosis, and potential patient harm.

Using Labs Wisely led to a reduction in unnecessary laboratory serum folate testing of over 90%

In 2022, Choosing Wisely Canada launched Using Labs Wisely to help curb unnecessary lab testing. To date, the program has 160+ participating Canadian hospitals who agree to implement a lab stewardship quality improvement initiative, routinely attend Using Labs Wisely webinars to learn with lab experts from across the country and share aggregated data for comparative reports on overused lab tests (CK-MB, Urea, ESR, PT/PTT/INR, Folate, Urine Culture, Vitamin B12, and AST/ALT).

When Using Labs Wisely participating hospital, Queen Elizabeth II Hospital in Nova Scotia, received their Using Labs Wisely comparative report, they noticed they were in the top 25% of users of serum folate compared to their Using Labs Wisely counterparts. This encouraged them to reconfigure their lab information system's back-end, including updating lab ordering requirements and cancellation rules. QEII's initiative led to a 92% reduction in unnecessary serum folate testing at their site. Thoughtful change management and stakeholder engagement ensured this was done with little/no negative feedback from clinical requestors.

Using Labs Wisely participants also help to identify gaps and resources to help promote lab stewardship, such as the need for guidance on minimum re-testing intervals for certain lab tests. To address this need, Choosing Wisely Canada partnered with Canada's Drug Agency to conduct a consensus expert panel to develop recommendations[15] to support Using Labs Wisely sites and other labs across Canada.

Source: Box authored by Wendy Levinson. Choosing Wisely Canada

Clinical guidelines and protocols need to be updated routinely

187. Diagnosis and diagnostics must feature in clinical standards, guidelines and protocols, and should ideally include advice for practitioners on how to manage patient expectations of diagnostic tests. In **Australia**, *clinical care standards* address issues relating to diagnostic safety include standards for colonoscopy, stillbirth, sepsis, heavy menstrual bleeding. The latter involves potential overdiagnosis and unnecessary use of imaging, and explicitly advises to "reserve imaging for suspected serious pathology" and details the signs, symptoms and red flags that may signal rare but serious problems. Moreover, the Standard includes information for patients and advice for clinicians on how to communicate why imaging

[13] http://www.rindapiearsta.lv/lv/mekle_isako https://www.vmnvd.gov.lv/lv/rindapiearstalv
[14] https://www.cdhowe.org/publication/what-doctor-ordered-improving-use-and-value-laboratory-testing/
[15] https://www.cda-amc.ca/minimum-retesting-intervals-lab-tests

may not be necessary for them[16]. Policy makers, payers and regulators need to work with professional societies, patients and experts to ensure that clinical guidelines, standards and protocols (1) specifically address diagnostic safety and error, and (2) are updated whenever new evidence emerges, and/or new diagnostic technology enters the market. Diagnostic safety is a component of the *Be a voice for safety Program,* implemented in New South Wales[17].

188. Other countries use clinical standards are used as the mechanism to improve diagnostics safety. In **Austria**, quality standards are being used to advance diagnostic safety—particularly those related to patient blood management and integrated care of adult patients for preoperative diagnostics for elective procedures. In Luxembourg, accreditation and certification guides for laboratory and hospital settings include diagnostic safety improvement. **Luxembourg** and **Germany** have implemented policies to promote patient involvement and participation in their care pathway, with mandatory double reading of mammograms in breast cancer screening to reduce diagnostic errors.

Box 4.2. WHO initiatives on diagnostic safety and error

Diagnostic Safety in the WHO Global Safety Action Plan

In the World Health Organization (WHO) report "Global safety action plan 2021-2030 Towards eliminating avoidable harm in health care", strategic objective 3 "Ensuring safety in every clinical process" encourages governments to establish a program to improve patient safety, including diagnostic safety, particularly in primary care (WHO, 2023[249])

WHO Safer Primary Care: Diagnostic errors

In the WHO technical series on primary care, the focus on the scope of diagnostic error in the primary care setting outlines the common causes and proposed solutions to improve diagnostic safety. These include a whole of process approach from improving clinical education and training to improving health systems and information technology (World Health Organization, 2016[90]).

World Patient Safety Day 2024 - improving diagnosis for patient safety

The WHO 2024 focus is on the importance of correct and timely diagnosis for improving patients safety (World Health Organization, 2024[250]). The campaign promotes patients and healthcare providers working together to navigate the complex and lengthy diagnostic process. It revisits the steps in diagnosis including diagnostic interpretation and communication of results, includes tips to improve diagnostic accuracy for clinicians, and encourages discussion and review of diagnostic results to inform final diagnosis and treatment decision.

Discussion, consensus and review of initial diagnosis

189. Many diagnostic errors occur in the outpatient and primary care setting, there is considerable scope for improving patient safety here. The WHO technical series on *Safer Primary care: Diagnostic errors*, proposes solutions to improve reduce diagnostic error and improve safety, such as improving clinical education and training and improving health systems and information technology for reviewing patient data (World Health Organization, 2016[90]). The theme of World Patient Safety Day 2024 focuses on improving diagnosis, specifically diagnostic review and discussion between provider and patient (World

[16] https://www.safetyandquality.gov.au/standards/clinical-care-standards/low-back-pain-clinical-care-standard
[17] https://www.cec.health.nsw.gov.au/improve-quality/system-safety-culture/be-a-voice-for-safety/diagnostic-error

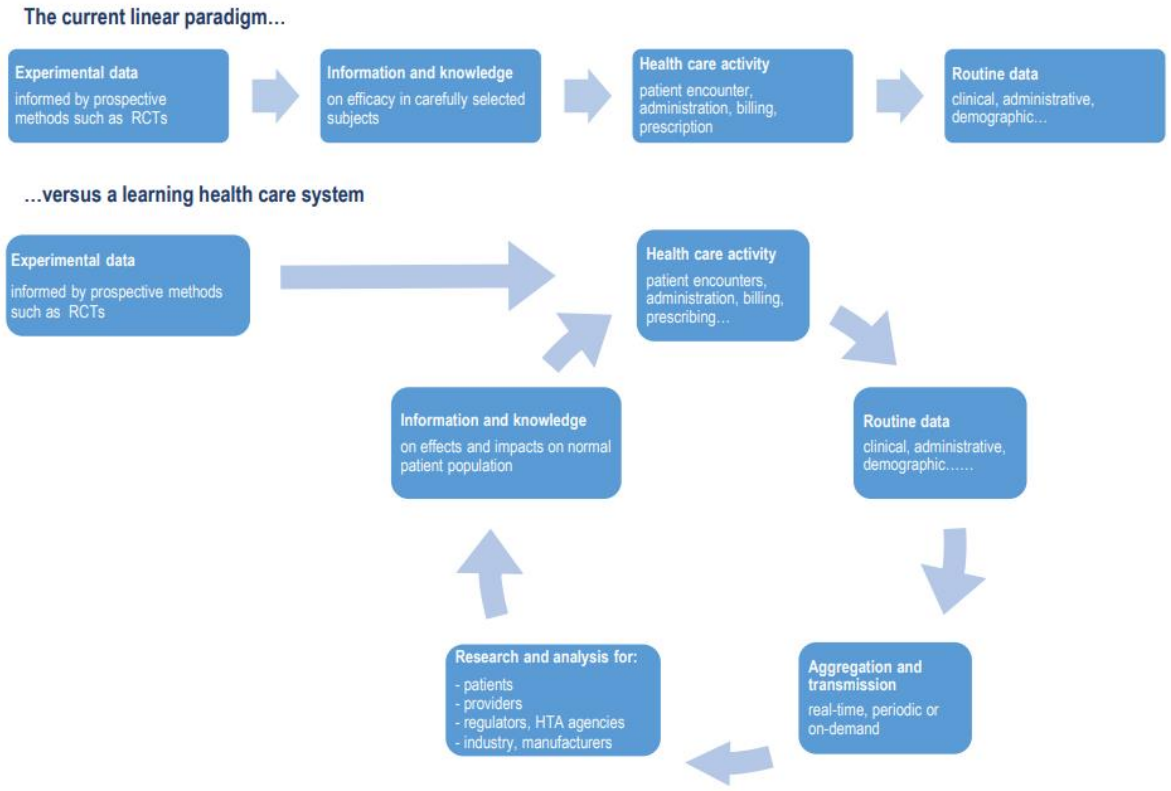
Health Organization, 2024[250]). The campaign includes steps to improve diagnostic accuracy and encourages discussion and review of diagnostic results to inform final diagnosis and treatment decision.

Medical device regulation

190. Medical devices – which include diagnostic tools and apparatus – have historically been subject to less robust market entry requirements than pharmaceuticals. Direct-to-consumer products such as wearable monitors face even lower regulatory requirements (OECD, 2017[7]). As outlined previously, some countries are taking steps to strengthen their regulatory frameworks. It is also important to improve post-market surveillance, harnessing the information infrastructure described above to enable routine, continuous monitoring of how technology performs in normal practice (Figure 4.6).

191. Moreover, regulation needs to adapt to new technology types, especially hybrid technologies that combine diagnosis and treatment and tracking devices marketed directly at consumers, by promoting co-ordination between entities that typically manage separately different types of technologies (OECD, 2017[7]; de Bienassis et al., 2022[251]).

Figure 4.6. The current linear approach versus the cycles of improvement where real world data complements experimental data



Source: (OECD, 2019[46])

Improving digital technology interfaces

192. Improving health information technology interfaces can support diagnostic decision-making. Systems should be designed and implemented to ensure that the most clinically relevant information (previous test results and diagnostic reasoning) are accessible and clearly contextualized (Agha, Skinner

and Chan, 2022[252]). A study of almost 250,000 emergency department visits suggested that some clinicians often overweigh salient symptoms, while neglecting to consider the full range of clinical information (Mullainathan and Obermeyer, 2021[253]). Technological supports, for example, machine learning could be more broadly incorporated to support the assessment of disease risk and diagnosis making processes (Mullainathan and Obermeyer, 2021[253]).

Rethinking healthcare funding models

193. Funding models influence practice. Addressing overdiagnosis and underdiagnosis especially will rely on financial incentives to be aligned with policy objectives (assuming that these objectives are to maximise population health with the available resources by optimising allocative and technical efficiency). Funding models must shift away from itemised funding such as fee-for-service towards bundled payments encompassing cycles of patient care and outcomes, shown to change practice patterns and reduce low-value care (Schwartz et al., 2015[254]).

194. Access to care must be based on need. This is not always the case even in countries that claim to provide 'universal access'. In the Australian example concerning inappropriate coronary angiography and unmet cardiovascular health need (Chew et al., 2016[189]), the problem can be ascribed to a two-tier healthcare system but more fundamentally a funding model that incentivises individual outputs -- and contributes to overdiagnosis in some and underdiagnosis in other populations – without regard for the entire diagnostic and care pathway, isn't designed to consider population health, and contains little external accountability for clinical decision making.

195. National policies to support more efficient diagnostic systems could be better adopted, including payment policies and publicly available quality metrics (Kocher and Emanuel, 2022[143]). In **Latvia** there are quality payments to general practitioners related to the stage of detection of cancer, as well as colon, prostate, cervical, and breast cancer screening coverage. Recent work from the US has also illustrated ways to promote payer and provider engagement in the design and adoption of diagnostic safety accountability mechanisms (Ali et al., 2023[255]).

National and international cooperation to drive best practice

196. A concerted national strategy must be informed by measurement (recognition of the problem) and commitment from national leaders as well as providers, patients and civil society. The evidence of the costs and burdens of diagnostic error is overwhelming and should galvanise any government that values the health of its population and the stewardship of its resources. Previous OECD Economics of Patient Safety reports have advocated for top-down leadership, and any existing national patient safety strategies should incorporate diagnostic safety. A separate coordinated initiative can be created, if this is what stakeholders prefer. The strategy would principally concern itself with creating the right policy settings for diagnostic safety to thrive and for creating support as well as accountability frameworks for healthcare system actors to work within.

197. Several countries, including **Australia**, **Czechia**, **France**, **Japan**, **Norway**, **Sweden**, and **Switzerland**, have identified diagnostic safety as a currently a priority with specific programmes or reviews to monitor and reduce diagnostic error. In **Belgium**, there is a specific project for diagnostic radiology, with the aim of increasing the appropriate use, ensuring that patients receive the most appropriate examination more quickly. This project, called "Prescription Search Support for Radiology" is planned for rollout in 2026. In 2024, the **Swiss** Patient Safety Foundation conducted a week of action on the topic of diagnostic safety, with a focus on good communication between all interfaces throughout the entire diagnostic process. In **France**, the National Authority for Health (HAS) issued a recent report on diagnostic errors in medicine, accompanied with recommendations for France to address to improve diagnostic safety across healthcare

system.[18] In **Japan**, incentives have been put in place that pay additional medical fees to medical institutions that take organisational measures—such as appointment of a person in charge of reviewing the report and evaluation of the status of report management and implementation of staff training—to prevent oversight of test result reports.[19]

International collaboration will also be required. For example, removing financial conflicts of interest in decisions about disease definitions and diagnostic thresholds (or at least making these conflicts more transparent to decision makers) can realistically only be achieved through by legislating requirements (self-regulation rarely works). Moreover, the international structure of many medical specialist societies make this a global concern. The World Medical Association is encouraged to be involved. As outlined above, additions to classification systems for diseases and interventions are required to improve measurement.

Safer diagnosis requires adaptability, rationalism … and courage

198. The Economics of Patient Safety series has, from the beginning, highlighted how risk-management practices in other high-risk industries should be adapted to health care– specifically 'Safety II' principles like proactive assessment and learning, adaptability in the face of uncertainty, and resilience (Slawomirski, Auraaen and Klazinga, 2017[256]; de Bienassis, Slawomirski and Klazinga, 2021[257]; Auraaen, Slawomirski and Klazinga, 2018[25]; de Bienassis, Llena-Nozal and Klazinga, 2020[258]). Diagnostic safety is no different. In fact, given the prominent role played by uncertainty and trade-offs between risks and benefits, diagnosis is perhaps more suited to this framework than other aspects of patient safety.

199. Greater recognition of diagnostic uncertainty is essential to fostering a more effective and resilient healthcare system. Diagnostic uncertainty, and low tolerance for risk among prescribers facing diagnostic uncertainty, have been identified as drivers of poor outcomes, including the overuse and misuse of antimicrobials (Özçelik, Chapman and Cecchini, 2024[259]). Penalizing systems and physicians for challenges in achieving a definitive or timely diagnosis may inadvertently contribute to over-testing and inefficiencies. Insufficient training for physicians in managing clinical uncertainty, coupled with the stigma associated with acknowledging the limits of medical science, exacerbates this issue. To address this, prioritizing education for both healthcare professionals and patients on strategies to navigate uncertainty is a critical area for policy focus.

200. The concept of diagnostic stewardship is defined as ordering the right tests for the right patient at the right time to provide information necessary to optimize clinical care (Agency for Healthcare Research and Quality, 2024[260]). It equally encompasses the reporting and interpretation of results at later stages of the diagnostic process, to guide optimal clinical management. Stewardship promotes discussion and feedback between clinicians on the choice of diagnostic testing, to discourage overtesting and reduce overdiagnosis. The Choosing Wisely initiatives (see Box 4.1) are a practical means of assessing diagnostic performance and instilling a culture of diagnostic safety and excellence at institutional or national level.

201. A certain amount of courage, tolerance for risk, and rationalism by individuals, professional groups and organisations is required. Acknowledgement and discussions of diagnostic performance may not always be comfortable, even if a safe, non-punitive environment is provided for such discussions, e.g. in multidisciplinary team meetings (see Teamwork and a conducive work culture are essential). One of the features of a learning health system is positive deviance, which essentially means trialling things different to the norm in a tightly managed fashion and recognising when this leads to better outcomes. This takes organisational / institutional courage. System-level policies should enable, indeed encourage, such experimentation on diagnosis on the proviso that it is rigorously measured, evaluated and the results fed

[18] https://www.has-sante.fr/upload/docs/application/pdf/2024-11/rapport_erreurs_diagnostiques_medecine.pdf
[19] https://www.mhlw.go.jp/content/12404000/001252053.pdf

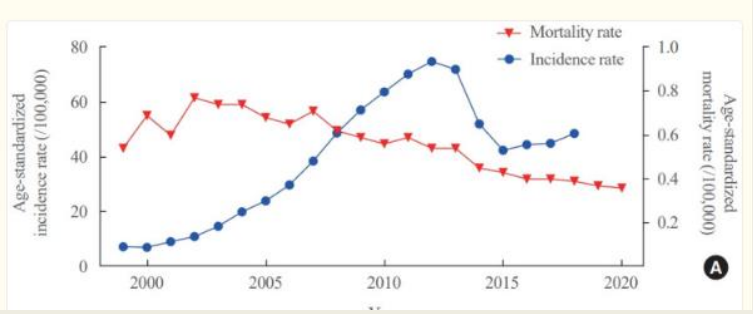
back to enable learning across the system (even if results are negative). Courage is also needed to call out and address poor diagnostic practices. A good illustration are the Korean physicians who formed a coalition to address egregious overdiagnosis of thyroid cancer in the 1990s and 2000s, resulting in changes to screening policy and practice (see Box 4.3).

Box 4.3. Change from within: Example of Overdiagnosis of Thyroid Cancer in Korea

The Korean example of thyroid cancer screening again serves as a useful example – this time in how improvement can be achieved at scale. After the problem of overdiagnosis and overtreatment was confirmed by data, eight Korean physicians -- Coalition for Prevention of Overdiagnosis of Thyroid Cancer -- wrote an open letter to the public highlighting the extraordinarily high thyroid cancer incidence and surgery rates and proposing that ultrasonography screening be discouraged. This was extensively reported by the media in investigative television and newspaper reports. The definition and diagnostic threshold for thyroid cancer, as well as treatment guidelines, were consequently revised (Yi et al., 2015[261]; Yi, 2016[262]).

After a decade of explosive growth, marked decrease in thyroid cancer diagnosis and surgery was observed. According to the authors of the study, the reduction in operations was "not primarily the result of more conservative surgical practice (e.g., opting for active surveillance instead); rather, it resulted from less screening — and less diagnosis" and that that the changes reflect patients' choices as opposed to physician recommendations (Ahn and Welch, 2015[215]). The Coalition was candid about the possible trade-off involved: that reduced diagnosis and surgeries may ultimately result in more deaths but considered this unlikely because mortality rates from thyroid cancer had remained unchanged, and that additional diagnoses from screening had a papillary histology, which is prevalent in the general population and considered a normal finding. They appear to have been correct. Thyroid cancer diagnosis decreased in Korea after 2014 but then rebounded after to become the most diagnosed cancer again in 2020 (Choi et al., 2023[263]). Most studies suggest that standardised thyroid cancer mortality was not affected (Choi et al., 2023[263]; The Community of Population-Based Regional Cancer Registries*, 2023[264]; Kim, 2024[265]). Case fatality rates increased but this would be expected if fewer low-level tumours are diagnosed (Kim et al., 2024[266]). In the context of overdiagnosis population-level mortality is the measure of interest.

Figure 4.7. Age-standardized incidence and mortality rates of thyroid cancer in Korea



This example highlights how, despite strong pushback from vested interests (as noted previously, the Korean Thyroid Association comprising endocrinologists and thyroid surgeons said that screening and treatment are "basic human rights")20 how policy and practice can change practice. The Coalition hopes that "this example will encourage other doctors to find their voice when medical trends run counter to their patients' interests" (Kim, 2024[265]).

Source: (Kim, 2024[265])

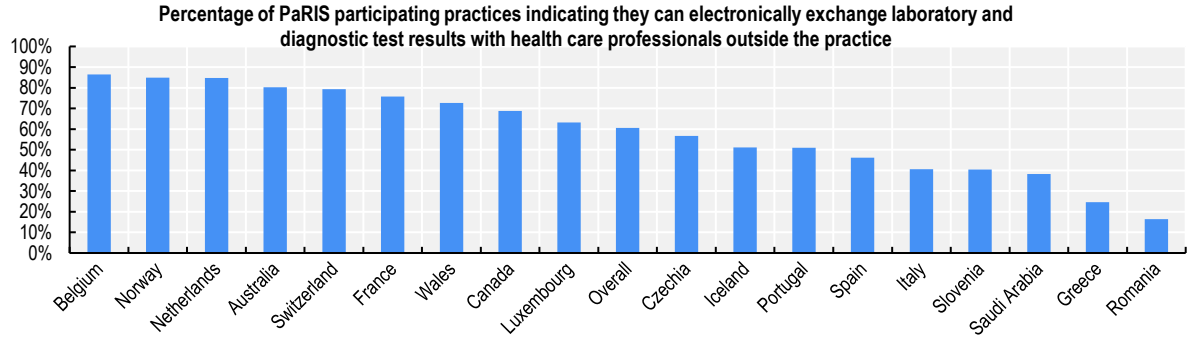
Tracking of test results and their follow-up can improve

202. Patient-reported data suggests that co-ordination problems between primary health care, specialists and hospitals are prevalent. And between 29% and 51% of people in 11 OECD countries reported having experienced problems of co-ordination between primary care and specialised care (OECD, 2020[267]). Co-ordination problems include medical tests not being available at the time of appointment or that duplicate tests being made; specialists do not have basic information from GPs (and vise-versa), or received conflicting information from different providers.

203. Policies to improve access to diagnostic test results have been implemented in six of 16 responding countries, including **Australia, Colombia, Latvia, Luxembourg, Germany,** and **Slovenia**. **Australia's** national My Health Record includes access to key health information, such as pathology and diagnostic imaging reports[21]. In **Colombia**, efforts to improve access to diagnostic test results are lead by Cuenta de Alto Costo, an entity of the health system that manages and provides comparable data and information to accelerate improvements in the care of people with high-cost diseases, evaluate the performance of the entities and compensate for deviations in the concentration of these risks in the insurance[22]. In **Latvia**, patient access to test results is available through laboratory websites, IT private companies (usually paid service) and freely through e-health system managed by the National Health Service—which includes laboratory and some radiology results as of 2024[23].

204. Poor uptake and use of digital tools for tracking and supporting communication of test results is one contributing factor. On average, only 61% of PaRIS participating primary care practices cited that they had capacity to electronically exchange laboratory and diagnostic tests with health care processionals outside the practice (see Figure 4.8).

Figure 4.8. Only 3 out of 5 of PaRIS participating primary care practices can electronically exchange laboratory and diagnostic tests outside of the practice



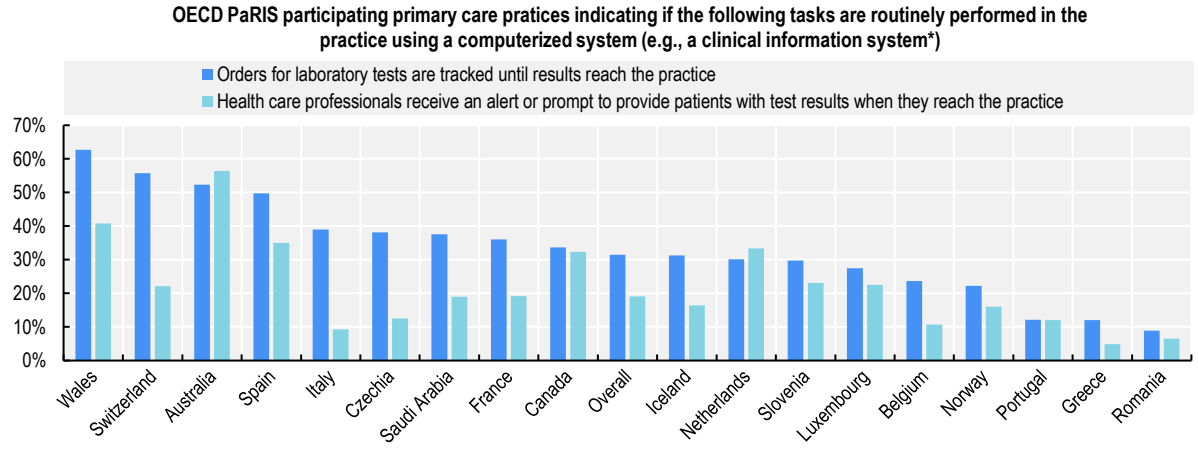
Note: *Data for Italy refer to patients enrolled in outpatient settings for specialist visits in selected regions.
Source: OECD PaRIS 2024 Database.

205. Among primary care practices that participated int the OECD's PaRIS Survey, only 31% of practices indicated that orders for laboratory tests were tracked until they reached the practice via a

[21] https://www.digitalhealth.gov.au/initiatives-and-programs/my-health-record/statistics
[22] https://cuentadealtocosto.org/
[23] https://www.maniveselibasdati.lv/

computerised system. Less than 20% of practices indicated that health care professionals received an alert or prompt to provide patients with test results when they reach their practice (see Figure 4.9)

Figure 4.9. Digital tools for tracking and communicating test results are uncommon in most countries



Note: *Data for Italy refer to patients enrolled in outpatient settings for specialist visits in selected regions. *Clinical information system (CIS) is a computer-based system that is designed for collecting, storing, using and making clinical information available to the health care delivery process.

206. Emergency department (ED) and hospital care often leads to investigations for which results are unavailable when the patient is discharged or leaves the facility. Estimates from the United States suggest that as many as 2 in 5 hospitalised adults have at least one test result pending at discharge (TPAD) (Shriner et al., 2021[268]). Access to pending test results following emergency care or hospitalization is important for several reasons. Firstly, some pending tests might point to urgent diagnoses that require immediate action (e.g., infections, haemorrhage) or follow up care (e.g. abnormal lab values). Ensuring patients are informed helps avoid diagnostic delays in addressing potential health problems.

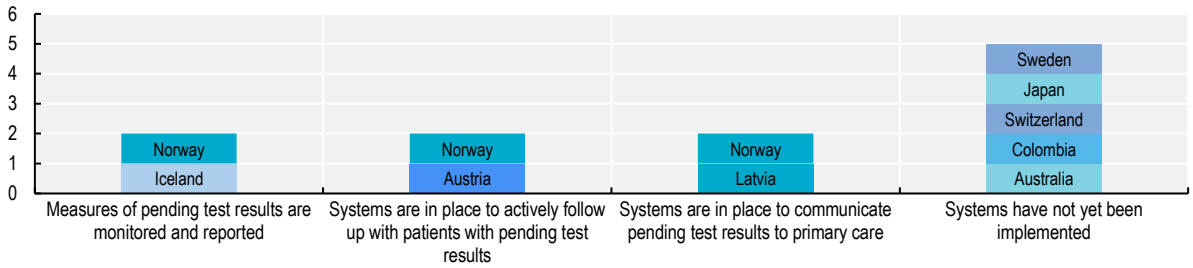
207. Follow-up of incidental findings detected by laboratory testing or radiological investigations in ED setting is a huge constraint on clinician's time and healthcare resources, for abnormalities that would have otherwise gone undetected until a later date, if at all. CT scan for investigating suspected pulmonary embolism in ED in Canada resulted in incidental findings in 13% of cases, with 88% of these clinically non-significant (Anjum, Bleeker and Ohle, 2019[269]). The overdiagnosis risk from extensive investigation, or diagnostic error from failure to follow-up and communicate these findings, cautions for judicious use of medical investigations and electronic systems to facilitate follow-up of results.

208. The percentage of pending test results that are followed up on vary widely but are frequently too high. Studies found that the rate of follow up on pending test rests in emergency settings ranged from an appalling 1% to 75%. For inpatients the follow-up rate varied from 20% to 69% (Whitehead et al., 2018[270]). In one study, over 40% of surveyed medical inpatients had a TPAD—over 40% of which were abnormal and almost 10% required action (Roy et al., 2005[271]). A Netherlands study found that delayed follow-up of (abnormal) hospital test results was the leading contributing factor in diagnostic error (Hooftman et al., 2023[272])

209. Follow up rates can be improved through the implementation of relatively simple quality assurance processes. For example, various studies looking at the impact of nurse/clerical staff follow up of pending results found that the interventions were able to increase successful follow up from below 10% to almost 60% (Mikhaeil et al., 2020[273]). Currently a number of countries are implementing systems for addressing

pending test results, including the reporting of measures related to pending test results, systems that actively follow up with patients, and systems that communicate pending results to primary care providers (see Figure 4.10).

Figure 4.10. Relatively few countries have systems in place for addressing pending test results



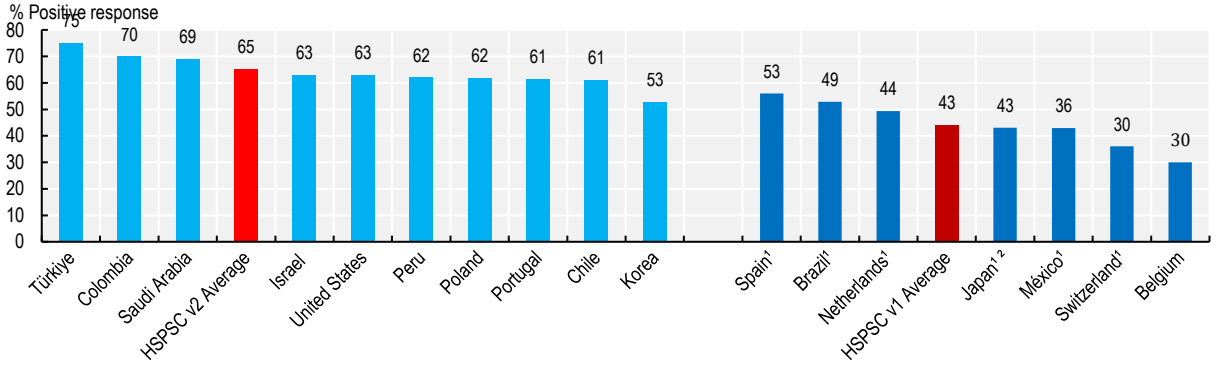
Source: OECD Survey on the Assessment of the Adoption of Systems and Interventions to Improve Diagnostic Safety, 2024

Teamwork and a conducive work culture are essential

210. Teamwork is a determinant of quality and safety of care, and reaching diagnosis is team effort, part of WHO World Patient Safety Day 2024 campaign (World Health Organization, 2024[250]). Diagnosis should be viewed as a process that leverages the strengths of other team members and the benefits of having different perspectives on diagnostic dilemmas. Best practices for diagnostic decision making should rely on strong communication and integration of clinical opinions to ensure that diagnostic decisions reflect the insight of the full clinical team (Agha, Skinner and Chan, 2022[252]). Good communication and multidisciplinary exchanges with peers ensures that required information is available and used in the diagnostic process, to improve diagnostic accuracy (World Health Organization, 2024[250]). Effective patient referral pathways and handovers – components of patient safety culture – are essential for ensuring this.

211. For several years, the OECD has been collecting data from hospital workers as part of the Hospital Survey of Patient Safety Culture (HSPSC). The survey asks clinical staff to evaluate the perceived quality of **handoffs and information exchange**—if important patient care information is transferred across hospital units and during shift changes (de Bienassis and Klazinga, 2024[274]). Results suggest tha the quality of care transitions could be improved. Across all countries, the average positive response was 65% for countries using the 2nd version (HSPSC v2) of the tool and 44% of countries using the 1st version (HSPSC v1) (see Figure 4.11).

Figure 4.11. Hospital Workers Perceptions of the Safety of Handoffs and Information Exchange



Note: Data from 2019-2024. 1. HSPSC v1 2. Data from previous PSC pilot data collection.
Source: OECD Pilot Data Collections on Patient Safety Culture (de Bienassis and Klazinga, 2024[274])

212. The data collection on patient safety culture discussed in the previous section likewise collects data on hospital workers perceptions of teamwork using the Hospital Survey of Patient Safety Culture (HSPSC). The teamwork domain aims to capture workers perceptions of if staff work together as an effective team, help each other during busy times, and are respectful. On average 76% of staff have a positive perception of teamwork at their workplace for countries using HSPSC v2 (eight OECD countries) and 69% among countries using HSPSC v1 (six OECD countries) (de Bienassis and Klazinga, 2024[274])

213. The ability to work well in a team is not an inherent human capability or attribute, and not all health practitioners will be natural team players. Training and education must instil the necessary skills to work in a team environment. The role of selection, education and ongoing professional development in diagnostic safety is discussed in a later section.

Active patient involvement and communication is an important factor in diagnostic safety

214. Patients are a key member of their healthcare team and play an important role in the diagnostic process. Providers need to be able to explain complex health diagnoses and treatment approaches to patients in a user-friendly, easily understood manner. In addition, engaged patients play a role in detecting and preventing diagnostic safety errors–for example by identifying inaccurate documentation or information gaps.

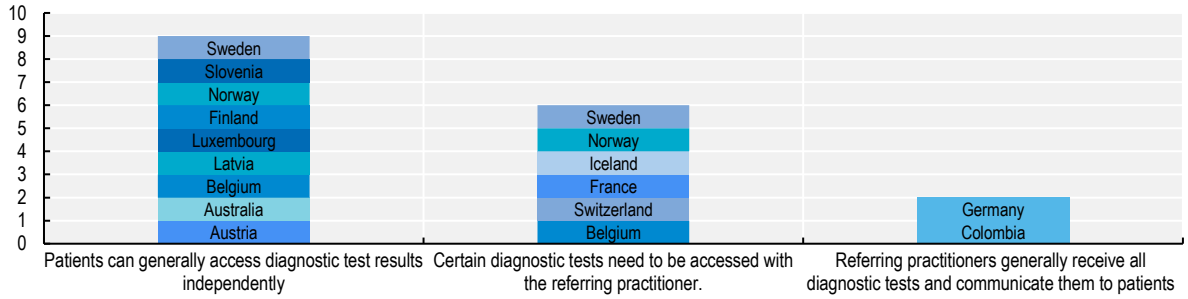
215. Patient reports can uncover diagnostic error that wasn't detected through other means (Kistler et al., 2010[275]). In addition, encouraging patient reporting of adverse events including diagnostic error has intrinsic value in that it can involve people in their care (Giardina et al., 2021[276]). It also signals and manifests a people-centred approach more broadly and should be a key component of any measurement framework. At a minimum, diagnostic safety should be included in the collection of patient-reported outcomes and experiences.[24] Recent initiatives, such as the PAIRED project, in the **United States**, have worked with patients and family members with lived experiences of diagnostic error to co-produce a curriculum that promoted patient engagement in diagnostic research (Sheridan et al., 2020[277]). Another example is the Safer Dx Patient Instrument, a proactive, structured evaluation of a patient's medical record

[24] https://www.oecd.org/en/about/programmes/patient-reported-indicator-surveys-paris.html

by that patient, which can potentially improve transparency in the diagnostic process (Giardina et al., 2022[278]).

216. Patients having access to their own test results empowers them to identify potential errors, seek timely clarification, and actively participate in their healthcare decisions. A large number of countries indicate that patients can either generally access test results independently—either at the same time as time as the provider (as in **Austria**, **Latvia**, **Luxembourg**, **Norway**, and **Sweden**) or with a delay (in **Australia**, **Belgium**, **Japan**, and **Finland**) (see Figure 4.12). In some cases, the diagnosis needs to be communicated by the referring practitioner, for example, the results of diagnostic imaging (**Sweden**), genetic testing (**Switzerland**), or diagnostic tests under development (**Norway**). **Germany** and **Colombia** indicate that test results are most frequently communicated to patients by referring practitioners. **Hungary** sets limits on the types of services that can be delivered remotely, so that physicians making a final diagnosis or a significant therapeutic change are required to do so in the presence of the patient (Oliveira Hashiguchi, 2020[279]).

Figure 4.12. In many countries, patients can access their diagnostic test results independently of a health care provider



Source: OECD Survey on the Assessment of the Adoption of Systems and Interventions to Improve Diagnostic Safety, 2024

217. Clear and documented communication of diagnosis by healthcare professionals to patients is essential to reduce diagnostic error caused by miscommunication or follow-up. Patients often lack the health literacy or digital literacy to access and fully understand testing results or the implications of a medical diagnosis. Communicating diagnosis using appropriate and accessible language, accompanied by printed plain language summaries, or supported by electronic records and telephone calls, can help improve patient engagement and adherence to planned follow-up (Atlas et al., 2023[126]). There is a need to provide training for healthcare professionals in health communication and professional health literacy (De Gani et al., 2023[280]), and should prioritised in education as a lever to improve diagnostic safety.

Digital technology can both help and hinder diagnostic safety

218. Data fuel (AI) models. Better data – all things being equal – will lead to better AI, which can help improve diagnosis but only if effectively implemented. The potential of artificial intelligence (AI) in healthcare practice and analytics can be overstated. However, an abundance of data and greater computation power mean that modern analytical approaches such as machine learning can certainly provide the means to develop measurement approaches not possible a decade ago (Oliveira Hashiguchi, Slawomirski and Oderkirk, 2021[281]). Natural language processing in particular carries the potential to change the game in this regard because much clinical documentation is recorded in narrative form (OECD, 2019[46]).

219. Experimentation with new analytic methods has already commenced with positive early signs. The FIND-AF algorithm, developed by **English** researchers using machine learning models, which scans GP medical records for red flags that a patient might develop atrial fibrillation, which undiagnosed can lead to stroke, is being piloted.[25] [26]

220. Caution is again urged, however, as cases of error and bias in AI diagnostic and treatment algorithms exist and must be minimised, and the patient safety implications of AI in diagnosis are yet to be fully understood (Oliveira Hashiguchi, Slawomirski and Oderkirk, 2021[281]; Matheny et al., 2025[282]). Nevertheless, human cognitive and cultural biases exist in the conventional practice of medical diagnosis without AI, which may be programmed to minimise these. Biased algorithms can result in disparities in patient care, for example AI algorithms applied to chest radiographs systematically missing diagnoses in under-represented patient populations (Seyyed-Kalantari et al., 2021[283]) (Tipton et al., 2023[284]). An algorithm to estimate glomerular filtration rate algorithm to assess kidney function for kidney transplant that disadvantaged black patients serves as an example (Williams, Hogan and Ingelfinger, 2021[285]).

221. The effectiveness of AI can be overestimated. A sepsis prediction tool integrated into EHRs by a national vendor in the United States was evaluated and found to have missed almost 70% patients identified by clinical teams (Classen, Longhurst and Thomas, 2023[286]). Pre-emptive diagnosis using large datasets and machine learning can lead to 'digital overdiagnosis' (Capurro, Coghlan and Pires, 2022[287]).

Integrating AI into routine practice

222. Human-machine interaction can present challenges. A recent study of physicians in the US, for example, found that Large Language Model (LLM), when used on its own, outperformed clinicians in identifying the correct diagnosis for information provided via clinical vignette. However, clinicians who used the LLM performed at the same (inferior) level as those who did not have access to the tool (Goh et al., 2024[288]). A study of computer-assisted diagnostic making in emergency departments in Switzerland demonstrated no improvement in diagnostic quality or errors (Hautz et al., 2025[289]). Other efforts have been found to be more effective. For example, an evaluation of AI algorithms to support breast cancer mammograms and TB X-ray screening, researchers demonstrated the potential for improved accuracy (Eisemann et al., 2025[290]), and reduced physician workload (Dvijotham et al., 2023[291]).

223. Even so, these findings demonstrate that despite access to supporting tools, their impact on improving diagnostic outcomes may be limited without sufficient workforce training and integration into workflow processes.

224. Health professionals will need to develop new skills for using AI and other technological advancements in a way that ensures and increases diagnostic safety. Health professionals are required to pursue ongoing education and training, with a particular focus on integrating emerging health discoveries and innovations. Recent recommendations from the OECD on the health workforce and AI note that, "The integration of AI can be likened to adopting a novel protocol for immunotherapy in cancer treatment who had to undergo specialised re-training to address the advancements in cancer treatment, including understanding the mechanisms of immunotherapy, patient selection criteria, potential side effects, and the management of these side effects. Similarly with AI, health workers need to comprehend the significance of this innovation, grasp the essential knowledge, and effectively implement the innovation" (Almyranti et al., 2024[292]). In addition, health professionals should be included to advise on assessments of the clinical validity of novel AI in health solutions and ways to optimize their use in clinical workflows.

[25] Algorithm could help prevent thousands of strokes in UK each year | Stroke | The Guardian

[26] https://www.bhf.org.uk/what-we-do/news-from-the-bhf/news-archive/2024/december/first-of-its-kind-algorithm-helping-to-identify-hidden-heart-condition

Tacit knowledge is difficult to instil in a machine

225. LLMs are not superior to clinicians in a number of scenarios—as evaluated in a primary care study in Sweden, where the use of ChatGPT was found to be inferior to physician performance in formulating responses to medical examination scenarios (Arvidsson et al., 2024[293]).

226. A key concern is that current AI cannot effectively account for the real-world, human aspects of diagnosis that are extremely difficult to codify. Like in all professions, part of medical knowledge can be 'tacit', extending beyond what can be expressed or articulated (Polanyi, 1967[294]). At this stage, this type of knowledge can only be acquired through real-world experience, not mining large datasets, It includes not only empathy and communication, but also fundamental clinical skills, history taking and physical examination as well as factoring in (in)accuracy and (in)completeness of EHR data, and general uncertainty involved in diagnosis (Kulkarni and Singh, 2023[295]), which can be related to contextual factors that may not be documented or even expressed verbally.

227. Trust is an important foundation for the effective and responsible use of AI tools into healthcare (Almyranti et al., 2024[292]). Improved e-health literacy can improve levels of trust among the public as it relates to digital health-related innovations and technologies (Paige, Krieger and Stellefson, 2016[296]). A lack of trust in the use of digital tools among providers, patients, and the public can result in hesitancy to use or scale AI solutions that could improve outcomes. A survey conducted in early 2023 indicated that 60% of the public in the United States indicated that they would not comfortable with their physician using AI to inform a diagnosis (Tyson et al., 2023[297]).

Developing knowledge, skill and acumen in the clinical workforce

Calibrating diagnostic skills through training and feedback

228. Good diagnostic practice is hard. Practitioners must balance undertesting, and possibly missing a diagnosis, with over testing, which could be harmful and costly. A good diagnostician must concurrently think about diagnostic accuracy, sensitivity and specificity (i.e. uncertainty) the patient's preferences and circumstances, as well as the costs of diagnostics and the costs and benefits of any consequent intervention. Much diagnostic error can be attributed to cognitive errors and poor decision making. Diagnostic calibration – when clinicians' confidence in their diagnostic ability aligns with the objective quality of diagnosis – is a useful concept here (Meyer et al., 2013[298]). Low confidence can lead to over testing and vice versa.

229. The groundwork for calibration should be done in undergraduate training, mainly through the development of transversal skills including cognitive bias, reflection, communication and teamwork. Calibration can be developed in the workplace through continuous feedback and learning. However, diagnosis goes to the heart of medical practice and many practitioners (especially those trained in a competitive, individualistic environment) may be uncomfortable with this feedback, and feel that such feedback undermines their professional self-image and status—despite the known influence of systems-related factors on diagnosis (Meyer and Singh, 2019[299]).

230. Unsurprisingly, feedback on diagnostic performance isn't yet the norm. Improving and normalising this should help. Calibrate Dx is a practical tool to help clinicians do advance this in their own practice[27]. Concurrent steps should be taken to implement feedback on diagnostic performance within healthcare organisations. Feedback will often need to involve teams because diagnosis is a team endeavour (Meyer and Singh, 2019[299]).

[27] https://www.ahrq.gov/diagnostic-safety/tools/calibrate-dx.html

231. Diagnostic safety heavily depends on individual clinicians, in particular, the awareness, and management of, cognitive and behavioural processes that can lead to incorrect diagnosis. Providing feedback can help standardise or optimise to acceptable norms the ordering of diagnostic tests. For example, in **Australia**, the National Prescribing Service's feedback to GPs on their rates of referrals for CT scans for low back pain was associated with an 11% fall in referral rate, equivalent to some 50 000 scans (OECD, 2017[24]).

How medical trainees are taught and socialised matters

232. As mentioned, erroneous thinking and cognitive bias influence diagnostic safety. Early intervention is among most effective way to address this. It needs to begin in the way students and selected, taught and socialised. In the United States, the Institute of Medicine, AHRQ and National Quality Forum advocate improved teaching of diagnosis (Balogh et al., 2015[1]; AHRQ, 2022[300]) (National Quality Forum, 2017[301]). The scientific literature suggests that recognising and communicating diagnostic, therapeutic and prognostic uncertainty, and assessing diagnostic calibration and performance should be taught early (Moulder, Harris and Santhosh, 2022[302]; Graber, Wachter and Cassel, 2012[303]; Meyer et al., 2013[298]).

233. Moreover, techniques such as deliberate reflection and metacognition that recognise and counteract cognitive limitations should be instilled in students and young practitioners (Croskerry, 2022[304]; Kuhn et al., 2023[305]; Royce, Hayes and Schwartzstein, 2019[306]). Of course, success can be influenced by how students are chosen and socialised. Reflecting on one's biases and on the way one goes about solving problems is easier for some than others, especially if these implicit biases concerns others' race, sexual orientation or social status (drivers of underdiagnosis) or if the learning environment is based on competition and the belief that some professions inherently superior to others.

234. The intention is not to disparage the medical professions and institutions, but to emphasise that – when it comes to diagnostic safety and its drivers – metacognitive attributes firmly in the domain of emotional intelligence are equally as instrumental as academic ability, not least because safe diagnosis, like safe healthcare, is a team sport, where no member consider themselves superior to others. In the end, a certain medical culture is among the root causes of failure (diagnostic and other). Progress will not be made without changing the culture away from hierarchies, authority and aggrandisement. Education – both overt and covert – plays an integral role in shaping professional culture. Student selection and training that promote diversity and soft/transversal skills are therefore highly important considerations for institutions to consider.

5 Conclusions

235. The process of diagnosis, owing to its complex, iterative and dynamic nature, presents many sources for potential error in accurately and timely identifying the underlying health problem, and communicating this to the patient. Diagnostic errors have underlying causes, which may be exacerbated or improved by cultural, behavioural, technological, and administrative drivers, at the level of individual clinical practice or across the organisation of the healthcare system. Changes in medical culture and clinical governance, routine information collection and reporting, and rationalisation and standardisation of diagnostic testing and interpretation can help identify and ultimately reduce harm from diagnostic error.

236. Reducing diagnostic harm from misdiagnosis, underdiagnosis or overdiagnosis has the potential for large cost savings through improvements in efficiency and reduction in patient harm. Expenditure on diagnostics, as well as the costs from resulting diagnostic harm, accounts for a considerable proportion of total healthcare expenditure. Increased use of AI in healthcare and availability of novel diagnostic technologies such as genomic testing, present challenges to safeguarding diagnostic safety, but also potential solutions to reducing diagnostic error and associated harms in clinical practice.

237. Improvement isn't free, of course, and any estimate of potential savings must also consider the costs of implementing the strategies and interventions to improve diagnostic safety outlined in this report. These predominantly concern reforms in policy, regulation and governance, which is difficult to price especially when dealing with a range of health systems that differ in their structure and policy settings (see Box 5.1). For example, changes in health information governance to allow routine linkage of various data sources will impart different financial costs in different countries. The same can be said for changing medical training curricula and student selection, or healthcare payment models – although it's safe to say that the price for such reforms will mostly be political. In fact, none of the strategies and interventions proposed here – with the possible exception of AI, whose carbon costs must also be part of the equation – are particularly expensive (financially speaking at least) especially compared to what countries spend on acute care and on biomedical technologies.

238. Investing financial and political capital on improving diagnostic safety will produce a very good return for patients, providers, and payers and insurers. A modest target of halving the rates of misdiagnosis, overdiagnosis and underdiagnosis would free up 8% of healthcare expenditure if the cost of improvement is assumed to be 0.8%. In the United States (where even greater improvements are probably on offer) 8% equates to USD 390 Billion a year. Across OECD countries, an 8% saving would equate to USD 676 Billion.[28] A more ambitious target of 80% -- the current estimate of the proportion of diagnostic error deemed preventable – would represent approximately 13% of healthcare expenditure,[29] which would equate to approximate savings of USD 1.1 Trillion a year.

[28] Combined GDP OECD countries (without USA) x 9% (healthcare expenditure) x 8%; plus USA componenet (USD390B) (2023 values).
[29] Assuming a higher improvement cost of 1% health expenditure.

Box 5.1. What can policymakers do to improve diagnostic safety?

Clinical directors should foster changes in medical work culture and clinical environment for peer feedback and multidisciplinary approach to patient diagnosis and review. Medical education and training should encourage a move away from silo or defensive practice of medicine towards a less risk averse approach to diagnostic uncertainty. The risks, harms and costs from diagnostic error, including overdiagnosis should be incorporated into undergraduate and continuous medical education.

Patient perspectives and preferences should be taken into account when making and reviewing a diagnosis. To avoid the harms of overdiagnosis and overtreatment, the health consequences and costs of diagnostic testing should be discussed prior to diagnostic testing, and their relative benefits explained to patients. Uncertainty in the interpretation of diagnostic and genomic testing results should be acknowledged by clinicians, and risk-benefit ratio of performing these tests assessed prior to testing.

Medical specialty associations should set national or international standards and guidelines for ordering diagnostic testing and interpreting results, to minimise diagnostic error, harms and wasteful healthcare expenditure. Specialty associations should engage with and prioritise issues for Diagnostic Safety initiatives such as *Choosing Wisely*, and champion the "Less but better" approach to diagnosis and diagnostic testing.

National patient safety agencies should routinely collect, report and publish quality assurance indicators for error and safety for diagnosis of common conditions such as cancer screening, mental health disorders and sepsis. These indicators, alongside international norms, should be used to set acceptable standards to audit diagnostic performance and drive diagnostic safety improvement initiatives.

Health financing should report on regional or institutional variations or anomalies in expenditure and reimbursement for diagnosis rates or diagnostic testing, indicative or poor quality care. Hospitals should set clear targets for reducing prescribing or ordering of radiology or laboratory diagnostic tests where they exceed norms, to promote diagnostic stewardship.

Healthcare insurers and providers should review policies for financing and reimbursement of diagnostic practices that do not conform to best international practice or guidelines in order to enable healthcare expenditure savings. Policy set at identifying excess expenditure on low value diagnostic healthcare is a means of reducing both diagnostic error and harms, and costs for patients, providers and insurers.

Healthcare systems should leverage digital health architecture to prioritise development of integrated health information flows between patients and different healthcare providers, to ensure timely and systematic follow-up and communication of diagnosis. Alerts and tracking of pending results should be made available to both healthcare providers and patients where possible and clinically appropriate, to avoid miscommunication or delays in delivering diagnosis.

The use of language learning models and AI to analyse multiple clinical, biomedical and radiological patient data sources to achieve a more accurate and timely diagnosis requires clinical validation and ongoing refinement, but may be of use in conditions where clinical diagnosis is currently challenging or reliable diagnostic testing is lacking.

References

Aaron, S. et al. (2024), "Early Diagnosis and Treatment of COPD and Asthma — A Randomized, Controlled Trial", *New England Journal of Medicine*, Vol. 390/22, pp. 2061-2073, https://doi.org/10.1056/nejmoa2401389. [181]

ACSQHC (2018), *Third Australian Atlas of Healthcare Variation: Myocardial perfusion scans, 18 years and over | Australian Commission on Safety and Quality in Health Care*, https://www.safetyandquality.gov.au/our-work/healthcare-variation/third-atlas-2018/atlas-2018-4-cardiac-tests/43-myocardial-perfusion-scans-18-years-and-over (accessed on 24 July 2024). [154]

Agency for Healthcare Research and Quality (2024), *Diagnostic Stewardship as a Model To Improve the Quality and Safety of Diagnosis*, https://www.ahrq.gov/diagnostic-safety/resources/issue-briefs/dxsafety-dx-stewardship2.html, https://www.ahrq.gov/diagnostic-safety/resources/issue-briefs/dxsafety-dx-stewardship2.html (accessed on 18 February 2025). [260]

Agha, L., J. Skinner and D. Chan (2022), "Improving Efficiency in Medical Diagnosis", *JAMA*, Vol. 327/22, p. 2189, https://doi.org/10.1001/jama.2022.8587. [252]

Ahmed, A. et al. (1997), "Screening for Osteopenia and Osteoporosis: Do the Accepted Normal Ranges Lead to Overdiagnosis?", *Osteoporosis International*, Vol. 7/5, pp. 432-438, https://doi.org/10.1007/s001980050029. [310]

Ahn, H., H. Kim and H. Welch (2014), "Korea's Thyroid-Cancer "Epidemic" — Screening and Overdiagnosis", *New England Journal of Medicine*, Vol. 371/19, pp. 1765-1767, https://doi.org/10.1056/nejmp1409841. [164]

Ahn, H. and H. Welch (2015), "South Korea's Thyroid-Cancer "Epidemic" — Turning the Tide", *New England Journal of Medicine*, Vol. 373/24, pp. 2389-2390, https://doi.org/10.1056/nejmc1507622. [215]

Aho, J. et al. (2022), "Rising syphilis rates in Canada, 2011–2020", *Canada Communicable Disease Report*, Vol. 47/2-3, pp. 52-60, https://doi.org/10.14745/ccdr.v48i23a01. [186]

AHRQ (2022), *Improving Education—A Key to Better Diagnostic Outcomes*, https://www.ahrq.gov/diagnostic-safety/resources/issue-briefs/education-dx-outcomes-1.html (accessed on 21 January 2025). [300]

Al-Aly, Z. et al. (2024), "Long COVID science, research and policy", *Nature Medicine*, Vol. 30/8, pp. 2148-2164, https://doi.org/10.1038/s41591-024-03173-6. [130]

Albarqouni, L. et al. (2022), "Overdiagnosis and overuse of diagnostic and screening tests in low-income and middle-income countries: a scoping review", *BMJ Global Health*, Vol. 7/10, p. e008696, https://doi.org/10.1136/bmjgh-2022-008696. [176]

Ali, K. et al. (2023), "The PRIDx framework to engage payers in reducing diagnostic errors in healthcare", *Diagnosis*, Vol. 11/1, pp. 17-24, https://doi.org/10.1515/dx-2023-0042. [255]

Almyranti, M. et al. (2024), "Artificial Intelligence and the health workforce: Perspectives from medical associations on AI in health", *OECD Artificial Intelligence Papers*, No. 28, OECD Publishing, Paris, https://doi.org/10.1787/9a31d8af-en. [292]

Ancochea, J. et al. (2013), "Infradiagnóstico de la enfermedad pulmonar obstructiva crónica en mujeres: cuantificación del problema, determinantes y propuestas de acción", *Archivos de Bronconeumología*, Vol. 49/6, pp. 223-229, https://doi.org/10.1016/j.arbres.2012.11.010. [185]

Angus, D. and A. Bindman (2022), "Achieving Diagnostic Excellence for Sepsis", *JAMA*, Vol. 327/2, pp. 117-118, https://doi.org/10.1001/jama.2021.23916. [101]

Angus, D. et al. (2016), "A Framework for the Development and Interpretation of Different Sepsis Definitions and Clinical Criteria", *Critical care medicine*, Vol. 44/3, p. e113, https://doi.org/10.1097/CCM.0000000000001730. [26]

Anjum, O., H. Bleeker and R. Ohle (2019), "Computed tomography for suspected pulmonary embolism results in a large number of non-significant incidental findings and follow-up investigations", *Emergency Radiology*, Vol. 26/1, pp. 29-35, https://doi.org/10.1007/s10140-018-1641-8. [269]

Arvidsson, R. et al. (2024), "ChatGPT (GPT-4) versus doctors on complex cases of the Swedish family medicine specialist examination: an observational comparative study", *BMJ Open*, Vol. 14/12, p. e086148, https://doi.org/10.1136/bmjopen-2024-086148. [293]

Askari, A. et al. (2017), "Who requires emergency surgery for colorectal cancer and can national screening programmes reduce this need?", *International Journal of Surgery*, Vol. 42, pp. 60-68, https://doi.org/10.1016/J.IJSU.2017.04.050. [18]

Atlas, S. et al. (2023), "A Multilevel Primary Care Intervention to Improve Follow-Up of Overdue Abnormal Cancer Screening Test Results: A Cluster Randomized Clinical Trial", *JAMA*, Vol. 330/14, pp. 1348-1358, https://doi.org/10.1001/jama.2023.18755. [126]

Auraaen, A., K. Saar and N. Klazinga (2020), "System governance towards improved patient safety: Key functions, approaches and pathways to implementation", *OECD Health Working Papers*, No. 120, OECD Publishing, Paris, https://doi.org/10.1787/2abdd834-en. [59]

Auraaen, A., L. Slawomirski and N. Klazinga (2018), "The economics of patient safety in primary and ambulatory care: Flying blind", *OECD Health Working Papers*, No. 106, OECD Publishing, Paris, https://doi.org/10.1787/baf425ad-en. [25]

Bachmann, C. et al. (2017), "Trends in ADHD medication use in children and adolescents in five western countries, 2005–2012", *European Neuropsychopharmacology*, Vol. 27/5, pp. 484-493, https://doi.org/10.1016/j.euroneuro.2017.03.002. [74]

Bach, P. (2015), "New Math on Drug Cost-Effectiveness", *New England Journal of Medicine*, Vol. 373/19, pp. 1797-1799, https://doi.org/10.1056/nejmp1512750. [178]

Baicker, K. et al. (2018), "The Effect of Medicaid on Management of Depression: Evidence From the Oregon Health Insurance Experiment", *The Milbank Quarterly*, Vol. 96/1, pp. 29-56, https://doi.org/10.1111/1468-0009.12311. [204]

Balogh, E., B. Miller and J. Ball (eds.) (2015), *Improving Diagnosis in Health Care*, National Academies Press, Washington, D.C., https://doi.org/10.17226/21794. [39]

Balogh, E., B. Miller and J. Ball (eds.) (2015), *Improving Diagnosis in Health Care*, National Academies Press, Washington, D.C., https://doi.org/10.17226/21794. [312]

Balogh, E. et al. (2015), "Improving Diagnosis in Health Care", *Improving Diagnosis in Health Care*, pp. 1-472, https://doi.org/10.17226/21794. [1]

Barlam, T. et al. (2016), "Unnecessary Antibiotics for Acute Respiratory Tract Infections: Association With Care Setting and Patient Demographics", *Open Forum Infectious Diseases*, Vol. 3/1, https://doi.org/10.1093/ofid/ofw045. [147]

Barratt, A. and F. McGain (2021), "Overdiagnosis is increasing the carbon footprint of healthcare", *BMJ*, p. n2407, https://doi.org/10.1136/bmj.n2407. [195]

Ben-Assuli, O. et al. (2020), "The use of analytic hierarchy process for measuring the complexity of medical diagnosis", *Health Informatics Journal*, Vol. 26/1, pp. 218-232, https://doi.org/10.1177/1460458218824708/ASSET/IMAGES/LARGE/10.1177_1460458218824708-FIG5.JPEG. [10]

Black, W. (1998), "Advances in radiology and the real versus apparent effects of early diagnosis", *European Journal of Radiology*, Vol. 27/2, pp. 116-122, https://doi.org/10.1016/s0720-048x(97)00160-5. [206]

Bleyer, A. and H. Welch (2012), "Effect of Three Decades of Screening Mammography on Breast-Cancer Incidence", *New England Journal of Medicine*, Vol. 367/21, pp. 1998-2005, https://doi.org/10.1056/nejmoa1206809. [160]

Boyum, J. et al. (2016), "Incidence and Risk Factors for Adverse Events Related to Image-Guided Liver Biopsy", *Mayo Clinic Proceedings*, Vol. 91/3, pp. 329-335, https://doi.org/10.1016/J.MAYOCP.2015.11.015. [158]

Bradford, A. et al. (2024), "Diagnostic error in mental health: a review", *BMJ Quality & Safety*, Vol. 33/10, p. 663, https://doi.org/10.1136/bmjqs-2023-016996. [61]

Braithwaite, J., R. Wears and E. Hollnagel (2015), "Resilient health care: turning patient safety on its head", *International Journal for Quality in Health Care*, Vol. 27/5, pp. 418-420, https://doi.org/10.1093/intqhc/mzv063. [55]

Brett, D. et al. (2016), "Factors Affecting Age at ASD Diagnosis in UK: No Evidence that Diagnosis Age has Decreased Between 2004 and 2014", *Journal of Autism and Developmental Disorders*, Vol. 46/6, pp. 1974-1984, https://doi.org/10.1007/s10803-016-2716-6. [85]

Bulliard, J. et al. (2021), "Breast cancer screening and overdiagnosis", *International Journal of Cancer*, Vol. 149/4, pp. 846-853, https://doi.org/10.1002/ijc.33602. [108]

Burke, J. et al. (2019), "Improving outcome of sepsis on the ward: introducing the 'Sepsis Six' bundle", *Nursing in Critical Care*, Vol. 24/1, pp. 33-39, https://doi.org/10.1111/nicc.12358. [100]

Burstin, H. and K. Cosby (2022), "Measuring Performance of the Diagnostic Process", *JAMA*, Vol. 328/2, p. 143, https://doi.org/10.1001/jama.2022.10166. [12]

Burstin, H. and E. Schneider (2022), "Building Connections Between Guidelines and Quality Improvement", *Annals of Internal Medicine*, Vol. 175/5, pp. 755-756, https://doi.org/10.7326/m22-0409. [243]

Callen, J. et al. (2012), "Failure to Follow-Up Test Results for Ambulatory Patients: A Systematic Review", *Journal of General Internal Medicine*, Vol. 27/10, pp. 1334-1348, https://doi.org/10.1007/s11606-011-1949-5. [123]

Callister, M., P. Sasieni and H. Robbins (2021), "Overdiagnosis in lung cancer screening", *The Lancet Respiratory Medicine*, doi: 10.1016/S2213-2600(20)30553-1, pp. 7-9, https://doi.org/10.1016/S2213-2600(20)30553-1. [110]

Cambridge Econometrics (2024), *The economic burden of Long COVID in the UK*, https://www.camecon.com/wp-content/uploads/2024/03/The-Economic-Burden-of-Long-Covid-in-the-UK_Cambridge-Econometrics_V1.1_March2024.pdf (accessed on 6 January 2025). [138]

Camidge, D. et al. (2022), "Quantifying The Medical Impact of A Missed Diagnosis of Non-Small Cell Lung Cancer on Chest Imaging", *Clinical Lung Cancer*, Vol. 23/5, pp. 377-385, https://doi.org/10.1016/j.cllc.2022.03.006. [116]

Camillo, C. (2023), "Addressing the ethical problem of underdiagnosis in the post-pandemic Canadian healthcare system", *Healthcare Management Forum*, Vol. 36/6, pp. 420-423, https://doi.org/10.1177/08404704231200113. [180]

Capurro, D., S. Coghlan and D. Pires (2022), "Preventing Digital Overdiagnosis", *JAMA*, Vol. 327/6, p. 525, https://doi.org/10.1001/jama.2021.22969. [287]

Car, L. et al. (2016), "Clinician-identified problems and solutions for delayed diagnosis in primary care: a PRIORITIZE study", *BMC Family Practice*, Vol. 17/1, https://doi.org/10.1186/S12875-016-0530-Z. [200]

Cecchini, M., J. Langer and L. Slawomirski (2015), "ANTIMICROBIAL RESISTANCE IN G7 COUNTRIES AND BEYOND: Economic Issues, Policies and Options for Action". [148]

Cheraghi-Sohi, S. et al. (2021), "Incidence, origins and avoidable harm of missed opportunities in diagnosis: longitudinal patient record review in 21 English general practices", *BMJ Quality & Safety*, Vol. 30/12, p. 977, https://doi.org/10.1136/bmjqs-2020-012594. [4]

Chew, D. et al. (2016), "Variation in coronary angiography rates in Australia: correlations with socio-demographic, health service and disease burden indices", *Medical Journal of Australia*, Vol. 205/3, pp. 114-120, https://doi.org/10.5694/mja15.01410. [189]

Choi, Y. et al. (2023), "Long-Term Changes in the Mortality Rates of Thyroid Cancer in Korea: Analysis of Korean National Data from 1985 to 2020", *Endocrinology and Metabolism*, Vol. 38/5, pp. 588-595, https://doi.org/10.3803/enm.2023.1723. [263]

Christensen, K. et al. (2012), "Use of general practice, diagnostic investigations and hospital services before and after cancer diagnosis - a population-based nationwide registry study of 127,000 incident adult cancer patients", *BMC Health Services Research*, Vol. 12/1, p. 224, https://doi.org/10.1186/1472-6963-12-224. [118]

Churruca, K. et al. (2021), "Patient-reported outcome measures (PROMs): A review of generic and condition-specific measures and a discussion of trends and issues", *Health Expectations : An International Journal of Public Participation in Health Care and Health Policy*, Vol. 24/4, p. 1015, https://doi.org/10.1111/HEX.13254. [13]

Ciello, A. et al. (2017), "Missed lung cancer: when, where, and why?", *Diagnostic and Interventional Radiology*, Vol. 23/2, pp. 118-126, https://doi.org/10.5152/dir.2016.16187. [120]

Classen, D., C. Longhurst and E. Thomas (2023), "Bending the patient safety curve: how much can AI help?", *npj Digital Medicine*, Vol. 6/1, https://doi.org/10.1038/s41746-022-00731-5. [286]

Clearstate (2023), *Point-of-care testing: promises and challenges*, https://www.clearstate.com/poct-promises-and-challenges/ (accessed on 19 December 2024). [49]

Cochrane COVID-19 Diagnostic Test Accuracy Group (2022), "Rapid, point-of-care antigen tests for diagnosis of SARS-CoV-2 infection", *Cochrane Database of Systematic Reviews*, Vol. 2022/7, https://doi.org/10.1002/14651858.cd013705.pub3. [37]

Coleman, A. et al. (2018), "The unsolved problem of otitis media in indigenous populations: a systematic review of upper respiratory and middle ear microbiology in indigenous children with otitis media", *Microbiome*, Vol. 6/1, https://doi.org/10.1186/s40168-018-0577-2. [188]

Collaborators, G. (2022), "Global, regional, and national burden of 12 mental disorders in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019", *The Lancet Psychiatry*, doi: 10.1016/S2215-0366(21)00395-3, pp. 137-150, https://doi.org/10.1016/S2215-0366(21)00395-3. [66]

Copp, T. et al. (2017), "Are expanding disease definitions unnecessarily labelling women with polycystic ovary syndrome?", *BMJ*, p. j3694, https://doi.org/10.1136/bmj.j3694. [170]

Correll, C. et al. (2017), "Prevalence, incidence and mortality from cardiovascular disease in patients with pooled and specific severe mental illness: a large-scale meta-analysis of 3,211,768 patients and 113,383,368 controls", *World Psychiatry*, Vol. 16/2, pp. 163-180, https://doi.org/10.1002/wps.20420. [87]

Croskerry, P. (2022), "The Rational Diagnostician and Achieving Diagnostic Excellence", *JAMA*, Vol. 327/4, p. 317, https://doi.org/10.1001/jama.2021.24988. [304]

Csanádi, M. et al. (2019), "Key indicators of organized cancer screening programs: Results from a Delphi study", *Journal of Medical Screening*, doi: 10.1177/0969141318820362, pp. 120-126, https://doi.org/10.1177/0969141318820362. [129]

Cundy, T. (2012), "Proposed new diagnostic criteria for gestational diabetes – a pause for thought?", *Diabetic Medicine*, Vol. 29/2, pp. 176-180, https://doi.org/10.1111/j.1464-5491.2011.03407.x. [208]

Dalton, A. (2018), "Incomplete diagnostic follow-up after a positive colorectal cancer screening test: a systematic review", *Journal of Public Health*, Vol. 40/1, pp. e46-e58, https://doi.org/10.1093/pubmed/fdw147. [124]

Danish Medicines Agency (n.d.), *Changes as a result of the EU regulation on in vitro diagnostic medical devices*, https://laegemiddelstyrelsen.dk/en/devices/new-regulations/changes-as-a-result-of-the-eu-regulation-on-in-vitro-diagnostic-medical-devices/ (accessed on 19 December 2024). [48]

Darzi (2024), *Independent Investigation of the National Health Service in England*, https://assets.publishing.service.gov.uk/media/66f42ae630536cb92748271f/Lord-Darzi-Independent-Investigation-of-the-National-Health-Service-in-England-Updated-25-September.pdf. [82]

Darzi, O. (2024), "Independent Investigation of the National Health Service in England", http://www.gov.uk (accessed on 21 January 2025). [83]

Davidovitch, M. et al. (2017), "Challenges in defining the rates of ADHD diagnosis and treatment: trends over the last decade", *BMC Pediatrics*, Vol. 17/1, p. 218, https://doi.org/10.1186/s12887-017-0971-0. [63]

de Bienassis, K. et al. (2022), "The economics of medication safety: Improving medication safety through collective, real-time learning", *OECD Health Working Papers*, No. 147, OECD Publishing, Paris, https://doi.org/10.1787/9a933261-en. [251]

de Bienassis, K. and N. Klazinga (2024), "Comparative assessment of patient safety culture performance in OECD countries: Findings based on the Hospital Survey on Patient Safety Culture versions 1 and 2", *OECD Health Working Papers*, No. 168, OECD Publishing, Paris, https://doi.org/10.1787/d0552328-en. [274]

de Bienassis, K., A. Llena-Nozal and N. Klazinga (2020), *The economics of patient safety Part III: Long-term care: Valuing safety for the long haul*, OECD Health Working Paper, https://www.oecd-ilibrary.org/social-issues-migration-health/the-economics-of-patient-safety-part-iii-long-term-care_be07475c-en (accessed on 1 April 2021). [258]

de Bienassis, K., L. Slawomirski and N. Klazinga (2021), "The economics of patient safety Part IV: Safety in the workplace: Occupational safety as the bedrock of resilient health systems", *OECD Health Working Papers*, No. 130, OECD Publishing, Paris, https://doi.org/10.1787/b25b8c39-en. [257]

De Gani, S. et al. (2023), "Professional health literacy - first results of a pilot study in Austria, Germany, and Switzerland", *Population Medicine*, Vol. 5/Supplement, https://doi.org/10.18332/popmed/164043. [280]

Dembek, C. et al. (2023), "The economic and humanistic burden of bipolar disorder in adults in the United States", *Annals of General Psychiatry*, Vol. 22/1, p. 13, https://doi.org/10.1186/s12991-023-00440-7. [67]

Devos, C. et al. (2019), "Performance of the Belgian health system – report 2019 – Supplement: technical sheets for indicators", https://www.pasderayonssansraisons.be/ihttps://www.health.belgium.be/en/node/27827 (accessed on 8 July 2024). [153]

Doshi, J. et al. (2012), "Economic Impact of Childhood and Adult Attention-Deficit/Hyperactivity Disorder in the United States", *Journal of the American Academy of Child & Adolescent Psychiatry*, Vol. 51/10, pp. 990-1002.e2, https://doi.org/10.1016/j.jaac.2012.07.008. [78]

Douglas, M. and A. Kumar (2022), "Analyzing Precision Medicine Utilization with Real-World Data: A Scoping Review", *Journal of Personalized Medicine*, Vol. 12/4, p. 557, https://doi.org/10.3390/jpm12040557. [45]

Ducatman, B. et al. (2020), "The Value Proposition for Pathologists: A Population Health Approach", *Academic Pathology*, Vol. 7, p. 2374289519898857, https://doi.org/10.1177/2374289519898857. [43]

Duncan, C. et al. (2021), "Diagnostic challenges in sepsis", *Current Infectious Disease Reports*, Vol. 23, pp. 1-14. [95]

Dvijotham, K. et al. (2023), "Enhancing the reliability and accuracy of AI-enabled diagnosis via complementarity-driven deferral to clinicians", *Nature Medicine*, Vol. 29/7, pp. 1814-1820, https://doi.org/10.1038/s41591-023-02437-x. [291]

Eisemann, N. et al. (2025), "Nationwide real-world implementation of AI for cancer detection in population-based mammography screening", *Nature Medicine*, https://doi.org/10.1038/s41591-024-03408-6. [290]

Enck, P. et al. (2016), "Irritable bowel syndrome", *Nature Reviews Disease Primers 2016 2:1*, Vol. 2/1, pp. 1-24, https://doi.org/10.1038/nrdp.2016.14. [27]

Esserman, L. et al. (2014), "Addressing overdiagnosis and overtreatment in cancer: a prescription for change", *The Lancet Oncology*, doi: 10.1016/S1470-2045(13)70598-9, pp. e234-e242, https://doi.org/10.1016/S1470-2045(13)70598-9. [106]

Evans, L. et al. (2021), "Surviving sepsis campaign: international guidelines for management of sepsis and septic shock 2021", *Critical care medicine*, Vol. 49/11, pp. e1063-e1143. [98]

Fayyad, J. et al. (2017), "The descriptive epidemiology of DSM-IV Adult ADHD in the World Health Organization World Mental Health Surveys", *ADHD Attention Deficit and Hyperactivity Disorders*, Vol. 9/1, pp. 47-65, https://doi.org/10.1007/s12402-016-0208-3. [69]

Fedyukova, A., D. Pires and D. Capurro (2021), "Quantifying machine learning-induced overdiagnosis in sepsis". [313]

Fedyukova, A., D. Pires and D. Capurro (2021), "Quantifying machine learning-induced overdiagnosis in sepsis", *ArXiv*, Vol. abs/2107.10399, https://api.semanticscholar.org/CorpusID:236170955. [245]

Ferket, B. et al. (2017), "Impact of total knee replacement practice: cost effectiveness analysis of data from the Osteoarthritis Initiative", *BMJ*, p. j1131, https://doi.org/10.1136/bmj.j1131. [175]

Fineberg, H. et al. (eds.) (2024), *A Long COVID Definition: A Chronic, Systemic Disease State with Profound Consequences*, The National Academies Press, Washington, DC, https://doi.org/10.17226/27768. [137]

Fletcher, T. et al. (2020), "Identifying psychiatric diagnostic errors with the Safer Dx Instrument", *International Journal for Quality in Health Care*, Vol. 32/6, pp. 405-411, https://doi.org/10.1093/intqhc/mzaa066. [62]

Friberg, N. et al. (2019), "Cause of death and significant disease found at autopsy", *Virchows Archiv*, Vol. 475/6, pp. 781-788, https://doi.org/10.1007/s00428-019-02672-z. [144]

Fryar, C. et al. (2024), "Hypertension Prevalence, Awareness, Treatment, and Control Among Adults Age 18 and Older: United States, August 2021–August 2023", https://www.cdc.gov/nchs/products/index.htm. (accessed on 2 January 2025). [142]

Gauer, R., D. Forbes and N. Boyer (2020), "Sepsis: diagnosis and management", *American family physician*, Vol. 101/7, pp. 409-418. [96]

Gentilotti, E. et al. (2023), "Clinical phenotypes and quality of life to define post-COVID-19 syndrome: a cluster analysis of the multinational, prospective ORCHESTRA cohort", *eClinicalMedicine*, Vol. 62, p. 102107, https://doi.org/10.1016/j.eclinm.2023.102107. [134]

Giardina, T. et al. (2022), "Inviting patients to identify diagnostic concerns through structured evaluation of their online visit notes", *Journal of the American Medical Informatics Association*, Vol. 29/6, pp. 1091-1100, https://doi.org/10.1093/jamia/ocac036. [278]

Giardina, T. et al. (2021), "Use of patient complaints to identify diagnosis-related safety concerns: a mixed-method evaluation", *BMJ Quality & Safety*, Vol. 30/12, p. 996, https://doi.org/10.1136/bmjqs-2020-011593. [276]

Giglio, A. et al. (2024), "Adult Code Sepsis: A Narrative Review of its Implementation and Impact", *Journal of Intensive Care Medicine*, https://doi.org/10.1177/08850666241293034. [314]

Gildea, T. et al. (2017), "A retrospective analysis of delays in the diagnosis of lung cancer and associated costs", *ClinicoEconomics and Outcomes Research*, doi: 10.2147/CEOR.S132259, pp. 261-269, https://doi.org/10.2147/CEOR.S132259. [122]

Ginsberg, Y. et al. (2014), "Underdiagnosis of Attention-Deficit/Hyperactivity Disorder in Adult Patients", *The Primary Care Companion For CNS Disorders*, https://doi.org/10.4088/pcc.13r01600. [76]

Goh, E. et al. (2024), "Large Language Model Influence on Diagnostic Reasoning", *JAMA Network Open*, Vol. 7/10, p. e2440969, https://doi.org/10.1001/jamanetworkopen.2024.40969. [288]

Gopal, D. et al. (2021), "Implicit bias in healthcare: clinical practice, research and decision making", *Future Healthcare Journal*, Vol. 8/1, pp. 40-48, https://doi.org/10.7861/fhj.2020-0233. [205]

Government of Ireland (2024), *Healthy Ireland Survey 2024*, https://www.gov.ie/en/publication/d1ab3-healthy-ireland-survey-2024/#covid-and-long-covid (accessed on 4 December 2024). [131]

Graber, M. (2013), "The incidence of diagnostic error in medicine", *BMJ Quality & Safety*, Vol. 22/Suppl 2, pp. ii21-ii27, https://doi.org/10.1136/bmjqs-2012-001615. [139]

Graber, M. (2013), "The incidence of diagnostic error in medicine", *BMJ Quality & Safety*, Vol. 22/Suppl 2, p. ii21, https://doi.org/10.1136/bmjqs-2012-001615. [3]

Graber, M. (1999), "Expanding the Goals of Peer Review to Detect Both Practitioner and System Error", *The Joint Commission Journal on Quality Improvement*, Vol. 25/8, pp. 396-407, https://doi.org/10.1016/s1070-3241(16)30454-0. [231]

Graber, M. et al. (2014), "The Next Organizational Challenge: Finding and Addressing Diagnostic Error", *The Joint Commission Journal on Quality and Patient Safety*, Vol. 40/3, pp. 102-110, https://doi.org/10.1016/s1553-7250(14)40013-8. [229]

Graber, M., R. Wachter and C. Cassel (2012), "Bringing Diagnosis Into the Quality and Safety Equations", *JAMA*, Vol. 308/12, p. 1211, https://doi.org/10.1001/2012.jama.11913. [303]

Graber, M. et al. (2024), "Interventions to improve timely cancer diagnosis: an integrative review", https://doi.org/doi:10.1515/dx-2024-0113. [125]

Grimmsmann, T. and W. Himmel (2021), "The 10-year trend in drug prescriptions for attention-deficit/hyperactivity disorder (ADHD) in Germany", *European Journal of Clinical Pharmacology*, Vol. 77/1, pp. 107-115, https://doi.org/10.1007/s00228-020-02948-3. [75]

Groome, P. et al. (2007), "The IASLC Lung Cancer Staging Project: Validation of the Proposals for Revision of the T, N, and M Descriptors and Consequent Stage Groupings in the Forthcoming (Seventh) Edition of the TNM Classification of Malignant Tumours", *Journal of Thoracic Oncology*, Vol. 2/8, pp. 694-705, https://doi.org/10.1097/JTO.0b013e31812d05d5. [121]

Gunderson, C. et al. (2020), "Prevalence of harmful diagnostic errors in hospitalised adults: a systematic review and meta-analysis", *BMJ Quality & Safety*, Vol. 29/12, p. 1008, https://doi.org/10.1136/bmjqs-2019-010822. [6]

Haase, C. and J. Brodersen (2023), "87 Data as symptom: doctors' responses to patient-generated data in general practice and the potential risks of overdiagnosis", *Preventing overdiagnosis meeting abstracts*, pp. A43.1-A43, https://doi.org/10.1136/ebm-2023-pod.87. [218]

Hackbarth, A. (2012), "Eliminating Waste in US Health Care", *JAMA*, Vol. 307/14, p. 1513, https://doi.org/10.1001/jama.2012.362. [173]

Halbisen, A. and C. Lu (2023), "Trends in Availability of Genetic Tests in the United States, 2012–2022", *Journal of Personalized Medicine*, Vol. 13/4, p. 638, https://doi.org/10.3390/jpm13040638. [223]

Hanley, A. et al. (2021), "The diagnostic odyssey of autism: a cross-sectional study of 3 age cohorts of children from the 2016–2018 National Survey of Children's Health", *Child and Adolescent Psychiatry and Mental Health*, Vol. 15/1, p. 58, https://doi.org/10.1186/s13034-021-00409-y. [86]

HAS (2024), *Les erreurs diagnostiques en médecine: État des lieux*, https://www.has-sante.fr/upload/docs/application/pdf/2024-11/rapport_erreurs_diagnostiques_medecine.pdf (accessed on 8 January 2025). [196]

Hautz, W. et al. (2025), "Diagnoses supported by a computerised diagnostic decision support system versus conventional diagnoses in emergency patients (DDX-BRO): a multicentre, multiple-period, double-blind, cluster-randomised, crossover superiority trial", *The Lancet Digital Health*, doi: 10.1016/S2589-7500(24)00250-4, pp. e136-e144, https://doi.org/10.1016/S2589-7500(24)00250-4. [289]

Heijnsdijk, E. et al. (2009), "Overdetection, overtreatment and costs in prostate-specific antigen screening for prostate cancer", *British Journal of Cancer*, Vol. 101/11, pp. 1833-1838, https://doi.org/10.1038/sj.bjc.6605422. [111]

Helm, B., M. Ayers and A. Kean (2018), "All Along the Watchtower: a Case of Long QT Syndrome Misdiagnosis Secondary to Genetic Testing Misinterpretation", *Journal of Genetic Counseling*, Vol. 27/6, pp. 1515-1522, https://doi.org/10.1007/s10897-018-0287-8. [221]

Henderson, A. et al. (2024), "Clinical coding of long COVID in primary care 2020–2023 in a cohort of 19 million adults: an OpenSAFELY analysis", *eClinicalMedicine*, Vol. 72, p. 102638, https://doi.org/10.1016/j.eclinm.2024.102638. [133]

Herndon, M. et al. (2007), "Implications Of Expanding Disease Definitions: The Case Of Osteoporosis", *Health Affairs*, Vol. 26/6, pp. 1702-1711, https://doi.org/10.1377/hlthaff.26.6.1702. [209]

Hodgkinson, J. et al. (2011), "Relative effectiveness of clinic and home blood pressure monitoring compared with ambulatory blood pressure monitoring in diagnosis of hypertension: systematic review", *BMJ*, Vol. 342/jun24 1, pp. d3621-d3621, https://doi.org/10.1136/bmj.d3621. [311]

Hollnagel, E., R. Wears and J. Braithwaite (2015), *From Safety-I to Safety-II: A White Paper*, National Library of Congress. [56]

Hooftman, J. et al. (2023), "Common contributing factors of diagnostic error: A retrospective analysis of 109 serious adverse event reports from Dutch hospitals", *BMJ Quality & Safety*, Vol. 33/10, pp. 642-651, https://doi.org/10.1136/bmjqs-2022-015876. [272]

Horný, M. et al. (2024), "Decomposition of medical imaging spending growth between 2010 and 2021 in the US employer–insured population", *Health Affairs Scholar*, Vol. 2/3, https://doi.org/10.1093/haschl/qxae030. [42]

Huon, J. et al. (2024), "The cost of potentially inappropriate medications for older adults in Canada: A comparative cross-sectional study", *Journal of the American Geriatrics Society*, Vol. 72/11, pp. 3530-3540, https://doi.org/10.1111/jgs.19164. [91]

Ibrahim T, A. (2023), "The Impact of Wearable IoT Devices on Early Disease Detection and Prevention", *International Journal of Applied Health Care Analytics*, Vol. 8/8, pp. 1-15. [216]

Kalish, H. et al. (2021), "Undiagnosed SARS-CoV-2 seropositivity during the first 6 months of the COVID-19 pandemic in the United States", *Science Translational Medicine*, Vol. 13/601, https://doi.org/10.1126/scitranslmed.abh3826. [192]

Kapadia, P. et al. (2024), "Development and Implementation of a Digital Quality Measure of Emergency Cancer Diagnosis", *Journal of Clinical Oncology*, Vol. 42/21, pp. 2506-2515, https://doi.org/10.1200/JCO.23.01523. [17]

Kapadia, P. et al. (2024), "Development and Implementation of a Digital Quality Measure of Emergency Cancer Diagnosis", *Journal of Clinical Oncology*, Vol. 42/21, pp. 2506-2515, https://doi.org/10.1200/JCO.23.01523. [316]

Kaplan, R. and M. Ong (2007), "Rationale and Public Health Implications of Changing CHD Risk Factor Definitions", *Annual Review of Public Health*, Vol. 28/1, pp. 321-344, https://doi.org/10.1146/annurev.publhealth.28.021406.144141. [210]

Karras, D. et al. (2003), "Antibiotic use for emergency department patients with acute diarrhea", *Annals of Emergency Medicine*, Vol. 42/6, pp. 835-842, https://doi.org/10.1016/s0196-0644(03)00602-4. [227]

Kendir, C. et al. (2023), "Patient engagement for patient safety: The why, what, and how of patient engagement for improving patient safety", *OECD Health Working Papers*, No. 159, OECD Publishing, Paris, https://doi.org/10.1787/5fa8df20-en. [233]

Kerber, K. et al. (2011), "Nystagmus Assessments Documented by Emergency Physicians in Acute Dizziness Presentations: A Target for Decision Support?", *Academic Emergency Medicine*, Vol. 18/6, pp. 619-626, https://doi.org/10.1111/j.1553-2712.2011.01093.x. [52]

Keteepe-Arachi, T. (2017), "Cardiovascular Disease in Women: Understanding Symptoms and Risk Factors", *European Cardiology Review*, Vol. 12/1, p. 10, https://doi.org/10.15420/ecr.2016:32:1. [183]

Kherad, O. and A. Carneiro (2023), "General health check-ups: To check or not to check? A question of choosing wisely", *European Journal of Internal Medicine*, Vol. 109, pp. 1-3, https://doi.org/10.1016/j.ejim.2022.12.021. [33]

Kim, K. et al. (2024), "Thyroid cancer-specific mortality during 2005–2018 in Korea, aftermath of the overdiagnosis issue: a nationwide population-based cohort study", *International Journal of Surgery*, Vol. 110/9, pp. 5489-5495, https://doi.org/10.1097/js9.0000000000001767. [266]

Kim, W. (2024), "Trends in Thyroid Cancer Mortality Rates in Korea: Insights from National Health Database", *Endocrinology and Metabolism*, Vol. 39/6, pp. 853-855, https://doi.org/10.3803/enm.2024.2251. [265]

Kistler, C. et al. (2010), "Patient Perceptions of Mistakes in Ambulatory Care", *Archives of Internal Medicine*, Vol. 170/16, https://doi.org/10.1001/archinternmed.2010.288. [275]

Kline, J. et al. (2020), "Over-Testing for Suspected Pulmonary Embolism in American Emergency Departments", *Circulation: Cardiovascular Quality and Outcomes*, Vol. 13/1, https://doi.org/10.1161/circoutcomes.119.005753. [167]

Kocher, B. and E. Emanuel (2022), "Aligning Incentives for Improving Diagnostic Excellence", *JAMA*, Vol. 327/16, p. 1543, https://doi.org/10.1001/jama.2022.4594. [143]

Krogsbøll, L., K. Jørgensen and P. Gøtzsche (2019), "General health checks in adults for reducing morbidity and mortality from disease", *Cochrane Database of Systematic Reviews*, Vol. 2019/1, https://doi.org/10.1002/14651858.CD009009.pub3. [31]

Krogsbøll, L., K. Jørgensen and P. Gøtzsche (2019), "General health checks in adults for reducing morbidity and mortality from disease", *The Cochrane Database of Systematic Reviews*, Vol. 2019/1, https://doi.org/10.1002/14651858.CD009009.PUB3. [36]

Kuhn, J. et al. (2023), "Teaching medical students to apply deliberate reflection", *Medical Teacher*, Vol. 46/1, pp. 65-72, https://doi.org/10.1080/0142159x.2023.2229504. [305]

Kulasekere, D. et al. (2024), "Appendicitis Hospitalization Care Costs Among Patients With Delayed Diagnosis of Appendicitis", *JAMA Network Open*, Vol. 7/4, p. e246721, https://doi.org/10.1001/jamanetworkopen.2024.6721. [38]

Kulkarni, P. and H. Singh (2023), "Artificial Intelligence in Clinical Diagnosis: Opportunities, Challenges, and Hype", *JAMA*, Vol. 330/4, pp. 317-318, https://doi.org/10.1001/jama.2023.11440. [295]

Lam, D. et al. (2022), "Use of e-triggers to identify diagnostic errors in the paediatric ED", *BMJ Quality & Safety*, Vol. 31/10, pp. 735-743, https://doi.org/10.1136/bmjqs-2021-013683. [54]

Laposata, M. (2022), "DIAGNOSTIC ERROR IN THE UNITED STATES: A SUMMARY OF THE REPORT OF A NATIONAL ACADEMY OF MEDICINE COMMITTEE", *Transactions of the American Clinical and Climatological Association*, Vol. 132, p. 194, https://pmc.ncbi.nlm.nih.gov/articles/PMC9480522/ (accessed on 28 January 2025). [29]

Lebrun, S. and R. Bond (2018), "Spontaneous coronary artery dissection (SCAD): The underdiagnosed cardiac condition that plagues women", *Trends in Cardiovascular Medicine*, Vol. 28/5, pp. 340-345, https://doi.org/10.1016/j.tcm.2017.12.004. [184]

Lee, C. et al. (2013), "Cognitive and System Factors Contributing to Diagnostic Errors in Radiology", *American Journal of Roentgenology*, Vol. 201/3, pp. 611-617, https://doi.org/10.2214/ajr.12.10375. [198]

Li, C. et al. (2020), "Mining Misdiagnosis Patterns from Biomedical Literature", *AMIA Summits on Translational Science Proceedings*, Vol. 2020, p. 360, http:///pmc/articles/PMC7233073/ (accessed on 20 June 2024). [28]

Liberman, A. and D. Newman-Toker (2018), "Symptom-Disease Pair Analysis of Diagnostic Error (SPADE): a conceptual framework and methodological approach for unearthing misdiagnosis-related harms using big data", *BMJ Quality & Safety*, Vol. 27/7, pp. 557-566, https://doi.org/10.1136/bmjqs-2017-007032. [193]

Liberman, A. et al. (2023), "Optimizing measurement of misdiagnosis-related harms using symptom-disease pair analysis of diagnostic error (SPADE): comparison groups to maximize SPADE validity", *Diagnosis*, Vol. 10/3, pp. 225-234, https://doi.org/10.1515/dx-2022-0130. [244]

Lichtenfeld, L. (2011), "Overdiagnosed", *Journal of Clinical Investigation*, Vol. 121/8, pp. 2954-2954, https://doi.org/10.1172/jci57171. [214]

Li, M. et al. (2024), "Evolving epidemiological patterns of thyroid cancer and estimates of overdiagnosis in 2013–17 in 63 countries worldwide: a population-based study", *The Lancet Diabetes & Endocrinology*, doi: 10.1016/S2213-8587(24)00223-7, pp. 824-836, https://doi.org/10.1016/S2213-8587(24)00223-7. [113]

Lim, J. et al. (2020), "The costs of an expanded screening criteria for COVID-19: A modelling study", *International Journal of Infectious Diseases*, Vol. 100, pp. 490-496, https://doi.org/10.1016/j.ijid.2020.08.025. [191]

Li, M. et al. (2023), "The Economic Cost of Thyroid Cancer in France and the Corresponding Share Associated With Treatment of Overdiagnosed Cases", *Value in Health*, Vol. 26/8, pp. 1175-1182, https://doi.org/10.1016/j.jval.2023.02.016. [115]

Lindblom, A. (2014), "Under-detection of autism among First Nations children in British Columbia, Canada", *Disability & Society*, Vol. 29/8, pp. 1248-1259, https://doi.org/10.1080/09687599.2014.923750. [187]

Lindsay, D. et al. (2024), "Estimating the magnitude and healthcare costs of melanoma in situ and thin invasive melanoma overdiagnosis in Australia", *British Journal of Dermatology*, Vol. 191/6, pp. 906-913, https://doi.org/10.1093/bjd/ljae296. [165]

Liu, B. et al. (2024), "Exploring treatment options in cancer: tumor treatment strategies", *Signal Transduction and Targeted Therapy*, Vol. 9/1, https://doi.org/10.1038/s41392-024-01856-7. [179]

Lo, D. et al. (2018), "Misdiagnosis of children's asthma is common in UK primary care and can be improved with objective tests", *Paediatric asthma and allergy*, p. PA1314, https://doi.org/10.1183/13993003.congress-2018.pa1314. [182]

Lortet-Tieulent, J. et al. (2018), "Thyroid cancer "epidemic" also occurs in low- and middle-income countries", *International Journal of Cancer*, Vol. 144/9, pp. 2082-2087, https://doi.org/10.1002/ijc.31884. [163]

Lyratzopoulos, G., P. Vedsted and H. Singh (2015), "Understanding missed opportunities for more timely diagnosis of cancer in symptomatic patients after presentation", *British Journal of Cancer*, Vol. 112/1, pp. S84-S91, https://doi.org/10.1038/bjc.2015.47. [119]

Malik, M. et al. (2022), "A structured approach to EHR surveillance of diagnostic error in acute care: an exploratory analysis of two institutionally-defined case cohorts", *Diagnosis*, Vol. 9/4, pp. 446-457, https://doi.org/10.1515/dx-2022-0032. [239]

Malone, J. et al. (2016), "Justification of CT for Individual Health Assessment of Asymptomatic Persons: A World Health Organization Consultation", *Journal of the American College of Radiology*, Vol. 13/12, pp. 1447-1457.e1, https://doi.org/10.1016/j.jacr.2016.07.020. [34]

Mangione-Smith, R. et al. (1999), "The Relationship Between Perceived Parental Expectations and Pediatrician Antimicrobial Prescribing Behavior", *Pediatrics*, Vol. 103/4, pp. 711-718, https://doi.org/10.1542/peds.103.4.711. [228]

Manos, M., K. Giuliano and E. Geyer (2017), "ADHD: Overdiagnosed and overtreated, or misdiagnosed and mistreated?", *Cleveland Clinic Journal of Medicine*, Vol. 84/11, p. 873, https://doi.org/10.3949/ccjm.84a.15051. [72]

Marino, A. and L. Lorenzoni (2019), "The impact of technological advancements on health spending: A literature review", *OECD Health Working Papers*, No. 113, OECD Publishing, Paris, https://doi.org/10.1787/fa3bab05-en. [8]

Matheny, M. et al. (2025), "Artificial Intelligence In Health And Health Care: Priorities For Action", *Health Affairs*, https://doi.org/10.1377/hlthaff.2024.01003. [282]

McCool, A. et al. (2022), "Antidepressant medication prescribing patterns in Irish general practice from 2016 to 2020 to assess for long-term use", *Irish Journal of Medical Science (1971 -)*, Vol. 191/5, pp. 2239-2246, https://doi.org/10.1007/s11845-021-02833-7. [64]

McCoy, R. et al. (2015), "HbA_{1c}overtesting and overtreatment among US adults with controlled type 2 diabetes, 2001-13: observational population based study", *BMJ*, p. h6138, https://doi.org/10.1136/bmj.h6138. [157]

McCracken, C. et al. (2024), "NHS Health Check attendance is associated with reduced multiorgan disease risk: a matched cohort study in the UK Biobank", *BMC Medicine*, Vol. 22/1, p. 1, https://doi.org/10.1186/s12916-023-03187-w. [32]

McIntyre, R. et al. (2022), "The real-world health resource use and costs of misdiagnosing bipolar I disorder", *Journal of Affective Disorders*, Vol. 316, pp. 26-33, https://doi.org/10.1016/j.jad.2022.07.069. [68]

McPhail, S. et al. (2022), "Risk factors and prognostic implications of diagnosis of cancer within 30 days after an emergency hospital admission (emergency presentation): an International Cancer Benchmarking Partnership (ICBP) population-based study", *The Lancet Oncology*, Vol. 23/5, pp. 587-600, https://doi.org/10.1016/S1470-2045(22)00127-9. [16]

MedTech Europe (2022), "European IVD Market Statistics Report". [47]

Mello, M. et al. (2010), "National Costs Of The Medical Liability System", *Health Affairs*, Vol. 29/9, pp. 1569-1577, https://doi.org/10.1377/hlthaff.2009.0807. [224]

Meyer, A. et al. (2013), "Physicians' Diagnostic Accuracy, Confidence, and Resource Requests", *JAMA Internal Medicine*, Vol. 173/21, p. 1952, https://doi.org/10.1001/jamainternmed.2013.10081. [298]

Meyer, A. and H. Singh (2019), "The Path to Diagnostic Excellence Includes Feedback to Calibrate How Clinicians Think", *JAMA*, Vol. 321/8, p. 737, https://doi.org/10.1001/jama.2019.0113. [299]

Meyer, N. and H. Prescott (2024), "Sepsis and Septic Shock", *New England Journal of Medicine*, Vol. 391/22, pp. 2133-2146. [97]

Mikhaeil, J. et al. (2020), "Quality Assurance Processes Ensuring Appropriate Follow-up of Test Results Pending at Discharge in Emergency Departments: A Systematic Review", *Annals of Emergency Medicine*, Vol. 76/5, pp. 659-674, https://doi.org/10.1016/j.annemergmed.2020.07.024. [273]

Ministère de la Santé (n.d.), "GENOMIC MEDICINE FRANCE 2025". [219]

Moise, I. (ed.) (2017), "Overtreatment in the United States", *PLOS ONE*, Vol. 12/9, p. e0181970, https://doi.org/10.1371/journal.pone.0181970. [171]

Morrow, R. et al. (2012), "Influence of relative age on diagnosis and treatment of attention-deficit/hyperactivity disorder in children", *Canadian Medical Association Journal*, Vol. 184/7, pp. 755-762, https://doi.org/10.1503/cmaj.111619. [169]

Moulder, G., E. Harris and L. Santhosh (2022), "Teaching the science of uncertainty", *Diagnosis*, Vol. 10/1, pp. 13-18, https://doi.org/10.1515/dx-2022-0045. [302]

Moustafa, Y., M. Chauhan and T. Rummans (2022), "Attention-Deficit/Hyperactivity Disorder Overdiagnosis and Overprescriptions: Medicalization of Distractions", *Mayo Clinic Proceedings*, Vol. 97/7, pp. 1339-1344, https://doi.org/10.1016/j.mayocp.2022.02.026. [73]

Moynihan, R. (2011), "A new deal on disease definition", *BMJ*, Vol. 342/may03 1, pp. d2548-d2548, https://doi.org/10.1136/bmj.d2548. [213]

Moynihan, R., J. Doust and D. Henry (2012), "Preventing overdiagnosis: how to stop harming the healthy", *BMJ*, Vol. 344/may28 4, pp. e3502-e3502, https://doi.org/10.1136/bmj.e3502. [172]

Mullainathan, S. and Z. Obermeyer (2021), "Diagnosing Physician Error: A Machine Learning Approach to Low-Value Health Care", *The Quarterly Journal of Economics*, Vol. 137/2, pp. 679-727, https://doi.org/10.1093/qje/qjab046. [253]

Murphy, D. et al. (2013), "Electronic health record-based triggers to detect potential delays in cancer diagnosis", *BMJ Quality & Safety*, Vol. 23/1, pp. 8-16, https://doi.org/10.1136/bmjqs-2013-001874. [234]

Murphy, D. et al. (2018), "Application of electronic trigger tools to identify targets for improving diagnostic safety", *BMJ Quality & Safety*, Vol. 28/2, pp. 151-159, https://doi.org/10.1136/bmjqs-2018-008086. [238]

Müskens, J. et al. (2022), "Overuse of diagnostic testing in healthcare: a systematic review", *BMJ Quality & Safety*, Vol. 31/1, p. 54, https://doi.org/10.1136/BMJQS-2020-012576. [151]

Müskens, J. et al. (2021), "Overuse of diagnostic testing in healthcare: a systematic review", *BMJ Quality & Safety*, Vol. 31/1, pp. 54-63, https://doi.org/10.1136/bmjqs-2020-012576. [23]

Nadpara, P., S. Madhavan and C. Tworek (2015), "Guideline-concordant timely lung cancer care and prognosis among elderly patients in the United States: A population-based study", *Cancer Epidemiology*, Vol. 39/6, pp. 1136-1144, https://doi.org/10.1016/j.canep.2015.06.005. [105]

National Academies (2013), *Best Care at Lower Cost*, National Academies Press, Washington, D.C., https://doi.org/10.17226/13444. [174]

National Quality Forum (2017), *Improving diagnostic quality and safety. Washington*. [301]

Newman-Toker, D. (2014), "A unified conceptual model for diagnostic errors: underdiagnosis, overdiagnosis, and misdiagnosis", *Diagnosis*, Vol. 1/1, pp. 43-48, https://doi.org/10.1515/dx-2013-0027. [21]

Newman-Toker, D. (2009), "Diagnostic Errors—The Next Frontier for Patient Safety", *JAMA*, Vol. 301/10, p. 1060, https://doi.org/10.1001/jama.2009.249. [22]

Newman-Toker, D., K. McDonald and D. Meltzer (2013), "How much diagnostic safety can we afford, and how should we decide? A health economics perspective", *BMJ Quality & Safety*, Vol. 22/Suppl 2, pp. ii11-ii20, https://doi.org/10.1136/bmjqs-2012-001616. [11]

Newman-Toker, D. et al. (2024), "Burden of serious harms from diagnostic error in the USA", *BMJ Qual Saf*, Vol. 33, pp. 109-120, https://doi.org/10.1136/bmjqs-2021-014130. [141]

Newman-Toker, D. et al. (2024), "Burden of serious harms from diagnostic error in the USA", *BMJ Quality & Safety*, Vol. 33/2, p. 109, https://doi.org/10.1136/bmjqs-2021-014130. [60]

Newman-Toker, D. et al. (2023), "Burden of serious harms from diagnostic error in the USA", *BMJ Quality & Safety*, Vol. 33/2, pp. 109-120, https://doi.org/10.1136/bmjqs-2021-014130. [2]

Newman-Toker, D. et al. (2019), "Serious misdiagnosis-related harms in malpractice claims: The "Big Three" – vascular events, infections, and cancers", *Diagnosis*, Vol. 6/3, pp. 227-240, https://doi.org/10.1515/dx-2019-0019. [246]

Newman-Toker, D. et al. (2021), "Rate of diagnostic errors and serious misdiagnosis-related harms for major vascular events, infections, and cancers: toward a national incidence estimate using the "Big Three"", Vol. 8/1, pp. 67-84, https://doi.org/doi:10.1515/dx-2019-0104. [307]

Newman-Toker, D. et al. (2020), "Rate of diagnostic errors and serious misdiagnosis-related harms for major vascular events, infections, and cancers: toward a national incidence estimate using the "Big Three"", *Diagnosis*, Vol. 8/1, pp. 67-84, https://doi.org/10.1515/dx-2019-0104. [309]

Novelli G, B. (2025), *Thyroid cancer monitor. Management of thyroid cancer in Italy: a snapshot of a complex and evolving disease*, ITCO, https://issuu.com/raffaelecreativagroupcom/docs/report_itco_draft (accessed on 26 February 2025). [114]

Oderkirk, J. (2021), "Survey results: National health data infrastructure and governance", *OECD Health Working Papers*, No. 127, OECD Publishing, Paris, https://doi.org/10.1787/55d24b5d-en. [242]

OECD (2025), *Does Healthcare Deliver?: Results from the Patient-Reported Indicator Surveys (PaRIS)*, OECD Publishing, Paris, https://doi.org/10.1787/c8af05a5-en. [248]

OECD (2023), *Embracing a One Health Framework to Fight Antimicrobial Resistance*, OECD Health Policy Studies, OECD Publishing, Paris, https://doi.org/10.1787/ce44c755-en. [99]

OECD (2023), *Health at a Glance 2023: OECD Indicators*, OECD Publishing, Paris, https://doi.org/10.1787/7a7afb35-en. [152]

OECD (2022), *Towards an Integrated Health Information System in Korea*, OECD Publishing, Paris, https://doi.org/10.1787/c4e6c88d-en. [241]

OECD (2020), *Realising the Potential of Primary Health Care*, OECD Health Policy Studies, OECD Publishing, Paris, https://doi.org/10.1787/a92adee4-en. [267]

OECD (2019), *Health in the 21st Century: Putting Data to Work for Stronger Health Systems*, OECD Health Policy Studies, OECD Publishing, Paris, https://doi.org/10.1787/e3b23f8e-en. [46]

OECD (2017), *New Health Technologies: Managing Access, Value and Sustainability*, OECD Publishing, Paris, https://doi.org/10.1787/9789264266438-en. [7]

OECD (2017), *Tackling Wasteful Spending on Health*, OECD Publishing, Paris, https://doi.org/10.1787/9789264266414-en. [24]

OECD (2017), *The Economics of Patient Safety in Primary and Ambulatory Care. Flying blind*. [9]

OECD (2016), *Recommendation of the Council on Health Data Governance*, http://legalinstruments.oecd.org (accessed on 20 April 2020). [247]

Office for National Statistics (ONS), U. (2024), *Self-reported coronavirus (COVID-19) infections and associated symptoms, England and Scotland: November 2023 to March 2024*. [132]

Oliva, F. et al. (2020), "Diagnostic delay in <scp>ADHD</scp>: Duration of untreated illness and its socio-demographic and clinical predictors in a sample of adult outpatients", *Early Intervention in Psychiatry*, Vol. 15/4, pp. 957-965, https://doi.org/10.1111/eip.13041. [77]

Oliveira Hashiguchi, T. (2020), "Bringing health care to the patient: An overview of the use of telemedicine in OECD countries", *OECD Health Working Papers*, No. 116, OECD Publishing, Paris, https://doi.org/10.1787/8e56ede7-en. [279]

Oliveira Hashiguchi, T., L. Slawomirski and J. Oderkirk (2021), "Laying the foundations for artificial intelligence in health", *OECD Health Working Papers*, No. 128, OECD Publishing, Paris, https://doi.org/10.1787/3f62817d-en. [281]

O'Neill, J. (2015), "RAPID DIAGNOSTICS: STOPPING UNNECESSARY USE OF ANTIBIOTICS". [149]

Ong, M. and K. Mandl (2015), "National Expenditure For False-Positive Mammograms And Breast Cancer Overdiagnoses Estimated At $4 Billion A Year", *Health Affairs*, Vol. 34/4, pp. 576-583, https://doi.org/10.1377/hlthaff.2014.1087. [161]

O'Nions, E. et al. (2025), "Life expectancy and years of life lost for adults with diagnosed ADHD in the UK: matched cohort study", *The British Journal of Psychiatry*, pp. 1-8, https://doi.org/DOI: 10.1192/bjp.2024.199. [70]

Oregon Health Study Group (2012), "The Oregon Health Insurance Experiment: Evidence from the First Year*", *The Quarterly Journal of Economics*, Vol. 127/3, pp. 1057-1106, https://doi.org/10.1093/qje/qjs020. [203]

Özçelik, E., S. Chapman and M. Cecchini (2024), "Tackling antimicrobial resistance in Indigenous, rural and remote communities", *OECD Health Working Papers*, No. 172, OECD Publishing, Paris, https://doi.org/10.1787/349bc0fc-en. [259]

Paige, S., J. Krieger and M. Stellefson (2016), "The Influence of eHealth Literacy on Perceived Trust in Online Health Communication Channels and Sources", *Journal of Health Communication*, Vol. 22/1, pp. 53-65, https://doi.org/10.1080/10810730.2016.1250846. [296]

Pettit, N., A. Al-Hader and C. Thompson (2021), "Emergency department associated lung cancer diagnosis: Case series demonstrating poor outcomes and opportunities to improve cancer care", *Current Problems in Cancer: Case Reports*, Vol. 3, p. 100059, https://doi.org/10.1016/J.CPCCR.2021.100059. [20]

Pew (n.d.), *Inappropriate Antibiotic Prescribing for Adults Comes With Increased Risks | The Pew Charitable Trusts*, 2023, https://www.pewtrusts.org/en/research-and-analysis/fact-sheets/2023/01/inappropriate-antibiotic-prescribing-for-adults-comes-with-increased-risks (accessed on 14 January 2025). [146]

Polanyi, M. (1967), *The Tacit Dimension*, Routledge & Kegan Paul, London. [294]

Prasad, V., J. Rho and A. Cifu (2012), "The Diagnosis and Treatment of Pulmonary Embolism", *Archives of Internal Medicine*, Vol. 172/12, https://doi.org/10.1001/archinternmed.2012.195. [168]

PRNewswire (2022), *United States Medical Diagnostics Market Report 2022-2026: Rising Integration with Artificial Intelligence & Growing Awareness on Personalized Medicine*, https://www.prnewswire.com/news-releases/united-states-medical-diagnostics-market-report-2022-2026-rising-integration-with-artificial-intelligence--growing-awareness-on-personalized-medicine-301498981.html (accessed on 18 December 2024). [41]

Pruitt, S. et al. (2013), "Do diagnostic and treatment delays for colorectal cancer increase risk of death?", *Cancer Causes & Control*, Vol. 24/5, pp. 961-977, https://doi.org/10.1007/s10552-013-0172-6. [103]

Puliti, D. et al. (2012), "Overdiagnosis in Mammographic Screening for Breast Cancer in Europe: A Literature Review", *Journal of Medical Screening*, doi: 10.1258/jms.2012.012082, pp. 42-56, https://doi.org/10.1258/jms.2012.012082. [107]

Rao, P. and M. Place (2011), "Prevalence of ADHD in four general adult outpatient clinics in North East England", *Progress in Neurology and Psychiatry*, Vol. 15/5, pp. 7-10, https://doi.org/10.1002/pnp.211. [80]

Read, J. et al. (2019), "Staying on, and coming off, antidepressants: The experiences of 752 UK adults", *Addictive Behaviors*, Vol. 88, pp. 82-85, https://doi.org/10.1016/j.addbeh.2018.08.021. [89]

Read, J. et al. (2020), "A survey of UK general practitioners about depression, antidepressants and withdrawal: implementing the 2019 Public Health England report", *Therapeutic Advances in Psychopharmacology*, doi: 10.1177/2045125320950124, p. 2045125320950124, https://doi.org/10.1177/2045125320950124. [88]

Romanello, M. et al. (2023), "The 2023 report of the Lancet Countdown on health and climate change: the imperative for a health-centred response in a world facing irreversible harms", *The Lancet*, Vol. 402/10419, pp. 2346-2394, https://doi.org/10.1016/S0140-6736(23)01859-7. [194]

Rothberg, M. et al. (2014), "The Cost of Defensive Medicine on 3 Hospital Medicine Services", *JAMA Internal Medicine*, Vol. 174/11, p. 1867, https://doi.org/10.1001/jamainternmed.2014.4649. [226]

Royce, C., M. Hayes and R. Schwartzstein (2019), "Teaching Critical Thinking: A Case for Instruction in Cognitive Biases to Reduce Diagnostic Errors and Improve Patient Safety", *Academic Medicine*, Vol. 94/2, pp. 187-194, https://doi.org/10.1097/acm.0000000000002518. [306]

Roy, C. et al. (2005), "Patient Safety Concerns Arising from Test Results That Return after Hospital Discharge", *Annals of Internal Medicine*, Vol. 143/2, p. 121, https://doi.org/10.7326/0003-4819-143-2-200507190-00011. [271]

Rudd, K. et al. (2020), "Global, regional, and national sepsis incidence and mortality, 1990–2017: analysis for the Global Burden of Disease Study", *The Lancet*, Vol. 395/10219, pp. 200-211. [94]

Ryser, M. et al. (2022), "Estimation of Breast Cancer Overdiagnosis in a U.S. Breast Screening Cohort", *Annals of Internal Medicine*, doi: 10.7326/M21-3577, pp. 471-478, https://doi.org/10.7326/M21-3577. [109]

Saha, S. and S. Saha (2021), "The impact of the undetected COVID-19 cases on its transmission dynamics", *Indian Journal of Pure and Applied Mathematics*, Vol. 52/4, pp. 1229-1234, https://doi.org/10.1007/s13226-021-00035-6. [190]

Sajid, I., A. Parkunan and K. Frost (2021), "Unintended consequences: quantifying the benefits, iatrogenic harms and downstream cascade costs of musculoskeletal MRI in UK primary care", *BMJ Open Quality*, Vol. 10/3, p. e001287, https://doi.org/10.1136/bmjoq-2020-001287. [150]

Saposnik, G. et al. (2016), "Cognitive biases associated with medical decisions: a systematic review", *BMC Medical Informatics and Decision Making*, Vol. 16/1, pp. 1-14, https://doi.org/10.1186/S12911-016-0377-1. [197]

Schein, J. et al. (2022), "Economic burden of attention-deficit/hyperactivity disorder among adults in the United States: a societal perspective", *Journal of Managed Care & Specialty Pharmacy*, Vol. 28/2, pp. 168-179, https://doi.org/10.18553/jmcp.2021.21290. [79]

Schiff, G. (2013), "Diagnosis and diagnostic errors: time for a new paradigm", *BMJ Quality & Safety*, Vol. 23/1, pp. 1-3, https://doi.org/10.1136/bmjqs-2013-002426. [236]

Schiff, G. (2008), "Minimizing Diagnostic Error: The Importance of Follow-up and Feedback", *The American Journal of Medicine*, Vol. 121/5, pp. S38-S42, https://doi.org/10.1016/j.amjmed.2008.02.004. [230]

Schulman, K. and S. Tunis (2010), "A policy approach to the development of molecular diagnostic tests", *Nature Biotechnology*, Vol. 28/11, pp. 1157-1159, https://doi.org/10.1038/nbt1110-1157. [44]

Schwartz, A. et al. (2015), "Changes in Low-Value Services in Year 1 of the Medicare Pioneer Accountable Care Organization Program", *JAMA Internal Medicine*, Vol. 175/11, p. 1815, https://doi.org/10.1001/jamainternmed.2015.4525. [254]

Schwartz, A. et al. (2012), "Uncharted territory: Measuring costs of diagnostic errors outside the medical record", *BMJ Quality and Safety*, Vol. 21/11, pp. 918-924, https://doi.org/10.1136/BMJQS-2012-000832. [53]

Schwartz, L. and S. Woloshin (1999), "Changing disease definitions: implications for disease prevalence. Analysis of the Third National Health and Nutrition Examination Survey, 1988-1994", *Effective clinical practice : ECP*, Vol. 2/2, pp. 76-85, https://pubmed.ncbi.nlm.nih.gov/10538480/ (accessed on 14 January 2025). [207]

Seyyed-Kalantari, L. et al. (2021), "Underdiagnosis bias of artificial intelligence algorithms applied to chest radiographs in under-served patient populations", *Nature Medicine*, Vol. 27/12, pp. 2176-2182, https://doi.org/10.1038/s41591-021-01595-0. [283]

Shaik, T. et al. (2024), "Lab testing overload: a comprehensive analysis of overutilization in hospital-based settings", *Proceedings (Baylor University. Medical Center)*, Vol. 37/2, p. 312, https://doi.org/10.1080/08998280.2023.2288788. [156]

Shaw, T. et al. (2023), "Missed diagnosis or misdiagnosis: common pitfalls in genetic testing", *Singapore Medical Journal*, Vol. 64/1, pp. 67-73, https://doi.org/10.4103/singaporemedj.smj-2021-467. [222]

Sheridan, S. et al. (2020), "What If?: Transforming Diagnostic Research by Leveraging a Diagnostic Process Map to Engage Patients in Learning from Errors", *NAM Perspectives*, https://doi.org/10.31478/202002a. [277]

Shih, P. et al. (2022), "Direct-to-consumer detection of atrial fibrillation in a smartwatch electrocardiogram: Medical overuse, medicalisation and the experience of consumers", *Social Science & Medicine*, Vol. 303, p. 114954, https://doi.org/10.1016/j.socscimed.2022.114954. [217]

Shreffler, J. and M. Huecker (2023), "Diagnostic Testing Accuracy: Sensitivity, Specificity, Predictive Values and Likelihood Ratios", *StatPearls*, https://www.ncbi.nlm.nih.gov/books/NBK557491/ (accessed on 14 June 2024). [14]

Shriner, A. et al. (2021), "Improving Follow-Up of Tests Pending at Discharge", *Hospital Pediatrics*, Vol. 11/12, pp. 1363-1369, https://doi.org/10.1542/hpeds.2021-006000. [268]

Sik, A., K. Jung and W. Gilbert (2025), "Korea's Thyroid-Cancer "Epidemic" — Screening and Overdiagnosis", *New England Journal of Medicine*, doi: 10.1056/NEJMp1409841, pp. 1765-1767, https://doi.org/10.1056/NEJMp1409841. [315]

Singer, M. et al. (2016), "The third international consensus definitions for sepsis and septic shock (Sepsis-3)", *Jama*, Vol. 315/8, pp. 801-810. [92]

Singh, H. et al. (2009), "Missed Opportunities to Initiate Endoscopic Evaluation for Colorectal Cancer Diagnosis", *Official journal of the American College of Gastroenterology | ACG*, Vol. 104/10, https://journals.lww.com/ajg/fulltext/2009/10000/missed_opportunities_to_initiate_endoscopic.23.aspx. [102]

Singh, H. et al. (2011), "Electronic health record-based surveillance of diagnostic errors in primary care", *BMJ Quality & Safety*, Vol. 21/2, pp. 93-100, https://doi.org/10.1136/bmjqs-2011-000304. [235]

Singh, H. et al. (2013), "Types and Origins of Diagnostic Errors in Primary Care Settings", *JAMA Internal Medicine*, Vol. 173/6, pp. 418-425, https://doi.org/10.1001/jamainternmed.2013.2777. [30]

Singh, H., M. Graber and T. Hofer (2019), "Measures to Improve Diagnostic Safety in Clinical Practice", *Journal of Patient Safety*, Vol. 15/4, https://journals.lww.com/journalpatientsafety/fulltext/2019/12000/measures_to_improve_diagnostic_safety_in_clinical.15.aspx. [232]

Singh, H. et al. (2010), "Characteristics and Predictors of Missed Opportunities in Lung Cancer Diagnosis: An Electronic Health Record–Based Study", *Journal of Clinical Oncology*, Vol. 28/20, pp. 3307-3315, https://doi.org/10.1200/JCO.2009.25.6636. [104]

Singh, H., A. Meyer and E. Thomas (2014), "The frequency of diagnostic errors in outpatient care: estimations from three large observational studies involving US adult populations", *BMJ Quality & Safety*, Vol. 23/9, pp. 727-731, https://doi.org/10.1136/bmjqs-2013-002627. [140]

Singh, H., A. Meyer and E. Thomas (2014), "The frequency of diagnostic errors in outpatient care: estimations from three large observational studies involving US adult populations", *BMJ Quality & Safety*, Vol. 23/9, p. 727, https://doi.org/10.1136/bmjqs-2013-002627. [5]

Singh, H. and D. Sittig (2015), "Advancing the science of measurement of diagnostic errors in healthcare: the Safer Dx framework", *BMJ Quality & Safety*, Vol. 24/2, p. 103, https://doi.org/10.1136/bmjqs-2014-003675. [40]

Singh, H. and D. Sittig (2015), "Advancing the science of measurement of diagnostic errors in healthcare: the Safer Dx framework", *BMJ Quality & Safety*, Vol. 24/2, pp. 103-110, https://doi.org/10.1136/bmjqs-2014-003675. [202]

Singh, H. et al. (2009), "Improving follow-up of abnormal cancer screens using electronic health records: trust but verify test result communication", *BMC Medical Informatics and Decision Making*, Vol. 9/1, https://doi.org/10.1186/1472-6947-9-49. [201]

Slawomirski, L., A. Auraaen and N. Klazinga (2017), *THE ECONOMICS OF PATIENT SAFETY Strengthening a value-based approach to reducing patient harm at national level*. [256]

Slawomirski, L. and N. Klazinga (2022), "The economics of patient safety: From analysis to action", *OECD Health Working Papers*, No. 145, OECD Publishing, Paris, https://doi.org/10.1787/761f2da8-en. [58]

Slawomirski, L. et al. (2023), "Progress on implementing and using electronic health record systems: Developments in OECD countries as of 2021", *OECD Health Working Papers*, No. 160, OECD Publishing, Paris, https://doi.org/10.1787/4f4ce846-en. [240]

Statista (2023), *Average cost genetic test for precision medicine worldwide by region 2027 | Statista*, https://www.statista.com/statistics/1420932/average-cost-genetic-test-used-for-precision-medicine-by-world-region/ (accessed on 18 December 2024). [220]

Stein, D. et al. (2010), "What is a mental/psychiatric disorder? From DSM-IV to DSM-V", *Psychological Medicine*, Vol. 40/11, pp. 1759-1765, https://doi.org/10.1017/s0033291709992261. [212]

Taylor, H. et al. (2023), "Economic Burden Associated With Untreated Mental Illness in Indiana", *JAMA Health Forum*, Vol. 4/10, p. e233535, https://doi.org/10.1001/jamahealthforum.2023.3535. [65]

te Marvelde, L. et al. (2019), "Alignment with Indices of A Care Pathway Is Associated with Improved Survival: An Observational Population-based Study in Colon Cancer Patients", *EClinicalMedicine*, Vol. 15, pp. 42-50, https://doi.org/10.1016/j.eclinm.2019.08.009. [19]

The Community of Population-Based Regional Cancer Registries* (2023), "Cancer Statistics in Korea: Incidence, Mortality, Survival, and Prevalence in 2020", *Cancer Research and Treatment*, Vol. 55/2, pp. 385-399, https://doi.org/10.4143/crt.2023.447. [264]

The International Agency for Research on Cancer, L. (2023), *Best practices in cervical screening programmes: Audit of cancers, legal and ethical frameworks, communication, and workforce competencies*, https://www.ncbi.nlm.nih.gov/books/NBK597783/ (accessed on 3 January 2025). [128]

Thomaides-Brears, H. et al. (2022), "Incidence of Complications from Percutaneous Biopsy in Chronic Liver Disease: A Systematic Review and Meta-Analysis", *Digestive Diseases and Sciences*, Vol. 67/7, pp. 3366-3394, https://doi.org/10.1007/S10620-021-07089-W/TABLES/4. [159]

Thomas, J., E. Ziller and D. Thayer (2010), "Low Costs Of Defensive Medicine, Small Savings From Tort Reform", *Health Affairs*, Vol. 29/9, pp. 1578-1584, https://doi.org/10.1377/hlthaff.2010.0146. [225]

Tipton, K. et al. (2023), *Impact of Healthcare Algorithms on Racial and Ethnic Disparities in Health and Healthcare*, Agency for Healthcare Research and Quality (AHRQ), https://doi.org/10.23970/ahrqepccer268. [284]

Tit Albreht, R. (2017), *Policy recommendations on governance, organization and evaluation of cancer screening*, European Guide on Quality Improvement in Comprehensive Cancer Control, https://cancercontrol.eu/archived/guide-landing-page/index.html (accessed on 3 January 2025). [127]

Tyson, A. et al. (2023), *60% of Americans Would Be Uncomfortable With Provider Relying on AI in Their Own Health Care*, https://www.pewresearch.org/science/2023/02/22/60-of-americans-would-be-uncomfortable-with-provider-relying-on-ai-in-their-own-health-care/ (accessed on 19 December 2024). [297]

Vaccarella, S. et al. (2016), "Worldwide Thyroid-Cancer Epidemic? The Increasing Impact of Overdiagnosis", *New England Journal of Medicine*, Vol. 375/7, pp. 614-617, https://doi.org/10.1056/nejmp1604412. [162]

Vaccarella, S. et al. (2024), "Prostate cancer incidence and mortality in Europe and implications for screening activities: population based study", *BMJ*, Vol. 386, p. e077738, https://doi.org/10.1136/bmj-2023-077738. [112]

Valdivieso-Martínez, B. et al. (2023), "Impact of Long-COVID on Health Care Burden: A Case Control Study", *Journal of Clinical Medicine*, Vol. 12/18, p. 5768, https://doi.org/10.3390/jcm12185768. [308]

Van den Berg, M. et al. (2022), "Hospital-related costs of sepsis around the world: A systematic review exploring the economic burden of sepsis", *Journal of Critical Care*, Vol. 71, p. 154096. [93]

van der Veer, E. et al. (2023), "Causes and consequences of delayed diagnosis in breast cancer screening with a focus on mammographic features and tumour characteristics", *European Journal of Radiology*, Vol. 167, p. 111048, https://doi.org/10.1016/j.ejrad.2023.111048. [117]

Verhagen, M. et al. (2022), "The problem with making Safety-II work in healthcare", *BMJ Quality & Safety*, Vol. 31/5, pp. 402-408, https://doi.org/10.1136/bmjqs-2021-014396. [57]

Versluis, A. et al. (2022), "Direct Access for Patients to Diagnostic Testing and Results Using eHealth: Systematic Review on eHealth and Diagnostics", *Journal of Medical Internet Research*, Vol. 24/1, p. e29303, https://doi.org/10.2196/29303. [50]

Visser, S. et al. (2014), "Trends in the Parent-Report of Health Care Provider-Diagnosed and Medicated Attention-Deficit/Hyperactivity Disorder: United States, 2003–2011", *Journal of the American Academy of Child & Adolescent Psychiatry*, Vol. 53/1, pp. 34-46.e2, https://doi.org/10.1016/j.jaac.2013.09.001. [71]

Vrijsen, B. et al. (2020), "Inappropriate laboratory testing in internal medicine inpatients: Prevalence, causes and interventions", *Annals of Medicine and Surgery*, Vol. 51, p. 48, https://doi.org/10.1016/J.AMSU.2020.02.002. [155]

Wang, Y. et al. (2025), "Refinement of post-COVID condition core symptoms, subtypes, determinants, and health impacts: a cohort study integrating real-world data and patient-reported outcomes", *eBioMedicine*, doi: 10.1016/j.ebiom.2024.105493, https://doi.org/10.1016/j.ebiom.2024.105493. [135]

Weibel, S. et al. (2020), "Practical considerations for the evaluation and management of Attention Deficit Hyperactivity Disorder (ADHD) in adults", *L'Encéphale*, Vol. 46/1, pp. 30-40, https://doi.org/10.1016/j.encep.2019.06.005. [81]

Welch, H. and W. Black (2010), "Overdiagnosis in Cancer", *JNCI Journal of the National Cancer Institute*, Vol. 102/9, pp. 605-613, https://doi.org/10.1093/jnci/djq099. [166]

Welch, H. and E. Fisher (2017), "Income and Cancer Overdiagnosis — When Too Much Care Is Harmful", *New England Journal of Medicine*, Vol. 376/23, pp. 2208-2209, https://doi.org/10.1056/nejmp1615069. [177]

Whitehead, N. et al. (2018), "Interventions to Improve Follow-Up of Laboratory Test Results Pending at Discharge: A Systematic Review", *Journal of Hospital Medicine*, Vol. 13/9, pp. 631-636, https://doi.org/10.12788/jhm.2944. [270]

WHO (2023), *Global Patient Safety Action Plan 2021-2030*, https://www.who.int/teams/integrated-health-services/patient-safety/policy/global-patient-safety-action-plan (accessed on 2 June 2021). [249]

WHO (2017), *Use of medical imaging in asymptomatic people for individual health assessment (IHA)*, https://www.who.int/publications/m/item/use-of-medical-imaging-in-asymptomatic-people-for-individual-health-assessment-(iha) (accessed on 31 January 2025). [35]

Williams, W., J. Hogan and J. Ingelfinger (2021), "Time to Eliminate Health Care Disparities in the Estimation of Kidney Function", *New England Journal of Medicine*, Vol. 385/19, pp. 1804-1806, https://doi.org/10.1056/nejme2114918. [285]

Winearls, C. and R. Glassock (2011), "Classification of Chronic Kidney Disease in the Elderly: Pitfalls and Errors", *Nephron Clinical Practice*, Vol. 119/Suppl. 1, pp. c2-c4, https://doi.org/10.1159/000328013. [211]

Winters, B. et al. (2012), "Diagnostic errors in the intensive care unit: a systematic review of autopsy studies", *BMJ Quality & Safety*, Vol. 21/11, pp. 894-902, https://doi.org/10.1136/bmjqs-2012-000803. [145]

World Health Organization (2024), *World Patient Safety Day 2024*, https://www.who.int/campaigns/world-patient-safety-day/world-patient-safety-day-2024. [250]

World Health Organization (2022), *Post COVID-19 condition (Long COVID)*, https://www.who.int/europe/news-room/fact-sheets/item/post-covid-19-condition#:~:text=Definition,months%20with%20no%20other%20explanation. (accessed on 4 November 2024). [136]

World Health Organization (2016), *Technical Series on Safer Primary Care: Diagnostic errors*, https://www.who.int/publications/i/item/9789241511636, Geneva, https://www.who.int/publications/i/item/9789241511636 (accessed on 17 January 2025). [90]

Yi, K. (2016), "The Revised 2016 Korean Thyroid Association Guidelines for Thyroid Nodules and Cancers: Differences from the 2015 American Thyroid Association Guidelines", *Endocrinology and Metabolism*, Vol. 31/3, p. 373, https://doi.org/10.3803/enm.2016.31.3.373. [262]

Yi, K. et al. (2015), "The Korean guideline for thyroid cancer screening", *Journal of the Korean Medical Association*, Vol. 58/4, p. 302, https://doi.org/10.5124/jkma.2015.58.4.302. [261]

Zeidan, J. et al. (2022), "Global prevalence of autism: A systematic review update", *Autism Research*, Vol. 15/5, pp. 778-790, https://doi.org/10.1002/aur.2696. [84]

Zhou, Y. et al. (2016), "Diagnosis of cancer as an emergency: a critical review of current evidence", *Nature Reviews Clinical Oncology 2016 14:1*, Vol. 14/1, pp. 45-56, https://doi.org/10.1038/nrclinonc.2016.155. [15]

Zimolzak, A. et al. (2024), "Machine Learning to Enhance Electronic Detection of Diagnostic Errors", *JAMA Network Open*, Vol. 7/9, pp. e2431982-e2431982, https://doi.org/10.1001/jamanetworkopen.2024.31982. [237]

Zwaan, L. et al. (2010), "Patient record review of the incidence, consequences, and causes of diagnostic adverse events", *Archives of internal medicine*, Vol. 170/12, pp. 1015-1021, https://doi.org/10.1001/ARCHINTERNMED.2010.146. [199]

Zwaan, L. and H. Singh (2015), "The challenges in defining and measuring diagnostic error", Vol. 2/2, pp. 97-103, https://doi.org/doi:10.1515/dx-2014-0069. [51]

OECD Health Working Papers

A full list of the papers in this series can be found on the OECD website: Reports and research papers

No. 175 – SMART SPENDING TO COMBAT GLOBAL HEALTH THREATS (March 2025) David Morgan, Chris James, Kerri Elgar, Caroline Penn and Yasmin Ahmad

No. 174 – ANALYSE DE LA CROISSANCE DES DÉPENSES DE SANTÉ EN FRANCE ET EN ALLEMAGNE (January 2025) Michael Mueller, Ricarda Milstein, Pauline Fron, David Morgan, Luca Lorenzoni

No. 173 – LEADING PRACTICES FOR THE FUTURE OF TELEMEDICINE (January 2025) Rishub Keelara, Eric Sutherland, Margarita Almyranti

No. 172 – TACKLING ANTIMICROBIAL RESISTANCE IN INDIGENOUS, RURAL AND REMOTE COMMUNITIES (September 2024) Ece Özçelik, Suzannah Chapman and Michele Cecchini

No. 171 – EXPLORINGTHE FEASIBILITY OF SHARING INFORMATION ON MEDICINE PRICES ACROSS COUNTRIES (September 2024) Marjolijn Moens, Eliana Barrenho and Valérie Paris

No. 170 – ACCESS TO ONCOLOGY MEDICINES IN EU AND OECD COUNTRIES (September 2024) Thomas Hofmarcher, Caroline Berchet and Guillaume Dedet

No. 169 - DIGITAL AND INNOVATIVE TOOLS FOR BETTER HEALTH AND PRODUCTIVITY AT THE WORKPLACE (August 2024) Pedro Isaac Vazquez-Venegas, Marion Devaux, Hikaru Aihara and Michele Cecchini

No. 168 - COMPARATIVE ASSESSMENT OF PATIENT SAFETY CULTURE PERFORMANCE IN OECD COUNTRIES (July 2024) Katherine de Bienassis and Niek Klazinga

No. 167 - THE IMPACTS OF LONG COVID ACROSS OECD COUNTRIES (June 2024) Ana Espinosa Gonzalez and Elina Suzuki

No. 166 - PARIS FIELD TRIAL REPORT: TECHNICAL REPORT ON THE CONDUCT AND RESULTS OF THE FIELD TRIAL OF THE INTERNATIONAL PARIS SURVEY OF PEOPLE LIVING WITH CHRONIC CONDITIONS (April 2024) Michael van den Berg, Candan Kendir, Diana Castelblanco, Nicolas Larrain, Frederico Guanais, Oliver Groene, Pilar Illarramendi, Jose Maria Valderas, Rachel Williams and Mieke Rijken

No. 165 - ADVANCED PRACTICE NURSING IN PRIMARY CARE IN OECD COUNTRIES: RECENT DEVELOPMENTS AND PERSISTING IMPLEMENTATION CHALLENGES (April 2024) Gaetan Lafortune and Ian Brownwood

No. 164 - FAST-TRACK ON DIGITAL SECURITY IN HEALTH (November 2023) Eric Sutherland, Rishub Keelara, Samuel Eiszele and June Haugrud

Recent related OECD publications

DOES HEALTHCARE DELIVER? (February 2025)

TACKLING THE IMPACT OF CANCER ON HEALTH, THE ECONOMY AND SOCIETY (November 2024)

HEALTH AT A GLANCE: ASIA/PASIFIC 2024 (November 2024)

HEALTH AT A GLANCE: EUROPE 2024 (November 2024)

IS CARE AFFORDABLE FOR OLDER PEOPLE? (October 2024)

TOBACCO TAXATION IN LATIN AMERICA AND THE CARIBBEAN (October 2024)

STRENGTHENING PANDEMIC PREPAREDNESS AND RESPONSE THROUGH INTEGRATED MODELLING (May 2024)

OECD HEALTH STATISTICS (2024) – July 2024. The 2024 edition of OECD Health Statistics is now available on OECD Data Explorer OECD Health Statistics 2024 - OECD

STRENGTHENING HEALTH SYSTEMS. A PRACTICAL HANDBOOK FOR RESILIENCE TESTING (March 2024)

SECURING MEDICAL SUPPLY CHAINS IN A POST-PANDEMIC WORLD (February 2024)

BEATING CANCER INEQUALITIES IN THE EU: SPECIAL FOCUS ON CANCER PREVENTION AND EARLY DETECTION (February 2024)

RETHINKING HEALTH SYSTEM PERFORMANCE ASSESSMENT: A RENEWED FRAMEWORK (January 2024)

COLLECTIVE ACTION FOR RESPONSIBLE AI IN HEALTH (January 2024)

FISCAL SUSTAINABILITY OF HEALTH SYSTEMS- HOW TO FINANCE MORE RESILIENT HEALTH SYSTEMS WHEN MONEY IS TIGHT? (January 2024)

EU COUNTRY HEALTH PROFILES 2023 (December 2023)

EVALUATION OF BELGIUM'S COVID-19 RESPONSES - FOSTERING TRUST FOR A MORE RESILIENT SOCIETY (December 2023)

HEALTH AT A GLANCE 2023 – OECD INDICATORS (November 2023)

PURCHASING FOR QUALITY CHRONIC CARE - SUMMARY REPORT (October 2023)

EMBRACING A ONE HEALTH FRAMEWORK TO FIGHT ANTIMICROBIAL RESISTANCE (September 2023)

For a full list, consult the OECD health web page