

Bridging theory and practice – implementation insights on artificial intelligence in health care

*Report of Working Group 1 of the Strategic
Partners' Initiative for Data and Digital Health*

Bridging theory and practice – implementation insights on artificial intelligence in health care

Report of Working Group 1 of the Strategic
Partners' Initiative for Data and Digital Health

Abstract:

This report represents the final deliverable of Working Group 1 of the WHO Regional Office for Europe's Strategic Partners' Initiative for Data and Digital Health addressing the topic of implementation insights on artificial intelligence in health care.

KEYWORDS:

DIGITAL HEALTH; DIGITAL TECHNOLOGY; ARTIFICIAL INTELLIGENCE; MEDICAL INFORMATICS; HEALTH INFORMATION SYSTEMS

Document number: WHO/EURO:2026-13225-52999-82657 (PDF)

© World Health Organization 2026

Some rights reserved. This work is available under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 IGO licence (CC BY-NC-SA 3.0 IGO; <https://creativecommons.org/licenses/by-nc-sa/3.0/igo>).

Under the terms of this licence, you may copy, redistribute and adapt the work for non-commercial purposes, provided the work is appropriately cited, as indicated below. In any use of this work, there should be no suggestion that WHO endorses any specific organization, products or services. The use of the WHO logo is not permitted. If you adapt the work, then you must license your work under the same or equivalent Creative Commons licence. If you create a translation of this work, you should add the following disclaimer along with the suggested citation: "This translation was not created by the World Health Organization (WHO). WHO is not responsible for the content or accuracy of this translation. The original English edition shall be the binding and authentic edition: Bridging theory and practice – implementation insights on artificial intelligence in health care: report of Working Group 1 of the Strategic Partners' Initiative for Data and Digital Health. Copenhagen: WHO Regional Office for Europe; 2026".

Any mediation relating to disputes arising under the licence shall be conducted in accordance with the mediation rules of the World Intellectual Property Organization (<http://www.wipo.int/amc/en/mediation/rules/>).

Suggested citation. Bridging theory and practice – implementation insights on artificial intelligence in health care: report of Working Group 1 of the Strategic Partners' Initiative for Data and Digital Health. Copenhagen: WHO Regional Office for Europe; 2026. Licence: CC BY-NC-SA 3.0 IGO.

Cataloguing-in-Publication (CIP) data. CIP data are available at <http://apps.who.int/iris>.

Sales, rights and licensing. To purchase WHO publications, see <http://apps.who.int/bookorders>. To submit requests for commercial use and queries on rights and licensing, see <https://www.who.int/about/policies/publishing/copyright>

Third-party materials. If you wish to reuse material from this work that is attributed to a third party, such as tables, figures or images, it is your responsibility to determine whether permission is needed for that reuse and to obtain permission from the copyright holder. The risk of claims resulting from infringement of any third-party-owned component in the work rests solely with the user.

General disclaimers. The designations employed and the presentation of the material in this publication do not imply the expression of any opinion whatsoever on the part of WHO concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. Dotted and dashed lines on maps represent approximate border lines for which there may not yet be full agreement.

The mention of specific companies or of certain manufacturers' products does not imply that they are endorsed or recommended by WHO in preference to others of a similar nature that are not mentioned. Errors and omissions excepted, the names of proprietary products are distinguished by initial capital letters.

All reasonable precautions have been taken by WHO to verify the information contained in this publication. However, the published material is being distributed without warranty of any kind, either expressed or implied. The responsibility for the interpretation and use of the material lies with the reader. In no event shall WHO be liable for damages arising from its use.

Designed by: Kristine Onarheim

Contents

Acknowledgements	v
Abbreviations	vi
Executive summary	vii
Background.....	1
1. Introduction	2
2. Governance and key regulatory and legislative considerations	4
3. Frameworks for AI categorization.....	10
4. Case series	13
5. Discussion and cross-case insights	32
6. Policy considerations	34
References	37

Acknowledgements

The WHO Regional Office for Europe would like to extend its sincere gratitude to the members of SPI-DDH Working Group 1 whose exceptional dedication, commitment and expertise were instrumental in the development of this report. This working group was led by Esti Shelly, Ministry of Health, Israel (Working Group 1 Chair), Elena Petelos, Health Technology Assessment International (Working Group 1 Co-Chair) and Adrian Jonas, National Health Service England (North-West Region), United Kingdom (Working Group 1 Co-Chair). Other members of the group (in alphabetical order with affiliations as at commencement of membership) were Liliana Buzdugan, National Agency for Public Health, Republic of Moldova; Šeila Cilović Lagarija, Institute for Public Health, Bosnia and Herzegovina; Megan Coder, Digital Medicine Society; Leah Dowdall, Department of Health, Ireland; Miriam El-Mahdi, Young Experts: Tech for Health; Javier Elkin, International Committee of the Red Cross; Rachel Fellner, Organisation for Economic Co-operation and Development; Giuseppe Giovanni Gatt, EBO.AI; Ketevan Goginashvili, Ministry of Internally Displaced Persons from the Occupied Territories, Labour, Health and Social Affairs, Georgia; Nigel Green, Healthsum; Oliver Gröne, OptiMedis; Zdravko Grubač, Ministry of Health and Social Welfare, Republic of Srpska, Bosnia and Herzegovina; Maria Hägglund, Uppsala University, Sweden; Marianna Imenokhoeva, Health Tech Without Borders; Chiaki Ito, International Organization for Migration; Ivan Ivanovic, Institute of Public Health, Serbia; Nora Markova, United Nations Industrial Development Organization; Sebastiaan Meijer, KTH Royal Institute of Technology, Sweden; Mamura Mirakhmedova, Ministry of Health, Uzbekistan; Diogo Nunes, Portuguese Shared Services for the Ministry of Health, Portugal; Ronan O'Connor, Healthcare Information and Management Systems Society Europe; Mert Özcan, Ministry of Health, Türkiye; Sari Palojoki, Ministry of Social Affairs and Health, Finland; Terje Peetso, North Estonia Medical Centre Foundation, Estonia; Farkhod Rakhimkulov, Ministry of Health, Uzbekistan; Carmen Pauline Rios Benton, Doctomatic; Yury Rogachev, University of Oslo, Norway; Elettra Ronchi, Sciences Po School of Public Affairs, France; Philip Russmeyer, Fitfile; Marlies Schijven, Amsterdam University Medical Centers, Netherlands (Kingdom of the); Özgür Sezer, Ministry of Health, Türkiye; Mahsa Shabani, University of Amsterdam, Netherlands (Kingdom of the); Marina Shakh-Nazarova, National Centre for Disease Control and Public Health, Georgia; Cátia Sousa Pinto, Shared Services for the Ministry of Health, Portugal; Eric Sutherland, Organisation for Economic Co-operation and Development; Miklós Szócska, Semmelweis University, Hungary; Natasa Terzic, Institute for Public Health, Montenegro; Tomi Thomo, Ministry of Health and Social Protection, Albania; Laurie Werner, PATH; Hal Wolf, Healthcare Information and Management Systems Society; Filip Wolf, Ministry of Health, Serbia; Peiling Yap, HealthAI; Ruslana Yegizbayeva, Ministry of Healthcare, Kazakhstan; Monica Åberg Yngwe, EIT Health.

Members of the WHO Secretariat supporting activities of the group and facilitating development of this report were Clayton Hamilton (WHO Regional Office for Europe), Indra Joshi (Consultant to the WHO Regional Office for Europe, Copenhagen, Denmark from January 2025 to December 2025), Giovanni Monti (Consultant to the WHO Regional Office for Europe, Copenhagen, Denmark from August 2024 to December 2024), Tyrone Reden Sy (WHO Regional Office for Europe) and Stein Olav Skrivseth (Consultant to the WHO Regional Office for Europe, Copenhagen, Denmark from March 2025 to December 2025).

The WHO Regional Office for Europe would like to thank the following internal reviewers of this document: Ivo Rakovac, Anna Sagan and Yerbol Spatayev, all from the WHO Regional Office for Europe and Sameer Pujari from WHO headquarters as well as external reviewers: Axel Kaehne, Edgehill University, United Kingdom; Giovanni Monti, Option 5 Health and Ran Milman, Digital Health Division, Israeli Ministry of Health.

The SPI-DDH is facilitated by the Data, Artificial Intelligence and Digital Health Unit in the Division of Health Systems under the overall technical conceptualization of Natasha Azzopardi-Muscat and David Novillo Ortiz both from the WHO Regional Office for Europe.

This publication was co-funded by the European Union. Its contents are the sole responsibility of WHO and do not necessarily reflect the views of the European Union.

For further information please contact the WHO Data, Artificial Intelligence and Digital Health Unit (EUDigitalHealth@who.int).

Abbreviations

AI	artificial intelligence
AISAP	AI As Soon As Possible (firm)
CHAI	Coalition for Health AI
CT	computerized tomography
DTAC	Digital Technology Assessment Criteria
EHRs	electronic health records
EU	European Union
FDA	United States Food and Drug Administration
GDPR	General Data Protection Regulation
ICD-10	International Classification of Disease version 10
IEC	International Electrotechnical Commission
ISO	International Organization for Standardization
IT	information technology
ITU	International Telecommunication Union
NHS	National Health Service
OECD	Organisation for Economic Co-operation and Development
OPTICA	Organisational PerspecTive Checklist for AI solutions adoption
PACS	picture archiving and communication system
RCT	randomized controlled trial
SPI-DDH	Strategic Partners' Initiative for Data and Digital Health
WG1	Working Group 1 of the Strategic Partners' Initiative for Data and Digital Health

Executive summary

This report, from Working Group 1 the WHO Strategic Partners Initiative on Data and Digital Health, addresses the critical gap between the transformative potential of artificial intelligence (AI) in health care and the practical realities of real-world deployment across Member States' health systems. With WHO estimating a global shortfall of 11 million health workers by 2030, AI offers strategic value for augmenting existing capacity, streamlining clinical and administrative tasks, and supporting earlier and more accurate diagnosis, but only when implemented through context-sensitive, ethically grounded and systemically sustainable approaches. This report provides a synthesized overview of governance frameworks spanning global initiatives to regional legislation and national frameworks across multiple jurisdictions, while presenting evaluation and implementation frameworks, and regional approaches that enable health-care organizations to systematically assess AI solutions with a view to adoption.

A series of 11 detailed case studies from seven countries – Finland, Israel, Latvia, Norway, Portugal, Spain and the United Kingdom – demonstrate real-world AI implementation across its capabilities for prediction, classification, association and optimization. These examples demonstrate that successful AI deployment requires adaptive, nonlinear processes that address data quality challenges, clinical workflow integration, stakeholder engagement and robust governance mechanisms.

The case studies emphasize that technical sophistication alone is insufficient and that success depends on strong clinical involvement, integration with existing workflows, alignment with organizational priorities and the presence of internal champions who can bridge technical and clinical perspectives. The findings demonstrate that, while AI can serve as a transformative force in health care delivery, realizing this promise requires deliberate investment in data infrastructure, evaluation capacity, workforce training and peer learning mechanisms that enable safer, more effective implementation across diverse health care contexts.

Background

The Strategic Partners' Initiative for Data and Digital Health (SPI-DDH) (1) was launched in June 2024 in response to WHO Regional Committee for Europe resolution EUR/RC72/R2 on *leveraging digital transformation for better health in Europe: regional digital health action plan for the WHO European Region 2023–2030* (2), which called for the WHO Regional Office for Europe to provide stewardship for a new “strategic partnership council for data and digital health to share information and encourage collaborative innovation in the European digital health ecosystem”. In response, the SPI-DDH was created to bring together a diverse constellation of public and private stakeholders and subject matter experts to tackle key issues relating to the digital transformation of health systems, and to pursue the goal of all people benefitting from safe, affordable and person-centred digital health.

At the commencement of the initiative, a voting process was launched to identify the most crucial topics to be addressed by its members. This resulted in the following topics being identified:

1. unlocking more capacity and access to health care with digitalization and responsible artificial intelligence (AI);
2. leveraging data and digital technologies to bring health care and prevention into the home;
3. an ecosystem approach to standards and interoperability in health care; and
4. digital solutions supporting the mental health of the health workforce (public health).

Working groups were established to discuss and address each of the above, each with a mandate to develop an output they considered would be relevant as a tool to assist health policy-makers in countries of the WHO European Region.

This report details the result of discussions and the development of the chosen output for working group 1 in addressing their topic of Digitalization and Responsible AI.

Further detail on case studies referenced in this report and those of other working groups are available via the SPI-DDH online case study repository (3).

1. Introduction

Why artificial intelligence in health care matters for Member States

Artificial intelligence (AI) is rapidly transforming global health systems, with growing evidence of its potential to enhance health care, at both the individual and population level, improve health outcomes and increase system efficiency. For WHO Member States, the question is no longer whether to engage with AI, but how to do so in a way that is context-sensitive, ethically grounded and systemically sustainable.

AI offers strategic value at a time of unprecedented strain on health systems. WHO has estimated that there will be a global shortfall of 11 million health workers by 2030 (4). This shortage underscores the need for new approaches that can augment the capacity of existing staff, streamline clinical and administrative tasks, and support earlier and more accurate diagnosis. According to research from the Organisation for Economic Co-operation and Development (OECD), early adopters of AI in health care have reported positive effects – such as job reorganization, not displacement – with gains in reducing tedious tasks and improving worker safety. When implemented responsibly, AI can support health professionals in working at the top of their licenses, rather than replacing them (5).

AI is not, however, a conventional medical technology. Its performance and safety are deeply tied to the data it learns from, the environments in which it is deployed and the workflows it supports. A model that performs well in one setting may fail in another if demographic, clinical or infrastructural conditions differ. This sensitivity to context creates both opportunity and risk: AI can help close care gaps in under resourced areas but, if applied without adaptation or oversight, it may exacerbate inequities or introduce new forms of harm.

Recognizing these challenges, WHO (6) provided critical guidance for Member States through its publication *Regulatory considerations on artificial intelligence for health*. This framework highlights six areas essential for responsible implementation: documentation and transparency, life cycle risk management, validation of intended use, data quality, privacy protection and stakeholder engagement. These principles are fundamental for building trust and ensuring that AI technologies deliver measurable benefits while minimizing potential harms.

The Global Initiative on AI for Health – launched by WHO with the International Telecommunication Union (ITU) and the World Intellectual Property Organization – further warns that, without deliberate strategies to support inclusion, AI adoption may concentrate in high-resource settings, leaving low-resource environments behind. Addressing this requires investment in cost-effective, scalable models of training, deployment, and governance that are fit for diverse health care systems (7).

Meanwhile, regulatory agencies such as the United States Food and Drug Administration (FDA) and national authorities operating under frameworks established by the European Union (EU) (8–11) have developed pathways for approving AI-based medical devices. In addition, the International Medical Device Regulators Forum published its working groups guidelines for regulating Software as Medical Devices (12), as well as guiding principles for Good Machine Learning Practice (13). While essential, these frameworks often focus on static product evaluation and may not address the broader challenges of real-world deployment, including bias detection, dynamic learning systems and long-term oversight. Countries must therefore go beyond certification and invest in their own national strategies for AI integration, adapted to local health priorities and capacities.

This report responds to that need by offering concrete insights into implementation. It presents 11 real-world case studies from health systems in seven countries that have deployed AI solutions in diverse settings and for varied health care challenges. These examples offer evidence of what works and what does not in specific contexts, and what lessons can inform future planning and

governance. AI in health can serve as a transformative force, but only if adopted with clear goals, robust governance and a commitment to equity and public value.

Working Group 1

SPI-DDH Working Group 1 (WG1) focused on addressing the rapidly evolving landscape of AI in health care. With the emergence of both significant opportunities and complex challenges across sectors and borders, there is a growing need to generate high-quality real-world evidence, to ensure that the potential of AI is explored and utilized across health systems and in the context of cross-sectoral collaboration to transform care modalities and delivery, to safeguard public health and ultimately to improve timely access to safe and effective health technology. In this context, emerging cases already indicate the value and effectiveness of AI solutions for purposes ranging from discovery, research and development, clinical development, regulatory pathways and market entry, to optimizing care delivery.

WG1 has sought to deliver practical, evidence-based and actionable outcomes. It therefore collected key case studies, analyzed them with a comprehensive framework and sought to ensure that the insights developed could support evidence-informed decision- and policy-making. Its added value has lain in facilitating the exchange of implementation experience and insights from real-world experience, enabling shared learning. Further, although AI solutions are emerging as critical enablers in discovery, development and other stages of the health technology lifecycle (14), WG1, at this early stage, has concentrated its efforts on health care delivery and solutions deployed in clinical practice. To this effect, WG1 focused on generating real-world evidence on AI solutions in health care, operationalizing its core strategic objectives to reinforce the resilience and modernization of health systems.

Report purpose and structure

To support its work, WG1 developed a series of AI case studies, using a structured methodology designed to ensure relevance, diversity and practical value. It began with an exchange of ideas among stakeholders, followed by the series of AI case studies through an open questionnaire distributed by SPI-DDH. To deepen the understanding of each case, the collected examples were further refined and structured through interviews and follow-up questionnaires with the contributors.

The series aims to capture a wide range of real-world experiences and insights from health care providers, public and private payers, industry partners, government bodies and non-profit-making organizations. By documenting and sharing these case studies, SPI-DDH seeks to accelerate learning across the health care ecosystem, enabling the adaptation of effective AI implementation strategies.

This report has four key sections, each designed to support a comprehensive understanding of how AI is being applied in health care settings. They focus on:

- regulatory frameworks, which provide a synthesized overview of existing laws and policies governing AI in health care across different contexts;
- frameworks and methodologies for evaluating and implementing AI solutions in health care, offering reference to practical tools and guidance that can be useful in assessing and implementing of AI technologies;
- an introduction to the series that outlines the motivation and rationale behind its creation; and
- presentation of the case studies providing detail on real-world examples and policy insights.

Each case study describes how a specific health challenge has been addressed, including an overview of the achieved or expected impact, a detailed description of the main challenges encountered during implementation and the resources required, and the key risks lessons learned, thus ultimately fostering knowledge exchange and supporting evidence-informed decision-making.

2. Governance and key regulatory and legislative considerations

This section reviews the governance and regulatory considerations relevant to the adoption of AI in health, with a focus on both global and regional initiatives of international organizations. Such frameworks are being developed to encourage innovation while safeguarding ethics, human rights, equity and data protection. In addition to these initiatives, some policy analyses have sought to document and compare emerging approaches, for example, the HealthAI global landscape report surveys international and national governance efforts, identifies areas of alignment and divergence, and presents four case studies from Colombia, Lebanon, Pakistan and Rwanda to illustrate how global frameworks are interpreted in practice (15).

Global initiatives

High-Level Advisory Board on AI

The report of the multistakeholder United Nations Secretary-General's High-level Advisory Board on AI (16) highlights the need to ensure sound governance to avoid widening digital divides that could limit the benefits of AI to a handful of states, companies and individuals. It links the concept of leaving no one behind to global AI governance, along with key considerations such as the supply chains, training data, and geopolitical and geoeconomic implications.

The report also emphasizes the need for the inclusion of national governments and regional organizations in such a governance system. It says that “the very nature of the technology itself – transboundary in structure and application – necessitates a global approach” and that “the development, deployment and use of such a technology cannot be left to the whims of markets alone”, and acknowledging that “governance can also be a key enabler for AI innovation for the SDGs [Sustainable Development Goals] globally” (16).

Framework Convention on Artificial Intelligence

The Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law (17) was initiated in 2019, with an ad-hoc committee examining the feasibility of its development, followed by the drafting and negotiation of the text. The Convention outlines principles, risks and safeguards for the development of AI, as shown in Table 1.

The Convention covers the use of AI systems by public authorities, including private actors acting on their behalf, and private actors. It does not apply to national defence matters or research and development activities, except when the testing of AI systems may have the “potential to interfere with human rights, democracy, or the rule of law” (17).

Table 1. Council of Europe Framework Convention on Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law

Fundamental principles	Remedies, procedural rights and safeguards	Risk and impact management requirements
Activities within the lifecycle of AI systems must comply with the following fundamental principles: - human dignity and individual autonomy - equality and nondiscrimination	- Document the relevant information regarding AI systems and their usage and make it available to affected persons - The information must be sufficient to enable people concerned to challenge the decision(s) made through the use of the system or based substantially on it and	- Carry out risk and impact assessments in respect of actual and potential impacts on human rights, democracy and the rule of law, in an iterative manner - Establishment of sufficient prevention and mitigation measures as a result of the implementation of these assessments

• respect for privacy and personal data protection • transparency and oversight • accountability and responsibility • reliability • safe innovation	• to challenge the use of the system itself • Effective possibility to lodge a complaint to competent authorities • Provide effective procedural guarantees, safeguards and rights to affected persons in connection with the application of an artificial intelligence system where an artificial intelligence system significantly impacts upon the enjoyment of human rights and fundamental freedoms • Provision of notice that one is interacting with an artificial intelligence system and not with a human being	• Possibility for the authorities to introduce ban or moratoria on certain application of AI systems (“red lines”)
---	---	--

Source: reproduced by permission from Council of Europe (2024).

WHO and ITU products

WHO and ITU have issued a series of publications focused on digital health governance, including AI. These stress the need for robust legal and regulatory foundations that ensure privacy, data protection, equitable access and interoperability.

They encourage countries to adopt adaptive and transparent regulatory models, and to invest in digital literacy, workforce training and ethics oversight to promote trust in AI-powered health technology. WHO’s Global Strategy on Digital Health 2020–2025 (18) urged Member States to embed ethical considerations into all stages of AI adoption.

OECD

The Recommendation of the Council on Artificial Intelligence from OECD (19) sets forth five recommendations to be implemented in national policies and international cooperation for trustworthy AI.

OECD promotes a human-centric, risk-based approach to regulating AI in health, grounded in its core AI principles of transparency, accountability, fairness and privacy. Its framework emphasizes sector-specific risk assessment, certification processes akin to medical device regulation, and clear governance and liability structures. OECD also advocates building workforce capacity, ensuring equitable access to AI technologies and modernizing data protection and cybersecurity standards. By encouraging global collaboration, shared regulatory models and interoperable infrastructures, OECD aims to foster responsible AI innovation that improves health outcomes while safeguarding public trust (20).

Timely access to high-quality health data is fundamental to the responsible adoption of AI in health. The Recommendation of the Council on Health Data Governance (21) sets forth 12 recommendations to strengthen health data foundations. It encourages better access to health data to improve the quality, safety and patient-centeredness of health care services; support for scientific innovation and the discovery of new treatments; and evaluation of new models of health service delivery. These goals are to be achieved by strengthening the foundations of health data governance, which includes: encouraging the availability, access and use of personal health data; promoting the protection of privacy and data security; and supporting legal, policy, process, and

health data interoperability. Key enablers included in the Recommendation are building human and technical capacity to use and protect data; involving key stakeholders in design and implementation to foster trust; and enabling incentives, co-operation and accountability to achieve common goals.

Regional legislation

EU Regulation on AI

The first comprehensive law on AI to be enacted was the EU's Regulation 2024/1689 (11), which adopts a risk-based approach to outlining a comprehensive framework, defining rules for different risk levels, setting down transparency requirements and aiming towards the establishment of a strong AI ecosystem by encouraging innovation and start-ups. AI can potentially threaten fundamental rights and user safety, so the Regulation adopts a risk-based approach.

It categorizes AI systems into four levels of risk: unacceptable, high, limited and low. This categorization helps determine the level of regulation that should be applied to AI systems.

European Health Data Space

EU Regulation 2025/327 on the European Health Data Space (22) establishes a common framework for the use and exchange of electronic health data across the EU. It strengthens individuals' access to and control over their personal health information while enabling certain categories of data to be reused for research, innovation and public policy. By introducing a harmonized legal and technical framework for electronic health record systems, the Regulation aims to foster interoperability, trust and a single market for digital health services and products. Entering into force in March 2025, the Regulation would be implemented gradually, with the support of further implementing acts.

In the context of AI, the Regulation on the European Health Data Space is particularly relevant: high-risk AI applications in health care rely on access to high-quality, interoperable and well governed data, making Regulation 2025/327 a key complement to Regulation 2024/1689 on AI.

Evaluation and implementation frameworks

Health care systems implementing AI-enabled solutions can leverage several structured evaluation frameworks to ensure responsible deployment and clinical effectiveness. They address the critical gap between AI technical development and real-world clinical implementation by providing systematic approaches to assess safety, efficacy, workflow integration and organizational readiness. They also facilitate the monitoring of implementation, safety, effectiveness and cost-effectiveness for different modalities.

In parallel, a growing body of academic literature has proposed methodological standards for evaluating AI studies in health care. A comprehensive review identified more than a dozen reporting and evaluation frameworks, each addressing different stages of the development, validation and clinical integration of AI models. These tools are essential for researchers, implementers and regulators aiming to ensure quality, transparency and reproducibility in AI-driven health care interventions (23).

Here are four complementary approaches that range from understanding and evaluating the transparency of AI solutions to a practical checklist on implementation of an AI model:

Organisational Perspective Checklist for AI solutions adoption (OPTICA)

OPTICA is a practical evaluation tool, developed by researchers from Clalit Health Services in Israel, to help health care organizations systematically assess AI solutions towards adoption. The structured checklist of 77 items guides health organizations through a stepwise evaluation of AI solutions, ensuring that each tool is rigorously assessed for clinical relevance, data representativeness and equity impact before deployment. The checklist is organized in 13 chapters and four domains: clinical appropriateness, data exploration, development and performance evaluation, and deployment and monitoring plan. It requires input from five designated stakeholders:

clinical experts, AI developers, data leaders, machine learning operations experts and organizational AI leaders, to assess all risks or concerns, from the availability and quality of required input data to the need for stratified performance evaluation across demographic groups.

The OPTICA framework also embeds requirements for organizational AI governance, which include continuing monitoring and impact assessment, ensuring that AI solutions continue to perform as intended as real-world conditions evolve (24).

Coalition for Health AI (CHAI): a tool for AI solution transparency

CHAI developed an applied model card that helps health care organizations systematically document and evaluate AI solutions before deployment. The model card acts as a standardized nutrition label that systematically documents AI systems' capabilities, limitations, risks and performance metrics before clinical implementation. This collaborative framework requires input from diverse stakeholders: AI developers, clinical experts, data scientists, security teams and organizational leaders, to ensure that no critical perspective is missed while creating shared accountability for AI system performance and safety.

Health care organizations can adopt the CHAI structured documentation approach to make informed, responsible decisions about AI adoption while building stakeholder trust and regulatory compliance (25).

Regional health care system approach

A collaboration of Integrated Care Systems – health authorities in London, the United Kingdom – created a step-by-step framework that enables health care organizations to deploy AI safely at scale. The framework solves the core problem that most health systems face: having AI solutions available but lacking the structured processes, governance and collaboration mechanisms needed to implement them safely and effectively across an entire health care system.

The framework establishes a systematic and sequential four-stage implementation process – idea evaluation, proof of concept, pilot-testing and deployment – with mandatory safety assessments and continuing monitoring to ensure that AI solutions remain safe and effective over time. It creates a collaborative regional model in which AI coordination leaders across health systems share expertise and pool resources, enabling smaller organizations to benefit from collective learning while reducing duplication and costs. The approach prioritizes comprehensive risk management through clinical safety reviews, data protection assessments and ethical evaluations, while addressing the human element through structured workforce training and transparent community engagement, to build trust in AI-assisted health care decisions (26).

Evaluating Generative AI

HealthBench offers policy-makers a physician-validated, open-source evaluation framework that provides a standardized methodology for testing generative AI models in health care settings. Developed by 262 physicians across 26 specialties globally, it delivers the evidence-based assessment tool that health systems need to evaluate AI performance during implementation. This method does not measure whether using AI improves outcomes in health care; it measures whether the AI/large language model gives good answers (27).

Box 1 below provides an example of how AI frameworks have been implemented in the English National Health System

Box 1. Example - AI frameworks in the English National Health Service

For any organization attempting to deploy AI in the National Health Service (NHS) in England, United Kingdom, two frameworks inevitably become central to the journey: the NHS AI knowledge repository and the Digital Technology Assessment Criteria (DTAC). They have different functions but are increasingly referenced in parallel, especially by integrated care systems, NHS trusts and public sector buyers tasked with making safe, evidence-informed decisions on digital investment.

This section outlines the purpose, scope, practical utility and real-world limitations of both frameworks. The perspective is informed by the operational experience of deploying AI and digital health solutions in health care settings in the United Kingdom, balancing policy intent with how these tools are applied in practice.

National Health Service AI knowledge repository: structured early-stage due diligence

The NHS AI knowledge repository is a resource library that was created to help nonspecialist NHS buyers navigate the early stages of assessing AI solutions. It is not a regulatory document, but a structured set of guiding documents and resources, with guiding questions and checklists focused on issues such as data quality, algorithm transparency, intended use, clinical risk and alignment with NHS infrastructure (28).

Its main value lies in creating a shared vocabulary between procurement leaders, clinical teams and digital strategy personnel. It encourages early and focused scrutiny of AI products before formal commissioning. This includes identifying whether a solution has appropriate validation evidence, fits within clinical workflows or presents risks due to opaque logic or poor generalizability across populations.

Nevertheless, the NHS AI knowledge repository does not replace the Medicines and Healthcare products Regulatory Agency, the National Institute for Health and Care Excellence or NHS England standards; neither does it independently validate the performance of AI solutions. Its use depends heavily on local capability; some organizations apply it rigorously and others may not use it at all.

Here is an illustrative scenario. In 2023, several NHS trusts evaluating AI radiology triage tools used the AI knowledge repository to identify gaps in explainability and audit trail capacity. This helped avoid procurement delays later. Uptake across trusts, however, varies significantly, depending on internal digital maturity and procurement capacity.

DTAC: baseline conformance for digital tools

The DTAC framework is the NHS' official framework for assuring that any digital health technology – AI or not – meets minimum safety and quality requirements before being used in NHS environments (29). It is structured around five domains:

1. clinical safety
2. data protection
3. cybersecurity
4. interoperability
5. usability and accessibility.

Unlike the NHS AI knowledge repository, the DTAC framework is increasingly embedded in procurement and supplier onboarding processes. While not legally mandatory, its status as a de facto requirement means that vendors unable to demonstrate DTAC readiness often struggle to progress through early-stage procurement discussions.

The DTAC framework is especially useful as a threshold measure for procurement and information governance teams, ensuring that proposed solutions do not introduce avoidable risk. It does not, however, assess clinical efficacy, economic value or long-term health outcomes, and does not comprise an endorsement or formal certification.

Here is a use case observation. Companies such as Accurx and DrDoctor integrated DTAC compliance into their product roadmaps early, to support smoother NHS adoption. In contrast, several AI suppliers passed DTAC but still faced additional barriers owing to limited clinical evidence or lack of stakeholder alignment.

How the frameworks interact and where gaps remain

While the two frameworks serve distinct functions, they are complementary (Table 2). The AI knowledge repository supports early-stage thinking and strategic alignment. The DTAC framework provides a minimum operating standard for safe and lawful deployment. Together, they form an informal first filter for NHS buyers.

Table 2. Characteristics of the National Health Service AI knowledge repository and the DTAC framework

Framework	Purpose	Best users	Limitation
NHS AI knowledge repository	Guide for initial assessment	Commissioners, clinical sponsors, integrated care systems	Nonbinding: depends on local application
DTAC	Minimum digital assurance check	Procurement teams, suppliers	Does not assess effectiveness or impact

Importantly, neither framework alone guarantees adoption or scale. Their effective use depends on wider factors, including:

- local digital maturity and workforce capability
- clarity around reimbursement or commissioning routes
- evidence of benefit in real-world settings.

Conclusion: enablers, not endpoints

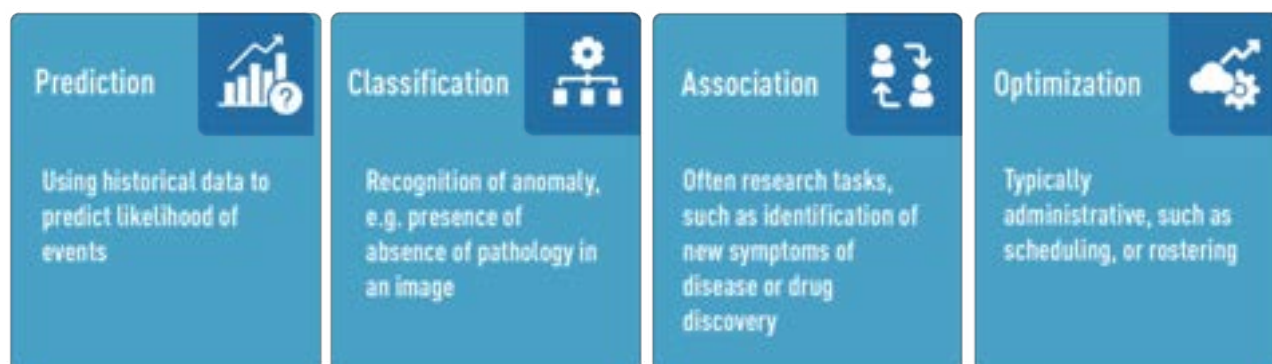
Both DTAC and the knowledge repository are increasingly central to NHS digital procurement, but they are best seen as enabling frameworks, not final gatekeepers. They help reduce risk, encourage consistent standards and structure decision-making. Successful AI deployment, however, still relies on implementation science, stakeholder engagement and evidence generation after a device has received regulatory approval and is in clinical use.

These frameworks are likely to remain fundamental as NHS digital maturity continues to grow and as AI capabilities evolve. They must be paired with local capacity building, evaluation funding and clearer commissioning pathways, however, if they are to drive equitable and sustained adoption across the NHS.

3. Frameworks for AI categorization

In its evaluation of potential categorization of AI frameworks, WG1 considered a number of resources, such as Morley (30), and selected four groupings (see Fig. 1) of AI capabilities (31) to organize the case studies in this series.

Fig. 1. Categorization of key AI capabilities in health care

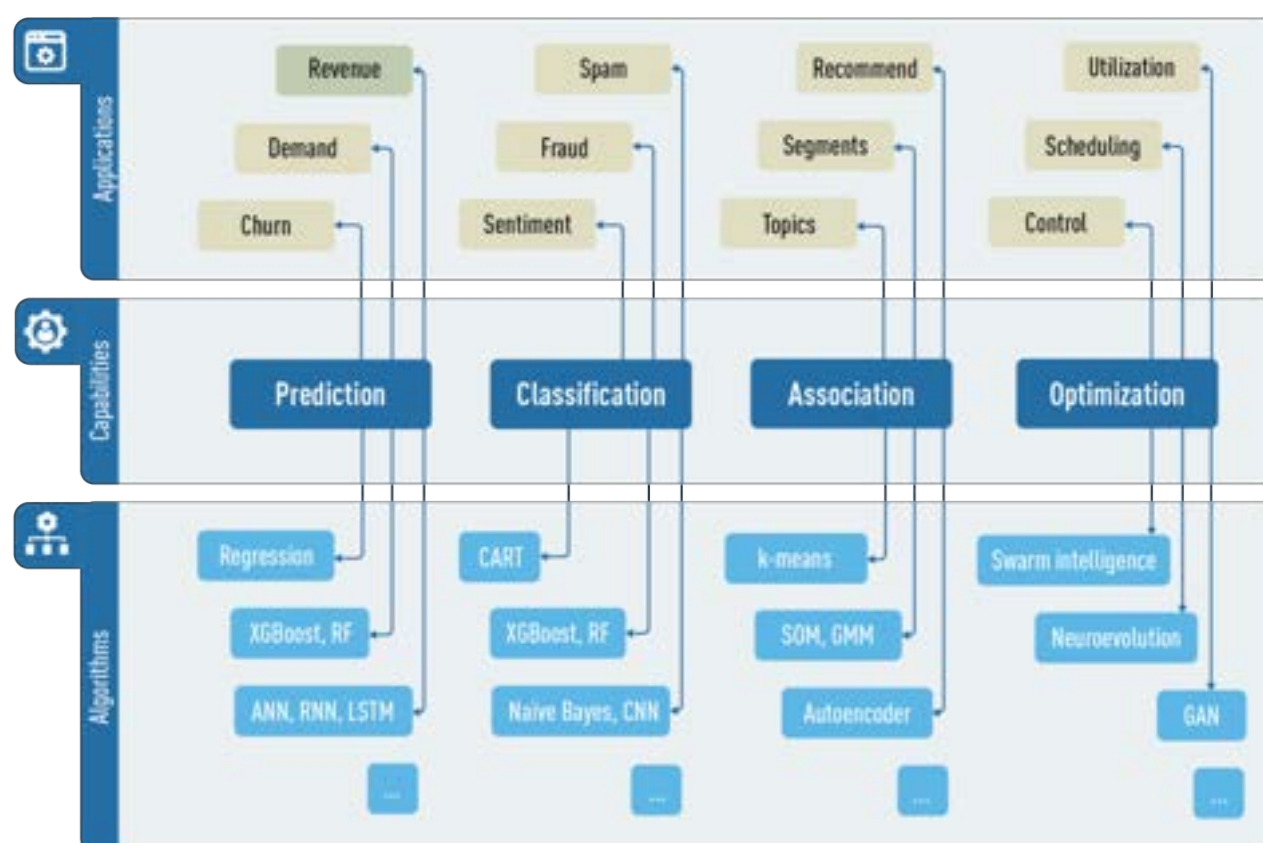


Source: adapted from DeSilva and Alahakoon (31). License: CC BY 4.0.

1. **Prediction capabilities** use historical data to predict the likelihood of events such as individuals' risk of developing a particular disease based on risk factors distilled from historical patterns in electronic health records (EHRs) (32).
2. **Classification capabilities** detect anomalies in the underlying data: for example, using computer vision to recognize the presence or absence of disease in a diagnostic image (33, 34).
3. **Association capabilities** create new knowledge (35) for health researchers, planners and practitioners through, for instance, drawing out correlations and causations around risk factors of disease or best responders to drug therapies.
4. **Optimization capabilities** focus on administrative efficiencies in the health care ecosystem. Apart from cost savings and resource optimization, they can also help with aspects such as health care practitioners' well-being (36).

It is important to note that each of the key AI capabilities described lies at the core of a technically oriented taxonomy that includes applications and indicative algorithms. In that way, users of this series may be able to link the why of artificial intelligence use (the application) with the what (the capability) and find a starting point for the how in terms of indicative algorithms (31) as seen in Fig. 2.

Fig. 2. Taxonomy of AI capabilities, applications and algorithms



Source: reproduced from DeSilva and Alahakoon (31). License: CC BY 4.0.

Note: ANN: Artificial Neural Network; CART: Classification and Regression Trees; CNN: Convolutional Neural Network; GAN: Generative Adversarial Network; GMM: Gaussian Mixture Models; k-means: k-means clustering algorithm; LSTM: Long Short-Term Memory Network; RF: Random Forest; RNN: Recurrent Neural Network; SOM: Self-Organizing Maps; XGBoost: Extreme Gradient Boosting.

Where case studies in this series straddle several of the above AI capabilities, WG1 made its best effort to allocate them into the group that represents the most significant aspect.

From a clinical perspective, it is possible to consider five core categories that highlight AI's transformative role in enhancing patient care, operational efficiency and clinical work.

1. To improve operational efficiency and resource management, AI can streamline administrative tasks, resource allocation and predictive capacity planning.
2. In clinical documentation and workflow management, solutions that leverage AI to automate routine clinical tasks, such as transcription, coding and scheduling, thereby reducing clinician burden and improving care delivery efficiency.
3. For personalized treatment and remote monitoring, AI applications can tailor health care to individual patient profiles and enable continuous monitoring of health status outside traditional care settings.
4. In predictive analytics and risk assessment, technologies use data-driven models to anticipate patient outcomes, identify high-risk individuals and support early intervention strategies.

5. In diagnostics and imaging, AI systems assist or automate the interpretation of medical images and diagnostic data to enhance accuracy, speed, and accessibility of diagnosis.

The Focus Group on Artificial Intelligence for Health was a partnership between WHO and ITU, the United Nations specialized agency for digital technologies. In seeking to establish a standardized assessment framework for the evaluation of AI-based methods for health, the Focus Group on Artificial Intelligence for Health published a thematic classification scheme (37) that distinguishes public health from clinical health. Clinical health comprises four main categories: prevention, diagnosis, treatment and research. Within any one of these categories, AI can then be classified as any one or more of 37 different clinical subclassifications that range from anaesthesiology to urology. In addition, OECD published a classification framework for AI, not specifically for health (34).

4. Case series

Introduction: bridging theory and practice

The landscape of AI in health care is now rich with strategic frameworks and theoretical potential, but a critical gap persists in the collective understanding of how these frameworks translate into practice. Despite the wealth of studies and reports on AI's transformative capabilities, detailed accounts of real-world implementation – the practical stories of how health care organizations are navigating the complex journey from AI potential to tangible patient and provider benefits – remain notably absent. This report addresses this crucial gap by providing context-specific case studies that reveal the nuanced realities of AI deployment across diverse health care settings.

These case studies serve as an essential complement of the conceptual foundations established by organizations such as WHO, OECD and regional and national regulatory bodies. Where existing literature effectively maps the strategic landscape and regulatory considerations, this series of implementation experiences illuminates the actual pathways that health care organizations are taking to operationalize these principles. By examining both successful deployments and instructive challenges, it offers health care providers and decision-makers concrete insights into how theoretical frameworks translate into clinical and operational realities.

The case series represents the collective expertise of WG1, capturing diverse implementation scenarios across varying resource contexts, clinical domains and organizational structures. Each case study unpacks the practical complexities that existing reports acknowledge but cannot fully explore, from securing stakeholder buy-in and addressing technical integration challenges, to managing the human–AI collaboration that was identified as essential for success. By grounding abstract concepts in tangible examples, it aims to accelerate the learning curve for health care systems worldwide, fostering a community of practice where both achievements and setbacks inform more effective AI implementation strategies. As AI continues its rapid evolution in health care, this series will be periodically updated to reflect emerging approaches, ensuring that Member States not only have access to the *why* of health care AI adoption but, crucially, also the *how* of successful implementation.

Methodology

Fig. 3 outlines the scope and purpose of a series of AI case studies in health care; the scope includes a problem statement, impact assessment, implementation challenges, resource considerations and knowledge-sharing aspects. The aims are to build a systematic evidence base and to focus on solutions actually deployed.

Fig. 3. Scope and purpose of a case series on AI in health care



WG1 members were invited to submit deployed solutions, known through their networks, that could be considered as suitable case studies. This was administered through a Microsoft Forms questionnaire, which enabled a good diversification of cases that represented different perspectives and contexts across the health care AI ecosystem. The initial responses were then analysed according to the six scope boxes in Fig. 3, each of which addressed a key aspect of the series:

1. identifying the health care issues the AI solution aims to address (problem);
2. briefly describing the AI-enabled solution being used (solution);
3. highlighting the need to quantify the positive outcomes of the AI implementation (impact);
4. acknowledging the various hurdles involved in implementing AI in health care, not all of which were relevant to all cases (implementation challenges: data, skills, ethics, clinical and regulatory challenges);
5. focusing on the financial and human aspects of the implementation (resources); and
6. underscoring the importance of collaboration and knowledge exchange within the health care AI community (knowledge sharing).

Almost 20 responses were collected and evaluated, and subsequent interactions by WG1 members explored them in more detail. WG1 used several criteria to decide whether to include specific case studies, including the maturity level of the project, the availability of real-world implementation experience and contributors' willingness to engage in follow-up discussions and share relevant details. A few case studies were discounted, due to their lack of relevance or close similarity to others. Eleven case studies were included in total (Table 3).

It is important to emphasize that the inclusion of any technology, vendor or institution in this report does not imply endorsement by WHO, SPI-DDH WG1 or WHO Member States. References to vendors or specific clinical implementations are included solely to provide practice-based insights into real-world challenges and enablers of AI adoption, such as governance models, regulatory navigation, workflow integration and measurable effects on care. These examples are intended to support shared learning that can guide future implementation, adoption, and policy-making.

Limitations

While the case studies methodology provides valuable real-world insights, it also has inherent limitations. The examples presented in this series reflect the experience of selected organizations that were willing and able to share information during a defined period. As such, the findings may not be representative of all AI implementations across Member States or health care settings. The insights derived are context dependent and influenced by factors such as organizational capacity, regulatory environment and local infrastructure.

Caution should be used in generalizing from these findings without consideration of the specific conditions in which each case was implemented. Further, because the process relied on self-reporting and follow-up interviews, the depth and accuracy of information provided may vary. This series should therefore be viewed as a starting point for learning and dialogue, rather than a definitive or exhaustive evaluation of AI adoption in health care. Table 3 summarizes the case studies in this report.

Table 3. Case studies presented in this report

Case	Title	Country
1	AI for pulmonary embolism detection	Finland
2	AI to transform health data into actionable insights	Spain
3	Monitoring system for deployed AI models reviewing chest X-rays	United Kingdom
4	AI for predictive coding of the International Statistical Classification of Diseases and Related Health Problems, tenth revision (ICD-10)	Portugal
5	AI implementation for point-of-care cardiac ultrasound	Israel
6	Centralized platform to scale multiple AI tools for diagnostic imaging	Israel
7	AI for automated organ contouring on computerized tomography (CT) images in radiotherapy	Norway
8	AI-driven risk stratification for hepatitis C screening	Israel
9	Agentic AI for emergency department administrative automation	United Kingdom
10	AI-Assisted Medication Risk Tool	Finland
11	AI for radiology assessment in orthopaedics	Latvia

Table 4, below, shows the classification of each case study according to the clinical use classification framework outlined earlier in this report.

Table 4. Classification of case studies by clinical use

					
CASE	Operational efficiency and resource management	Clinical documentation and workflow management	Personalized treatment and remote monitoring	Predictive analytics and risk assessment	Diagnostics and imaging
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					
11					

 denotes major classification(s) of case study
 denotes minor (subsidiary) classification(s) of case study

Case study 1. AI for pulmonary embolism detection (Finland)

Classification(s): Diagnostics and imaging; Operational efficiency and resource management

Problem

A team at the Helsinki University Hospital Diagnostic Centre in Finland implemented an AI solution that detected pathology – a pulmonary embolism – in CT scans. The solution addressed two areas: enhancing radiologists' diagnostic accuracy and improving patient access to timely imaging results and treatment.

Impact: diagnostic

The solution demonstrated a 25% increase in detection of incidental findings of pulmonary embolism in cancer patients, leading to a two weeks' faster access to treatment and improved patient outcomes. The team also noted a growing confidence in the workforce in utilizing AI solutions, which meant that further tools could be used to further enhance diagnostics and patient outcomes.

Implementation challenges

- As to data challenges, the team ensured that vendors met the privacy and security requirements for integrating with existing hospital computer systems. This maintained data protection and system compatibility throughout the process.
- As to skills barriers, the project identified specific gaps in regulatory/legal understanding and domain expertise among project managers and support staff. These were systematically addressed through internal training programmes and strategic collaboration with external partners, rather than recruitment, demonstrating a pragmatic approach to knowledge development.
- As to ethical considerations, understanding that the technology was not designed for the most critical cases, the team limited certain features, prioritizing local patients' needs.
- Clinical integration presented significant challenges during the pilot phase. The team emphasized that deployment decisions were based on patient safety and clinical functionality, with the planning of triage processes tailored to the team's specific patient populations and existing clinical pathways.
- As to regulatory complexities: the system was implemented as Software as a Medical Device, which is regulated like traditional medical devices requiring initial approval (FDA or European Conformity marking) and compliance with International Organization for Standardization (ISO)/International Electrotechnical Commission (IEC) standards, including IEC 82304-1 (on safety), IEC 62304 (on software lifecycle), ISO 13485 (on quality management) and ISO 14971 (on risk management). Team members noted that regulatory requirements aligned effectively with their established practices, although continuing requirements include maintaining quality management systems, conducting post market surveillance and submitting regulatory notifications for significant updates, which can delay software improvements compared to nonmedical applications.

Resources and knowledge sharing

The implementation required substantial organizational commitment involving a multidisciplinary team of data scientists, clinicians, information technology (IT) developers, regulatory experts and project managers to systematically evaluate 60 AI products, with 15 (25%) reaching production deployment. The team published its evidence-based testing methodology (38) and continues developing deployment and surveillance documentation. This programmatic approach demonstrates the significant human and intellectual resources needed for rigorously integrating AI while maintaining clinical standards and patient safety.

Case study 2. AI to transform health data into actionable insights (Spain)

Classification(s): Predictive analytics and risk assessment; Clinical documentation and workflow management; Personalized treatment and remote monitoring.

Problem

In multiple care settings, delays in manually recording data from medical devices can increase the risk of patient harm and reduce care efficiency. Medical devices can include glucose meters, blood pressure monitors, scales, pulse oximeters, thermometers and other devices. In long-term care facilities run by the Clariane Group, staff shortages limited the time available for entering clinical data. At Hospital Sant Joan de Déu in Barcelona, the challenge centred on supporting remote monitoring for patients receiving palliative care and care for rare diseases without overwhelming existing clinical teams or infrastructure.

Solution: clinical operations

To address these challenges, both institutions partnered with Doctomatic, to use an AI-based solution that accurately transcribes health data from medical devices and transforms them into actionable insights in patients' digital care records. The solution enables real-time monitoring while reducing the burden on staff. It was therefore used to automate data transcription to enhance operational efficiency in Clariane nursing homes and remote monitoring of patients of Hospital Sant Joan de Déu.

Impact

The deployment at Clariane demonstrated administrative time savings of 123.74 hours per month per facility and high staff satisfaction scores (8.9/10). These results underscore the system's value in improving resource utilization and patient safety in long-term care settings.

The deployment at Hospital Sant Joan de Déu continues; early-stage integration highlights the system's potential to extend care to home-based patients with complex conditions.

Implementation challenges

- As to data challenges, the system encountered significant variation in image quality that affected recognition accuracy. The team implemented a tiered verification system that automatically routed low-confidence images to human reviewers, along with photography guidelines during implementation training. Its iterative approach improved algorithm accuracy from 95% to 98%, with human beings providing verification for uncertain readings.
- As to skills barriers, health care staff in both settings demonstrated widely varying comfort levels with technology, which was addressed through comprehensive training materials, recorded sessions and peer-to-peer training via internal champions.
- As to ethical considerations, the team encountered instances where patients inadvertently included sensitive information in images, particularly with weight scales. This was addressed through the use of established protocols for staff and patients, including privacy guidance during measurements and robust data anonymization processes.
- As to clinical integration, the implementation faced resistance from established clinical workflows, leading the team to design integration with existing processes rather than replacement of them, emphasizing simplicity and reduced administrative burden. At Hospital Sant Joan de Déu, diversity in EHRs and integration of the application programming interface remain active challenges, requiring technical alignment across multiple hospital systems.
- As to regulatory complexities, the solution was operated in an international market, and understanding local regulatory requirements ensured compliance.

Resources and knowledge sharing

The use of a resource-efficient deployment model, with costs scaled by patient volume, allowed for smaller upfront investment and the demonstration of value before expansion. Key factors for success included champion-led adoption, integration-focused implementation, and alignment with current clinical workflows.

Case study 3. Monitoring system for deployed AI models reviewing X-rays (United Kingdom)

Classification(s): Operational efficiency and resource management; Diagnostics and imaging.

Problem

A health care provider working across England and Scotland needed a system for effective oversight of commercial AI diagnostic systems for chest X-rays, operating across multiple locations. The provider needed to monitor the AI models' system performance and clinical usage patterns after deployment, something often cited in literature but rarely done in health care.

Solution: clinical operations

By implementing a centralized real-time Federated AI Monitoring System, powered by Newton's Tree, a global health care AI company, the hospitals were able to better manage their resource allocation and identify potential overreliance on AI recommendations, in addition to technically monitoring the algorithm performance. This ability to monitor not only technical algorithm performance but also its influence on clinical decision-making patterns represents an advance in responsible AI governance.

Implementation challenges

- As to data challenges, the team prioritized data privacy/security and bias/fairness concerns by implementing an on-premise deployment model with a federated monitoring infrastructure. This architecture preserved local data governance while enabling centralized oversight.
- As to skills barriers, the project identified a specific gap in clinical expertise among data scientists, which was addressed through internal training programmes and strategic recruitment of personnel with relevant domain knowledge.
- As to ethical considerations, the team encountered limitations in the quality of data on patients' ethnicity, which ultimately prevented comprehensive demographic analysis.
- As to clinical integration, the implementation leveraged existing systems rather than creating standalone applications, minimizing workflow disruption while maximizing utility.
- As to regulatory complexities, the team recognized that data privacy regulations, often perceived as implementation barriers, actually provided a framework for developing its federated monitoring approach. The project secured local ethical approval, rather than pursuing traditional regulatory pathways for medical devices, demonstrating adaptability to the regulatory landscape.

Resources and knowledge sharing

By focusing on monitoring commercially available AI products, rather than developing new algorithms, and by embedding the system into existing technical architecture, the project optimized the hospitals' limited resources while maximizing clinical utility and governance capabilities. This implementation demonstrates the value of cross-disciplinary collaboration involving data scientists, clinicians, IT developers, regulatory experts and project managers. Clear communication around technical implementation and design was important in preventing confusion about data flows and information governance frameworks.

Case study 4. AI for predictive coding of ICD-10 (Portugal)

Classification(s): Clinical documentation and workflow management; Operational efficiency and resource management.

Problem

The Shared Services of the Ministry of Health of Portugal implemented a predictive coding system that assigned ICD-10 diagnostic codes to hospital episodes. The project developed a coding aid tool that suggests the most likely ICD-10 codes associated with each case, targeting improvements in both the efficiency and accuracy of the medical coding process. This implementation addresses a high-volume administrative function that directly affects health care reimbursement, clinical documentation quality and health system analytics capabilities.

Impact: administrative operations

The implementation delivered measurable improvements in medical coding processes, specifically reducing coding time while enhancing accuracy. The system's performance was evaluated through quantitative metrics, including the number of correctly suggested codes compared to the actual required codes, and the accuracy percentage of those suggestions.

Implementation challenges

- As to data challenges, the team prioritized data usability assessment and implemented high-quality data standards with regulatory compliance, emphasizing controlled vocabularies, data minimization principles and continuous international privacy measures for deidentification.
- As to skills barriers, the implementation revealed a significant gap between clinical experts' understanding of AI capabilities and what the technology could realistically deliver. This knowledge disparity created misaligned expectations that required continuing communication and education to dispel.
- As to ethical considerations, the team utilized the General Data Protection Regulation (GDPR) as an appropriate instrument to operationalize European principles for ethical AI in the absence of a formal regulation. This approach demonstrates how existing regulatory frameworks can be adapted to guide responsible AI implementation.
- As to clinical integration, the implementation faced challenges with practitioner expectations, particularly among medical coders who anticipated that the system would learn from their feedback. The team invested significant effort in explaining the necessary balance between accuracy and precision in algorithm development, highlighting the communication requirements for successful AI adoption in clinical environments.
- In the absence of AI-specific regulation, the team leveraged GDPR as its primary regulatory framework.

Resources and knowledge sharing

The implementation demonstrated organizational commitment through the Advanced Analytics and Intelligence unit of the Shared Services, following a structured approach to development and deployment. Key insights emphasized the critical importance of expectation management, with the team noting that the importance of initially defining the expectations of what the algorithm could or could not provide, so that the medical coders did not have the idea that the project would perform all of the coding for them. This highlights how projects on health care AI must navigate both technical challenges and human factors, requiring transparent communication about AI capabilities and limitations while balancing technical accuracy against usability considerations.

Case study 5. AI implementation for point-of-care cardiac ultrasound (Israel)

Classification(s): Diagnostics and Imaging.

Problem

At Sheba Medical Center, clinicians in emergency departments, internal medicine wards and outpatient settings faced significant barriers to timely cardiac assessment. A key challenge was the limited availability of experts who could both perform and interpret cardiac ultrasound. This led to delays in the diagnosis of conditions such as heart failure and valvular disease. The absence of immediate, expert-level diagnostic capabilities at the bedside, particularly for non-cardiologists, was identified as a critical gap in routine clinical practice.

Impact: diagnostics

Sheba Medical Center collaborated with the AI As Soon As Possible (AISAP) firm to implement a cloud-based AI platform for point-of-care ultrasound, featuring 12 FDA-cleared modules that assist non-cardiologists with real-time interpretation of ultrasounds and automated quantitative assessments for structural heart disease.

A prospective study of 660 patients demonstrated significant clinical impact: 55% of patients had direct changes in clinical decision-making based on AI-supported findings, with over 30% of cases prompting treatment escalation or de-escalation, early discharge or care transitions. Scans were completed in under 5 minutes, and there was high agreement between observers and formal echocardiography reports. As of the second quarter of 2025, the platform had performed over 3 000 scans across six departments and the emergency room, with nearly 150 trained clinicians actively using the system integrated into routine workflows and electronic medical records.

Implementation challenges

- As to the deployment model, implementation followed a phased rollout beginning with two clinicians in a single internal medicine department. Based on user feedback and clinical impact, it expanded incrementally to three departments with six users, and then to six departments with 15 users. After this, 10 clinicians in the emergency department were onboarded.
- As to skills barriers, a standardized training programme of 2–3 hours was provided for all users. It included instruction on the AISAP platform, focused education on cardiac ultrasound and hands-on scanning sessions tailored to clinicians' prior experience. While the training provided the necessary foundation, clinicians required additional time and repeated use to build confidence and proficiency in applying the tool effectively in clinical settings.
- Workflow integration required iterative adaptation of the platform based on real-world use. Technical adjustments were made to reduce the number of clicks, simplify the user interface and shorten the time from image acquisition to interpretation.
- As to cultural change, transitioning from reliance on delayed echocardiography interpretation to real-time, AI-supported diagnosis required clinician engagement and departmental alignment.
- As to stakeholder alignment, adoption depended on support from department heads and institutional leadership to embed the platform within standard clinical policies.

Resources and knowledge sharing

The implementation was supported by the innovation ecosystem of the ARC (Accelerate, Redesign, Collaborate) Center for Digital Innovation at Sheba Medical Center. It enabled collaboration between clinicians and technology developers through real-time clinical feedback and phased deployment for gradual scaling. This approach, combined with structured training programmes and

workflow integration, provides a replicable model for other hospitals considering AI-based diagnostic support systems, particularly when establishing co-development partnerships with technology companies.

Real-world evidence was generated through internal studies and operational monitoring to assess clinical, operational, and economic value. The AISAP platform has since expanded across multiple departments at Sheba and other institutions, although this case study focuses specifically on Sheba Medical Center's implementation experience.

Case study 6. Centralized platform to scale multiple AI solutions for diagnostic imaging (Israel)

Classification(s): Operational efficiency and resource management, Diagnostics and imaging.

Problem

Clalit Health Services, Israel's largest integrated health care organization, is facing radiologist shortages and growing imaging demand. While seeking to deploy multiple AI algorithms across radiology and pathology, Clalit discovered that evaluating and implementing each AI solution separately was both too slow and resource intensive to scale effectively. The organization needed a systematic approach to rapidly deploy multiple AI solutions while ensuring clinical safety, performance monitoring and seamless workflow integration.

Solution: clinical operations

Clalit adopted an AI orchestration platform from the Aidoc firm that integrates directly with its picture archiving and communication system (PACS), allowing centralized management of multiple AI solutions from various vendors and internal teams. This unified approach replaced the need for individual AI implementations, significantly simplifying onboarding, monitoring and maintenance processes across the imaging network. The platform reduced operational burden while enhancing Clalit's ability to efficiently evaluate and deploy new AI applications. Following successful implementation, system use was expanded to emergency medicine departments with further expansion being considered for additional clinical areas.

Implementation challenges

- As to the deployment model, the rollout followed a phased, site-specific approach, with each hospital coordinating implementation through its imaging department and deputy director for technology. Initial adoption was voluntary, allowing clinical teams to engage based on readiness.
- As to skills barriers, dedicated Aidoc teams supported each site during onboarding, working closely with Clalit staff to ensure technical and clinical integration.
- As to algorithm monitoring, in line with recommendations from Israel's National Imaging Council, Clalit addressed the need for continuing performance monitoring. The annual re-evaluation of algorithms requires substantial clinical and technical resources. To manage this, monitoring processes were embedded into routine workflows to allow early detection of performance degradation, such as issues following equipment upgrades, without relying solely on periodic audits.
- The organizational insight was that AI implementation often necessitates revisiting and adjusting existing workflows. A key lesson was the importance of appointing clinical leads who understand local processes and can guide both clinical and operational adaptation. These roles are critical for successful and sustainable integration.

Resources and knowledge sharing

The initiative required coordination among clinical, imaging, digital health, innovation, cybersecurity, and IT teams, with funding provided partly by Israel's Ministry of Health as part of a national advanced imaging initiative. Clalit's implementation followed its internal AI integration framework, OPTICA, and was executed in collaboration with the platform vendor as part of their broader digital health strategy.

The experience demonstrated the importance of structured workflows, integrated monitoring, and clearly defined clinical ownership. Key lessons include the need for workflow adaptation, sustainable monitoring mechanisms and new clinical roles to support AI integration. These are critical elements for other health systems adopting AI at scale.

Case study 7. AI for automated organ contouring on CT images in radiotherapy (Norway)

Classification(s): Diagnostics and imaging.

Problem

At Haukeland University Hospital's radiotherapy unit, clinicians had to delineate organs at risk on every patient's CT images to plan precise cancer treatments. This contouring process was tedious and time consuming, as experts drew organ boundaries slice by slice on the CT scan. It also suffered from inconsistencies, as different specialists may draw slightly different contours. With around 2500 cancer patients receiving radiotherapy at Haukeland each year, this manual workload created a significant operational challenge.

Solution: diagnostics

A project was conceived to address these issues by using AI to automatically outline organs on CT scans, thereby streamlining the process. Haukeland purchased an AI system for automatic delineation of CT images – the ART-Plan™ software from TheraPanacea – and procured a different system for its proton beam therapy facility.

Impact

The solution led to time savings of up to 80% per CT scan of head and neck patients and 50% for breast and prostate patients, with significantly reduced variability between users. The quality of the contours is much more consistent. This uniformity in contouring is expected to make radiation targeting more reliable and could translate into better patient outcomes. The introduction of AI auto-contouring enhances the radiotherapy department's capacity and quality of care, freeing up staff resources, standardizing planning practices and potentially enabling patients to start therapy faster with a higher degree of safety.

Haukeland University Hospital treats 2500 patients annually with radiotherapy for cancer of the breast, prostate, and head and neck. Initially, the technology was purchased for 1000 patients in the first year, but the allocation was quickly exhausted, and the Hospital purchased an additional 1000 patient licenses.

Implementation challenges

- As to data challenges, during procurement data had to be sent to the manufacturer. CT scans of head-and-neck patients can only be pseudo-anonymized, and the project therefore needed a contract for GDPR compliance with all manufacturers.
- As to skills barriers, dosimetrists initially demanded near perfection from the technology. To address this, a peer review process was introduced to highlight the variations and biases between different dosimetrists' plans. A clinical trial involving 20 patients was conducted, using imaging data to make comparisons.
- As to clinical integration, costs were the main criterion in the procurement process, weighted at 50%, so quality aspects, such as the precision of the delineation, had a lower weighting. This could have led to the cheapest offer being preferred if the quality and potential time savings had not been carefully evaluated. At the time of procurement, cloud-based solutions were disregarded due to uncertainties about legal aspects, and a local implementation was chosen. A cloud-based solution, however, could have led to easier maintenance and technical implementation. While two different systems are in use, they are likely to be consolidated into one as a result of financial constraints and maintenance burden, quality issues notwithstanding.

Resources and knowledge sharing

The time savings from automatic delineation meant that staff could be reassigned within the organization. The team does extensive outreach to other hospitals in Norway, communicating the solution and benefits, has worked with other hospitals and has published its results (39).

Case study 8. AI-driven risk stratification for hepatitis C screening (Israel)

Classification(s): Predictive analytics and risk assessment.

Problem

Despite the availability of highly effective antiviral treatments, hepatitis C remains underdiagnosed worldwide. Clalit Health Services faced a significant challenge: traditional screening approaches, even when aligned with guidelines from WHO and the Centers for Control and Disease Prevention, yielded low detection rates. Of the approximately 50 000 individuals screened annually, only a few dozen carriers were diagnosed, reflecting poor efficiency in identifying undiagnosed carriers. Such a gap in early identification impedes timely treatment and increases the risk of liver complications such as cirrhosis and cancer.

Solution: clinical operations

To overcome this challenge, researchers from the Clalit Research Institute developed an AI model using 25 years of patient data from Clalit's electronic health records to identify the patients most likely to have undiagnosed hepatitis C. The system generates personalized risk scores and integrates directly into clinician workflows to guide targeted screening decisions.

The pilot results were striking: the AI identified approximately 500 high-risk individuals, 38 of whom tested positive for hepatitis C, matching the number found through previous mass screening of 50 000 people and representing a hundredfold improvement in efficiency. The system provides explainable recommendations to clinicians, enabling faster treatment initiation and supporting more precise public health screening. As of the middle of 2025, Clalit was scaling up the solution across its health care system, with plans to apply the approach to other diseases with high rates of undiagnosed cases.

Implementation challenges

- As to clinical integration, integrating AI outputs into clinicians' daily workflow required embedding the risk scores into existing EHR interfaces. The system not only flags high-risk patients but also provides a transparent rationale to support physician trust and decision-making.
- Gaining clinician acceptance meant addressing the so-called black box nature of machine learning models. By focusing on explainable features and clinical logic, the system reinforced trust.
- The model was developed under strict data governance frameworks, ensuring regulatory compliance and the ethical use of personal health data.
- As to skills barriers, clinicians needed to adapt to a new model of proactive case finding. Efforts were made to streamline alerts and reduce false positives, minimizing unnecessary burden on clinical teams.

Resources and knowledge sharing

The project was led by Clalit Innovation in collaboration with internal AI research teams from Clalit Research Institute, clinical leaders and IT professionals. Its unified, longitudinal EHR system provided a strong foundation for training machine learning models, while Clalit's innovation culture and digital health experience enabled rapid development and deployment. The implementation used a phased approach emphasising explainable AI results.

Key factors in success included engaging clinicians early with continuing feedback, embedding AI insights into existing clinical workflows, rather than creating separate alerts, and setting realistic expectations for AI as a decision support tool rather than automation. This approach demonstrates how combining data science, clinical expertise and robust infrastructure can achieve significant

population health improvements. The model is applicable to other silent or underdiagnosed conditions where targeted, data-driven screening can deliver substantial benefits (40).

Case study 9. Agentic AI for emergency department administrative automation (United Kingdom)

Classification(s): Operational efficiency and resource management.

Problem

Emergency departments face severe staffing shortages worsened by overwhelming administrative burdens. Clinical staff must divert time from patient care to handle manual tracking and documentation, while administrators spend hours on repetitive tasks across multiple systems. A hospital in the United Kingdom was facing great pressure on resources, leading to treatment delays, poor patient experience and missed breach targets, breaking the NHS standard requiring 95% of patients in emergency departments to be admitted, transferred or discharged within four hours of arrival. The hospital wanted an AI solution to automate repetitive administrative tasks and free staff time for patient care.

Impact: administrative operations

In collaboration with an AI partner, the hospital implemented an agentic AI system that was able to identify when critical steps occur for patients and what important steps are missing, and analyse the capacity bottlenecks in ways not possible for human beings. Early results show significant time savings: a 50% reduction in time spent by senior operational managers on performance reporting and a 30% reduction in time spent by administrators in the department. The system monitors key departmental metrics including time from triage to first clinical assessment, time to specialty referral, time between diagnostic order and clinical decision, and four-hour standard compliance. These early results align with initial expectations and demonstrate increased operational efficiency in emergency department workflows.

Implementation challenges

- As to data challenges, data quality was the primary concern, due to inconsistent documentation practices across staff groups. The agentic system addressed this by recognizing inconsistent data entry and proactively flagging variations, with human-in-the-loop feedback improving performance over time.
- No specific ethical challenges were encountered. The focus on operational tasks, rather than clinical decision-making, combined with human-in-the-loop design, minimized ethical complexities.
- As to clinical integration, the system integrated alongside existing tools without requiring new workflows or significant process changes. One compromise involved excluding AI interpretation of clinical data from the initial scope, owing to ongoing considerations of medical device classification.
- As to regulatory complexities, the operational nature allowed adherence to best practices under the United Kingdom AI guidelines and ISO 42001, rather than complex medical device pathways. The focus on administrative applications involved simpler regulatory requirements than clinical decision-making systems.

Resources and knowledge sharing

The project required a multidisciplinary team including data scientists, clinicians, IT developers, project managers and data protection officers. No significant skill gaps were identified that required additional training or hiring.

The key lesson learned was the importance of optimizing training; future implementations should invest in a train-the-trainer approach to reduce training burden during rollout. The system's design incorporates data quality education into user training, while the AI's ability to flag variations and improve through feedback creates a scalable model for other emergency departments. The

technology's system-agnostic design and focus on operational rather than clinical tasks enable rapid deployment across different hospital environments while maintaining compliance through established AI governance frameworks.

Case study 10. AI-assisted medication risk tool (Finland)

Classification(s): Predictive analytics and risk assessment; Operational efficiency and resource management; Clinical documentation and workflow management.

Problem

In acute care, clinicians face time pressure, fragmented data and complex medication histories, making timely identification of risks difficult. The project at Wellbeing Services in the county of North Savo addresses the challenge of identifying medication-related risks in emergency patients, particularly risks related to falls, gastrointestinal bleeding and delirium. The aim is to enhance patient safety and reduce health care costs by providing more efficient and reliable detection of medication-related harm.

To meet this need, an AI-assisted tool was developed that combines generative AI with rule-based algorithms to compile and summarize risk-related information from clinical data. AI was chosen because it enables rapid synthesis of structured and unstructured information, supporting clinicians in emergency settings where manual review is not feasible.

Impact

The project aims to improve patient safety, optimize workflows and support cost-effectiveness in emergency care. While a randomized controlled trial (RCT) began in October 2025 and results are not yet available, early implementation and user feedback have met expectations for feasibility and usability.

Key indicators for evaluation include the tool's effect on identifying risks, improving safety, streamlining workflows and reducing costs. Quantitative results will be analysed in 2026 and 2028, and the tool is expected to demonstrate positive clinical and economic impact

Implementation challenges

- Regulatory requirements significantly shaped the implementation. To comply with the EU regulations on medical devices and AI (10,11), the first version was developed as a nonmedical device, limiting AI functionality to data aggregation and visualization and not allowing inferential functions. This ensured compliance and data safety while an RCT was conducted, with inference to be added in the next, regulated phase.
- A multidisciplinary team completed a data protection impact assessment and implemented secure methods for handling sensitive data. A favourable ethical opinion was obtained from the research ethics committee to ensure proper safeguards for data use, safety and consent. Technical integration with clinical workflows required process adjustments and close collaboration between the pharmacy, emergency services, IT and legal teams. If the project were to be repeated, the regulatory and ethical processes and technical integration testing would start earlier, and clinical user involvement would begin even sooner.

Resources and knowledge sharing

The project brought together data science, clinical, IT, and regulatory and project management expertise, with support from public funders and DigiFinland, which develops and supports national digital solutions for the needs of cross-sectoral public administration. Skill gaps in AI and clinical domains were addressed through collaboration across professional groups. The tool was pilot-tested as part of clinical workflows, using an iterative approach based on user feedback. Knowledge will be shared through publishing of the results of the RCT and future clinical and regulatory work to support adoption in other health systems.

Case study 11. AI for radiology assessment in orthopaedics (Latvia)

Classification(s): Diagnostics and imaging; Operational efficiency and resource management; Predictive analytics and risk assessment.

Problem

The Children's Clinical University Hospital, in Riga, adopted AI in radiology for two purposes: for identification of bone fractures and automatic determination of bone age. In line with the Hospital's medium-term strategy for 2020–2025 to improve health care capacity, it adopted these solutions to optimize efficiency in medical imaging. The overarching goals included reducing the length of diagnostic episodes, enhancing workforce efficiency and shortening the time required for image analysis and reporting.

Impact: diagnostics

All trauma patients' radiology scans are analysed for fractures and automatically transferred to the Hospital's PACS. The software proved to be valuable, particularly after hours, when immediate radiologist interpretation is unavailable. In these situations, it provides clinicians with rapid diagnostic support that is particularly helpful when their experience may not be sufficient for precise fracture analysis. The AI is a useful addition to the radiologists' work, particularly in for initial assessment before a radiologist's review. AI does, however, make the radiologists' assessment process longer, as they are required to wait a few minutes for output.

The verification system for bone age is used in endocrinological and legal settings, where either a child's growth is impaired and it is necessary to verify the age of the bones, or for legal purposes, where it is used to determine a child's age where it is uncertain, for example, in immigration cases.

Implementation challenges

- The team did not report any significant challenges in terms of the technical implementation of the solution, and the automatic transfer to the PACS was especially helpful.
- Bone structures in children differ from those of adults, which has led to occasional inaccuracies, especially in paediatric cases where fractures present atypically and growth plates can mimic fractures.
- The bone age programme sometimes fails to produce results if the images are improperly captured or if incorrect data are entered into the hospital's information system. These challenges have provided opportunities to improve imaging practices and enhance the accuracy of data entry, ultimately leading to better overall workflow quality.
- During initial testing, the system produced incorrect results more often than anticipated. This led to some groups, such as orthopaedic surgeons, having low trust in the solution.
- If the software mistakenly fails to identify a fracture, this is later discovered in the radiologists' examination of the images, and the patient is called back if discharged.

Resources and knowledge sharing

When errors in the software are identified, the team actively communicates with the providers and provides feedback to help refine the solutions. If the solution fails to produce results, owing to improperly captured images, the team focuses on ensuring that imaging practices are improved. Errors in data entry are identified and addressed within the Hospital's information system to prevent similar issues in the future.

The Hospital is actively communicating with other hospitals in Latvia; as the only specialist paediatric hospital in the country, it receives cases from all over the country.

5. Discussion and cross-case insights

The integration of AI into health care is not merely a technological challenge but also involves a complex interplay of policy, practice and human resource dynamics. Real-world case studies can help to reveal the nuances of AI implementation, providing evidence that can inform both operational strategies and policy reforms. As health care organizations navigate workforce challenges and regulatory uncertainties, policy-makers have a pivotal role in creating environments that nurture experimentation that leads to true innovation. By encouraging collaborative learning and continuous policy refinement, stakeholders can work together to unlock the full potential of AI in health care, ultimately leading to safer, more efficient and more equitable health care systems. Ultimately, this collaborative approach – enriched by the collective experience of WG1 – aims to transform theoretical potential into measurable improvements in patient care and operational efficiency. The members of WG1 hope that this case series will evolve and be periodically updated with contributions from practitioners and experts, to drive continuing innovation and learning in the field of AI in health care.

AI implementation was consistently described as a nonlinear, adaptive process. The projects evolved in response to internal and external factors, with teams adjusting strategies around data access, clinical integration and oversight as challenges emerged. This highlights the need for flexible, iterative approaches that allow for learning and refinement over time.

A recurring challenge across cases was data quality and availability. Although many institutions had digital records, issues such as fragmentation, limited interoperability, privacy challenges in accessing record-level training and validation data, and a lack of structured or labelled data limited the ability to develop or validate AI models effectively. These constraints point to the importance of investment in data governance, infrastructure and interoperability standards. Another challenge that starts to emerge in health organizations that implement several AI solutions refers to the data transportation over the network. Although maximal data will improve algorithm accuracy, organizations tend to minimize the data being sent out, owing to performance and network limitations.

Projects that progressed beyond their pilot stages often shared similar features: strong clinical involvement, integration with clinical and operational workflows, alignment with organizational priorities and the presence of internal champions who could bridge technical and clinical perspectives. These factors appeared to be more important to success than technical sophistication alone.

Relatively few projects incorporated formal evaluation frameworks, and even fewer linked evaluation to health outcomes. This limits opportunities for shared learning and accountability and suggests a need to implement feasible frameworks that fit care settings. As such, this report highlights relevant evaluation and implementation frameworks.

Governance mechanisms varied substantially. While some projects operated under clear institutional or national oversight, others relied on vendor-led protocols or internal guidance. The absence of consistent frameworks for transparency, accountability and human oversight presents a risk, especially as AI systems become more deeply embedded in clinical workflows. Innovative governance mechanisms, such as *regulatory sandboxes* for AI solutions in health care, offer a controlled environment where real-world testing of AI solutions can take place under supervision of a regulator for a limited period. These can accelerate the necessary ecosystem learning in parallel, which helps to transform health care and maximize the potential of using AI solutions so that health care workers can operate at the top of their licenses. This is valuable considering the forecast shortages in most of the health professions globally (4), high burnout rates and the high costs of unutilized resources.

Finally, several organizations noted the value of peer learning and exchange across institutions. The willingness to share experiences, including challenges and setbacks, reflects a recognition that AI in health care is still an emerging field. Supporting mechanisms for structured knowledge sharing may

help reduce duplication of effort and enable safer, more effective implementation across diverse contexts.

Going forward, countries and health systems aiming to scale up AI solutions are likely to need to address a combination of data readiness, evaluation capacity, clinical and managerial engagement, governance clarity and peer learning. It is paramount that AI solutions address challenges related to transparency, equity and interoperability. This case series offers one contribution towards that collective effort.

Key learning points

The case studies presented above show a set of applications of AI across health care settings, each addressing specific clinical and operational challenges. Collectively, they highlight the potential of AI to enhance diagnostic accuracy, operational efficiency, and patient outcomes. Benefits can be foreseen in five areas.

1. As to enhanced diagnostic capabilities, AI can improve diagnostic accuracy and speed, particularly in cases where human expertise is limited or delayed. Successful implementations show measurable improvements in clinical decision-making and treatment timelines.
2. For operational efficiency, AI can streamline administrative processes, reducing manual workload and errors. This leads to improved resource utilization, staff satisfaction and quality of data for analytics.
3. In the area of governance and monitoring, monitoring after deployment is critical for ensuring that AI systems remain effective and ethically sound. Continuous oversight can identify potential overreliance on AI and facilitate adjustments to clinical workflows.
4. As to integration and adaptation, the effective integration of AI into existing workflows requires iterative adaptation and continuing clinician training. Phased rollouts based on user feedback can enhance acceptance and utilization.
5. Ethical frameworks and compliance with regulations such as GDPR and ISO standards are essential for the responsible deployment of AI in health care. Addressing data privacy, bias and user expectations must be prioritized throughout the implementation process.

6. Policy considerations

The following policy considerations are targeted at health authorities and governments to support implementation of AI across different health system contexts. They are drawn from the case studies together with implementation experience of the WG1 members. Where indicated, supporting actions refer to those listed in tables 5-7 below.

- 1. Establish programmes to ensure health data quality and interoperability**
High-quality, standardized data are the foundation for reliable AI applications. By improving data accuracy, consistency, and interoperability, health care providers can better support AI integration. These data programmes also enable effective monitoring systems (Action 4) and interoperable pilot programmes (Action 7), ensuring that AI solutions are built and evaluated on solid, trusted data.
- 2. Foster collaborative sandboxes**
Creating innovation sandboxes encourages active collaboration among health care providers, developers and regulators. These environments are ideal for testing solutions based on standardized data (Action 1) and serve as proving grounds for scalable pilot models (Action 5). They also foster discussions around ethics (Action 6) and policy development (Action 7) in real-world contexts.
- 3. Promote continuous education and training**
Continuing education enhances AI literacy among health care professionals, helping them engage meaningfully in sandbox environments (Action 2) and contribute to pilot projects (Action 5). Training also supports the effective use of monitoring tools (Action 4) and builds understanding around ethical practices (Action 6) and policy standards (Action 7).
- 4. Invest in monitoring and evaluation tools**
Real-time monitoring and evaluation systems rely on high-quality data (Action 1) and provide critical feedback for iterative development in sandbox initiatives (Action 2) and pilot programmes (Action 5). These tools also inform education programmes (Action 3) by identifying knowledge gaps and are key to verifying alignment with ethical standards (Action 6).
- 5. Support pilot programmes with scalable models**
Pilot programmes provide a controlled way to test AI solutions using validated data (Action 1) and within collaborative sandboxes (Action 2). They benefit from trained personnel (Action 3) and are guided by insights from monitoring tools (Action 4). The results from pilot evaluations inform the creation of comprehensive guidelines (Action 7).
- 6. Embed ethics at every stage of AI development**
Ethics should be embedded at every stage of AI development, from data collection (Action 1) to pilot-testing (Action 5) and staff training (Action 3). Ethical oversight in sandboxes (Action 2) and through monitoring systems (Action 4) ensures AI systems are fair, transparent and accountable. These ethical principles form the backbone of integration guidelines (Action 7).
- 7. Establish comprehensive guidelines for AI integration**
Draw on insights from data quality initiatives (Action 1), sandbox experiments (Action 2), training programmes (Action 3), monitoring systems (Action 4), pilot evaluations (Action 5), and ethical frameworks (Action 6) to create clear, adaptable policies. These guidelines will ensure responsible, effective, and scalable AI deployment in health care settings.

Potential prioritization of policy considerations

Based on their available resources and capacity to deliver policy changes, countries may wish to prioritize different actions in implementing the above policy considerations in three different phases: foundational, developmental and advanced. Several themes cut across these actions: Every phase

includes ethical considerations; Feedback loops between monitoring, pilot programmes and guidelines ensure continuous learning. In addition, capacity building through education underpins all other components. The actions presented in the tables below refer to those policy considerations described above.

Foundational phase

The goal of the **foundational phase** is to build the core infrastructure for safe and ethical AI exploration (Table 5).

Table 5. Priority actions for low-capacity countries in the foundational phase

Actions	What it enables	Links to policy consideration
Establish standards for accurate, consistent, and interoperable data	Reliable input for all future AI systems	1 (Health data quality and interoperability)
Build AI literacy among health care staff and leaders	Informed participation in development and oversight	3 (Continuous education and training)
Introduce ethical frameworks addressing fairness, transparency and accountability	Public trust and safe experimentation	6 (Ethical AI development)

Developmental phase

The goal of the **developmental phase** is to begin controlled, collaborative testing and performance monitoring (Table 6).

Table 6. Priority actions for medium-capacity countries in the developmental phase

Actions	Content	Links to policy consideration
Create innovation hubs with providers, developers and regulators to test AI in controlled settings	Foundation of trained personnel and trusted data (foundational and advanced phases)	2 (Collaborative sandboxes)
Deploy systems to track AI performance, clinical outcomes and user feedback	Enabler of safe scaling up and learning	4 (Monitoring and evaluation tools)

Advanced phase

The goal of the **advanced phase** is to scale up, evaluate and standardize AI systems at the national level (Table 7).

Table 7. Priority actions for high-capacity countries in the advanced phase

Actions	Is informed by	Links to policy consideration
Implement phased, real-world testing with feedback loops to refine AI solutions	Data, training, monitoring and collaboration readiness	5 (Scalable pilot programmes)
Develop national/regional policies based on insights from pilot programmes, monitoring and ethics	Content from the two previous phases that feeds into robust policy-making	7 (Comprehensive guidelines for integration)

References¹

1. Partners Initiative for Data and Digital Health. In: World Health Organization [website]. World Health Organization (<https://www.who.int/europe/initiatives/strategic-partners-initiative-for-data-and-digital-health>).
2. Seventy-second Regional Committee for Europe: Tel Aviv, 12–14 September 2022: resolution: leveraging digital transformation for better health in Europe: Regional digital health action plan for the WHO European Region 2023–2030. Copenhagen: WHO Regional Office for Europe (EUR/RC72/R2; <https://iris.who.int/handle/10665/362959>).
3. Strategic Partners Initiative for Data and Digital Health: Online case study repository. In: WHO Collaborating Centre for Digital Health and Telemedicine [website]. Norwegian Centre for eHealth Research (<https://ehealthresearch.no/spiddh>).
4. World health statistics 2025: monitoring health for the SDGs, Sustainable Development Goals. Geneva: World Health Organization, 2025 (<https://iris.who.int/handle/10665/381418>). License: CC BY-NC-SA 3.0 IGO.
5. Filippucci F, Gal P, Jona-Lasinio C, Leandro A, Nicoletti G. The impact of artificial intelligence on productivity, distribution and growth: key mechanisms, initial evidence and policy challenges. Paris: OECD Publishing, 2024 (OECD Artificial Intelligence Papers No. 15; <https://doi.org/10.1787/8d900037-en>).
6. Regulatory considerations on artificial intelligence for health. Geneva: World Health Organization, 2023 (<https://iris.who.int/handle/10665/373421>). License: CC BY-NC-SA 3.0 IGO.
7. Ethics and governance of artificial intelligence for health: WHO guidance. Geneva: World Health Organization, 2021 (<https://iris.who.int/handle/10665/341996>). License: CC BY-NC-SA 3.0 IGO.
8. US Food and Drug Administration Artificial Intelligence/Machine Learning (AI/ML)-Based Software as a Medical Device (SaMD) Action Plan. Silver Spring: US Food and Drug Administration, 2021 (<https://www.fda.gov/media/145022/download>).
9. Clinical decision support software. Guidance for industry and Food and Drug Administration staff. Silver Spring: US Food and Drug Administration, 2026 (<https://www.fda.gov/media/109618/download>).
10. Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 on medical devices, amending Directive 2001/83/EC, Regulation (EC) No 178/2002 and Regulation (EC) No 1223/2009 and repealing Council Directives 90/385/EEC and 93/42/EEC. In: EUR-Lex [website]. European Union, 2017 (<https://eur-lex.europa.eu/eli/reg/2017/745/2025-01-10/eng>).
11. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act). In: EUR-Lex [website]. European Union, 2024 (<https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng>).

¹ All electronic references were accessed on 1 June 2026.

12. Characterization Considerations for Medical Device Software and Software-Specific Risk. International Medical Device Regulators Forum, 2025 (IMDRF/SaMD WG/N81 FINAL:2025; <https://www.imdrf.org/documents/characterization-considerations-medical-device-software-and-software-specific-risk>).
13. Good machine learning practice for medical device development: Guiding principles. International Medical Device Regulators Forum, 2025 (IMDRF/AIML WG/N88 FINAL:2025; <https://www.imdrf.org/documents/good-machine-learning-practice-medical-device-development-guiding-principles>).
14. Medical devices. Definitions. In: World Health Organization [website]. World Health Organization (<https://www.who.int/teams/health-product-policy-and-standards/assistive-and-medical-technology/medical-devices>).
15. Mapping AI governance in health: from global regulatory alignments to LMICs' policy developments. Geneva: HealthAI, 2024 (https://healthai.agency/app/uploads/2025/05/HealthAI_GlobalLandscapeReport_Oct.2024.pdf). License: CC BY-NC-SA 3.0 IGO.
16. Governing AI for humanity: final report. New York: United Nations, 2024(https://www.un.org/sites/un2.un.org/files/governing_ai_for_humanity_final_report_en.pdf).
17. Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law. Strasbourg: Council of Europe, 2024 (Council of Europe Treaty Series – No. 225; <https://www.coe.int/en/web/artificial-intelligence/the-framework-convention-on-artificial-intelligence>).
18. Global strategy on digital health 2020–2025. Geneva: World Health Organization, 2021 (<https://iris.who.int/handle/10665/344249>). License: CC BY-NC-SA 3.0 IGO.
19. Recommendation of the Council on Artificial Intelligence. In: OECD [website]. OECD, 2019(OECD Legal Instruments No. 449; <https://legalinstruments.oecd.org/en/instruments/oecd-legal-0449>).
20. AI watch: global regulatory tracker – OECD. In: White & Case [website].| White & Case LLP, 2025 (<https://www.whitecase.com/insight-our-thinking/ai-watch-global-regulatory-tracker-oecd>).
21. OECD (2017). Recommendation of the Council on Health Data Governance. Paris: OECD Publishing (OECD Legal Instruments No. 433; <https://legalinstruments.oecd.org/public/doc/348/348.en.pdf>).
22. European Union (2025). Regulation (EU) 2025/327 of the European Parliament and of the Council of 11 February 2025 on the European Health Data Space and amending Directive 2011/24/EU and Regulation (EU) 2024/2847. In: EUR-Lex [website]. European Union (<https://eur-lex.europa.eu/eli/reg/2025/327/oj/eng>).
23. Chew B-H, Ngiam KY (2025). Artificial intelligence tool development: What clinicians need to know? BMC Med. 23(1):244 (<https://doi.org/10.1186/s12916-025-04076-0>).
24. Dagan N, Devons-Sberro S, Paz Z, Zoller L, Sommer A, Shaham G (2024). Evaluation of AI solutions in health care organizations – the OPTICA tool. NEJM AI. 1(9). Massachusetts Medical Society, 2024 (<https://doi.org/10.1056/AIcs2300269>).
25. The CHAI applied model card. Boston: Coalition for Health AI, 2024(<https://assets.ctfassets.net/7s4afyr9pmov/5DcjUQWKrM9aGYmAttvtFK/8aedcfc6d>

[26d020ea615be32e80fc752/AMC - Complete Documentation.pdf](#)). License (CC BY-ND 4.0).

26. A framework for the safe, efficient and effective implementation, use and maintenance of AI in health and care in London. London: One London, 2025 (<https://www.onelondon.online/wp-content/uploads/2025/03/A-Framework-for-the-safe-efficient-and-effective-implementation-use-and-maintenance-of-AI-in-health-and-care-in-London.pdf>).
27. Arora RK, Wei J, Hicks RS, Bowman P, Quiñonero-Candela J, Tsimplouras F et al.. HealthBench: evaluating large language models towards improved human health. In: arXiv [website]. Cornell University, 2025 (<https://doi.org/10.48550/arXiv.2505.08775>). License CC BY 4.0.
28. AI knowledge repository. In: NHS England [website]. HMSO (<https://digital.nhs.uk/services/ai-knowledge-repository>).
29. Digital Technology Assessment Criteria (DTAC). In: NHS England [website]. HMSO (<https://transform.england.nhs.uk/key-tools-and-info/digital-technology-assessment-criteria-dtac/>).
30. Morley J. Thinking critically about AI in healthcare. In: Rogers Wannell [website]. Rogers Wannell, 2023 (<https://rogerswannell.com/wp-content/uploads/2023/12/wp-1703519698138.pdf>).
31. De Silva D, Alahakoon D. An artificial intelligence life cycle: from conception to production. Patterns. 3(6):100489, 2022 (<https://doi.org/10.1016/j.patter.2022.100489>). License CC BY 4.0.
32. Hall PS, Morris A. Predictive analytics and population health. In: Sheikh A, Cresswell KM, Wright A, Bates DW, editors. Key advances in clinical informatics. Transforming health care through health information technology. Amsterdam: Elsevier:217–25, 2017 (<https://doi.org/10.1016/B978-0-12-809523-2.00015-7>).
33. Voulodimos A, Doulamis N, Doulamis A, Protopapadakis E. Deep learning for computer vision: a brief review. Comput Intell Neurosci, 2018 (<https://doi.org/10.1155/2018/7068349>). Licence CC BY 4.0.
34. OECD framework for the classification of AI systems. Paris: OECD Publishing, 2022 (OECD Digital Economy Papers, No. 323; <https://doi.org/10.1787/cb6d9eca-en>).
35. Chen P-H C, Liu Y, Peng L How to develop machine learning models for health care. Nat Mater. 18(5):410–4, 2019 (<https://doi.org/10.1038/s41563-019-0345-0>).
36. Alowais SA, Alghamdi SS, Alsuhebany N, Alqahtani T, Alshaya AI, Almohareb SN et al.. Revolutionizing health care: the role of artificial intelligence in clinical practice. BMC Med Educ. 23(1):689, 2023 (<https://doi.org/10.1186/s12909-023-04698-z>).
37. ITU-T Focus Group on Artificial Intelligence for Health . Updated draft thematic classification scheme. FG-AI4H-C-104. Geneva: International Telecommunication Union, 2019 (<https://www.itu.int/en/ITU-T/focusgroups/ai4h/Documents/FGAI4H-C-104-DraftThemClassifScheme.pdf>)
38. Ketola JHJ, Inkinen SI, Mäkelä T, Syväranta S, Peltonen J, Kaasalainen T et al. Testing process for artificial intelligence applications in radiology practice. Physica Medica: European Journal of Medical Physics, 128:104842, 2024(doi:[10.1016/j.ejmp.2024.104842](https://doi.org/10.1016/j.ejmp.2024.104842)).

39. Søbstad JM, Sulen TH, Pettersen HES, Engeseth GM, Hirschi LA, Stokkevåg CH Time efficiency, geometric accuracy, and clinical impact of AI-assisted contouring of organs at risk in head and neck cancer radiotherapy. *Acta Oncol.* 64:1194–201, 2025 (<https://doi.org/10.2340/1651-226X.2025.44015>).
40. Dagan N, Magen O, Leshchinsky M, Makov-Assif M, Lipsitch M, Reis BY et al. Prospective evaluation of machine learning for public health screening: identifying unknown hepatitis C carriers. *NEJM AI.* 1(2), Massachusetts Medical Society, 2024 (<https://doi.org/10.1056/Aloa2300012>).

The WHO Regional Office for Europe

The World Health Organization (WHO) is a specialized agency of the United Nations created in 1948 with the primary responsibility for international health matters and public health. The WHO Regional Office for Europe is one of six regional offices throughout the world, each with its own programme geared to the particular health conditions of the countries it serves.

Member States

Albania	Greece	Portugal
Andorra	Hungary	Republic of Moldova
Armenia	Iceland	Romania
Austria	Ireland	Russian Federation
Azerbaijan	Israel	San Marino
Belarus	Italy	Serbia
Belgium	Kazakhstan	Slovakia
Bosnia and Herzegovina	Kyrgyzstan	Slovenia
Bulgaria	Latvia	Spain
Croatia	Lithuania	Sweden
Cyprus	Luxembourg	Switzerland
Czechia	Malta	Tajikistan
Denmark	Monaco	Türkiye
Estonia	Montenegro	Turkmenistan
Finland	Netherlands (Kingdom of the)	Ukraine
France	North Macedonia	United Kingdom
Georgia	Norway	Uzbekistan
Germany	Poland	

World Health Organization Regional Office for Europe

UN City, Marmorvej 51,
DK-2100, Copenhagen Ø, Denmark
Tel.: +45 45 33 70 00
Fax: +45 45 33 70 01
Email: eurocontact@who.int
Website: who.int/europe