Implementation, Adoption, and Scaling Workgroup: Landscape Assessment on the Use of Artificial Intelligence To Scale PC CDS

Agency for Healthcare Research and Quality
5600 Fishers Lane
Rockville, MD 20857

www.ahrq.gov

Contract No: 75Q80120D00018

Prepared by:
Kensaku Kawamoto, MD, PhD, MHS, FACMI, FAMIA
Sofia Ryan, MSPH
Krysta Heaney-Huls, MPH
Andrew B. Chiao, MPH
David Lobach, MD, PhD, MS
Priyanka J. Desai, PhD, MSPH
Prashila M. Dullabh, MD, FAMIA
CDSiC Implementation, Adoption, and Scaling Workgroup

AHRQ Publication No. 24-0069-1
June 2024

Agency for Healthcare Research and Quality

Clinical Decision Support Innovation Collaborative

PURPOSE

The Clinical Decision Support Innovation Collaborative (CDSiC) Implementation, Adoption, and Scaling Workgroup is charged with advancing the adoption and use of safe and effective patient-centered clinical decision support (PC CDS) by identifying barriers, opportunities, and resources to achieving PC CDS at scale. The Workgroup is composed of 16 experts and stakeholders representing diverse perspectives related to CDS. This report is intended to be utilized broadly by those interested in the use of artificial intelligence to scale PC CDS. All qualitative research activities conducted by the CDSiC are reviewed by the NORC at the University of Chicago Institutional Review Board (FWA00000142).

FUNDING STATEMENT

This project was funded under contract number 75Q80120D00018 from the Agency for Healthcare Research and Quality (AHRQ), U.S. Department of Health and Human Services (HHS). The opinions expressed in this document are those of the authors and do not reflect the official position of AHRQ or HHS.

SUGGESTED CITATION

Kawamoto K, Ryan S, Heaney-Huls K, Chiao AB, Lobach D, Desai PJ, Dullabh PM, CDSiC Implementation, Adoption, and Scaling Workgroup. Implementation, Adoption, and Scaling Workgroup: Landscape Assessment on the Use of Artificial Intelligence To Scale PC CDS. Prepared under Contract No. 75Q80120D00018. AHRQ Publication No. 24-0069-1. Rockville, MD: Agency for Healthcare Research and Quality; June 2024.

Contributors: CDSiC Implementation, Adoption, and Scaling Workgroup

Members of the Implementation, Adoption, and scaling workgroup who contributed to this report are listed below.

Name	Affiliation
Deborah Collyar	Patient Advocates in Research (PAIR)
Supriya Doshi, MBA, MPH	VDO Solutions
Christopher Harle, PhD	Indiana University
Dave Little, MD	Epic
Steve Miller, MD	First Databank
Tiffany Peterson	LupusChat
Prableen Singh, MBBS, MHA, MS	Frederick Health

Table of Contents

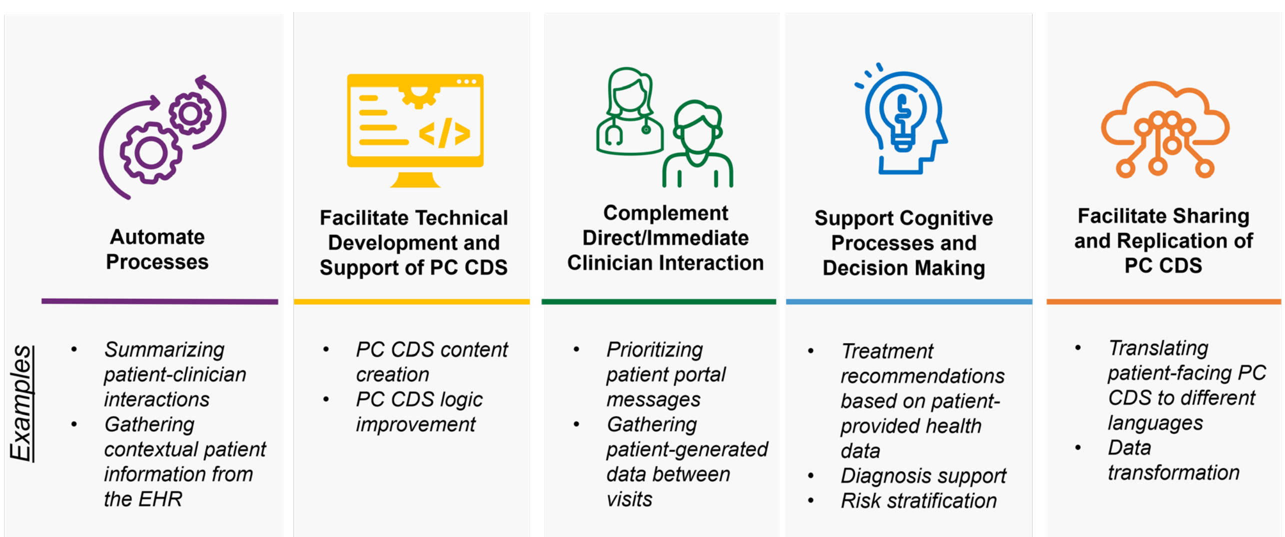
Executive Summary

Patient-centered clinical decision support (PC CDS) is a form of CDS that significantly incorporates patient-centered factors related to knowledge, data, delivery, and use. PC CDS has the potential to improve healthcare delivery by 1) facilitating shared decision making between patients and their care team, 2) ensuring that patient preferences are considered in clinical decision making, and 3) leveraging patient-generated health data and patient-reported outcomes.

Artificial intelligence (AI) has been used with CDS to improve healthcare delivery. For example, AI can assist with analyzing patient data, support clinical decision making with patients, and process large amounts of clinical data to provide recommendations more quickly and accurately compared with conventional methods. The benefits and drawbacks to using AI in clinical decision making have been noted in the literature; however, no reports have explored opportunities, considerations, and recommendations for the use of AI to scale PC CDS.

In this report, we focus on five dimensions of PC CDS to which AI can be applied to facilitate scaling of PC CDS (shown below) as well as crosscutting considerations. The scope of AI considered in this report includes machine learning techniques (including large language models and generative AI) and sophisticated rule-driven approaches.

Exhibit 1. PC CDS Scaling Dimensions and Examples



Key Findings

Below we describe the five scaling dimensions and summarize key findings specific to each. We describe unintended consequences and strategies to ensure patient safety when applicable. We also identify findings that cut across all five scaling dimensions (subsequently referred to as cross-cutting considerations).

Automate processes. AI can be used to automate processes, such as gathering or synthesizing information from a patient's record or from patient–clinician interactions, to make PC CDS tools more efficient and reduce the need for clinician time and effort on certain tasks, providing more time to focus on patient interaction and safety.

- **Challenges.** Some clinicians fear that AI may replace their job functions.
- **Unintended consequences.** Automation bias and deskilling of clinicians due to overreliance on AI-generated output are concerns that arise when using AI to automate processes.
- **Promising practices.** AI-based PC CDS should be viewed as a partner to clinicians rather than a substitute for human judgment. Clinicians and patients should continuously and thoroughly review AI-generated content to detect errors arising from AI-automated processes.

Facilitate technical development and support of PC CDS. AI can be used to facilitate technical development and support of PC CDS by making creation of PC CDS less costly and time intensive. Examples of how AI can be used in this scaling dimension include writing code, mapping variables, creating values sets, and/or creating or improving PC CDS logic.

- **Challenges.** Limited involvement of clinicians and patients in the process of developing AI-based PC CDS can hamper acceptance and use of AI-generated PC CDS content in practice.
- **Promising practices.** Clinicians and patients should be actively involved in AI-based PC CDS development, particularly to verify underlying PC CDS logic, ensure usability, and confirm the PC CDS output adds value to patient care and clinician workflows. Additionally, AI should be viewed as a partner, rather than a competitor, to human input when developing PC CDS. Clinicians and developers should review AI-generated code or suggestions for PC CDS logic improvement.

Complement direct or immediate clinician interaction. AI can complement direct or immediate clinician interaction by making information exchange between patients and their care team more efficient, freeing up more time for clinicians to focus on quality patient interaction and care. Examples of how AI can be used to scale PC CDS include gathering information from patients between visits, facilitating and prioritizing messaging between visits, or notifying clinicians of changes to a patient's health status.

- **Challenges.** Patients generally prefer to hear information from their clinicians. AI-generated responses between visits may lack contextual awareness and may not be tailored to patients' specific needs. Clinicians may lack training and experience on how to use AI-generated content, particularly on how to best convey AI-generated suggestions to patients during decision making.
- **Promising practices.** AI should not replace critical communication functions between patients and clinicians, and clinicians should review all information gathered between visits by AI chatbots for completeness and accuracy. Medical students and practicing clinicians should be trained on how best to incorporate AI into shared decision making with patients.

Support cognitive processes and decision making. To support cognitive processes and decision making, AI can provide recommendations to assist clinicians and/or patients in deciding next steps for patient care, with a goal of decreasing the time and resources required for decision making. Examples of how AI can be used in this scaling dimension include treatment recommendations based on patient-provided health data, risk assessments, or diagnostic suggestions.

- **Challenges.** The added cognitive burden from another source of information may contribute to alert fatigue. AI-based PC CDS may not fully consider context such as past medical history or patient preferences, which can affect treatment or diagnosis recommendations. Because many AI systems lack transparency regarding how recommendations are generated, clinicians may not be able to explain to patients the basis for AI-derived advice.
- **Unintended consequences.** AI can generate errors when providing diagnoses, risk prediction, or treatment recommendations, especially in uncommon or complex situations, which can pose liability concerns. Reliance on AI-based tools for decision making can limit clinicians' ethical sensitivity or decision making skills.
- **Promising practices.** AI should be viewed as a partner rather than a substitute for clinician expertise. Clinicians should always doublecheck and review AI-generated recommendations for accuracy and relevance to the patient. Training on how to safely and appropriately leverage AI within decision making should be integrated into medical education and continued education opportunities for practicing clinicians.

Facilitate sharing and replication of PC CDS. AI can be used to encourage the wider use of shareable PC CDS, ensuring that PC CDS reaches a variety of health systems, geographic locations, or patient populations and is used for a diverse set of use cases. For example, AI can be used in PC CDS to facilitate the gathering or sharing of information across health systems.

- **Challenges.** Development of interoperable AI-based PC CDS that can be easily and reliably integrated across healthcare organizations is difficult. Reports describing the development and performance of AI-based PC CDS tools are often incomplete or of poor quality. Information on the external validation of AI-based tools is limited. Migrating AI-based PC CDS across settings can result in concerns about data security. AI-based tools could potentially exacerbate the digital divide.
- **Promising practices.** Engagement of multi-disciplinary implementation teams, including clinician and patient champions, can promote the use of AI-based tools and evaluate tool performance. Reporting about the development and performance of AI-based PC CDS can be improved by setting minimum standards for describing how the AI model was developed, how clinicians were involved in development, and what evaluation metrics were used to assess performance. AI-based PC CDS tools should be trained on high-quality datasets, subject to cross-cultural/racial/ethnic validation, and subject to peer review in order to improve external validity.
- **Strategies for ensuring patient safety.** Developers can consider using synthetic data to train data models to protect patient-identifiable data from being inadvertently shared.

Crosscutting considerations. Across the literature and key informant interviews, several findings were relevant to most or all aspects of the five scaling dimensions.

- **Challenges.** The "black box" nature of AI and the fact that AI-generated output varies even when similar prompts are used erodes trust among clinicians and hinders reproducibility of results. Bias can be introduced in AI-generated output through several avenues, which can introduce—or exacerbate existing—biases and disparities in healthcare. Seamless integration of AI-based PC CDS into the workflow is often difficult and may burden clinicians with another source of information and limit time for patient interaction. The limited regulatory standards pertaining to AI-based PC CDS tools hinder scaling of PC CDS because of ill-defined expectations and processes.
- **Unintended consequences.** AI hallucinations, also known as confabulations—defined as incorrect or misleading results introduced by generative AI models—can impede shared decision making, foster clinician distrust of PC CDS recommendations, erode patient confidence in their care teams, and have negative implications for patient outcomes.
- **Promising practices.** Black box AI systems should be discouraged in favor of explainable AI. Developers can utilize a variety of techniques to optimize output quality and avoid AI hallucinations. AI-based tools should be rigorously evaluated with a defined set of metrics across their lifecycle to ensure safety and effectiveness for use in healthcare settings.
- **Strategies for ensuring patient safety.** Consistent approaches are needed to address bias in AI-generated output, such as training AI models on large diverse datasets and training clinicians to recognize when AI-based tools are being inappropriately used on specific patient populations. Research ethics principles based on best practices should be consistently applied when using AI to scale PC CDS.

Future Directions

This report 1) provides an overview of the range of approaches for using AI to scale PC CDS that are currently in use, 2) outlines existing challenges when using AI to scale PC CDS, and 3) highlights promising practices for mitigating these challenges to ensure patient safety and privacy. The findings reveal several gap areas that can be addressed by future research. The key informants provided ample suggestions for future opportunities to leverage AI with PC CDS to facilitate scaling. These suggestions include the following:

- **Patient agency**
  - Development of encounter summaries or treatment options in plain language for patients
  - Provision of additional information about symptoms or diagnosis to patients
  - Incorporation of patient preferences into PC CDS
- **Clinical decision making**
  - Development of data collection guidelines to improve clinical care
  - Synthesis of existing literature on medical conditions to aid clinician decision making
  - Management of clinical documentation (e.g., transforming PDFs into actionable information)
  - Maintenance of PC CDS (e.g., diagnosing broken logic, reviewing/updating CDS evidence base, summarizing user feedback on alerts)
- **Equity**
  - Identification and inclusion of social determinants of health factors for decision making

1. Introduction

Patient-centered clinical decision support (PC CDS) is a form of CDS that significantly incorporates patient-centered factors related to knowledge, data, delivery, and use.[1] PC CDS has potential to improve healthcare delivery by facilitating shared decision making between patients and their care team, ensuring that patient preferences are considered in clinical decision making, and leveraging patient -generated health data and patient-reported outcomes, among others.

While promising, the adoption of PC CDS remains relatively limited compared with its potential due to a range of technical implementation challenges, costs involved in creating and maintaining effective PC CDS interventions, and a lack of evidence for clinical, workload, financial, and efficiency outcomes.[1 2 3]

Artificial intelligence (AI) has been used with CDS to improve healthcare delivery. For example, AI can assist with analyzing patient data, support clinical decision making with patients, and process large amounts of clinical data to provide recommendations more quickly and accurately compared with conventional methods.[4 5 6 7] Interest in continuing to use AI to improve healthcare delivery is increasing.

Perceived benefits among clinicians about AI used in PC CDS include improved workflow efficiency, quality assurance, standardization in interpretation of results, and increased time for clinicians to interact with patients due to process automation, all of which can support efforts to scale up the implementation and adoption of PC CDS within and across settings.[4] However, barriers to realizing the potential of AI in PC CDS have been noted in the literature, including limited universal regulatory standards, variability in requirements for reporting and external validation, and mixed perceptions about the utility and safety of the technology among clinicians and patients.[8 9 10]

Although benefits and drawbacks to the use of AI in clinical decision making have been noted in the literature, there has not yet been an exploration of opportunities, considerations, and recommendations for the use of AI to scale PC CDS.

1.1 What Does the Landscape Assessment Cover?

This landscape assessment outlines key opportunities and considerations for the use of AI to scale PC CDS, including recommendations for how to use AI to scale PC CDS in a patient-centered way. In addition, this landscape assessment presents illustrative use cases for how AI is currently being used to scale PC CDS. The assessment concludes with a discussion of the opportunities to advance the use of AI in PC CDS as well as a future research agenda and recommendations for addressing barriers to the use of AI to scale PC CDS.

The landscape assessment was informed by a rapid analysis of systematic reviews related to CDS and AI, two discussions with Clinical Decision Support Innovation Collaborative (CDSiC) team members, and nine key informant interviews with health systems leaders, researchers, clinicians, industry representatives, and patient partners. A summary of methods used to develop the landscape assessment is provided in the Appendix.

While a majority of the findings in this assessment are relevant to both traditional CDS and CDS that is patient centered, we present the findings in the context of PC CDS. Several findings are particularly relevant to PC CDS, such as those focused on patient-facing aspects of AI-based tools, the effect of AI-based PC CDS on shared decision making, incorporation of patient input into AI-based PC CDS development, mitigation of bias in AI-generated output to protect patients, and privacy and security of patient-generated health data.

The findings in this landscape assessment will enable PC CDS stakeholders to better understand and leverage AI to scale PC CDS more widely and encourage the use of AI in PC CDS among clinicians, patients, and their care teams.

Section 2 of the report defines AI and scaling of PC CDS in the context of our assessment. Section 3 of the report describes key findings for using AI to scale PC CDS across each of the five dimensions of scaling, as well as cross-cutting key findings. **A summary of findings for each scaling dimension is presented in the textbox with a lightbulb icon at the beginning of each section**. Readers can view these text boxes for the key points related to each scaling dimension and read the text that follows for more detail. Additionally, future directions for research related to each scaling dimension are presented in the textbox with a map icon at the end of each section. Section 4 of the report highlights continued research gaps and future directions for the use of AI to scale PC CDS.

2. Defining Artificial Intelligence and Scaling

2.1 Conceptualizing Artificial Intelligence

The literature contains a variety of definitions of AI.[11 12 13] For the purposes of this report, we have based our concept of AI on the definition put forth by the National Artificial Intelligence Initiative Act of 2020: “A machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations or decisions influencing real or virtual environments.”[11]

Given the heterogeneity of AI systems and approaches, we present a delineation of the types of AI we considered in this report in the call-out box.

Conceptualization of Artificial Intelligence

- In this report, we ***included*** AI that utilizes advanced machine learning techniques (e.g., large language models [LLMs], natural language processing, speech recognition) and/or sophisticated rule-driven approaches to address one or more of the scaling dimensions described in section 3.2.
- In this report, we ***excluded*** AI that is used for image interpretation such as radiology and pathology diagnostics.

2.2 Key PC CDS Scaling Dimensions

Scaling has been defined as “deliberate efforts to increase the impact of successfully tested health interventions so as to benefit more people and to foster policy and program development on a lasting basis.”[14] In terms of PC CDS, this definition can be understood as efforts to expand the wider use of PC CDS across health systems and patient populations.

Shareable PC CDS is defined as PC CDS that is created using preexisting scientific or technical resources.[15] In this report, we focus on five dimensions of PC CDS where AI can be applied to facilitate scaling of shareable PC CDS more widely, shown in Exhibit 1. Examples of activities that fall under each scaling dimension are included in the Exhibit.

Exhibit 1. PC CDS Scaling Dimensions and Examples

	Automate Processes	**Facilitate Technical Development and Support of PC CDS**	**Complement Direct/Immediate Clinician Interaction**	**Support Cognitive Processes and Decision Making**	**Facilitate Sharing and Replication of PC CDS**
Examples	• *Summarizing patient-clinician interactions* • *Gathering contextual patient information from the EHR*	• *PC CDS content creation* • *PC CDS logic improvement*	• *Prioritizing patient portal messages* • *Gathering patient-generated data between visits*	• *Treatment recommendations based on patient-provided health data* • *Diagnosis support* • *Risk stratification*	• *Translating patient-facing PC CDS to different languages* • *Data transformation*

3. Key Findings

Below we describe challenges, unintended consequences (when relevant), promising practices, and strategies to ensure patient safety when using AI to scale PC CDS. We begin by describing findings specific to each of the five scaling dimensions, including examples of how AI is currently being used to scale PC CDS. Then we describe cross-cutting findings that are relevant across all five scaling dimensions.

Several of our findings apply to one or more (but not all) scaling dimensions. In these instances, we present findings in their respective sections. Examples of findings that show up frequently across dimensions include challenges with deskilling, lack of contextual awareness of AI-generated output, and clinician education as well as promising practices relating to involving AI as a partner and improving education about AI.

3.1 Automate Processes

Summary of Key Findings

- **Use Cases.** AI is currently being used to generate summaries of patients' medical records, identify information hidden in patients' records, and conduct ambient listening.
- **Challenges.** Some clinicians fear that AI may replace their job functions.
- **Unintended Consequences.** Automation bias and deskilling of clinicians due to overreliance on AI-generated output are concerns that arise when using AI to automate processes.
- **Promising Practices.** AI-based PC CDS should be viewed as a partner to clinicians rather than a substitute for human judgment. Clinicians and patients should continuously and thoroughly review AI-generated content to detect errors arising from AI-automated processes.

AI can be used to automate processes, such as gathering or synthesizing information from a patient's record or patient–clinician interactions, to make PC CDS tools more efficient and reduce the need for clinician time and effort on certain tasks, providing them more time to focus on patient interaction and safety. For example, AI can be used in PC CDS to provide descriptions of lab results, create summaries of patient medical history or treatment options, conduct ambient note taking, or flag relevant information from the electronic health record (EHR).

Researchers have examined how generative AI can support information retrieval and synthesis in healthcare.[16 17] This includes using generative AI to produce text-like encounter notes and mining EHR data to create patient summaries. By using AI to automate these processes, clinician burnout could be reduced, in turn creating efficiencies for the scale-up of PC CDS and improvement of patient outcomes.[17]

Box 1 highlights examples of how AI is currently being used to automate processes within PC CDS tools.

Box 1. Using AI To Generate Summaries and Capture Information From the Patient Record

A study explored the use of large language model (LLM) chatbots, powered by ChatGPT, to generate history of present illness summaries for patients based on interview scripts.[18] When compared with history of present illness summaries written by senior internal medicine residents, this study found the summaries generated by a chatbot were graded similarly by senior internal medicine attending physicians.

Several key informants indicated they are experimenting with ambient listening at their organizations. AI codifies words from a recorded patient encounter and creates a clinical progress note, which the clinician reviews and edits before integrating it into the medical record. One organization is conducting pilot testing of ambient documentation with over 100 providers, targeting users who are less technologically savvy to improve their efficiency and accuracy in documenting patient encounters. Another key informant highlighted reductions in clinician burnout up to 70% from using the technology.

A study assessed the use of several LLMs to identify references to social determinants of health (SDOH) information in visit notes for cancer patients receiving radiation therapy.[19] The best -performing model in the study was able to identify 45 out of 48 patients with an SDOH challenge hidden in their medical record, compared with just 1 patient identified by a clinician through International Classification of Diseases 10th Revision (ICD-10) diagnostic codes. Similar pilot studies are also in place within health systems to test the ability of LLMs to extract free-text SDOH information from medical records [20] or to identify high-risk patients in rural areas to better utilize limited resources.[21]

Challenges, unintended consequences, and promising practices for using AI to automate processes for PC CDS are summarized in Exhibit 2. These findings are expanded upon below.

Exhibit 2. Summary of Findings Relevant to the Process Automation Scaling Dimension

Challenges	Unintended Consequences	Promising Practices
• Clinician fear of replacement	• Automation bias • Deskilling of clinicians *	• Involve AI as a partner *

*Indicates a finding that is relevant to one or more scaling dimensions

3.1.1 Challenges

We identified one key challenge with using AI to automate processes.

Clinician fear of replacement. Clinicians' fear that AI may replace their job functions was noted in the literature and key informant interviews.[4] In terms of process automation, clinicians may be concerned that automation of certain tasks, such as interpretation of results, will lead to workforce loss.

3.1.2 Unintended Consequences

We also identified two unintended consequences related to the use of AI to automate processes for PC CDS.

Automation bias. Automation bias—or the tendency to over rely on automation[22 23]—is an unintended consequence of AI-based PC CDS, wherein clinicians may begin to accept all AI-generated output, even when false, just because the output is AI generated.[24]

Deskilling of clinicians. Deskilling due to overreliance on AI-generated output in PC CDS was also raised as a potential unintended consequence for process automation.[25] Key informants raised the possibility of potential loss in clinician skill or comfort with information synthesis, stating, *"If we provide too much on AI, we lose some of the knowledge ourselves. I still have to know how to be a doctor … I have to know when AI is wrong. I can't rely on AI to be my end all be all."*

3.1.3 Promising Practices

We identified the following promising practice to mitigate concerns with deskilling or automation bias associated with using AI to automate processes in PC CDS.

Involve AI as a partner. To alleviate challenges related to deskilling and automation bias, several sources and key informants suggest that AI-based PC CDS should be viewed as a partner to clinicians rather than a substitute for human judgment. This might involve continuous and thorough clinician review of AI-generated summaries using a "human in the loop" process to identify and address errors with the tools when used for process automation.[24] Patients should also be involved in review of AI -generated output. In the context of ambient listening, a key informant suggested that patients should have the opportunity to review AI-generated summaries about their visit to ensure accuracy.

Areas for Future Research

To support the practical application of AI-based PC CDS to automate processes in a safe and patient-centered way, research is needed in the following areas:

- The effects of automation on quality of care [25]
- Clinician sentiments about the potential for AI to replace their job functions [4]
- Comparison of AI-generated summaries with clinician-generated summaries

3.2 Facilitate Technical Development and Support of PC CDS

Summary of Key Findings

- **Use Cases.** AI is currently being used to generate suggestions to improve PC CDS logic and create PC CDS content (e.g., content for PC CDS alerts, order sets).
- **Challenges.** Limited involvement of clinicians and patients in the process of developing AI-based PC CDS can hamper acceptance and use of AI-generated PC CDS content in practice.
- **Promising Practices.** Clinicians and patients should be actively involved in AI-based PC CDS development, particularly to verify underlying PC CDS logic, ensure usability, and confirm the PC CDS output adds value to patient care and clinician workflows. Additionally, AI should be viewed as a partner, rather than a competitor, to human input when developing PC CDS. Clinicians and developers should review AI-generated code or suggestions for PC CDS logic improvement.

AI can be used to facilitate technical development and support of PC CDS by making creation of PC CDS less costly and time intensive. Examples of how AI can be used in this scaling dimension include writing code, mapping variables, creating value sets, and/or creating or improving PC CDS logic.

Many PC CDS alerts are overridden or ignored by clinicians, contributing to alert fatigue that can have negative downstream impacts on patient safety.[26,27] PC CDS alerts need to be regularly optimized to ensure accuracy and promote their wider use. However, human review of PC CDS alerts to eliminate unnecessary firings is resource intensive and prone to cognitive bias, which introduces barriers to the efficient scaling of PC CDS.[28,29]

Box 2 highlights examples of how AI is currently being used to facilitate technical development and support of PC CDS tools.

Box 2. Using AI To Improve PC CDS Logic and Create PC CDS Content

LLMs can automatically analyze PC CDS logic and generate suggestions for improvement, such as additional inclusion or exclusion criteria.[30] A study using ChatGPT to generate suggestions for improving the logic of CDS alerts found the AI-generated suggestions had high relevance and understanding scores and moderate usefulness scores when compared with human-generated suggestions. With additional research and refinement, AI may be useful for identifying and improving problematic PC CDS alerts that contribute to alert fatigue, allowing for more sustainable and scalable PC CDS maintenance.[31] However, continued human review and modification of AI-generated suggestions will be needed.

A key informant shared their experience using GPT-4 to create PC CDS content, such as creating order sets and care plans and writing content for PC CDS alerts. The quality of PC CDS content developed by GPT-4 was of similar quality to content generated by clinician experts and was more efficient in terms of cost and time needed for development. The key informant also used GPT-4 to write code in clinical quality language (CQL) for PC CDS but noted that the AI made more mistakes when doing so.

Challenges and promising practices for using AI to facilitate technical development and support of PC CDS are summarized in Exhibit 3. We did not identify unintended consequences for this scaling dimension. These findings are expanded upon below.

Exhibit 3. Summary of Findings Relevant to the Facilitation of Technical Development and Support of PC CDS Scaling Dimension

Challenges	Promising Practices
• Lack of patient/clinician involvement in development	• Involve patients/clinicians in development • Involve AI as a partner *

*Indicates a finding that is relevant to one or more scaling dimensions

3.2.1 Challenges

We identified one challenge associated with using AI to facilitate technical development and support of PC CDS.

Lack of patient/clinician involvement in development. Clinician involvement in the process of developing AI-based tools is limited.[32] Similarly, key informants noted limited patient involvement in the development of AI-based PC CDS. This lack of patient and clinician involvement can hamper the acceptance and use of the PC CDS code, logic, or variables created by AI in clinical practice.[9]

3.2.2 Promising Practices

We identified promising practices to address challenges associated with using AI to facilitate technical development and support of PC CDS.

Involve patients/clinicians in development. Clinicians and patients should be actively involved when developing PC CDS models using AI, particularly to verify the underlying PC CDS logic and ensure the

PC CDS output adds value to patient care and clinician workflows.[8 32 33 34] Suggested strategies for doing so include the use of think-aloud assessment methods on early prototypes to identify and address misalignments,[8] or inclusion of clinicians in the model predictor selection process.[33] Key informants particularly highlighted the need for more concerted efforts to gather information from patients about AI-based PC CDS, noting, *"People need to think like compassionate caregivers and not like scientists […] We need to think about the patient's viewpoint. It's hard to get that information sometimes, but we need to gather it."* For example, developers can conduct user experience and human factors testing with patients during the design stage to make sure tools are usable and accessible (e.g., color coding for people with color blindness), particularly for AI-based PC CDS that supports self-management among patients.

Involve AI as a partner. Viewing AI as a partner rather than competitor to human input is another promising practice for scaling. This partnership may manifest through clinicians and developers reviewing code or suggestions for PC CDS logic improvement generated by AI, rather than complete automation of these processes.[35] Similarly, respondents of a cross-sectional survey of clinician attitudes toward AI found that many advocated for involvement of clinicians in system design, procurement, and updating of AI-based tools.[4]

Areas for Future Research

To support the practical application of AI to develop PC CDS in a safe and patient-centered way, research is needed in the following area:

- Approaches for expanding involvement of patients in the development and testing of AI-based PC CDS

3.3 Complement Direct or Immediate Clinician Interaction

Summary of Key Findings

- **Use Cases.** AI is currently being used to draft responses to patient portal messages for clinicians to review before sending, triage and prioritize patient portal messages, and assist with proactive management of patients between visits.
- **Challenges.** Patients generally prefer to hear information from their clinicians. AI-generated responses between visits may lack contextual awareness and may not be tailored to patients' specific needs. Clinicians may lack training and experience on how to use AI-generated content, particularly on how to best convey AI-generated suggestions to patients during decision making.
- **Promising Practices.** AI should not replace critical communication functions between patients and clinicians, and clinicians should review all information gathered between visits by AI chatbots for completeness and accuracy. Medical students and practicing clinicians should be trained on how best to incorporate AI into shared decision making with patients.

AI can complement direct or immediate clinician interaction by making information exchange between patients and their care team more efficient, freeing up more time for clinicians to focus on quality patient interaction and care. Examples of how AI can be used to scale PC CDS include gathering information from patients between visits, facilitating and prioritizing messaging between visits, or notifying clinicians of changes to a patient's health status.

While the increased use of patient portals has facilitated communication between patients and their care teams, it has also led to a significant increase in electronic patient messages that contribute to workflow burden and clinician burnout.[7 36] LLMs may be able to facilitate patient portal communications, serving as a first stop for communication and question answering to reduce the burden on overwhelmed clinicians or for enabling the routing of important patient messages to the right member of the care team at the right time.[16 37]

Box 3 highlights examples of how AI is currently being used to complement direct or immediate clinician interaction.

Box 3. Using AI To Communicate With Patients Between Visits, Prioritize Patient Portal Messages, and Identify Patients Who Need Additional Support

A pilot program is underway at University of California San Diego Health System that integrates Microsoft's GPT-4 generative AI system into the MyChart health portal.[38] Clinicians are testing the use of GPT-4 to generate draft responses to patient portal messages using their past medical history, which the clinician can then edit for content and tone before sending to the patient. While full pilot results are still needed to determine whether the tool is effective in reducing clinicians' workload by facilitating replies to patient messages, early results are promising.

A few key informants shared that their organizations are using AI for triage and prioritization of patient portal messages. One organization is using AI to understand the urgency of a patient message and prioritizing more urgent messages for clinician attention. Another organization is using LLMs to filter in-basket patient messages. The AI tool identifies the content of the patient message and routes the message to the appropriate decision maker, improving clinician workflow.

The Mercy Healthcare system is leveraging an AI-based texting tool to help patients struggling with chemotherapy side effects.[39] Patients report and rate their symptoms into a smart texting platform, which uses AI to predict how likely a patient's symptoms may lead to hospitalization. If the AI predicts a patient is likely to be hospitalized based on their symptoms, the tool will send the patient's symptom information to the clinician and alert them that follow up is needed. The tool creates efficiencies in the proactive management of patients and allows clinicians to be more involved in the process of chemotherapy recovery.

Challenges and promising practices for using AI to complement direct and/or immediate clinician interaction are summarized in Exhibit 4. We did not identify unintended consequences for the use of AI within this scaling dimension. These findings are expanded upon below.

Exhibit 4. Summary of Findings Relevant to the Complementing Direct or Immediate Clinician Interaction Scaling Dimension

Challenges	Promising Practices
• Lack of contextual awareness * • Lack of clinician education *	• Involve AI as a partner * • Integrate AI in medical education and clinician training *

*Indicates a finding that is relevant to one or more scaling dimensions

3.3.1 Challenges

We identified two challenges associated with using AI to complement direct or immediate clinician interaction.

Lack of contextual awareness. Lack of contextual awareness of AI-generated output can pose an issue to its use in facilitating clinician interaction and shared decision making. Views about the level of empathy supplied by AI-generated output are mixed. While participants in one study rated chatbot responses in a social media forum as significantly higher quality and empathetic compared with clinician responses,[7] clinicians in a separate study perceived that AI is unable to act as empathetically as a human and has difficulty grasping context.[4] Additionally, key informants noted that patients prefer to hear information from their clinicians first, and information provided by AI may be too generic and not tailored to the patient's specific needs. Therefore, AI-based PC CDS may perform poorly or lengthen the care process when placed in a shared decision-making environment where knowledge of context and compassionate communication is key.

Lack of clinician education. Challenges related to machine learning adoption and scaling include a lack of clinician education and a steep learning curve for the use of AI-based tools.[8] A literature review of AI acceptance among clinicians found that a majority of clinicians lack basic knowledge about AI, and many felt that current training and educational tools about the use of AI in clinical settings was inadequate.[4] This lack of training and experience can negatively affect how AI-based PC CDS is used to augment clinician interaction with patients, as clinicians may not understand how to best leverage the tool in their workflow or how to best convey AI-generated output to patients during decision making.

3.3.2 Promising Practices

We identified promising practices to address challenges associated with using AI-based PC CDS to complement direct or immediate clinician interaction.

Involve AI as a partner. AI-based PC CDS should augment rather than undermine the patient–clinician relationship to address challenges with lack of contextual awareness.[9] Clinicians should be involved in reviewing all information gathered from patients between visits by AI chatbots for completeness and accuracy. Additionally, key informants emphasized that AI should not replace critical communication functions between clinicians and patients, as AI may not fully understand the experiences of patients. Further, they noted that patients value interactions with their clinician:

> *"Patients don't want to just send a message to their doctor and get a message back from the computer. They want that personal engagement of their clinician as well as the learned intermediary, the person who has the clinical experience and eyes on the data that moves in and out to verify that it is good quality information that is being transmitted."*

Therefore, AI may best be used during clinician–patient interactions as a trigger rather than substitute for direct clinician communication, exemplified by identifying overt or subtle statements from a patient that require immediate clinician interaction.[40] Importantly, clinicians must consider whether and how to disclose to a patient that AI has been used during decision making to encourage patient trust and buy-in for the tool for this purpose.[24] Some key informants emphasized that, to maintain trust in the clinician–patient relationship, patients should be notified if an AI-based tool has been used in their care.

Integrate AI in medical education and clinician training. To address challenges related to a lack of clinician education about how to use AI-based PC CDS for interaction with patients, integration of AI education into medical school curriculum and residency programs will be important.[4] In the context of this scaling dimension, improved training on how to translate AI predictions into meaningful, personalized information for patients will be critical, alongside training in communication skills to compassionately deliver this information to patients.[34] Additionally, it may be beneficial for professional organizations like the American Medical Association, the American Medical Informatics Association, or the National Association for Healthcare Quality to provide training for practicing clinicians on how best to incorporate AI-based tools into their interactions with patients.

Areas for Future Research

To support the practical application of AI-based PC CDS to complement direct clinician interaction, research is needed in the following areas:

- Patient perspective on AI-generated chatbot responses and how to ensure they meet patients' needs and preferences (e.g., by incorporating context and previous medical history) [8]
- Strategies for training clinicians on how to best engage with patients when leveraging AI-based PC CDS (e.g., how to disclose use of AI-based tools in decision making) [26]
- Approaches for incorporating AI-based tools into shared decision making with patients

3.4 Support Cognitive Processes and Decision Making

Summary of Key Findings

- **Use Cases.** AI is currently being used to manage patients' treatment and allocate resources more effectively to patients who need care the most.
- **Challenges.** The added cognitive burden from another source of information may contribute to alert fatigue. AI-based PC CDS may not fully consider context such as past medical history or patient preferences, which can affect treatment or diagnosis recommendations. Because many AI systems lack transparency regarding how recommendations are generated, clinicians may not be able to explain to patients the basis for AI-derived advice.
- **Unintended Consequences.** AI can generate errors when providing diagnoses, risk prediction, or treatment recommendations, especially in uncommon or complex situations, which can pose liability concerns. Reliance on AI-based tools for decision making can limit clinicians' ethical sensitivity or decision making skills.
- **Promising Practices.** AI should be viewed as a partner rather than a substitute for clinician expertise. Clinicians should always doublecheck and review AI-generated recommendations for accuracy and relevance to the patient. Training on how to safely and appropriately leverage AI within decision making should be integrated into medical education and continued education opportunities for practicing clinicians.

To support cognitive processes and decision making, AI can provide recommendations to assist clinicians and/or patients in deciding next steps for patient care, with the goal of decreasing the time and resources required for decision making. Examples of how AI can be used in this scaling dimension include treatment recommendations based on patient-provided health data,[25 35] risk assessments,[34] or diagnostic suggestions.[34]

AI-based PC CDS tools have the potential to quickly review large amounts of individual patient data, both found within EHRs and provided by patients, to give patient-specific recommendations.[8 35 41] Machine learning tools can compare these compiled data into effective healthcare actions for similar patient populations in order to provide tailored recommendations to care teams and patients for decision making, potentially reducing workload, optimizing clinicians' time, and improving efficiency and safety in healthcare.[41]

Box 4 highlights examples of how AI is currently being used to support cognitive processes and decision making.

Box 4. Using AI To Manage Treatment of Patients and Allocate Resources

Researchers developed a web-based platform, Sinedie, to manage treatment of patients with gestational diabetes.[42] Patients upload their data for remote monitoring into the app, which generates diet recommendations for patients and insulin therapy recommendations for clinicians. Results from a clinical evaluation of Sinedie indicated the system detected all cases requiring a therapy adjustment and generated safe recommendations with a high patient satisfaction rating. The tool reduced the clinical evaluation time required per patient, leading to quicker decision making, as well as facilitated improved access to specialized healthcare assistance to patients who needed it.[43]

A key informant's organization is using predictive modeling PC CDS to prioritize services for patients with heart failure who need rehabilitation, home health, or other services. The AI-based tool uses historical data from heart failure patients to predict hospital readmissions and identify services to support care management interventions. The tool uses these data combined with patient-reported outcome data from the Kansas City Cardiomyopathy Questionnaire to improve the predictions and optimally allocate human resources to offer care management services to more patients.

Challenges, unintended consequences, and promising practices for using AI to facilitate cognitive processes and decision making are summarized in Exhibit 5. These findings are expanded upon below.

Exhibit 5. Summary of Findings Relevant to the Support Cognitive Processes and Decision Making Scaling Dimension

Challenges	Unintended Consequences	Promising Practices
• Alert fatigue • Lack of contextual awareness * • Lack of clinician education *	• Deskilling of clinicians * • Potentially harmful errors	• Integrate AI in medical education and clinician training * • Involve AI as a partner *

*Indicates a finding that is relevant to one or more scaling dimensions

3.4.1 Challenges

We identified three challenges related to using AI-based PC CDS to support cognitive processes and decision making.

Alert fatigue. Alert fatigue has been identified as a key potential challenge when using AI-based tools among healthcare clinicians.[8] Increased work burden on care teams introduced by input from predictive PC CDS models has been shown in some instances to discourage use of AI-based tools among clinicians. Clinicians may see the review of AI-generated recommendations as extra work in their decision making process, which could negatively affect their responses to PC CDS alerts.[44]

Lack of contextual awareness. The lack of contextual awareness of AI-generated output can also pose a challenge to its use in decision making. For example, AI-based PC CDS may not fully consider context such as past medical history or patient preferences, which can affect treatment or diagnosis recommendations. Additionally, AI-based PC CDS tools may not be appropriate for complex decision making, such as in end-of-life care, because it is difficult to reduce the complex deliberations needed for these situations into a set of equations in the tool's underlying algorithm.[25]

Lack of clinician education. The lack of clinician education about AI is a challenge for its use in decision making processes. A lack of training or familiarity on how underlying AI models in PC CDS come to a decision combined with the lack of transparency in AI systems regarding how recommendations are generated can affect clinicians' willingness to accept AI-derived advice.[44] Additionally, it remains to be seen how an AI-based tool can best explain its output—one study collecting feedback on an AI model that predicted lung cancer recurrence risk found that clinicians were confused or misled by AI-generated results, even when provided with example-based explanations.[8]

3.4.2 Unintended Consequences

We identified two unintended consequences for using AI to support cognitive processes and decision making.

Deskilling of clinicians. Deskilling can be a potential unintended consequence of incorporating AI into decision making processes. For example, reliance on AI-based PC CDS for decision making in complex patient situations, such as end-of-life care, can limit ethical sensitivity and decision making skills among clinicians.[25]

Potentially harmful errors. AI can generate potentially harmful errors when providing diagnoses, risk prediction, or treatment recommendations, especially in uncommon or complex situations in which the algorithm has not been trained.[4] These potentially harmful errors can have impacts on patient care outcomes if used by clinicians to make a decision about care.[45] Additionally, potentially harmful errors pose liability concerns for clinicians because of ambiguity regarding who is held responsible for errors made by AI-based tools.[4 46]

3.4.3 Promising Practices

We identified the following promising practices to address challenges associated with using AI to support cognitive processes and decision making.

Integrate AI in medical education and clinician training. Clinicians should be taught about how the underlying models used by AI aggregate and analyze data to generate recommendations as well as their strengths and limitations.[32] Clinicians should also be specifically trained on how to recognize when AI-based PC CDS needs to be updated to reflect the latest best practices in clinical care.[34] In addition to education during medical school or residency, continuing education opportunities should be offered to practicing clinicians to refresh their skills with AI-based tools and ensure they are aware of the most up-to-date guidance and best practices.

Involve AI as a partner. AI should be viewed as a partner rather than a substitute for clinician expertise in the decision making process to address concerns related to deskilling and potentially harmful errors. As noted by a key informant, "*[we must] make sure what we're providing is decision support and not decision making … [We must] understand the interplay between the AI or PC CDS and the user's own intelligence and skills.*" Output from AI-based PC CDS should make up just one of many inputs in the decision making process and should play a role in augmenting, rather than automating, clinicians' decisions.[45 46] To avoid unintended patient outcomes or liability concerns with using AI in decision making, clinicians should always doublecheck and review AI-generated recommendations for accuracy and relevance to the patient.[4]

Areas for Future Research

To promote the practical application of AI-based PC CDS to support cognitive processes and decision making, research is needed in the following areas:

- Identification and testing of strategies for explaining AI-generated output to clinicians (e.g., examples, visual explanations) [9]
- Utility of AI-based PC CDS in complex decision making situations (e.g., end-of-life care, rare conditions) [25]
- Pilot studies focused on identifying potentially harmful errors when using LLMs and testing risk mitigation strategies
- Strategies for reducing impact of potentially harmful errors in AI-generated output on patients

3.5 Facilitate Sharing and Replication of PC CDS

Summary of Key Findings

- **Use Cases.** AI is currently being used to facilitate language translation during visits for patients with limited English proficiency and to promote semantic interoperability of PC CDS between health systems.
- **Challenges.** Development of interoperable AI-based PC CDS that can be easily and reliably integrated across healthcare organizations is difficult. Reports describing the development and performance of AI-based PC CDS tools are often incomplete or of poor quality. Information on the external validation of AI-based tools is limited. Migrating AI-based PC CDS across settings can result in concerns about data security. AI-based PC CDS tools could potentially exacerbate the digital divide.
- **Promising Practices.** Engagement of multi-disciplinary implementation teams, including clinician and patient champions, can promote the use of AI-based tools and evaluate tool performance. Reporting about the development and performance of AI-based PC CDS can be improved by setting minimum standards for describing how the AI model was developed, how clinicians were involved in development, and what evaluation metrics were used to assess performance. AI-based PC CDS tools should be trained on high-quality datasets, subject to cross-cultural/racial/ethnic validation, and subject to peer review in order to improve external validity.
- **Strategies for Ensuring Patient Safety.** Developers can consider using synthetic data to train data models to protect patient-identifiable data from being inadvertently shared.

AI can be used to encourage the wider use of shareable PC CDS, ensuring that PC CDS reaches a variety of health systems, geographic locations, or patient populations and is used for a diverse set of use cases. For example, AI can be used in PC CDS to facilitate the gathering or sharing of information across health systems.

While illustrative examples of the use of AI to facilitate sharing and replication of PC CDS across systems were lacking in the literature, key informants noted opportunities exist for AI to deliver healthcare more efficiently to larger populations. For example, AI can be used to adapt PC CDS to different sets of recipients through language translation at scale (i.e., AI can be used to translate PC CDS into different languages to address patients' language preferences). AI has already been used to transform the way content is translated and localized in social media and business,[47] suggesting that similar improvements in accessibility, cost-effectiveness, and efficiency could be realized when using AI for language translation in PC CDS.

Box 5 highlights examples of how AI is currently being used to facilitate sharing and replication of PC CDS.

Box 5. Using AI for Language Translation and To Conduct Data Transformations

An Agency for Healthcare Research and Quality (AHRQ)–funded project explored the use of AI to build and integrate an automated asynchronous interpretation tool for use within telepsychiatry visits with Spanish-speaking patients.[48] This tool allowed for multilanguage clinical evaluations to occur without the use of human interpreters. The study team compared the effectiveness of two different automated speech recognition and machine translation systems and found that both were similar in accuracy to in-person translators. However, the tools were not as accurate as humans for figurative language translation.[49]

A key informant shared that their organization is exploring the use of AI to promote semantic interoperability across health systems. The organization is using LLMs to conduct data profiling and data transformation tasks—the AI tool is used to understand how a new system's data model wraps onto the host health system's data model in order to promote interoperability between healthcare centers. The use of LLMs is more efficient to scale this process versus conducting data profiling tasks manually or through traditional AI methods (e.g., machine learning).

Challenges, promising practices, and strategies for ensuring patient safety/privacy for using AI to facilitate sharing and replication of PC CDS are summarized in Exhibit 6. We did not identify unintended consequences for the use of AI within this scaling dimension. These findings are expanded upon below.

Exhibit 6. Summary of Findings Relevant to the Sharing Facilitation and Replication of PC CDS Scaling Dimension

Challenges	Promising Practices	Strategies To Ensure Patient Safety and Patient Privacy
• Interoperability • Poor/limited reporting • Limited external validation • Data security concerns • Potential to exacerbate the digital divide	• Deploy implementation teams • Improve reporting • Improve external validation	• Use of synthetic data in • training datasets

3.5.1 Challenges

We identified five challenges for using AI to facilitate sharing and replication of PC CDS.

Interoperability. For PC CDS to be shared across healthcare organizations, systems must be interoperable, meaning that data are captured, organized, and measured using standardized methods that can be integrated across systems.[50] However, many healthcare systems use their own information systems, making it difficult to develop AI-based systems that can be easily and reliably integrated across healthcare organizations.[51] Additionally, some data or variables included in AI models are not routinely acquired or quickly available, which can impede efficiency in integrating AI-based PC CDS tools across disparate systems.[52]

Poor/limited reporting. According to the literature, reports describing the development and performance of AI-based PC CDS tools are often incomplete or of poor quality. Comprehensive reporting on AI model development and validation is lacking,[53 54] which can hinder the replication and implementation of AI-based PC CDS across settings. Standardized performance measures across studies of AI-based PC CDS tools are also lacking, which makes meaningful comparison of tools difficult.[10 55] According to a key informant, monitoring of AI-based tools centers on performance, issues with drift (i.e., alterations to inputs over time which change outputs), and bias, with less focus on monitoring clinical outcomes. Additionally, there is insufficient guidance on how to measure and report clinical outcomes associated with the use of AI-based tools.

Limited external validation. Limited reporting on external validation of AI-based tools poses a challenge to its use to scale PC CDS.[52 56] Reporting on the performance of AI-based tools in the real world or in representative data samples is limited, thus restricting the generalizability of AI across different settings and applicability to different patient populations.[56] The availability of high-quality datasets to conduct external validation of AI-based PC CDS is insufficient.[4] Key informants noted that errors in training datasets due to missing, excluded, or imputed data may pose issues for generalizability of models to wider patient populations.

Data security concerns. Data security concerns can arise when attempting to reproduce AI-based PC CDS across settings due to risk of privacy disclosure when sharing raw training datasets (which may contain identifiable information) and programming code.[4 9] Key informants raised concerns with the unintentional capture of sensitive information from AI-based tools (e.g., in transcripts for ambient listening or chart summaries) in databases for future use. Other key informants discussed the potential for exploitation of AI-generated data by bad actors and highlighted the need for backend security safeguards and appropriate data management to promote ongoing data security.

Potential to exacerbate the digital divide. Some key informants expressed concern about the potential of AI-based tools to exacerbate the digital divide between larger or academic medical centers and smaller healthcare organizations with limited resources. Key informants elaborated that, each time an AI-based tool generates an output, there is an associated cost. While competition between AI vendors can keep prices in check, the potential exists for costs to remain consistently high due to resources required to support AI development, implementation, and maintenance. Therefore, smaller institutions may not be able to afford this cost and may fall behind their peers at larger institutions. As stated by one key informant:

"I actually think the academic medical centers will probably do fine [with using AI] … you're not going to have the capabilities within the smaller hospitals and smaller centers that don't have the capabilities, and they won't be able to do AI responsibly unless you build these tools out."

A key informant also noted that it can be difficult to scale AI-based PC CDS from one organization to another without automated or semi-automated governance—while better-funded organizations have the ability to implement AI-based PC CDS, smaller organizations may lack the resources to implement all the necessary guardrails to effectively scale AI-based PC CDS.

3.5.2 Promising Practices

We identified some promising practices to address challenges associated with using AI to facilitate the sharing and replication of PC CDS.

Deploy implementation teams. To address challenges with integrating AI-based PC CDS across systems, several sources suggested engaging multidisciplinary implementation teams consisting of clinicians, health systems leaders, and AI experts to encourage the use of AI-based tools and evaluate their performance.[24,46] Additionally, the presence of a clinical and patient champion to initiate and encourage implementation is recommended to facilitate widespread use of AI-based PC CDS tools.[4,64] Since clinicians are likely to be the earliest adopters and most direct AI operators, understanding their views on the use of AI and leveraging them as champions to encourage widespread acceptance and use of AI-based PC CDS across systems is important.[4]

Improve reporting. To address challenges related to poor or incomplete reporting and increase clinical adoption of AI-based PC CDS, sources have suggested setting minimum standards for reporting to help other researchers judge the tool. This includes standards for reporting on model development (i.e., model construction, algorithms used, underlying variables, training datasets), reporting on how clinicians were involved in development, and reporting of evaluation metrics used to compare the new AI-based tool with existing models or the current standard of practice.[4,24,32,45,54] Examples of reporting standards to adopt or modify for reporting on AI-based PC CDS include the Consolidated Standards of Reporting Trials—Artificial Intelligence (CONSORT-AI)[57] or the Transparent Reporting of Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD) checklist.[32] PC CDS–focused checklists, such as the Implementation, Adoption, and Reporting tool developed by the CDSiC[58] and the Guideline Implementation with Decision Support (GUIDES) checklist,[59] can also be modified to consider the use of AI when reporting about PC CDS tools. When reporting on AI-based PC CDS development, researchers must also consider how to ethically and securely share data, methodologies, and algorithms used in development to ensure reproducibility of findings while safeguarding patient data and intellectual property.[9]

Improve external validation. Improved reporting and external validation of AI-based PC CDS are essential to facilitate replication across systems and encourage use among diverse patient populations in a safe way. Tools should be trained on high-quality datasets, subject to cross–cultural/racial/ethnic validation, and subject to peer review before used widely in clinical practice to improve external validity.[43,45] If validation of an AI-based tool across different populations is not possible, researchers should clearly report the cultural, ethnic, and racial distribution of the model, specify for which populations the AI is applicable, and report possibilities for bias in the data.[43] When reporting about the external validation of AI-based PC CDS, researchers need to provide information on the data used for training the AI, the model construction process, and the variables underlying AI models.[4]

3.5.3 Strategies To Ensure Patient Safety and Privacy

We identified the following strategy to protect patient privacy when using AI to replicate PC CDS across systems.

Use of synthetic data in training datasets. To protect patient-identifiable data from being inadvertently shared when replicating AI-based PC CDS across systems, developers can consider using synthetic data to train AI models.[43]

Areas for Future Research

To support the practical application of AI to facilitate replication and sharing of PC CDS, more research is needed in the following areas:

- Identification of robust external validation processes and approaches for clearer reporting of these processes to ensure that AI-based tools are unbiased, reliable, and generalizable across settings[4,24,33,46]
- Optimization of AI training datasets for external validation

3.6 Crosscutting Considerations

Summary of Key Findings

- **Challenges.** The "black box" nature of AI and the fact that AI-generated output varies even when similar prompts are used erodes trust among clinicians and hinders reproducibility of results. Bias can be introduced in AI-generated output through several avenues, which can introduce or exacerbate existing biases and disparities in healthcare. Seamless integration of AI-based PC CDS into the workflow is often difficult and may burden clinicians with another source of information and limit time for patient interaction. The limited regulatory standards pertaining to AI-based PC CDS tools may hinder scaling of PC CDS because of ill-defined expectations and processes.
- **Unintended Consequences.** Hallucinations, also known as confabulations—defined as incorrect or misleading results introduced by generative AI models—can impede shared decision making, foster clinician distrust of PC CDS recommendations, erode patients' confidence in their care teams, and have negative implications for patient outcomes.
- **Promising Practices.** Black box AI systems should be discouraged in favor of explainable AI. Developers can utilize a variety of techniques to optimize output quality and avoid hallucinations. AI-based tools should be rigorously evaluated with a defined set of metrics across their lifecycle to ensure safety and effectiveness for use in healthcare settings.
- **Strategies for Ensuring Patient Safety.** Consistent approaches are needed to address bias in AI-generated output, such as training AI models on large diverse datasets and training clinicians to recognize when AI-based tools are being inappropriately used on specific patient populations. Research ethics principles based on best practices should be consistently applied when using AI to scale PC CDS.

Below we present challenges, unintended consequences, promising practices, and strategies to ensure patient safety and privacy when using AI to scale PC CDS that are relevant across all five scaling dimensions. Across the literature and key informant interviews, these findings were relevant to most or all aspects of the five scaling dimensions. These findings are summarized in Exhibit 7.

Exhibit 7. Summary of Crosscutting Findings

Challenges	Unintended Consequences	Promising Practices	Strategies To Ensure Patient Safety and Patient Privacy
• Black box nature of AI • Varied output with similar prompts • Bias in AI-generated output • Workflow integration • Limited regulatory guidelines	• AI hallucinations (also known as confabulations)	• Explainable AI • Effectively evaluate AI-based PC CDS • Optimize output during AI-based PC CDS development	• Address biases in AI-generated output • Apply ethics principles

3.6.1 Challenges

We identified five crosscutting challenges to the use of AI for scaling PC CDS that are relevant across all five scaling dimensions.

Black box nature of AI. The "black box" nature of AI—referring to the lack of transparency of the inputs and operations used by AI to generate output[60]—is a key challenge to the use of AI-based tools in healthcare.[4 24 35 43 55] Clear guidelines for addressing the lack of transparency in black box AI models are not available, which in turn limits clinicians' ability to understand, explain to patients, or act upon AI-generated predictions or recommendations.[24 32 34 55] This void can lead to a lack of trust in AI-generated output among clinicians, which can have downstream consequences for using the AI's guidance in patient–clinician interactions.[35 45]

Varied output with similar prompts. According to key informants, AI-generated output can vary, even when identical prompts are used, which can erode confidence in automatically generated content and hinder reproducibility of results across systems. While some variance in output can be explained by randomness in the sampling process, AI-generated outputs are also heavily dependent on prompt quality, and resulting variations in AI-generated summary length, organization, or tone across AI-based tools can lead to variances in clinical decision making, affecting wider scale-up.[61] Additionally, LLMs can demonstrate "sycophancy" bias, wherein the AI tailors responses to perceived user expectations.[61]

Bias in AI-generated output. The presence of bias in AI-generated output is another commonly cited challenge to the use of AI to scale PC CDS. Bias can be introduced in AI-generated output through several avenues, including insufficient demographic information in training models,[52 62] lack of representative training datasets (e.g., lack of diverse patient populations, lack of rare or uncommon health outcomes),[4 24 45] low availability of high-quality training datasets,[4] and issues with duplicate, missing, or imputed data in training datasets.[35,62] Additionally, key informants highlighted the effect of out-of-date or poor quality training datasets on biased output. As noted by a key informant:

> *"Understanding the context and the source of the data, how it was collected, is important to create models ... For example, if there's an AI model for early diagnosis of some condition, and you train it based on EHR data from a health system that does a bad job of diagnosing the condition early, then your model is going to learn to diagnose the condition late."*

Potential bias in AI-generated output is concerning because it can introduce and exacerbate existing biases and disparities in healthcare and contribute to poor outcomes among underrepresented populations.[25,35,46,63]

Workflow integration. To promote adoption of AI-based PC CDS among clinicians, the systems must not introduce additional workflow burden. However, determining the right place to integrate the AI in the workflow is often challenging.[6,64] Clinician time is scarce, and information from AI-based PC CDS can increase cognitive load on clinicians if not presented at the correct time.[8,32] Additionally, AI-generated output will have to be reviewed for accuracy, which can increase burden on clinicians, who may rather use that time to speak to a patient face to face. As a key informant noted:

> *"I don't want to read [AI-generated output] and then have to go spend all this time figuring out if its right or wrong, and who to believe and what to trust, and I don't want my provider to be so busy double-checking documents that I can't have an extra three minutes with them."*

Limited regulatory guidelines. Regulations surrounding the use of AI in healthcare are evolving. Currently, regulatory standards surrounding AI-based tools are limited, which can pose a challenge to widespread use of AI in scaling PC CDS because of ill-defined expectations and processes.[4] The development of regulatory guidelines is underway, such as the Office of the National Coordinator for Health Information Technology (ONC) Health Data, Technology, and Interoperability (HTI-1) Final Rule that advocates for algorithm transparency in AI.[65] However, regulatory guidelines often focus more narrowly on the performance of AI-based tools and offer limited insight on approaches for continued monitoring or integration of AI models within and between larger healthcare systems.[10] Additionally, a key informant noted that regulatory agencies may not be able to keep up with the rapid increase of generative AI use in healthcare, which may require public–private partnerships that allow academia to work with other federal stakeholders to evaluate AI tools. Several sources in the literature noted the need for more robust regulatory approval processes for AI-based tools and slower rollout into clinical practice to protect patient safety.[24] Key informants shared a range of views about AI regulation. While some informants advocated for increased guardrails and regulation to ensure patient safety and quality, others expressed concern that regulation could stifle AI innovation. Given that AI in healthcare is a developing area with potential risks to patient safety, a balance needs to be struck between innovation and a need for regulation.

3.6.2 Unintended Consequences

We identified one crosscutting unintended consequence to the use of AI for scaling PC CDS that is relevant across all five scaling dimensions.

AI hallucinations/confabulations. AI hallucinations, also known as confabulations—defined as incorrect or misleading results introduced by generative AI models[66]—pose a significant challenge for using AI-based PC CDS.[18 30 67] Some AI-generated output may introduce completely fabricated or partially false information with the same confidence that it shares correct information, which has negative implications for patient care.[30] Generation of false information in chatbot responses can impede the shared decision making process, foster clinician distrust of PC CDS recommendations, and erode patient confidence in their care teams. According to key informants, AI hallucinations can also become costly and pose liability concerns when clinicians unintentionally use false information generated by AI-based PC CDS to make care decisions. For example, even small, one-word mistakes (e.g., replacing indications of nonproductive cough and chills with the word "fever" in a summary) can lead a clinician to a completely incorrect diagnosis that they might not have reached otherwise.[61] Importantly, a key informant noted that patients and clinicians may not be able to easily identify hallucinations in AI-generated output.

3.6.3 Promising Practices

We identified the following promising practices to address the crosscutting challenges to using AI to scale PC CDS described above.

Explainable AI. Explainable AI aims to make AI processes more transparent and interpretable, which can address challenges associated with the black box nature of AI tools.[9 35 43] Explainable AI makes the patterns underlying AI decisions clearer to researchers and clinicians, such as through example-based explanations that accompany output, which can build trust and provide a better understanding of the reasoning behind AI-generated recommendations.[8 35] Explainable AI also introduces opportunities to understand potential biases within algorithms.[46] The use of black box AI systems should be strongly discouraged in favor of explainable AI as advised in several United States and international regulations.[43] These include the World Health Organization AI Guidelines for Health,[43 68] the European Commission's legal framework for regulating AI,[8] and the ONC HTI-1 Final Rule.[65]

Effectively evaluate AI-based PC CDS. AI-based PC CDS should be rigorously evaluated with a defined set of metrics to ensure safety and effectiveness for use in healthcare settings.[9] One key informant emphasized the idea of continuous monitoring and assessment of AI tools across their life cycle:

> *"Now we are thinking of it in terms of AI as a life cycle where we need to continuously assess, and we also need to think about how implementers play an active role in this process. So, it's now developers, implementers, and regulators all working together to assess AI in a life cycle process where there's continuous measurement and assessment."*

This is in line with advocacy from key informants for slower, more gradual rollout of AI-based PC CDS tools, starting with increased pilot studies and additional comparative effectiveness and implementation science research to assess effectiveness and safety before widespread use in clinical practice.

Optimize output during AI-based PC CDS development. To avoid AI hallucinations and variable AI output, AI-based PC CDS tools must be engineered to optimize output accuracy and meet the needs of patients.[18 35] Key informants highlighted several strategies that can be undertaken during AI-based PC CDS development to optimize output, including 1) testing different versions of prompts and their associated outputs to identify preferred patterns in output and compare error rates, 2) requiring the AI to provide evidence for its output, 3) continuously testing and refining prompts to fine-tune output, and 4) training AI on larger, or more tailored datasets. Additionally, several informants suggested the retrieval automated generation approach, which allows only AI-based tools to generate output based on a specific document or set of documents in order to reduce errors. For example, clinicians can provide developers with proprietary information (e.g., order sets, guidelines) to train the AI model and instruct it to generate only output derived from those sources. Importantly, a key informant noted standards should be stricter for avoiding errors or hallucinations in AI-generated output for patient-facing tools versus clinician-facing tools.

3.6.4 Strategies To Ensure Patient Safety and Privacy

We also identified the following strategies to safeguard patient safety and privacy when using AI to scale PC CDS.

Address biases in AI-generated output. Unconscious biases, which can be exacerbated by AI-generated output, can lead to differential treatment of patients during clinical decision making and lead to patient harm.[69] For example, decisions made about a patient population using AI-generated output that was not trained using a representative dataset can result in adverse treatment outcomes for said population. Consistent approaches are needed to address bias in AI-generated output to ensure that these tools are not harmful to underrepresented or vulnerable patient populations. The Coalition for Health AI, an organization focused on providing guidelines for the responsible use of AI in healthcare, is developing a framework to reduce algorithmic bias and promote equity in AI tools, which developers can refer to when creating AI-based PC CDS.[70] Other strategies to reduce bias include training AI models on diverse datasets, entering increased amounts of data or continuous data into training datasets, and accounting for unintentional discrimination in algorithms.[34 35 46 51] Additionally, AHRQ has recommended strategies for mitigating racial disparities in algorithms in a recently published systematic review.[63] Finally, clinicians should be educated on how to recognize when AI-based PC CDS is being used on inappropriate patient populations and/or when data inputted into the tool are biased.[34]

Apply ethics principles. To protect patient safety and alleviate patient privacy concerns, research ethics principles based on best practices should be consistently applied when using AI to scale PC CDS. Modifications to traditional medical research ethics principles may be needed to guide the use and governance of AI-based PC CDS, including informed consent to use AI in patient care, recognition of individual and group-level harms and benefits, patient empowerment within the patient-clinician relationship, data protection regulations for patient-provided data, and access to inputs and outputs of AI-supported PC CDS.[9] Formal processes for training clinicians and developers on AI-related ethics as well as guidelines for ethically incorporating AI recommendations into practices are needed.[24]

Areas for Future Research

To support the practical application of AI-based PC CDS across all five scaling dimensions, research is needed in the following areas:

- Reporting guidelines to address the lack of transparency of black box AI
- Identification and testing of approaches to fine tune prompts for AI-based tools to optimize output accuracy [19, 20]
- Effectiveness of approaches for mitigating hallucinations to ensure AI-based tools can be safely and widely used [19]
- Rigorous studies to assess the effectiveness of AI-based PC CDS, including randomized controlled trials exploring model performance and the effect of tools on clinical outcomes
- Application of newer generative AI technology, such as LLMs, in conjunction with more traditional machine learning–based AI
- Identification of bias in AI-generated output [33, 46]

4. Future Directions for the Use of AI To Scale PC CDS

In this report, we provide an overview of the range of approaches for using AI to scale PC CDS that are currently in use, identify existing challenges when using AI to scale PC CDS, and highlight promising practices for mitigating these challenges to ensure patient safety and privacy. Our findings reveal gap areas for future research to address, and key informants provided ample suggestions for future opportunities to leverage AI with PC CDS to facilitate scaling (callout box).

Future Opportunities To Leverage AI in PC CDS

Patient agency

- Development of encounter summaries or treatment options in plain language for patients
- Provision of additional information about symptoms or diagnosis to patients
- Incorporation of patient preferences into PC CDS

Clinical decision making

- Development of data collection guidelines to improve clinical care
- Synthesis of existing literature on medical conditions to aid clinician decision making
- Management of clinical documentation (e.g., transforming PDFs into actionable information)
- Maintenance of PC CDS (e.g., diagnosing broken logic, reviewing/updating PC CDS evidence base, summarizing user feedback on alerts)

Equity

- Identification and inclusion of SDOH factors for decision making

However, these future use cases cannot be fully realized without further research to address the crosscutting challenges across scaling dimensions, particularly those surrounding AI hallucinations and bias in AI-generated output as well as the black box nature of AI. Several challenges that are relevant across one or more scaling dimensions need to be addressed, namely the potential for deskilling of clinicians, lack of clinician education and training about AI, and lack of contextual awareness of AI-generated output. Importantly, this assessment has shown that reporting and external validation of AI-based PC CDS tools is lacking and improvements in successful scale-up of AI in PC CDS across settings and patient populations will be difficult until standardized reporting and validation processes are in place. Additional foundational research is also needed on the effectiveness, efficiency, and acceptability of AI-based PC CDS.

In this landscape assessment, we identified various promising practices across scaling dimensions that can be applied broadly to begin to address challenges and optimize the use of AI for scaling PC CDS. In particular, key informants were adamant about viewing AI as a partner that augments human expertise and judgment rather than full automation and replacement of human skill—a point that was echoed in the literature. This “human in the loop” approach when using AI-based PC CDS systems will be critical to build trust in this emerging technology and catch potentially harmful errors that are not flagged by other guardrails. Increased use of explainable AI techniques has potential to build trust in the technology among patients and clinicians, thus facilitating interactions between humans and AI-based systems. In addition, increased clinician training on how to best leverage AI-based tools in their workflow will be beneficial to improve efficiency and cost savings for PC CDS processes while also ensuring these types of tools are used in such a way that benefits rather than harms patients. While we have identified several promising practices related to the use of AI-based PC CDS that require clinician action, more work is needed to understand how these practices can be incorporated into the workflow while minimizing clinician burden.

Finally, we identified several areas for future study that PC CDS researchers should focus on to build trust in AI-generated outputs and improve the ability of AI-based PC CDS to make processes more efficient and scalable. For example, our review identified a variety of practices that can address AI hallucinations/confabulations in AI-generated results. More research can be conducted to systematically evaluate these techniques and their effectiveness in reducing errors in AI-generated output, as well as exploration of other strategies for reducing the risk of potentially harmful errors and optimizing prompts to increase output quality. The field can also benefit from additional study on how to best optimize AI-based PC CDS into clinician workflows and patient lifeflows,[71] particularly identifying strategies for increased involvement of clinicians and patients in tool development. This should be complemented by further analyses of clinician and patient sentiments about the use of AI in healthcare and the effects of AI automation on quality of care. Importantly, identification of improved processes for reporting on AI-based tool development and external validation, alongside more robust evaluations of AI-based PC CDS tools, will be critical to scale up this technology more widely.

5. Conclusion

As the use of AI continues to become more ubiquitous in the healthcare field, its promise for facilitating the scaling of PC CDS will become more prominent. Our literature review and key informant interview findings underscore an optimistic, yet reserved, attitude about the use of AI to scale PC CDS, and we have highlighted key challenges, unintended consequences, promising practices, and strategies to ensure patient safety inherent to this process. Our findings highlight key areas where future research is needed to better understand the limitations of AI-based systems and test strategies to ameliorate these limitations in order to improve the utility of AI for scaling PC CDS. With continued testing and evaluation of AI-based PC CDS, along with sufficient regulatory guardrails to protect patients, the potential to leverage this emerging technology to scale PC CDS and make patient-centered care more efficient, safe, and accessible for patients will become increasingly clear.

Appendix. Landscape Assessment Methodology

This tool was developed collaboratively through extensive interactions between the Clinical Decision Support Innovation Collaborative (CDSiC) Implementation, Adoption, and Scaling Workgroup leads, Workgroup members, and the Workgroup support team. The methods that guided development of the landscape assessment within this collaboration are described below.

Research Questions

The following research questions informed development of the landscape assessment:

1. What are exemplary use cases for using artificial intelligence (AI) to scale patient-centered clinical decision support (PC CDS)?
2. What are the key considerations (e.g., potential pitfalls, solutions) for the use of AI to scale PC CDS?
   a. What are the challenges to using AI to scale PC CDS?
   b. How can AI be implemented in a patient-centered way, particularly at scale?
3. How should AI be used to scale PC CDS moving forward?

Scoping Literature Review

We identified peer-reviewed and grey literature to inform the landscape assessment in a rapid, multiphased approach.

We conducted one PubMed search to identify systematic reviews related to CDS and AI (Exhibit A1). After deduplication, our search yielded 83 peer-reviewed articles. We conducted two levels of screening—a title/abstract review and a full-text review. At each level, we assessed whether the reviewed records appeared to meet our eligibility criteria (Exhibit A2).

Records deemed *eligible* at the title/abstract level were screened again at the full-text review. We conducted a full-text review of 25 peer-reviewed articles identified from the PubMed search. We then determined the final list of eligible records for data abstraction, and for ineligible records, documented the reason(s) they were excluded. In total, 18 articles were included from the PubMed literature search performed.

In addition to the PubMed search, we identified relevant grey literature and additional peer-reviewed literature through recommendations by Workgroup members and CDSiC project team members. We also identified relevant literature from industry newsletters, such as the American Medical Informatics Association Informatics SmartBrief. Through these methods, we identified 30 additional peer-reviewed articles and grey literature identified for full-text review. After excluding additional records upon full-text review, 21 records were included through this mechanism.

In total, we screened 113 peer-reviewed journal articles and grey literature and included 39 articles.

Exhibit A1. Key Search Terms for the Scoping Literature Review

#1 CDS Search String	#2 Artificial Intelligence Search String	#3 Systematic Review Search String
decision support systems, clinical[Majr] OR “clinical decision support”[tiab]	Intelligence, Artificial[Majr] OR “artificial intelligence”[tiab] OR “machine learning”[tiab] OR “large language models”[tiab] OR “chatbots”[tiab] OR “chat bot”[tiab] NOT “radiology”[tiab] OR “imaging studies”[tiab]	("Cochrane database syst rev"[Journal] AND "review"[Publication Type]) OR "systematic review"[Publication Type] OR ("systematic review"[Title] OR "systematic literature review"[Title] OR "systematic scoping review"[Title] OR "meta-analysis"[Title])) NOT ("comment"[Publication Type] OR "protocol*"[Title])

Exhibit A2. Literature Search Inclusion and Exclusion Criteria

Inclusion Criteria	Exclusion Criteria
• Published/developed in 2018 or later • Research is United States-based • Discusses implementation of clinical decision support in inpatient, outpatient/ambulatory, or primary care setting • Discusses patient-facing and clinician-facing CDS • Explores patient and/or clinician perspectives of the use of AI to scale PC CDS	• Discusses AI used in imaging studies (unless the AI supports expedited review/scheduling). • Discusses AI used in radiology. • Discusses the use of AI for health plan decision making. • Discusses research applications of AI (i.e., no real-world implementation or scaling component). • Does not describe a CDS/PC CDS implementation (i.e., theoretical frameworks for using AI). • Describes only the development/evaluation of an AI technique (e.g., training datasets for machine learning) and excludes considerations for scaling (e.g., adoption, clinician perspectives, challenges). • Study focuses only on healthcare access and cost. • Study compares the outcome/performance of two AI techniques and excludes considerations for scaling (e.g., adoption, clinician perspectives, challenges). • Blog, book, study protocol, discussion forum, webinar.

Key Informant Interviews

We conducted key informant interviews with health systems leaders, researchers, clinicians, industry representatives, and patient representatives to further understand common implementation challenges, promising practices, and recommended approaches to ensuring patient safety when using AI to scale PC CDS.

We gathered feedback from two CDSiC team members and nine other stakeholders on key considerations, common implementation challenges, and recommended approaches to ensuring patient safety when implementing AI-based technologies to scale PC CDS. Key informants also described initiatives they were involved with or aware of that use AI to scale PC CDS.

We developed semistructured discussion guides that allowed the interviewer to steer the conversation toward each key informant's expertise. Each interview was conducted via Zoom, audio recorded, and lasted approximately 60 minutes. Transcript-style notes were created for each interview to support analysis.

Analysis and Synthesis

Three independent reviewers extracted the following data from the included literature from the scoping review: type of AI technique, scaling use case, implementation challenges and/or unintended consequences, risks and barriers to using AI to scale PC CDS, opportunities or promising practices for implementing AI to scale PC CDS, regulatory landscape insights, and approaches to ensuring patient safety.

After abstracting data from the literature, we qualitatively synthesized literature review findings using qualitative content analysis to identify key findings for the landscape assessment. We captured relevant challenges/risks, promising practices/opportunities, and example use cases for the following scaling dimensions: 1) automating processes, 2) reducing design and development barriers, 3) complementing direct or immediate clinician interaction, 4) facilitating sharing and replication of PC CDS, and 5) supporting cognitive processes and decision making. We synthesized input from key informant interviews to augment findings from the literature review and showcase illustrative use cases of AI being used to scale PC CDS along the five scaling dimensions.

References

[1] Dullabh P, Sandberg SF, Heaney-Huls K, Hovey LS, Lobach DF, Boxwala A, Desai PJ, Berliner E, Dymek C, Harrison MI, Swiger J, Sittig DF. Challenges and opportunities for advancing patient-centered clinical decision support: findings from a horizon scan. *J Am Med Inform Assoc.* 2022 Jun 14;29(7):1233-1243. doi: 10.1093/jamia/ocac059.

[2] Bright TJ, Wong A, Dhurjati R, Bristow E, Bastian L, Coeytaux RR, Samsa G, Hasselblad V, Williams JW, Musty MD, Wing L, Kendrick AS, Sanders GD, Lobach D. Effect of clinical decision-support systems: a systematic review. *Ann Intern Med.* 2012 Jul 3;157(1):29-43. doi: 10.7326/0003-4819-157-1-201207030-00450.

[3] Marcial LH, Blumenfeld B, Harle C, Jing X, Keller MS, Lee V, Lin Z, Dover A, Midboe AM, Al-Showk S, Bradley V, Breen J, Fadden M, Lomotan E, Marco-Ruiz L, Mohamed R, O'Connor P, Rosendale D, Solomon H, Kawamoto K. Barriers, facilitators, and potential solutions to advancing interoperable clinical decision support: multi-stakeholder consensus recommendations for the opioid use case. *AMIA Annu Symp Proc*. 2020 Mar 4;2019:637-646.

[4] Chen M, Zhang B, Cai Z, Seery S, Gonzalez MJ, Ali NM, Ren R, Qiao Y, Xue P, Jiang Y. Acceptance of clinical artificial intelligence among physicians and medical students: A systematic review with cross-sectional survey. *Front Med (Lausanne).* 2022 Aug 31;9:990604. doi: 10.3389/fmed.2022.990604.

[5] Zusterzeel R, Goldstein B, Evans B, Roades T, Mercon K, Silcox C. Evaluating AI-enabled clinical decision and diagnostic support tools using real-world data. Duke Margolis Center for Health Policy. March 11, 2022. Accessed March 31, 2024. https://healthpolicy.duke.edu/publications/evaluating-ai-enabled-clinical-decision-and-diagnostic-support-tools-using-real-world.

[6] Lenharo M. An AI revolution is brewing in medicine. What will it look like? Nature. October 2023. Accessed March 31, 2024. https://www.nature.com/articles/d41586-023-03302-0.

[7] Ayers JW, Poliak A, Dredze M. Comparing physician and artificial intelligence chatbot responses to patient questions posted to public social media forum. *JAMA Intern Med.* 2023 April;183(6):589-596. doi: 10.1001/jamainternmed.2023.1838.

[8] Anjara SG, Janik A, Dunford-Stenger A, Mc Kenzie K, Collazo-Lorduy A, Torrente M, Costabello L, Provencio M. Examining explainable clinical decision support systems with think aloud protocols. *PLoS One.* 2023 Sep 14;18(9):e0291443. doi: 10.1371/journal.pone.0291443.

[9] Liaw ST, Liyanage H, Kuziemsky C, Terry AL, Schreiber R, Jonnagaddala J, de Lusignan S. Ethical use of electronic health record data and artificial intelligence: recommendations of the primary care informatics working group of the international medical informatics association. *Yearb Med Inform.* 2020 Aug;29(1):51-57. doi: 10.1055/s-0040-1701980.

[10] Kennedy S. New healthcare AI framework incorporates medical knowledge, values. TechTarget Health IT Analytics. December 1, 2023. Accessed March 31, 2024. https://healthitanalytics.com/news/new-healthcare-ai-framework-incorporates-medical-knowledge-values#:~:text=December%2001%2C%202023%20%2D%20A%20novel,the%20technology%20into%20clinical%20settings.

[11] Congress.gov. "H.R.6216 - 116th Congress (2019-2020): National Artificial Intelligence Initiative Act of 2020." March 12,2020. Accessed March 31, 2024. https://www.congress.gov/bill/116th-congress/house-bill/6216.

[12] National AI Advisory Committee (NAIAC). FAQs on foundation models and generative AI. August 8,2023 . Accessed March 31, 2204. https://ai.gov/wp-content/uploads/2023/09/FAQs-on-Foundation-Models-and-Generative-AI.pdf.

[13] United States Department of Health and Human Services. Trustworthy AI Playbook. September, 2021. Accessed March 31, 2024. https://www.hhs.gov/sites/default/files/hhs-trustworthy-ai-playbook.pdf.

[14] Milat A, Lee K, Conte K, Grunseit A, Wolfenden L, van Nassau F, Orr N, Sreeram P, Bauman A. Intervention scalability assessment tool: A decision support tool for health policy makers and implementers. *Health Res Policy Syst.* 2020 Jan 3;18(1):1. doi: 10.1186/s12961-019-0494-2.

[15] MedStar Health research institute. Quantifying efficiencies gained through shareable clinical decision support resources - final report. (prepared under contract 233-20-1500022I.) AHRQ Publication No. 20-0018. Rockville, MD: Agency for Healthcare Research and Quality. December 2019.

[16] Li R, Kumar A, Chen JH. How chatbots and large language model artificial intelligence systems will reshape modern medicine: fountain of creativity or pandora's box? *JAMA Intern Med.* 2023;183(6):596–597. doi:10.1001/jamainternmed.2023.1835.

[17] Allen M. Are we missing the mark with generative AI? MedPage Today's KevinMD. October 24, 2023. Accessed Mach 31, 2024. https://www.kevinmd.com/2023/10/are-we-missing-the-mark-with-generative-ai.html.

[18] Nayak A, Alkaitis MS, Nayak K, Nikolov M, Weinfurt KP, Schulman K. Comparison of history of present illness summaries generated by a chatbot and senior internal medicine residents. *JAMA Intern Med.* 2023 Sep 1;183(9):1026-1027. doi: 10.1001/jamainternmed.2023.2561.

[19] Guevara M, Chen S, Thomas S, et al. Large language models to identify social determinants of health in electronic health records. *npj Digit. Med.* 2024 July 6. https://doi.org/10.1038/s41746-023-00970-0.

[20] Palmer K. Patients' social needs often get lost in health records. Generative AI could help. Stat+. January 11, 2024. Accessed April 12, 2024. https://www.statnews.com/2024/01/11/health-records-ai-tools-social-determinants-health/.

[21] Afridi MJ, Farooq M. OG-Miner: an intelligent health tool for achieving millennium development goals (mdgs) in m-health environments. Presented at: *44th Hawaii International Conference on System Sciences*. January 7-11, 2022. Kauai, HI. doi: 10.1109/HICSS.2011.320.

[22] Goddard K, Roudsari A, Wyatt JC. Automation bias: a systematic review of frequency, effect mediators, and mitigators. *J Am Med Inform Assoc.* 2012 Jan-Feb;19(1):121-7. doi: 10.1136/amiajnl-2011-000089..

[23] US Food and Drug Administration. Clinical Decision Support Software: Guidance for Industry and Food and Drug Administration Staff. September 2022. Accessed April 12, 2024. https://www.fda.gov/media/109618/download.

[24] Clement J, Maldonado AQ. Augmenting the transplant team with artificial intelligence: toward meaningful ai use in solid organ transplant. *Front Immunol.* 2021 Jun 11;12:694222. doi: 10.3389/fimmu.2021.694222.

[25] Benzinger L, Ursin F, Balke WT, Kacprowski T, Salloch S. Should artificial intelligence be used to support clinical ethical decision-making? a systematic review of reasons. *BMC Med Ethics.* 2023 Jul 6;24(1):48. doi: 10.1186/s12910-023-00929-6.

[26] Weingart SN, Toth M, Sands DZ, et al. Physicians' decisions to override computerized drug alerts in primary care. *Arch Intern Med 2003*; 163 (21): 2625–31.

[27] Powers EM, Shiffman RN, Melnick ER, et al. Efficacy and unintended consequences of hard-stop alerts in electronic health record systems: a systematic review. *J Am Med Inform Assoc.* 2018; 25 (11): 1556–66.

[28] McCoy AB, Russo EM, Johnson KB, et al. Clinician collaboration to improve clinical decision support: the Clickbusters initiative. *J Am Med Inform Assoc.* 2022; 29 (6): 1050–9.

[29] Croskerry P. From mindless to mindful practice – cognitive bias and clinical decision making. *N Engl J Med.* 2013; 368 (26): 2445–8.

[30] Liu S, Wright AP, Patterson BL, Wanderer JP, Turer RW, Nelson SD, McCoy AB, Sittig DF, Wright A. Using AI-generated suggestions from ChatGPT to optimize clinical decision support. *J Am Med Inform Assoc*. 2023 Jun 20;30(7):1237-1245. doi: 10.1093/jamia/ocad072.

[31] Liu S, Kawamoto K, Del Fiol G, et al. The potential for leveraging machine learning to filter medication alerts. *J Am Med Informatics Assoc.* 2022; 29 (5): 891–9.

[32] Schwartz JM, Moy AJ, Rossetti SC, Elhadad N, Cato KD. Clinician involvement in research on machine learning-based predictive clinical decision support for the hospital setting: A scoping review. *J Am Med Informatics Assoc.* 2021;28(3); 653–663. https://doi.org/10.1093/jamia/ocaa296.

[33] Kennedy EE, Bowles KH, Aryal S. Systematic review of prediction models for postacute care destination decision-making. *J Am Med Inform Assoc*. 2021.

[34] Giordano C, Brennan M, Mohamed B, Rashidi P, Modave F, Tighe P. Accessing artificial intelligence for clinical decision-making. *Front Digit Health.* 2021 Jun 25;3:645232. doi: 10.3389/fdgth.2021.645232.

[35] Balla Y, Tirunagari S, Windridge D. Pediatrics in artificial intelligence era: a systematic review on challenges, opportunities, and explainability. *Indian Pediatr.* 2023 Jul 15;60(7):561-569.

[36] Akbar F, Mark G, Warton EM, et al. Physicians' electronic inbox work patterns and factors associated with high inbox work duration. *J Am Med Inform Assoc*. 2021;28(5):923-930. doi:10.1093/jamia/ocaa229.

[37] Milne-Ives M, de Cock C, Lim E, et al. The effectiveness of artificial intelligence conversational agents in health care: systematic review. *J Med Internet Res.* 2020;22(10):e20346. doi:10.2196/20346.

[38] Bock S. Introducing Dr. Chatbot. UC San Diego Today. June 15, 2023. Accessed April 12, 2024. https://today.ucsd.edu/story/introducing-dr-chatbot.

[39] Defreitas M. Can AI help monitor chemotherapy side effects? Healthleaders. January 8. 2024. Accessed April 12, 2024. https://www.healthleadersmedia.com/technology/can-ai-help-monitor-chemotherapy-side-effects.

[40] Perna G. Virtua Health is bringing on AI therapists amid clinician shortage. Modern Healthcare. November 7, 2023. Accessed April 12, 2024. https://www.modernhealthcare.com/digital-health/virtua-health-adopts-ai-chatbots-mental-health#:~:text=The%20academic%20health%20system%20Virtua,care%20to%20behavioral%20health%20patients.

[41] Davenport T, Kalakota R. The potential for artificial intelligence in healthcare. *Future Healthc J.* 2019 Jun;6(2):94-98. doi: 10.7861/futurehosp.6-2-94.

[42] Caballero-Ruiz E, García-Sáez G, Rigla M, Villaplana M, Pons B, Hernando ME. A web-based clinical decision support system for gestational diabetes: automatic diet prescription and detection of insulin needs. *Int J Med Inform.* 2017 Jun;102:35-49. doi: 10.1016/j.ijmedinf.2017.02.014.

[43] Du Y, McNestry C, Wei L, Antoniadi AM, McAuliffe FM, Mooney C. Machine learning-based clinical decision support systems for pregnancy care: a systematic review. *Int J Med Inform*. 2023 May;173:105040. doi: 10.1016/j.ijmedinf.2023.105040.

[44] Lee TC, Shah NU, Haack A, Baxter SL. Clinical implementation of predictive models embedded within electronic health record systems: a systematic review. *Informatics (MDPI).* 2020;7(3). doi: 10.3390/informatics7030025.

[45] Buchlak QD, Esmaili N, Leveque JC, Farrokhi F, Bennett C, Piccardi M, Sethi RK. Machine learning applications to clinical decision support in neurosurgery: an artificial intelligence augmented systematic review. *Neurosurg Rev.* 2020 Oct;43(5):1235-1253. doi: 10.1007/s10143-019-01163-8.

[46] Davidson L, Boland MR. Towards deep phenotyping pregnancy: a systematic review on artificial intelligence and machine learning methods to improve pregnancy outcomes. *Brief Bioinform.* 2021 Sep 2;22(5):bbaa369. doi: 10.1093/bib/bbaa369.

[47] AIContentfy. The impact of AI on content translation and localization. November 6, 2023. Accessed April 30, 2024. https://aicontentfy.com/en/blog/impact-of-ai-on-content-translation-and-localization.

[48] AHRQ Digital Healthcare Research. A clinical trial to validate an automated online language interpreting tool with Hispanic patients who have limited English proficiency. Accessed April 30, 2024. https://digital.ahrq.gov/ahrq-funded-projects/clinical-trial-validate-automated-online-language-interpreting-tool-hispanic.

[49] Yellowlees PM, Burke Parish M, Iosif A-M, Gonzalez A, Fisher A, Chan S, Martini J, Sciolla A. Chun R, Tougas H, Shahrvini T. AHRQ grantee final progress report: A clinical trial to validate an automated online language interpreting tool with Hispanic patients who have limited English proficiency. 2022. Accessed April 30, 2024. https://digital.ahrq.gov/sites/default/files/docs/citation/r01hs024949-yellowlees-final-report-2022.pdf.

[50] Dullabh PM, Desai PJ, Gordon JR, Leaphart D, Wilson KS, Richesson RL, Boxwala AA, and the CDSiC Standards and Regulatory Frameworks Workgroup. Standards and Regulatory Frameworks Workgroup: Environmental Scan. Prepared under Contract No. 75Q80120D00018. AHRQ Publication No. 23-0029. Rockville, MD: Agency for Healthcare Research and Quality; January 2023.

[51] Oehring R, Ramasetti N, Ng S, Roller R, Thomas P, Winter A, Maurer M, Moosburner S, Raschzok N, Kamali C, Pratschke J, Benzing C, Krenzien F. (2023). Use and accuracy of decision support systems using artificial intelligence for tumor diseases: a systematic review and meta-analysis. *Frontiers in Oncology.* 2023 Oct 3;13, 1224347. https://doi.org/10.3389/fonc.2023.1224347.

[52] Akay EMZ, Hilbert A, Carlisle BG, Madai VI, Mutke MA, Frey D. Artificial intelligence for clinical decision support in acute ischemic stroke: a systematic review. *Stroke.* 2023 Jun;54(6):1505-1516. doi: 10.1161/STROKEAHA.122.041442.

[53] Oikonomidi T, Ravaud P, LeBeau J, Tran VT. (2023). A systematic scoping review of just-in-time, adaptive interventions finds limited automation and incomplete reporting. *Journal of Clinical Epidemiology*, 154, 108–116. https://doi.org/10.1016/j.jclinepi.2022.12.006.

[54] Yang C, Kors JA, Ioannou S, John LH, Markus AF, Rekkas A, de Ridder MAJ, Seinen TM, Williams RD, Rijnbeek PR. Trends in the conduct and reporting of clinical prediction model development and validation: a systematic review. *Journal of the American Medical Informatics Association. 2022;*29(5), 983–989. https://doi.org/10.1093/jamia/ocac002.

[55] Muralitharan S, Nelson W, Di S, McGillion M, Devereaux PJ, Barr NG, Petch J. Machine learning–based early warning systems for clinical deterioration: systematic scoping review. *J Med Internet Res.* 2021 April;23(2):e25187. doi: 10.2196/25187.

56 Moazemi S, Vahdati S, Li J, Kalkhoff S, Castano LJV, Dewitz B, et al. Artificial intelligence for clinical decision support for monitoring patients in cardiovascular ICUs: a systematic review. *Front Med (Lausanne).* 2023 March;10:1109411. doi: 10.3389/fmed.2023.1109411.

57 Liu X, Cruz Rivera S, Moher D, Calvert MJ, Denniston AK; SPIRIT-AI and CONSORT-AI Working Group. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. *Lancet Digit Health*. 2020 Oct;2(10):e537-e548. doi: 10.1016/S2589-7500(20)30218-1.

58 Desai PJ, Osheroff JA, Ryan S, Heaney-Huls K, Jiménez F, McCoy AB, Dullabh PM, and the CDSiC Scaling, Measurement, and Dissemination of CDS Workgroup. Scaling, Measurement, and Dissemination of CDS Workgroup: PC CDS Planning, Implementation, and Reporting User Guide. Prepared under Contract No. 75Q80120D00018. AHRQ Publication No. 23-0066. Rockville, MD: Agency for Healthcare Research and Quality; August 2023.

59 Van de Velde S, Kunnamo I, Roshanov P, Kortteisto T, Aertgeerts B, Vandvik PO, Flottorp S; GUIDES expert panel. The GUIDES checklist: development of a tool to improve the successful use of guideline-based computerised clinical decision support. *Implement Sci.* 2018 Jun 25;13(1):86. doi: 10.1186/s13012-018-0772-3.

60 Yasar K. Black Box AI. TechTarget. Accessed April 30, 2024. https://www.techtarget.com/whatis/definition/black-box-AI#:~:text=Black%20box%20AI%20is%20any,to%20how%20they%20were%20reached.

61 Goodman KE, Yi PH, Morgan DJ. AI-generated clinical summaries require more than accuracy. *JAMA.* Published online January 29, 2024. doi:10.1001/jama.2024.0555.

62 Jeong HK, Park C, Henao R, Kheterpal M. Deep learning in dermatology: a systematic review of current approaches, outcomes, and limitations. *JIV Innov.* 2023; 3(1):100150. doi:10.1016/j.xjidi.2022.

63 Systematic Review: Impact of healthcare algorithms on racial and ethnic disparities in health and healthcare. Content last reviewed. Effective Health Care Program, Agency for Healthcare Research and Quality, Rockville, MD. December 2023. Accessed April 30, 2024. https://effectivehealthcare.ahrq.gov/products/racial-disparities-health-healthcare/research.

64 Strohm L, Hehakaya C, Ranschaert ER, Boon WPC, Moors EHM. Implementation of artificial intelligence (AI) applications in radiology: hindering and facilitating factors*.* *Eur Radiol.* 2020 Oct;30(10):5525-5532. doi: 10.1007/s00330-020-06946-y.

65 United States Department of Health and Human Services. HHS finalizes rule to advance health it interoperability and algorithm transparency. December 13, 2023. Accessed May 1, 2024. https://www.hhs.gov/about/news/2023/12/13/hhs-finalizes-rule-to-advance-health-it-interoperability-and-algorithm-transparency.html.

66 Google Cloud. What are AI hallucinations? Accessed May 1, 2024. https://cloud.google.com/discover/what-are-ai-hallucinations#:~:text=AI%20hallucinations%20are%20incorrect%20or,used%20to%20train%20the%20model

67 Hatem R, Simmons B, Thornton JE. A call to address AI “hallucinations” and how healthcare professionals can mitigate their risks. *Cureus.* 2023 Sep 5;15(9):e44720. doi: 10.7759/cureus.44720.

68 World Health Organization. WHO calls for safe and ethical AI for health. May 16, 2023. Accessed May 1, 2024. https://www.who.int/news/item/16-05-2023-who-calls-for-safe-and-ethical-ai-for-health#:~:text=The%206%20core%20principles%20identified,AI%20that%20is%20responsive%20and.

[69] The Joint Commission. Quick safety 23: implicit bias in health care. April 11, 2016. Accessed May 1, 2024. https://www.jointcommission.org/resources/news-and-multimedia/newsletters/newsletters/quick-safety/quick-safety-issue-23-implicit-bias-in-health-care/implicit-bias-in-health-care/.

[70] Coalition for Health AI. Providing guidelines for the responsible use of AI in healthcare: virtual workgroup session #1: bias, equity, and fairness. June 21, 2022. Accessed May 1, 2024. https://www.coalitionforhealthai.org/papers/Virtual_Workgroup%20Session%201-Bias_Fairness_and_Equity.pdf.

[71] Desai PJ, Osheroff JA, Jiménez F, Heaney-Huls K, Ryan S, McCoy AB, Dullabh PM, CDSiC Scaling, Measurement, and Dissemination of CDS Workgroup. Scaling, Measurement, and Dissemination of CDS Workgroup: Approaches to Measuring Patient-Centered CDS Workflow and Lifeflow Impact. Prepared under Contract No. 75Q80120D00018. AHRQ Publication No. 23-0061. Rockville, MD: Agency for Healthcare Research and Quality; July 2023.