

Primer for artificial intelligence in primary care

Jacqueline K. Kueper MSc

Artificial intelligence (AI) and machine learning^{1,2*} are transforming clinical practice.³⁻⁹ There is a need for tailored education for end users and a need for more primary care stakeholder engagement during development and testing of AI-driven tools.³ An in-depth or technical understanding of AI is not required for clinicians to be able to make meaningful contributions to AI-related initiatives; however, most existing AI education resources focus on methods or are not specific to primary care. This primer is divided into 4 sections with the goal to fill this gap and support brainstorming and early-stage interdisciplinary research team discussions about potential uses and applications of AI for clinical practice and research. A secondary purpose is for this primer to serve as an introductory resource on AI for primary care stakeholders.

Section 1 introduces AI and machine learning; section 2 describes 3 main types of tasks that AI methods are currently capable of performing (**Figure 1**), with examples of primary care applications; section 3 touches on specialized techniques that might be needed to perform tasks with language and image data; and section 4 mentions earlier approaches to AI, before large data sources were available.

Section 1: AI and machine learning

The field of AI is broad and rapidly expanding. The field is centred on how computers might be able to perform humanlike “intelligent tasks,” such as summarizing large amounts of information or making inferences about a situation. The 3 main types of tasks that AI methods are currently capable of performing are summarized in **Figure 1** and will be elaborated on with examples in section 2.

Details on how AI is able to perform these tasks are not needed to brainstorm potential uses of AI or to see its potential value in practice; however, it might be interesting to know that most AI functions by using math to find and fill in patterns, whereby parts of the math equations are specified by humans (ie, AI developers) and parts of the equations are learned from data. Data might originally consist of numbers, letters, audio, or images; these data can be carefully transformed into formats that AI is able to process.

Today, most of the AI functionality that will be described in sections 2 and 3 is achieved using machine learning, which represents a collection of methods able to perform intelligent tasks by identifying patterns in data. The

potential for machine learning to support research and clinical care comes largely from its ability to capture much more complex relationships among data elements than a human or simple algorithm would typically be capable of capturing. Although machine learning can be applied to small amounts of data, it is large data sources, such as electronic medical record (EMR) databases storing years of patient information, that provide enough data for machine-learning methods to be able to uncover these complex relationships. Artificial intelligence has been receiving so much more attention lately because increases in available data over the past years have sparked advancements in machine-learning method development, so that good performance is now feasible for real-world applications.

Although this primer presents tasks separately, in practice the tasks and the AI methods used to perform them might overlap. Any given task might be performed by several different types of machine-learning methods, and any given method might be applicable for different tasks. For example, neural networks, which are a type of AI method belonging to the machine-learning subfield of deep learning, have been used to perform all 3 types of tasks described above. Knowledge of specific methods is not necessary for the remainder of this primer; however, those interested in more technical aspects might look to the first 3 cited references as a starting point.¹⁻³

Section 2: key tasks that can be supported by AI

Predicting prespecified outcomes. The first task is concerned with using existing, known information (ie, data inputs) to predict or estimate unknown information, or *outcomes*. These outcomes might be current (eg, does this person have undiagnosed diabetes right now?) or future events (eg, will this person develop diabetes in the next year?). **Figure 2** depicts this key task.

These types of tasks are primarily performed by supervised machine-learning methods, which look for patterns that relate input data to a prespecified outcome. Patterns are learned during a training phase whereby an AI algorithm has access to labeled data for which both inputs (eg, clinical histories captured through EMRs) and outcomes (eg, existing diabetes diagnosis status) are known. The end product would be a tool that relies on known input data (eg, EMR data from patients new to the clinic who do not have an existing diabetes diagnosis) to make useful predictions in situations where the outcome is truly unknown (eg, whether they will develop diabetes). These types of AI-driven tools can be classified as *locked* or *adaptive*, depending on whether the patterns underlying predictions are allowed to actively change as a tool is used over time.

*Examples cited throughout the text do not represent endorsements nor speak to the quality of the tools.

Predictions might be made at an individual or group level. For example, a model that predicts an individual's risk of progressing from prediabetes to diabetes might support clinical decisions related to prevention or screening,¹⁰ while a model that predicts incidence of diabetes in a population might inform resource planning or policy decisions.^{11,12} The intended level of use should be identified early in the model development process.

Example of prediction problem application: A tool¹³ with potential relevance to primary care was created to optimize clinic scheduling and reduce wait times for patients. Inputs include basic demographic and appointment data from the scheduling system; predicted outcomes include no-shows and appointment cancellations. Administrative

Figure 1. Key tasks supported by AI

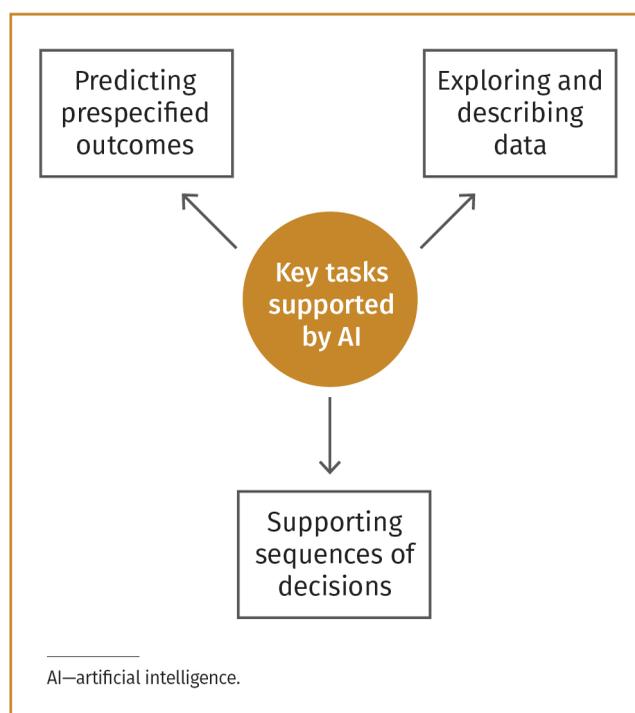
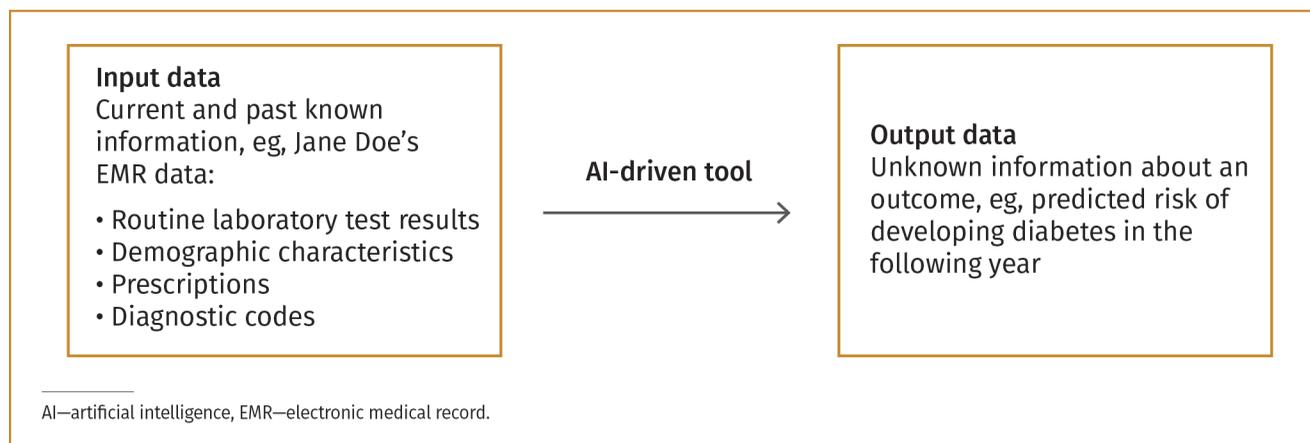


Figure 2. Predicting prespecified outcomes



staff can use the tool to identify appointments where there is a high risk of the patient missing or canceling, which provides an opportunity to follow up with the patient well in advance of the scheduled appointment date.

Exploring and describing data. The second type of task can be performed on similar data sources as the first task, but the patterns of interest are no longer related to predicting a prespecified outcome. Rather, the objective here is to capture overall structure, trends, or other kinds of descriptive information about the data. A common task is clustering, which could be used to better understand a data set and generate hypotheses about the population represented in those data (**Figure 3**).¹⁴

The types of AI methods used to perform exploratory and descriptive data tasks are often called *unsupervised machine-learning methods*.

Supporting sequences of decisions. The third type of task is to support sequences of decisions for both short- and long-term outcomes. This task is performed using data over time. At any given time point (eg, a clinical appointment) there might be multiple possible actions to take (eg, treatment options); AI methods can be used to predict what action will optimize a “reward function” that represents the outcome or outcomes of interest. After a decision is made about what action to take, feedback is captured and used to update predictions about what the next best action or actions might be (**Figure 4**).

Tasks related to supporting decisions or actions over time are usually performed with methods falling within the machine-learning subfield of reinforcement learning.

Section 3: specialized methods

There are specialized subfields of AI concerned with analyzing and interpreting data types, such as language and image data, which might require extra or different methods and approaches to be able to support what would be considered an “intelligent task” for that type of data.

Figure 3. Exploring and describing data

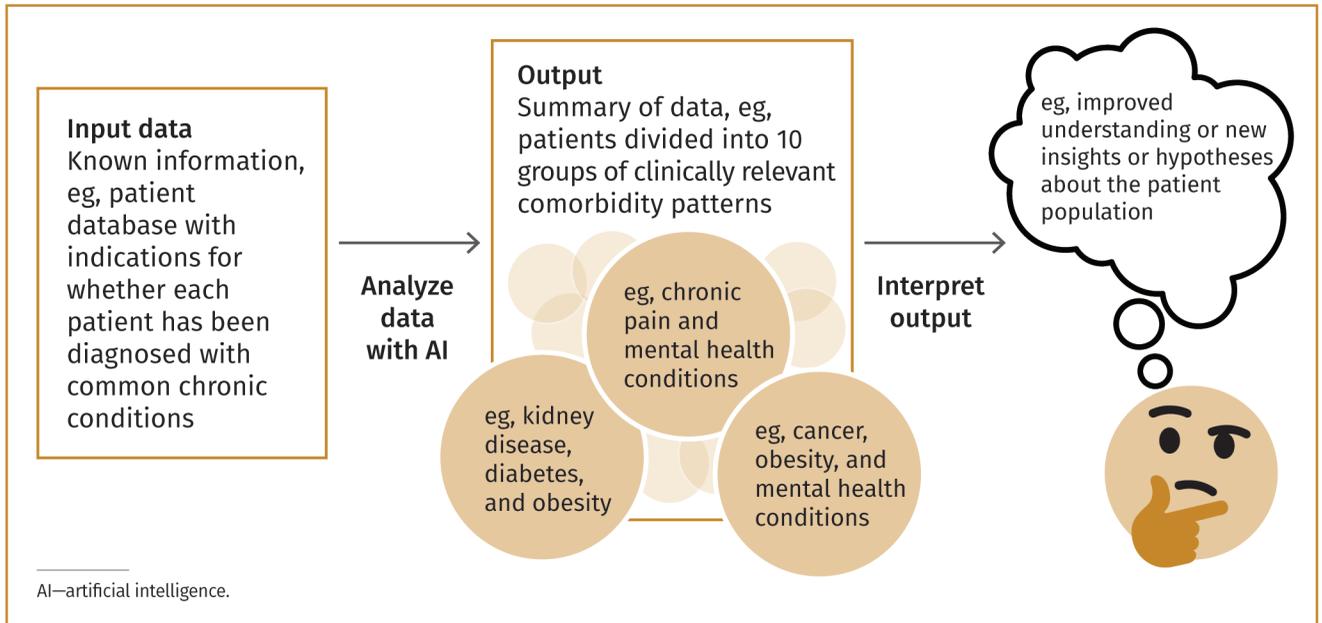
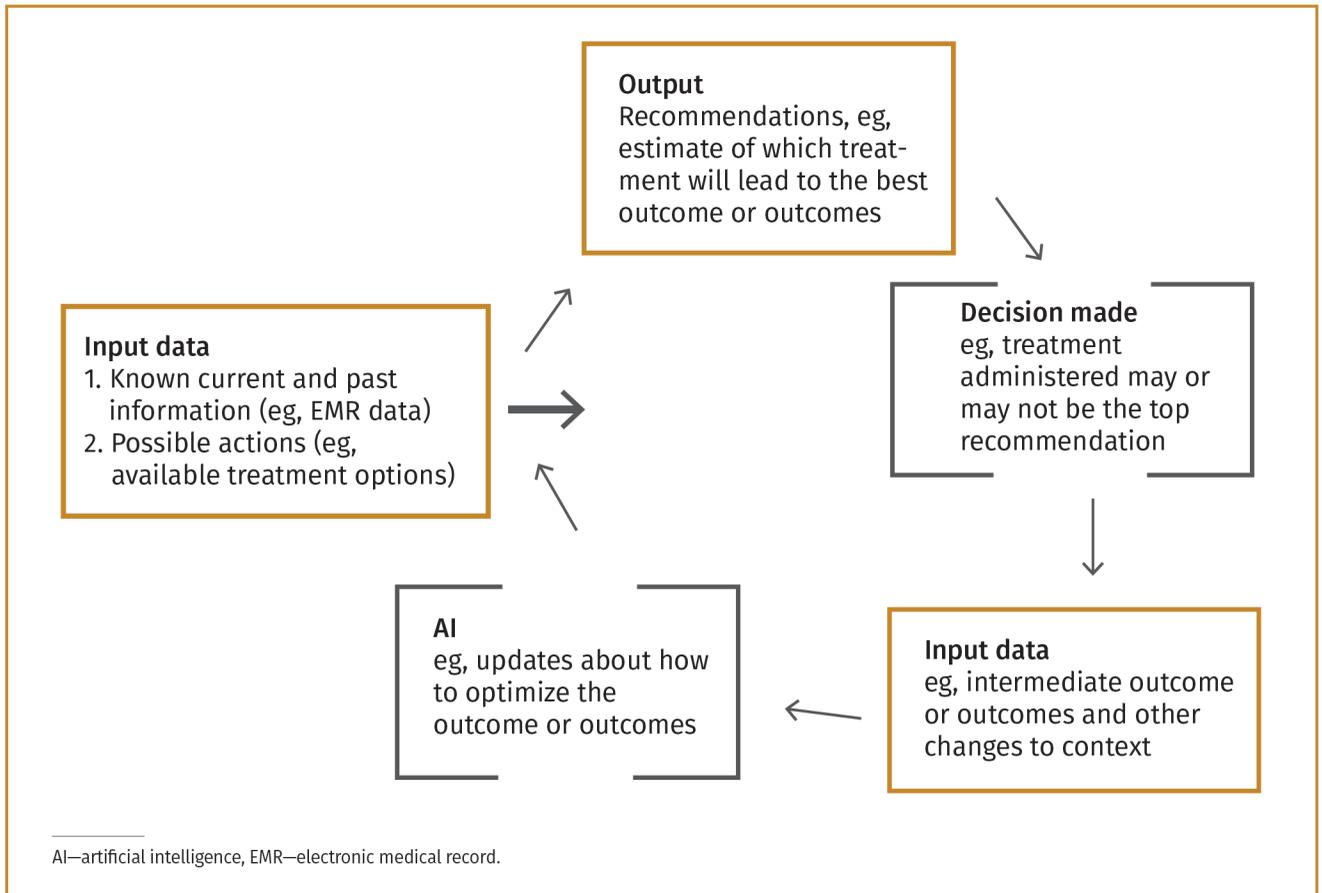


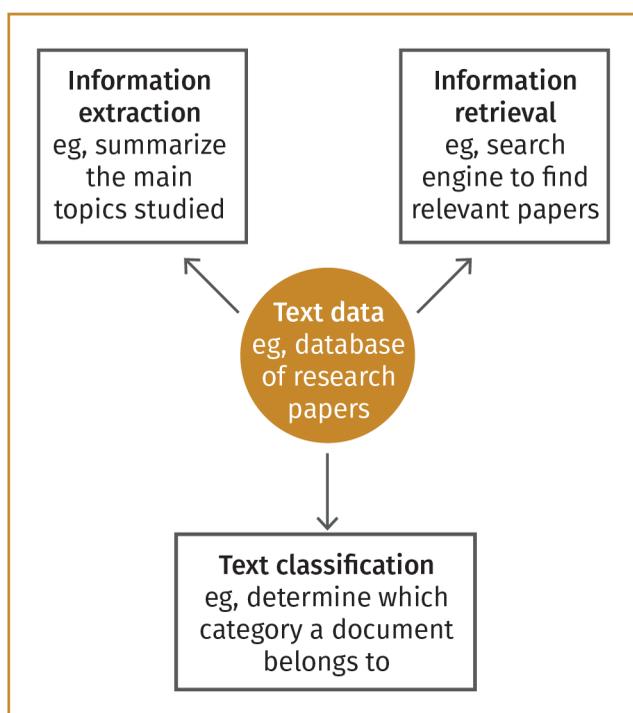
Figure 4. Supporting sequences of decisions



Analyzing language data. The subfield of AI called *natural language processing* is concerned with identifying and acting on patterns in language data. A common application for oral language data is speech recognition, such as question-and-answer tasks in automated telephone calls. Common applications with written language data are presented in **Figure 5**.

Example of natural language processing clinical application: A tool¹⁵⁻¹⁷ that relies on natural language processes was developed to assist with physician documentation during a clinical encounter. It can automatically extract and record key notes and diagnoses from a conversation between a clinician and a patient.

Figure 5. Common applications with written language data



Analyzing image data. The subfield of AI called *computer vision* focuses on processing and interpreting image data. Clinical applications of computer vision usually involve the end goal of classification, such as predicting whether the contents of a picture are within expected healthy variation or are concerning and warrant further investigation.

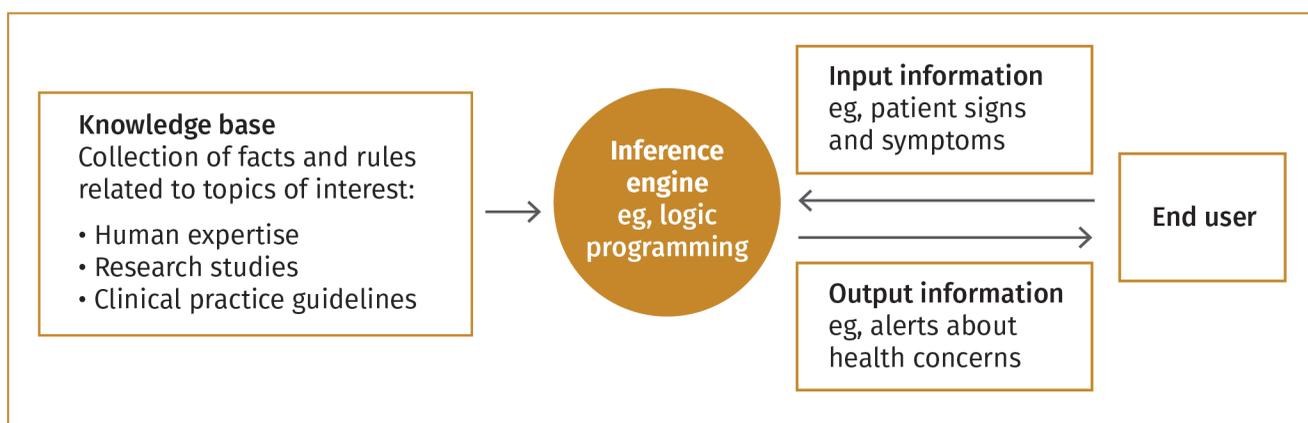
Example of computer vision application: An example tool¹⁸ was designed to identify diabetic retinopathy within 60 seconds and works with most specialized eye cameras that provide high-quality images. It makes recommendations regarding whether to consult with an eye care specialist.

Section 4: before and beyond data-driven systems

From data-centric to rule-centric methods. Often AI is thought about in tandem with the type and quality of available data. Data are integral to most cutting-edge AI methods, including those currently being used to perform the tasks presented above; however, this was not always the case. The field of AI can be viewed as existing along a spectrum from data-centric methods to rule-centric methods. At one end of the spectrum could be a heavily “data-driven” machine-learning tool whereby decisions throughout the entire development process, from what data should be considered as inputs to what outputs will look like, are derived almost entirely from data. At the other end of the spectrum would be some of the earliest AI tools, including expert systems from before the “big data” era. To perform tasks such as suggesting diagnoses or treatments, expert systems could use logic programming to work through collections of rules, such as if-then statements extracted from clinical guidelines (**Figure 6**). Modern expert systems might rely on a combination of rule-centric and data-centric methods.

Historical example of expert system application: The earliest example of AI developed to support clinical practice, MYCIN, was an expert system developed in the 1970s to support treatment decisions for bacterial infections. The knowledge base included hundreds of rules extracted from

Figure 6. Expert system structure



human experts, textbooks, and case reports. Each rule had 2 main parts: a precondition and an action. Precondition information was obtained through a question-and-answer process with a clinician until all relevant information had been entered, at which point MYCIN would provide suggested diagnoses and treatments.¹⁹

Conclusion

This primer provided a conceptual overview of the types of tasks AI can perform with examples of applications in primary care. It might serve as an educational resource to be used for personal interest or to facilitate early-stage discussions between interdisciplinary teams. As discussions progress, it will be important to introduce additional considerations necessary for safe and effective applications of AI, such as bias, equity, explainability, generalizability, and data quality and privacy. 

Jacqueline K. Kueper is CFPC-AMS TechForward Fellow at the College of Family Physicians of Canada and a PhD candidate in the Department of Epidemiology and Biostatistics and the Department of Computer Science at Western University in London, Ont.

Acknowledgment

I acknowledge **Ravninder Bahniwal**, MPH, who helped with diagram formatting and primary care examples; and **Dan Lizotte**, PhD, **Ravninder Bahniwal**, MPH, **Amanda Terry**, PhD, **Leslie Meredith**, and **Jason Black**, MSc, all of whom further helped improve this resource through editing and constructive criticism. AMS Healthcare supports the CFPC-AMS TechForward Fellow position that I hold.

Competing interests

None declared

Correspondence

Jacqueline K. Kueper; e-mail jkueper@uwo.ca

The opinions expressed in commentaries are those of the authors. Publication does not imply endorsement by the College of Family Physicians of Canada

References

1. Russell S, Norvig P. *Artificial intelligence: a modern approach*. 3rd ed. Upper Saddle River, NJ: Pearson Education Inc; 2010.

2. Hastie T, Tibshirani R, Friedman J. *The elements of statistical learning: data mining, inference, and prediction*. 2nd ed. New York, NY: Springer; 2008.
3. Kueper JK, Terry AL, Zwarenstein M, Lizotte DJ. Artificial intelligence and primary care research: a scoping review. *Ann Fam Med* 2020;18(3):250-8.
4. Liaw W, Kakadiaris IA. Primary care artificial intelligence: a branch hiding in plain sight. *Ann Fam Med* 2020;18(3):194-5.
5. Wiens J, Shenoy ES. Machine learning for healthcare: on the verge of a major shift in health-care epidemiology. *Clin Infect Dis* 2018;66(1):149-53.
6. Sendak M, Gao M, Nichols M, Lin A, Balu S. Machine learning in health care: a critical appraisal of challenges and opportunities. *EGEMs (Wash DC)* 2019;7(1):1.
7. Beam AL, Kohane IS. Big data and machine learning in health care. *JAMA* 2018;319(13):1317-8.
8. Celi LA, Fine B, Stone DJ. An awakening in medicine: the partnership of humanity and intelligent machines. *Lancet Digit Health* 2019;1(6):e255-7. Epub 2019 Sep 26.
9. Lin SY, Mahoney MR, Sinsky CA. Ten ways artificial intelligence will transform primary care. *J Gen Intern Med* 2019;34(8):1626-30. Epub 2019 May 14.
10. Cahn A, Shoshan A, Sagiv T, Yescharim R, Goshen R, Shalev V, et al. Prediction of progression from pre-diabetes to diabetes: development and validation of a machine learning model. *Diabetes Metab Res Rev* 2020;36(2):e3252. Epub 2020 Jan 14.
11. Rosella LC, Manuel DG, Burchill C, Stukel TA; PHAT-DM Team. A population-based risk algorithm for the development of diabetes: development and validation of the Diabetes Population Risk Tool (DPoRT). *J Epidemiol Community Health* 2011;65(7):613-20. Epub 2010 Jun 1.
12. Rosella LC, Peirson L, Bornbaum C, Kotnowski K, Leberbaum M, Fransoo R, et al. Supporting collaborative use of the Diabetes Population Risk Tool (DPoRT) in health-related practice: a multiple case study research protocol. *Implement Sci* 2014;9(1):35.
13. *Optimize your schedule with Attendance Predictor*. Orlando, FL: Mend. Available from: <https://www.mend.com/no-show-predictor/>. Accessed 2021 Nov 5.
14. Newcomer SR, Steiner JF, Bayliss EA. Identifying subgroups of complex patients with cluster analysis. *Am J Manag Care* 2011;17(8):e324-32.
15. Khattak FK, Jebblee S, Crampton N, Mamdani M, Rudzicz F. AutoScribe: extracting clinically pertinent information from patient-clinician dialogues. *Stud Health Technol Inform* 2019;264:1512-3.
16. Jebblee S, Khattak FK, Crampton N, Mamdani M, Rudzicz F. Extracting relevant information from physician-patient dialogues for automated clinical note taking. In: *Proceedings of the Tenth International Workshop on Health Text Mining and Information Analysis (LOUHI 2019)*. Hong Kong: Association for Computational Linguistics; 2019. p. 65-74.
17. *AutoScribe*. Toronto, ON: Mutuo Health Solutions. Available from: <https://mutuohealth.com>. Accessed 2021 Apr 23.
18. *EyeArt AI eye screening system*. Woodland Hills, CA: Eyenuk, Inc. Available from: <https://www.eyenuk.com/en/products/eyeart/>. Accessed 2021 Feb 20.
19. Shortliffe EH, Davis R, Axline SG, Buchanan BG, Green CC, Cohen SN. Computer-based consultations in clinical therapeutics: explanation and rule acquisition capabilities of the MYCIN system. *Comput Biomed Res* 1975;8(4):303-20.

Can Fam Physician 2021;67:889-93. DOI: 10.46747/cfp.6712889

La traduction en français de cet article se trouve à www.cfp.ca dans la table des matières du numéro de décembre 2021 à la page e317.