

CASE AND COMMENTARY

Should Watson Be Consulted for a Second Opinion?

David D. Luxton, PhD, MS

Abstract

This article discusses ethical responsibility and legal liability issues regarding use of IBM Watson™ for clinical decision making. In a case, a patient presents with symptoms of leukemia. Benefits and limitations of using Watson or other intelligent clinical decision-making tools are considered, along with precautions that should be taken before consulting artificially intelligent systems. Guidance for health care professionals and organizations using artificially intelligent tools to diagnose and to develop treatment recommendations are also offered.

Case

Ms L is a 63-year-old woman who visits her primary care physician, Dr R, with new-onset fatigue and gum bleeding. After a thorough history and physical examination, Dr R orders a complete blood count, the results of which show anemia, thrombocytopenia, and leukocytosis. Dr R urgently refers Ms L to Dr O in hematology-oncology. Following more testing, Dr O concludes that Ms L has acute myeloid leukemia, admits her to the hospital, and schedules her for induction chemotherapy.

After several weeks of therapy, Ms L seems to be benefitting little from treatment and her condition has worsened since admission. Stumped by what once seemed like a routine case, Dr O recommends to Ms L genetic testing of her cancerous cells, which could offer additional information that could potentially lead to a different diagnosis or treatment plan. Dr O also considers consulting colleagues at the University of Tokyo, who recently used the IBM Watson™ artificial intelligence (AI) system to correctly diagnose a rare form of leukemia in a very similar case.¹ Dr O recalls Watson being able to sift through millions of pages of clinical literature as well as being able to incorporate a patient's genetic background and clinical history to come up with a diagnosis and treatment plan, all in a fraction of the time that it would take a physician to do.

Dr O wonders: *Perhaps Watson would come to the same conclusion I did. Or maybe Watson would find something I missed and help save Ms L's life. Or maybe Watson would be totally unhelpful and waste my and Ms L's time.* Dr O also wonders whether to explain Watson to Ms L and invite her to consider these questions.

Commentary

Watson is an advanced question-answering computer system developed by IBM that can be used as a clinical decision support system (CDSS) to assist health care professionals in making decisions about diagnoses and treatment options.² The system uses a variety of artificial intelligence (AI) approaches including natural language processing, information retrieval, semantic analysis, automated reasoning, and machine learning.² IBM calls its software architecture DeepQA, where QA stands for “question and answering.”² The DeepQA system famously beat 2 Jeopardy game show contestants in a televised exhibition match in 2011.²

Watson is an example of [augmented intelligence](#), whereby normal human intelligence is supplemented through use of technology in order to help people become faster and more accurate at the tasks that they’re performing.² The system works essentially like this: massive amounts of unstructured and semistructured data such as that from the clinical literature, health records, and test results (eg, pathology reports) are fed into the Watson system database. A physician poses a query to the system describing symptoms of a specific patient and other related factors. Watson first parses the input to identify the most important pieces of information and then mines a patient’s data to find relevant facts about the patient’s clinical and hereditary history. The system then examines available data (that were previously inputted) to form and test hypotheses and finally provides a list of individualized, confidence-scored recommendations, such as a patient’s eligibility for specific treatments. The system uses numerous scoring methods and sophisticated algorithms to determine the degree of certainty that retrieved evidence supports the candidate answers. The system can then describe the supporting evidence in text form for its ranked responses.³ Because information is constantly being fed to Watson, the system can learn over time to optimize its recommendations.

IBM Watson Health™ presently offers commercialized applications of the Watson system for genomics, drug discovery, health care management, and oncology.⁴ IBM has partnered with several academic and private institutions to apply Watson to patient care research and treatment. For example, in 2013, IBM partnered with a company called WellPoint to train Watson in utilization management and partnered with Memorial Sloan Kettering Cancer Center to train Watson in extracting and interpreting data related to lung cancer.⁵ And, in 2015, IBM and Manipal Hospitals (a large hospital system in India) announced the launch of IBM Watson for Oncology, which sorts through information and provides insights to physicians and cancer patients to help them identify personalized, evidence-based cancer care options. The service is also made available directly to patients through Manipal Hospitals’ website as a physician-mediated expert second opinion.⁶

Given the potential that Watson and any other intelligent CDSS has for clinical care and research, it’s essential that physicians such as Dr L consider the ethical (and legal)

ramifications regarding their use. Some essential questions are these: (1) Should Watson ever replace the clinical judgment of a physician? (2) What are the liability concerns of professionals who use Watson? (3) What are the limitations of Watson and their ethical implications?

Watson's Role

According to IBM, Watson is intended to assist and enhance the decision making of health care professionals by giving them greater confidence in their diagnostic and treatment decisions for their patients.¹ Thus, the system is not intended to replace the judgment of health care professionals, nor should it be viewed as any kind of authoritative decision-making tool. In the United States, the Food and Drug Administration (FDA) regulates the safety and effectiveness of devices and drugs. Because Watson is considered a management tool under the control of physicians (like a peer-reviewed publication) and not a device, the system does not presently require [regulatory oversight](#).⁷ However, regulatory requirements could change as Watson and other emerging AI systems are used to make diagnoses or treatment decisions with little or no supervision from physicians.⁸

Liability

While Watson aims to assist the accuracy of clinical judgment and improve health outcomes, use of the system as an assistant also has potential to [increase liability](#) for health care professionals and organizations. As noted by Jacobson,⁹ technological innovations create opportunities for error in diagnosis and treatment, and those errors could result in more visible and potentially detrimental outcomes than what might have happened without the new technology. As a hypothetical example, Watson could recommend a particular medication regimen that a physician decides to pursue while ignoring other contraindicating patient data because of the physician's assumption that Watson (or any other CDSS like it) had evaluated that information. Such a scenario could result in a malpractice claim against the physician.

While we can hope that advances in technologies such as Watson can improve outcomes for patients, they also have the potential to prematurely contribute to a higher legal standard of care that could put health care professionals at greater risk for negligence.⁵ This is because expectations of the standard of care can shift while the impact of the technology on health outcomes is not yet fully known. For example, if Watson is shown to improve diagnostic accuracy and treatment recommendations for leukemia, then expectations that clinicians who consult Watson will get the diagnosis and treatment recommendation "right" could be raised to a higher level.

Unfortunately, a paucity of clinical trials evaluating every possible diagnosis and treatment approach can limit the reliability and usefulness of Watson. That is, recommendations provided by the system might not be supported by sufficient research

to instill confidence in health care professionals, who could be found liable if their diagnoses or treatment recommendations are shown to be incorrect or possibly prove harmful. Thus, health care professionals who use Watson, such as Dr O in the case example, should do so with an awareness of potential harm that overreliance on the system could cause in the individual case, but also with appreciation for how the system can also improve their decision making.

Understanding Watson's Limitations

There are precautions that should be taken into consideration before consulting Watson. First, it's important for physicians such as Dr O to understand the technical challenges of accessing quality data that the system needs to analyze in order to derive recommendations. Idiosyncrasies in patient health care record systems is one culprit, causing missing or incomplete data. If some of the data that is available to Watson is inaccurate, then it could result in diagnosis and treatment recommendations that are flawed or at least inconsistent. An advantage of using a system such as Watson, however, is that it might be able to identify inconsistencies (such as those caused by human input error) that a human might otherwise overlook. Indeed, a primary benefit of systems such as Watson is that they can discover patterns that not even human experts might be aware of, and they can do so in an automated way. This automation has the potential to reduce uncertainty and improve patient outcomes.

It is possible, however, that Watson might make a recommendation that is inconsistent with current clinical standards or that contradicts what a physician considers to be the appropriate decision. For example, a clinical standard might be always to prescribe a particular medication with a particular diagnosis, but an intelligent system such as Watson could recommend an alternative (eg, a nonstandard medication or no medication at all). In such a scenario, physicians must be able to support their decision to follow or not to follow the alternative and to understand the potential clinical and legal consequences. However, systems such as Watson can and should be designed to use a rule base that can limit recommendations to current clinical standards, thus ensuring that recommendations are consistent with treatment guidelines and currently accepted practices.

Inconsistency is associated with another consideration sometimes referred to as the "black-box" problem, whereby developers and users are unable to demonstrate how the system operates or derived its decisions for a particular course of action.⁸ For example, Watson's machine learning algorithms can derive conclusions that are not consistent with a physician's judgment regarding the diagnosis or prognosis, yet why it derived particular solutions might not be obvious. From an ethical point of view, it is therefore essential that both the developers and the users of AI systems understand (or at least [be able to explain](#)) the basis of how the algorithms work to reduce risk of harm to patients.

Requirements for an audit trail with a minimum level of detail to describe the decision process might be one way to address the black-box issue and help ensure public trust.⁴

Technologies such as IBM Watson also have potential to create unrealistic patient expectations regarding outcomes. For example, a patient such as Ms L might be overly optimistic about her new treatment because her physician consulted Watson, a presumably superintelligent machine that can do things better than humans. Because the foundation of the patient-physician relationship is trust, patients should be informed—by the health care professionals who treat them—about the tools and tests used to make decisions about their health. It is therefore of great importance that clinical computing tools be presented as decision assistants, rather than as decision makers, and that their limitations be communicated effectively.

Meeting Challenges of Clinical Decision Support Systems

Historically, the deployment and adoption of technological clinical decision-making tools has been met with some challenges.^{10,11} Some of the challenges are due to technological limitations (such as those associated with the data problems mentioned in this article), technology adoption issues (eg, usability and workflow integration), and physicians' perception of the technology when assured capabilities and timelines have not been achieved.¹⁰ For example, IBM Watson Health has been criticized for not living up to promises of the system's ability to transform cancer treatment and outcomes.^{12,13} Regardless, the need for intelligent automated systems such as Watson is evident given the exponential expansion and complexity of clinical data. For example, IBM has suggested that a person will generate 1 million gigabytes of health-related data in a lifetime—which is equivalent to more than 300 million books.¹⁴ Given the amount and complexity of patient data, physicians would be remiss not to consult intelligent systems such as Watson. In the future, it may very well be considered unethical (and create liability) not to consult Watson or intelligent systems like it for a second opinion, assuming that such systems prove effective in what they purport to do.

In conclusion, the emergence of innovative technologies raises familiar and sometimes new legal and ethical ramifications for the health care profession. Health care organizations must educate and train their staff on the capabilities and limitations of technological tools while ensuring that patients are adequately informed of how these tools are used to make decisions about their care. Many of the challenges regarding the adoption and deployment of systems such as Watson have solutions that can be addressed in time. When new technologies become available, it inevitably requires time for the study of their safety and clinical effectiveness. It's unlikely that intelligent systems such as Watson will one day displace health professionals, but instead they will advance patient care and clinical research beyond its present limits.

References

1. Monegain B. IBM Watson pinpoints rare form of leukemia after doctors misdiagnosed patient. *Healthcare IT News*. August 8, 2016. <https://www.healthcareitnews.com/news/ibm-watson-pinpoints-rare-form-leukemia-after-doctors-misdiagnosed-patient>. Accessed November 16, 2018.
2. Luxton DD. An introduction to artificial intelligence in behavioral and mental health care. In: Luxton DD, ed. *Artificial Intelligence in Behavioral and Mental Health Care*. San Diego, CA: Elsevier Science; 2015:1-26.
3. Ferrucci D, Brown E, Chu-Carroll J, et al. Building Watson: an overview of the DeepQA Project. *AI Magazine*. 2010;31(3):59-79.
4. IBM. IBM Watson health: empowering heroes, transforming health. <https://www.ibm.com/watson/health/>. Accessed November 20, 2018.
5. IBM Watson hard at work: new breakthroughs transform quality care for patients [press release]. New York, NY: Memorial Sloan Kettering Cancer Center; February 8, 2013.
6. Manipal Hospitals. Watson for Oncology. <https://watsononcology.manipalhospitals.com/>. Accessed November 2, 2018.
7. US Food and Drug Administration; Federal Communications Commission; Office of the National Coordinator for Health Information Technology. FDASIA health IT report: proposed strategy and recommendations for a risk-based framework. <https://www.fda.gov/downloads/AboutFDA/CentersOffices/OfficeofMedicalProductsandTobacco/CDRH/CDRHReports/UCM391521.pdf>. Published April 2014. Accessed December 4, 2018.
8. Luxton DD, Anderson SL, Anderson M. Ethical issues and artificial intelligence technologies in behavioral and mental health care. In: Luxton DD, ed. *Artificial Intelligence in Behavioral and Mental Health Care*. San Diego: Elsevier Science; 2015:255-276.
9. Jacobson PD. *Medical Liability and the Culture of Technology*. Pew Project on Medical Liability. http://www.pewtrusts.org/-/media/legacy/uploadedfiles/wwwpewtrustsorg/reports/medical_liability/medmal092204pdf.pdf. Published 2004. Accessed August 7, 2018.
10. McCullagh LJ, Sofianou A, Kannry J, Mann DM, McGinn TG. User centered clinical decision support tools: adoption across clinician training level. *Appl Clin Inform*. 2014;5(4):1015-1025.
11. Garg AX, Adhikari NK, McDonald H, et al. Effects of computerized clinical decision support systems on practitioner performance and patient outcomes: a systematic review. *JAMA*. 2005;293(10):1223-1238.
12. Ross C, Swetlitz I. IBM pitched Watson as a revolution in cancer care. It's nowhere close. *STAT*. September 5, 2017. <https://www.statnews.com/2017/09/05/watson-ibm-cancer/>. Accessed November 20, 2018.

13. Freedman DH. A reality check for IBM's AI ambitions. *MIT Technology Review*. June 27, 2017. <https://www.technologyreview.com/s/607965/a-reality-check-for-ibms-ai-ambitions/>. Accessed August 8, 2018
14. IBM and Partners to Transform Personal Health with Watson and Open Cloud [press release]. Armonk, NY: IBM; April 13, 2015. <https://www-03.ibm.com/press/uk/en/pressrelease/46609.wss>. Accessed on August 15, 2018.

David D. Luxton, PhD, MS is an affiliate associate professor in the Department of Psychiatry and Behavioral Sciences at the University of Washington School of Medicine in Seattle. His research focuses on the development and study of health care technology, artificial intelligence, and ethics.

Editor's Note

The case to which this commentary is a response was developed by the editorial staff.

Citation

AMA J Ethics. 2019;21(2):E131-137.

DOI

10.1001/amajethics.2019.131.

Conflict of Interest Disclosure

The author(s) had no conflicts of interest to disclose.

The people and events in this case are fictional. Resemblance to real events or to names of people, living or dead, is entirely coincidental. The viewpoints expressed in this article are those of the author(s) and do not necessarily reflect the views and policies of the AMA.