Opinion

Elementary, My Dear Doctor Watson

Ramy Arnaout*

Mr. Watson—come here—I want to see you.

—Alexander Graham Bell, March 1876

Many readers will recognize these as the first words ever transmitted by telephone. The Watson on the other end of the line was Thomas A. Watson, Bell’s assistant. Over a century later, we may soon be hearing these words again, this time at the hospital. But this time the Watson on the other end may not be a person but a machine: IBM’s Watson, a state-of-the-art computer system capable of answering questions posed to it in natural human language. No one knows exactly what the arrival of the new Watson will mean for medicine, or for laboratory medicine in particular, but it is clear that change is coming, and coming soon enough to warrant speculation on what to expect.

First, some background. For most people, Watson entered popular consciousness on Valentine’s Day 2011 with its now-legendary victory over human competitors on the TV game show Jeopardy! (1, 2). In a 3-day, nationally televised performance, Watson defeated the show’s 2 greatest champions, Ken Jennings and Brad Rutter, prompting Jennings to quip, “I for one welcome our new computer overlords” (3). Watson’s victory did not come out of the blue, however. It was the culmination of a 4-year, multimillion-dollar effort by a 25-member team as part of an IBM project called DeepQA (“question answering”) (3, 4). The project’s goal: to build a computer system that can sort through massive amounts of data—terabytes, or hundreds of millions of Web pages, in the case of the Jeopardy! performance (3)—to answer virtually any question a human might ask, however he or she might ask it, instantly. Because the questions on Jeopardy! can be about anything, because they can involve puns, slang, and other wordplay, and because contestants have only seconds to answer them, winning on the show presented a fitting challenge. But victory was only ever meant to be a proof of principle. The ultimate goal was—and is—to use Watson’s capability to help real-world decision-making in a number of important areas, including medicine.

The year since Watson’s big win has seen several announcements relevant to healthcare. Last February, IBM announced a collaboration with Nuance, makers of medical dictation software, with the goal of making it possible to interact with Watson by speaking, as opposed to just via the written word (5). In September, IBM announced an agreement with the 34 million-member multiregional health insurer WellPoint to build a commercial system around healthcare applications (6). The 2 companies have mentioned several general uses for Watson, including analyzing patient records and sifting through the latest published research. Just this month IBM announced formation of a 9-person Watson healthcare advisory board, with members drawn from hospital administration as well as practicing MDs to align the Watson project with “industry needs” (7). Board member Herbert Chase, MD, a professor at Columbia University Medical School and whose research is in clinical-decision support, has said that Watson can already answer medical questions better than his residents (4) (which, given the Jeopardy! performance, is no ding against residents).

The Watson team has clearly been busy, and IBM clearly means for Watson to earn more than just Jeopardy! winnings. But what exactly will the new technology do for medicine? The announcements so far have been relatively light on specifics, suggesting that IBM is either not completely sure or not ready to share. It is clear, however, that a “Doctor” Watson will aim to serve at least 2 groups within healthcare: providers (such as doctors, nurses, and trainees) and insurers. In addition, Watson may well be available for consultation to patients directly. As for what it will do, one can expect much the same as what it did on Jeopardy!: provide answers to questions on the basis of large amounts of data—in this case any data set that is too large or complex for a single person to process comprehensively on the spot.

A basic application to expect is a more goal-oriented way of querying individual patients’ medical records. The typical record contains no more than a few megabytes of text, largely consisting of visit, admit, consult, and progress notes; and radiology, surgical, and pathology reports. Still, “knowing your patient” is a challenge that takes time and training, and not only because current electronic interfaces leave much to be desired (8). Watson should be able to help. Does the
doctor want to know whether the patient’s first documented anemia preceded travel to the Far East? As long as the information is in the record—a big if—the computer should be able to offer an opinion much faster than the doctor can with the chart. Note the difference between an opinion and a definitive answer.Those who saw Watson on Jeopardy! may recall that for each question a box appeared on the TV screen that listed Watson’s top 3 guesses, along with a percentage indicating its confidence in each guess and a white stripe indicating its threshold for buzzing in. During its development, Watson’s evidence—its “thought process”—for any guess was available to the Watson team for analysis. This evidence, especially when Watson was wrong, was instrumental in making Watson better. Judging from Dr. Chase’s residents, Watson’s medical training is already well under way. One can expect a Dr. Watson to be able to present several possible answers, along with its confidence in and the evidence for each one, in whatever medical context it will appear.

More applications become apparent as one imagines Watson with access to thousands or millions of patient records at once. Only a year ago, Watson had already incorporated a substantial amount of medical background information when it downloaded all of Wikipedia (the team considered Jeopardy! too fast-paced to rely on an Internet connection) (4). For years, sites like WolframAlpha (http://www.wolframalpha.com) have drawn on US government data, such as the National Ambulatory Medical Care Survey and the National Hospital Ambulatory Medical Care Survey, to present medical statistics in a human-friendly manner (those who have not tried WolframAlpha can get a feel for what it can do by searching it for, say, “colon cancer”). I would not be surprised if Watson already incorporates these and similar training resources, such as UpToDate (http://www.uptodate.com), in its database. Thus, if Watson’s method of answering questions on Jeopardy! reminds some readers of a differential diagnosis, they may not be far off: Given a history and a physical, Watson can probably already offer a reasonable differential diagnosis and suggest steps in a workup. Indeed, Watson’s next victory may well be against the US medical-licensing exams, but although training material can teach Watson the current state of the art, giving the system large numbers of patient records should help improve on it. The Watson team and its collaborators have cited performing association studies to look for drug side effects and comparing treatment outcomes as possible applications (4). These possibilities seem plausible.

There are also clear applications for laboratory medicine. If, as futurist Ray Kurzweil has said (9), health and medicine are now information technologies, then laboratory medicine should be the poster child for what Watson can offer—for 3 reasons. First, clinical laboratories speak computers’ language. The lion’s share of the information that laboratories generate is numerical—a glucose concentration of 112 mg/dL (6.22 mmol/L), a factor VIII level of 12%, a zone of inhibition of 14 mm—and that information is usually sent and stored in standard, structured formats. Second, laboratory data influence a wide range of clinical decisions involving diagnosis, prognosis, treatment, and monitoring. Inevitably, therefore, many of the questions that a Dr. Watson would be asked would involve laboratory data. Third, laboratories simply produce a lot of data: A midsized academic hospital can produce >10 million results per year. One can imagine a proud IBM engineer objecting that answering questions on the basis of laboratory data would be too easy for Watson. After all, the pride of the system is in its ability to handle natural (e.g., written) language and ambiguous and unstructured data. A skeptic might grant that Watson was impressive on Jeopardy! but point out that nothing on the show suggested it can do mathematics, at least the kind required to understand a reference interval or a positive predictive value. That is possible but, one suspects, easily fixed. The fact is that, at minimum, a host of important clinical questions boil down to “Should I order this test?” or “How should I interpret this test?” These questions lie at the heart of laboratory medicine, bear on medicine in general, and seem reasonable for expecting a Dr. Watson to be able to weigh in on in the not-too-distant future.

The above discussion may sound threatening to some laboratorians. Looking back on the Jeopardy! match, Jennings wrote that “quiz show contestant” might be the first profession that Watson makes redundant but that it is was unlikely to be the last (10). In a future in which Watson or something like it knows the answers—and soon maybe even the questions—practitioners from across medicine may wonder what space that leaves for them. Until relatively recently, the monopoly that humans have had on activities related to thought, speech, touch, and empathy looked as though it would last forever. It is easy to interpret Watson’s arrival as hastening a day when it will not. But roles always change, and at any rate there are obstacles. Watson was built to handle text but not pathology slides, radiographs, electrocardiogram strips, electroencephalography recordings, surgical videos, electrophoresis gels, dermatologic photographs, and other non–text-based medical information (genomes fall somewhere between text and nontext). That ability still lies over the horizon. Even with text, however, one should not forget how many medical records are still on paper—probably the greatest obstacle to realizing nationwide learning systems from health records (11). Finally, it is easy to underestimate the work still neces-
sary to turn proof of principle into product. Even IBM’s programmers can code only so fast.

In the end, one can expect pragmatism to trump philosophy in determining Watson’s role in the hospital. Like other medical specialists, laboratorians are aware of how much more information is available about a patient, diagnosis, treatment, or outcome than a single person can process in a reasonable time. The Watson team’s view (4), which I largely share (12), is that any help will be welcome. Moreover, the ability to simply speak a question into a smartphone (e.g., Apple’s Siri)—or at an order-entry or results viewer terminal or to an office workstation—and hear or see a useful answer instantly will free the inquisitive person to ask many more questions, thereby allowing a more comprehensive understanding of the care we provide.

That is, of course, if Watson ends up being available. IBM hopes you, like Alexander Graham Bell, will want Watson. The company has done very well selling large-price-tag licenses to multibillion-dollar organizations. Will your hospital be able to afford Dr. Watson? Will insurers? Will a version be made available to practitioners and other individuals? Place your bets.

Author Contributions: All authors confirmed they have contributed to the intellectual content of this paper and have met the following 3 requirements: (a) significant contributions to the conception and design, acquisition of data, or analysis and interpretation of data; (b) drafting or revising the article for intellectual content; and (c) final approval of the published article.

Authors’ Disclosures or Potential Conflicts of Interest: No authors declared any potential conflicts of interest.

References