



Did IBM blow Watson's use in Medical AI?

By *Bill Moran and Rich Ptak*

Introduction

The Wall Street Journal (WSJ) published an article last year discussing Watson, IBM's artificial intelligence (AI) product. The article focuses on Watson's status in the marketplace, and specifically on IBM's initial target market, cancer care. The WSJ reporters spent significant time on the article in order to present a comprehensive view of Watson. Our interest was stirred by our on-going curiosity about how IBM's Watson efforts were working. IBM's Dr. John Kelly¹ who has overall responsibility for Watson, responded to the article in a blog.

This exchange motivated us to analyze both the WSJ article and Kelly's response. We believed that mining IBM's Watson experience to-date could reveal interesting lessons for AI vendors, their customers, and users as AI technology begins to rollout in multiple vendors' products. Finally, we offer some tentative conclusions about Watson and the future.

Background

Before proceeding, we offer some relevant background. First, we, like many US adults, have personal experiences with cancer. We are fully aware how terrible this disease is and how much suffering it causes. Secondly, a bit less important, but still significant is the amount spent worldwide dealing with cancer and its fallout. It is massive. This drives a bias in favor of virtually any activity to mitigate cancer's consequences.

Next is the issue of the nature of the research process. We are computer industry technical analysts with over 50 years' experience, not experts in AI technicalities or medical cancer. We understand the research effort as well as why and how projects fail. Research and development are uncertain processes. Typically, many failures precede success. Successfully applying AI to creatively treat cancer is extremely challenging. That it takes hard work to succeed, should surprise no one. Also note, the current emphasis is on treatment, not cure.

As a commercial enterprise, IBM deserves credit for pioneering this field. Naturally, they also expect success to yield a business opportunity. Only guesstimates exist of IBM's investment in Watson. However, it is safe to assume it runs into multiple billions. Considering the great contribution resulting from a reduction in cancer's effects, IBM's efforts deserve to be rewarded. A topic for a different article. There is another point to make. IBM initially developed Watson as a research project; successfully winning Jeopardy² and chess competitions; eventually beating reigning world champions³. Such wins were thought to suggest that applying Watson to cancer treatment held a reasonable chance of

¹ Dr. John Kelly was the Director of IBM Research during Watson's development. He is currently the IBM Senior VP responsible for both Research and Cognitive System, including both Watson and Power systems hardware.

² The details of IBM's efforts to prepare for Jeopardy are in the Wikipedia article. The article seems credible to us, but caution is necessary using Wikipedia references.

³ Unlike chess there is no world champion in Jeopardy, but the two best players were Watson's opponents.



success. Unfortunately, this overlooks how different learning how to treat cancer is from learning a game.

A Thomas Edison story explains some things to keep in mind about the research process.

While Edison was trying to develop a working electric light bulb, a reporter asked him about his progress. Edison responded he had over 100 failed attempts. The reporter then asked what he had learned. Edison responded he'd learned 100 ways that wouldn't work.

Edison continued his efforts until he had a solution.

Games exist in a very well-defined, bounded universe. Strict rules apply. A limited number of pieces make only well-defined moves. All potential moves are known. All games between Chess masters are recorded in a standard way that can be digitized. In fairness, the Jeopardy effort did involve natural language processing that is critical in the cancer effort as well. On the other hand, there is a great deal unknown about cancer. What is known changes frequently and rapidly. No standard exists to record and report information⁴; much resides in paper databases not accessible by Watson. Some digitized data may be the property of other organizations and unavailable to IBM.

Product discussion

IBM has three products addressing cancer treatment sharing the Watson title. Each focused on different aspects of the problem. The first, probably flagship product is Watson for Oncology⁵. It is designed to help oncologists diagnose and treat the various forms of cancer. The oncologist provides Watson with the patient's symptoms. Watson searches its database, then suggests possibilities for the doctor to consider⁶.

The second is Watson for Genomics⁷. The doctor collects a tumor sample and sends it to a lab for a genetic analysis. The doctor submits a report of the lab findings to Watson for Genomics for analysis. Watson creates a report of treatment options. One copy of the report stays with lab and another goes to the doctor.

The final product is Watson for Clinical Trials⁸. This compares a patient's condition to on-going clinical trials. The doctor receives a report of matches. This capitalizes on Watson's ability to rapidly search massive databases.

The WSJ article focuses primarily on Watson for Oncology. Watson for Clinical Trials is mentioned, but not discussed. Watson for Genomics product is mainly discussed in relation to the work at the US

⁴ There exists a Federal project to digitize medical data which we have not explored.

⁵ We recommend the 5-minute video available on the IBM website for a quick product overview. Note the interaction possible between the physician and Watson.

⁶ It is important to emphasize that the doctor makes the decision about a diagnosis. Watson makes suggestions.

⁷ See <https://www.ibm.com/us-en/marketplace/watson-for-genomics/details> for more information.

⁸ See <https://www.ibm.com/us-en/marketplace/watson-for-clinical-trial-matching> for more information.



Veterans Administration. We maintain their focus. Dr. John Kelly of IBM adds a discussion of Watson for Clinical Trials. We include this in his responses to the WSJ article.

Summary & discussion of the WSJ article

The WSJ article “IBM has a Watson dilemma” appeared in August of 2018. Written by Daniela Hernandez and Ted Greenwald⁹. It is an excellent piece of journalism; we piggyback on much of their work.

Acknowledging our debt to the authors, their approach was very different from ours. Quoting the article’s opening: “Big Blue promised its AI platform would be a big step forward in treating cancer. But after pouring billions into the project the diagnosis is gloomy.” They evaluate and draw conclusions about Watson from a strictly business context; entirely appropriate for the WSJ and its readers.

As IT industry analysts, our focus is on and about successfully applying technology. We have no breakdown of IBM’s investments in Watson Cancer products nor on IBM’s multi-year investments in developing the technology for Chess, and Jeopardy versions or applications in other industries. In our view the diagnosis is gloomy only if you believe Watson products should/could have achieved far greater results than they have.

Here’s an analogy to clarify the difference between approaches. Consider someone evaluating a house strictly as an investment versus as a home. Investors see the house either as a rental property or future resale. On the other hand, someone evaluating a family home has other concerns. They want a fair idea of its value, but their fundamental evaluation includes many other factors that the investor may not care about. Neither approach is wrong; each serves its different purpose.

Putting aside the business evaluation, we examine the technological progress and lessons learned from IBM’s experience to date. Let’s turn to specific points in the WSJ article.

The main points in the WSJ article appear in the text box (right). They are listed logically, not as they appear in the article. We will discuss only what we regard as the main points.

First, consider the challenges to any medical software recommending personal medical treatment¹⁰. Watson’s knowledge base requires on-going training to stay current on the history and evolution of cancer research. This includes

Challenges IBM faced developing Watson medical software and discussed in WSJ article:

- 1. IBM financials***
- 2. Watson financials***
- 3. IBM CEO on Watson***
- 4. IBM Watson strategy***
- 5. Watson Goals***
- 6. Customer usage of Watson***
- 7. IBM partners***
- 8. AI competitors***
- 9. Updating Watson***
- 10. Watson for Genomics***
- 11. VA experience***
- 12. Other, miscellaneous***

⁹ The article is available (to subscribers) on the Wall Street Journal’s web site. A general search using: “ibm Watson wsj August 2018” may work.

¹⁰ This has been adapted from and expanded from the WSJ article.



symptoms, treatment types, details of patient history and condition, outcomes – successes and failures, new treatments, etc.

Unfortunately, no standard format or form exists for storing medical data. There are no standard formats for digitized text, non-digitized text, hardcopy paper, images, audio, etc.

Next, major inconsistencies exist in the content, quality and integrity of data. Leading oncologists admit that there remains much to learn about cancer. Important details in patient demographics, and environmental data may be incomplete or totally lacking. Some cancers are caused or influenced by chemical exposure when handling or present in the environment.

Data is owned, and access controlled by many different organizations. Content, i.e. what is collected, varies over time and by location.

Cancer research developments and treatments are changing and evolving rapidly. The result is that it is very difficult to keep Watson's learning databases up-to-date.

IBM had multiple goals when they decided to embark on the Watson effort in the cancer field. The financial ones are neither public nor our interest. IBM knew cancer treatments would be a very high-profile, market-interesting topic. Success would yield favorable Watson publicity and build on the Jeopardy success.

The WSJ quotes Dr John Kelly, IBM senior VP, as saying IBM had bet the ranch¹¹ on Watson, and therefore a key strategy for creating new Cloud and AI-based businesses. In a 2017 MSNBC interview, IBM's CEO Ginni Rometty was quoted saying Watson will be able to diagnose and treat what causes 80% of the cancer in the world¹². Neither Ms. Rometty, nor Dr. John Kelly provided a time-frame for realizing that goal. Our analysis suggests the goal should be restated to reflect current realities.

Clearly, there are many challenges. These fundamental limits need to be kept in mind when discussing what has been achieved. Potentially, advances in AI will help to circumvent some of these limitations. Unfortunately, currently these do not exist. Finally, on a more practical level we see that point 6 (customer usage of Watson) directly impacts Point 1 (IBM financials).

[Implementing strategy](#)

IBM's initial strategy for Watson involved approaching leading US cancer research and treatment centers to persuade them to add Watson to their cancer treatment efforts. IBM planned to convince the researchers and doctors at these centers that they would benefit from Watson.

¹¹ The WSJ quoted Dr. Kelly with this comment. To our knowledge, he has not denied making this or an equivalent comment. We accept WSJ quotes at face value.

¹² A search of the MSNBC web site for the video failed to find it. There is a story describing the interview. The quote is from that story. The WSJ repeated the quote.



Not being present at these meetings, our opinions are based on analysis with the benefit of hindsight. Subsequent events lead us to believe the initial strategy failed because it set incorrect expectations. We believe that evidence/experience did not support the idea that Watson would directly benefit leading-edge doctor/researchers, at least not in the minds of those that count.

In its early trials Watson served to validate/confirm existing treatment programs. Not surprising as that data served as the basis for training Watson. Watson could not help in dealing with rare types of cancer where most help was needed. Also not be surprising, Watson made some errors early on. Watson was not set-up to do the creative analysis needed in rare or abnormal cases, nor was it set-up for cancer diagnosis.

Researchers quickly dismissed Watson as of no value to them. It is unclear how much IBM consulted with researchers when planning Watson development and training programs. In retrospect, they should have been intimately involved.

We believe Watson's real early benefit will be to the thousands of practicing oncologists in the US (and worldwide), not leading research institution staffs. Watson could be an ideal vehicle to disseminate knowledge and maintain practitioner currency on the state of the art coming from researchers.

In such case, lead doctors at MSK, Anderson and the others¹³ would see that real benefit would be delivered to millions of cancer patients who will never set foot in their institution. This is a challenging task, but it is a challenge that IBM may be uniquely qualified to take on¹⁴. The joint MIT and IBM research project on AI supplies evidence of IBM's ability to partner with other institutions.

Also, there exists a natural human bias to resent & resist AI (any change) when it is suspected that the long-range plan to replace human efforts. Much research supports this. IBM has been careful to emphasize that Watson was to assist NOT replace doctors. Nevertheless, we suspect the underlying resentment bias persists.

Another demonstrated benefit is Watson's ability to help oncologists stay up-to-date on the latest literature. Again, an even greater benefit to those oncologists not working at Sloan Kettering or MD Anderson, etc. A benefit to cutting-edge research institutions exists but isn't viewed as major.

The Watson experience at India's Manipal Hospitals is instructive. It is mentioned by both the WSJ article and Dr. John Kelly. Manipal decided to only use Watson in the minority of breast cancer cases where their review board had questions, about 30% of the cases. In most cases, Watson confirmed the existing diagnosis. Watson was accurate in over 93% of the cases. Watson affected the diagnosis in

¹³ Other leading institutions appear in the article available here: <https://www.wsj.com/articles/ibm-bet-billions-that-watson-could-improve-cancer-treatment-it-hasnt-worked-1533961147>.

¹⁴ See our blog on the joint IBM MIT project for AI research for evidence of this.



around 9% of the patients¹⁵. The results are impressive testimony to Watson for Oncology. Ironically, results that pleased the Manipal researchers was seen by others as a negative.

US VA and Watson for Genomics

The WSJ devotes some space to this topic. The largest user of the Genomics product is the US VA. The WSJ quotes Doctor Michael Kelly, an oncologist for the VA. (This is NOT IBM's Dr. John Kelly.)

The WSJ says that IBM and the VA signed an agreement in 2016 that the product would be used on 10,000 patients over the next couple of years. Dr. M. Kelly provides details on the approximately 3,000 cases where it has been used. He found it useful in finding relevant medical literature the doctors weren't always aware of¹⁶. The product is fast, but sometimes errs even in cases requiring standard treatments. It does save time. In another case, Watson found a genetic mutation that Dr. M. Kelly said he would probably not have found on his own. He informed the attending physician, but the resulting treatment was ineffective. In another case, where they followed the Watson recommended treatment, the results are not yet known.

The VA decided to continue the study for another year. Dr. M. Kelly says that they are not paying IBM anything. He says that AI has lots of potential, but that the potential is not yet realized.

We believe that the VA case studies cited are inconclusive. Dr. M. Kelly is correct when he says that AI and Watson have enormous currently unrealized potential. The fact that IBM and the VA decided to continue the study indicates that both see value in the relationship.

Should anyone be surprised by the results of this pilot project so far? We say "No". It is excellent that the product has been used in the number of cases that it has. Initial targets were probably too aggressive. But note that the 30% attainment coincides with the Manipal results. The whole point of a pilot is to uncover errors and fix them. Exactly what IBM seems be doing.

Other comments in the WSJ article indicate that some oncologists lack trust in the results from Watson for Genomics. The only way to fix this problem is to drive up usage, identify and fix the problems encountered.

Earlier work on the general question of customer adoption of new technology reveals wide differences among customers (early versus late adopters). Some are eager to try new things. Others prefer to wait until a new technology is more widespread and proven. There is no reason to expect the medical field to be different. Efforts to persuade a late adopter to an early one is highly problematic with any reasonable incentive. They will always find reasons, which may be perfectly valid, to wait.

¹⁵ See <https://academic.oup.com/annonc/article/29/2/418/4781689> (Note: IBM provided financial support for the article). We should mention that when Watson & the Manipal board agreed it is possible that they were both in error. Only a significant follow up study of results could detect this situation.

¹⁶ This tends to confirm our earlier point that Watson's results would be most valuable to physicians not located in the major cancer research centers.



Kelly's response¹⁷ to the WSJ article

Dr. John Kelly, IBM senior VP, responded to the article in a blog. He states that media reports, including the WSJ article were inaccurate. He stated the WSJ ignored clear testimony from doctors and others to imply that IBM had not made “enough” progress with Watson.

Now, the WSJ never actually said that IBM had not made enough progress. We do agree with Dr. J. Kelly that they certainly implied this particularly in the discussion of Watson vis-à-vis IBM financials. As the article's general drift is somewhat negative to IBM and its Watson Healthcare products, Kelly's reaction is understandable.

Near the beginning of the WSJ article it is stated that “No published research shows Watson improving patient outcomes.” In one sense one could see the point of this statement – “patient outcomes” means what happens at the very end of the treatment. Was the patient cured of cancer or not?

However, Watson is not responsible for the entire process of the patient's treatment. The doctor makes the diagnosis not Watson. We see this distinction most clearly in the Watson clinical trials product. We conclude that if Watson identifies a clinical trial with a potential benefit for a patient that the doctor was unaware of, Watson has done its job.¹⁸ When the doctor evaluates the situation and the patient joins the clinical trial the ultimate outcome may be in doubt for a long time. This does not reduce the value of the recommendation. Clinical trials fail but that is not a failure of Watson.

When an influential part of the media portrays a negative attitude about a product, discussions with potential customers about the product become more difficult. We believe that the WSJ article could quite fairly have included more commentary, e.g. around the distinctions we've made. They could have also commented on some unrealistic financial expectations.

Dr. Kelly provides some background as a level set. He discusses five (5) specific Watson benefits. He asserts that the WSJ indicated no clear Watson benefits. The WSJ did discuss Manipal and the VA results (#6 in the list below). Dr. Kelly's Watson successes include:

1. Mayo Clinic physicians presented a poster presentation at the ASCO Annual Meeting, reporting that Watson for Clinical Trial Matching boosted enrollment in breast cancer trials by 80% following implementation (to 6.3 patients/month, up from 3.5 patients/month in the immediate 18 months prior). (http://abstracts.asco.org/214/AbstView_214_218403.html)
2. Dr. Thaddeus Beck and the group at Highland Oncology Group reported that Watson Clinical Trial Matching reduced the time for clinical trials matching by 78%. (http://ascopubs.org/doi/abs/10.1200/JCO.2017.35.15_suppl.6501)
3. Working with training partner Memorial Sloan Kettering, Dr. Mark Kris and the oncologists at MSK helped to train Watson for Oncology on thirteen (13) cancers, representing up to 80% of the global cancer incidence and prevalence.

¹⁷ Kelly's blog is available at <https://www.ibm.com/blogs/watson-health/author/johnkelly/>

¹⁸ One could also argue that selection for a clinical trial might also offer hope to a stage 4 cancer patient. We think that even if this Hope is ultimately disappointed it is not bad.



4. Dr. Somashekhar and Manipal Hospital published a 93% concordance rate in breast cancer for their multidisciplinary tumor board in the Annals of Oncology in 2018. They subsequently stated they use Watson for Oncology with all complex cases in their multidisciplinary tumor board. Significantly, in 9-11% of patient cases this results in changes to treatment recommendations (<https://academic.oup.com/annonc/article/29/2/418/4781689>).
5. Dr. William Kim and the University of North Carolina Lineberger Cancer Center published a study where Watson for Genomics found new, actionable mutations in 32% of patients. (<http://theoncologist.alphamedpress.org/content/early/2017/11/20/theoncologist.2017-0170.abstract>).
6. Dr. Michael Kelley and the Department of Veterans Affairs extended their contract with IBM for Watson for Genomics, meaning nearly 3,000 veterans facing Stage 4 cancer will be supported.

Interested readers can follow the links for more information. Kelly's first two examples refer to the Watson for Clinical Trials product which was not addressed in detail in the WSJ article.

The WSJ did discuss the MSK example, the Manipal case and the VA example. Manipal includes a serious study of patient usage and benefits. Perhaps the WSJ should have given more weight to and details on the Manipal results. The Linberger study documents significant patient benefits from the use of Watson Genomics.

Dr. Kelly makes a strong case for the value of the overall Watson Healthcare project. He makes a very significant point "*Together they (Watson healthcare products) are now in use at 230 hospitals and health organizations globally and have nearly doubled the number of patients they've reached in the first six months of the year to 84,000*".

Kelly concludes with a discussion of the "Grand Challenges" IBM has faced and overcome in the past, e.g. building the world's fastest supercomputer. He states that IBM intends to conquer cancer's challenges as well. Clearly, a lot of that work remains.

Some comments on status & future strategy

IBM has made some significant progress toward their publicly stated goals for Watson Healthcare products in the cancer sector. They are delivering documented and documentable patient and doctor benefits. On the other hand, IBM still has much to do to meet the ambitious goals including the 80% goal articulated by their CEO. We will not estimate how long this might take.

Finances are not our area of expertise or interest. However, there is a relevant financial aspect, pricing. It is in IBM's interest to increase usage of these products in order to discover/fix problems and elicit suggestions for product improvements. We suggest IBM reexamine their pricing to encourage use of Watson. Customers willingly pay for a product when it delivers a benefit; especially if cumulative benefits exceed the cost. Price should not discourage potential users who know that 90% of the time Watson only confirms an existing diagnosis. The remaining 10%



Lessons for the AI Industry

The VCs, at least the more sophisticated ones, have already drawn conclusions about AI. Profits from AI products in the healthcare field will be slow in coming¹⁹. As a rule, caution is advisable when using unproven technology. Today, all the factors involved in the success or failure of a given AI installation are not well understood. Success cannot be predicted based on small samples. Indeed, many IT projects fail using current technology. We don't expect AI projects will be any different. All these considerations signal caution when evaluating any AI project.

Secondly, AI's history demonstrates that exaggerated expectations lead to failure and disillusionment. One example, years ago machine language translation was the fad. The US was competing with the (then) Soviet Union. It was thought translation of Russian into English (using AI-driven devices) would help build knowledge of Russian activities. The results were disappointing. Decades later, such translation programs are widely available on mobile devices. It took a long time. It is reasonable to expect that today's more ambitious hopes of AI will also take some time to realize.

A final lesson is the critical importance to work with high quality data in any AI application. IBM earned this lesson. Today, they strongly advise their clients that collecting and validating data is an absolute requirement before beginning any AI efforts. In a recent briefing, IBM showcased a client who spent more than one year gathering and validating their data before beginning their AI project.

Conclusions

It is notable that WSJ's research²⁰ as reported does not reveal any even slightly questionable ethical behavior by IBM. Now, "*absence of evidence is not evidence of absence*". However, given the WSJ's past success in uncovering corporate misdoings, the point is worth mentioning.

Our bottom-line is echoed by Dr. Mark Kris of MSK, who works with IBM on updating Watson for Oncology. He also coordinates Watson's cancer training for MSK. He described it as still a work in progress. He said that keeping pace with rapidly evolving cancer treatments has proven to be tougher than he imagined. Dr. Kris still believes in the technology. His comment was "Is it as nuanced as we'd like; Is it 100% up to date? No. But for what it is it's pretty darned good"²¹.

IBM has much more work to do. Dr. John Kelly acknowledged as much as he noted that IBM knew the cancer challenge was extreme when they committed to resolve it. We support and admire that decision. If the Watson effort even partially succeeds, IBM will have significantly contributed to cancer treatments worldwide.

¹⁹ Kocher MD (and partner in VC firm Venrock) says: Oncology is not a great space for AI products until there is better patient data including genetic, environmental, lifestyle & health information. In the near-term AI benefits in health care come from admin tasks like billing.

²⁰ The authors interviewed numerous individuals including ex-IBM employees who are quoted anonymously.

²¹ We can't resist saying that we aren't actually sure that he really said "darned".



Final note

Please note that IBM did not commission, request or compensate us for this article. We have done business with IBM in the past. To identify and correct any errors of fact and assure technical accuracy this article was provided to IBM, both authors, and the WSJ before publication. It does not mean they agree with our conclusions or statements. One author provided comments. WSJ declined comment. IBM's Dr. Kelly did not respond to our request.

Addendum

We added this appendix to discuss issues raised in a number of recent articles critical of IBM Watson & cancer, including one published by the IEEE. The intent is to determine if information in them impacts our conclusions. We made a Google search using "IBM Watson Cancer". There are too many articles for us to cover each individually. However, we think that this list is a fair sample of general topics raised.

- The financial condition of IBM as it relates to Watson.
- The bias introduced into Watson by Memorial Sloan Kettering updating the Watson database.
- The hype around Watson and IBM's responsibility for the hype.
- The accuracy of Watson's results.
- The need for Watson customization for local conditions.
- The claim that IBM underestimated the difficulties of building and updating a database of medical information.
- The MD Anderson experience.
- The IEEE article (reference #4)
- Miscellaneous items

Point 1, financials. We already discussed this topic earlier in our article. While important, we have no special expertise nor access to anything other than publicly available data. We believe IBM's investment is consistent with what would be expected in a cutting-edge venture.

Point 2, single data source. We agree with the critics. Watson needs more input from other sources. Data/information collected via a single U.S. institution provides neither the breadth or the depth of information/data needed for useful knowledge. The comments from oncologists from both Holland and Germany support this. IBM has decades-long, extensive experience at internationalizing products. However, this case includes many, more complex issues than exist in just adapting for language, alpha/numeric characters, etc. IBM has announced plans for adapting Watson for the Chinese market. We believe that this limitation will be addressed with time.

Point 3, hype. Watson for Oncology has been the subject of a great deal of hype in many media reports. We did not find any direct statements from senior IBM executives containing irresponsible hype. Many quotes attributed to unnamed IBM people do hype Watson. It may be that individual salespeople attempt to leverage Watson's Jeopardy success with claims of success dealing with cancer. And, it is fair to say that the tech industry has frequently been guilty of over-hyping products. In Watson's case, it was more the Media (at large) and some partners making exaggerated claims about



“Watson curing cancer”. To be fair, IBM could have been more precise at clarifying what to expect from early versions of Watson Oncology. Also, IBM management are responsible for reining-in employees using hype in sales efforts. We agree IBM did make statements about “a new era in computing” resulting from the combination of Watson and AI. While this may be a bit of an exaggeration, it is a long way from the degree of hype in some accusations. We believe it is true that widespread use of AI can change computing in fundamental ways.

Point 4, accuracy. We believe that Watson’s impact is inconclusive at this time. An accurate assessment of the real impact and diagnostic deliverables from Watson for Oncology will be determined only after in-depth field trials. The process should have patients diagnosed by both a panel of doctors and by Watson. The patients should be tracked until a proper assessment can be made of the accuracy and appropriateness of the Watson diagnosis versus the experts’ diagnosis. We do think that Watson should be compared against two types of oncologists, experts and ordinary practitioners. The results might be quite different.

In the reported comparisons Watson is in the unfortunate position of being criticized no matter what happens. If the Watson and doctors agree on a diagnosis, critics say that Watson is worthless because it is redundant. If Watson disagrees with the doctors, the critics assume Watson is wrong. Note that the IEEE article makes clear that current physician diagnostic techniques are not perfect. There is no guarantee that current best practices are infallible.

For the results to be useful and accurate, timing is critical. If done too early with unreasonable expectations, (as it appears today) evidence is inconclusive and weak. Early evaluations of the product’s accuracy seem to us to indicate that IBM’s release of Watson for Oncology as a product was premature. Since then, the technology has matured significantly which Manipal’s study appears to support.

It remains our opinion that Watson needs more work before a realistic study can be done. It may take several years even assuming IBM continues significant product development.

Point 5, customization. Unquestionably, Watson needs to be adapted for local conditions including government regulations, environment, economics, cultural mores and limitations, etc. In Korea, for example, certain treatments that Watson suggests are not covered by the local insurance. Korean doctors have developed a work-around which is standard for the country. More needs to be done.

Point 6, data consumability. We agree, IBM underestimated the difficulty of including the different types of medical information in Watson. The MIT article (# 2) makes this point. It also points out that this is a problem for all AI products. The article then states that IBM may have a better chance to correct this problem than some competitors because of IBM’s ability to partner. We agree.

Point 7, MD Anderson. The MD Anderson Cancer Center story is sad and left Watson with a black eye. We believe some of the criticism is unjustified. Here’s why. From a technical perspective, it is reasonable to assume that Watson has two logical parts. The Watson Base, which we call Part 1, includes database management, natural language processing and other such functions. Part 2 consists



of the Watson's industry specific functions accessible thru an API. These functions change depending on the targeted industry. For Watson for Oncology, these are medical- and cancer-related functions.

MD Anderson wanted to develop a product that would compete with Watson for Oncology. Not being privy to the details of the relationships among IBM, PricewaterhouseCoopers and MD Anderson it is hard to judge the results. An audit stated that IBM realized \$39.2M and PwC \$21.2M in the deal. The net result for MD Anderson is a failure. We do not think that this failure reflects the true value of Watson for Oncology.

The details of who was responsible for what are not public knowledge. We don't know if IBM acted merely as a consultant, or if they had significant responsibility for parts of the project. A logical approach would have MD Anderson develop Part 2-type functions to compete with IBM. Watson-Part 1 would provide basic functions. Without full information, we neither defend nor condemn IBM for its business decision to form a partnership with Anderson. Given its failure, both IBM and MD Anderson might have been better off not partnering in the first place.

It is worth noting that MD Anderson underestimated the difficulty in extracting usable data from the enormous variations in content, format and (media) sources used for storing/recording medical data. As a medical institution, they have perhaps less excuse than IBM has for this error.

However, there is one point about this whole situation that has puzzled us. Why would IBM help someone build a competitive product? We could not understand what the answer was to this question. Of course, financial considerations would play a part but that did not seem to be a sufficient answer. But this question lead us to the following speculation. It is possible that IBM agreed to help Anderson because Anderson was going to use Watson functions to construct the Anderson database. IBM might have thought that the Anderson database would be very compatible with the database used by Watson for Oncology. This compatibility would mean that in the future the two databases could be merged. This would mean a major addition to Watson in the future and incidentally a good first step in answer to the problem that we discussed in point 2 above.

Finally, a couple of reflections on the Anderson story.

- Anderson undertook an IT project to build a product. As noted in our article, many IT projects fail, so no one should be too surprised at their failure. We cautioned venture firms to expect more AI projects including in the medical field to fail. History teaches that such failures are necessary, and we can learn from them as the Edison example shows.
- We discussed the Anderson project and given our opinion. Given the lack of public information, it is conceivable we are wrong.

Point 8, IEEE Spectrum. Eliza Strickland's IEEE Spectrum article is a major contribution. We rate it highly and encourage readers to review it. The IBM development process and decisions are discussed in detail. As we made clear in earlier comments, any complex development process will be messy with many failures. The article provides a fair discussion of the issues. It highlights the problems that any AI system will face in the medical field. We found no major surprises in the article. We stand by our conclusion that IBM should continue to work in this area.



Point 9, miscellaneous. We found multiple snarky comments about Watson from IBM competitors and others with an axe to grind. Lacking details and based only on anecdotal evidence, we have ignored them.

Finally, Watson for Oncology can have a future in the cancer field. It will take time and a great deal of work and investment from IBM. Frankly, we were unaware of the amount of bad press surrounding Watson. Much of this criticism relates not to any technical failures of Watson, but reflects a natural disappointment when Watson fails to rapidly deliver on its promoted potential. IBM should keep the bad press in mind.

We think that the potential exists and can be realized. How long it might take, no one really knows. We suggest that IBM should clarify a possible timeline. Without any hype IBM should carefully highlight any Watson successes in the medical field.

Of course, it is possible for a competitor to leapfrog IBM. Only time will answer that question. As stated earlier, our bias favors any progress in the battle against this terrible disease that impacts and takes so many lives.

References

To pursue these issues further, these articles are a good starting point:

- 1) ***MD Anderson Cancer Center's IBM Watson project fails, and so did the journalism related to it*** - <https://www.healthnewsreview.org/2017/02/md-anderson-cancer-centers-ibm-watson-project-fails-journalism-related/>
- 2) ***A Reality Check for IBM's AI Ambitions*** *MIT Technology Review* - <https://www.technologyreview.com/s/607965/a-reality-check-for-ibms-ai-ambitions/>
- 3) ***MONEYBOX What Went Wrong With IBM's Watson, The A.I. project was supposed to change the state of cancer treatment. Here's what happened instead*** - <https://slate.com/.../ibms-watson-how-the-ai-project-to-improve-cancer-treatment-went-wrong.html>
- 4) ***IEEE Spectrum Watson report How IBM Watson Overpromised and Underdelivered on AI Health Care*** – <https://spectrum.ieee.org/biomedical/diagnostics/how-ibm-watson-overpromised-and-underdelivered-on-ai-health-care>
- 5) ***IBM Watson recommended unsafe and Incorrect cancer treatments, STAT report finds*** - <https://www.beckershospitalreview.com/artificial-intelligence/ibm-s-watson-recommended-unsafe-and-incorrect-cancer-treatments-stat-report-finds.html>



Publication Date: November 8, 2019

This document is subject to copyright. No part of this publication may be reproduced by any method whatsoever without the prior written consent of Ptak Associates LLC.

To obtain reprint rights contact associates@ptakassociates.com

All trademarks are the property of their respective owners.

While every care has been taken during the preparation of this document to ensure accurate information, the publishers cannot accept responsibility for any errors or omissions. Hyperlinks included in this paper were available at publication time.

About Ptak Associates LLC

We monitor a breadth of areas to provide a complete picture on technology trends across the industry. Whether it's mainframes, Open Systems, Blockchain, Cognitive Computing, DevOps, Quantum Computing, or other emerging trend, are covered with a uniquely deep and broad perspective.

Our clients include industry leaders and dynamic newcomers. We help IT organizations understand and prioritize their needs within the context of present and near-future IT trends, enabling them to apply IT technology to enterprise challenges. We help technology vendors refine strategies and provide them with both market insight and deliverables that communicate the enterprise values of their services. Clients benefit from our understanding of how competitors play in the market space as well as from our actionable recommendations.

www.ptakassociates.com