The Value of Artificial Intelligence in Laboratory Medicine

Current Opinions and Barriers to Implementation

Ketan Paranjape, MS, MBA,1,6 Michiel Schinkel, MD,2 Richard D. Hammer, MD,3 Bo Schouten, MSc,1,4 R. S. Nannan Panday, MD, PhD,2 Paul W. G. Elbers, MD, PhD,5 Mark H. H. Kramer, MD,6 and Prabath Nanayakkara, MD, PhD2

From the 1Amsterdam UMC, 2Section Acute Medicine, Department of Internal Medicine, Amsterdam UMC, 3Department of Intensive Care Medicine, Amsterdam Medical Data Science, Amsterdam Cardiovascular Science, Amsterdam Infection and Immunity Institute, Amsterdam UMC, and 4Board of Directors, Amsterdam UMC, Vrije Universiteit, Amsterdam, The Netherlands; 5Department of Pathology and Anatomical Sciences, University of Missouri School of Medicine, Columbia; and 6Department of Public and Occupational Health, Amsterdam Public Health Research Institute, Amsterdam, The Netherlands.

Key Words: Artificial intelligence; Laboratory medicine; Diagnostics; Medical education

ABSTRACT

Objectives: As laboratory medicine continues to undergo digitalization and automation, clinical laboratorians will likely be confronted with the challenges associated with artificial intelligence (AI). Understanding what AI is good for, how to evaluate it, what are its limitations, and how it can be implemented are not well understood. With a survey, we aimed to evaluate the thoughts of stakeholders in laboratory medicine on the value of AI in the diagnostics space and identify anticipated challenges and solutions to introducing AI.

Methods: We conducted a web-based survey on the use of AI with participants from Roche’s Strategic Advisory Network that included key stakeholders in laboratory medicine.

Results: In total, 128 of 302 stakeholders responded to the survey. Most of the participants were medical practitioners (26%) or laboratory managers (22%). AI is currently used in the organizations of 15.6%, while 66.4% felt they might use AI in the future. Most had an unsure attitude on what they would need to adopt AI in the diagnostics space. High investment costs, lack of proven clinical benefits, number of decision makers, and privacy concerns were identified as barriers to adoption. Education in the value of AI, streamlined implementation and integration into existing workflows, and research to prove clinical utility were identified as solutions needed to mainstream AI in laboratory medicine.

Conclusions: This survey demonstrates that specific knowledge of AI in the medical community is poor and that AI education is much needed. One strategy could be to implement new AI tools alongside existing tools.

Advances in our understanding of biology, disease, and molecular medicine have created a central role for laboratory medicine in the diagnostic workup of many, if not most, diseases. It is estimated that 70% of decisions regarding a patient’s diagnosis, treatment, and discharge are in part based on results of laboratory tests.1 Unfortunately, the main cause of medical errors in the United States is inaccurate diagnosis.2-5 The ever-increasing workload, high health care costs, and need for improved precision call for continuous optimization of the laboratory processes.6 With both health care and laboratory medicine7 transitioning into an era of big data and artificial intelligence (AI), the ability to provide...
accurate, readily available, and contextualized data is crucial. AI in health care is the use of complex algorithms and software to emulate human cognition in the analysis of complicated medical data generated from diagnostics, medical records, claims, clinical trials, and so on. AI algorithms can only properly function with reliable and accurate laboratory data. Automation and AI can fundamentally change the way medicine is practiced, as demonstrated by the recent applications in ophthalmology and dermatology. Some possible applications of AI specific to laboratory medicine are presented in Table 1. With the increasing role of laboratory medicine in many diseases, AI has the potential to improve diagnostics through more accurate detection of pathology, better laboratory workflows, improved decision support, and reduced costs, leading to higher efficiencies.

As laboratory medicine continues to undergo digitalization and automation, clinical laboratorians will likely be confronted with the challenges associated with evaluating, implementing, and validating AI algorithms, both inside and outside their laboratories. Understanding what AI is good for and where it can be applied, along with the state-of-the-art and limitations, will be useful to practicing laboratory professionals and clinicians. On the other hand, the introduction of new technologies requires willingness to change the current structure and mindset toward these technologies, which are not always well understood. Historically, there has been resistance to the adoption of new technologies in the medical community.

With a web-based survey among stakeholders in laboratory medicine in the United States, we aimed to evaluate their current perspectives on the value of AI in the diagnostics space and identify anticipated challenges with the introduction of AI in this field, as well as resistance to introduction of this new technology in today’s practice.

Today, AI is occasionally used in laboratory medicine for enabling the effective use of resources, avoiding unnecessary tests, improving patient safety, and alerting for abnormal results. AI is also being used in limited clinical usage for molecular/genomic testing by accurately identifying variants and matching it to possible treatments.

Materials and Methods

Survey Development

A web-based survey on the use of AI in laboratory medicine was designed in several independent steps. First, 98 stakeholders participated in a 2-week online discussion board on AI in diagnostics in August 2019. These participants were part of Roche’s Strategic Advisory Network (SAN), a group consisting of laboratory medicine decision makers, practicing physicians and surgeons, point-of-care coordinators, anatomic pathologists, laboratory management, information technology (IT) management, and senior leadership. Roche does not know the identity of the community members to protect their privacy. The online discussion board was moderated by C-space and developed questions to gain insights on potentially important topics to discuss in the survey. Open-ended as well as multiple-choice questions were formulated based on content of the discussion board.

Next, two 1-hour online group chats were organized on October 2 and 3, 2019, to discuss these questions and fine-tune their phrasing and to refine the answer possibilities to the multiple-choice questions. In these group chats, a total of 11 practitioners in laboratory medicine were asked to answer the initial survey questions one at the time, after which they could comment on each other’s answers and discuss their opinions on and interpretations of the questions.
This thoroughly discussed survey was fielded to a group of 302 laboratory medicine practitioners who are part of SAN via email. These individuals were both available for completing surveys and in a position to decide on embracing or refraining from using technologies such as AI in their respective organizations. The survey was available from October 21 until November 1, 2019. The data were collected using a software platform called Confirmit, and all participants gave informed consent for their input to be used for research purposes.

Finally, as there are multiple different definitions of AI, for the sake of the survey we defined AI as follows: “Artificial intelligence (AI) in health care is the use of complex algorithms and software to emulate human cognition in the analysis of complicated medical data generated from diagnostics, medical records, claims, clinical trials, and so on. AI is truly the ability for computer algorithms to approximate conclusions without direct human input.”

Questions

In the survey, we posed 21 questions that ranged from collecting demographic information to answering questions about if the respondents used AI in their organizations, what kind of improvements they would like to see in the current use of AI, how valuable they think AI will be in their practice, and what challenges they feel exist. A full list of all the questions can be found in the Supplement (all supplemental material can be found at American Journal of Clinical Pathology online).

Data Analysis

Quantitative Data

Categorical data were analyzed using a Pearson $\chi^2$ test, considering a $P$ value less than .05 to be statistically significant. The perceived value of AI was compared for different age groups and different experience levels with AI. The analyses were performed in Microsoft Excel. Data from the multiple-choice questions were presented as percentages per category.

Qualitative Data

An inductive approach of direct content analysis was used to analyze the open-ended questions. First, two researchers, B.S. (psychologist) and M.S. (internal medicine doctor), independently screened the answers and drafted a rough framework of themes. After consensus on the overarching themes, answers were independently coded with this framework by both researchers. Then conflicts were resolved by consensus to account for different interpretations of the answers. Coding was performed and bar charts were created in Excel.

Results

Demographics

The survey was fielded to 302 stakeholders in laboratory medicine, of whom 128 (42%) responded. The modal age group was aged between 41 and 50 years (32.0%), while 23 (18.0%) of the respondents were younger than 41 years. The top three participants were physicians, laboratory managers, and pathologists. See Table 1 for further details on demographic information.

Qualitative Analysis

Based on the data, six main themes were derived (attitude, quality of care, organizational value, data analysis, prerequisites, and education). The “attitude” theme was further categorized into three subthemes (positive, unsure, and negative). To prevent losing valuable information, multiple themes could be assigned to an answer. The specific content of these themes is presented in Table 2.

<table>
<thead>
<tr>
<th>Theme</th>
<th>Examples of Content</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attitude—positive</td>
<td>Respondent showed a positive attitude toward AI rather than giving a really specific answer to the question.</td>
</tr>
<tr>
<td>Attitude—negative</td>
<td>Respondent showed a negative attitude toward AI rather than giving a really specific answer to the question.</td>
</tr>
<tr>
<td>Attitude—unsure</td>
<td>Respondent generally was not sure about the influence of AI in a certain area.</td>
</tr>
<tr>
<td>Quality of care</td>
<td>Accessibility of care, accuracy of diagnoses, and early recognition of certain disease states</td>
</tr>
<tr>
<td>Organizational value</td>
<td>Providing quick results, reducing redundancy, and resource management</td>
</tr>
<tr>
<td>Data analysis</td>
<td>Analyzing large data sets (big data)</td>
</tr>
<tr>
<td>Prerequisites</td>
<td>Workable user interface, IT support, and better software</td>
</tr>
<tr>
<td>Education</td>
<td>Education specific to tools and AI in general</td>
</tr>
</tbody>
</table>

AI, artificial intelligence; IT, information technology.
as well as in subsequent paragraphs, along with quotes from the survey participants. It should be noted that the attitude theme could be interpreted as being a separate sentiment analysis. However, this is not the case. Attitude was merely a theme in which many answers could be categorized according to both researchers.

In 173 (73%) of the 237 coded cases, there was an initial agreement on the codes to be assigned. In 64 (27%) cases, when there was a discrepancy between codes assigned by the two different coders, an extensive consensus procedure was followed. This resulted in a 100% agreement between coders after this consensus procedure.

Current Uses of AI in Laboratory Medicine

AI is currently used in the organizations of 20 (15.6%; 95% confidence interval [CI], 9.8%-23.1%) of the participants, while 85 (66.4%; 95% CI, 57.5%-74.5%) felt they might use it in the future and 8 (6.3%; 95% CI, 2.7%-11.9%) felt that they would never use AI. Respondents who have AI in their practice use it for diagnosing diseases from images (30%), reviewing patients’ risk profiles for certain conditions (40%), preempting rapid response solutions (30%), and automatically releasing laboratory results and financial analytics (10%). Examples of specific use cases include AI to perform digital cell analysis for peripheral blood, analyze medical record laboratory data to determine which patients are at risk of infection, improve patients’ outcomes and length of stays and readmissions, or preempt rapid response situations in hospitals and automated sepsis alerts to identify patients early.

When asked how respondents felt that these current applications could be improved, most of them answered that certain prerequisites (eg, user interface, IT support, and better software) were needed. All 20 participants who currently use AI answered this question; see Figure 1 for counts against themes. For example, respondent 119 answered, “reduce the number of pop-ups in the EMR [electronic medical record],” and respondent 70 said, “We use AI for chatbots about common questions for diagnostics. The AI chatbots are not very intelligent. Need to make AI smarter.”

Value of AI in Practice

Regarding the potential use of AI in the diagnostics space, of the 90 (81%; 95% CI, 72.6%-87.9%) participants who felt AI will be valuable in their organization within the next 5 years, 20 (18%; 95% CI, 11.4%-26.5%) labeled it as expected to be extremely valuable. Twenty-one (19%; 95% CI, 12.1%-27.5%) of the respondents felt like AI will not be valuable in their organization within the next 5 years. There were missing data on this question for 17 participants. To further examine whether the results were different in subgroups of respondents, we dichotomized the answers to finding AI valuable (including extremely valuable, very valuable, and somewhat valuable; n = 90) or not valuable at all (n = 21). A χ² test showed no significant difference in how participants in the different age groups valued AI in the diagnostics space (χ² = 5.0947 [4 degrees of freedom]; P = .28). Also, there was no significant difference in how AI was valued between respondents who currently use AI in their practice (n = 17) compared with respondents who have not used it yet (χ² = 0.6698 [1 degree of freedom]; P = .41). Table 3 shows the number of respondents to find AI valuable or not valuable in the different subgroups.

Valuable

Respondents were also given the opportunity to elaborate on why they thought AI would or would not be valuable in their organizations. A plethora of reasons were given, and all were coded by theme, as shown in Figure 2. Most answers indicated that AI could be valuable because

<table>
<thead>
<tr>
<th>Table 3</th>
<th>Categorized Subgroup Results for Finding AI Valuable or Not Valuable in the Diagnostics Space</th>
</tr>
</thead>
<tbody>
<tr>
<td>Characteristic</td>
<td>Valuable, No. (%)</td>
</tr>
<tr>
<td>Age, y</td>
<td></td>
</tr>
<tr>
<td>31-40</td>
<td>17 (85)</td>
</tr>
<tr>
<td>41-50</td>
<td>22 (69)</td>
</tr>
<tr>
<td>51-60</td>
<td>25 (83)</td>
</tr>
<tr>
<td>61-70</td>
<td>23 (88)</td>
</tr>
<tr>
<td>70+</td>
<td>3 (100)</td>
</tr>
<tr>
<td>Experience</td>
<td></td>
</tr>
<tr>
<td>Use AI</td>
<td>15 (88)</td>
</tr>
<tr>
<td>Do not use AI</td>
<td>75 (80)</td>
</tr>
</tbody>
</table>

AI, artificial intelligence.
*Percentage calculated row-wise.
of the “organizational value” (eg, quicker results, reduced redundancy, and resource management). As an example, respondent 87 answered, “it could make the lab more efficient by streamlining work flow.” Another frequently reported theme was “quality of care” (eg, accessibility of care, accuracy, and early recognition). This was illustrated by respondent 100, who said, “might help in keeping patients informed of test results/appointments/follow up more efficiently,” and respondent 47 believed “could have some useful clinical algorithms to identify problems before they are known by humans, but the technology is still in early development.” Another substantial part of respondents, who thought AI would be valuable, were not sure about the reasons for this (“attitude—unsure”), as suggested by respondent 68: “I’m not entirely sure, I just know something is there!”

Not Valuable

The 19% (95% CI, 12.1%-27.5%) of respondents who did not consider AI to be valuable in their organizations in the next 5 years had more uniform responses. The answers were largely split between the themes’ “pre-requisites” (eg, budget and strategic plan) and an unsure attitude. See Figure 3 for more details. The missing pre-requisites, for example, were presented by respondent 47: “very expensive and we have very limited capital dollars that we need to use to refresh old technology,” and respondent 106 said, “it’s not in our strategic plan to implement AI at this time.” The unsure attitude toward AI was summed up by respondent 75: “I’m not sure about the use of AI.”

Requirements for Implementing AI

Participants were asked what they would need to feel comfortable with using AI in the diagnostics space. Most respondents had an “unsure attitude” toward what they needed most to adopt AI in their practice. For example, respondent 69 said, “this seems like too forward thinking of a question that we aren’t yet prepared to answer.” Others felt like they need education (eg, specific to tools and on AI in general). Respondent 40 answered “specific training to device” and respondent 46 said “AI short course training.” Most of the remaining group felt like they needed various pre-requisites (eg, support systems, certifications, and evidence of benefits) to feel comfortable to adopt AI in their practice. There were missing data for 23 of the participants; see Figure 4 for the counts per theme.

In the next question, participants were asked to specify how they would like to be trained to use AI. Twenty-three (22%) of 105 (95% CI, 14.4%-31.0%) participants felt that they were not able to speak to how they should be educated on these new technologies. Finally, we asked participants to select persons within their organizations who they felt should be involved in the selection of AI equipment. Up to 10 individuals across an organization were identified who could be involved in evaluating a potential AI diagnostic solution. Respondent 56 said, “Medical staff committees and physicians and mid-level providers that use AI, utilization review staff
that monitor provider performance, IT department and leadership that maintain AI software.”

Discussion

With this survey, we aimed to evaluate the thoughts of stakeholders in laboratory medicine on the value of AI in the diagnostics space and identify anticipated challenges with the introduction of AI in this field. About four in five respondents believed AI will be valuable in their organization within the next 5 years, mostly because of organizational and patient-oriented benefits. One in five respondents do not see any value in AI, which is often because of the lack of prerequisites such as budget for the implementation and not being adopted as part of an overall strategic initiative. Maybe this is because of the bias management might have toward AI.

The Value of AI

The quantitative results in this survey are similar to those found in surveys on this subject among other health care professionals. Similar to our findings, surveys on AI for pathologists, medical students, physicians, and radiologists all found that about 80% of participants believe AI will be influential or valuable in their practice in the upcoming years. Interestingly enough, a 2019 survey shows that 84% of the general population in the United States thinks that AI will be at the center of the next technological revolution.

All surveys on this subject in the medical community seem to show similar results regarding the perceived value of AI, independent of age or experience with AI in clinical practice. The fact that these results also overlap with the value of AI as perceived by the general population raises concerns that specific knowledge on this subject has not yet penetrated the medical community at large and that the surveys on this subject just reflect the ongoing AI hype. Our survey adds to this concern by showing that many respondents are unsure about why AI would or would not be valuable, what is needed to comfortably adopt AI, or how to be educated on AI.

There certainly seems to be a disconnect between the more positive views of information experts on AI and views of the medical community. To get all the benefits AI presents, while keeping its drawbacks to a minimum, drastic changes are needed in the medical community. There is a need for general AI training to the various health care stakeholders as identified in a recent publication on the need to introduce AI training in medical education. In the meantime, training on new AI tools should also be the responsibility of the companies who provide the algorithms through extensive web-based training, along with on-site hands-on training.

Health Care Costs

Another highlight from this survey was the potential of AI to target high health care expenditure, since it can reduce and replace repetitive manual labor. Recent study has shown that AI can help reduce the waste in the US health care system in the range from $760 billion to $935 billion in 2019. Respondents believe AI can make the diagnostics process more efficient and decrease costs. For example, safely reducing the number of laboratory tests ordered or the frequency of ordering repeat tests is illustrated with a quote from respondent 94: “Alert me to the fact that a lab test I ordered was already completed at another hospital system in the past week.”

Impact on Jobs

Finally, we learned from the discussion board conducted prior to developing the survey that advisers also mentioned that they know AI is rapidly increasing in importance and value and want it to evolve their roles rather than replace them. “AI will become an integral concept for health care. Whether diagnosis or process improvement in medicine, AI will impact the industry. For personalized medicine and improving diagnostic accuracy AI will drive decision making in the hands of providers,” said respondent 12, an executive at a large integrated health network. Laboratory managers similarly think that AI could create efficiencies that expedite their workflows but want to ensure that they are still in control. Respondent 31 believes, “AI needs to be used in the right spaces and not to eliminate med techs but to supplement them.”
Implementation Strategies

In this survey, 19% of respondents did not see the added value of implementing AI in laboratory medicine, partially because of the high initial investment costs. This will be a limiting factor as long as the return on investment and clinical benefits of these tools are not well understood. A recent narrative review on the clinical applications of AI for sepsis validates this idea by identifying that a large gap still remains between the development of AI algorithms and their clinical implementation. The question remains whether this gap between development and clinical implementation might be caused by the resistance to implementing new technologies. Unfortunately, this can hold back research on this subject and thereby delay the gathering of evidence on whether AI tools can be beneficial and cost-effective in clinical practice on a large scale.

Although it was not mentioned in the survey itself, an interesting strategy to implement AI in laboratory medicine was discussed in one of the group chats that was used to shape the final survey. One of the participants disclosed that in their hospital, a new AI tool was introduced alongside an existing tool that was used in routine clinical practice. The old tool was still used but as a backup for when the AI tool failed. Practitioners were encouraged to try the AI tool but could choose either of the available options. They gradually got familiar with the AI tool and could see the added value firsthand. They ended up switching to the AI tool completely. This illustrates a viable way to integrate AI tools in health care. Although more expensive, this provides an opportunity to compare these tools in practice and allow the practitioners to feel comfortable with the tool before having to rely on it completely. See Table 4 for our key recommendations.

Patient Viewpoint

The overarching goal of implementing AI in clinical practice is to benefit the patient. Therefore, the patient’s perspective should also be discussed. One of the respondents posed an interesting question in the online group chat: “Should the patients be informed that some of the decisions are being recommended by AI?” Another question is whether we should inform patients when an AI recommendation is not followed. Unfortunately, this burdens the patient as they now have to choose between the physician and the computer. Many algorithms are already being used in medicine, like the YEARS criteria for pulmonary embolisms. Their role in the diagnostic process is rarely explained to or discussed with the patient, as only trained physicians can interpret the results of these algorithms. We therefore believe that a similar approach might be best when using more advanced algorithms, in which explainability and interpretation are an even larger problem. These tools and algorithms are an aid to complement the healthcare practitioners who are eventually responsible for the diagnostic process and decision making. Finally, from a provider’s viewpoint, they will need to know details of the algorithms they use to make decisions.

Strengths of the Survey

We addressed a target population of participants who are currently in a position to influence organizational policies to either embrace new technologies or refrain from using them in their laboratories. Any specific intervention to encourage the introduction of AI in the diagnostics space should be tailored to such a population of decision makers. Another strength is that the results were independently analyzed by two researchers with different backgrounds, thereby minimalizing the chance of interpretation bias. Finally, the questions were extensively scrutinized in the initial discussion board and group chats prior to fielding the final survey.

Limitations of the Survey

The participants did not represent the entire population of practitioners who will be using AI in a diagnostic setting. We cannot generalize these findings to all laboratory medicine practitioners across multiple types of settings. Finally, the study population (n = 128) was relatively small for quantitative analyses, perhaps causing the nonsignificance of the χ² tests. Only a large difference in how AI was valued between groups would have shown significant results in the quantitative analysis.

Conclusions

This survey shows that many stakeholders in laboratory medicine think that AI will be valuable to them in the near
future, mostly given the “organizational value” and expected improvements in “quality of care,” although vital prerequisites such as support systems, strategic plans, and budgets need to be provided. The overall response to this and other similar surveys raises the concern that specific knowledge on AI in the medical community at large is still poor. AI education in the medical community is much needed. As suggested by one respondent, one strategy to implement new AI tools could be to implement it alongside existing tools, so that practitioners can feel comfortable with the new tools and experience their added value in practice firsthand while awaiting further research studies on the clinical evidence, implementation, and benefits of AI.

References


