

# A Review on the Current Applications of Artificial Intelligence in the Operating Room

Surgical Innovation  
2021, Vol. 28(5) 611–619  
© The Author(s) 2021



Article reuse guidelines:  
[sagepub.com/journals-permissions](https://sagepub.com/journals-permissions)  
DOI: 10.1177/1553350621996961  
[journals.sagepub.com/home/sri](https://journals.sagepub.com/home/sri)



David C. Birkhoff, BSc<sup>1</sup>, Anne Sophie H.M. van Dalen, MD<sup>1</sup> , and Marlies P. Schijven, MD, PhD, MHS<sup>2,3</sup> 

## Abstract

**Background.** Artificial intelligence (AI) is an era upcoming in medicine and, more recently, in the operating room (OR). Existing literature elaborates mainly on the future possibilities and expectations for AI in surgery. The aim of this study is to systematically provide an overview of the current actual AI applications used to support processes inside the OR. **Methods.** PubMed, Embase, Cochrane Library, and IEEE Xplore were searched using inclusion criteria for relevant articles up to August 25th, 2020. No study types were excluded beforehand. Articles describing current AI applications for surgical purposes inside the OR were reviewed. **Results.** Nine studies were included. An overview of the researched and described applications of AI in the OR is provided, including procedure duration prediction, gesture recognition, intraoperative cancer detection, intraoperative video analysis, workflow recognition, an endoscopic guidance system, knot-tying, and automatic registration and tracking of the bone in orthopedic surgery. These technologies are compared to their, often non-AI, baseline alternatives. **Conclusions.** Currently described applications of AI in the OR are limited to date. They may, however, have a promising future in improving surgical precision, reduce manpower, support intraoperative decision-making, and increase surgical safety. Nonetheless, the application and implementation of AI inside the OR still has several challenges to overcome. Clear regulatory, organizational, and clinical conditions are imperative for AI to redeem its promise. Future research on use of AI in the OR should therefore focus on clinical validation of AI applications, the legal and ethical considerations, and on evaluation of implementation trajectory.

## Keywords

artificial intelligence, black box, machine learning, surgery, operating room, innovation, deep learning, neural networks, computer vision

## Introduction

The last few years have seen a tremendous growth in the use of sensors, video, and digital devices in the operating room (OR).<sup>1–3</sup> These applications generate large amounts of data in various formats, often referred to as “big data.”<sup>4</sup> Big data sets are complex and may be analyzed computationally to reveal patterns, trends, and associations, especially relating to human behavior and interactions. Big data has the potential to become progressively useful in both guiding surgical care and optimizing clinical patient outcomes, if handled well.<sup>5–8</sup> A limitation often overseen in analyzing big data is that traditional data processing techniques are not able to handle these vast amounts of complex data.<sup>9</sup> The solution may lie in a research area that became popularly known as “artificial intelligence (AI).” The term AI is often used to describe the study of algorithms that enables machines to reason and perform cognitive functions such as learning, problem-solving, and decision-making.<sup>10,11</sup> Recently, AI has made its introduction into medicine and, even more

recently, into the OR.<sup>2</sup> This is of interest as these high-risk environments are considered to be one of the most error-prone areas in the hospital, where outcome is highly dependent on use of modern technology generating multisource data.<sup>12,13</sup> As such, if properly used, AI may have great impact on surgical workflow and outcome. It may provide context-aware perioperative decision support, predict patterns in patient parameters, monitor progress, and develop new in situ training tools.<sup>14–17</sup> These are just a few examples. To date, AI applications

<sup>1</sup>Department of Surgery, Amsterdam UMC, University of Amsterdam, The Netherlands

<sup>2</sup>Department of Surgery, Amsterdam Gastroenterology and Metabolism, University of Amsterdam, The Netherlands

<sup>3</sup>Li Ka Shing Knowledge Institute, St Michaels Hospital, Toronto, Canada

### Corresponding Author:

Marlies P. Schijven, Department of Surgery, Amsterdam UMC, Location AMC, Meibergdreef 9, Amsterdam | 105 AZ, The Netherlands. Email: [m.p.schijven@amsterdamumc.nl](mailto:m.p.schijven@amsterdamumc.nl)

are painting and predicting a promising future surgical landscape. Yet, as is often the case with new innovations, AI may become lost in its promise when it is unclear what the actual baseline and best use case is.<sup>18–20</sup>

The current medical literature fixates predominantly on the future possibilities of AI in surgery, or more specifically, inside the OR. However, it is important to know the current situation—where does AI in the OR stand?—in order to validly decide on areas worthy of further exploration. The aim of this study is to systematically provide an overview of the current AI applications in surgery, used to support various processes inside the OR.

## Methods

### Literature Search

A systematic literary search was performed up to August 25th, 2020 using the following online databases: PubMed, Embase, Cochrane Library, and IEEE Xplore. The terms AI, OR, and surgery, including synonyms or equivalent terms, were used in certain combinations to obtain the relevant literature. The full search strategy can be found in [Supplemental Appendix A](#).

Article screening was done independently by 2 reviewers (DCB and AvD). The inclusion criteria were as follows: (1) AI, (2) in surgery, and (3) in the OR. The exclusion criteria, next to duplicates and articles older than 10 years, were the following: (1) articles published in any language other than Dutch or English, (2) articles containing future applications of AI only, (3) AI used outside the OR, and (4) no full-text availability. Any study design may benefit the study, so no specific study designs were excluded beforehand. Disagreement between the two reviewers in study selection was resolved by healthy discussions concluding in consensus.

The studies that were included after full-text screening were critically appraised, with the use of an Evidence-Based Medicine Critical Appraisal Checklist (see [Supplemental Appendix B](#)).

### Data Extraction

The included articles were extracted of data on study design, publication year, country of origin, and the specific researched applications of AI. The outcomes of these studies were analyzed and described and, if possible, defined in numbers. A clear overview of the different studies, their applications of AI and their specifically used subfield of AI, and their data type/source is provided. AI, while not easily defined, is a machine's capability to mimic intelligent human behavior.<sup>21</sup> AI is a broad field to be distinguished by multiple subfields. In order to better understand the analyses and outcomes of the studies, it was decided to explain some of the different subfields in

AI beforehand. The subfields that are of importance to this systematic review are explained and elaborated on in [Table 1](#).

## Results

### Search Results and Study Selection

The literature search yielded 193 articles from PubMed database, 50 articles from Embase database, 5 articles from the Cochrane Library, and 27 articles from IEEE Xplore database. Finally, 9 articles were included. The flowchart with a more detailed description of the selection procedure may be viewed in [Figure 1](#). The nine included studies are the following: Bodenstedt et al.,<sup>22</sup> Cho et al.,<sup>23</sup> Devi et al.,<sup>24</sup> Hashimoto et al.,<sup>25</sup> Jermyn et al.,<sup>26</sup> Kassahun et al.,<sup>27</sup> Padoy,<sup>17</sup> Zhao et al.,<sup>28</sup> and Liu et al.<sup>29</sup>

### Critical Appraisal

Only the 2 included review studies by Padoy<sup>17</sup> and Kassahun et al.,<sup>27</sup> were critically appraised. As a consequence of inhomogeneity in study design, the additional seven included studies did not contain a sufficient amount of checklist characteristics and were therefore not suitable for critical appraisal. Although both review studies scored negatively on many criteria, indicating that the quality of the studies should be considered relatively low, these studies were not of a regular review design either and were therefore difficult to classify.

### Applications of AI

The included articles respectively researched one or multiple applications of AI in surgery. [Table 2](#) shows an overview of the different studies, their researched application(s), and the specific AI subfield(s) the application is based on. Additionally, [Table 2](#) specifies the data type/source that was used by the AI application.

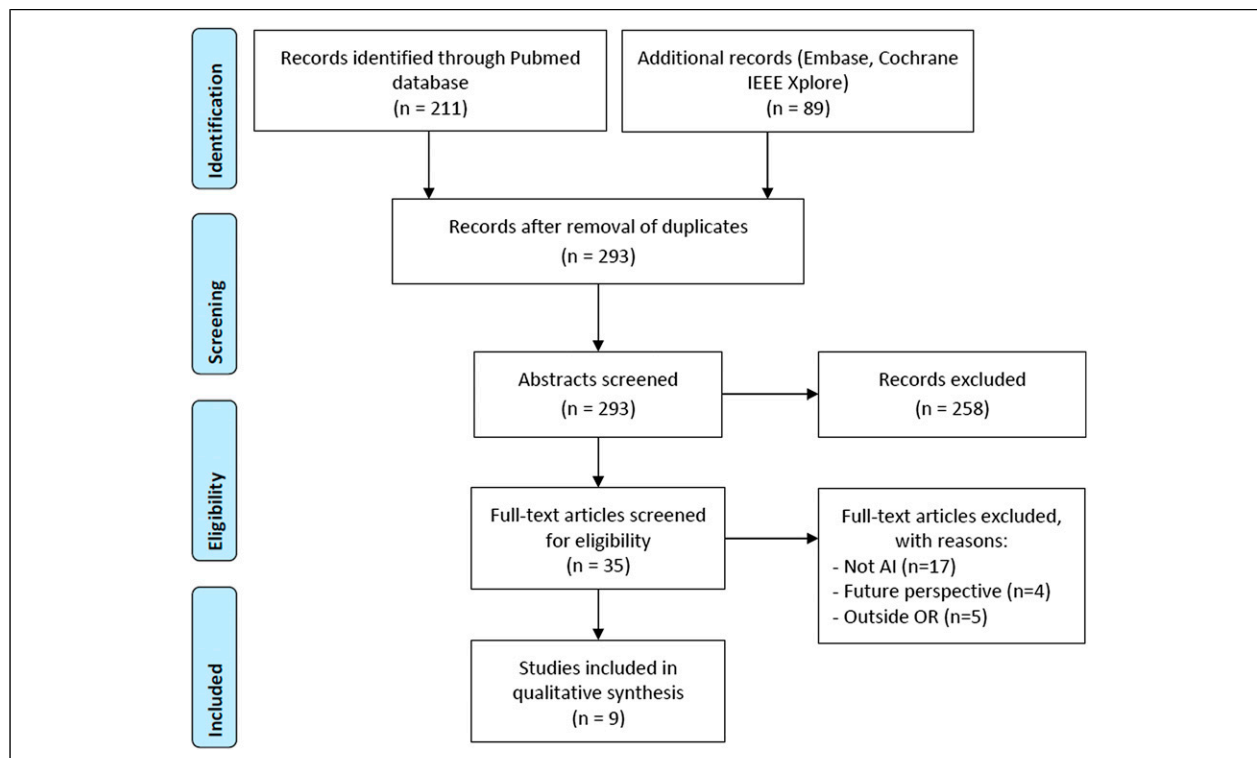
### Procedure Duration Prediction

Due to the high density and non-singularity of information in a video stream, extracting its data for evaluation purposes is a challenging process. In comparison to the video stream, data from surgical instruments provide information that is easier to quantify. Whether or not such data provide sufficient information to make presumptive predictions on surgery duration is uncertain to date. Bodenstedt et al.<sup>22</sup> proposed and compared methods, based on CNNs to predict procedure duration based on data from surgical devices or video streams. The input was acquired from 80 recorded laparoscopic interventions of which the necessary data were available. Overall, the combined method (both video and surgical device data)

**Table 1.** Definitions of major subfields in artificial intelligence.

Subfield	Definition
Machine learning (ML)	Gives computers via algorithms the ability to modify its processing when exposed to more information, without being specifically programmed to do so. <sup>57,58</sup> In this way, computers are capable of “learning from experience.” <sup>21,58,59</sup> ML is considered to be promising in pattern recognition in large cohorts of data by using more complex techniques than traditional statistical analysis does. <sup>60,61</sup>
Artificial neural networks (ANNs)	Are tools used in ML. In function, they are imitating the human brain by connecting and finding interrelated complex relationships and patterns between data. <sup>2</sup> Basically, ANNs are composed of many computational units (neurons) that receive inputs, perform calculations, and direct output to the next computational unit. In other words, the input is being processed as signals through layers of algorithms that create certain patterns as final output; these patterns are interpreted and used in decision-making. <sup>62</sup> ANNs are commonly composed of simple 1- or 2-layered neural networks.
Deep learning	Deep learning networks consist of many layers and are able to recognize and learn more subtle and complex patterns. <sup>63</sup> Deep learning networks may take one or more datasets into account, which are evaluated multiple times in many different layers, until reaching the desired output. <sup>21</sup>
Convolutional neural networks (CNNs)	A convolutional neural network (CNN) is a class of ANNs that specializes in processing data in visualized imagery. In deep learning, “convolution” is a specialized kind of linear operation used in analyzing images, and in CNNs, the ANN employs this mathematical operation in at least one of its layers, hence the name convolutional. <sup>64</sup>
Computer vision (CV)	Focuses on how computers can gain high-level understanding of digital images and videos such as object and scene recognition, comparable to the human visual system. <sup>25,65,66</sup> The processed data may consist of video sequences, views from multiple cameras, or multidimensional data from a medical scanning device. <sup>67,68</sup>

Abbreviations: ML = machine learning; ANN = artificial neural networks; CV = computer vision; CNN = convolutional neural network

**Figure 1.** Flowchart of literature search.

**Table 2.** Overview of included studies with specific AI application(s).

Application(s)	Study	AI Subfield(s) <sup>a</sup>	Data Type/Source
Procedure duration prediction	Bodenstedt et al. <sup>22</sup>	CNN, deep learning, and ML	Video stream
	Zhao et al. <sup>28</sup>	ANN and ML	Case characteristics <sup>b</sup>
	Devi et al. <sup>24</sup>	ANN and ML	Surgical environment <sup>c</sup>
Gesture recognition	Cho et al. <sup>23</sup>	CNN and ML	Depth video stream and radiological images
Intraoperative cancer detection	Jermyn et al. <sup>26</sup>	ANN and ML	Spectral light
Intraoperative video analysis	Hashimoto et al. <sup>25</sup>	ANN, CV, and ML	Video stream
Workflow recognition	Padoy <sup>17</sup>	CNN, CV, RNN, deep learning, and ML	Video stream and images
Endoscopic guidance system and knot-tying	Kassahun et al. <sup>27</sup>	RNN and ML	Video stream
Automatic registration and tracking of the bone in orthopedic surgery	Liu et al. <sup>29</sup>	ANN, deep learning, and ML	Depth camera images

<sup>a</sup>Abbreviations: CNN = convolutional neural network, ML = machine learning, ANN = artificial neural network, CV = computer vision, RNN = recurrent neural network.

<sup>b</sup>Scheduled duration, age, gender, and comorbidities of the patient, tumor location, month of year, time of day, day of the week, etc.

<sup>c</sup>Experience of surgeon in years, experience of anesthetist in years, staff experience in years, type of anesthesia, etc. The actual set of environment variables depends on the type of surgery.

performed best with an average error of 37% and an average halftime error of approximately 28%. This is an improvement to the baseline method with an average error and average halftime error of both 124%.<sup>22</sup>

Zhao et al. sought to accurately predict procedure duration of robot-assisted surgery cases using multiple machine learning (ML) models, using case characteristics (scheduled duration, age, gender, and comorbidities of the patient, tumor location, month of year, time of day, day of the week etc.) as data input. They compared the ML models to the baseline model, which is the time scheduled for the procedure determined by former case duration averages and changes by the surgeon. The following ML models were used: (1) multivariable linear regression, (2) ridge regression, (3) lasso regression, (4) random forest, (5) boosted regression tree, and (6) ANNs. The average root-mean-squared error (RMSE), a measure for the imperfection of the fit of the estimator to the data, was lower for all the ML models than the baseline model. The average RMSE was lowest with the boosted regression tree (80.2 minutes, 95% confidence interval 74.0–86.4), which was significantly lower than the baseline model (100.4 minutes, 95% Confidence interval 90.5–110.3). The use of a boosted regression tree, a predictive modeling approach used in ML, increased the amount of correctly booked procedures from 148 to 219 (34.9% to 51.7%,  $P < .001$ ).<sup>28</sup>

Devi et al. researched several techniques to estimate procedure duration in an ophthalmology department by taking the surgical environment into account (experience of surgeon in years, experience of anesthetists in years, type of anesthesia, etc.). Three techniques were

**Table 3.** Comparison of techniques to estimate procedure duration.<sup>24</sup>

Type of Surgery	Root-Mean-Squared Error (RMSE)		
	ANFIS	ANN	Regression
Corneal transplant	.1557	.1895	.2755
Cataract	.0697	.1427	.1768
Oculoplastic	.1431	.1668	.2123

Abbreviations: ANN = artificial neural networks; ANFIS = adaptive neuro-fuzzy inference systems.

researched, namely, adaptive neuro-fuzzy inference systems (ANFISs), multiple linear regression analysis (MLRA), and ANNs. However, ANFIS is a fusion between the adaptive learning capability of ANNs and the intuitive logic of human reasoning, formulated as a feed-forward neural network. The results of procedure duration prediction were compared between the three techniques, and the ANFIS model came out to be performing better than the other 2 as portrayed in [Table 3](#).<sup>24</sup>

### Gesture Recognition

To decrease the risk of contamination during surgical procedures, Cho et al.<sup>23</sup> researched a noncontact interface based on ML models in order to enhance the accuracy of gesture recognition. Support vector machines (SVMs) and naive Bayes classifiers, ML models with associated algorithms used for classification, were used in the study.<sup>30</sup> Cho et al. used 30 features, including hand and finger data, as input for these ML models to predict and train 5 types

of gestures. The overall accuracy of the 5 gestures was  $99.58\% \pm .06$  and  $98.74\% \pm 3.64$ , respectively, for SVM and naive Bayes classifiers. Self-training methods of SVMs and naive Bayes classifiers improved accuracies by about 5–10%.<sup>23</sup>

### *Intraoperative Cancer Detection*

During brain tumor removal it is important yet very difficult to detect and remove all cancer cells. As a consequence, when not completely removed, the patient is at risk for recurrence of cancer. With certain types of brain cancer in vivo, Raman spectroscopy can detect these invasive cancer cells. A downside to this technique is the fact that the Raman signal is weakened by spectral artifacts generated by the regular lights in the OR. Jermyn et al. found that ANNs are able to improve the detection of invasive brain cancer cells by overcoming the negative impact of spectral artifacts. Despite the inclusion of light artifacts, ANNs keep the detection of invasive cancer cells at almost the same level, improving sensitivity by 19% and specificity by 7% compared to the standard technique.<sup>26</sup>

### *Intraoperative Video Analysis*

Video data of laparoscopic procedures are used for both education and quality improvement purposes. In order to decrease the required time for analysis and review of video data, Hashimoto et al. investigated the possibility of automatic video segmentation using CV and ML techniques. Their research demonstrated that CV and ML techniques were able to differentiate between specific steps of laparoscopic surgery procedures with an accuracy of  $82\% \pm 4\%$ .<sup>25</sup>

### *Workflow Recognition*

The long-term vision of Padoy<sup>17</sup> is to develop a surgical control tower (SCT) that, using AI, can monitor and support many processes, providing overall awareness of what is happening in the OR. Key for such an SCT is the requirement of an AI system that can recognize the surgical workflow and is aware of the surgical context. Workflow is often described as the sequence of tasks, interactions, or other processes through which a piece of work passes from initiation to completion.<sup>31</sup> In their review, Padoy<sup>17</sup> researched several recent ML and deep learning applications that can add to the workflow recognition system. These applications include phase recognition, tool detection and localization, and human detection and pose estimation and are described below.<sup>17</sup>

### *Phase Recognition*

Phase recognition, the task of instantly determining the current phase of surgery at any time  $t$  from video data, was researched both in laparoscopic videos and external videos. In laparoscopic videos, a study showed that the combination of a CNN and a recurrent neural network (RNN) was able to recognize the different phases automatically and in real time, with an accuracy of 86%. In a study using external videos, a combination of a CNN and hidden Markov models (HMMs), a popular application for ordinal or temporal data within AI, recognized different phases in the surgical procedure with an accuracy of 90%.<sup>17</sup>

### *Tool Detection and Localization*

Tool detection and localization adds to the precision of phase recognition. By recognizing more subtle and detailed activities, tool detection and localization may be informative for predicting operative steps and length of operation. Deep learning techniques were used to research tool detection and localization in laparoscopic images and videos. Using a CNN, results show a mean average precision of 87% in tool detection and 88% in tool localization.<sup>17</sup>

### *Human Detection and Pose Estimation*

Since the people are the main actors in the OR, detecting their position and estimating their poses by localizing their body parts can provide useful information for optimizing workflow. With the use of external videos, the ability to estimate the specific body poses of the people in the OR was investigated. The mean per joint position error (MPJPE) was used as a quantitative measure for 2D and 3D body part localization. Deep learning approaches yielded the best results in both 2D and 3D pose estimation with an average MPJPE of 17 and 5 cm, respectively.<sup>17</sup>

### *Endoscopic Guidance System*

Weede et al. described an autonomous endoscopic guidance system based on ML. The system is capable of collecting and processing data on the movements of surgical instruments in recorded videos of surgical procedures. Subsequently, with the use of trajectory clustering, maximum likelihood classification, and HMMs, the system uses this information to predict trajectories that are used to guide the endoscope. The results show a hit rate of over 89% for predicting the movement of the surgeon's instruments, leading to 29.2% less camera movements and improved visibility.<sup>27,32</sup>

## Knot-Tying

Although in open surgery, knot-tying is part of basic skills and a relative fast procedure, in minimally-invasive surgery, laparoscopic knot-tying can take up to three minutes for a single knot to complete. Mayer et al.<sup>32</sup> described a system to speed up the knot-tying based on RNNs in robotic heart surgery. The surgeon presents a sequence (eg, examples of human-performed knot-tying) to the network and, an RNN with long-term storage learns the task. The preprogrammed controller was able to construct a knot in 33.7 seconds, whereas the use of an RNN provided—after learning from 50 previous runs—a speed improvement of almost 25%, producing a knot in 25.8 sec.<sup>27,33</sup>

## Automatic Registration and Tracking of the Bone in Orthopedic Surgery

In computer-assisted orthopedic surgery, registration of the bone plays a vital role as it describes the position of the patient in regard to the surgical system. This way, the surgical site can be correctly aligned according to the preoperative plan. Therefore, the precision of the registration has influence on all the following steps in the procedure. Liu et al.<sup>29</sup> describe a new way of automatic registration and tracking of the bone, based on depth imaging and deep learning. During surgery, a depth camera repeatedly captures depth images of exposed bone. Using these images, deep neural networks learn to localize, segment, and extract the surface geometry of the target bone. The extracted surface geometry is then compared to a preoperative model of the same bone for registration, making surgical intervention or invasive optical markers superfluous. Ex vivo experiments show a mean translational and rotational error of 2.74 mm and 6.66°, respectively. However, these accuracies are currently lower than conventional intraoperative registration methods based on optical markers.<sup>34,29</sup>

## Discussion

The results of this systematic review study provide an overview of various AI applications currently used for surgical purposes inside the OR. The great majority, of the AI applications have shown superior results in comparison to their non-AI alternatives. However, studies are set up in various pilot settings. The various applications are an indication of multi-field interest in finding use cases for AI in the OR, paired with a need for more clinical research across user settings. Many studies have shown significant technological performance in the field of AI, but only a small minority has been able to situate their impacts and associated changes in current health systems.<sup>35</sup>

According to Rogers'<sup>36</sup> widely used *Diffusion of Innovations* theory, adoption of innovative technology always involves early and late adopters. During the innovation process, where an individual is motivated to reduce uncertainty about the advantages and disadvantages of an innovation, it is important to emphasize the ethical and legal challenges.<sup>37,38</sup> Yet, sufficient political, regulatory, organizational, and clinical conditions for AI development and ethical use of sensitive information are still lacking and hence needed to implement AI applications safely and sustainably in the future.<sup>35,39,40</sup> Additional barriers for the widespread implementation of AI in health care may be unawareness on the topic or solutions, lack of user or implementation knowledge by the medical professionals and their workplace supporters, unresolved questions about ethics or privacy from management, or an insufficient IT infrastructure. Most likely, it will be a combination of these barriers.<sup>41</sup>

While AI, and ML in particular, is receiving more attention in surgery, it is obviously not the only field of medicine in which the use of AI is growing.<sup>27</sup> The surgical field may be able to learn from the use of AI in other medical fields. For example, in oncology, research has demonstrated that ML applications can be of great help for the diagnosis or detection of cancer.<sup>42-44</sup> In cardiology, AI techniques are capable of reading electrocardiograms, and by integration with electronic medical records of patients, heart failure can be detected early on with reduced mortality as outcome.<sup>45-47</sup> In anesthesiology, ANNs are used to monitor the depth of anesthesia, and ML techniques are able to predict hypotension during surgery.<sup>48,49</sup> And now, during the current COVID-19 pandemic, more AI applications and studies have been initiated.<sup>19,50,51</sup> The Guangdong Second Provincial General Hospital, for example, plans to incorporate AI image recognition into their infection control system to provide real-time monitoring and an aid for minimizing the risk nosocomial COVID-19 infection. The observing system aims to enhance the sensitivity and accuracy of instant detection in negative pressure isolation wards, which offers creative assistance to combat the COVID-19 outbreak.<sup>50</sup> This application may also be used in the OR to minimize the risk of surgical infection.

Indeed, AI in health care has presented some promising and impressive results and is a fertile area of research, as Challen et al.<sup>52</sup> concluded in their review. However, as this study shows the multilingual character of AI in surgery, AI is a complicated and comprehensive field of study. The rapid pace of change, diversity of different techniques, and multiplicity of tuning parameters make it difficult to get a clear picture of how accurate these systems might be in clinical practice or how reproducible they are in different environments.<sup>52</sup> A realistic perspective is needed, balancing the potential for improvement against the risk of negative outcomes. As Yu et al.<sup>8</sup>

also concluded, we need to acknowledge the brittleness of these systems, the importance of defining the correct frameworks for their application, and ensure rigorous quality control, including human supervision, to unwanted outcomes. Rigorous prospective trials in a diverse patient population and clinical review of atypical feature statistics are needed, to safeguard the value and coherency of the collected data.<sup>8,52</sup> It is therefore wise to attract knowledge coming from ML experts, ethicists, and lawyers, next to healthcare professionals, to decide on proper fit of use case and safety of AI systems.

This study has some limitations to take into account. First, as this is a review study, unpublished data and gray literature, such as technical reports, are not included, which may have strengthened the results. Moreover, the results may have been influenced by a publication bias, especially, because—as this study shows—AI assistance in the OR is still in its infancy. Park et al.<sup>53</sup> acknowledged the problem of irregular research designs in medical AI studies. This is also displayed by the significant variability in the way results are reported, making it very difficult to combine and compare data across studies. This results in the realization that before any AI tool can be used in clinical practice, it requires confirmation of its clinical utility by undergoing thorough research. In their article, they therefore described and reviewed essential methods on the design of such studies, like the importance of using an adequate external dataset, crucial to the clinical evaluation of AI in medicine.<sup>53</sup>

Second, the applications of AI discussed in this study are, although interesting in their pilot effort, not ready for large-scale clinical practice.<sup>54</sup> AI is not yet able to detect causal relationships in data at a necessary level for clinical implementation to rely on, nor is it able to produce truly automated interpretations of its analyses.<sup>54</sup> Before these implications can be clinically and safely applied in the OR on a bigger scale, future studies should focus on clinical studies, with data from actual patients.<sup>39</sup>

## Conclusion

AI systems inside the OR, if well-designed, embedded, and researched, may have a promising future in the OR environment. It may support surgical decision-making, improve surgical precision, reduce manpower, improve workflow, increase surgical safety, and some day it may even carry out some autonomous functions.<sup>6–8,16,21</sup> In the not so distant future, evolving technology like the OR black box, with integrated deep learning algorithms, may prove to be of great help in analyzing and optimizing workflow and outcome in real time.<sup>55</sup> Indeed, the application and implementation of AI inside the OR still has several challenges to overcome. However, evidence-based research adding to the body of knowledge concerning applications of AI inside the OR is moving quickly.

Healthcare professionals ought to accept the fact that we need AI in order to optimize future circumstances in the OR and ultimately, surgical quality and safety.<sup>14,55,56</sup>

## Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

## ORCID iDs

Anne Sophie H.M. van Dalen  <https://orcid.org/0000-0002-5468-9371>

Marlies P. Schijven  <https://orcid.org/0000-0001-7013-0116>

## Supplemental Material

Supplemental material for this article is available online.

## References

1. Al Hajj H, Lamard M, Conze P-H, Cochener B, Quellec G. Monitoring tool usage in surgery videos using boosted convolutional and recurrent neural networks. *Med Image Anal.* 2018;47:203-218.
2. Hashimoto DA, Rosman G, Rus D, Meireles OR. Artificial intelligence in surgery: Promises and perils. *Ann Surg.* 2018; 268(1):70-76. doi:10.1097/sla.0000000000002693
3. Malpani A, Lea C, Chen CCG, Hager GD. System events: readily accessible features for surgical phase detection. *Int J Comput Assist Radiol Surg.* 2016;11(6):1201-1209.
4. Languages OL. Definition of big data. Available from [https://www.lexico.com/definition/big\\_data](https://www.lexico.com/definition/big_data). accessed July 05, 2020.
5. Alderson D. The future of surgery. *BJS.* 2019;106(1):9-10. doi:10.1002/bjs.11086
6. Loftus TJ, Upchurch GR Jr, Bihorac A. Use of artificial intelligence to represent emergent systems and augment surgical decision-making. *JAMA Surgery.* 2019;154(9): 791-792. doi:10.1001/jamasurg.2019.1510.
7. Panesar S, Cagle Y, Chander D, Morey J, Fernandez-Miranda J, Kliot M. Artificial intelligence and the future of surgical robotics. *Ann Surg.* 2019;270(2):223-226. doi: 10.1097/sla.0000000000003262
8. Yu K-H, Kohane IS. Framing the challenges of artificial intelligence in medicine. *BMJ Qual Saf.* 2019;28(3): 238-241. doi:10.1136/bmjqs-2018-008551
9. Murdoch TB, Detsky AS. The inevitable application of big data to health care. *JAMA.* 2013;309(13):1351-1352. doi: 10.1001/jama.2013.393
10. Russell SJ, Norvig P. *Artificial Intelligence: A Modern Approach.* Malaysia: Pearson Education Limited; 2016.
11. Loftus TJ, Tighe PJ, Filiberto AC, et al. Artificial intelligence and surgical decision-making. *JAMA Surg.* 2020; 155(2):148-158. doi:10.1001/jamasurg.2019.4917
12. WHO. Safe. Available from [https://www.who.int/patientsafety/safesurgery/faq\\_introduction/en/](https://www.who.int/patientsafety/safesurgery/faq_introduction/en/). accessed May 23, 2020.

13. Makary MA, Daniel M. Medical error-the third leading cause of death in the US. *BMJ*. 2016;353, i2139. doi:10.1136/bmj.i2139
14. Gordon L, Grantcharov T, Rudzicz F. Explainable artificial intelligence for safe intraoperative decision support. *JAMA Surg*. 2019;154(11):1064-1065. doi:10.1001/jamasurg.2019.2821
15. Mirnezami R, Ahmed A. Surgery 3.0, artificial intelligence and the next-generation surgeon. *Br J Surg*. 2018;105(5):463-465. doi:10.1002/bjs.10860
16. Navarrete-Welton AJ, Hashimoto DA. Current applications of artificial intelligence for intraoperative decision support in surgery. *Front Med*. 2020;14, 369. doi:10.1007/s11684-020-0784-7
17. Padoy N. Machine and deep learning for workflow recognition during surgery. *Minim Invasive Ther Allied Technol*. 2019;28(2):82-90. doi:10.1080/13645706.2019.1584116
18. Chen JH, Asch SM. Machine learning and prediction in medicine - Beyond the peak of inflated expectations. *N Engl J Med*. 2017;376(26):2507-2509. doi:10.1056/NEJMp1702071
19. Desai AN. Artificial intelligence: Promise, pitfalls, and perspective. *J Am Med Assoc*. 2020;323, 2448. doi:10.1001/jama.2020.8737
20. Matheny ME, Whicher D, Thadaney Israni S. Artificial intelligence in health care: A report from the national academy of medicine. *J Am Med Assoc*. 2020;323(6):509-510. doi:10.1001/jama.2019.21579
21. Mintz Y, Brodie R. Introduction to artificial intelligence in medicine. *Minim Invasive Ther Allied Technol*. 2019;28(2):73-81. doi:10.1080/13645706.2019.1575882
22. Bodenstedt S, Wagner M, Mündermann L, et al. Prediction of laparoscopic procedure duration using unlabeled, multimodal sensor data. *Int J Comput Assist Radiol Surg*. 2019;14(6):1089-1095. doi:10.1007/s11548-019-01966-6
23. Cho Y, Lee A, Park J, Ko B, Kim N. Enhancement of gesture recognition for contactless interface using a personalized classifier in the operating room. *Comput Methods Progr Biomed*. 2018;161:39-44. doi:10.1016/j.cmpb.2018.04.003
24. Devi SP, Rao KS, Sangeetha SS. Prediction of surgery times and scheduling of operation theaters in ophthalmology department. *J Med Syst*. 2012;36(2):415-430. doi:10.1007/s10916-010-9486-z
25. Hashimoto DA, Rosman G, Witkowski ER, et al. Computer vision analysis of intraoperative video: Automated recognition of operative steps in laparoscopic sleeve gastrectomy. *Ann Surg*. 2019;270(3):414-421. doi:10.1097/sla.0000000000003460
26. Jermyn M, Desroches J, Mercier J, et al. Neural networks improve brain cancer detection with Raman spectroscopy in the presence of operating room light artifacts. *J Biomed Optic*. 2016;21(9):094002. doi:10.1117/1.Jbo.21.9.094002
27. Kassahun Y, Yu B, Tibebu AT, et al. Surgical robotics beyond enhanced dexterity instrumentation: A survey of machine learning techniques and their role in intelligent and autonomous surgical actions. *Int J Comput Assist Radiol Surg*. 2016;11(4):553-568. doi:10.1007/s11548-015-1305-z
28. Zhao B, Waterman RS, Urman RD, Gabriel RA. A machine learning approach to predicting case duration for robot-assisted surgery. *J Med Syst*. 2019;43(2):32. doi:10.1007/s10916-018-1151-y
29. Liu H, Baena FRY. Automatic markerless registration and tracking of the bone for computer-assisted orthopaedic surgery. *IEEE Access*. 2020;8:42010-42020. doi:10.1109/ACCESS.2020.2977072
30. Cortes C, Vapnik V. Support-vector networks. *Mach Learn*. 1995;20(3):273-297.
31. Adi A, Hadash K, Kerem O, et al. *Definition of Workflow Patterns Using Complex Event Processing*. Armonk, NY: Google Patents; 2006.
32. Weede O, Mönnich H, Müller B, Wörn H. An intelligent and autonomous endoscopic guidance system for minimally invasive surgery. Paper presented at: 2011 IEEE International Conference on Robotics and Automation; 9-13 May 2011; Shanghai, China. IEEE.
33. Mayer H, Gomez F, Wierstra D, et al. A system for robotic heart surgery that learns to tie knots using recurrent neural networks. *Adv Robot*. 2008;22(13-14):1521-1537.
34. Nishihara S, Sugano N, Ikai M, et al. Accuracy evaluation of a shape-based registration method for a computer navigation system for total knee arthroplasty. *J Knee Surg*. 2003;16(2):98-105
35. Alami H, Lehoux P, Auclair Y, et al. Artificial intelligence and health technology assessment: Anticipating a new level of complexity. *J Med Internet Res*. 2020;22(7):e17707.
36. Rogers EM. *Diffusion of Innovations*. New York City, NY: Simon & Schuster; 2010.
37. Sahin I. Detailed review of Rogers' diffusion of innovations theory and educational technology-related studies based on Rogers' theory. *TOJET*. 2006;5(2):14-23.
38. Dalen ASHM, Legemaate J, Schlack WS, Legemate DA, Schijven MP. Legal perspectives on black box recording devices in the operating environment. *Br J Surg*. 2019;106(0):1433. doi:10.1002/bjs.11198
39. Pesapane F, Volonté C, Codari M, Sardanelli F. Artificial intelligence as a medical device in radiology: Ethical and regulatory issues in Europe and the United States. *Insights Imaging*. 2018;9(5):745-753. doi:10.1007/s13244-018-0645-y
40. MacNeil M, Koch M, Kuspinar A, Juzwishin D, Lehoux P, Stolee P. Enabling health technology innovation in Canada: Barriers and facilitators in policy and regulatory processes. *Health Pol*. 2019;123(2):203-214.
41. Cath C. Governing artificial intelligence: Ethical, legal and technical opportunities and challenges. *Phil Trans Math Phys Eng Sci*. 2018;376:20180080. doi:10.1098/rsta.2018.0080
42. Bejnordi BE, Veta M, Van Diest PJ, et al. Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer. *JAMA*. 2017;318(22):2199-2210.
43. Somashekhar SP, Sepúlveda M-J, Puglielli S, et al. Watson for oncology and breast cancer treatment recommendations: Agreement with an expert multidisciplinary tumor board. *Ann Oncol*. 2018;29(2):418-423.



44. Rahman SA, Walker RC, Lloyd MA, et al. Machine learning to predict early recurrence after oesophageal cancer surgery. *BJS*. 2020;107(8):1042-1052. doi:10.1002/bjs.11461
45. Buzaev IV, Plechev VV, Nikolaeva IE, Galimova RM. Artificial intelligence: Neural network model as the multidisciplinary team member in clinical decision support to avoid medical mistakes. *Chronic Dis Transl Med*. 2016; 2(3):166-172.
46. Choi E, Schuetz A, Stewart WF, Sun J. Using recurrent neural network models for early detection of heart failure onset. *J Am Med Inf Assoc*. 2016;24(2):361-370.
47. Slomka PJ, Dey D, Sitek A, Motwani M, Berman DS, Germano G. Cardiac imaging: Working towards fully-automated machine analysis & interpretation. *Expet Rev Med Dev*. 2017;14(3):197-212.
48. Gu Y, Liang Z, Hagihira S. Use of multiple EEG features and artificial neural network to monitor the depth of anesthesia. *Sensors*. 2019;19(11):2499. doi:10.3390/s19112499
49. Wijnberge M, Schenk J, Terwindt LE, et al. The use of a machine-learning algorithm that predicts hypotension during surgery in combination with personalized treatment guidance: study protocol for a randomized clinical trial. *Trials*. 2019;20(1):582. doi:10.1186/s13063-019-3637-4
50. Chen X, Tian J, Li G, Li G. Initiation of a new infection control system for the COVID-19 outbreak. *Lancet Infect Dis*. 2020; 20(4):397-398. doi:10.1016/S1473-3099(20)30110-9
51. McCall B. COVID-19 and artificial intelligence: Protecting health-care workers and curbing the spread. *The Lancet Digit Health*. 2020;2(4):e166-e167.
52. Challen R, Denny J, Pitt M, Gompels L, Edwards T, Tsaneva-Atanasova K. Artificial intelligence, bias and clinical safety. *BMJ Qual Saf*. 2019;28(3):231-237.
53. Park SH, Han K. Methodologic guide for evaluating clinical performance and effect of artificial intelligence technology for medical diagnosis and prediction. *Radiology*. 2018; 286(3):800-809. doi:10.1148/radiol.2017171920
54. Pearl J. *Causality: Models, Reasoning and Inference*. New York City, NY: Springer; 2000.
55. Goldenberg MG, Jung J, Grantcharov TP. Using data to enhance performance and improve quality and safety in surgery. *JAMA Surg*. 2017;152(10):972-973.
56. Khalid S, Goldenberg M, Grantcharov T, Taati B, Rudzicz F. Evaluation of deep learning models for identifying surgical actions and measuring performance. *JAMA Netw Open*. 2020;3(3):e201664. doi:10.1001/jamanetworkopen.2020.1664
57. Bishop CM. *Pattern Recognition and Machine Learning*. New York City, NY: Springer; 2006.
58. Mitchell TM. Does machine learning really work? *AI Mag*. 1997;18(3):11.
59. Fürnkranz J, Gamberger D, Lavrač N. *Foundations of Rule Learning*. New York City, NY: Springer Science & Business Media; 2012.
60. Cruz JA, Wishart DS. Applications of machine learning in cancer prediction and prognosis. *Canc Inf*. 2006;2: 117693510600200030.
61. Miller RA, Pople HE Jr, Myers JD. Internist-I, an experimental computer-based diagnostic consultant for general internal medicine. *N Engl J Med*. 1982;307(8):468-476.
62. Deo RC. Machine learning in medicine. *Circulation*. 2015;132(20):1920-1930. doi:10.1161/circulationaha.115.001593
63. Hinton GE, Osindero S, Teh Y-W. A fast learning algorithm for deep belief nets. *Neural Comput*. 2006;18(7): 1527-1554.
64. Goodfellow I, Bengio Y, Courville A. *Deep Learning*. Cambridge, MA: MIT press; 2016.
65. Ballard DH, Brown CM. *Computer Vision*. Upper Saddle River, NJ: Prentice-Hall; 1982.
66. Szeliski R. *Computer Vision: Algorithms and Applications*. New York City, NY: Springer Science & Business Media; 2010.
67. Huang T. *Computer Vision: Evolution and Promise*. Champaign, IL: University of Illinois Press; 1996.
68. Sonka M, Hlavac V, Boyle R. *Image Processing, Analysis, and Machine Vision*. Boston, MA: Cengage Learning; 2014.