Review

Machine Learning Approaches for the Image-Based Identification of Surgical Wound Infections: Scoping Review

Juan Pablo Tabja Bortesi[1], HBSc; Jonathan Ranisau[1], MASc; Shuang Di[1,2], MSc; Michael McGillion[3], PhD; Laura Rosella[2], PhD; Alistair Johnson[4], DPhil; PJ Devereaux[3], PhD; Jeremy Petch[1,3,5,6], BA, MA, PhD

[1]Centre for Data Science and Digital Health, Hamilton Health Sciences, Hamilton, ON, Canada

[2]Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada

[3]Population Health Research Institute, Hamilton, ON, Canada

[4]SickKids Research Institute, Toronto, ON, Canada

[5]Institute for Health Policy, Management and Evaluation, University of Toronto, Toronto, ON, Canada

[6]Division of Cardiology, McMaster University, Hamilton, ON, Canada

Corresponding Author:
Jeremy Petch, BA, MA, PhD
Centre for Data Science and Digital Health
Hamilton Health Sciences
175 Longwood Road South
Suite 207
Hamilton, ON, L8P 0A1
Canada
Phone: 1 9055212100
Email: petchj@hhsc.ca

Abstract

Background: Surgical site infections (SSIs) occur frequently and impact patients and health care systems. Remote surveillance of surgical wounds is currently limited by the need for manual assessment by clinicians. Machine learning (ML)–based methods have recently been used to address various aspects of the postoperative wound healing process and may be used to improve the scalability and cost-effectiveness of remote surgical wound assessment.

Objective: The objective of this review was to provide an overview of the ML methods that have been used to identify surgical wound infections from images.

Methods: We conducted a scoping review of ML approaches for visual detection of SSIs following the JBI (Joanna Briggs Institute) methodology. Reports of participants in any postoperative context focusing on identification of surgical wound infections were included. Studies that did not address SSI identification, surgical wounds, or did not use image or video data were excluded. We searched MEDLINE, Embase, CINAHL, CENTRAL, Web of Science Core Collection, IEEE Xplore, Compendex, and arXiv for relevant studies in November 2022. The records retrieved were double screened for eligibility. A data extraction tool was used to chart the relevant data, which was described narratively and presented using tables. Employment of TRIPOD (Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis) guidelines was evaluated and PROBAST (Prediction Model Risk of Bias Assessment Tool) was used to assess risk of bias (RoB).

Results: In total, 10 of the 715 unique records screened met the eligibility criteria. In these studies, the clinical contexts and surgical procedures were diverse. All papers developed diagnostic models, though none performed external validation. Both traditional ML and deep learning methods were used to identify SSIs from mostly color images, and the volume of images used ranged from under 50 to thousands. Further, 10 TRIPOD items were reported in at least 4 studies, though 15 items were reported in fewer than 4 studies. PROBAST assessment led to 9 studies being identified as having an overall high RoB, with 1 study having overall unclear RoB.

Conclusions: Research on the image-based identification of surgical wound infections using ML remains novel, and there is a need for standardized reporting. Limitations related to variability in image capture, model building, and data sources should be addressed in the future.

Introduction

Postoperative complications are associated with significant morbidity and mortality [1,2]. Wound-related issues following surgery remain common and represent a considerable cost to patients and health care systems [3,4]. The global incidence of surgical site infections (SSIs)—which include superficial or deep infections occurring at the incision site as well as organ-space infections related to the surgery [5]—has been estimated to be 11% [6]. Many of these events occur after hospital discharge, highlighting the need for remote posthospital discharge monitoring. Early research suggests that remote postoperative wound follow-up is associated with high patient satisfaction and reduced costs [7,8].

Artificial intelligence tools have been applied to various aspects of health care and are contributing to the shift toward precision medicine [9-11]. Specifically, machine learning (ML) techniques can leverage health data and develop predictive models to assist in clinical decision-making [12], and can be used in conjunction with computer vision. An important medical task is the classification and detection of various objects, ranging from skin lesions to cell nuclei [13]. Recently, ML-enabled computer vision methods have been used to contribute to the automation of wound segmentation [14,15], evaluation of postoperative outcomes [16,17], and improvement of wound assessment practices [18,19], often outperforming existing approaches.

Wound care involves cleaning and dressing, monitoring healing, addressing possible infection, and other wound type-specific measures [20]. Current image-based wound management practices, often involving manual wound photography and assessment carried out by nurses, are time- and labor-intensive [21]. In contrast, models of care augmented with ML-enabled methods can be automated [22,23]. The portability of these methods might also be employed to conduct such assessments remotely [24], reducing patient travel burden and improving access to wound care in rural areas [25,26]. A recent clinical trial (Post-Discharge After Surgery Virtual Care With Remote Automated Monitoring-1) found that virtual care with remote monitoring that included wound evaluation shows promise in improving outcomes important to patients and to optimal health system function [27]. These results highlight the utility of digital approaches to care, which can be integrated with automated ML systems to increase scalability.

The research landscape of ML-based methods for wound surveillance is evolving rapidly. Several reviews have addressed the use of ML for various aspects of wound care from different perspectives. One scoping review focused on mapping the use cases for ML in the management of various types of chronic wounds (eg, visual assessment and predicting evolution) [28]. Another review addressed image-based chronic wound assessment from a technical standpoint, characterizing existing rule-based and ML methods for wound feature extraction and classification, as well as systems for wound imaging [29]. However, chronic and acute wounds differ in terms of the clinical signs associated with infection as those in chronic wound infections are often less discernible [30], and there is a need to establish the state of the science with respect to how ML-based tools are being used for postoperative wounds. One systematic review specifically characterized the effectiveness of ML algorithms that use textual or structured data for the detection and prediction of SSIs [31], though a survey of image-based methods has not been undertaken. Likewise, other systematic reviews have found that reporting in ML-based prediction model studies is generally poor and that most are at high risk of bias (RoB) [32,33]. Considering these results, assessments of RoB and the employment of reporting guidelines—which have not been included in previous reviews of image-based ML for wound care—can further provide insights into the current state of research in this field.

The scope and purpose of this review was to provide an in-depth overview of ML approaches that use visual data for the identification of SSIs. Specifically, this review describes the nature of the methods used in this context, the ways in which they have been validated, the extent to which the reporting of these studies follows guideline recommendations, and their RoB.

Methods

Review Methodology

This scoping review was conducted in accordance with the appropriate JBI (Joanna Briggs Institute) methodology [34]. The PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) checklist was used to guide the writing of this review [35]. We opted for a scoping review approach as we sought to analyze the methods employed in conducting research in this field, an indication for scoping reviews [36], rather than synthesize model performance.

Search Strategy and Study Selection

Following our protocol [37], participants of any age (or other demographic variable) who underwent any type of surgery were considered. The main concept being addressed is the use of ML-based computer vision in the image-based identification of surgical wound infections. Only wounds that were directly the product of surgery were included. Other types of wounds, such as pressure ulcers, were excluded. We included studies that described detection of infection of such wounds (as defined by study authors). Studies solely focusing on tasks other than identification (eg, segmentation) and using sources other than images or videos for prediction were not considered. Studies conducted in any postoperative context, including postdischarge settings, were included.

Studies that developed or validated one or more prediction models were included in this review, including those that gathered data from experimental, quasi-experimental, and observational studies (eg, randomized controlled trials, and prospective and retrospective studies). Only primary sources

were considered. Select grey literature sources, such as conference proceedings and preprints, were also considered. Animal studies were excluded.

An initial limited search of MEDLINE (Ovid) and CINAHL (EBSCO) was undertaken to identify relevant papers. Text words used in the titles and abstracts of retrieved records, as well as index terms used to describe them, were used to develop the full search strategy (Multimedia Appendix 1), which was adapted for each database. The databases we searched were MEDLINE (Ovid), CENTRAL (Ovid), Embase (Ovid), CINAHL (EBSCO), Web of Science Core Collection, IEEE Xplore, and Compendex. We also searched arXiv for relevant preprints. All databases were searched from inception to November 24, 2022. Reference lists of all included records were likewise searched for other records. Only English-language records were considered.

After the search was completed, duplicate citations were removed and all identified citations were uploaded into Rayyan [38] for title and abstract and full-text screening by 2 independent reviewers. An abstract screening tool was used to aid in the screening process (Multimedia Appendix 2). The texts of potentially relevant records were retrieved in full and assessed in the same manner. Disagreements were resolved through discussion or by consultation with an additional reviewer.

Assessment of the Employment of Reporting Guidelines and RoB

A data extraction tool (Multimedia Appendix 3)—that had been piloted with 20% (2/10) of the included reports by 2 independent reviewers—was used to abstract the relevant data. After piloting the tool, a single reviewer extracted data from the remaining sources with validation by an additional reviewer. The data were summarized using tables and presented narratively.

We determined the extent to which the included reports employed TRIPOD (Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis) guidelines using the TRIPOD adherence assessment form [39], and used the PROBAST (Prediction Model Risk of Bias Assessment Tool) to conduct critical appraisal [40]. Further, 2 reviewers assessed both employment of reporting guidelines and RoB for 20% (2/10) of the included reports; the remaining assessments were carried out by 1 reviewer (with an additional reviewer available for validation). In studies that developed multiple models, we only evaluated reporting and RoB for those that were image-based. To facilitate comparison between the reporting level of TRIPOD items, we chose arbitrary thresholds to denote high (≥70%), moderate (40%-69%), and low (1%-39%) adherence.

The TRIPOD adherence form and PROBAST were modified as needed for the purposes of this review. As has been noted in other reviews [33,41-43], it is difficult to assess RoB in the predictors of deep learning (DL) models that use images for prediction, as the image features are automatically selected by the algorithm. Still, we deemed image capture considerations important (eg, whether images were systematically captured) and altered the relevant TRIPOD and PROBAST items accordingly. The full list of modifications can be found in Multimedia Appendix 4.

Results

Study Inclusion

The search retrieved 796 records, or 699 records after duplicates were removed (Figure 1). We excluded 684 records during initial screening and full-text screened 15 reports. We identified 10 reports that met the eligibility criteria. The reference lists of these reports had an additional 16 potentially relevant records, though none met the eligibility criteria.

Figure 1. PRISMA flow diagram showing study selection process. ML: machine learning; PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses; SSI: surgical site infection.
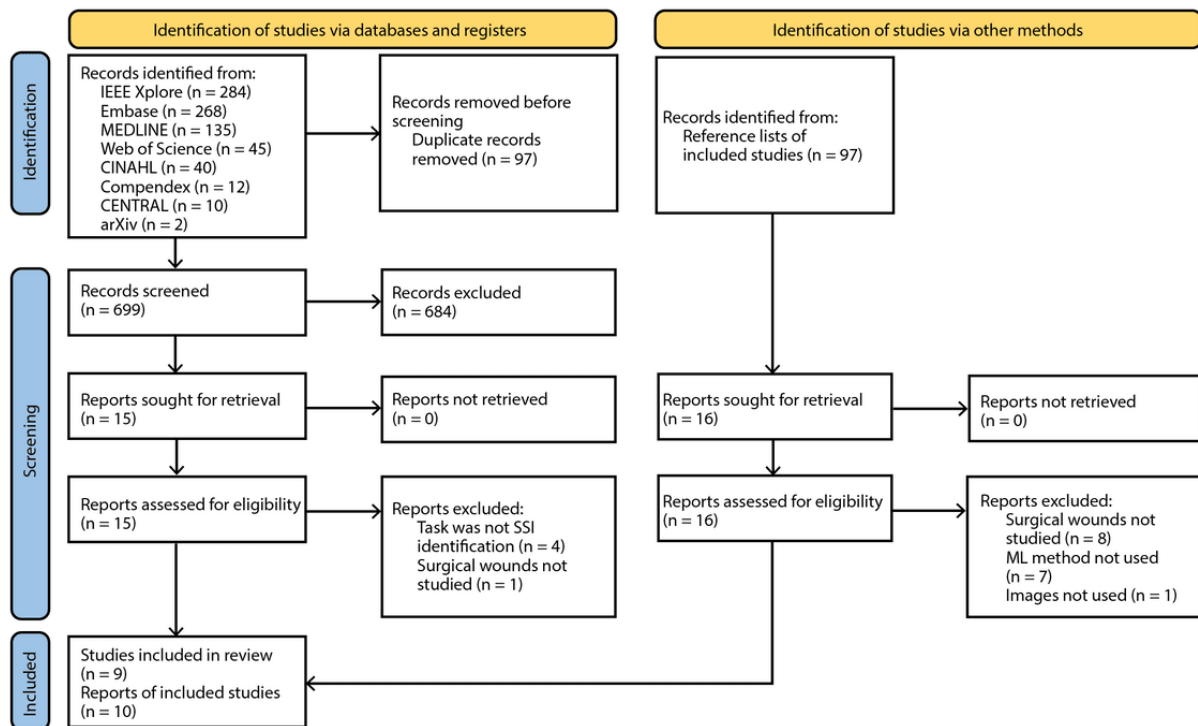
Review Findings

The included studies took place in a variety of settings, across a wide range of cohort sizes (Table 1). Important study characteristics were sometimes unclear or not reported. The full data extraction sheet can be found in Multimedia Appendix 5.

Table 1. Study characteristics.

Author	Purpose	Setting	Cohort	Events
Fletcher et al [44]	To develop a model for predicting SSI[a] in C-section wounds from thermal images taken with smartphones	Women who underwent C-section at a particular hospital in Kigali, Rwanda, between September 2019 and February 2020 prospectively enrolled on postoperative day 1	In total, 530 participants	In total, 30 participants with infected wounds
Fletcher et al [45]	To develop a model for predicting SSI in C-section wounds from color images taken with mobile devices	Women aged >18 years who underwent C-section at a particular hospital in Kigali, Rwanda, between March and October 2017 enrolled prior to discharge	In total, 572 participants (out of 729) that returned for follow-up	In total, 62 participants with infected wounds
Wu et al [46]	To develop an automatic monitoring tool for surgical wounds based on smartphone images	Prospectively collected wound database of patients who had undergone laparotomy, minimal invasive surgery, or hernia repair at an Asian academic center	In total, 480 wound images from 100 patients	In total, 136 images of infected wounds
Fletcher et al [47]	To develop models for predicting SSI in C-section wounds from questionnaire and image data	Women aged 18+ years who underwent C-section at a particular hospital in Kigali, Rwanda, between March and October 2017 enrolled prior to discharge	In total, 572 participants (out of 729) that returned for follow-up; images available for 568 patients	In total, 62 participants with infected wounds
Hsu et al [48]	To develop an automatic wound interpretation app for automated wound monitoring	Images of chest, abdomen, back, hand, and podiatry wounds collected from the Department of Surgery and Department of Internal Medicine of National Taiwan University Hospital	In total, 293 wound images	In training set, 27 infection images; total number unclear
Lüneburg et al [49]	To explore ML[b] approaches for remote LVAD[c] patient monitoring using images	Images of LVAD driveline exit sites obtained from Schüchtermann-Schiller'sche Kliniken and Hannover Medical School	In total, 745 images from 61 patients, though only 732 are labeled	In total, 212 images of mild infection and 37 images of severe infection
Shenoy et al [50]	To develop a model that can identify the onset of wound ailments from smartphone images	Images collected primarily from patients and surgeons at the Palo Alto Veterans Affairs Hospital and the Washington University Medical Center	In total, 1335 images	In total, 355 images of infection
Hsu et al [51]	To develop a model for recognizing SSI	Images collected from the Department of Surgery of National Taiwan University Hospital	In total, 42 images	In total, 30 images of infection
Zeng et al [52]	To develop a system for automatic wound detection and subsequent infection detection	Not reported	Total unclear; 6 images for testing	Unclear
Wang et al [53]	To develop an integrated system for automatic wound segmentation and analysis of wound conditions from wound images	Images collected from New York University Wound Database	In total, 3400 images	In total, 155 images of infection

[a]SSI: surgical site infection.

[b]ML: machine learning.

[c]LVAD: left ventricular assist device.

The earliest included paper was published in 2015 [53], 6 papers were published between 2017 and 2019 [47-52], and 3 papers were published between 2020 and November 2022 [44-46].

The objective of the included studies was generally to develop models for identifying surgical wound infection from images. In some cases, the purpose was broader; 2 studies sought to identify the presence of various wound attributes (eg, granulation) [46,50] and 4 studies developed models for automatic wound segmentation [48,49,51,53]. Other objectives included healing progress prediction [53], surface area estimation [53], and wound detection [52].

Patients, Procedures, and Image Capture

The types of patients and surgical procedures studied varied. In total, 3 papers focused on C-section patients in rural Rwanda [44,45,47], while another study examined patients implanted with a left ventricular assist device in Germany [49]. Further,

2 studies conducted in Asia described the surgical procedures more broadly; for instance, 1 paper included patients that had undergone laparotomy, minimal invasive surgery, or hernia repair [46], while another included surgical wounds of the chest, abdomen, back, hands, and feet [48]. In 4 papers, this information was not specified [50-53].

The context of image capture likewise varied (Table 2). Most studies simply stated that images were obtained from one or more sites or data sets [48-51,53], without further details on how the images were selected; though 1 study additionally indicated that the data were "prospectively collected" [46] and the studies conducted in Rwanda described their cohorts in the greatest detail [44,45,47].

Table 2. Study data collection and ML[a] methodology.

Author	Time of image capture	Imaging modality	Outcome determination	Modeling methods	Performance metric for best-performing model
Fletcher et al [44]	Approximately 10 days after surgery	Thermal images taken by community health workers with a thermal camera module connected to smartphone that produces a JPG thermal image and a separate 2D temperature array	Physical examination performed by general practitioner	CNN[b]	Median AUC[c]: 0.90
Fletcher et al [45]	Approximately 10 days after surgery	Color images taken by community health workers with Android tablets	Physical examination performed by general practitioner	CNN	Median AUC: 0.655
Wu et al [46]	Just after surgery, during hospitalization, and in outpatient clinic follow-up	Color images taken by surgeons with smartphones	Annotation of abnormal wound features on images performed by surgeons	CNN, SVM[d], RF[e], GB[f]	Median AUC: 0.833
Fletcher et al [47]	Approximately 10 days after surgery	Color images taken by community health workers with Android tablets	Physical examination performed by general practitioner	Unclear; potentially both SVM and logistic regression	Median AUC: 1.0
Hsu et al [48]	Not reported	Color images taken with smartphones	Unclear, but likely annotation of images by 3 physicians	SVM	Overall accuracy: 0.8358
Lüneburg et al [49]	Not reported	Color images; device not reported	Unclear, but likely based on physical examination performed by physicians	CNN	Overall accuracy: 0.670
Shenoy et al [50]	Not reported	Color images taken by patients and surgeons with smartphones	Not reported	CNN	AUC: 0.82
Hsu et al [51]	Not reported	Color images; device not reported	Not reported	SVM	Overall accuracy: 0.9523
Zeng et al [52]	Not reported	Color images; device not reported	Not reported	SVM	AUCs varied depending on the infection-related wound attribute, but ranged from 0.7682 to 0.9145.
Wang et al [53]	Not reported	Color images; device not reported	Not reported	SVM using CNN features	AUC: 0.847

[a]ML: machine learning.

[b]CNN: convolutional neural network.

[c]AUC: area under curve.

[d]SVM: support vector machine.

[e]RF: random forest.

[f]GB: gradient boosting.

In the studies conducted with C-section patients, the wounds were photographed approximately 10 days after surgery, with infection assessment taking place on the same day [44,45,47]. Another study collected images at multiple time points: immediately after surgery, during hospitalization, and at a later follow-up, though the number of days post surgery was not indicated [46]. However, the time at which the images were taken relative to surgery and the time at which infection was assessed relative to image capture were not reported in 6 records [48-53].

In terms of the images themselves, 9 studies used color images [45-53], and 1 used thermal images [44]. Further, 6 studies used a mobile device (either smartphone or tablet) to capture the image [44-48,50], while others did not report the device used [49,51-53]. Across studies that reported the persons responsible for capturing the images, community health workers were typically responsible [44,45,47]; 1 study used images taken by surgeons [46]; and another used images collected by both patients and surgeons [50].

Assessment of surgical wound infection establishes the model ground truth and occurred mainly through face-to-face physical examination [44,45,47,49], through manual annotation of the wound images [46,48,51], or was not reported [50,52,53].

ML Approaches

All the included records were model development studies (ie, no external validation). In total, 4 papers used convolutional neural networks (CNNs) [44,45,49,50], 3 used support vector machines (SVMs) developed using handcrafted features [48,51,52], 1 trained an SVM classifier using CNN-derived features [53], 1 used a CNN, an SVM, a random forest model, and a gradient boosting classifier [46], and 1 paper's methods were not entirely clear but may have involved both logistic regression and SVMs [47]. Additional technical details are available in Multimedia Appendix 6.

The number of images used for developing an infection detection model ranged from just 42 [51] to 3400 [53]. Likewise, the proportion of images of infected wounds ranged from 4.6% (155/3400) [53] to 71.4% (30/42) [51]. In some cases, there was 1 image per patient [44,45,47], while in others, there were multiple per patient [46,49] or the number of patients was not reported [48,50-53].

In 5 papers, the classification task was binary [44-47,53], while in most others, the task was multiclass. In 1 paper, multiclass classification entailed distinguishing between mild, severe, and no infection [49], while in 3 others, the model differentiated between various infection-related wound attributes, such as granulation and swelling [48,51,52]. In contrast, 1 paper addressed a multilabel task in which the model identified the presence of a wound, infection, granulation, and drainage per image [50].

All studies reported model performance. In total, 7 studies reported area under the receiver operating characteristic curve values, which ranged from 0.655 [45] to 1.0 [47] for the best-performing models. The remaining studies reported overall accuracies, ranging from 0.670 [49] to 0.952 [51] for the best-performing models, as well as other performance metrics (eg, F_1-scores).

Employment of Reporting Guidelines

There were a few TRIPOD items that were highly employed (ie, employed by at least 7 out of the 10 included studies). For instance, all papers reported their objectives, and most reported background information, overall interpretations of study results,

and descriptions of whether actions were taken to standardize image capture or otherwise systematically identify wounds from the images. In addition, 6 TRIPOD items had moderate employment (employed by between 4 and 6 studies); namely, the reporting of data sources and study setting, descriptions of model-building procedures, the number of participants or images (and the number showing infection), study limitations, as well as the potential clinical use of the models and future research directions.

Employment of 8 TRIPOD items was low (employed by between 1 and 3 studies), including items related to the reporting of participant selection methods, descriptions of how and when images were taken, rationales for sample sizes, the flow of participants within the paper, explanations of how to use the models, and funding details. Most studies did not completely use these guidelines in terms of outcome assessment, as there was often no indication of the criteria used to diagnose surgical wound infection or the time interval between surgery and assessment was unclear.

An additional 7 TRIPOD items were not reported in any of the included studies. Titles and abstracts did not employ reporting guidelines, and participant demographics were not reported. Similarly, model calibration was not discussed, and in studies that did not exclusively use DL methods for infection detection [46-48,51-53], reporting of feature modeling details did not meet TRIPOD guidelines.

About RoB

The RoB assessment led to 9 studies being identified as having an overall high RoB, while the remaining study was determined to have overall unclear RoB (Table 3). The participants domain was determined to be unclear in terms of RoB because little information about the source of data and recruitment methods was reported [46,48-53]. The 3 papers on C-section patients in Rwanda were at low RoB for this domain, as the nature of these works was cross-sectional and the cohorts were well defined [44,45,47]. In terms of predictors, we identified 5 papers as being at high RoB since there was variability in image capture conditions without later accounting for this variability [45-47,50,53]. In contrast, other papers were judged to be at low RoB for this domain because they segmented the wound prior to infection detection [48,49,51] or placed a frame around the wound prior to image capture [44], improving the uniformity of images processed for model training. Likewise, most studies were rated as having unclear RoB in the outcome domain, largely because the specific criteria used to gauge the presence of surgical wound infection were not reported. In other cases, the outcome domain was determined to be at high RoB because the presence of infection was determined solely from images, as opposed to by face-to-face review. In 8 studies, the analysis domain was assessed as being at high RoB for many reasons [45,47-53], including omission of participants in model development, an absence of both discrimination and calibration measures, and failure to appropriately account for overfitting.

Table 3. PROBAST[a] RoB[b] assessment of the 10 included reports.

Study	RoB				
	Participants	Predictors	Outcome	Analysis	Overall
Fletcher et al [44]	+[c]	+	?[d]	?	?
Fletcher et al [45]	+	_[e]	?	–	–
Wu et al [46]	?	–	–	?	–
Fletcher et al [47]	+	–	?	–	–
Hsu et al [48]	?	+	–	–	–
Lüneburg et al [49]	?	+	?	–	–
Shenoy et al [50]	?	–	?	–	–
Hsu et al [51]	?	+	–	–	–
Zeng et al [52]	?	?	?	–	–
Wang et al [53]	?	–	?	–	–

[a]PROBAST: Prediction Model Risk of Bias Assessment Tool.

[b]RoB: risk of bias.

[c]+ indicates low RoB.

[d]? indicates unclear RoB.

[e]– indicates high RoB.

Discussion

Principal Findings

This scoping review aimed to characterize the available research on ML approaches for the image-based identification of surgical wound infections. Such research is important as it can be integrated with remote patient monitoring, which enables improved health care decision-making and management, with additional benefits such as reduced travel burden. Initial work has suggested that remote image-based monitoring of wounds is feasible and associated with higher patient satisfaction [54-56], and is at least comparable to routine in-person care in terms of time to infection diagnosis [57]. Other aspects of wound assessment targeted by image-based remote patient monitoring include identification of dehiscence and surface area and temperature measurements [58,59], though much has not been automated or ML-based.

Despite the extensive body of ML-based work using medical images in other specialties [60,61], there is scarce ML research on the identification of surgical wound infections from digital images. We identified only 10 such papers, 7 of which were conference papers, which limits the space for reporting and likely contributed to the low reporting of TRIPOD items. In contrast, a recent review of ML for SSI detection identified 32 papers that used structured electronic health record, free-text, or administrative data for prediction [31], suggesting that ML-based SSI detection research has mostly used these more readily available data sources. While models based on such in-hospital data perform well in the context of inpatient SSI detection, they may be limited in their practical application during clinical care, as visual inspection is the essential mode by which infection is identified. In terms of incorporating innovative imaging techniques, thermal imaging has recently emerged as a potentially valuable tool in the management of surgical wounds [62-64]. Thermal imaging can be used with mobile devices [44,65], which facilitates its application for postdischarge monitoring, and may better generalize to different skin colors. On the other hand, the utility of electronic health record– or text-based models for postdischarge surveillance is perhaps less clear. Current postdischarge surgical wound surveillance largely depends on evaluation at follow-up visits, which may be infrequent and not timely [66], or on patient self-assessment, which is not reliable [67,68]. ML for the image-based identification of surgical wound infections presents the opportunity to automate this practice.

Reporting Data Collection Details

ML hinges on effective data collection, which can be challenging in outpatient or remote monitoring settings; hence, this type of research is still in early development. Although virtual care as a model of health care is relatively new, progress has been made in terms of data collection technology [64,69], similar telemedicine research without ML [70-72], and monitoring of other wound types [73,74]. As almost three-fourths of individuals worldwide own a mobile phone [75], leveraging this technology for remote monitoring holds potential. Still, it is worth noting that mobile phone ownership and mobile network coverage is lower in certain geographical areas and in low-income groups. In these contexts, alternative approaches, such as in-hospital follow-up with pictures taken by a community health worker [44], may be more appropriate. In terms of the data used in the included studies, it has mainly been collected in non-Western settings, and there are no publicly available data sets of surgical wound infection images, which presents a challenge to reproducibility and further development in the field. Likewise, the lack of reporting on image metadata (eg, gender and age distributions, procedures received, and occurrence of surgical complications) and eligibility criteria

XSL•FO

RenderX

limits the understanding of the populations that this research can be generalized to and contributes to RoB in terms of participant selection. Reporting of such details needs improvement for the progression of different prototypes for different subpopulations in this domain.

Transparency and Standardization in Model Development

The nature of the models developed in the included studies was diagnostic rather than prognostic. Similarly, none of the included papers performed out-of-sample external validation, highlighting the newness of this field of work and opportunity for further maturity. Interestingly, 4 papers published between 2017 and 2019 did not use DL methods, perhaps because the expertise required for development of such models was not yet widely available. Model performance is likewise not well-standardized in its reporting, as no papers reported on calibration and some did not include discrimination measures, which gives rise to RoB in analysis methods. Many papers did not report on measures to address overfitting, which calls the developed models' generalizability into question. Despite the partial reporting, the performance of the models in the included papers suggest that image-based ML for identification of surgical wound infection holds promise. In order to better understand their generalizability and reliability, future studies should externally validate and calibrate the developed models and report areas under the curve (as opposed to solely reporting other measures such as accuracy), and provide transparent documentation (eg, open-source code) to promote reproducibility and collaboration. Considering that interpretability and explainability support clinician trust [76], researchers may likewise wish to explore these concepts in future work.

Employment of Reporting Guidelines and RoB

Standardization of Image Capture

Use of TRIPOD guidelines was mostly low and RoB was generally unclear or high. This was in part due to participant- and analysis-related considerations discussed above; however, there were also concerns with the images themselves. In most studies, the way in which the images were taken, environmental conditions, the persons responsible for taking them, and the time of image capture relative to surgery, were not reported in detail. Still, there was often variability in the conditions of image capture, which might be attributed to unique challenges associated with collection and standardization in this particular modality. As opposed to other modalities, surgical wound infection images are largely taken by hand, without explicit training or guidance, which makes for considerable differences among images and introduces RoB in terms of model predictors. Efforts to standardize image capture help reduce RoB by minimizing systematic differences between images of infected versus noninfected wounds. Recent approaches such as instructions for patient-generated surgical wound images [77] or automated color calibration, scaling, and rotation correction [78] suggest that these considerations are receiving attention. Some studies created segmentation algorithms to capture the wound more reliably from the nonuniform images, which may have hindered the development of infection detection models. Segmentation and classification represent distinct areas of research, though many studies developed their own segmentation models rather than using or building on existing segmentation algorithms. In future work, specific directions detailing the time (relative to surgery), method, and conditions of image capture should be provided in order to reduce unwanted variability, and image processing steps can be undertaken for further standardization.

Transparency in Outcome Assessment

Outcome assessment was also not well reported in most papers. While there is no universally accepted and objective gold standard for SSI detection [79], clinical examination (involving direct observation of the wound) is frequently used as a reference standard [68,70,72,80,81]. Although some studies did perform in-person clinical examination, none reported the specific criteria used to gauge the presence of infection. Considering that there are differences in the rates of reported SSIs depending on the criteria used [82], specifying these criteria is important to more accurately assess the RoB arising from outcome assessment. It is worth noting, however, that there are challenges associated with in-person postoperative wound assessment. Surgical wound infections progress variably, with some only apparent after the 30-day postoperative surveillance benchmark [83,84]. However, extended in-person follow-up timeframes may require additional administrative resources. In practice, the criteria employed for SSI assessment typically consider both feasibility and validity [79]. This may necessitate striking a balance between resources, time constraints, and quality of assessment, which can pose challenges to the comprehensive evaluation of surgical wound infections. On a smaller scale, interrater reliability of in-person SSI assessment using established criteria can be modest [85,86], and in rural areas, there may be limited access to high-quality in-person wound care. Where feasible, determination of ground truth should use established criteria for infection and employ multiple independent assessors to minimize RoB.

Limitations

There are some limitations to this review. For instance, additional searching (eg, forward citation searching) may have led to more relevant reports being identified, as may have searching grey literature sources, which would reduce selection bias. We may have missed other relevant non–English-language papers, potentially excluding valuable studies. The included studies are from diverse locations (eg, Rwanda, Germany, and Taiwan), though this does not fully compensate for the potential language bias. Similarly, data extraction and the TRIPOD and PROBAST assessments were mainly completed by 1 reviewer, which introduces a potential source of bias in our findings. The modifications made to the TRIPOD and PROBAST tools may limit the ability to compare the results of our assessments to those of other reviews. Artificial intelligence–oriented extensions of both tools are in development [87] and will facilitate their use in appraising ML-based studies.

Conclusions

The use of ML for the image-based identification of surgical wound infections remains in the early stages, with only 10 studies available and a need for reporting standardization. Future

development and validation of such models should carefully consider image variability, overfitting concerns, and criteria for determination of infection. These considerations are important to advance the state of image-based ML for wound management, which has the potential to automate traditionally labor-intensive practices.

Acknowledgments

Conflicts of Interest

JP declares research funding support from Roche Canada. PJD has received grants from Abbott Diagnostics, Siemens Canada, and Roche Diagnostics; received consulting fees from Abbott Laboratories, Renibus Therapeutics, Roche Diagnostics, and Trimedic Canada; received monitoring devices from CloudDX and Philips Healthcare; participated in advisory board meetings for Bayer AG; and is a member of the Data Safety Monitoring Board for the PEPPER (Comparative Effectiveness of Pulmonary Embolism Prevention After Hip and Knee Replacement) Study, New Hampshire. The other authors have no conflicts of interest to disclose.

Multimedia Appendix 1

Search strategy.
[DOCX File , 22 KB-Multimedia Appendix 1]

Multimedia Appendix 2

Abstract screening tool.
[DOCX File , 14 KB-Multimedia Appendix 2]

Multimedia Appendix 3

Data extraction tool.
[DOCX File , 14 KB-Multimedia Appendix 3]

Multimedia Appendix 4

TRIPOD and PROBAST altered or excluded items.
[DOCX File , 17 KB-Multimedia Appendix 4]

Multimedia Appendix 5

Complete results of data extraction, assessment of TRIPOD employment, and PROBAST assessment.
[DOCX File , 70 KB-Multimedia Appendix 5]

Multimedia Appendix 6

Additional technical details.
[DOCX File , 28 KB-Multimedia Appendix 6]

Multimedia Appendix 7

PRISMA-ScR checklist.
[PDF File (Adobe PDF File), 101 KB-Multimedia Appendix 7]

References

1. Tevis SE, Kennedy GD. Postoperative complications and implications on patient-centered outcomes. J Surg Res. 2013;181(1):106-113. [FREE Full text] [doi: 10.1016/j.jss.2013.01.032] [Medline: 23465392]
2. Endo I, Kumamoto T, Matsuyama R. Postoperative complications and mortality: are they unavoidable? Ann Gastroenterol Surg. 2017;1(3):160-163. [FREE Full text] [doi: 10.1002/ags3.12045] [Medline: 29863116]
3. Monahan M, Jowett S, Pinkney T, Brocklehurst P, Morton DG, Abdali Z, et al. Surgical site infection and costs in low- and middle-income countries: a systematic review of the economic burden. PLoS One. 2020;15(6):e0232960. [FREE Full text] [doi: 10.1371/journal.pone.0232960] [Medline: 32497086]

4. Guest JF, Fuller GW, Vowden P. Costs and outcomes in evaluating management of unhealed surgical wounds in the community in clinical practice in the UK: a cohort study. BMJ Open. 2018;8(12):e022591. [FREE Full text] [doi: 10.1136/bmjopen-2018-022591] [Medline: 30552253]

5. Onyekwelu I, Yakkanti R, Protzer L, Pinkston CM, Tucker C, Seligson D. Surgical wound classification and surgical site infections in the orthopaedic patient. J Am Acad Orthop Surg Glob Res Rev. 2017;1(3):e022. [FREE Full text] [doi: 10.5435/JAAOSGlobal-D-17-00022] [Medline: 30211353]

6. Gillespie BM, Harbeck E, Rattray M, Liang R, Walker R, Latimer S, et al. Worldwide incidence of surgical site infections in general surgical patients: a systematic review and meta-analysis of 488,594 patients. Int J Surg. 2021;95:106136. [FREE Full text] [doi: 10.1016/j.ijsu.2021.106136] [Medline: 34655800]

7. Wang SC, Au Y, Ramirez-GarciaLuna JL, Lee L, Berry GK. The promise of smartphone applications in the remote monitoring of postsurgical wounds: a literature review. Adv Skin Wound Care. 2020;33(9):489-496. [doi: 10.1097/01.ASW.0000694136.29135.02] [Medline: 32810062]

8. Gunter RL, Chouinard S, Fernandes-Taylor S, Wiseman JT, Clarkson S, Bennett K, et al. Current use of telemedicine for post-discharge surgical care: a systematic review. J Am Coll Surg. 2016;222(5):915-927. [FREE Full text] [doi: 10.1016/j.jamcollsurg.2016.01.062] [Medline: 27016900]

9. Mesko B. The role of artificial intelligence in precision medicine. Expert Rev Precis Med Drug Dev. 2017;2(5):239-241. [FREE Full text] [doi: 10.1080/23808993.2017.1380516]

10. Johnson KB, Wei WQ, Weeraratne D, Frisse ME, Misulis K, Rhee K, et al. Precision medicine, AI, and the future of personalized health care. Clin Transl Sci. 2021;14(1):86-93. [FREE Full text] [doi: 10.1111/cts.12884] [Medline: 32961010]

11. Amisha; Malik P, Pathania M, Rathaur VK. Overview of artificial intelligence in medicine. J Family Med Prim Care. 2019;8(7):2328-2331. [FREE Full text] [doi: 10.4103/jfmpc.jfmpc_440_19] [Medline: 31463251]

12. Ghassemi M, Naumann T, Schulam P, Beam AL, Chen IY, Ranganath R. A review of challenges and opportunities in machine learning for health. AMIA Jt Summits Transl Sci Proc. ;2020:191-200. [FREE Full text] [Medline: 32477638]

13. Esteva A, Chou K, Yeung S, Naik N, Madani A, Mottaghi A, et al. Deep learning-enabled medical computer vision. npj Digit Med. 2021;4(1):5. [FREE Full text] [doi: 10.1038/s41746-020-00376-2] [Medline: 33420381]

14. Carrión H, Jafari M, Bagood MD, Yang HY, Isseroff RR, Gomez M. Automatic wound detection and size estimation using deep learning algorithms. PLoS Comput Biol. 2022;18(3):e1009852. [FREE Full text] [doi: 10.1371/journal.pcbi.1009852] [Medline: 35275923]

15. Wang C, Anisuzzaman DM, Williamson V, Dhar MK, Rostami B, Niezgoda J, et al. Fully automatic wound segmentation with deep convolutional neural networks. Sci Rep. 2020;10(1):21897. [FREE Full text] [doi: 10.1038/s41598-020-78799-w] [Medline: 33318503]

16. Sikka K, Ahmed AA, Diaz D, Goodwin MS, Craig KD, Bartlett MS, et al. Automated assessment of children's postoperative pain using computer vision. Pediatrics. 2015;136(1):e124-e131. [FREE Full text] [doi: 10.1542/peds.2015-0029] [Medline: 26034245]

17. Şimşek İ, Şirolu C. Analysis of surgical outcome after upper eyelid surgery by computer vision algorithm using face and facial landmark detection. Graefes Arch Clin Exp Ophthalmol. 2021;259(10):3119-3125. [doi: 10.1007/s00417-021-05219-8] [Medline: 33963919]

18. Barakat-Johnson M, Jones A, Burger M, Leong T, Frotjold A, Randall S, et al. Reshaping wound care: evaluation of an artificial intelligence app to improve wound assessment and management amid the COVID-19 pandemic. Int Wound J. 2022;19(6):1561-1577. [FREE Full text] [doi: 10.1111/iwj.13755] [Medline: 35212459]

19. Mukherjee R, Manohar DD, Das DK, Achar A, Mitra A, Chakraborty C. Automated tissue classification framework for reproducible chronic wound assessment. Biomed Res Int. ;2014:851582. [FREE Full text] [doi: 10.1155/2014/851582] [Medline: 25114925]

20. Dreifke MB, Jayasuriya AA, Jayasuriya AC. Current wound healing procedures and potential care. Mater Sci Eng C. 2015;48:651-662. [FREE Full text] [doi: 10.1016/j.msec.2014.12.068] [Medline: 25579968]

21. Aldaz G, Shluzas LA, Pickham D, Eris O, Sadler J, Joshi S, et al. Hands-free image capture, data tagging and transfer using Google Glass: a pilot study for improved wound care management. PLoS One. 2015;10(4):e0121179. [FREE Full text] [doi: 10.1371/journal.pone.0121179] [Medline: 25902061]

22. Zahia S, Garcia-Zapirain B, Elmaghraby A. Integrating 3D model representation for an accurate non-invasive assessment of pressure injuries with deep learning. Sensors (Basel). 2020;20(10):2933. [FREE Full text] [doi: 10.3390/s20102933] [Medline: 32455753]

23. Lau CH, Yu KHO, Yip TF, Luk LYF, Wai AKC, Sit TY, et al. An artificial intelligence-enabled smartphone app for real-time pressure injury assessment. Front Med Technol. 2022;4:905074. [FREE Full text] [doi: 10.3389/fmedt.2022.905074] [Medline: 36212608]

24. Zoppo G, Marrone F, Pittarello M, Farina M, Uberti A, Demarchi D, et al. AI technology for remote clinical assessment and monitoring. J Wound Care. 2020;29(12):692-706. [FREE Full text] [doi: 10.12968/jowc.2020.29.12.692] [Medline: 33320742]

25. Sood A, Granick MS, Trial C, Lano J, Palmier S, Ribal E, et al. The role of telemedicine in wound care: a review and analysis of a database of 5,795 patients from a mobile wound-healing center in Languedoc-Roussillon, France. Plast Reconstr Surg. 2016;138(Suppl 3):248S-256S. [doi: 10.1097/PRS.0000000000002702] [Medline: 27556769]

26. Zhang J, Mihai C, Tüshaus L, Scebba G, Distler O, Karlen W. Wound image quality from a mobile health tool for home-based chronic wound management with real-time quality feedback: randomized feasibility study. JMIR mHealth uHealth. 2021;9(7):e26149. [FREE Full text] [doi: 10.2196/26149] [Medline: 34328440]

27. McGillion MH, Parlow J, Borges FK, Marcucci M, Jacka M, Adili A, et al. Post-discharge after surgery Virtual Care with Remote Automated Monitoring-1 (PVC-RAM-1) technology versus standard care: randomised controlled trial. BMJ. 2021;374:n2209. [FREE Full text] [doi: 10.1136/bmj.n2209] [Medline: 34593374]

28. Dabas M, Schwartz D, Beeckman D, Gefen A. Application of artificial intelligence methodologies to chronic wound care and management: a scoping review. Adv Wound Care (New Rochelle). 2023;12(4):205-240. [doi: 10.1089/wound.2021.0144] [Medline: 35438547]

29. Anisuzzaman DM, Wang C, Rostami B, Gopalakrishnan S, Niezgoda J, Yu Z. Image-based artificial intelligence in wound assessment: a systematic review. Adv Wound Care (New Rochelle). 2022;11(12):687-709. [doi: 10.1089/wound.2021.0091] [Medline: 34544270]

30. Hurlow J, Bowler PG. Acute and chronic wound infections: microbiological, immunological, clinical and therapeutic distinctions. J Wound Care. 2022;31(5):436-445. [FREE Full text] [doi: 10.12968/jowc.2022.31.5.436] [Medline: 35579319]

31. Wu G, Khair S, Yang F, Cheligeer C, Southern D, Zhang Z, et al. Performance of machine learning algorithms for surgical site infection case detection and prediction: a systematic review and meta-analysis. Ann Med Surg (Lond). 2022;84:104956. [FREE Full text] [doi: 10.1016/j.amsu.2022.104956] [Medline: 36582918]

32. Navarro CLA, Damen JAA, Takada T, Nijman SWJ, Dhiman P, Ma J, et al. Risk of bias in studies on prediction models developed using supervised machine learning techniques: systematic review. BMJ. 2021;375:n2281. [FREE Full text] [doi: 10.1136/bmj.n2281] [Medline: 34670780]

33. Nagendran M, Chen Y, Lovejoy CA, Gordon AC, Komorowski M, Harvey H, et al. Artificial intelligence versus clinicians: systematic review of design, reporting standards, and claims of deep learning studies. BMJ. 2020;368:m689. [FREE Full text] [doi: 10.1136/bmj.m689] [Medline: 32213531]

34. Peters MDJ, Godfrey C, McInerney P, Munn Z, Trico AC, Khalil H. Chapter 11: scoping reviews. In: Aromataris E, Munn Z, editors. JBI Manual for Evidence Synthesis. Adelaide, Australia. Joanna Briggs Institute; 2020.

35. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA extension for Scoping Reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med. 2018;169(7):467-473. [FREE Full text] [doi: 10.7326/M18-0850] [Medline: 30178033]

36. Munn Z, Peters MDJ, Stern C, Tufanaru C, McArthur A, Aromataris E. Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. BMC Med Res Methodol. 2018;18(1):143. [FREE Full text] [doi: 10.1186/s12874-018-0611-x] [Medline: 30453902]

37. Machine learning approaches for computer vision tasks related to the identification of surgical site infections: a scoping review protocol. OSF Registries. URL: https://osf.io/3k9xq [accessed 2023-03-23]

38. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. Syst Rev. 2016;5(1):210. [FREE Full text] [doi: 10.1186/s13643-016-0384-4] [Medline: 27919275]

39. Heus P, Damen JAAG, Pajouheshnia R, Scholten RJPM, Reitsma JB, Collins GS, et al. Uniformity in measuring adherence to reporting guidelines: the example of TRIPOD for assessing completeness of reporting of prediction model studies. BMJ Open. 2019;9(4):e025611. [FREE Full text] [doi: 10.1136/bmjopen-2018-025611] [Medline: 31023756]

40. Wolff RF, Moons KGM, Riley RD, Whiting PF, Westwood M, Collins GS, et al. PROBAST: a tool to assess the risk of bias and applicability of prediction model studies. Ann Intern Med. 2019;170(1):51-58. [FREE Full text] [doi: 10.7326/M18-1376] [Medline: 30596875]

41. Roberts M, Driggs D, Thorpe M, Gilbey J, Yeung M, Ursprung S, et al. Common pitfalls and recommendations for using machine learning to detect and prognosticate for COVID-19 using chest radiographs and CT scans. Nat Mach Intell. 2021;3(3):199-217. [FREE Full text] [doi: 10.1038/s42256-021-00307-0]

42. Wynants L, Van Calster B, Collins GS, Riley RD, Heinze G, Schuit E, et al. Prediction models for diagnosis and prognosis of COVID-19: systematic review and critical appraisal. BMJ. 2020;369:m1328. [FREE Full text] [doi: 10.1136/bmj.m1328] [Medline: 32265220]

43. Frizzell TO, Glashutter M, Liu CC, Zeng A, Pan D, Hajra SG, et al. Artificial intelligence in brain MRI analysis of Alzheimer's disease over the past 12 years: a systematic review. Ageing Res Rev. 2022;77:101614. [doi: 10.1016/j.arr.2022.101614] [Medline: 35358720]

44. Fletcher RR, Schneider G, Bikorimana L, Rukundo G, Niyigena A, Miranda E, et al. The use of mobile thermal imaging and deep learning for prediction of surgical site infection. In: Annu Int Conf IEEE Eng Med Biol Soc. Presented at: 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC); Mexico, 2021;5059-5062; November 1-5, 2021. URL: https://ieeexplore.ieee.org/document/9630094 [doi: 10.1109/EMBC46164.2021.9630094]

45. Fletcher RR, Schneider G, Hedt-Gauthier B, Nkurunziza T, Alayande B, Riviello R, et al. Use of convolutional neural nets and transfer learning for prediction of surgical site infection from color images. In: Annu Int Conf IEEE Eng Med Biol

Soc. Presented at: 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC); November 1-5, 2021, 2021;5047-5050; Mexico. [doi: 10.1109/EMBC46164.2021.9630430]

46. Wu JM, Tsai CJ, Ho TW, Lai F, Tai HC, Lin MT. A unified framework for automatic detection of wound infection with artificial intelligence. Appl Sci. 2020;10(15):5353. [FREE Full text] [doi: 10.3390/app10155353]

47. Fletcher RR, Olubeko O, Sonthalia H, Kateera F, Nkurunziza T, Ashby JL, et al. Application of machine learning to prediction of surgical site infection. In: Annu Int Conf IEEE Eng Med Biol Soc. Presented at: 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC); July 23-27, 2019, 2019;2234-2237; Berlin, Germany. [doi: 10.1109/EMBC.2019.8857942]

48. Hsu JT, Chen YW, Ho TW, Tai HC, Wu JM, Sun HY, et al. Chronic wound assessment and infection detection method. BMC Med Inform Decis Mak. 2019;19(1):99. [FREE Full text] [doi: 10.1186/s12911-019-0813-0] [Medline: 31126274]

49. Lüneburg N, Reiss N, Feldmann C, van der Meulen P, van de Steeg M, Schmidt T, et al. Photographic LVAD driveline wound infection recognition using deep learning. Stud Health Technol Inform. 2019;260:192-199. [doi: 10.3233/978-1-61499-971-3-192] [Medline: 31118337]

50. Shenoy VN, Foster E, Aalami L, Majeed B, Aalami O. Deepwound: automated postoperative wound assessment and surgical site surveillance through convolutional neural networks. In: IEEE Int Conf Bioinform Biomed. Presented at: 2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM); December 3-6, 2018, 2018;1017-1021; Madrid, Spain. [doi: 10.1109/bibm.2018.8621130]

51. Hsu JT, Ho TW, Shih HF, Chang CC, Lai F, Wu JM. Automatic wound infection interpretation for postoperative wound image. Presented at: Proceedings SPIE 10225, Eighth International Conference on Graphic and Image Processing (ICGIP 2016); February 8, 2017, 2017;1022526; Tokyo, Japan. [doi: 10.1117/12.2266110]

52. Zeng YC, Liao KH, Wang CH, Lin Y, Chang WT. Implementation of post-operative wound analytics. In: IEEE Int Conf Consum Electron-Taiwan. Presented at: 2017 IEEE International Conference on Consumer Electronics—Taiwan (ICCE-TW); June 12-14, 2017, 2017;93-94; Taipei, Taiwan. [doi: 10.1109/icce-china.2017.7991011]

53. Wang C, Yan X, Smith M, Kochhar K, Rubin M, Warren SM, et al. A unified framework for automatic wound segmentation and analysis with deep convolutional neural networks. In: Annu Int Conf IEEE Eng Med Biol Soc. Presented at: 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC); August 25-29, 2015, 2015;2415-2418; Milan, Italy. [doi: 10.1109/EMBC.2015.7318881]

54. Gunter RL, Fernandes-Taylor S, Rahman S, Awoyinka L, Bennett KM, Weber SM, et al. Feasibility of an image-based mobile health protocol for postoperative wound monitoring. J Am Coll Surg. 2018;226(3):277-286. [FREE Full text] [doi: 10.1016/j.jamcollsurg.2017.12.013] [Medline: 29366555]

55. Gunter R, Fernandes-Taylor S, Mahnke A, Awoyinka L, Schroeder C, Wiseman J, et al. Evaluating patient usability of an image-based mobile health platform for postoperative wound monitoring. JMIR mHealth uHealth. 2016;4(3):e113. [FREE Full text] [doi: 10.2196/mhealth.6023] [Medline: 27683059]

56. Mousa AY, Broce M, Monnett S, Davis E, McKee B, Lucas BD. Results of telehealth electronic monitoring for post discharge complications and surgical site infections following arterial revascularization with groin incision. Ann Vasc Surg. 2019;57:160-169. [doi: 10.1016/j.avsg.2018.09.023] [Medline: 30500646]

57. McLean KA, Mountain KE, Shaw CA, Drake TM, Pius R, Knight SR, et al. Remote diagnosis of surgical-site infection using a mobile digital intervention: a randomised controlled trial in emergency surgery patients. npj Digit Med. 2021;4(1):160. [FREE Full text] [doi: 10.1038/s41746-021-00526-0] [Medline: 34795398]

58. Wang SC, Anderson JAE, Evans R, Woo K, Beland B, Sasseville D, et al. Point-of-care wound visioning technology: reproducibility and accuracy of a wound measurement app. PLoS One. 2017;12(8):e0183139. [FREE Full text] [doi: 10.1371/journal.pone.0183139] [Medline: 28817649]

59. Wiseman JT, Fernandes-Taylor S, Gunter R, Barnes ML, Saunders RS, Rathouz PJ, et al. Inter-rater agreement and checklist validation for postoperative wound assessment using smartphone images in vascular surgery. J Vasc Surg Venous Lymphat Disord. 2016;4(3):320-328.e2. [FREE Full text] [doi: 10.1016/j.jvsv.2016.02.001] [Medline: 27318052]

60. Aggarwal R, Sounderajah V, Martin G, Ting DSW, Karthikesalingam A, King D, et al. Diagnostic accuracy of deep learning in medical imaging: a systematic review and meta-analysis. npj Digit Med. 2021;4(1):65. [FREE Full text] [doi: 10.1038/s41746-021-00438-z] [Medline: 33828217]

61. Liu X, Faes L, Kale AU, Wagner SK, Fu DJ, Bruynseels A, et al. A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis. Lancet Digit Health. 2019;1(6):e271-e297. [FREE Full text] [doi: 10.1016/S2589-7500(19)30123-2] [Medline: 33323251]

62. Childs C, Wright N, Willmott J, Davies M, Kilner K, Ousey K, et al. The surgical wound in infrared: thermographic profiles and early stage test-accuracy to predict surgical site infection in obese women during the first 30 days after caesarean section. Antimicrob Resist Infect Control. 2019;8:7. [FREE Full text] [doi: 10.1186/s13756-018-0461-7] [Medline: 30637101]

63. Siah CJR, Childs C, Chia CK, Cheng KFK. An observational study of temperature and thermal images of surgical wounds for detecting delayed wound healing within four days after surgery. J Clin Nurs. 2019;28(11-12):2285-2295. [doi: 10.1111/jocn.14832] [Medline: 30791157]

64. Li F, Wang M, Wang T, Wang X, Ma X, He H, et al. Smartphone-based infrared thermography to assess progress in thoracic surgical incision healing: a preliminary study. Int Wound J. 2023;20(6):2000-2009. [FREE Full text] [doi: 10.1111/iwj.14063] [Medline: 36542868]

65. Kanazawa T, Nakagami G, Goto T, Noguchi H, Oe M, Miyagaki T, et al. Use of smartphone attached mobile thermography assessing subclinical inflammation: a pilot study. J Wound Care. 2016;25(4):177-182. [doi: 10.12968/jowc.2016.25.4.177] [Medline: 27064366]

66. Saunders RS, Fernandes-Taylor S, Rathouz PJ, Saha S, Wiseman JT, Havlena J, et al. Outpatient follow-up versus 30-day readmission among general and vascular surgery patients: a case for redesigning transitional care. Surgery. 2014;156(4):949-958. [FREE Full text] [doi: 10.1016/j.surg.2014.06.041] [Medline: 25239351]

67. Whitby M, McLaws ML, Doidge S, Collopy B. Post-discharge surgical site surveillance: does patient education improve reliability of diagnosis? J Hosp Infect. 2007;66(3):237-242. [doi: 10.1016/j.jhin.2007.04.014] [Medline: 17582652]

68. Richter V, Cohen MJ, Benenson S, Almogy G, Brezis M. Patient self-assessment of surgical site infection is inaccurate. World J Surg. 2017;41(8):1935-1942. [doi: 10.1007/s00268-017-3974-y] [Medline: 28271262]

69. Maddah E, Beigzadeh B. Use of a smartphone thermometer to monitor thermal conductivity changes in diabetic foot ulcers: a pilot study. J Wound Care. 2020;29(1):61-66. [doi: 10.12968/jowc.2020.29.1.61] [Medline: 31930943]

70. Nkurunziza T, Williams W, Kateera F, Riviello R, Niyigena A, Miranda E, et al. mHealth-community health worker telemedicine intervention for surgical site infection diagnosis: a prospective study among women delivering via caesarean section in rural Rwanda. BMJ Glob Health. 2022;7(7):e009365. [FREE Full text] [doi: 10.1136/bmjgh-2022-009365] [Medline: 35902205]

71. Hedt-Gauthier B, Miranda E, Nkurunziza T, Hughes O, Boatin AA, Gaju E, et al. Telemedicine for surgical site infection diagnosis in rural Rwanda: concordance and accuracy of image reviews. World J Surg. 2022;46(9):2094-2101. [doi: 10.1007/s00268-022-06597-8] [Medline: 35665833]

72. Totty JP, Harwood AE, Wallace T, Smith GE, Chetter IC. Use of photograph-based telemedicine in postoperative wound assessment to diagnose or exclude surgical site infection. J Wound Care. 2018;27(3):128-135. [doi: 10.12968/jowc.2018.27.3.128] [Medline: 29509108]

73. Wang L, Pedersen PC, Strong DM, Tulu B, Agu E, Ignotz R. Smartphone-based wound assessment system for patients with diabetes. IEEE Trans Biomed Eng. 2015;62(2):477-488. [FREE Full text] [doi: 10.1109/TBME.2014.2358632] [Medline: 25248175]

74. Howell RS, Liu HH, Khan AA, Woods JS, Lin LJ, Saxena M, et al. Development of a method for clinical evaluation of artificial intelligence-based digital wound assessment tools. JAMA Netw Open. 2021;4(5):e217234. [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.7234] [Medline: 34009348]

75. Facts and figures 2022—mobile phone ownership. International Telecommunication Union. 2022. URL: https://www.itu.int/itu-d/reports/statistics/2022/11/24/ff22-mobile-phone-ownership/ [accessed 2023-10-03]

76. Tonekaboni S, Joshi S, McCradden MD, Goldenberg A. What Clinicians Want: Contextualizing Explainable Machine Learning for Clinical End Use. In: Proceedings of the 4th Machine Learning for Healthcare Conference. Presented at: Machine Learning for Healthcare; August 2019, 2019;359-380; Ann Arbor, MI. URL: https://proceedings.mlr.press/v106/tonekaboni19a.html

77. Macefield RC, Blazeby JM, Reeves BC, King A, Rees J, Pullyblank A, et al. Remote assessment of surgical site infection (SSI) using patient-taken wound images: development and evaluation of a method for research and routine practice. J Tissue Viability. 2023;32(1):94-101. [FREE Full text] [doi: 10.1016/j.jtv.2023.01.001] [Medline: 36681617]

78. Alayande BT, Prasad S, Abimpaye M, Bakorimana L, Niyigena A, Nkurunziza J, et al. Image-based surgical site infection algorithms to support home-based post-cesarean monitoring: lessons from Rwanda. PLOS Glob Public Health. 2023;3(2):e0001584. [FREE Full text] [doi: 10.1371/journal.pgph.0001584] [Medline: 36963016]

79. Bruce J, Russell EM, Mollison J, Krukowski ZH. The quality of measurement of surgical wound infection as the basis for monitoring: a systematic review. J Hosp Infect. 2001;49(2):99-108. [doi: 10.1053/jhin.2001.1045] [Medline: 11567554]

80. Lathan R, Sidapra M, Yiasemidou M, Long J, Totty J, Smith G, et al. Diagnostic accuracy of telemedicine for detection of surgical site infection: a systematic review and meta-analysis. npj Digit Med. 2022;5(1):108. [FREE Full text] [doi: 10.1038/s41746-022-00655-0] [Medline: 35922663]

81. Guerra J, Guichon C, Isnard M, So S, Chan S, Couraud S, et al. Active prospective surveillance study with post-discharge surveillance of surgical site infections in Cambodia. J Infect Public Health. 2015;8(3):298-301. [FREE Full text] [doi: 10.1016/j.jiph.2014.09.007] [Medline: 25466596]

82. Ashby E, Haddad FS, O'Donnell E, Wilson APR. How will surgical site infection be measured to ensure "high quality care for all"? J Bone Joint Surg Br. 2010;92(9):1294-1299. [FREE Full text] [doi: 10.1302/0301-620X.92B9.22401] [Medline: 20798451]

83. Hopkins B, Eustache J, Ganescu O, Ciopolla J, Kaneva P, Fiore JF, et al. At least ninety days of follow-up are required to adequately detect wound outcomes after open incisional hernia repair. Surg Endosc. 2022;36(11):8463-8471. [doi: 10.1007/s00464-022-09143-8] [Medline: 35257211]

84. Holihan JL, Flores-Gonzalez JR, Mo J, Ko TC, Kao LS, Liang MK. How long is long enough to identify a surgical site infection? Surg Infect (Larchmt). 2017;18(4):419-423. [doi: 10.1089/sur.2016.132] [Medline: 28004989]

85.   Carpini GD, Giannella L, Di Giuseppe J, Fioretti M, Franconi I, Gatti L, et al. Inter-rater agreement of CDC criteria and ASEPSIS score in assessing surgical site infections after cesarean section: a prospective observational study. Front Surg. 2023;10:1123193. [FREE Full text] [doi: 10.3389/fsurg.2023.1123193] [Medline: 37675248]
86.   Allami MK, Jamil W, Fourie B, Ashton V, Gregg PJ. Superficial incisional infection in arthroplasty of the lower limb. Interobserver reliability of the current diagnostic criteria. J Bone Joint Surg Br. 2005;87-B(9):1267-1271. [FREE Full text] [doi: 10.1302/0301-620X.87B9.16672] [Medline: 16129756]
87.   Collins GS, Dhiman P, Navarro CLA, Ma J, Hooft L, Reitsma JB, et al. Protocol for development of a reporting guideline (TRIPOD-AI) and risk of bias tool (PROBAST-AI) for diagnostic and prognostic prediction model studies based on artificial intelligence. BMJ Open. 2021;11(7):e048008. [FREE Full text] [doi: 10.1136/bmjopen-2020-048008] [Medline: 34244270]

Abbreviations

CNN: convolutional neural network
DL: deep learning
JBI: Joanna Briggs Institute
ML: machine learning
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews
PROBAST: Prediction Model Risk of Bias Assessment Tool
RoB: risk of bias
SSI: surgical site infection
SVM: support vector machine
TRIPOD: Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis