Natural Resources for Human Health
eISSN: 2583-1194

Original Research

Received 1 Feb 2026
Revised 13 Feb 2026
Accepted 21 Feb 2026

Natural Resources for Human Health 2026, Vol. 6 (1s): 818-831

KEYWORDS:
Digital health; teledentistry; tele-intensive care; artificial intelligence; patient safety; systematic review.

Digital Health Technologies in Dentistry and Intensive Care: Clinical Applications, Patient Safety, and Outcomes

Awatif Fahad Hassan Albalawi[1*], Alaa Eidhah Al Thomali[2], Hassan Amer A Alsuayri[3], Abduljalil Jassem Almahfoudh[4], Masoumh Jafar Mohammed Aldgham[5], Hussam Mustafa M Rawas[6], Zainab Abdullah Ali Al Bensaad[7], Maryam Ali Allayli[8], Abrar Abdulhamid Almousa[9], Bushra Ahmed Alrebh[10], Abrar Ali Al Abazaid[11]

[1] Technician Dental Assistant, King Salman Military Hospital, Saudi Arabia
[2] Technician Dental Assistant, King Salman Military Hospital, Saudi Arabia
[3] General Surgery, King Abdullah Hospital, Bisha, Saudi Arabia, [4] IT Technician, Maternity and Children Hospital, Dammam, Saudi Arabia
[5] Health Administration Technician, Maternity and Children Hospital, Dammam, Saudi Arabia
[6] Critical Care Physician, King Fahad Armed Forces Hospital, Saudi Arabia
[7] Dentist, Primary Health Care, Saudi Arabia
[8] Dentist, Primary Health Care, Saudi Arabia
[9] Dentist, Alfaisal Primary Health Care, Alahsa Health Cluster, Saudi Arabia
[10] Dentist, Alfaisal Primary Health Care, Alahsa Health Cluster, Saudi Arabia
[11] Dentist, Alfaisal Primary Health Care, Alahsa Health Cluster, Saudi Arabia
* Corresponding author: Awatif Fahad Hassan Albalawi, King Salman Military Hospital, Saudi Arabia

ABSTRACT: Digital health technologies are used increasingly in dentistry and intensive care, but their effects on safety and outcomes vary by clinical task and implementation. This systematic review synthesized comparative evidence from English-language studies published from January 2009 through December 2024. PubMed, Scopus, Web of Science, and Embase were searched, and reference lists were screened. Risk of bias was assessed using RoB 2, ROBINS-I, and QUADAS-2. Heterogeneous findings were synthesized narratively. Thirteen studies were included, comprising five dental and eight intensive care studies. In dentistry, telephone follow-up produced symptom and complication findings comparable to in-person review. Smartphone photographic screening had high specificity (97–98%) and modest sensitivity (60–68%). Artificial intelligence assistance improved caries-detection sensitivity (0.81 versus 0.72) but increased invasive treatment decisions. Six of seven tele-ICU studies reported associations with lower mortality or shorter stays, with effects varying by program design and engagement. A small randomized trial of machine-learning sepsis alerts reported shorter stays and lower mortality. No study was judged at low overall risk of bias, and observational findings remained vulnerable to confounding. Digital technologies may improve selected diagnostic and care processes when integrated into clinical workflows. Their implementation should include monitoring of missed diagnoses, overtreatment, and downstream patient outcomes. Larger controlled evaluations are needed to establish sustained safety and clinical benefit.

1. INTRODUCTION

Digital health technologies, including telehealth platforms, remote monitoring, mobile applications, and AI-based decision support, are being absorbed into routine clinical practice across specialties (Irving et al., 2018). Their promise is easy to state: wider access to expertise, earlier detection of deterioration or disease, and more reliable adherence to evidence-based processes. Whether that promise translates into safer care and better outcomes, however, depends on the clinical task, the technology, and the system in which it is embedded.

Dentistry illustrates the access side of this argument. Dental caries is the most prevalent noncommunicable disease worldwide, and many affected people face geographic, financial, or logistical barriers to regular examination (World Health Organization, 2022). Teledentistry, the remote delivery of dental care, advice, or treatment through information technology, has been proposed as a way to extend screening, triage, and follow-up beyond the dental chair (Irving et al., 2018). Syntheses of diagnostic accuracy studies suggest that image-based remote caries detection is feasible but varies with the method used (Estai et al., 2016b).

A second strand of dental digital health concerns computer vision. Convolutional neural networks have been trained to detect periodontal bone loss and proximal caries on radiographs, with performance broadly comparable to clinicians in experimental settings (Krois et al., 2019). Randomized evidence on clinician use of such tools is still sparse, and early work indicates that gains in sensitivity may be accompanied by more invasive treatment decisions, which can erode clinical and economic value (Mertens et al., 2021; Schwendicke et al., 2022).

Intensive care faces a different problem: a persistent mismatch between the number of critically ill patients and the availability of intensivists. Tele-intensive care units (tele-ICUs) allow off-site teams to watch physiological data, respond to alarms, and audit adherence to best practice across several ICUs at once. Early observational studies reached discordant conclusions, with one academic center reporting lower mortality and shorter stays (Lilly et al., 2011) and a multi-hospital health system reporting no overall improvement (Thomas et al., 2009).

Subsequent work has attributed this variability to implementation rather than to the technology itself. Large multi-system analyses have linked outcomes to specific program components such as early intensivist case review and rapid alarm response (Lilly et al., 2014), and national analyses show a small average mortality benefit that conceals wide differences between hospitals (Kahn et al., 2016). In parallel, machine-learning alerts for sepsis have entered ICU workflows, with a small randomized trial reporting favorable outcomes (Shimabukuro et al., 2017) at a time when external validation of deployed sepsis models has been uneven (Fleuren et al., 2020; Wong et al., 2021).

Despite these parallel developments, the dental and intensive care literatures have grown largely apart, and existing reviews are confined to one field or one technology (Young et al., 2011; Estai et al., 2016b). Yet the two fields share the same underlying questions: does a digital tool detect problems as accurately as conventional care, does it change what clinicians do next, and does it improve or endanger patients? A cross-field comparison can show which determinants of success are specific to a clinical domain and which generalize.

This review addressed the following question: how do digital health technologies, compared with conventional care or unaided assessment, affect clinical applications, patient safety, and healthcare outcomes in dentistry and intensive care?

2. MATERIALS AND METHODS

2.1 Protocol and registration

The review was conducted and reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 statement and the PRISMA-S extension for literature searching.

2.2 Eligibility criteria

The research question was framed using the PICO format:

Population: patients receiving dental care, or critically ill adults treated in intensive care units (including clinicians interpreting patient data in controlled comparisons).

Intervention or exposure: digital health technologies, defined as teledentistry or tele-ICU programs, remote monitoring, mobile health applications, and AI-based or electronic clinical decision support.

Comparator: conventional in-person care, unaided clinician judgment, a pre-implementation period, or non-adopting sites.

Outcomes: primary, patient safety (complications, mortality, diagnostic accuracy, and diagnostic or treatment errors); secondary, healthcare outcomes (length of stay, adherence to care processes, patient satisfaction, access, and cost).

Eligible designs were randomized controlled trials, cohort, case-control, cross-sectional (including diagnostic accuracy), and case-series studies published in English between January 2009 and December 2024. Exclusions were systematic or narrative reviews, editorials, letters, conference abstracts, protocols without results, qualitative-only studies, case reports, in vitro or animal studies, algorithm-development studies without a clinician or conventional-care comparator, and reports lacking extractable outcome data.

2.3 Information sources and search strategy

PubMed, Scopus, Web of Science, and Embase were searched from 1 January 2009 to 15 December 2024. Reference lists of included studies and relevant reviews were hand-searched. The PubMed strategy combined three concept blocks: (1) technology terms (telemedicine, teledentistry, tele-ICU, ICU telemedicine, remote monitoring, digital health, mHealth, artificial intelligence, machine learning, deep learning, clinical decision support); (2) setting terms (dentistry, dental, oral health, oral surgery, intensive care units, critical care, critically ill); and (3) outcome terms (patient safety, diagnostic accuracy, mortality, length of stay, complications, outcomes). Controlled vocabulary and free-text terms were combined with Boolean operators, limited by date and language, and adapted to each database.

2.4 Selection process and data extraction

Records were exported to reference-management software and de-duplicated. Two reviewers independently screened titles and abstracts, then full texts, resolving disagreements by discussion or a third reviewer. Two reviewers independently extracted study design, setting, sample size, technology, comparator, outcomes, and main results using a piloted form.

2.5 Risk of bias and certainty of evidence

Randomized trials were appraised with the Cochrane RoB 2 tool, non-randomized comparative studies (cohort and case-control) with ROBINS-I, and diagnostic accuracy studies with QUADAS-2. Domain judgments were harmonized into an overall rating using each tool's standard terminology (RoB 2: low, some concerns, high; ROBINS-I: low, moderate, serious, critical; QUADAS-2: low, unclear, high). Certainty of evidence for each main outcome was assessed using GRADE.

2.6 Synthesis methods

Meta-analysis was planned if studies were sufficiently homogeneous in population, intervention, comparator, and outcome. Because the technologies, settings, outcome definitions, and designs differed substantially, a pooled estimate was judged inappropriate and results were synthesized narratively according to the Synthesis Without Meta-analysis (SWiM) guideline, grouped by clinical field and technology type, with the direction and size of effects summarized in tables.

3. RESULTS AND DISCUSSION

3.1 Study selection

The database searches identified 4,212 records, of which 1,534 remained after duplicate removal. After title and abstract screening, 54 reports were retrieved in full text and 41 were excluded for the reasons shown in **Figure 1**. Thirteen studies met the inclusion criteria: five in dentistry and eight in intensive care.

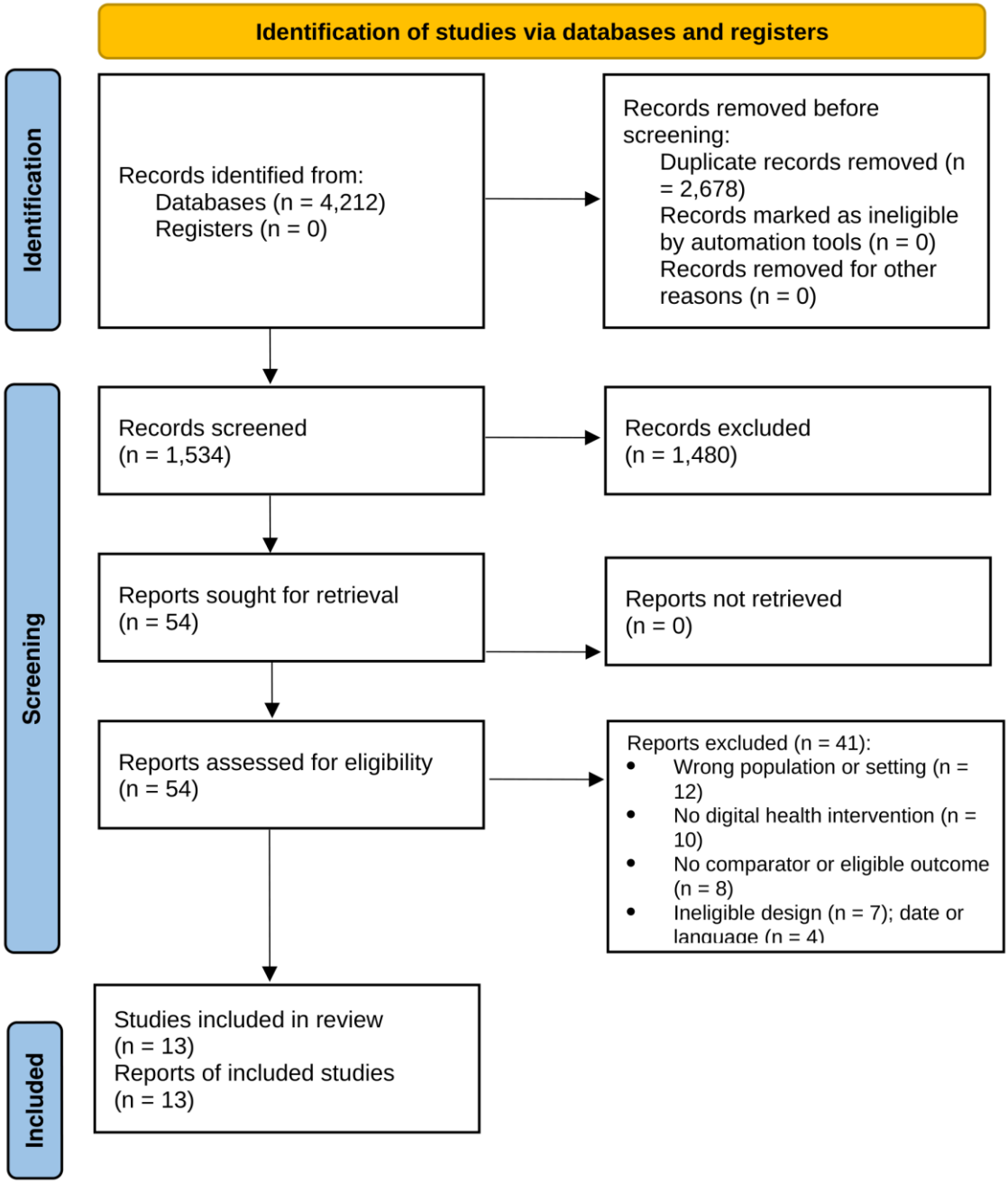

Figure 1. PRISMA 2020 flow diagram of study selection. Database records comprised PubMed (n = 1,450), Scopus (n = 1,120), Web of Science (n = 980), and Embase (n = 662). Adapted from Page et al. (2021), CC BY 4.0.

3.2 Study characteristics

The included studies were published between 2009 and 2022 and originated from the United States (eight), Germany (three), Australia (one), and Saudi Arabia (one). Designs comprised four randomized trials, six cohort or pre-post studies, one case-control study, and two diagnostic accuracy (cross-sectional) studies. Sample sizes ranged from 60 participants to 118,990 ICU patients. Characteristics are summarized in Table 1.

Table 1. Characteristics of the included studies

Study	Country	Design	Sample and setting	Digital technology	Comparator
Dentistry					
Aboalshamat et al. (2022)	Saudi Arabia	Single-blind parallel RCT	70 adult dental patients, university dental hospital, COVID-19 period	WhatsApp-based remote dental examination	Waiting-list control; baseline in-person examination as reference
Heimes et al. (2022)	Germany	Monocentric prospective RCT	60 patients after dentoalveolar surgery under local anesthesia	Telephone follow-up with structured questionnaire	Conventional in-person follow-up
Mertens et al. (2021)	Germany	Cluster-randomized crossover RCT	22 dentists reading 20 bitewing radiographs each	AI caries-detection software (dentalXrai Pro)	Unaided radiograph reading
Estai et al. (2016a)	Australia	Cross-sectional diagnostic accuracy	100 attendees of routine dental check-ups	Smartphone intraoral photographs with store-and-forward review	Face-to-face visual examination
Krois et al. (2019)	Germany	Cross-sectional diagnostic accuracy	2,001 panoramic radiograph segments; 6 dentists as readers	Deep convolutional neural network for bone loss	Dentist assessment
Intensive care					
Thomas et al. (2009)	USA	Pre-post cohort	4,142 patients (2,034 pre, 2,108 post); 6 ICUs in 5 hospitals	Tele-ICU remote monitoring	Pre-implementation period
Lilly et al. (2011)	USA	Pre-post cohort	6,290 adults; 7 ICUs, one academic center	Tele-ICU with reengineered care processes	Pre-implementation period
Lilly et al. (2014)	USA	Non-randomized pre-post cohort with controls	118,990 adults; 56 ICUs, 32 hospitals, 19 systems	Multicomponent ICU telemedicine	Control ICUs (11,558 patients) and pre-implementation
Kahn et al. (2016)	USA	Retrospective matched case-control (difference-in-differences)	132 adopting hospitals and 389 matched non-adopting hospitals	ICU telemedicine adoption	Matched non-adopting hospitals

Study	Country	Design	Sample and setting	Digital technology	Comparator
Shimabukuro et al. (2017)	USA	Open-label RCT	142 adults; 2 mixed ICUs (75 control, 67 intervention)	Machine-learning severe-sepsis prediction alerts	Existing rule-based sepsis detector
O'Shea et al. (2022)	USA	Retrospective cohort, propensity-matched	Veterans Health Administration ICU admissions, 2005 to 2018	Tele-critical care by interaction frequency	Matched control facilities
Fusaro et al. (2021)	USA	Retrospective pre-post cohort	16,165 patient-stays (1,581 pre, 14,584 post)	Tele-ICU implementation	Pre-implementation period
Hawkins et al. (2016)	USA	Retrospective comparative cohort	Mixed medical-surgical ICUs served by one tele-ICU center; patient total not verified in accessible report	Direct intervention with timely notification	Mixed or monitor-and-notify comanagement

RCT, randomized controlled trial; AI, artificial intelligence; ICU, intensive care unit.

3.3 Dental applications

Two trials examined remote care delivery. Telephone follow-up after dentoalveolar surgery did not differ from conventional follow-up in symptom frequency or complication rate, and 83.3% of patients allocated to telephone follow-up preferred it, compared with 16.7% who preferred conventional aftercare ($p = 0.047$) (Heimes et al. (2022)). In the trial of WhatsApp-based examination, accuracy was good for chief complaint and missing teeth (74.3% each), filled teeth (71.4%), and oral hygiene status (65.7%) but moderate for health complaints (51.4%) and for decayed teeth; decayed-tooth counts and DMF scores obtained remotely were significantly higher than at the baseline clinical examination ($p < 0.05$), indicating overestimation (Aboalshamat et al. (2022)).

In the smartphone-photography accuracy study, sensitivity against face-to-face examination was 60% to 68% and specificity 97% to 98%, with intra-rater agreement of 0.89 and inter-rater agreement of 0.57 to 0.61 (kappa) (Estai et al. (2016a)). For AI, dentists using caries-detection software achieved a higher area under the curve (0.89 versus 0.85) and sensitivity (0.81 versus 0.72) than unaided readers, without a significant change in specificity, but made more non-invasive and invasive treatment decisions ($p < 0.05$) (Mertens et al. (2021)). A neural network for periodontal bone loss reached an accuracy of 0.81 (sensitivity 0.81, specificity 0.81), against 0.76 for six dentists (sensitivity 0.92, specificity 0.63); the accuracy difference was not significant ($p = 0.067$) (Krois et al. (2019)).

3.4 Intensive care applications

Among tele-ICU cohorts, hospital mortality fell from 13.6% to 11.8% (adjusted OR 0.40, 95% CI 0.31 to 0.52), hospital stay from 13.3 to 9.8 days, and ventilator-associated pneumonia from 13% to 1.6% after tele-ICU reengineering at one academic center (Lilly et al. (2011)). The multi-system study of 118,990 patients found lower hospital mortality (HR 0.84, 95% CI 0.78 to 0.89) and ICU mortality (HR 0.74, 95% CI 0.68 to 0.79), with benefit linked to early intensivist case review, use of performance data, best-practice adherence, and fast alarm response (Lilly et al. (2014)). By contrast, a six-ICU system found crude hospital mortality of 12.0% before and 9.9% after implementation, with no significant adjusted change in mortality or length of stay, although survival improved in the sickest patients (Thomas et al. (2009)).

In a national matched analysis, adoption was associated with a small relative reduction in 90-day mortality (ratio of odds ratios 0.96, 95% CI 0.95 to 0.98), but only 16 of 132 adopting hospitals (12.2%) showed a significant reduction (Kahn et al. (2016)).

Within the Veterans Health Administration, facilities with frequent tele-critical care interactions had shorter ICU stays (RR 0.39, 95% CI 0.23 to 0.65), whereas facilities with infrequent interactions did not (O'Shea et al. (2022)). In the sepsis trial, machine-learning alerts reduced length of stay from 13.0 to 10.3 days ($p = 0.042$) and in-hospital mortality by 12.4 percentage points ($p = 0.018$), with no adverse events reported (Shimabukuro et al. (2017)).

In a separate pre-post cohort, Fusaro et al. (2021) reported risk-adjusted ICU mortality of 8.7% before and 6.5% after tele-ICU implementation ($p < 0.01$). Hawkins et al. (2016) found that direct intervention with timely notification was associated with shorter acuity-adjusted hospital length of stay than mixed or monitor-and-notify comanagement.

Table 2. Quantified key findings of the included studies.

Study	Primary outcomes	Key quantitative findings
Aboalshamat et al. (2022)	Diagnostic agreement; perceptions	Accuracy: chief complaint 74.3%, missing teeth 74.3%, filled teeth 71.4%, oral hygiene 65.7%, health complaints 51.4%. Decayed-tooth count and DMF score higher with teledentistry than at baseline ($p < 0.05$). Group knowledge and attitudes similar; more favorable experience in the teledentistry group.
Heimes et al. (2022)	Symptoms, complications, preference	No significant differences in symptom frequency or complications. Preference for telephone follow-up 83.3% versus 16.7% for conventional ($p = 0.047$).
Mertens et al. (2021)	Diagnostic accuracy; treatment decisions	AUC 0.89 (95% CI 0.87 to 0.90) with AI versus 0.85 (0.83 to 0.86) without ($p < 0.05$). Sensitivity 0.81 (0.74 to 0.87) versus 0.72 (0.64 to 0.79). Specificity not significantly different. More non-invasive and invasive treatment decisions with AI ($p < 0.05$).
Estai et al. (2016a)	Sensitivity, specificity, agreement	Sensitivity 60% to 68%; specificity 97% to 98%; intra-rater kappa 0.89; inter-rater kappa 0.57 to 0.61.
Krois et al. (2019)	Accuracy for periodontal bone loss	CNN accuracy 0.81 (SD 0.02), sensitivity 0.81, specificity 0.81. Dentists: accuracy 0.76 (SD 0.06), sensitivity 0.92, specificity 0.63. Difference not significant ($p = 0.067$). Dentist agreement moderate (Fleiss kappa 0.52).
Thomas et al. (2009)	Hospital and ICU mortality, complications, LOS	Hospital mortality 12.0% to 9.9% (crude); ICU mortality 9.2% to 7.8%. No significant adjusted differences in mortality or LOS. Improved survival in sicker patients (interaction $p < 0.001$).
Lilly et al. (2011)	Adjusted hospital mortality, LOS, complications	Hospital mortality 13.6% to 11.8% (adjusted OR 0.40; 0.31 to 0.52); ICU mortality 10.7% to 8.6%; hospital LOS 13.3 to 9.8 days; ICU LOS 6.4 to 4.5 days; VAP 13% to 1.6% (OR 0.15; 0.09 to 0.23); catheter-related bloodstream infection 1.0% to 0.6% (OR 0.50; 0.27 to 0.93).
Lilly et al. (2014)	Hospital and ICU mortality; LOS; components	Hospital mortality HR 0.84 (0.78 to 0.89); ICU mortality HR 0.74 (0.68 to 0.79); lower LOS. Components linked to benefit: intensivist case review within 1 hour, use of performance data, best-practice adherence, faster alert response.
Kahn et al. (2016)	90-day mortality (difference-in-differences)	Ratio of odds ratios 0.96 (0.95 to 0.98). Significant reduction in 16 of 132 adopting hospitals (12.2%); larger gains in high-volume urban hospitals (median ratio of odds ratios 1.01; range 0.45 to 2.54).

Study	Primary outcomes	Key quantitative findings
Shimabukuro et al. (2017)	LOS; in-hospital mortality	LOS 13.0 to 10.3 days (p = 0.042). In-hospital mortality reduced by 12.4 percentage points (58.0% relative; p = 0.018). No adverse events.
O'Shea et al. (2022)	ICU LOS and mortality by interaction frequency	Top-tertile interaction facilities: ICU LOS RR 0.39 (0.23 to 0.65) versus matched controls. Bottom-tertile facilities: no significant difference.
Fusaro et al. (2021)	Risk-adjusted ICU mortality	8.7% pre vs 6.5% post (p < 0.01); reduction concentrated in nighttime admissions.
Hawkins et al. (2016)	Acuity-adjusted hospital LOS	Direct intervention: 0.68 (95% CI 0.65–0.70); mixed methods: 0.70 (0.69–0.72; P = 0.01); monitor-and-notify: 0.83 (0.80–0.86; P < 0.001 vs mixed). These are adjusted LOS measures, not days.

CI, confidence interval; HR, hazard ratio; OR, odds ratio; RR, rate ratio; LOS, length of stay; VAP, ventilator-associated pneumonia; AUC, area under the receiver-operating characteristic curve; CNN, convolutional neural network.

3.5 Patient safety and healthcare outcomes

Safety signals differed by field. In dentistry, the principal risks were misclassification and overtreatment: remote examination overestimated decay, photographic screening missed a considerable share of lesions, and AI assistance increased invasive treatment. In intensive care, the main safety outcomes were mortality and preventable complications, which improved in well-integrated programs but not consistently across settings. Healthcare outcomes such as length of stay and process adherence were reported mainly in intensive care, whereas satisfaction and access dominated the dental trials. Table 3 summarizes these signals and the authors' conclusions.

Table 3. Patient-safety signals, healthcare-outcome signals, and authors' conclusions.

Study	Patient-safety signal	Healthcare-outcome signal	Authors' conclusion (paraphrased)
Aboalshamat et al. (2022)	Overestimation of decay; remote findings need clinical confirmation	Good acceptance; photo-taking problems reported	Teledentistry is acceptable and useful for preliminary examination but not a substitute for clinical examination.
Heimes et al. (2022)	No difference in complications detected	Patient preference for phone follow-up; potential time and cost savings	Telephone-only follow-up after minor dentoalveolar surgery is well accepted and appears safe.
Mertens et al. (2021)	More invasive treatment decisions	Higher diagnostic accuracy for enamel lesions	AI improves accuracy but may increase invasive therapy.
Estai et al. (2016a)	Missed lesions (modest sensitivity)	Feasible remote screening by mid-level providers	Mobile teledentistry shows potential for remote caries screening.
Krois et al. (2019)	Dentists over-called disease (low specificity); CNN more balanced	Possible reduction of diagnostic effort	CNN performs at least similarly to dentists on panoramic radiographs.

Study	Patient-safety signal	Healthcare-outcome signal	Authors' conclusion (paraphrased)
Thomas et al. (2009)	No overall change in complications; benefit in sicker patients	No overall LOS benefit; limited use of delegated authority	Remote monitoring was not linked to overall improvement; further study with outcome and cost monitoring is needed.
Lilly et al. (2011)	Lower VAP and catheter-related infections	Shorter ICU and hospital stay; higher best-practice adherence	Tele-ICU reengineering was linked to lower mortality, shorter stays, and fewer preventable complications.
Lilly et al. (2014)	Lower hospital and ICU mortality	Shorter LOS; component-level drivers identified	Interventions that add early intensivist involvement, best-practice adherence, and fast alarm response were linked to better outcomes.
Kahn et al. (2016)	Small average mortality reduction with large variation	Greater effect in large urban hospitals	Effect is heterogeneous across hospitals.
Shimabukuro et al. (2017)	Lower mortality; no adverse events	Shorter LOS	The algorithm was associated with improved outcomes.
O'Shea et al. (2022)	Mortality not the main signal	Shorter ICU stay only with frequent interactions	More frequent tele-critical care interactions may improve outcomes for the most acute cases.
Fusaro et al. (2021)	Lower adjusted ICU mortality	Mortality reduction concentrated in nighttime admissions	Tele-ICU implementation was associated with lower mortality, particularly for nighttime admissions.
Hawkins et al. (2016)	Specific safety outcomes not reported in the accessible abstract	Shorter acuity-adjusted LOS with direct intervention	Direct intervention with timely notification was associated with shorter LOS.

3.6 Risk of bias

No study was at low overall risk of bias (Table 4). Among randomized trials, concerns arose from small samples, unblinded delivery, subjective outcomes, optional use of the intervention, and, in the sepsis trial, unblinding after alerts. The tele-ICU cohorts were exposed to confounding by secular trends and concurrent quality initiatives despite severity adjustment. The two diagnostic accuracy studies relied on imperfect reference standards or non-representative image samples.

Table 4. Risk-of-bias assessment of the included studies.

Study	Design	Tool	Principal concerns	Overall
Aboalshamat et al. (2022)	RCT	RoB 2	Software randomization (low); waiting-list control and single blinding; enrolment below target (70 of 102)	Some concerns
Heimes et al. (2022)	RCT	RoB 2	Small sample; participants and clinicians not blinded; subjective questionnaire outcomes	Some concerns

Study	Design	Tool	Principal concerns	Overall
Mertens et al. (2021)	Cluster crossover RCT	RoB 2	Dentists could switch AI augmentation on or off; artificial reading task rather than patient outcomes	Some concerns
Shimabukuro et al. (2017)	RCT	RoB 2	Open-label; allocation revealed once alerts fired; 142 patients over about 3 months; single center; large effect from few events	High
Estai et al. (2016a)	Diagnostic accuracy	QUADAS-2	Sampling of attendees unclear; visual examination as imperfect reference; flow and timing low risk	Unclear
Krois et al. (2019)	Diagnostic accuracy	QUADAS-2	Image segments rather than whole radiographs (applicability); limited dataset; reference by repeated measurement	High
Thomas et al. (2009)	Pre-post cohort	ROBINS-I	Overlapping periods and secular trends; most physicians delegated minimal authority (exposure misclassification)	Serious
Lilly et al. (2011)	Pre-post cohort	ROBINS-I	Single center; concurrent quality initiatives; severity-adjusted, objective outcomes	Serious
Lilly et al. (2014)	Non-randomized pre-post with controls	ROBINS-I	Large sample with concurrent controls and adjustment; unblinded; residual confounding	Moderate
Kahn et al. (2016)	Case-control (DiD)	ROBINS-I	Matched controls and difference-in-differences; claims data limit severity adjustment	Moderate
O'Shea et al. (2022)	Retrospective cohort	ROBINS-I	Facility-level exposure defined by utilization tertile; unmeasured confounding	Serious
Fusaro et al. (2021)	Pre-post cohort	ROBINS-I	Secular trends; implementation-related changes; residual confounding despite adjustment	Moderate
Hawkins et al. (2016)	Retrospective cohort	ROBINS-I	ICU-level exposure; potential differences in case mix and organization; residual confounding	Moderate

The original 11-study judgments are reported as consensus following independent dual review. The two added assessments are provisional and require full-text, domain-level confirmation. Overall ratings use tool-specific terminology. RoB 2 uses low risk, some concerns, or high risk. ROBINS-I uses low, moderate, serious, or critical risk, and QUADAS-2 judgments use low, unclear, or high risk.

3.7 Synthesis and certainty of evidence

A meta-analysis was not undertaken because the technologies, populations, comparators, outcome metrics, and designs were too different to pool meaningfully; the tele-ICU mortality estimates, for example, derived from different adjustment methods

and effect measures. Certainty of evidence was judged low for teledentistry accuracy outcomes, low to moderate for AI-assisted caries detection, and low for tele-ICU effects on mortality and length of stay owing to observational designs, inconsistency, and confounding. The additional observational findings did not increase confidence in a causal effect of tele-ICU implementation on mortality or length of stay.

3.8 Discussion

Digital health technologies are redistributing diagnostic and monitoring work among clinicians, patients, and algorithms, and this review compared how that redistribution has performed in two very different clinical environments. Across both fields, the evidence indicates that digital tools can preserve or enhance clinical performance when the task is well defined and the workflow is rebuilt around the tool, but that benefit is conditional rather than automatic.

In dentistry, remote follow-up after minor surgery performed comparably to in-person review for detecting symptoms and complications and was preferred by most patients assigned to it. This agrees with Susarla et al. (2011), who reported high satisfaction with telephone follow-up after dentoalveolar surgery. Pippi et al. (2018), however, questioned whether telephone contact is adequate for early recognition of inflammatory complications, a reminder that safety depends on careful case selection. Photographic screening showed a recurring pattern in which specificity was high but sensitivity modest, meaning lesions were missed more often than healthy teeth were mislabeled, whereas the randomized trial of messaging-based examination found overestimation of decay. These findings are reconcilable. Standardized image capture by trained providers produces conservative reads, while patient-captured photographs suffer from poor lighting, angle, and plaque or staining that can be mistaken for cavitation. Estai et al. (2016b) similarly described substantial variation in teledentistry accuracy across methods, and Irving et al. (2018) framed teledentistry mainly as an enabler of access. Taken together, remote dental assessment is best positioned as a triage and monitoring tool rather than a replacement for clinical examination.

Artificial intelligence changed the balance of detection in a different way. Software assistance improved sensitivity for early enamel lesions, where human readers depend on faint changes in radiographic density, but offered little advantage for lesions that were already obvious. The mechanism is pattern recognition below the threshold of unaided perception, and its cost is a greater tendency to treat lesions that may never progress. Schwendicke et al. (2022) showed, using the same trial, that this shift toward invasive treatment can undermine economic value unless AI also guides subsequent management, and Schwendicke et al. (2021) reported higher sensitivity than dentists for proximal caries in the underlying data. The periodontal network matched but did not clearly surpass clinicians, and its errors were more evenly distributed, whereas dentists favored sensitivity at the expense of specificity. Because algorithm and human errors differ in kind, combined use may be more defensible than competition between the two.

In intensive care, the evidence for tele-ICU is mixed but patterned. Studies in which remote teams were tightly integrated with quality improvement reported better survival, shorter stays, and fewer preventable complications, whereas the multi-hospital system in which most physicians delegated minimal authority found no overall effect, with benefit confined to the sickest patients. National analyses showed a small average benefit that concealed wide variation between hospitals. The additional cohorts by Fusaro et al. (2021) and Hawkins et al. (2016) reinforce the pattern that tele-ICU outcomes vary with program design and the way remote teams participate in care. This heterogeneity agrees with Young et al. (2011), whose meta-analysis found that benefit was clearer for ICU-level than for hospital-level outcomes, and with Fusaro et al. (2021), who reported differences in risk-adjusted mortality according to the time of day of tele-ICU coverage, underscoring that implementation details matter.

Tele-ICU is better understood as an organizational intervention than as a device. Its effects plausibly operate through earlier intensivist review of new admissions, continuous surveillance with faster response to alarms, standardized auditing of preventive practices such as thromboprophylaxis and ventilator bundles, and routine use of performance data. The component analysis by Lilly et al. (2014) supports this view, Hawkins et al. (2016) linked the method of co-management to length of stay, and Kahn et al. (2019) identified organizational and relational determinants, such as acceptance by bedside staff, that shape whether remote input is actually used. Fortis et al. (2018) reported fewer interhospital transfers after tele-ICU introduction in a national health system, consistent with expanded local capability. The finding that benefit concentrated in facilities with frequent interactions suggests a dose-response relationship, and that sicker patients gain most is biologically coherent, because the margin for early detection of deterioration is widest in those with the least physiological reserve.

The machine-learning sepsis trial suggests that algorithmic surveillance can accelerate recognition and bundle initiation, but it was small, open-label, brief, and conducted at a single center, and its large effect should be interpreted with caution. External validation of a widely deployed proprietary sepsis model by Wong et al. (2021) showed weaker discrimination than developers reported, and Fleuren et al. (2020) found that models predicting sepsis accurately on retrospective data were too heterogeneous to support firm pooled conclusions. The mechanism is therefore plausible, but its generalizability to other ICUs, populations, and alert thresholds remains unproven, and alert fatigue is an unquantified risk.

Comparing the two fields yields a shared lesson: digital tools change the error profile of care, and their value depends on integration. In dentistry the characteristic hazard is overdiagnosis and overtreatment when sensitivity is raised without a corresponding management pathway, while in intensive care the characteristic hazard is variable and incomplete use of a tool whose benefit depends on engagement. Patient safety outcomes were measured far more consistently in the ICU studies, whereas dental trials emphasized diagnostic agreement and satisfaction, and few trials in either field tracked downstream treatment or harm. For clinicians and health systems, this argues for pairing technology with workflow redesign, training, careful case selection, and post-deployment audit, and for evaluation designs that capture false positives and subsequent interventions rather than accuracy alone.

Several limitations should be acknowledged. Clinical and methodological heterogeneity precluded meta-analysis, and conclusions rest on narrative synthesis. Most ICU evidence is observational and vulnerable to confounding by secular trends, while the randomized dental and sepsis trials were small with short follow-up and, in part, used artificial tasks. The technologies evaluated span more than a decade and may not represent current systems. Restriction to English-language reports and possible publication bias may have favored positive findings, and the risk-of-bias judgments, although structured, depend on published descriptions.

4. CONCLUSIONS

The available evidence suggests that digital health technologies can support selected diagnostic, follow-up, and monitoring tasks in dentistry and intensive care. Dental applications showed benefits in access and detection, accompanied by risks of missed lesions, overestimation of disease, and more invasive treatment decisions. Tele-ICU programs were associated with improved outcomes in several settings, with variation linked to the organization and intensity of remote clinical participation. Evidence for machine-learning sepsis alerts remains preliminary. Heterogeneous designs and substantial risk of bias limit causal inference and the generalizability of these findings. Implementation should therefore be accompanied by clinical oversight, workflow integration, and prospective monitoring of patient safety and downstream treatment. Larger pragmatic trials and controlled implementation studies are needed to establish sustained clinical benefit, safety, and cost-effectiveness.

FUNDING

No external funding was received for this work.

ACKNOWLEDGEMENTS

None.

AUTHOR CONTRIBUTIONS

All authors equally contributed to this work (Conceptualization, Methodology, Data Curation, Formal Analysis, Writing – Original Draft, Writing – Review & Editing, and Supervision).

CONFLICT OF INTEREST

The authors declare no conflict of interest.

Ethics Statement

This study is a systematic review of previously published, de-identified data and did not involve new human or animal subjects; formal ethics approval was therefore not required.

Data Availability Statement

The data extraction tables supporting this review are available from the corresponding author upon reasonable request. All included studies are publicly available in the cited peer-reviewed sources.

REFERENCES

[1] Aboalshamat, K.T., Althagafi, T.K., Alsaeedi, S.A., Alhumaidi, S.N., Alemam, A.A., 2022. Accuracy and perceptions of teledentistry in KSA during the COVID-19 pandemic: a single-centre randomised controlled trial. Journal of Taibah University Medical Sciences 17, 506-515. https://doi.org/10.1016/j.jtumed.2021.11.015

[2] Estai, M., Kanagasingam, Y., Huang, B., Checker, H., Steele, L., Kruger, E., 2016a. The efficacy of remote screening for dental caries by mid-level dental providers using a mobile teledentistry model. Community Dentistry and Oral Epidemiology 44, 435-441. https://doi.org/10.1111/cdoe.12232

[3] Heimes, D., Luhrenberg, P., Langguth, N., Kaya, S., Obst, C., Kammerer, P.W., 2022. Can teledentistry replace conventional clinical follow-up care for minor dental surgery? A prospective randomized clinical trial. International Journal of Environmental Research and Public Health 19, 3444. https://doi.org/10.3390/ijerph19063444

[4] Kahn, J.M., Le, T.Q., Barnato, A.E., Hravnak, M., Kuza, C.C., Pike, F., Angus, D.C., 2016. ICU telemedicine and critical care mortality: a national effectiveness study. Medical Care 54, 319-325. https://doi.org/10.1097/MLR.0000000000000485

[5] Krois, J., Ekert, T., Meinhold, L., Golla, T., Kharbot, B., Wittemeier, A., Dorfer, C., Schwendicke, F., 2019. Deep learning for the radiographic detection of periodontal bone loss. Scientific Reports 9, 8495. https://doi.org/10.1038/s41598-019-44839-3

[6] Lilly, C.M., Cody, S., Zhao, H., Landry, K., Baker, S.P., McIlwaine, J., Chandler, M.W., Irwin, R.S., 2011. Hospital mortality, length of stay, and preventable complications among critically ill patients before and after tele-ICU reengineering of critical care processes. JAMA 305, 2175-2183. https://doi.org/10.1001/jama.2011.697

[7] Lilly, C.M., McLaughlin, J.M., Zhao, H., Baker, S.P., Cody, S., Irwin, R.S., 2014. A multicenter study of ICU telemedicine reengineering of adult critical care. Chest 145, 500-507. https://doi.org/10.1378/chest.13-1973

[8] Mertens, S., Krois, J., Cantu, A.G., Arsiwala, L.T., Schwendicke, F., 2021. Artificial intelligence for caries detection: randomized trial. Journal of Dentistry 115, 103849. https://doi.org/10.1016/j.jdent.2021.103849

[9] O'Shea, A.M.J., Reisinger, H.S., Panos, R., Goede, M., Fortis, S., 2022. Association of interactions between tele-critical care and bedside with length of stay and mortality. Journal of Telemedicine and Telecare 30, 961-968. https://doi.org/10.1177/1357633X221107993

[10] Shimabukuro, D.W., Barton, C.W., Feldman, M.D., Mataraso, S.J., Das, R., 2017. Effect of a machine learning-based severe sepsis prediction algorithm on patient survival and hospital length of stay: a randomised clinical trial. BMJ Open Respiratory Research 4, e000234. https://doi.org/10.1136/bmjresp-2017-000234

[11] Thomas, E.J., Lucke, J.F., Wueste, L., Weavind, L., Patel, B., 2009. Association of telemedicine for remote monitoring of intensive care patients with mortality, complications, and length of stay. JAMA 302, 2671-2678. https://doi.org/10.1001/jama.2009.1902

[12] Fusaro, M.V., Becker, C., Miller, D., Hassan, I.F., Scurlock, C., 2021. ICU telemedicine implementation and risk-adjusted mortality differences between daytime and nighttime coverage. Chest 159, 1445-1451. https://doi.org/10.1016/j.chest.2020.10.055

[13] Hawkins, H.A., Lilly, C.M., Kaster, D.A., Groves Jr, R.H., Khurana, H., 2016. ICU telemedicine comanagement methods and length of stay. Chest 150, 314-319. https://doi.org/10.1016/j.chest.2016.03.030

[14] Estai, M., Bunt, S., Kanagasingam, Y., Kruger, E., Tennant, M., 2016b. Diagnostic accuracy of teledentistry in the detection of dental caries: a systematic review. Journal of Evidence-Based Dental Practice 16, 161-172.

[15] Fleuren, L.M., Klausch, T.L.T., Zwager, C.L., et al., 2020. Machine learning for the prediction of sepsis: a systematic review and meta-analysis of diagnostic test accuracy. Intensive Care Medicine 46, 383-400.

[16] Fortis, S., Sarrazin, M.V., Beck, B.F., et al., 2018. ICU telemedicine reduces interhospital ICU transfers in the Veterans Health Administration. Chest 154, 69-76.

[17] Irving, M., Stewart, R., Spallek, H., Blinkhorn, A., 2018. Using teledentistry in clinical practice as an enabler to improve access to oral health care: a qualitative systematic review. Journal of Telemedicine and Telecare 24, 129-146.

[18] Kahn, J.M., Rak, K.J., Kuza, C.C., et al., 2019. Determinants of intensive care unit telemedicine effectiveness: an ethnographic study. American Journal of Respiratory and Critical Care Medicine 199, 970-979.

[19] Pippi, R., Pietrantoni, A., Patini, R., Santoro, M., 2018. Is telephone follow-up really effective in early diagnosis of inflammatory complications after tooth extraction? Medicina Oral, Patologia Oral y Cirugia Bucal 23, e707-e715.

[20] Schwendicke, F., Rossi, J.G., Gostemeyer, G., Elhennawy, K., Cantu, A.G., Gaudin, R., Chaurasia, A., Gehrung, S., Krois, J., 2021. Cost-effectiveness of artificial intelligence for proximal caries detection. Journal of Dental Research 100, 369-376.

[21] Schwendicke, F., Mertens, S., Cantu, A.G., Chaurasia, A., Meyer-Lueckel, H., Krois, J., 2022. Cost-effectiveness of AI for caries detection: randomized trial. Journal of Dentistry 119, 104080. https://doi.org/10.1016/j.jdent.2022.104080

[22] Susarla, S.M., Black, R., Dodson, T.B., 2011. After dentoalveolar surgery, most patients are satisfied with telephone follow-up. Journal of Oral and Maxillofacial Surgery 69, 2099-2105. https://doi.org/10.1016/j.joms.2010.12.013

[23] Wong, A., Otles, E., Donnelly, J.P., et al., 2021. External validation of a widely implemented proprietary sepsis prediction model in hospitalized patients. JAMA Internal Medicine 181, 1065-1070.

[24] World Health Organization, 2022. Global oral health status report: towards universal health coverage for oral health by 2030. World Health Organization, Geneva.

[25] Young, L.B., Chan, P.S., Lu, X., Nallamothu, B.K., Sasson, C., Cram, P.M., 2011. Impact of telemedicine intensive care unit coverage on patient outcomes: a systematic review and meta-analysis. Archives of Internal Medicine 171, 498-506.