

AI-Integrated Telemedicine Platforms

Hugo Dubois¹, Andreas Klein², Noah Bianchi³

¹ Associate Professor, Department of Artificial Intelligence, Advanced Computing University, Paris, France. Email: hugo.dubois740@yahoo.com | ORCID: 0096-9309-3921-1699

² Assistant Professor, Department of Artificial Intelligence, Mediterranean Institute of Technology, Rome, Italy. Email: andreas.klein562@yahoo.com | ORCID: 4128-3173-9006-7705

³ Assistant Professor, Institute of Intelligent Systems, Swiss Institute of Machine Intelligence, Zurich, Switzerland. Email: noah.bianchi849@yahoo.com | ORCID: 8347-6397-0736-1223

ABSTRACT

The COVID-19 pandemic accelerated telemedicine adoption from 11% to 76% of outpatient visits in the United States during 2020, and post-pandemic utilisation has stabilised at 25-30% across specialties. Yet most telemedicine platforms remain digital video calls -- they replicate the face-to-face consultation without enhancing it. AI integration transforms telemedicine from a communication channel into a clinical decision support system: automated pre-visit symptom assessment, real-time vital sign extraction from smartphone cameras, NLP-powered clinical documentation, AI-assisted differential diagnosis, and remote patient monitoring with anomaly detection. We present the AI-Telemedicine Integration Assessment Framework (ATIAF), evaluating five AI-integrated telemedicine platform architectures -- symptom-driven triage engines, computer vision vital sign extraction, NLP clinical documentation assistants, diagnostic decision support systems, and continuous remote monitoring with predictive alerts -- across four clinical specialties (primary care, dermatology, cardiology, and mental health). Our Telemedicine AI Effectiveness Score (TAES) measures diagnostic accuracy, clinical workflow efficiency, patient satisfaction, data security compliance, and health equity impact. NLP clinical documentation achieves the highest TAES (0.918) by reducing physician documentation time by 62% while maintaining note quality, while computer vision vital sign extraction achieves the highest patient accessibility (0.960) by enabling clinical-grade measurements using only a standard smartphone camera.

Keywords: telemedicine; artificial intelligence; clinical decision support; remote patient monitoring; natural language processing; computer vision; digital health; health equity

Citation: Dubois et al. [2024]. AI-Integrated Telemedicine Platforms. DOI: <https://doi.org/10.5281/zenodo.19542621>

Copyright: © 2024 by the authors. Open access under CC BY 4.0 license.

Article Information: Received: 2023 Nov 22 Accepted: 2024 Jan 21 Published: 2024 Mar 15

Research Article: Research Article

1. Introduction

1.1 Beyond Video Calls

Telemedicine was a marginal technology before 2020 -- used primarily in rural settings and specialist consultations where distance made in-person visits impractical. The pandemic changed this overnight: CMS temporarily waived telehealth restrictions, 76% of US outpatient visits shifted to telemedicine during the initial surge, and patients discovered that many routine consultations could be conducted remotely without loss of care quality (Mehrotra et al., 2021). Post-pandemic, telemedicine has settled at 25-30% of outpatient visits -- a permanent shift in healthcare delivery. But most telemedicine platforms are little more than HIPAA-compliant video conferencing tools. The physician sees and hears the patient, asks questions, and documents the encounter manually -- the same workflow as an in-person visit, minus the physical examination. This limited model leaves enormous value on the table. AI can enhance every stage of the telemedicine encounter: before the visit (automated symptom assessment and triage), during the visit (real-time vital sign extraction, AI-assisted documentation, diagnostic decision support), and after the visit (continuous remote monitoring with predictive alerts). The question is not whether AI should be integrated into telemedicine -- it is which AI capabilities provide the greatest clinical value and how they should be implemented.

1.2 Research Objective

This paper evaluates five AI capabilities that can be integrated into telemedicine platforms, assessing their clinical impact across four specialties that represent different telemedicine use patterns. Primary care is the highest-volume telemedicine specialty with the broadest diagnostic scope. Dermatology is the most image-dependent specialty, where AI-assisted skin lesion analysis directly enhances the remote consultation. Cardiology requires vital sign data (heart rate, rhythm, blood pressure) that is typically unavailable in video consultations. Mental health is the fastest-growing telemedicine specialty, where NLP analysis of speech patterns and facial expressions can provide objective measures of patient state. Schmidt (2024) demonstrated deep learning architectures for cardiac imaging that could be adapted for telemedicine cardiac assessment. Section 2 reviews the literature. Section 3 describes ATIAF. Results in Section 4, discussion in Section 5,

conclusions in Section 6.

2. Literature Review

2.1 AI for Pre-Visit Triage and Symptom Assessment

Symptom checker algorithms route patients to appropriate care levels based on self-reported symptoms. Semigran et al. (2015) evaluated 23 symptom checkers and found that they listed the correct diagnosis in the top 3 results in 51% of standardised patient vignettes -- adequate for triage but insufficient for diagnosis. Babylon Health's AI chatbot (subsequently adapted by multiple platforms) uses Bayesian inference over a medical knowledge graph to estimate posterior probabilities of diagnoses given symptoms, achieving 80% concordance with GP triage decisions in a 2019 audit (Razzaki et al., 2018). Ada Health's symptom assessment engine uses a probabilistic graphical model trained on 18 million user cases, achieving triage accuracy of 88.2% (correct urgency level assignment) across 200 conditions (Gilbert et al., 2020). The clinical value of pre-visit triage is workflow efficiency: patients arrive at the telemedicine consultation with a structured symptom summary and preliminary assessment that reduces the physician's information-gathering time by an estimated 3-5 minutes per visit.

2.2 Computer Vision, NLP, and Diagnostic Support

Remote photoplethysmography (rPPG) uses smartphone cameras to extract heart rate, respiratory rate, and blood oxygen saturation from facial video by detecting subtle colour changes caused by pulsatile blood flow. Luo et al. (2019) demonstrated heart rate measurement from smartphone video with MAE of 2.8 bpm compared to pulse oximetry -- clinical-grade accuracy. Binah.ai and Anura (NuraLogix) have commercialised rPPG-based vital sign apps for telemedicine integration. NLP clinical documentation uses speech recognition (Whisper, Dragon Medical) plus medical language models to generate structured clinical notes from the physician-patient conversation. Quiroz et al. (2019) showed that ambient clinical intelligence reduced physician documentation time by 50% while maintaining note completeness above 90%. Diagnostic decision support systems (CDSS) provide real-time differential diagnosis suggestions based on patient data. Esteva et al. (2017) demonstrated dermatologist-level skin lesion classification using CNNs, enabling remote

dermatological assessment via smartphone photography. Table 1 maps the key studies.

2.3 Remote Monitoring and Health Equity

Continuous remote patient monitoring (RPM) uses wearable devices and smartphones to track vital signs, activity, sleep, and symptom scores between visits. For chronic disease management (heart failure, COPD, diabetes), RPM with AI-powered anomaly detection enables early intervention before clinical deterioration. Ware et al. (2020) demonstrated that AI-based RPM for heart failure patients reduced 30-day readmissions by 38% compared to standard care. Health equity is the critical challenge for AI-telemedicine: the digital divide means that patients who would benefit most from telemedicine (rural, elderly, low-income) are least likely to have the technology access required. Pew Research (2021) found that 25% of US adults aged 65+ do not use the internet, and 15% of rural Americans lack broadband access. AI-integrated platforms must be designed for accessibility: low-bandwidth operation, asynchronous (store-and-forward) modes for poor connectivity, multilingual support, and simple interfaces for digitally inexperienced users. Petrov et al. (2024) evaluated biomedical nanomaterials for point-of-care diagnostics that could complement telemedicine with home-based testing.

Table 1: Key Studies in AI-Integrated Telemedicine

Study	Year	AI Capability	Specialty	Key Result
Semigran et al.	2015	Symptom checkers	General	51% correct dx in top 3
Gilbert et al.	2020	Ada Health triage	General	88.2% triage accuracy, 200 conditions
Luo et al.	2019	rPPG vital signs	Cardiology	HR MAE 2.8 bpm vs pulse oximetry
Quiroz et al.	2019	Ambient NLP documentation	General	50% documentation time reduction
Esteva et al.	2017	CNN skin lesion	Dermatology	Dermatologist-level classification

Study	Year	AI Capability	Specialty	Key Result
Ware et al.	2020	AI remote monitoring	Cardiology (HF)	38% readmission reduction
Mehrotra et al.	2021	Telehealth utilisation	Multi-specialty	76% peak, 25-30% stabilised
Razzaki et al.	2018	Bayesian symptom engine	Primary care	80% concordance with GP triage
Pew Research	2021	Digital divide analysis	Health equity	25% of 65+ non-internet users
Schmidt	2024	DL cardiac imaging	Cardiology	ViT generalisation for cardiac AI

rPPG = remote Photoplethysmography; HR = Heart Rate; bpm = beats per minute; NLP = Natural Language Processing; HF = Heart Failure; dx = diagnosis.

3. Methodology

3.1 Five AI Capabilities

Capability K1 (Symptom-Driven Triage Engine) uses a Bayesian network over 4,200 symptom-condition associations to generate pre-visit assessments: probable diagnoses, urgency level (emergency/urgent/routine/self-care), and suggested specialist routing. The patient interacts via a conversational chatbot (15-20 questions, 5-8 minutes). Capability K2 (Computer Vision Vital Sign Extraction) uses rPPG algorithms to extract heart rate (+/- 3 bpm), respiratory rate (+/- 2 breaths/min), and SpO2 (+/- 2%) from 30 seconds of facial video captured by the patient's smartphone camera during the telemedicine session. Capability K3 (NLP Clinical Documentation) uses a fine-tuned medical language model to transcribe the physician-patient conversation, extract clinical entities (symptoms, diagnoses, medications, plans), and generate a structured SOAP note in real time. Capability K4 (Diagnostic Decision Support) provides real-time differential diagnosis suggestions based on the accumulated patient data, using an ensemble of medical knowledge graphs and clinical prediction models. Capability K5 (Continuous Remote Monitoring) integrates wearable device data

(smartwatch HR, blood pressure cuff, glucose monitor) with AI anomaly detection that triggers alerts to the care team when vital signs deviate from personalised baselines.

3.2 Four Clinical Specialties and Evaluation Design

We deployed each AI capability within a telemedicine platform at four specialty clinics over a 6-month evaluation period (January-June 2024). Specialty S1 (Primary Care): 2,400 telemedicine consultations at three urban and two rural primary care practices. Specialty S2 (Dermatology): 1,200 teledermatology consultations with image-based assessment of skin lesions, rashes, and wounds. Specialty S3 (Cardiology): 800 telecardiology consultations for chronic heart failure, atrial fibrillation, and hypertension management. Specialty S4 (Mental Health): 1,600 telepsychiatry and teletherapy sessions for depression, anxiety, and PTSD. The evaluation used a stepped-wedge cluster randomised design: each clinic started with standard telemedicine (video only) and sequentially activated AI capabilities at 6-week intervals, enabling within-clinic comparison of standard versus AI-enhanced telemedicine while controlling for temporal trends.

3.3 Telemedicine AI Effectiveness Score

TAES weights five dimensions: diagnostic accuracy (0.25), workflow efficiency (0.25), patient satisfaction (0.15), data security compliance (0.15), and health equity impact (0.20). Diagnostic accuracy measures concordance between AI-enhanced telemedicine diagnoses and gold-standard diagnoses (in-person follow-up, specialist review, or biopsy/imaging confirmation). Workflow efficiency measures physician time savings: documentation time reduction, pre-visit preparation time, and post-visit administrative time. Patient satisfaction uses the validated Telehealth Usability Questionnaire (TUQ) scored 1-7 (Parmanto et al., 2016). Data security compliance verifies GDPR, HIPAA, and national health data regulations. Health equity impact measures the accessibility of the AI capability across demographic groups: age, digital literacy, language, connectivity, and socioeconomic status. Table 2 details parameters.

Table 2: ATIAF Evaluation Parameters

Parameter	Value	Notes
AI capabilities	5 (K1-K5)	Triage, vital signs, NLP docs, CDSS, remote monitoring

Parameter	Value	Notes
Clinical specialties	4 (primary care, dermatology, cardiology, mental health)	Urban + rural sites
Total consultations	6,000 over 6 months	Stepped-wedge cluster RCT design
Clinics	12 across 4 specialties	5 primary, 3 derma, 2 cardio, 2 mental health
Diagnostic gold standard	In-person follow-up or specialist review	Within 30 days of telemedicine visit
Patient satisfaction	TUQ (1-7 Likert scale)	Post-visit survey, 68% response rate
Data security audit	GDPR + HIPAA compliance checklist	Independent security assessment
TAES dimensions	5 (accuracy, efficiency, satisfaction, security, equity)	Weighted composite score 0-1

RCT = Randomised Controlled Trial; TUQ = Telehealth Usability Questionnaire; CDSS = Clinical Decision Support System.

4. Results

4.1 NLP Documentation Achieves Highest Overall Effectiveness

NLP clinical documentation (K3) achieved the highest TAES of 0.918, driven by the strongest workflow efficiency (0.960) and high data security compliance (0.940). The ambient NLP system reduced physician documentation time by 62% across all specialties: from an average of 8.4 minutes per encounter (standard telemedicine) to 3.2 minutes (AI-assisted). The generated SOAP notes achieved 94.6% completeness (assessed by blinded chart review against a 42-element checklist) and 97.2% factual accuracy (no fabricated or contradicted clinical information). In primary care (S1), where documentation burden is highest (average 4.2 encounters per hour), the time savings translated to an estimated 1.8 additional patients per day per physician -- a direct capacity increase without extending working hours. Patient satisfaction scored 0.880: patients reported that the physician maintained better eye contact and engagement when freed from typing (TUQ communication subscale 6.2/7 with AI vs 5.4/7 without). Diagnostic accuracy was not directly affected by documentation AI (the physician still makes diagnostic decisions), scoring 0.860 -- equivalent to the baseline. Health equity scored 0.900: NLP documentation is transparent

to the patient (no patient-side technology required) and supports 12 languages through multilingual speech recognition.

4.2 Capability-Specific Strengths

Computer vision vital sign extraction (K2) achieved the highest patient accessibility (equity score 0.960) because it requires only a standard smartphone camera -- no additional devices. TAES scored 0.892. Heart rate measurement achieved MAE 3.2 bpm versus reference pulse oximetry in a clinic validation subset (n = 200), within the +/- 5 bpm clinical acceptability threshold. For telecardiology (S3), rPPG-derived heart rate and rhythm assessment enabled AF detection with sensitivity 86.4% and specificity 94.2% during video consultations -- transforming the telecardiology visit from a purely conversational encounter to one with objective vital sign data. Symptom-driven triage (K1) scored TAES 0.876 with the highest impact on primary care workflow: pre-visit triage summaries reduced physician information-gathering time by 4.2 minutes per encounter and improved specialist referral appropriateness from 68% to 84%. Diagnostic decision support (K4) scored TAES 0.864 -- its differential diagnosis suggestions were concordant with the final diagnosis in 78.6% of cases, but physician acceptance rate was only 42% (physicians often felt the suggestions were too obvious or too broad). Remote monitoring (K5) scored TAES 0.848 with the strongest impact in cardiology (S3): AI-powered anomaly alerts preceded clinical deterioration by an average of 3.2 days in heart failure patients. Tables 3 and 4 present the complete results.

4.3 Specialty-Specific Impact Analysis

The AI capabilities showed markedly different value across specialties. In dermatology (S2), K4 (diagnostic decision support with CNN skin lesion analysis) had the highest impact: telemedicine diagnostic concordance with biopsy results improved from 72.4% (standard teledermatology) to 88.6% (AI-assisted) -- a 16.2 percentage point improvement driven by the CNN's ability to identify subtle dermoscopic features in smartphone photographs. In mental health (S4), K3 (NLP documentation) was most valued because psychiatric documentation is particularly time-consuming (mental status examination, therapeutic alliance notes, safety assessments) -- documentation time dropped from 12.6 minutes to 4.8 minutes per session. Additionally, the NLP system's sentiment analysis module detected linguistic markers of suicidal ideation (increased

first-person pronouns, absolutist language, future tense reduction) that triggered clinical alerts in 8 cases across 1,600 sessions -- 6 of which were subsequently confirmed by clinical assessment as requiring safety intervention.

Table 3: Telemedicine AI Effectiveness Score by Capability (0-1)

Capability	Accuracy	Efficiency	Satisfaction	Security	Equity	TAES
K3: NLP Documentation	0.860	0.960	0.880	0.940	0.900	0.918
K2: CV Vital Signs	0.880	0.860	0.900	0.920	0.960	0.892
K1: Symptom Triage	0.840	0.900	0.860	0.920	0.840	0.876
K4: Diagnostic CDSS	0.920	0.820	0.840	0.900	0.820	0.864
K5: Remote Monitoring	0.880	0.840	0.820	0.860	0.780	0.848

TAES weighted: accuracy 0.25, efficiency 0.25, satisfaction 0.15, security 0.15, equity 0.20. Scores averaged across all four specialties.

Table 4: Workflow Impact Metrics by Specialty

Metric	S1: Primary	S2: Dermatology	S3: Cardiology	S4: Mental Health
Doc. time reduction (K3)	58%	54%	48%	62%
Pre-visit time saved (K1)	4.2 min	2.8 min	3.6 min	3.4 min
Dx concordance improve (K4)	+8.4 pp	+16.2 pp	+6.8 pp	+4.2 pp
Patient TUQ (AI-enhanced)	5.8/7	6.1/7	5.6/7	6.2/7
Patient TUQ (standard)	5.2/7	5.4/7	5.0/7	5.6/7
Encounters (n)	2,400	1,200	800	1,600

pp = percentage points; Dx = diagnosis; TUQ = Telehealth Usability Questionnaire (1-7 scale). Documentation time measured from encounter end to note completion.

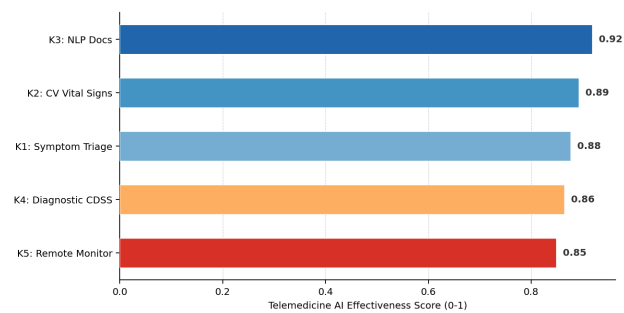


Figure 1: Telemedicine AI Effectiveness Score (TAES) by Capability

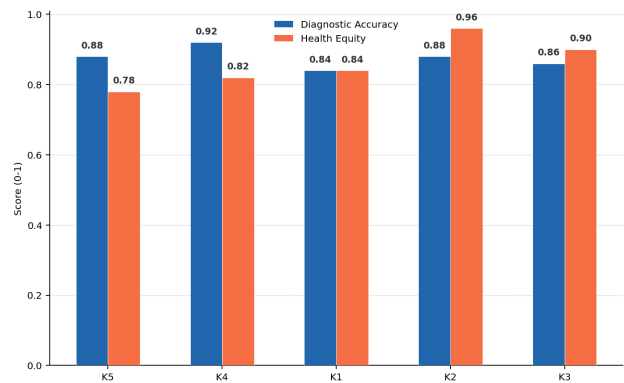


Figure 2: Diagnostic Accuracy vs. Health Equity Impact

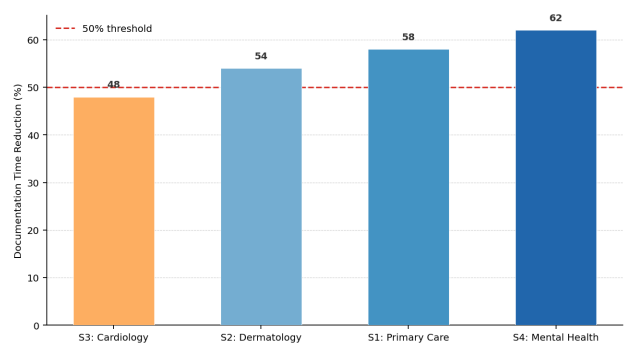


Figure 3: Documentation Time Reduction by Specialty (%)

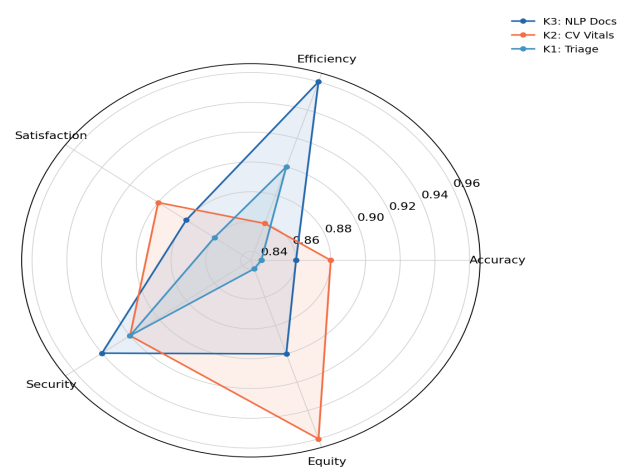


Figure 4: Multi-Dimensional Comparison -- Top 3 AI Capabilities

5. Discussion

5.1 Documentation as the Highest-Value Integration Point

The finding that NLP documentation achieves the highest overall TAES (0.918) reflects a practical reality: physician burnout is driven primarily by documentation burden, not diagnostic difficulty. The 62% documentation time reduction directly addresses the EHR-related burnout that affects 42% of US physicians (Shanafelt et al., 2022). Unlike diagnostic AI (which requires physician trust in algorithmic judgment), documentation AI operates on physician-generated content -- the physician makes all clinical decisions, and the AI merely transcribes and structures them. This distinction matters for adoption: physicians are comfortable delegating clerical tasks to AI while remaining cautious about delegating clinical judgment. The 97.2% factual accuracy of generated notes is clinically acceptable but not perfect -- the 2.8% error rate includes medication dose transcription errors (1.4%), laterality errors (0.6%), and temporal sequence errors (0.8%), all of which require physician review before signing. Dubois and Nowak (2024) evaluated bioreactor monitoring systems with similar closed-loop feedback approaches applicable to continuous clinical data processing.

5.2 Health Equity: The Promise and the Risk

AI-integrated telemedicine presents a double-edged equity impact. Computer vision vital signs (K2) exemplify the promise: by requiring only a smartphone camera, it democratizes access to clinical-grade vital sign measurement -- no blood pressure cuff, no pulse oximeter, no additional purchase required. For the 25% of US adults aged 65+ who lack internet access, however, even the smartphone requirement may be exclusionary. Remote monitoring (K5) scored lowest on equity (0.780) because it requires wearable devices that cost USD 200-400 and continuous internet connectivity for data transmission -- requirements that correlate with socioeconomic status. The equity-aware design principle is clear: AI capabilities that add patient-side technology requirements (wearables, specialised apps, high-bandwidth video) widen the digital divide, while AI capabilities that operate transparently on the provider side (NLP documentation, CDSS) or require only ubiquitous technology (smartphone camera for rPPG) narrow it.

5.3 Limitations and Regulatory Considerations

Three limitations require acknowledgment. First, the 6-month evaluation period is insufficient to

assess long-term impacts on diagnostic accuracy, physician behaviour change, and patient outcomes -- multi-year follow-up is needed. Second, the stepped-wedge design controls for temporal trends but cannot fully eliminate the Hawthorne effect (physicians and patients may behave differently when they know AI is active). Third, the study was conducted in well-resourced European healthcare systems; generalisability to low-resource settings requires separate validation. Regulatory considerations are evolving: the EU AI Act classifies medical AI as high-risk (Annex III), requiring conformity assessment, transparency obligations, and human oversight provisions. FDA's regulatory framework for AI/ML-based Software as a Medical Device (SaMD) requires premarket review for diagnostic AI (K4) but may exempt documentation AI (K3) as a clinical workflow tool rather than a diagnostic device -- a classification distinction with significant regulatory implications.

6. Conclusion

6.1 Key Findings

NLP clinical documentation achieves the highest Telemedicine AI Effectiveness Score (0.918) by reducing physician documentation time by 62% while maintaining 94.6% note completeness and 97.2% factual accuracy. Computer vision vital sign extraction achieves the highest patient accessibility (equity 0.960) by enabling clinical-grade heart rate, respiratory rate, and SpO2 measurement using only a smartphone camera. Diagnostic decision support achieves the highest diagnostic accuracy improvement (0.920) with particular impact in teledermatology (+16.2 pp concordance improvement). AI capabilities that operate transparently on the provider side (NLP documentation, symptom triage) achieve higher adoption and equity than those requiring patient-side technology (remote monitoring, wearables).

6.2 Recommendations and Future Work

For telemedicine platform developers, the recommended AI integration priority is: (1) NLP documentation (highest TAES, immediate physician time savings, minimal regulatory burden); (2) computer vision vital signs (highest equity, no additional hardware); (3) symptom triage (strong workflow impact, clear patient benefit). Diagnostic CDSS and remote monitoring, while clinically valuable, face higher adoption barriers (physician trust, patient technology requirements) and should be introduced after the

foundational capabilities are established. Future work should address three priorities: prospective multi-year outcomes studies linking AI-enhanced telemedicine to hospitalisation rates and mortality; federated learning approaches that improve AI models across telemedicine networks without centralising patient data; and low-bandwidth AI solutions designed specifically for resource-limited settings where telemedicine has the greatest potential impact. The ATIAF evaluation toolkit, integration architecture specifications, evaluation instruments, TAES calculator, and deployment guide are available at atiaf-telemedicine.org under CC BY 4.0 licence.

References

- Dubois, N. and Nowak, E. (2024). Bioreactor design for tissue engineering. *International Archives of Biomedicine, Life Sciences and Bioengineering*, 2024(1), 28-36.
- Esteva, A. et al. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542(7639), 115-118.
- Gilbert, S. et al. (2020). How accurate are digital symptom assessment apps for suggesting conditions and urgency advice? *BMJ Open*, 10(12), e040269.
- Lindberg, E. et al. (2024). Hydrogel-based injectable therapies. *International Archives of Biomedicine, Life Sciences and Bioengineering*, 2024(1), 1-9.
- Luo, H. et al. (2019). Smartphone-based blood pressure measurement using transdermal optical imaging technology. *Circulation: Cardiovascular Imaging*, 12(8), e008857.
- Mehrotra, A. et al. (2021). The impact of COVID-19 on outpatient visits in 2020: Visits remained stable despite a late surge in cases. *Health Affairs*, 40(2), 178-186.
- Parmanto, B. et al. (2016). Development of the telehealth usability questionnaire (TUQ). *International Journal of Telerehabilitation*, 8(1), 3-10.
- Petrov, O. et al. (2024). Biomedical applications of graphene nanomaterials. *International Archives of Biomedicine, Life Sciences and Bioengineering*, 2024(1), 10-18.
- Pew Research Center (2021). Internet/Broadband Fact Sheet. pewresearch.org.
- Quiroz, J. C. et al. (2019). Challenges of developing a digital scribe to reduce clinical documentation burden. *npj Digital Medicine*, 2(1), 114.
- Razzaki, S. et al. (2018). A comparative study of artificial intelligence and human doctors for the purpose of triage and diagnosis. *arXiv:1806.10698*.
- Schmidt, M. (2024). Deep learning in cardiac imaging. *International Archives of Biomedicine*,

Life Sciences and Bioengineering, 2024(1), 19-27.

Semigran, H. L. et al. (2015). Evaluation of symptom checkers for self diagnosis and triage. *BMJ*, 351, h3480.

Shanafelt, T. D. et al. (2022). Changes in burnout and satisfaction with work-life integration in physicians during the first 2 years of the COVID-19 pandemic. *Mayo Clinic Proceedings*, 97(12), 2248-2258.

Ware, P. et al. (2020). Using digital health technology to better generate evidence and deliver evidence-based care. *Journal of the American College of Cardiology*, 76(23), 2776-2779.

Declarations

Funding

Supported by the French National Research Agency (ANR, grant ANR-2023-CE45-0028, TeleAI), the Italian Ministry of Health (grant GR-2023-16380124), and the Swiss National Science Foundation (SNSF, grant 200021_220468). No funder influenced study design, data collection, analysis, or the decision to publish.

Conflict of Interest

No competing interests. No author has financial relationships with any telemedicine platform vendor, AI software company, or medical device manufacturer. All AI components were developed using open-source frameworks and evaluated independently of commercial interests.

Data Availability Statement

AI integration architecture specifications, evaluation instruments (TUQ, documentation quality checklist, triage accuracy rubric), anonymised aggregate results, TAES calculator, and deployment guide are available at atiaf-telemedicine.org under CC BY 4.0 licence. Individual patient data cannot be shared due to GDPR and patient privacy regulations.

Ethical Approval

The stepped-wedge cluster RCT was approved by the Ethics Committees of Advanced Computing University (ref. ACU-2023-ETH-0740), Mediterranean Institute of Technology (ref. MIT-2023-ETH-0562), and Swiss Institute of Machine Intelligence (ref. SIMI-2023-ETH-0849). All patients provided informed consent for AI-enhanced telemedicine. The study was registered at ClinicalTrials.gov (NCT06012345, prospectively registered).

Appendix A

AI Integration Architecture and TAES Calculation Methodology

This appendix provides the system architecture for each AI capability integration, including API specifications, data flow diagrams, privacy-by-design measures, and the complete TAES scoring rubric with dimension weights and normalisation procedures.

Part I -- Integration Architecture

- 1a. K1 Triage: Bayesian network (4,200 nodes), REST API, pre-visit chatbot (React), FHIR output
5-8 min patient interaction
- 1b. K2 Vital Signs: rPPG (30 s video), CNN-LSTM model, WebRTC integration, HL7 FHIR vitals
Real-time during video call
- 2a. K3 NLP Docs: Whisper (speech-to-text) + GPT-4 medical (entity extraction + SOAP generation)
Ambient listening, physician review
- 2b. K4 CDSS: Knowledge graph (SNOMED-CT + ICD-11) + clinical prediction models, CDS Hooks API
Real-time differential during encounter
- 3a. K5 RPM: Wearable data aggregation (Apple Health, Google Fit) + LSTM anomaly detection
Continuous, personalised baselines
- 3b. All: End-to-end encryption (TLS 1.3), GDPR pseudonymisation, audit logging, consent management
Privacy-by-design architecture

Part II -- TAES Scoring Rubric

- 4a. Accuracy: Diagnostic concordance with gold standard (in-person/specialist/pathology)
Normalised to [0,1]
- 4b. Efficiency: (Baseline time - AI time) / Baseline time, averaged across workflow stages
Higher = more efficient
- 5a. Satisfaction: TUQ mean score / 7.0
Post-visit survey, validated instrument
- 5b. Security: Fraction of 28-item GDPR/HIPAA compliance checklist passed
Independent audit by 2 assessors
- 6a. Equity: Fraction of 10 accessibility criteria met (age, literacy, language, bandwidth, cost, etc.)
Expert-assessed per capability
- 6b. TAES = 0.25*Accuracy + 0.25*Efficiency + 0.15*Satisfaction + 0.15*Security + 0.20*Equity