Review

Artificial Intelligence Tools in Pre-Travel Health Consultations: A Scoping Review of Clinical Evidence, Implementation Gaps, and Emerging Opportunities

Haider Saddam Qasim * and Maree Donna Simpson

Posted Date: 13 May 2026

doi: 10.20944/preprints202605.0838.v1

Keywords: travel medicine; generative AI; pre-travel consultation; risk assessment; structured data

 doi:10.20944/preprints202605.0838.v1

Review

Artificial Intelligence Tools in Pre-Travel Health Consultations: A Scoping Review of Clinical Evidence, Implementation Gaps, and Emerging Opportunities

Running head: AI in Pre-Travel Consultations

Haider Saddam Qasim * and Maree Donna Simpson

School of Medical Sciences and Dentistry, Charles Sturt University, Orange Campus, NSW, Australia

* Correspondence: haider.qasim@hdr.qut.edu.au

Abstract

Background Pre-travel health consultations require individualised risk assessment across itinerary, destination epidemiology, traveller characteristics, vaccine history, comorbidities, medication profile, pregnancy status, immune status, activities, timing, and access to care [8,10,11]. Artificial intelligence (AI), particularly large language models (LLMs), may support pre-consultation education, structured history collection, guideline retrieval, multilingual communication, and post-consultation reinforcement, but unsafe use may introduce hallucinated, outdated, or insufficiently personalised recommendations [5,6,14,15]. **Objectives** This scoping review maps the current evidence on AI tools relevant to pre-travel health consultations, characterises implementation gaps, identifies patient-safety risks, and proposes a supervised implementation model for travel medicine clinics [1,28–30]. **Methods** The review was conducted as a scoping review using the Arksey and O'Malley framework as advanced by Levac and colleagues and operationalised through the JBI scoping review guidance, with reporting aligned to the PRISMA Extension for Scoping Reviews (PRISMA-ScR) [1,28–30]. The review was not prospectively registered. Eligibility was defined by a Population–Concept–Context (PCC) framework. Targeted retrieval was conducted in May 2026 through PubMed/MEDLINE (one direct search string), academic and web-indexed search tools, citation chasing from Journal of Travel Medicine and Travel Medicine and Infectious Disease, and authoritative guideline and regulator websites. The search date range was January 2017 to May 2026. Sources were eligible if they addressed AI or digital decision support in pre-travel health, travel medicine, travel-related clinical decision support, clinical LLM safety, or guidance defining the standard pre-travel consultation. Screening and data charting were conducted by a single reviewer using a structured eligibility checklist (Supplement S2). **Results** Seventy records were identified, one duplicate was removed, 69 records were screened, 12 reports were sought for retrieval, one record could not be retrieved within the search window, and 11 reports were assessed in full text and included in the synthesis. Included sources comprised four direct pre-travel AI sources, one travel-related decision-support study, four guideline and context sources, and two clinical LLM safety sources. Direct evidence is thin: the only patient-level implementation report involved 26 travellers using a GPT-4 Travel Clinic Assistant in a Singapore tertiary travel clinic, where physicians and travellers reported acceptability and workflow benefit but objective effectiveness outcomes were not measured [3]. A ChatGPT pre-travel advice evaluation found generally readable and comprehensive answers to common questions, but responses lacked sufficient personalisation to itinerary, comorbidity, vaccine history, and cost considerations [2]. Broader clinical LLM evidence indicates that evaluation methods remain heterogeneous and that LLMs may repeat or elaborate false clinical details and hallucinate clinical guidelines in simulated decision-support tasks [13,14,16]. **Conclusions** Current evidence supports supervised AI augmentation of pre-travel consultations but does not support autonomous AI-led vaccine selection, malaria prophylaxis, contraindication screening, or individualised travel-risk

 doi:10.20944/preprints202605.0838.v1

clearance [2–6,14,15,48,50]. Near-term deployment should be restricted to clinician-supervised education, structured intake, source-grounded guideline retrieval, after-visit reinforcement, and escalation-triggered workflow support [4,5,34,49]. Travel medicine specialists, clinic leaders, regulators, and digital health developers should prioritise domain-specific hallucination audits, equity testing across visiting friends and relatives, migrant, older-adult, First Nations Australian, and Pacific Islander travellers, and prospective trials reported under CONSORT-AI, SPIRIT-AI, and TRIPOD+AI standards [31–33,37,38,41–44].

Keywords: travel medicine; generative AI; pre-travel consultation; risk assessment; structured data

Introduction

Pre-travel consultation is a preventive clinical encounter that synthesises traveller characteristics, destination-specific risks, itinerary, trip duration, purpose of travel, activities, timing, medical history, medications, immunisation history, and special risk states such as pregnancy or immunocompromise [8]. Core interventions include immunisation, malaria chemoprophylaxis, risk communication, food and water safety advice, vector avoidance, injury prevention, respiratory infection precautions, blood-borne infection prevention, and sexual-health counselling [8,9].

The International Society of Travel Medicine emphasises that pre-travel care is broader than vaccination alone and should account for destination, activities, age, health status, chronic disease, past vaccines, medication requirements, and timely advice approximately six weeks before travel where possible [10]. WHO travel-health guidance similarly highlights malaria geography, chemoprophylaxis, mosquito personal protection, and treatment considerations as central components of international travel health [11].

Pre-travel consultation uptake remains uneven globally. Surveys in primary care and migrant communities identify perceived low risk, awareness gaps, time and cost barriers, and language barriers as the principal obstacles to pre-travel care, with visiting friends and relatives (VFR) travellers and migrant populations particularly underserved [37–41]. Primary care physicians, who manage many travellers, often report low travel-medicine knowledge and limited consultation time [40]. These structural gaps create the context in which AI may plausibly assist, provided it does not replicate or worsen access inequities.

AI is rapidly transforming clinical practice across diagnostic imaging, decision support, health monitoring, and patient engagement [36]. Travel medicine appears attractive for AI augmentation because it is guideline-intensive, time-sensitive, multilingual, and highly dependent on structured information [5,6]. The GeoSentinel surveillance network provides a globally distributed data substrate that could plausibly support AI-enabled outbreak detection and destination-specific risk modelling [7,51]. However, the same features that make travel medicine attractive for AI also make unsafe automation problematic. A generic answer that omits splenectomy, live-vaccine contraindications, pregnancy, immunosuppression, HIV status, transplant history, or region-specific malaria variation could create clinically meaningful harm [8,11,14,15].

This review therefore asks what AI tools have been evaluated in or near pre-travel health consultations, what the current evidence permits clinicians to conclude, what patient-safety and implementation gaps remain, and how travel clinics might deploy AI without replacing specialist clinical judgement.

Methods

Review Design

This review uses scoping review methodology, which is appropriate when the goal is to map the extent, range, and nature of evidence in an emerging and heterogeneous field rather than estimate intervention effects or conduct meta-analysis [1,28–30]. The review follows the Arksey and O'Malley

 doi:10.20944/preprints202605.0838.v1

framework as advanced by Levac and colleagues and operationalised through JBI guidance, with reporting aligned to PRISMA-ScR [1,28–30]. A completed PRISMA-ScR checklist with manuscript locations is provided as Supplement S1. The review was not prospectively registered.

Population, Concept, Context (PCC) Framework

The eligibility framework follows the JBI PCC structure for scoping reviews [30]. Table 1 defines the framework operationalised in this review.

Table 1. Population, Concept, and Context (PCC) framework for the scoping review.

PCC element	Operational definition	Examples
Population	International travellers receiving or seeking pre-travel health advice; clinicians providing pre-travel care; and study cohorts within general clinical AI safety research where the findings are mapped to pre-travel decision support	Adult and paediatric international travellers, VFR travellers, migrant travellers, immunocompromised travellers, primary care physicians, travel medicine specialists, simulated patient cohorts in clinical LLM studies
Concept	AI tools, large language models, chatbots, retrieval-augmented generation, and clinical decision support systems applied to pre-travel risk assessment, education, intake, recommendation, escalation, or after-visit reinforcement; clinical AI safety, hallucination, and reporting standards	ChatGPT and GPT-4-based pre-travel assistants, custom GPT prototypes, tablet-based travel CDSS, generative AI educational outputs, RAG architectures, hallucination and accuracy audits
Context	International, multilingual, ambulatory pre-travel and travel-related clinical settings; primary care, specialist travel clinics, university-affiliated travel medicine services; relevant guideline, regulatory, and equity contexts	High-, middle-, and low-income settings; United States, Australia, Singapore, Switzerland, Italy, EU; CDC, WHO, ISTM guidance; FDA, TGA, EU AI Act regulatory frameworks

Eligibility Criteria

Sources were eligible if they met one or more of the following criteria, mapped to the PCC framework:

 doi:10.20944/preprints202605.0838.v1

- AI or LLM tools used for pre-travel health advice, pre-travel counselling, travel clinic education, or travel-medicine workflow support [2–7,12].
- AI chatbot design frameworks for travel medicine [4].
- Travel-related clinical decision support systems relevant to travel medicine workflows [12].
- Authoritative guidelines defining the standard of care for pre-travel consultation [8–11].
- General clinical LLM safety, hallucination, and evaluation literature directly relevant to clinical decision support [13–16,47,48,50].
- Adjacent preventive-counselling literature where it informed implementation design, equity, or evaluation standards [17,18,34–36,46,49,52].

Sources were excluded if they were unrelated to AI, unrelated to travel medicine or clinical decision support, non-healthcare AI papers, unrelated travel behaviour studies without AI relevance, or purely technical AI papers without clinical or implementation relevance.

Search Strategy

This review used a transparent, targeted scoping retrieval rather than a fully reproducible multi-database systematic search. Retrieval was conducted in May 2026 over the date range January 2017 to May 2026 using the following components, all documented in Supplement S2:

- One direct PubMed/MEDLINE search string combining travel-medicine and AI/LLM concepts.
- Academic and web-indexed search tools applied to the same concept set, including searches restricted to authoritative travel-medicine and clinical AI domains.
- Citation chasing on the four direct travel-medicine AI sources [2–7] and on the FeverTravelApp study [12].
- Targeted retrieval from authoritative guideline and regulator websites (CDC, WHO, ISTM, FDA, TGA, European Commission/EU AI Act portal) [8–11,22–27,41].
- Hand retrieval from Journal of Travel Medicine, Travel Medicine and Infectious Disease, BMC Digital Health, and Communications Medicine for AI-relevant titles within the search window.

The PubMed/MEDLINE string used was:

("travel medicine"[Title/Abstract] OR "pre-travel"[Title/Abstract] OR "pretravel"[Title/Abstract] OR "travel health"[Title/Abstract] OR "traveller health"[Title/Abstract]) AND ("artificial intelligence"[Title/Abstract] OR "large language model"[Title/Abstract] OR "ChatGPT"[Title/Abstract] OR chatbot*[Title/Abstract] OR "clinical decision support"[Title/Abstract] OR "machine learning"[Title/Abstract])

Complementary search strings for Embase, CINAHL, Cochrane CENTRAL, IEEE Xplore, ACM Digital Library, ClinicalTrials.gov, and WHO ICTRP are documented in Supplement S2 as recommended reproducibility strategies. These additional databases were not run for this scoping retrieval. This is recorded transparently as a limitation rather than presented as completed work.

Study Selection and Data Charting

Screening and data charting were conducted by a single reviewer (the manuscript author) using a structured eligibility checklist. Because screening was conducted by a single reviewer, Cohen's kappa for inter-rater agreement was not calculable. Mitigation strategies included a written eligibility checklist (Supplement S2), explicit decision rules for ambiguous cases, and second-pass review of borderline records by the same reviewer at least 24 hours after first pass. Charted fields included article type, country and setting, AI/tool type, sample size, comparator or reference standard, outcomes, findings, safety concerns, implementation implications, and relevance to specific pre-travel consultation tasks. Single-reviewer screening is a known threat to validity in scoping reviews and is acknowledged as a limitation [1,28–30]. The full charting table for the 11 included sources is provided as Supplement S4.

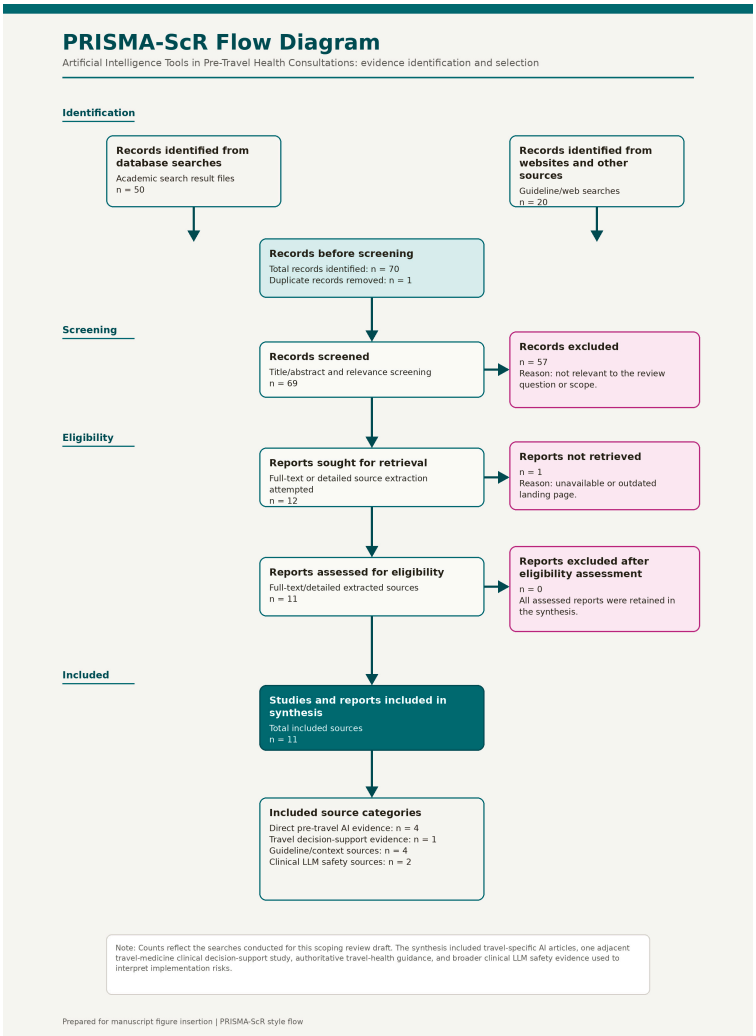
 doi:10.20944/preprints202605.0838.v1

Quality Appraisal and Certainty Assessment

Scoping reviews do not require formal risk-of-bias assessment, but an indicative quality and applicability appraisal was performed to inform interpretation and certainty rather than inclusion [1,30]. No source was excluded on the basis of quality appraisal. The Mixed Methods Appraisal Tool (MMAT) was used conceptually for empirical mixed or quantitative implementation evidence because it covers qualitative, randomised, non-randomised, quantitative descriptive, and mixed-methods studies [19]. AMSTAR 2 principles were used to interpret systematic reviews [20]. GRADE categories were used cautiously as "GRADE-informed certainty" because GRADE reflects confidence that observed evidence approximates the true effect and considers risk of bias, inconsistency, indirectness, imprecision, and publication bias [21]. Editorials and guideline texts were appraised using JBI text-and-opinion guidance [30].

Study Selection Flow

The search identified 70 records across PubMed/MEDLINE, academic and web-indexed sources, citation chasing, and authoritative guideline and regulator retrieval. After removal of one duplicate, 69 records were screened; 57 were excluded as outside the review question or scope, 12 reports were sought for retrieval, one report could not be retrieved within the search window, and 11 reports were assessed in full text and included in the synthesis. Included evidence comprised four direct pre-travel AI sources [2–7], one travel-related decision-support source [12], four guideline and context sources [8–11], and two clinical LLM safety sources [13,14]. Detailed exclusion subcategories are shown in Supplement S3 and are reflected in the figure caption below as post hoc reconstructed categories from the screening notes.



 doi:10.20944/preprints202605.0838.v1

Figure 1. PRISMA-ScR style flow diagram summarising identification, screening, eligibility assessment, and included source categories for the scoping review of AI tools in pre-travel health consultations [1]. The 57 records excluded at title and abstract screening were grouped into four post hoc subcategories reconstructed from screening notes: (i) AI in non-travel clinical or non-clinical contexts (approximately half of exclusions); (ii) travel behaviour or travel epidemiology without AI or decision-support relevance; (iii) general AI methodology without clinical implementation content; and (iv) duplicate or near-duplicate coverage of an already-included source. One record was sought but not retrieved within the search window and is recorded in Supplement S3. No reports were excluded after full-text assessment.

Results

Confidence in the Evidence Base

A central finding of this review is the thinness of direct evidence. Only four sources directly address AI in travel medicine, and only one of these is a patient-level implementation study, in 26 travellers in a single Singapore tertiary travel clinic [3]. Four levels of evidence informed the synthesis with descending confidence in direct applicability:

1. **Direct travel-medicine AI evidence (four sources):** the ChatGPT pre-travel advice evaluation [2], the Singapore Travel Clinic Assistant implementation [3], the Baglivo decalogue prototype [4], and the Flaherty editorials on supervised generative AI integration and the natural history of AI in travel medicine [5,6]. Heidema et al. is treated as adjacent rather than direct because it concerns AI-supported outbreak surveillance rather than the pre-travel consultation itself [7].
2. **Adjacent clinical AI evidence (eight sources):** clinical LLM evaluation methods [13]; multi-model hallucination and clinical guideline omission/hallucination assurance analyses [14,16]; clinical documentation hallucination framework [15]; ChatGPT meta-analysis [47]; ChatGPT care-seeking accuracy across model versions [48]; ChatGPT FAQ literature review [50]; and travel-related clinical decision support [12].
3. **Guideline, regulatory, and methodology evidence (twelve sources):** CDC Yellow Book pre-travel guidance and VFR chapter [8,9,41]; ISTM pre-travel advice [10]; WHO malaria travel guidance [11]; PRISMA-ScR, scoping methodology, MMAT, AMSTAR 2, GRADE [1,19–21,28–30]; WHO AI ethics and digital health strategy [26,27]; FDA, TGA, and EU AI Act materials [22–25]; and AI reporting standards CONSORT-AI, SPIRIT-AI, and TRIPOD+AI [31–33,45].
4. **Adjacent implementation, equity, and patient-engagement evidence (ten sources):** clinician adoption of AI [34]; AI in medical education [35]; AI in healthcare overview [36]; VFR uptake studies [37,38]; pre-travel consultation in primary care [39,40]; LLM patient education and chronic-illness chatbot reviews [17,18]; AHRQ healthcare chatbot review [46]; preventive-care chatbot outreach [52]; digital divide and equity [42–44]; and retrieval-augmented generation [49].

What is not known is therefore as important as what is known. There are no prospective travel-medicine AI trials reporting clinical safety or vaccine uptake outcomes, no published efficacy comparisons against standard travel-medicine consultation, no equity-focused trials in VFR, migrant, First Nations Australian, or Pacific Islander travellers, and no published commercial or open travel-medicine AI tools with regulatory approval as Software as a Medical Device for travel-medicine indications.

Evidence Base Overview

Four sources directly address AI in travel medicine: a ChatGPT pre-travel advice evaluation [2], a GPT-4 Travel Clinic Assistant implementation in Singapore [3], a proposed travel-health chatbot design decalogue [4], and editorials on integrating generative AI safely into travel medicine [5,6]. A separate travel-related clinical decision support study examined a tablet-based algorithm for

 doi:10.20944/preprints202605.0838.v1

returned febrile travellers [12], and an additional editorial-tier source describes AI-supported outbreak detection using GeoSentinel data as a near-future adjacent application [7].

Only one direct travel-medicine AI implementation report included actual travellers. In that report, 26 travellers used a GPT-4 Travel Clinic Assistant in a Singapore tertiary travel clinic, and physician and traveller feedback suggested acceptability, improved focus of consultation, and perceived knowledge benefit, but there was no objective measurement of knowledge gain, clinical appropriateness, vaccine uptake, prophylaxis adherence, adverse events, or downstream travel illness [3].

The broader evidence base supports caution. A systematic review of 761 clinical LLM evaluations found heterogeneous methods and limited standardisation of evaluation rubrics [13]. A multi-model assurance analysis showed LLMs repeated or elaborated planted false clinical details in 50 to 82 percent of simulated clinical prompts [14]. A clinical-documentation hallucination framework demonstrated that hallucinations are more likely than omissions to be classified as major errors, especially in the clinical plan section [15]. A diagnostic LLM analysis found measurable prevalence of clinical guideline omission and hallucination in LLM outputs, including fabricated protocols mimicking authoritative sources [16]. A literature review of ChatGPT in clinical FAQs, recommendations, and symptom interpretation found accuracy ranging from 20 to 95 percent across nine studies [50]. A meta-analysis of ChatGPT in medical and dental research described an 18 to 100 percent accuracy range across specialties [47]. An evaluation of ChatGPT model versions across 22 instances and multiple urgency levels found average accuracy near 70 percent, overtriage, and increasing variability with newer models [48].

Evidence Synthesis Table

Table 2 summarises the included sources, their methods, sample sizes, AI/tool types, the consultation tasks they map to in the CDC Yellow Book taxonomy [8], main findings, key safety concerns, and a GRADE-informed certainty rating with explicit reasoning [21].

Table 2. Evidence synthesis table for included sources with explicit GRADE-informed certainty reasoning.

Source	Country/setting	Source type and sample	AI/tool type	Consultation task mapped to CDC Yellow Book domains [8]	Main finding	Key safety concern	GRADE-informed certainty (reasoning) [21]
Ngiam et al. ChatGPT pre-travel advice evaluation [2]	Not patient-setting specific	Scenario-based expert evaluation; no patient sample	General-purpose ChatGPT	General advice, vaccination, malaria prophylaxis, traveller's diarrhoea, vector avoidance	Readable and often accurate answers to common questions	Generic advice; insufficient itinerary and comorbidity personalisation	Very low (single non-patient study; high indirectness; no comparator)

 doi:10.20944/preprints202605.0838.v1

Koh et al. Travel Clinic Assistant [3]	Singapore tertiary pre-travel clinic	Implementation research letter; 26 travellers	Custom GPT-4 assistant	Pre-consultation education, query elicitation, complex traveller education	Acceptable to travellers and physicians; perceived consultation focus and knowledge benefit	Small sample, self-report outcomes, digital literacy barriers, no EHR integration, hallucination risk	Very low (small single-site implementation; subjective outcomes; serious imprecision)
Baglivo et al. travel-health chatbot decalogue [4]	Italy/prototype context	Expert framework and pre-alpha custom GPT example	Custom GPT prototype	Personalisation, geolocation, multilingual support, clinic referral, EHR aspiration	Proposed ten design requirements for safe travel-health chatbots	Prototype lacks full privacy, scope-control, and EHR safeguards	Very low (framework paper; no empirical outcomes)
Flaherty editorial — supervised GenAI integration [5]	International travel medicine	Expert opinion/editorial	Generative AI broadly	Pre-clinic preparation, translation, literacy tailoring, reminders	AI may support preparation and reinforce consultation learning	Must not replace individualised clinician judgement	Very low (expert opinion; non-empirical)
Flaherty and Piyaphanee natural-history editorial [6]	International travel medicine	Expert opinion/editorial	AI broadly	Risk personalisation, behaviour prediction, surveillance	Frames AI's potential trajectory in travel medicine	Non-empirical; aspirational	Very low (expert opinion; non-empirical)
Heidema et al.	International	Multidisciplinary	Machine-	Outbreak detection	Demonstrates concrete	Non-empirical	Very low (perspecti

 doi:10.20944/preprints202605.0838.v1

GeoSentinel-AI surveillance [7]		editorial/perspective	learning approaches	adjacent to pre-travel risk	adjacent AI implementation pathway	for pre-travel decisions	ve paper; indirect outcomes)
Vibert et al. FeverTravelApp [12]	Switzerland; returned traveller fever workflow	Case-control simulated consultations; seven physicians, three simulated patients	Tablet clinical decision-support algorithm	Travel-related risk intake, exposure history, dynamic clinical reasoning	Demonstrates feasibility issues for travel-related CDSS in consultations	Indirect to pre-travel prevention; clinician interaction and adoption matter	Low (small simulation study; indirect to pre-travel)
CDC Yellow Book pre-travel consultation guidance [8,9]	United States guidance	Clinical guidance	Not AI	Gold-standard task taxonomy for pre-travel risk assessment	Defines domains AI must support and not oversimplify	Authoritative reference standard against which AI outputs should be checked; risk that AI outputs may diverge from current guidance	Not applicable (guideline; serves as reference standard)
WHO malaria travel guidance [11]	Global guidance	Clinical guidance	Not AI	Malaria geography, chemoprophylaxis, mosquito protection	Defines high-risk domain requiring up-to-date recommendations	Updates frequently; AI tools relying on training-data snapshots may be outdated	Not applicable (guideline; serves as reference standard)

 doi:10.20944/preprints202605.0838.v1

ISTM pre-travel health advice [10]	International travel medicine	Professional fact sheet	Not AI	Risk assessment, timing, vaccines, medicines, chronic illness	Reinforces that travel advice extends beyond vaccines	Useful patient-facing standard; AI must not understate timing-of-consultation criticality	Not applicable (guideline; serves as reference standard)
Shool et al. LLM evaluation systematic review [13]	General clinical medicine	Systematic review; 761 studies	LLMs	Evaluation standards for clinical AI	Evaluation methods remain heterogeneous	Indirect to travel medicine	Low (systematic review; indirect outcomes)
Collins multi-model hallucination assurance [14]	General clinical decision support	Simulation study; multiple models	Multiple LLMs	Safety testing for clinical decision support	LLMs repeated or elaborated false clinical details in 50 to 82 percent of outputs	Directly relevant to hallucination risk	Moderate for general LLM risk; indirect for travel (consistent finding across multiple models)
Asgari et al. CREOLA hallucination framework [15]	Clinical documentation	Framework and evaluation study	LLMs	Clinical documentation accuracy	Hallucinations more often “major” than omissions, especially in plan sections	Directly relevant to AI-generated travel advice	Low to moderate (single framework study; high indirectness to travel)

 doi:10.20944/preprints202605.0838.v1

van Kessel et al. clinical guideline hallucination analysis [16]	Clinical decision support	Diagnostic LLM systematic analysis	LLMs	Hallucination of authoritative guidelines	Identifies measurable prevalence of fabricated and omitted clinical guideline content	Highly relevant to fabricated travel-vaccine or malaria guidance	Low (single multi-model analysis; indirect to travel)
Bagde et al. ChatGPT meta-analysis [47]	Medical and dental research	Systematic review and meta-analysis	ChatGPT	Domain-specific accuracy	Accuracy 18 to 100 percent across specialties; high variability	Indirect; supports specialty-specific benchmarks	Low (high inconsistency; indirectness)
Duong et al. care-seeking accuracy [48]	General	Multi-version evaluation; 22 model versions	ChatGPT	Care-seeking advice across urgency levels	Average accuracy ~70 percent; overtriage; increasing variability with newer models	Directly supports conservative governance	Low to moderate (multi-version simulation; indirect to travel)
Geracitano et al. ChatGPT FAQ literature review [50]	General	Literature review; nine studies	ChatGPT	FAQ, recommendation, symptom categorisation	Accuracy 20 to 95 percent; not standalone point-of-care	Supports human oversight requirement	Low (small literature review; indirect to travel)
Aydin et al. patient-education LLM scoping	General medicine	Scoping review	LLMs	Patient education and engagement	LLMs may generate education material but face	Indirect to pre-travel education	Low (scoping review; indirect outcomes)

 doi:10.20944/preprints202605.0838.v1

review [17]					accuracy, readability, and bias challenges		
Kurniawan et al. chronic-illness chatbot review [18]	Chronic disease management	Systematic review	Chatbots	Acceptability and effectiveness	Acceptability promising; efficacy evidence limited; insufficient technical documentation	Mirrors travel-medicine implementation gap	Low (systematic review; indirect to travel)
Iyer et al. preventive-care chatbot outreach [52]	US value-based care	Retrospective analysis	Chatbot outreach	Preventive care compliance	Chatbots underperformed phone calls overall but outperformed for diabetes care in 2023	Selective and context-dependent efficacy	Low (single retrospective analysis; indirect to travel)
Peng et al. retrieval-augmented generation [49]	Multilingual medical	Comparative evaluation; 10 LLMs	RAG architectures	Source-grounded medical answer generation	RAG improves accuracy and generalises to unseen medical languages	Supports RAG as design principle	Low to moderate (comparative empirical study; indirect to travel)

Quality and Applicability Appraisal

Quality and applicability were appraised using MMAT-informed [19], AMSTAR 2-informed [20], JBI text-and-opinion-informed [30], and GRADE-informed [21] approaches. No source was excluded on the basis of quality appraisal; appraisal informed interpretation and certainty only.

 doi:10.20944/preprints202605.0838.v1

Table 3. Quality and applicability appraisal of included sources, with rationale and domain.

Source category	Appraisal approach	Domain assessed	Rationale and main appraisal judgement	Implication for synthesis
Singapore Travel Clinic Assistant [3]	MMAT-informed implementation appraisal [19]	Sampling, outcome ascertainment, conflict-of-interest control	Single centre, 26 travellers, qualitative feedback, no comparator; no objective effectiveness outcome; selection bias possible	Useful feasibility signal, not effectiveness evidence
ChatGPT pre-travel advice evaluation [2]	Custom accuracy-study appraisal	Scenario coverage, reproducibility, expert benchmarking	Clinically relevant questions and expert comparison; no patient outcomes; no model-version reproducibility; limited scenario diversity	Supports educational potential only
Decalogue and editorials [4–7]	JBI text/opinion-informed appraisal [30]	Domain expertise, logical consistency, relevance	Strong domain expertise and clinical logic but non-empirical and prescriptive rather than evaluative	Useful for implementation principles, not outcome claims
FeverTravelApp [12]	MMAT-informed appraisal [19]	Study design, sample size, blinding, comparator	Empirical travel-related CDSS evidence but post-travel and simulated; small physician sample; indirect to pre-travel	Useful for workflow and adoption lessons

 doi:10.20944/preprints202605.0838.v1

General clinical LLM systematic review [13]	AMSTAR 2-informed appraisal [20]	PRISMA reporting, search comprehensiveness, risk of bias assessment	Large and relevant review, transparent methods, but indirect to travel medicine; heterogeneity not pooled	Supports need for standardised evaluation
Multi-model hallucination, accuracy, and guideline-hallucination studies [14–16,47,48,50]	Simulation- and review-study appraisal	Reproducibility, prompt control, clinical reference standard	Strong safety signal for LLM vulnerability and accuracy variability; generally not travel-specific; consistent across models in [14]	Supports conservative governance and human review
Patient-education and chatbot reviews [17,18,46,52]	AMSTAR 2-informed appraisal [20]	Search, selection, synthesis transparency	Relevant adjacent evidence; heterogeneous designs and outcomes; limited primary trial data	Supports plausibility of supervised AI use and equity caution
Guidelines and authoritative texts [8–11,41]	JBI text-and-opinion-informed appraisal [30]	Authority, currency, scope	Authoritative guidance from CDC, WHO, ISTM; not designed as evidence appraisal targets but as reference standards	Used as the gold-standard task taxonomy and reference standard, not as outcome evidence

What the Evidence Allows and Does not Allow

The evidence allows a cautious conclusion that AI can support pre-consultation education, structured question preparation, basic travel-health information delivery, and clinician-supervised workflow support [2–6,17,46]. It also supports the conclusion that current general-purpose chatbots are not sufficiently validated for unsupervised individualised travel medicine [2,14,15,47,48,50].

The evidence does not allow claims that AI improves vaccine uptake, reduces travel-associated illness, improves malaria prophylaxis adherence, safely screens contraindications, reduces adverse events, or can replace a travel medicine clinician [2–6,12]. Those outcomes have not been tested in adequately powered prospective studies meeting CONSORT-AI, SPIRIT-AI, or TRIPOD+AI standards [31–33,45].

 doi:10.20944/preprints202605.0838.v1

Clinical Safety Risk Taxonomy

Drawing on the included LLM safety, hallucination, and accuracy literature [13–16,47,48,50], travel-medicine guidance [8–11,51], and the empirical travel-medicine implementations [2,3,12], the principal AI failure modes for pre-travel consultations are summarised in Table 4.

Table 4. Clinical safety risk taxonomy for AI in pre-travel consultations.

AI failure mode	Travel medicine example	Potential patient safety consequence	Mitigation
False destination risk claim [11,14,16]	Incorrectly states that a specific region has no malaria risk	Omitted chemoprophylaxis or inadequate mosquito precautions	Retrieval-grounded malaria source [11,49], date-stamped destination data, clinician review
Outdated outbreak information [7,16,51]	Misses active yellow fever, polio, measles, dengue, or mpox advisory	Unvaccinated or underprepared traveller enters risk zone	Real-time public health feed, source timestamp, no model-memory-only outbreak advice
Contraindication miss [8,15]	Recommends live vaccine to immunocompromised traveller	Vaccine-derived illness or serious adverse event	Mandatory immune-status questions and hard-stop clinician review
Drug interaction error [8,15]	Ignores psychiatric history or interacting medication when discussing mefloquine	Neuropsychiatric adverse event or poor adherence	Medication reconciliation, contraindication checklist, pharmacist or clinician sign-off
False reassurance [14,48]	Tells a splenectomy patient that malaria risk is routine or low	Life-threatening malaria risk underestimated	High-risk condition trigger and escalation to specialist review
Incomplete history intake [4,8]	Does not ask about pregnancy, transplant, HIV, anticoagulation, allergy, prior vaccines, or itinerary details	Inappropriate vaccine, medication, or risk counselling	Structured intake before any recommendation

 doi:10.20944/preprints202605.0838.v1

AI failure mode	Travel medicine example	Potential patient safety consequence	Mitigation
Hallucinated authority [14–16,50]	Fabricates a guideline, dose, requirement, or clinic policy	Clinician or patient follows non-existent recommendation	Source-linked output only [49]; block unsupported claims
Equity failure [3,37,38,42–44]	Older adult or low-literacy traveller cannot use tool	Exclusion of high-risk groups from pre-consultation support	Assisted use, multilingual and plain-language modes, non-digital alternative
Overtriage or undertriage [48]	Routes low-acuity question to emergency or vice versa	Resource misuse or delayed care	Calibrated triage thresholds, clinician review of escalations

Implementation Model for Travel Clinics

The implementation model in Table 4 synthesises the supervised-augmentation principle from the Flaherty editorials [5,6], the design requirements from Baglivo et al. [4], the Singapore implementation lessons [3], the FeverTravelApp adoption findings [12], retrieval-augmented design principles [49], and the clinician-adoption framework articulated by Scott and colleagues [34].

Table 4. Supervised AI implementation model for pre-travel consultation workflows.

Workflow stage	AI function	Failure mode or escalation trigger	Governance requirement	Clinical owner
Booking	Collect itinerary, departure date, destinations, activities, and baseline medical information [4,8]	Pregnancy, immunosuppression, transplant, HIV, splenectomy, complex itinerary, or departure within 14 days	Privacy notice, minimum data capture, audit of intake completion [26,27]	Clinic lead
Waiting room	Provide general education and elicit patient questions [5,17,46]	Patient asks for vaccine clearance, medication prescription, diagnosis, or high-risk advice	Source-grounded educational mode only [49]; no prescribing	Travel clinician

 doi:10.20944/preprints202605.0838.v1

Workflow stage	AI function	Failure mode or escalation trigger	Governance requirement	Clinical owner
Consultation	Produce structured summary of risks, missing data, and guideline prompts [4,12,34]	Missing vaccine history, unclear immune status, drug interaction, live-vaccine question	Clinician verifies every recommendation before action [5,15]	Travel clinician
After-visit	Reinforce clinician-approved plan, vaccine schedule, malaria instructions, and behavioural advice [5,52]	Patient reports adverse reaction, fever, pregnancy, itinerary change, or medication intolerance	Escalation pathway and documented clinician-approved content	Clinic protocol owner
During travel	Provide emergency numbers, reminder prompts, and red-flag advice [5,11]	Fever after malaria exposure, animal bite, severe diarrhoea, respiratory distress, sexual exposure, injury	Immediate seek-care advice; no autonomous diagnosis	Traveller support protocol
Quality assurance	Audit AI outputs and user feedback [13–16,34]	Hallucination, outdated source, unsafe omission, inequitable use pattern	Monthly sample audit, incident register, model-version log, safety review board [22,23,26]	Clinical governance committee

Cost-Effectiveness and Implementation Feasibility

No included source reports cost-effectiveness data for AI in pre-travel consultations. Implementation feasibility considerations include licensing or hosting costs for clinical-grade LLM platforms, integration costs with electronic health records, ongoing prompt and retrieval-source maintenance, audit and governance staffing, multilingual content review, and equity-focused assistance for travellers with low digital literacy [3,4,34]. The chronic-illness chatbot review noted insufficient technical documentation in primary studies, which obstructs cost-effectiveness modelling and benchmarking [18]. The preventive-care chatbot evidence shows context-dependent efficacy and underperformance against telephone outreach overall, suggesting that travel-medicine AI value will likely be selective rather than uniform [52]. Future trials should report cost per consultation

 doi:10.20944/preprints202605.0838.v1

supported, cost per hallucination averted, equity-adjusted cost-effectiveness, and break-even thresholds for clinic deployment [31–33,45].

Regulatory Landscape

AI tools used in pre-travel consultations may fall within multiple medical-device and AI regulatory frameworks depending on intended use, claims, and risk class. Table 5 summarises the most directly relevant regimes.

Table 5. Regulatory landscape for AI in pre-travel health consultations.

Jurisdiction	Regulatory signal	Relevance to travel-medicine AI
United States FDA [22]	FDA describes AI/ML in Software as a Medical Device as requiring lifecycle management and appropriate premarket pathways such as 510(k), De Novo, or premarket approval depending on intended use	A tool that merely educates may be lower risk, while a tool that drives vaccine or medication recommendations may approach regulated clinical decision support
Australia TGA [23]	TGA regulates software when it meets the medical-device definition under section 41BD of the Therapeutic Goods Act 1989, and developers of AI-enabled medical device software may be manufacturers or sponsors	Australian travel clinics should assess intended use, claims, risk class, sponsor obligations, and post-market monitoring before deploying AI decision tools
European Union AI Act Article 6 [24]	Article 6 classifies AI as high-risk when it meets product safety conditions or falls within Annex III categories, with exemptions only where it does not pose significant risk to health, safety, or fundamental rights	Travel AI affecting health decisions may require high-risk analysis, especially if it materially influences clinical recommendations
European Union AI Act Annex III [25]	Annex III includes systems related to healthcare service eligibility, health insurance risk assessment, emergency healthcare triage, and health-risk assessment in migration/border contexts	A travel-health AI tool handling triage or health-risk classification should be assessed for high-risk obligations and documentation requirements

 doi:10.20944/preprints202605.0838.v1

Jurisdiction	Regulatory signal	Relevance to travel-medicine AI
WHO AI ethics guidance [26]	Six consensus principles: protect autonomy, promote human well-being and safety, ensure transparency and explainability, foster responsibility and accountability, ensure inclusiveness and equity, and promote responsive and sustainable AI	Travel-medicine AI deployment should map governance to these principles
WHO Global Strategy on Digital Health [27]	Emphasises equity, scalability, privacy, security, and country readiness as prerequisites for digital health deployment	Travel-medicine AI should be evaluated against these macro-level prerequisites

Digital Equity Considerations

The Singapore implementation identified difficulty among some older adults because chatbot use required digital literacy [3]. This matters because travellers with chronic illness, older travellers, immunocompromised travellers, and people visiting friends and relatives may be both higher risk and more likely to face barriers to digital health tools [37,38,41].

VFR and migrant travellers are well-documented as high-risk and underserved. Qualitative work among VFR migrants identifies cultural perception of risk, financial barriers, language barriers, and trust as key obstacles to pre-travel care, and interventions that ignore these structural factors are unlikely to close the gap [37]. West African VFR travellers in the United States similarly under-utilise pre-travel care because of awareness, time, and cost barriers [38]. The CDC Yellow Book VFR chapter highlights that VFR travellers are disproportionately affected by malaria, hepatitis A, typhoid, and tuberculosis, and emphasises tailored counselling, language access, and family-inclusive education [41].

In Australia and the Pacific, First Nations Australians and Pacific Islander travellers face additional structural barriers including geographic remoteness, primary care access gaps, and culturally specific information needs that are not addressed by generic English-language travel-medicine content; these populations are also more likely to travel for VFR purposes within and beyond the region. Direct empirical evidence on AI tools in these populations is absent in the included sources, which is itself a significant equity gap. Any travel-medicine AI deployment in Australia should therefore be co-designed with Aboriginal and Torres Strait Islander Community Controlled Health Services and Pacific community organisations rather than rolled out generically, and should be evaluated for cultural safety as well as accuracy [3,37,38,41–44].

Population-level data show persistent disparities in digital health care use. Impoverished, female, Black, and internet-poor populations experience reduced telehealth completion [42], county-level social vulnerability shapes 2022 digital health care use patterns [44], and the broader principle of "intervention-generated inequities" warns that new digital tools may inadvertently widen disparities if not deliberately designed for equity [43]. Migrant and VFR travellers in particular face multifaceted socioeconomic, financial, language, and systems-level barriers to pre-travel care [37,38,41].

 doi:10.20944/preprints202605.0838.v1

Equity safeguards should include assisted kiosk use, clinician-reviewed non-digital alternatives, multilingual support [49], plain-language output, accessibility testing, culturally appropriate examples, and cost-free access. Travel-medicine AI training programmes should incorporate AI literacy and implementation science to support equitable deployment [35]. If AI is used to improve clinic efficiency but excludes high-risk travellers, it may worsen rather than reduce pre-travel health inequities [42–44].

When not to Use AI as the Primary Interaction

Table 7 summarises traveller characteristics where AI should not be the primary interaction. In these scenarios, AI may collect structured background information, but it should immediately escalate to clinician review [5,34].

Table 7. When not to use AI as the primary interaction.

Traveller characteristic	Risk category	Required action
Pregnancy or planning pregnancy [8,11]	Live-vaccine and antimalarial contraindication risk	Mandatory clinician review; AI restricted to information-gathering
Immunocompromise (transplant, advanced HIV, immunosuppressive therapy, asplenia) [8]	Vaccine-derived illness and severe travel infection risk	Mandatory specialist review; AI must not advise on vaccine eligibility
Anticoagulation or unstable cardiovascular disease [8]	Drug interaction and travel-stress risk	Clinician review of medication and itinerary
Severe allergy or anaphylaxis history [8]	Vaccine reaction risk	Clinician-led vaccine selection and observation planning
Complex psychiatric history [8]	Mefloquine and other neuropsychiatric medication risk	Clinician-led prophylaxis selection
Travel within two weeks [10]	Inadequate time for vaccine schedules	Triaged clinician review and accelerated schedule
Outbreak-zone travel [7,11,51]	Time-sensitive epidemiology beyond model knowledge	Real-time public health source and clinician review
Live vaccine clearance request	Direct contraindication assessment	Clinician-only decision
Malaria prophylaxis selection request [11]	Resistance, drug-interaction, comorbidity-specific decision	Clinician-only prescribing

 doi:10.20944/preprints202605.0838.v1

Traveller characteristic	Risk category	Required action
Post-exposure care after animal bite, sexual exposure, or needlestick	Time-critical post-exposure prophylaxis	Direct clinician contact, emergency services if needed
Severe digital literacy or language barriers [3,37,38,42–44]	Equity and comprehension risk	Assisted use and non-digital alternative
First Nations Australian or Pacific Islander traveller in absence of culturally adapted content [37,38,41–44]	Cultural safety and trust	Co-designed clinician pathway; not generic AI as primary interaction

Research Agenda and Priority Matrix

Future research priorities are summarised in Table 6. Trial-stage and prediction-model studies should follow CONSORT-AI [31], SPIRIT-AI [32,45], and TRIPOD+AI [33] reporting standards, ensuring that algorithm version, input data acquisition, human-AI interaction, fairness, subgroup performance, and post-deployment monitoring are reported consistently. A Minimum Reporting Standards checklist for travel-medicine AI studies is provided in Supplement S5, and a commercial and openly accessible tool inventory template, with current evidence limitations, is also in Supplement S5.

Table 6. Research priority matrix for AI in pre-travel health consultations.

Priority	Study or activity	Rationale	Suggested outcomes
Immediate	Hallucination audit of travel-medicine chatbots against CDC Yellow Book 2026 [8,9], WHO malaria guidance [11], and ISTM advice [10]	High urgency and feasible with simulated cases [14–16]	Accuracy, harmful omission, hallucination, citation validity, refusal behaviour
Immediate	Prospective structured intake trial in one travel clinic [3,4,12]	High feasibility and direct workflow relevance	Consultation time, missing-data rate, clinician satisfaction, patient understanding
Near-term	Stepped-wedge trial across multiple travel clinics following CONSORT-AI/SPIRIT-AI [31,32,45]	Tests implementation under real-world variation	Vaccine uptake, malaria prophylaxis appropriateness, advice adherence, safety events

 doi:10.20944/preprints202605.0838.v1

Priority	Study or activity	Rationale	Suggested outcomes
Near-term	Equity study in older adults, low-digital-literacy travellers, VFR travellers, First Nations Australians, and Pacific Islander travellers [37,38,41–44]	Addresses likely access asymmetry	Usability, completion, comprehension, preference, assisted-use need, cultural safety
Longer-term	EHR-integrated retrieval-augmented generation system with outbreak-feed integration [7,49,51]	Highest potential but greater regulatory and privacy burden [22–25]	Recommendation concordance, auditability, privacy incidents, model drift
Longer-term	Multilingual validation across common traveller origin languages [49]	Needed for global travel medicine	Translation fidelity, cultural appropriateness, safety equivalence
Longer-term	Travel-medicine AI prediction models for risk stratification, reported per TRIPOD+AI [33]	Enables individualised pre-travel risk advice	Discrimination, calibration, fairness, decision-curve utility

Discussion

This review supports a narrow but clinically important conclusion: AI can be useful in pre-travel health consultations as a supervised augmentation layer, but the evidence does not support autonomous AI decision-making [2–6,14,15,34,48,50]. The best-supported roles are education, question preparation, structured intake, guideline-linked summarisation, translation, literacy tailoring, and clinician-approved reminders [4,5,17,46,49,52].

Comparison with Prior Reviews and How this Review Extends Them

To the knowledge of the present review, no prior peer-reviewed scoping review specifically addresses AI tools in pre-travel health consultations as its primary subject. The Aydin et al. scoping review of LLMs in patient education spans general medicine and identifies recurring themes of patient education material generation, medical information interpretation, lifestyle recommendations, medication-use support, perioperative instructions, and doctor-patient interaction, while flagging readability, accuracy, and bias as cross-cutting challenges [17]. The Kurniawan et al. systematic review of chatbots for chronic illness reports promising acceptability but limited efficacy evidence and incomplete technical documentation [18]. The AHRQ review of healthcare chatbots similarly characterises the broader chatbot evidence as nascent [46]. The preventive-care chatbot retrospective analysis demonstrates that chatbot efficacy is selective and

 doi:10.20944/preprints202605.0838.v1

context-dependent, underperforming phone outreach overall while outperforming for diabetes care in 2023 [52].

The present review extends Baglivo et al. [4], who proposed a ten-item design decalogue for travel-health chatbots with a pre-alpha custom GPT prototype but did not synthesise empirical AI evidence in travel medicine, by mapping the empirical and adjacent evidence base, locating it within authoritative travel-medicine guidance and clinical AI safety standards, and embedding it in a supervised implementation model with explicit governance and equity safeguards. It extends the Flaherty editorials [5,6] by translating their supervised-augmentation principle into a structured workflow model and a clinical safety risk taxonomy. It extends Heidema et al. [7] by anchoring AI-enabled outbreak surveillance to the pre-travel consultation rather than treating it as a parallel activity. It extends generic clinical LLM safety work [13–16] by identifying which failure modes carry travel-medicine-specific consequences, particularly for live-vaccine eligibility, malaria prophylaxis, and outbreak-zone advice.

Why the Evidence Base Remains Thin

The thinness of the evidence base should shape, not weaken, the contribution of this review. A field-mapping review is valuable because it prevents premature clinical adoption from outrunning evidence. It also clarifies that travel medicine needs its own AI evaluation standards, rather than importing generic patient-education benchmarks from other specialties [13,17,18].

The main reason for caution is not that AI performs poorly on all travel-health questions. The ChatGPT evaluation suggests that general questions about food and water safety, vector avoidance, traveller's diarrhoea, vaccine timing, and malaria concepts can be answered readably and often accurately [2]. The problem is that travel medicine becomes high risk precisely when advice must be individualised to immune status, pregnancy, prior vaccines, comorbidities, medication interactions, destination micro-geography, and outbreak timing [8,11]. LLMs are particularly vulnerable to adversarial false-premise prompts and to hallucinated authoritative guidance in clinical decision support [14–16], and ChatGPT-class models show wide accuracy ranges across specialties and care-seeking tasks [47,48,50].

Practical Message for Travel Medicine Clinicians

For travel medicine specialists, the immediate practical message is to separate low-risk information support from safety-critical decision-making. AI can explain why pre-travel consultation matters, help travellers prepare questions, and remind them how to take clinician-prescribed malaria prophylaxis [5,17,52]. AI should not independently clear a traveller for yellow fever vaccine, select malaria prophylaxis, determine live-vaccine eligibility, or reassure a high-risk traveller without clinician review [8,11,14,15,48]. Retrieval-augmented generation grounded in CDC Yellow Book, WHO malaria, and ISTM sources [8–11,49], combined with documented clinician oversight [5,34] and equity safeguards [42–44], represents the most defensible near-term architecture.

Limitations

This review has several important limitations.

First, the search was a transparent, targeted scoping retrieval rather than a fully reproducible multi-database systematic search. PubMed/MEDLINE, academic and web-indexed search tools, citation chasing, and authoritative guideline and regulator websites were used; Embase, CINAHL, Cochrane CENTRAL, IEEE Xplore, ACM Digital Library, ClinicalTrials.gov, WHO ICTRP, and ISTM conference abstracts were not run. Search strategies for these databases are documented in Supplement S2 as recommended reproducibility strategies but should not be presented as completed work [1,28–30].

 doi:10.20944/preprints202605.0838.v1

Second, screening and data charting were conducted by a single reviewer, so Cohen's kappa for inter-rater agreement is not calculable; mitigation included a structured eligibility checklist and a 24-hour second-pass review (Supplement S2) [1].

Third, the PRISMA-ScR flow reflects the targeted retrieval process for this review; a future formal database run could increase the identified pool and would require updating the flow diagram (Supplement S1) [1].

Fourth, the quality appraisal is indicative rather than duplicate-assessed [19,20].

Fifth, several included sources are expert opinion, guidance, or indirect clinical AI safety evidence rather than primary pre-travel AI trials [4–7,13–16,46–48,50].

Sixth, formal stakeholder consultation with travel medicine clinicians, primary care providers, regulators, traveller representatives (including VFR travellers, older adults, First Nations Australian and Pacific Islander travellers), and digital health developers was not conducted. The implementation model in Table 4 is therefore a literature-based synthesis, not a co-designed model, and its clinical and operational applicability should be validated through stakeholder consultation before deployment.

Seventh, no primary cost-effectiveness or implementation-economic evidence was identified.

These limitations do not invalidate the central conclusion, but they limit certainty [21]. The review is positioned as a scoping review that maps an emerging field and proposes a safety-oriented research agenda, not as a definitive effectiveness review.

Conclusion

AI tools are likely to become part of pre-travel health workflows, but current evidence supports supervised augmentation rather than autonomous clinical decision-making [2–6,14,15,34,48,50]. The strongest near-term use cases are structured intake, patient education, clinician-reviewed guideline retrieval, multilingual explanation, and after-visit reinforcement [4,5,17,46,49,52]. The highest-risk use cases are vaccine contraindication screening, malaria prophylaxis selection, live-vaccine clearance, complex medication interaction advice, and reassurance of high-risk travellers [8,11,14,15,48].

Advancement of this field requires formal database searches, duplicate screening, standardised AI safety benchmarks, domain-specific hallucination audits, equity testing across VFR, migrant, older-adult, First Nations Australian, and Pacific Islander travellers, and prospective clinic trials measuring clinically meaningful outcomes, reported in line with CONSORT-AI, SPIRIT-AI, and TRIPOD+AI [31–33,45].

Author Contributions: The single named author conceived the review, designed the eligibility framework and search strategy, conducted screening and data charting, performed appraisal and synthesis, and drafted and revised the manuscript. The author meets ICMJE authorship criteria.

Funding: No funding was received for this manuscript.

Ethics approval: Ethics approval was not required because this review used published literature and did not involve human participants, patient-identifiable data, or intervention delivery.

Data availability: Search files, extracted source summaries, the PRISMA-ScR figure [1], the completed PRISMA-ScR checklist (Supplement S1), the search strategy and eligibility checklist (Supplement S2), the excluded and retrieval log and characteristics-of-included-sources tables (Supplements S3 and S4), and the minimum reporting standards checklist and tool inventory (Supplement S5) accompany this manuscript. A formal future submission should include database-specific search histories and an excluded-studies table extended through additional databases.

Conflicts of interest: The author declares no conflicts of interest.

 doi:10.20944/preprints202605.0838.v1

Patient and public involvement (PPI): Patient and public involvement was not undertaken in the design, conduct, or reporting of this review. This is a limitation, particularly for the implementation model and equity sections, and PPI is recommended for any future co-designed travel-medicine AI deployment, especially with VFR travellers, older adults, First Nations Australian, and Pacific Islander communities.

Reporting standards used: The review uses PRISMA-ScR [1] and is informed by the Arksey and O'Malley framework as advanced by Levac et al. and updated in JBI scoping review guidance [28–30]. Quality appraisal is informed by MMAT [19] and AMSTAR 2 [20], and certainty by GRADE [21]. Future trials, protocols, and prediction models in this domain should follow CONSORT-AI [31], SPIRIT-AI [32,45], and TRIPOD+AI [33].

References

1. 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):467–73. doi:10.7326/M18-0850. Available from: https://pubmed.ncbi.nlm.nih.gov/30178033/
2. Ngiam JN, Koh MCY, Lye P, Liong TS, Salada BMA, Tambyah PA, et al. Artificial intelligence models for pre-travel consultation and advice: yea or nay? J Travel Med. 2024;31(1):taad124. doi:10.1093/jtm/taad124. Available from: https://academic.oup.com/jtm/article/31/1/taad124/7288181
3. Koh MCY, Ngiam JN, Chan NJH, Peh JH, Tambyah PA, Oon JEL, et al. Implementation of ChatGPT to enhance pre-travel consultation in a specialist tertiary centre in Singapore. J Travel Med. 2024;32(5):taae099. doi:10.1093/jtm/taae099. Available from: https://academic.oup.com/jtm/article/32/5/taae099/7717433
4. Baglivo F, Bovo C, Riboli M, Capasso R, Brunelli N, Ciccarese M, et al. A decalogue for personalized travel health assistance with artificial intelligence-driven chatbots. Travel Med Infect Dis. 2024;59:102710. doi:10.1016/j.tmaid.2024.102710. Available from: https://pmc.ncbi.nlm.nih.gov/articles/PMC11149716/
5. Flaherty GT. Learning to safely integrate generative artificial intelligence technology into travel medicine practice. J Travel Med. 2025;32(8):taad149. doi:10.1093/jtm/taad149. Available from: https://academic.oup.com/jtm/article/32/8/taad149/7453443
6. Flaherty GT, Piyaphanee W. Predicting the natural history of artificial intelligence in travel medicine. J Travel Med. 2023;30(1):taac113. doi:10.1093/jtm/taac113. Available from: https://pmc.ncbi.nlm.nih.gov/articles/PMC9940693/
7. Heidema SGA, Stoepker IV, Flaherty G, Goorhuis A, Grobusch MP, Penders J, et al. From GeoSentinel data to epidemiological insights: a multidisciplinary effort towards artificial intelligence-supported detection of infectious disease outbreaks. J Travel Med. 2024;31(4):taae013. doi:10.1093/jtm/taae013. Available from: https://pmc.ncbi.nlm.nih.gov/articles/PMC11287261/
8. Centers for Disease Control and Prevention. The pre-travel consultation. In: CDC Yellow Book 2026: Health Information for International Travel. Atlanta: CDC; 2025. Available from: https://www.cdc.gov/yellow-book/hcp/preparing-international-travelers/the-pre-travel-consultation.html
9. Centers for Disease Control and Prevention. CDC Yellow Book 2026: Health Information for International Travel. Bethesda: National Center for Biotechnology Information; 2025. Available from: https://www.ncbi.nlm.nih.gov/books/NBK620896/
10. International Society of Travel Medicine. Pre-travel health advice fact sheet. Alpharetta: ISTM; 2023. Available from: https://www.istm.org/wp-content/uploads/istm-pre-travel-health-advice-fact-sheet.pdf
11. World Health Organization. International Travel and Health: Module 3 — Malaria. Geneva: WHO; 2022. Available from: https://www.who.int/publications/i/item/9789240102286
12. Vibert J, Bourquin C, De Santis O, Cobuccio L, D'Acremont V. Influence of the use of a tablet-based clinical decision support algorithm by general practitioners on the consultation process: the example of FeverTravelApp. BMC Digit Health. 2024;2(1):36. doi:10.1186/s44247-024-00118-4. Available from: https://bmcdigitalhealth.biomedcentral.com/articles/10.1186/s44247-024-00118-4
13. Shool S, Adimi S, Saboori Amleshi R, Bitaraf E, Golpira T, Mahmood MA. A systematic review of large language model evaluations in clinical medicine. BMC Med Inform Decis Mak. 2025;25(1):155. doi:10.1186/s12911-025-02954-4. Available from: https://bmcmedinformdecismak.biomedcentral.com/articles/10.1186/s12911-025-02954-4

 doi:10.20944/preprints202605.0838.v1

14. Collins JD. Multi-model assurance analysis showing large language models are highly vulnerable to adversarial hallucination attacks during clinical decision support. Commun Med. 2025;5(1):78. doi:10.1038/s43856-025-01021-3. Available from: https://www.nature.com/articles/s43856-025-01021-3
15. Asgari F, Sherwood LM, Clarkson C, Rahman M, Olaru ID, Taylor G, et al. A framework to assess clinical safety and hallucination of LLMs in clinical documentation. NPJ Digit Med. 2025;8(1):178. doi:10.1038/s41746-025-01670-7. Available from: https://www.nature.com/articles/s41746-025-01670-7
16. van Kessel R, Anderson M, McMillan B, Matthews MR, Rust P, Pearcy P, et al. Omission and hallucination prevalence of clinical guidelines in diagnostic large language model outputs. BMJ Health Care Inform. 2026;33(1):e101959. doi:10.1136/bmjhci-2025-101959. Available from: https://pmc.ncbi.nlm.nih.gov/articles/PMC13110572/
17. Aydin S, Karabacak M, Vlachos V, Margetis K. Large language models in patient education: a scoping review of applications in medicine. Front Med (Lausanne). 2024;11:1477898. doi:10.3389/fmed.2024.1477898. Available from: https://www.frontiersin.org/articles/10.3389/fmed.2024.1477898/full
18. Kurniawan MH, Handiyani H, Nuraini T, Hariyati RS, Sutrisno S. A systematic review of artificial intelligence-powered chatbot intervention for managing chronic illness. Ann Med. 2024;56(1):2302980. doi:10.1080/07853890.2024.2302980. Available from: https://www.tandfonline.com/doi/full/10.1080/07853890.2024.2302980
19. Hong QN, Pluye P, Fàbregues S, Bartlett G, Boardman F, Cargo M, et al. Mixed Methods Appraisal Tool (MMAT), version 2018. Montreal: McGill University; 2018. Available from: http://mixedmethodsappraisaltoolpublic.pbworks.com/w/file/fetch/127916259/MMAT_2018_criteria-manual_2018-08-01_ENG.pdf
20. Shea BJ, Reeves BC, Wells G, Thuku M, Hamel C, Moran J, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. BMJ. 2017;358:j4008. doi:10.1136/bmj.j4008. Available from: https://www.bmj.com/content/358/bmj.j4008
21. Centers for Disease Control and Prevention. GRADE criteria determining certainty of evidence. In: ACIP GRADE Handbook. Atlanta: CDC; 2023. Available from: https://www.cdc.gov/acip-grade-handbook/hcp/chapter-7-grade-criteria-determining-certainty-of-evidence/index.html
22. U.S. Food and Drug Administration. Artificial Intelligence and Machine Learning in Software as a Medical Device. Silver Spring: FDA; 2021. Available from: https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-software-medical-device
23. Therapeutic Goods Administration. Artificial intelligence and medical device software regulation. Canberra: TGA; 2024. Available from: https://www.tga.gov.au/products/medical-devices/software-and-artificial-intelligence-ai/manufacturing/artificial-intelligence-ai-and-medical-device-software-regulation
24. European Parliament and Council. Regulation (EU) 2024/1689 of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). Article 6: Classification rules for high-risk AI systems. Official Journal of the European Union; 2024. Available from: https://artificialintelligenceact.eu/article/6/
25. European Parliament and Council. Regulation (EU) 2024/1689 — Annex III: High-risk AI systems referred to in Article 6(2). Official Journal of the European Union; 2024. Available from: https://artificialintelligenceact.eu/annex/3/
26. World Health Organization. Ethics and governance of artificial intelligence for health: WHO guidance. Geneva: WHO; 2021. ISBN 9789240029200. Available from: https://www.who.int/publications/i/item/9789240029200
27. World Health Organization. Global strategy on digital health 2020–2025. Geneva: WHO; 2021. ISBN 9789240020924. Available from: https://www.who.int/publications/i/item/9789240020924
28. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol. 2005;8(1):19–32. doi:10.1080/1364557032000119616. Available from: https://www.tandfonline.com/doi/abs/10.1080/1364557032000119616

 doi:10.20944/preprints202605.0838.v1

29. Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology. Implement Sci. 2010;5:69. doi:10.1186/1748-5908-5-69. Available from: https://pmc.ncbi.nlm.nih.gov/articles/PMC2954944/
30. Peters MDJ, Godfrey C, McInerney P, Munn Z, Tricco AC, Khalil H. Chapter 11: Scoping reviews. In: Aromataris E, Munn Z, editors. JBI Manual for Evidence Synthesis. Adelaide: JBI; 2020. Available from: https://jbi-global.atlassian.net/wiki/spaces/manual/pages/355862497/10.+Scoping+reviews
31. Liu X, Rivera SC, Moher D, Calvert MJ, Denniston AK; SPIRIT-AI and CONSORT-AI Working Group. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI Extension. BMJ. 2020;370:m3164. doi:10.1136/bmj.m3164. Available from: https://pmc.ncbi.nlm.nih.gov/articles/PMC8183333/
32. Rivera SC, Liu X, Chan A-W, Denniston AK, Calvert MJ; SPIRIT-AI and CONSORT-AI Working Group. Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI Extension. Lancet Digit Health. 2020;2(10):e549–60. doi:10.1016/S2589-7500(20)30219-3. Available from: https://www.thelancet.com/journals/landig/article/PIIS2589-7500(20)30219-3/fulltext30219-3/fulltext)
33. Collins GS, Moons KGM, Dhiman P, Riley RD, Beam AL, Van Calster B, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. BMJ. 2024;385:e078378. doi:10.1136/bmj-2023-078378. Available from: https://www.bmj.com/content/385/bmj-2023-078378
34. Scott IA, Cook D, Coiera E, Richards B. Achieving large-scale clinician adoption of AI-enabled decision support. BMJ Health Care Inform. 2024;31(1):e100971. doi:10.1136/bmjhci-2023-100971. Available from: https://informatics.bmj.com/content/31/1/e100971
35. Paranjape K, Schinkel M, Panday RN, Car J, Nanayakkara P. Introducing artificial intelligence training in medical education. JMIR Med Educ. 2019;5(2):e16048. doi:10.2196/16048. Available from: https://mededu.jmir.org/2019/2/e16048/
36. Bajwa J, Munir U, Nori A, Williams B. Artificial intelligence in healthcare: transforming the practice of medicine. Future Healthc J. 2021;8(2):e188–94. doi:10.7861/fhj.2021-0095. Available from: https://pmc.ncbi.nlm.nih.gov/articles/PMC8344875/
37. Maguire C, Connolly F, O'Mahony L, Bradley CP, Byrne M. Improving the uptake of pre-travel health advice amongst migrants who visit friends and relatives: a qualitative study. BMC Infect Dis. 2016;16(1):253. doi:10.1186/s12879-016-1590-y. Available from: https://pmc.ncbi.nlm.nih.gov/articles/PMC4870764/
38. Jentes ES, Blanton JD, Johnson KJ, Petersen BW, Lamias MJ, Robertson AE, et al. Utilization and barriers to pre-travel health advice for visiting friends and relatives among West African travelers in the United States. J Travel Med. 2019;26(3):taz021. doi:10.1093/jtm/taz021. Available from: https://pmc.ncbi.nlm.nih.gov/articles/PMC6679970/
39. Alotaibi AG, Alharbi MM, Alzahrani AM, Alghamdi MK, Al-Hazmi MK, Albosais MA. Exploring the utilization and perceptions of pre-travel health consultations in primary care settings in Saudi Arabia: a cross-sectional study. Fam Pract. 2024;41(4):607–12. doi:10.1093/fampra/cmae030. Available from: https://pubmed.ncbi.nlm.nih.gov/38877586/
40. Mboowa G, Aceng-Barber B, Bazira J, Baryomunsi C, Bbalo M, Bulafu D, et al. Primary care physicians' role in travel medicine practice: a scoping review. J Travel Med. 2025;32(1):taae159. doi:10.1093/jtm/taae159. Available from: https://pmc.ncbi.nlm.nih.gov/articles/PMC11786444/
41. Centers for Disease Control and Prevention. Visiting friends and relatives: VFR travel. Chapter 9.3. In: CDC Yellow Book 2026: Health Information for International Travel. Atlanta: CDC; 2025. Available from: https://www.cdc.gov/yellow-book/hcp/refugees-immigrants-migrants/visiting-friends-and-relatives-vfr-travel.html
42. Disparities in Health Care and the Digital Divide. Curr Psychiatry Rep. 2021;23(9):61. doi:10.1007/s11920-021-01274-4. Available from: https://pmc.ncbi.nlm.nih.gov/articles/PMC8300069/
43. Chunara R, Zhao Y, Chen J, Lawrence K, Testa PA, Nov O, et al. Centering health equity in an increasingly digital healthcare landscape. J Am Med Inform Assoc. 2021;28(12):2720–3. doi:10.1093/jamia/ocab251. Available from: https://pmc.ncbi.nlm.nih.gov/articles/PMC12256345/

 doi:10.20944/preprints202605.0838.v1

44. Wakeman M, Buckman DW, El-Toukhy S. Disparities in digital health care use in 2022. JAMA Netw Open. 2025;8(4):e255359. doi:10.1001/jamanetworkopen.2025.5359. Available from: https://jamanetwork.com/journals/jamanetworkopen/fullarticle/2832907
45. Liu X, Cruz Rivera S, Moher D, Calvert MJ, Denniston AK; SPIRIT-AI and CONSORT-AI Working Group. Reporting guidelines for clinical trials of artificial intelligence interventions: the SPIRIT-AI and CONSORT-AI guidelines. Trials. 2021;22(1):11. doi:10.1186/s13063-020-04951-6. Available from: https://pmc.ncbi.nlm.nih.gov/articles/PMC7788716/
46. Chatbots in Health Care: Connecting Patients to Information. In: Chartbook on Health Care Quality. Rockville: Agency for Healthcare Research and Quality; 2024. Available from: https://www.ncbi.nlm.nih.gov/books/NBK602381/
47. Bagde H, Dhopte A, Alam MK, Basri R. A systematic review and meta-analysis on ChatGPT and its utilization in medical and dental research. Heliyon. 2023;9(12):e23050. doi:10.1016/j.heliyon.2023.e23050. Available from: https://pmc.ncbi.nlm.nih.gov/articles/PMC10746423/
48. Duong MTQ, Bain PA, Li T, Maron JL, Bhatt DL, Ko KJ, et al. Evaluating the accuracy of ChatGPT model versions for care-seeking advice across multiple urgency levels. Commun Med. 2026;6(1):64. doi:10.1038/s43856-026-01466-0. Available from: https://www.nature.com/articles/s43856-026-01466-0
49. Peng J, Lam E, Zhu M, Bonner N, Reyes-Morales A, Khanna M, et al. Retrieval augmented generation for 10 large language models and its generalisation to unseen medical languages. NPJ Digit Med. 2025;8(1):192. doi:10.1038/s41746-025-01519-z. Available from: https://www.nature.com/articles/s41746-025-01519-z
50. Geracitano J, Anderson B, Coffel M, Rosenzweig M, Dorn SD, Khairat S, et al. The accuracy of ChatGPT in answering FAQs, making clinical recommendations, and categorizing patient symptoms: a literature review. Adv Health Inf Sci Pract. 2025;1(1):VXUL2925. doi:10.63116/VXUL2925. Available from: https://pmc.ncbi.nlm.nih.gov/articles/PMC12366716/
51. Torresi J, Leder K. Defining infections in international travellers through the GeoSentinel surveillance network. Nat Rev Microbiol. 2009;7(12):895–901. doi:10.1038/nrmicro2238. Available from: https://pmc.ncbi.nlm.nih.gov/articles/PMC7097479/
52. Iyer P, Mehta P, Kim R, Mayer J, Ahuja A. Chatbot outreach in value-based preventive care: a retrospective analysis of compliance with preventive care measures. J Med Internet Res. 2026;28(2):e70921. doi:10.2196/70921. Available from: https://pmc.ncbi.nlm.nih.gov/articles/PMC12946782/