Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

The current roles of Artificial Intelligence in public health: A scoping review

Sean Harrison

s.harrison4@exeter.ac.uk

University of Exeter Medical School, University of Exeter

Claire Tatton

University of Exeter Medical School, University of Exeter

G. J. Melendez-Torres

University of Exeter Medical School, University of Exeter

Joelle Kirby

University of Exeter Medical School, University of Exeter

Sophie Robinson

University of Exeter Medical School, University of Exeter

Daniel Mutanda

University of Exeter Medical School, University of Exeter

Alisha R Davies

Public Health Wales

Jessica M Armitage

University of Exeter Medical School, University of Exeter

Rabeea'h Waseem Aslam

Cardiff University

Tom Arthur

University of Exeter Medical School, University of Exeter

Joht Singh Chandan

University of Oxford

Ruth Garside

University of Exeter Medical School, University of Exeter

Jo Thompson Coon

University of Exeter Medical School, University of Exeter

Rhiannon Evans

Cardiff University

Systematic Review

Keywords:

Posted Date: July 8th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10117761/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Background

Artificial intelligence (AI) is a broad field encompassing various techniques, which include algorithms that learn from data to perform automated tasks without explicit human programming. Public health aims to protect and improve the health of communities and populations at local, regional, national, and global level, but typically suffers from limited resources. While AI may lead to improved efficiency or effectiveness of public health interventions, adoption of AI interventions should be informed by evidence of effective use in practice. This review summarises how AI-assisted interventions are currently used by public health professionals.

Methods

We conducted a scoping review to identify the current uses of AI in public health. We searched Medline (Ovid), Embase (Ovid), Web of Science (Core Collection), and Scopus for relevant reports published between 01 January 2020 and 24 July 2025. Two reviewers independently screened reports. Eligible reports were primary studies (experimental or observational), commentaries, and editorials that discussed how AI tools have been used in practice in public health contexts relevant to high income countries. Data were extracted by one reviewer and checked by a second.

Results

We screened 5,400 records and 155 full texts, including 8 reports. These reports discussed a variety of uses of AI tools across public health domains: three reports used chatbots in different settings, one report used machine learning to forecast health service demand, two reports used AI for symptom checking, and two reports used AI for evidence synthesis. While assessing the effectiveness of AI tools for public health was beyond the scope of this review, and most reports did not robustly evaluate effectiveness, the evidence for effectiveness of AI tools was nonetheless mixed. One report described a chatbot that users found helpful, while another found a specific chatbot to be ineffective, highlighting its potential to undermine trust in the organisation that developed it. Another report highlighted both benefits and drawbacks of a large language model used to communicate with socially isolated individuals. The main finding of this review was that even with a comprehensive search across multiple databases, there were few reports of how AI tools are being used in practice in public health contexts.

Conclusion

The lack of evidence from public health practice makes it difficult to know which uses of AI, if any, would be useful to adopt into practice. In the absence of evidence, new interventions, including those using AI,

should be evaluated to ensure they are effective (and cost-effective). These evaluations should be widely and accessibly disseminated to create an evidence base that other public health teams could use to inform their own decisions around AI tools.

Registration

Zenodo 15971109

Background

Artificial intelligence (AI) is a broad field encompassing various techniques, including machine learning, natural language processing, and generative AI (1). AI learns from data to perform automated tasks without explicit human programming, using algorithms integrated into systems and tools (2–4). While the theoretical groundwork for machine learning was laid in the 1980s, recent advancements in computing power and the development of large language models (LLMs) like OpenAI's ChatGPT and Google's Gemini have sparked renewed interest in AI (5). These models, with more user-friendly interfaces and general-purpose applications, offer the potential to tackle a wide range of problems (6). However, the rapid pace of development in AI generally, and in LLMs specifically, means new evidence assessing the effectiveness of interventions using AI tools is constantly required, as newer versions of these tools may have differential effectiveness.

Public health is an area where AI is increasing being used with the aim of harnessing potential improvements in efficiency and effectiveness. This could be particularly useful, given the broad scope, important aims, and often limited resources of public health. In our recent reviews, we found several examples of chatbots, machine learning, and other AI-assisted interventions aiming to reduce engagement with alcohol drinking, smoking, and physical activity (7, 8), and 44 potential uses of AI within public health (9).

Public health activities are often organised and described under three interconnected domains: health improvement; health protection; and healthcare public health (10). Health improvement focusses on improving health and wellbeing and addressing inequalities through action on the wider determinants of health and encouraging individual behaviour change. Health protection focusses on the prevention and control of infectious diseases and emergency responsiveness. Healthcare public health focusses on ensuring that healthcare services are evidence-based, safe, effectively and equitably delivered. AI tools could potentially be used across all these public health activities to improve services.

However, many reports on the potential effectiveness of AI interventions in public health use evidence from trials, or from simulated or retrospective data. When moving into practice, the effectiveness of an intervention may be reduced, particularly in non-ideal settings (11). There may also be insurmountable barriers to adopting an intervention: for example, a lack of capacity, resources, or knowledge, particularly

for technologically complex interventions (11). Thus, it is important to evaluate how an intervention performs in routine use and in real-world contexts.

The purpose of this scoping review is therefore to identify how AI is currently used in public health, that is, how AI-assisted interventions are currently used by public health professionals. In doing so, the review will provide a snapshot of the evidence base for using AI tools in public health practice, helping to inform which tools could be used to shape future public health practice. Scoping reviews are ideally placed to map the research landscape, detailing what has already been examined to hopefully illuminate promising areas.

Methods

This scoping review systematically summarised evidence from primary studies, commentaries, and editorials to answer the following review question:

1. How is AI currently used in public health?

This review formed part of a broader programme of work evaluating the past, current, and potential uses of AI in public health (7–9). The protocol was registered on Zenodo: 15971109 (12). This review is reported using the PRISMA Extension for Scoping Reviews (PRISMA-ScR) (13). All data produced in this review are contained in the manuscript and its supplementary information files. We did not use AI in any part of the review process or when writing this article.

Stakeholder Engagement

We conducted stakeholder engagement across our broader programme of work across two phases of activity. The first phase involved protocol refinement and confirmation, with the aim of defining concepts related to AI and understanding comprehension of different models and platforms. The second phase gained feedback on initial synthesis findings, with consideration of how public health systems may support diffusion and delivery of AI-based interventions.

We conducted consultation meetings with two UK-based communities of practice at each phase. One community comprised nine practitioners from the NIHR funded Health Determinants Research Collaborations (HDRCs), with experience of public health service delivery. HDRCs are Local Government funded infrastructures intended to develop research capacity and capability to improve evidence-based decision-marking. The second community, also comprising nine people, included strategic directors with significant leadership roles for AI in public health across data science institutes, public health agencies, and national Local Government networks.

Eligibility criteria

We included primary studies (experimental or observational), commentaries, and editorials published or preprinted in English from 01 January 2020 that discussed how AI tools are currently used in public

health contexts relevant to high income countries (as defined by the World Bank (14)). We defined public health contexts as those where the primary aim was to improve public health, typically involving local and national Government-funded public health organisations delivering interventions or making policy decisions. We excluded pandemic contexts and clinical settings, for instance reports where reports only considered roles of AI that used NHS resources. Reports were only included if there was evidence the AI tool had been used in public health practice; reports of potential uses of AI in public health contexts without evidence of having been used in practice were excluded, even if, for example, the AI tool was evaluated in a randomised controlled trial. We defined AI broadly as algorithms integrated into systems and tools that can learn from data to be able to perform automated tasks without explicit human programming.

In the sole post-hoc change to the protocol, we refined the eligibility criteria to exclude outbreaks, including local disease modelling and prediction, as AI use in outbreaks has been extensively researched (15).

Search strategy

We used one AI search facet in the search strategy. This facet drew upon an unpublished filter created by members of the InterTASC (Technology Appraisal Support Collaboration) Information Specialists' Sub-Group (ISSG), as well as filters hosted on the associated ISSG Search Filter Resource (16). We combined the AI search facet with broad public health terms, as well as those reflecting activities linked to the research question, such as policy making and surveillance. These terms were derived from iterative searching, with keywords and MeSH terms harvested from relevant papers and reference lists. Search strategies for all databases are presented in **Supplementary Information A**.

Information sources

We conducted searches in four bibliographic databases: Medline (Ovid); Embase (Ovid); Web of Science (Core collection); and Scopus. Searches were automatically deduplicated using DOI matching using Stata 19 (17). Searches were conducted up to 24 July 2025.

Selection of sources of evidence

Two reviewers (SH and CT) independently screened titles and abstracts in Covidence (18), and full texts in a standardised Excel sheet. Titles and abstracts included by at least one reviewer on eligibility assessment progressed to full-text screening. Conflicts in assessment of full texts were resolved through discussion, with no necessity to involve a third party.

Data charting process and items

One reviewer (SH) completed all data extraction, which was checked by a second reviewer (CT). For each included report, we extracted information into standardised Excel sheets about the report characteristics

and methodology, the public health context, the type of AI and how it was used, and any outcomes associated with the performance of the AI intervention.

Synthesis of results

We summarised the evidence narratively, separately for each report.

All data generated or analysed during this review are included in this published article and its supplementary information files.

Ethics

Ethical approval

for this review was not required. We conducted consultations with stakeholder groups in accordance with any ethical requirements stipulated by the organisations and research studies that recruited participating members.

Results

Figure 1 shows the PRISMA diagrams for the flow of references. **Supplementary Information B** contains full text screening results for reports and full data extractions for included reports.

We screened 5,400 titles and abstracts and 155 full texts, ultimately including 8 reports that discussed AI tools used in public health practice. Four reports involved health improvement (19–22), four involved healthcare public health (21–24), and five involved health protection (20, 22, 24–26); four of these reports involved multiple public health contexts (20–22, 24), see Fig. 2.

Two reports were from the UK (22, 23), two from Canada (24, 25), and one each from Singapore (26), South Korea (21), the US (19), and worldwide (20). Five reports assessed the accuracy of AI tools (22–26), two reports conducted interviews to qualitatively assess the utility of AI tools (19, 21), and one report used structured interactions to assess the conversational and informational performance of an AI tool (20).

Table 1 provides an overview of the AI tools identified, while narrative summaries of included reports are provided below.

Table 1
Report characteristics

Reference	Country, report dates, participants	Report type, public health context(s)	Outcome measurement(s)	Intervention description
Baez, 2025 (19)	US June to July 2022 19 participants interviewed	Qualitative Health improvement	Open coding of interview recordings	**Intervention**: WIC chatbot (dialog-based): Maya **Intervention description**: The chatbot helped clients to search for local WIC offices and approved stores, report a lost or stolen card, and upload documentation, and answered client questions.
Germani, 2025 (20)	Worldwide 2024 or 2025 No participants	Structured interaction Health improvement & Health protection	Inductive, open, and deductive coding of S.A.R.A.H's responses to ChatGPT prompts and responses	**Intervention**: WHO chatbot (generative AI): S.A.R.A.H **Intervention description**: The chatbot provided accessible and reliable health information while combating misinformation.
Hyun Lim, 2024 (24)	Canada (PHAC) 2020 to 2023 No participants	Accuracy Health protection & Healthcare public health	Correct screens, validated by human screening	**Intervention**: Evidence review software (natural language processing): DistillerSR Artificial Intelligence System (DAISY) **Intervention description**: To screen citations based on a pre-specified eligibility criteria.
Jo, 2025 (21)	South Korea 2023 21 public agency workers	Qualitative Health improvement & Healthcare public health	Open coding of interview recordings	**Intervention**: CareCall, a phone call chatbot (LLM): HyperCLOVA **Intervention description**: CareCall mimicked the conversational style of a social worker to check on care recipients. The LLM accessed long-term memory to personalise open-ended conversations, including by asking follow-up questions. Public

Acronyms: AI = artificial intelligence, BERT = Bidirectional Encoder Representations from Transformers, LLM = large language model, PHAC = Public Health Agency of Canada, RSV = respiratory syncytial virus, UK = United Kingdom, UKHSA = UK Health Security Agency, US = United States, WIC = Special Supplemental Nutrition Program for Women, Infants, and Children

Reference	Country, report dates, participants	Report type, public health context(s)	Outcome measurement(s)	Intervention description
				health agency workers were notified of potential concerns.
Morbey, 2024 (23)	UK (UKHSA) 2022 to 2023 No participants	Accuracy Healthcare public health	Peak intensity and timing errors	**Intervention**: Short-term forecasting models for syndromic surveillance (machine learning): 2 machine-learning pipelines, to: 1) select and train forecast models, and 2) create daily forecasts **Intervention description**: Machine learning for forecasting RSV from NHS 111 calls for "cough" and emergency department attendance for "acute bronchitis", restricted to children aged 5 years and younger, with the aim of forecasting peak health care demand in the short-term. The random forest learning method with seasonality modelled by Fourier transformations was ultimately selected from the 252 trained models as it had the least error. This model was used to produce 1 to 28-day ahead forecasts using latest available UKHSA data (winter of 2022 to 2023).
Siedlikowski, 2021 (25)	Canada Launched March 2020 No participants	Accuracy Health protection	Overall accuracy and out-of-distribution accuracy	**Intervention**: COVID-19 chatbot (machine learning): Chloe **Intervention description**: Chloe was designed to help users self-assess COVID-19 symptoms, and to answer common questions in both English and French about COVID-19, which were clustered using unsupervised

Acronyms: AI = artificial intelligence, BERT = Bidirectional Encoder Representations from Transformers, LLM = large language model, PHAC = Public Health Agency of Canada, RSV = respiratory syncytial virus, UK = United Kingdom, UKHSA = UK Health Security Agency, US = United States, WIC = Special Supplemental Nutrition Program for Women, Infants, and Children

Reference	Country, report dates, participants	Report type, public health context(s)	Outcome measurement(s)	Intervention description
				machine learning. COVID-19 information used to train the chatbot was scraped from vetted online information sources. A BERT model was used to detect and remove out-of-distribution questions, which were not answerable with available information sources used by the chatbot. Users could freely ask questions, which the chatbot parsed through natural language understanding.
Simmons, 2025 (22)	UK (UKHSA) 2022 to 2025 Deduklick: 8 reviews; Rayyan and Eppi-Reviewer: 1 review	Accuracy Health improvement, Health protection & Healthcare public health	Deduklick: comparison with Leeds Method and general feedback; Rayyan and EPPI-Reviewer: how many records had to be screened to find all records in the prioritised list	**Intervention**: Deduklick (AI, not specified); Rayyan Prediction Classifier (Support Vector Machine classifier); EPPI-Reviewer Priority Screening (machine learning) **Intervention description**: Deduklick: removing duplicate records in reviews (no further details); Rayyan and Eppi-Reviewer: prioritise records when screening based on previous inclusion decisions.
Tan, 2024 (26)	Singapore 2022 200 images	Accuracy Health protection	Sensitivity and specificity of symptomatic Mpox infection	**Intervention**: HeHealth app (modified convolutional neural network) **Intervention description**: To examine images of penises, based on training from 1,000 images of penises with mpox positive rashes and 1,000 images of penises without mpox rashes. Additional information about the geographical prevalence of mpox and latest positive

Acronyms: AI = artificial intelligence, BERT = Bidirectional Encoder Representations from Transformers, LLM = large language model, PHAC = Public Health Agency of Canada, RSV = respiratory syncytial virus, UK = United Kingdom, UKHSA = UK Health Security Agency, US = United States, WIC = Special Supplemental Nutrition Program for Women, Infants, and Children

Reference	Country, report dates, participants	Report type, public health context(s)	Outcome measurement(s)	Intervention description
				patient statistical data was used to predict the likelihood of mpox.
Acronyms: AI = artificial intelligence, BERT = Bidirectional Encoder Representations from Transformers, LLM = large language model, PHAC = Public Health Agency of Canada, RSV = respiratory syncytial virus, UK = United Kingdom, UKHSA = UK Health Security Agency, US = United States, WIC = Special Supplemental Nutrition Program for Women, Infants, and Children				

The current roles of AI in public health

Baez et al interviewed 19 mothers from the United States (US) in 2022 to determine whether a chatbot on the Texas Special Supplemental Nutrition Program for Women, Infants, and Children (WIC) website was useful, and how it could be improved (19). The chatbot was called Maya, and helped clients to search for local WIC offices and approved stores, report a lost or stolen card, and upload documentation, while answering client questions. Participants stated that Maya helped them navigate and find information on the Texas WIC website, and that Maya was more efficient for obtaining information (particularly about their benefits) than other practices, for example, by visiting a clinic, calling a hotline, or emailing. As of January 2026, Maya is still active on the Texas WIC website (27).

Germani et al assessed the conversational and informational performance of the World Health Organization (WHO) chatbot, S.A.R.A.H, by presenting it with 10 statements generated by GPT- 3.5-turbo that contained health-related misinformation (20). S.A.R.A.H was developed to provide accessible and reliable health information and combat misinformation. The authors noted that some of S.A.R.A.H's responses failed to directly address user concerns (redirecting the conversation instead), lacked conversational skills associated with empathy, and failed to provide relevant evidence-based information to debunk misinformation. However, the authors also noted that one of S.A.R.A.H's responses acknowledged user concerns, suggested actional points, and showed artificial empathy. Overall, the authors concluded that S.A.R.A.H had "all the ingredients to serve not only as a useless tool in the quest to improve the information ecosystem and reduce the dangers infodemics pose to public health but also as a potential threat to erode trust in the WHO itself" (20). As of January 2026, S.A.R.A.H is no longer active on the WHO website (28).

Hyum Lim et al, within the Public Health Agency of Canada, used the DistillerSR Artificial Intelligence System (DAISY) to assist in screening of reports, and compared the accuracy of screening for relevant COVID-19 citations with human reviewers between 2020 and 2023 (24). DAISY uses natural language processing and human-coded training data to automatically screen the titles and abstracts of citations. DAISY correctly classified 98% of the 21,729 citations that had undergone human review, and found relevant citations missed by human reviewers (around 3% of included citations). As of January 2026, DistillerSR continues to use AI classifiers, although potentially not DAISY specifically (29).

Jo et al interviewed 21 public agency workers in South Korea in 2023 to determine their expectations of, and experiences with, CareCall, an LLM-based chatbot that spoke with socially isolated people through phone calls (21). CareCall is being used in South Korea to supplement the role of public agency workers in monitoring socially isolated people, including their health. The chatbot mimics the conversational style of a social worker to check on care recipients, and accessed long-term memory to personalise open-ended conversations, including by asking follow-up questions. Public health agency workers were notified of potential concerns, which could then be followed up. Overall, adoption of CareCall seems to have been viewed positively by the workers interviewed. They felt that CareCall expanded public reach and allowed for more frequent monitoring, but often without the necessary resources at implementation, which increased and diversified workload. Unlike previously, there was no need for dedicated devices (such as sensor systems and smart speakers) as CareCall uses existing phone lines. In this regard, CareCall was therefore cheaper and reduced barriers to adoption. Different (additional) care needs were also communicated, allowing workers to provide necessary support that may otherwise have been missed. However, some workers felt some participants who did not answer calls from CareCall did so as they did not like talking to an AI chatbot, and this added to workloads by necessitating callbacks to check participants were safe.

Morbey et al, within the UK Health Security Agency (UKHSA), created and piloted pipelines to select and use machine learning models to forecast healthcare service demand in the UK in 2022 and 2023 (23). Specifically, the machine learning models forecasted the peak timing and intensity of NHS 111 calls for "cough" and emergency department attendance for "acute bronchitis" over 28 days, restricted to children aged 5 years and younger. Seven machine learning methods were trained on data from before October 2022 (80% training, 20% validation), with four options for seasonality, secular trend, quadratic terms for current activity, and 3-day moving average, for a total of 252 alternative model specifications, each trained to forecast one to 28 days ahead inclusive. The model with the smallest weighted combination of intensity error (difference between the predicted and observed peak demand) and timing error (difference between the predicted and observed peak day) was selected and re-trained on all available data, and historical errors incorporated into forecast uncertainty. The random forest learning method with seasonality modelled by Fourier transformations was ultimately selected. This model was then used to produce one to 28-day ahead forecasts using latest available UKHSA data (winter of 2022 to 2023), which was compared to the observed peaks in NHS 111 calls and emergency department attendance. As there was an unprecedented demand for NHS 111 calls for cough, the forecasts could not predict peak demand until more data had been acquired, though the forecasts were able to correctly predict peak emergency department acute bronchiolitis attendance.

Siedlikowski et al discussed the development of a chatbot released in March 2020, Chloe, that was designed to help users self-assess COVID-19 symptoms and to answer common questions in both English and French about COVID-19 (25). The COVID-19 information used to train the chatbot was scraped from vetted online information sources. A Bidirectional Encoder Representations from Transformers model was used to detect and remove out-of-distribution questions, which were not answerable with available information sources used by the chatbot. Questions themselves were

clustered using unsupervised machine learning. Users could freely ask questions, which the chatbot parsed through natural language understanding. The overall accuracy of Chloe was 81.8% in English and 71.3% in French, and the out-of-distribution accuracy was 90.5% in English and 70.6% in French. As of January 2026, Chloe is no longer active on its website (30).

Simmons et al, within the UKHSA, tested and subsequently adopted 3 different AI-assisted interventions to help with conducting evidence reviews in the UK between 2022 and 2025 (22). They compared Deduklick, software that uses AI to identify and remove duplicate records before screening, with manual deduplication in Endnote (using the Leeds Method) across 8 reviews, concluding that although Deduklick removed slightly fewer references, it saved considerable time and was adopted into use. They subsequently used both Rayyan's Prediction Classifier (a support vector machine classifier) and EPPI Reviewer's Priority Screening (machine learning) to conduct priority screening in one review each. For Rayyan, all eventually included records were in the prioritised list after 32% of records had been screened, and for EPPI-Reviewer, 99% of eventually included records were in the prioritised list after 30% of records had been screened. Both tools have been adopted into use.

Tan et al developed and tested an AI-assisted app that detected the presence of symptomatic mpox in images of penises in 2022 (26). The app used a modified convolutional neural network trained on 1,000 images of penises with mpox rashes and 1,000 images of penises without mpox rashes, and additional information about the geographical prevalence of mpox and latest positive patient statistical data, to predict the likelihood of symptomatic mpox in new images of penises. The authors stated that the app had a sensitivity of 90% and specificity of 87%, and that between June and October 2022, the app had analysed 21,000 images of penises for mpox. The app is also used to predict the likelihood of other symptomatic sexually transmitted infections, and as of January 2026, is still available on its website.

Discussion

This scoping review summarised evidence for the current use of AI in public health. We identified a total of eight reports that discussed current uses of AI across all three public health domains. Uses of AI included using AI tools for symptom checking and conducting evidence reviews for public health, using AI tools to address health concerns and misinformation and connecting individuals to services and support, and using machine learning to forecast health service demand.

While assessing the effectiveness of AI tools was beyond the scope of this review, the evidence from included reports was mixed. For example, while the Maya chatbot seemed helpful to participants (19), the S.A.R.A.H chatbot was considered a "potential threat to erode trust in the WHO" (20), and CareCall had both advantages and disadvantages for public agency workers and wasn't evaluated by those at whom it was targeted. (21). Equally, in our reviews of the effectiveness of public health AI-assisted interventions for alcohol use, smoking, and physical inactivity, there was ultimately little evidence to suggest that AI-assisted interventions were any more or less effective than non-AI-assisted interventions (7, 8).

All reports included in this review were authored by larger public health organisations; for example, national and global public health bodies, who may be more likely than smaller and more local public health teams to possess the infrastructure and capabilities to use and evaluate AI technologies. These authorship imbalances may also reflect varying levels of organisational readiness to use AI tools, relating to divergences in workforce capacity, training, infrastructure, equipment, and resources. If so, this variability may perpetuate differences in how AI is used and evaluated across all aspects of public health, potentially leading to inequities if AI tools are more effective or efficient than alternatives. However, the observed imbalances may also simply reflect differential publishing practices between larger and smaller public health organisations: smaller public health teams may be using AI tools, but not publishing their experiences with academic journals. Indeed, our PPI groups noted that larger organisations may be slower to adopt AI due to more rigid governance systems than smaller organisations, which may be more flexible.

The rationale for identifying current uses of AI in this review was to determine how much real-world evidence exists, and how easily accessible AI tools are in public health practice. Using a comprehensive search across four databases, we were only able to identify eight reports. On the other hand, in our review of the potential roles of AI in public health, we identified 67 reports covering 44 potential uses of AI across eight public health contexts (9). This may mean that there are plenty potential uses of AI, but little accessible evidence supporting those uses of AI in current practice.

The limitations in research evidence are important: indeed, the effectiveness of interventions tends to reduce when moving into practice, and unintended consequences may also become apparent. For example, the CareCall AI expanded public reach to socially isolated people, but this also increased the workload of public agency workers, as their case load increased and some people did not want to communicate with AI (21). In our review of the potential roles of AI in public health contexts, we identified a large number of risks, including the perpetuation or reinforcement of health inequalities and discrimination, concerns around transparency, data management, privacy, security, and confidentiality, and ethical, legal, and regulatory concerns, all of which may contribute to unintended and unpredictable consequences of AI adoption within public health (9).

Without underlying evidence of effectiveness in practice, and the numerous and critical risks, it would seem extremely difficult for public health teams to know which uses of AI, if any, would be useful to adopt into practice.

Stakeholder Consultations

Stakeholders recognised the potential of AI-assistance in delivering behaviour change interventions but highlighted a number of key knowledge gaps that need to be attended to moving forward. First, stakeholders noted limited reporting of the extent to which interventions were co-produced with the target audience, which raises important considerations of how co-production methods can be optimally used in AI-based research. Second, stakeholders discussed the need for participants to have readiness and motivation to engage with AI, and if preparty skill development may be required. Third, stakeholders

queried the strength of implementation science in this area. For example, they considered if apps would be hosted by developers or Local Authorities, and what resources and capabilities would local public health teams require to support AI-assisted approaches. Fourth, stakeholders commented on the methodological limitations of evaluation, and whether new appraisal tools would need to be developed to account for the specific challenges and limitations associated with AI. Specifically, there was discussion of whether AI approaches have the potential to exacerbate health inequities and causes unintended harms, which have not yet been fully theorised or translated into appraisal criteria.

Limitations of this review

Despite the comprehensive search strategies, it is possible that relevant reports may have been missed from searches. Equally, despite duplication of screening, relevant reports may have been missed during screening, and relevant information may have been missed during data extraction of included reports. However, the purpose of this scoping review was only to determine how much real-world evidence exists and is easily accessible for using AI tools in public health: if we missed relevant evidence because it was not accessible, then public health teams may also miss that evidence when searching for the effectiveness of AI tools in public health.

Additionally, we focussed solely on AI tools used within public health. These tools, or entirely different AI tools that could or have been used within public health but were not included in this review, may have evidence of use in other fields. However, the utility of tools may differ across fields, in a similar way to how effectiveness of interventions in trials may differ from effectiveness in practice.

Finally, defining and identifying both public health and which interventions are AI-assisted was challenging, as the definitions for both public health and AI are broad, and some reports provide little or no information about the AI aspect of AI-assisted interventions. We aimed to avoid missing relevant reports by taking a sensitive approach to screening and using broad inclusion criteria.

Implications for research and practice

When considering adopting an AI intervention, public health professionals should carefully consider the potential effectiveness and cost-effectiveness alongside any potential risks of harm, even for non-clinical interventions (31). This review identified some current uses of AI tools that look beneficial (the WIC chatbot, some uses in evidence synthesis), though the description of the S.A.R.A.H chatbot as both ineffective and a “potential threat to erode trust in the WHO” provides a stark warning of the risks of harm.

One key finding in this review is that while there are a multitude of potential uses of AI tools in public health, there is limited evidence for their use in practice. In the absence of evidence, new interventions, including those using AI, should ideally be evaluated by public health teams to ensure they are effective (and cost-effective). These evaluations should be widely and accessibly disseminated to create an evidence base that other public health teams could use to inform their own decisions around AI tools.

This poses the opportunity and need for researchers and practitioners to work more closely to share skills and resources.

Conclusion

This scoping review summarised evidence for the current uses of AI in public health. Across four public health contexts, we identified only eight reports of current uses of AI, which ultimately had mixed evidence of effectiveness. This lack of evidence from public health practice makes it difficult to know which uses of AI, if any, would be useful to adopt. In the absence of evidence, new interventions, including those using AI, should ideally be evaluated to ensure they are effective (and cost-effective). These evaluations should be widely and accessibly disseminated to create an evidence base that other public health teams could use to inform their own decisions around AI tools.

Declarations

Disclosure of Interests:

Author Contribution

Sean Harrison: Conception and protocol, screening, data extraction, risk of bias assessment, narrative synthesis, first draft and editing of manuscriptClaire Tatton: Screening, editing of manuscriptG.J. Melendez-Torres: Conception and protocol, editing of manuscriptJoelle Kirby: Conception and protocol, searching, editing of manuscriptSophie Robinson: Conception and protocol, editing of manuscriptDaniel Mutanda: Conception and protocol, stakeholder engagement, editing of manuscriptAlisha Davies: Stakeholder engagement, editing of manuscriptJessica M Armitage: Conception and protocol, editing of manuscriptRabeea'h Waseem Aslam: Conception and protocol, editing of manuscriptTom Arthur: Editing of manuscriptJoht Singh Chandan: Conception and protocol, editing of manuscriptRuth Garside: Conception and protocol, editing of manuscriptJo Thompson Coon: Conception and protocol, editing of manuscriptRhiannon Evans: Conception and protocol, stakeholder engagement, editing of manuscript

Acknowledgement

Jenny Lowe provided professional services support for the review. National Institute for Health and Care Research Public Health Research Programme (NIHR175419). This research was supported by the National Institute for Health and Care Research Applied Research Collaboration South West. Prof G.J. Melendez-Torres is an NIHR Senior Investigator. Prof JS Chandan is an NIHR Research Professor (NIHR306365). The views expressed in this publication are those of the author(s) and not necessarily those of the National Institute for Health and Care Research. This work was supported by DECIPHer, funded by Welsh Government through Health and Care Research Wales. We are grateful to the following

advisory groups for their support during this review: the NIHR Health Determinants Research Collaborations (HDRCs) in England and Wales (including Zoe Lancelott, Sophia Looney, Sarah Markham, Sarah McGarrol, and Francis McKay); the Health Foundation; the Faculty of Public Health “AI and Digital Public Health” Special Interest Group (including Elliott Roy-Highley); the Local Government Association AI Network; and public contributors (including Sian Harding).

References

1. Sarker IH. Deep Learning: A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions. SN Comput Sci. 2021;2(6):420. https://doi.org/10.1007/s42979-021-00815-1
2. World Health Organization. The role of artificial intelligence in sexual and reproductive health and rights. World Health Organization Technical Brief. 2024. https://www.who.int/publications/i/item/9789240090705
3. Russell S. Artificial intelligence: a modern approach, fourth edition. Upper Saddle River (NJ): Prentice Hall. 2022. https://aima.cs.berkeley.edu/
4. World Health Organization. Ethics and governance of artificial intelligence for health: WHO guidance. 2021. https://iris.who.int/handle/10665/341996
5. Mohammedabadi SMS, Kara BC, Eyupoglu C, Uzay C, Tosun MS, Karakuş O. A Survey of Large Language Models: Evolution, Architectures, Adaptation, Benchmarking, Applications, Challenges, and Societal Implications. Electronics. 2025;14(18):3580. https://doi.org/10.3390/electronics14183580
6. Jungwirth D, Haluza D. Artificial Intelligence and Public Health: An Exploratory Study. Int J Environ Res Public Health. 2023;20(5). https://doi.org/10.3390/ijerph20054541
7. Harrison S, Kirby J, Tatton C, Armitage JM, Evans R, Lee SI, et al. The role of Artificial Intelligence in interventions to reduce physical inactivity: A systematic review Unpublished.
8. Harrison S, Tatton C, Kirby J, Das A, Robinson S, Evans R, et al. The role of Artificial Intelligence in interventions to reduce alcohol drinking and smoking: A systematic review Unpublished.
9. Harrison S, Tatton C, Kirby J, Robinson S, Mutanda D, Davies A, et al. The potential roles and risks of Artificial Intelligence in public health: A scoping review Unpublished.
10. Griffiths S, Jewell T, Donnelly P. Public health in practice: the three domains of public health. Public Health. 2005;119(10):907–13. https://doi.org/10.1016/j.puhe.2005.01.010
11. Victora CG, Habicht JP, Bryce J. Evidence-based public health: moving beyond randomized trials. Am J Public Health. 2004;94(3):400–5. https://doi.org/10.2105/ajph.94.3.400
12. Harrison S, Kirby J, Robinson S, Evans R, Aslam R, Chandan J, et al. The current role of Artificial Intelligence in public health: A scoping review. Zenodo. 2025. https://doi.org/10.5281/zenodo.15971108
13. 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.

https://doi.org/10.7326/M18-0850

14. The World Bank. World Bank Country and Lending Groups. 2025. https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups
15. Sadr H, Nazari M, Khodaverdian Z, Farzan R, Yousefzadeh-Chabok S, Ashoobi MT, et al. Unveiling the potential of artificial intelligence in revolutionizing disease diagnosis and prediction: a comprehensive review of machine learning and deep learning approaches. Eur J Med Res. 2025;30(1):418. https://doi.org/10.1186/s40001-025-02680-7
16. Glanville J, Lefebvre C, Manson P, Robinson S, Brbre I, Woods L. ISSG Search Filter Resource [Internet] York (UK): The InterTASC Information Specialists' Sub-Group. 2006. https://sites.google.com/a/york.ac.uk/issg-search-filters-resource/home
17. StataCorp. Stata Statistical Software: Release 19. College Station, TX: StataCorp LLC. 2025.
18. Veritas Health Innovation. Covidence systematic review software. Melbourne, Australia. www.covidence.org
19. Baez K, Biediger-Friedman L, Johnson CM, Stubblefield E, Escalera L, Markides BR. Exploring Client Perceptions on Gaining Infant Feeding Information Through the Texas Women, Infants, and Children (WIC) Chatbot. Int J Environ Res Public Health. 2025;22(2). https://doi.org/10.3390/ijerph22020193
20. Germani F, Spitale G, Biller-Andorno N. S.A.R.A.H. and the decline of trust in health information: a case study. Health Policy and Technology. 2025;14(3):101032. https://doi.org/10.1016/j.hlpt.2025.101032
21. Jo E, Kim YH, Ok SH, Epstein DA. Understanding Public Agencies' Expectations and Realities of AI-Driven Chatbots for Public Health Monitoring. 2025. https://doi.org/10.1145/3706598.3713593
22. Simmons Z, Bruce C, Thomas S, Lacey P, Marsh W, Rosenberg S, et al. Exploring the potential of artificial intelligence usage in the knowledge and evidence services of a public health body: A working group approach. IFLA Journal. 2025;51(3):750–65. https://doi.org/10.1177/03400352251342513
23. Morbey RA, Todkill D, Moura P, Tollinton L, Charlett A, Watson CH, et al. Using machine learning to forecast peak health care service demand in real-time during the 2022-23 winter season: a pilot in England, UK. medRxiv. 2024. https://doi.org/10.1101/2024.09.20.23296441
24. Hyun Lim S, Hersi M, Krishnan R, Montroy J, Rook B, Farrah K, et al. COVID-19 vaccine evidence monitoring assisted by artificial Intelligence: An emergency system implemented by the Public Health Agency of Canada to capture and describe the trajectory of evolving pandemic vaccine literature. Vaccine X. 2024;21:100575. https://doi.org/10.1016/j.jvacx.2024.100575
25. Siedlikowski S, Noel LP, Moynihan SA, Robin M. Chloe for COVID-19: Evolution of an Intelligent Conversational Agent to Address Infodemic Management Needs During the COVID-19 Pandemic. J Med Internet Res. 2021;23(9):e27283. https://doi.org/10.2196/27283
26. Tan RKJ, Perera D, Arasaratnam S, Kularathne Y. Adapting an artificial intelligence sexually transmitted diseases symptom checker tool for Mpox detection: the HeHealth experience. Sex

Health. 2024;21(3):SH23197. https://doi.org/10.1071/SH23197
27. TexasWIC. Texas WIC. 2026. https://texaswic.org/
28. World Health Organization. S.A.R.A.H. Smart AI Resource Assistant for Health. 2026. https://www.who.int/campaigns/s-a-r-a-h
29. DistillerSR. The DistillerSR Platform. 2026. https://www.distillersr.com/products/distillersr-systematic-review-software
30. Dialogue. Chloe's mission is now over. 2026. https://covid19.dialogue.co/#/
31. Bonell C, Jamal F, Melendez-Torres GJ, Cummins S. 'Dark logic': theorising the harmful consequences of public health interventions. J Epidemiol Community Health. 2015;69(1):95–8. https://doi.org/10.1136/jech-2014-204671

Figures

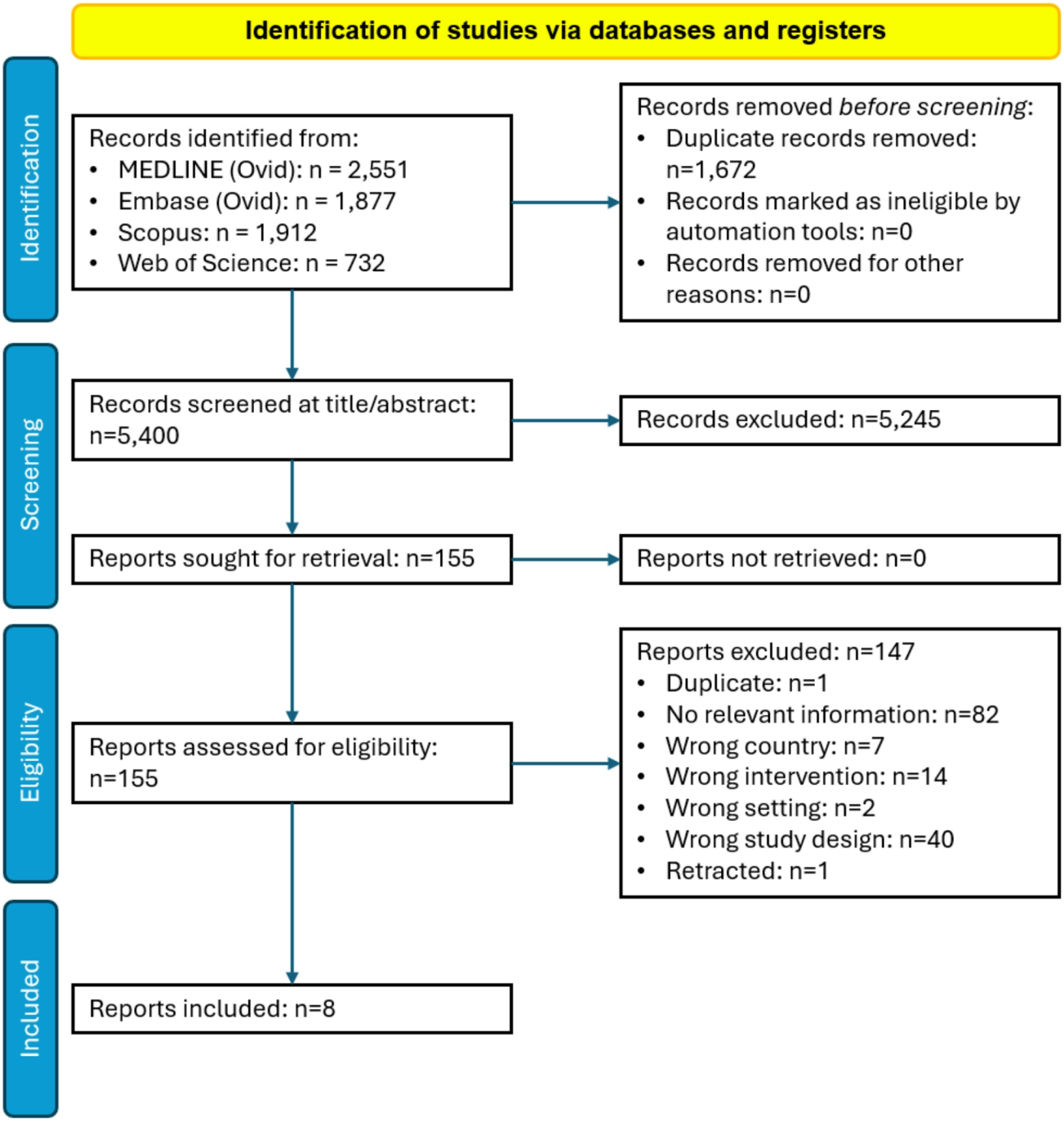


Figure 1

PRISMA diagram

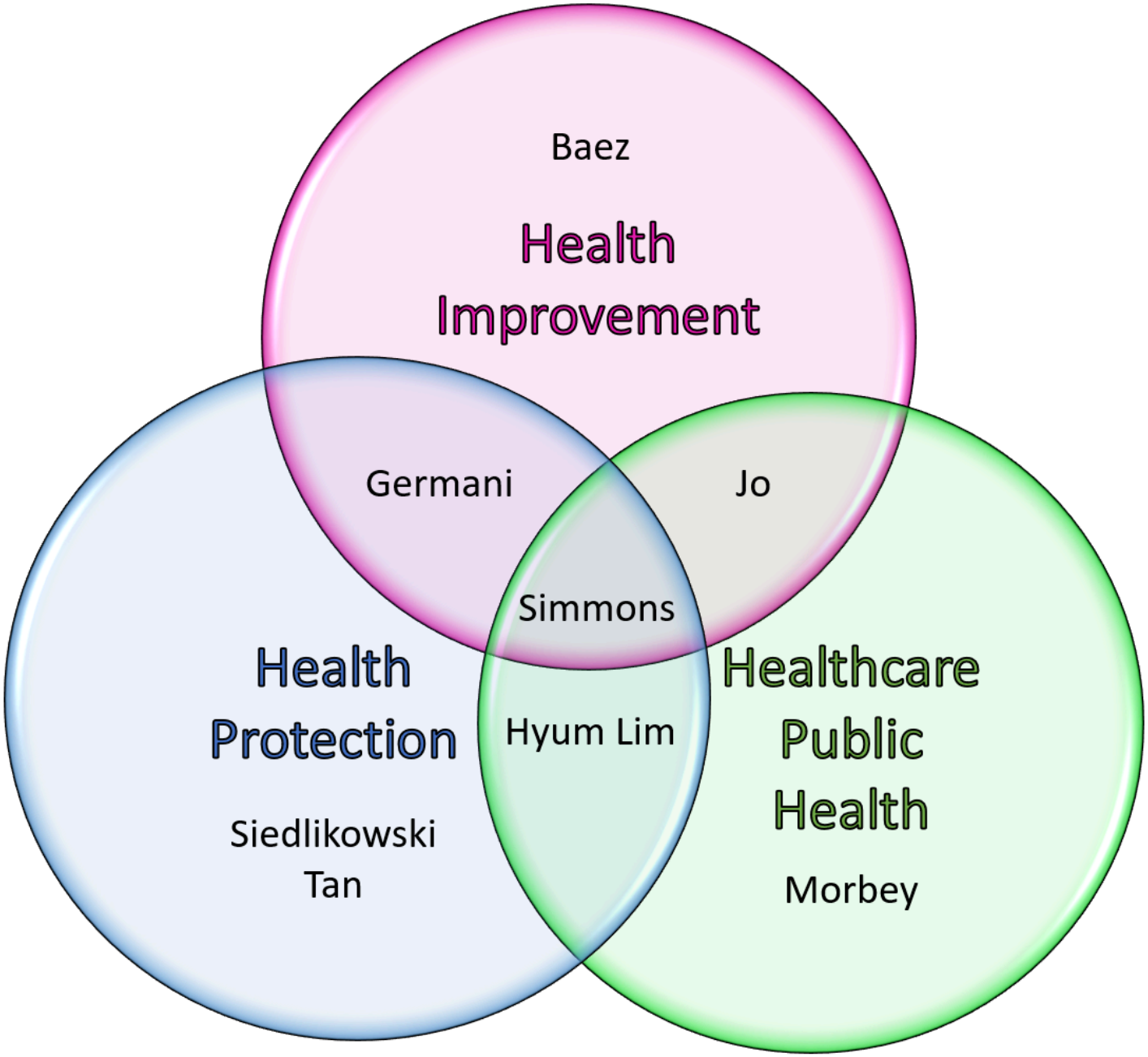


Figure 2

Public health contexts of included reports

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryInformationA.docx
- SupplementaryInformationB.xlsx