A Horizon Scan of Emerging Generative AI-Enabled Technologies for use in Healthcare

Authors: **Kate Lanyi, Katie Twentyman, Imogen Forsythe, Elizabeth Green, Jabez Barnabas and Andrew Mkwashi.**

Date: **May 2025**

Professor Dawn Craig
Director
T: +44(0) 191 208 2259
E: info@io.nihr.ac.uk
www.io.nihr.ac.uk

Copyright © National Institute for Health and Care Research Innovation Observatory (NIHR IO), Newcastle University

How to Cite This Document

Lanyi, K., Twentyman, K., Forsythe, I., Green, E., Barnabas, J., & Mkwashi, A. A horizon scan for emerging generative AI-enabled technologies for use in healthcare. NIHR Innovation Observatory; Newcastle upon Tyne, 2025. DOI: 10.6084/m9.figshare.29042408

Acknowledgements and Disclaimers

This study/project is funded by the National Institute for Health and Care Research (NIHR) [NIHR IO/project reference HSRIC-2016-10009]. The views expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care.

We would like to thank our stakeholders including the Department of Health and Social Care (DHSC) and NHS England who provided valuable expertise throughout all stages of this horizon scan and provided feedback on draft versions of this report.

Conflicts of Interest

No conflicts of interest to declare.

© 2024 This work is openly licensed via CC BY 4.0.

NIHR | Innovation Observatory

Newcastle University

Contents

How to Cite This Document 2

Acknowledgements and Disclaimers 2

Conflicts of Interest 2

Abstract 4

Glossary 5

Background 7

Aims 9

- Objectives 9

Methods 9

- Scope 10
- Eligibility criteria 10
- Inclusion criteria 10
- Exclusion criteria 11
- Search Strategy 11
- Sources: 11
- Search limits 12
- Record selection 12
- Data Extraction 12

Results 14

- Prisma diagram 14
- Type of Generative AI models 14
- Type of generative AI content 15
- Type of algorithm used 16
- NICE disease categories 17
- Developer type 20
- Developer location 21
- Technology readiness level 22
- Named technologies (UK) 23
- Generative AI content type 24

Implications and Recommendations 26

Conclusion 27

Appendix A: Clinical trial and funding search strategy 28

References 31

Abstract

Background:

Generative artificial intelligence (GenAI) is a type of AI that uses input data to generate new content such as images, text, videos or sounds. GenAI is rapidly emerging as a transformative tool in healthcare, offering advanced tools for diagnostics, decision support, image analysis, and personalised care. To support strategic planning and timely clinical adoption we undertook a horizon scan to identify GenAI-based technologies currently in clinical development. Additionally, we aimed to provide an overview of the GenAI landscape.

Methods:

We conducted searches in October 2024 across clinical trial registries (ClinicalTrials.gov, EudraCT, ICTRP, Scan Medicine) and funding portals (NIHR, UKRI) to retrieve records published between October 2021 and October 2024. Eligibility criteria required technologies to be regulated as a medical device, digital health technology or diagnostic, and explicitly utilising generative AI (e.g., LLMs, GANs, transformers) in a patient-facing clinical setting. We manually screened clinical trial and funding records against the eligibility criteria, then extracted data from eligible records. Extracted data were standardised across fields, including algorithm type, content generated, Generative AI (GenAI) model, developer, technology readiness level (TRL), and NICE disease area.

Results:

In total, 161 GenAI technologies were identified, with the majority (82%) being developed by non-industry entities and primarily classified as digital tools (55%). Text generation was the most common output (43%), followed by data augmentation (17%) and image generation (16%). Technologies generating text were most often at advanced development stages (TRL 7–9), suggesting they may be among the first GenAI applications to be widely implemented in clinical settings. In contrast, technologies generating 3D models tended to be at earlier TRL stages, reflecting a more exploratory phase of development. The majority of technologies did not specify the GenAI model or algorithm used; among those that did, large language models (LLMs) and reinforcement learning were most frequently mentioned. Key disease areas that GenAI technologies were being developed for included mental health (13%) and cancer (12%), though over a quarter of the technologies were not disease-specific. The USA led in clinical trial development activity, followed by China and the UK. Three UK-based technologies—Flok, Archer, and Spectroscopic Probe—were identified at higher TRLs (7–9).

Conclusions:

This horizon scan identified a diverse and expanding innovation landscape in Gen-AI technologies, being developed across various disease areas, TRLs, GenAI models, and types of content generated. These technologies demonstrated potential benefits across all aspects of patient care, including the most frequently generated content—text-based chatbots and virtual assistants providing mental health support. While these applications may help reduce pressures on mental health services, risks related to patient safety, hallucinations, and reduced human contact must be carefully considered. Additionally, poor reporting of model and algorithm types poses challenges for regulatory transparency and evaluation. These findings highlight the

need for ongoing surveillance and improved reporting standards to navigate this rapidly evolving landscape effectively.

Glossary

Bias	A systematic error that can lead to incorrect conclusions.
Artificial Intelligence (AI)	technology capable of performing advanced tasks that usually require human intelligence
Foundation Models	A type of generative AI that generates outputs from one or more inputs (prompts) in the form of different modalities such as human language instructions
Generative adversarial networks (GANs)	A machine learning model that uses two neural networks to make new data that resembles real data
Generative AI (GenAI)	A class of AI models that emulate the structure and characteristics of input data in order to generate derived synthetic content. This can include images, videos, audio, text, and other digital content.
Horizon scanning	A systematic examination of potential threats, opportunities and likely future developments including but not restricted to those at the margins of current thinking and planning
Large Language Model (LLM)	A type of AI program that has been trained on large datasets and can recognise and generate text, among other tasks
Deep Learning (DL)	A type of machine learning that uses artificial neural networks to learn from data.
Machine Learning (ML)	A type of AI that involves algorithms that are able to analyse data and improve performance over time
Neural networks	A neural network is a machine learning algorithm, or model, capable of making decisions in a manner similar to the human brain.
Reinforcement learning	A machine learning (ML) technique that trains software to make decisions to achieve the most optimal results.

Semi-supervised learning	A machine learning technique that uses labeled data alongside a large amount of unlabeled data to train a model
Supervised learning	Machine learning algorithms that are given labeled training to learn the relationship between the input and the outputs
Technology readiness levels (TRLs)	A tool used for measuring technology maturity. They define nine levels, ranging from basic research (TRL 1) to full-scale deployment (TRL 9)
Training data	A set of information used to "teach" a machine learning model how to perform a task by providing it with labeled examples which allows the model to learn patterns and make accurate predictions on new data
Unsupervised learning	Unsupervised machine learning algorithms are given unlabeled data and allowed to discover patterns and insights without any explicit guidance or instruction.
Variational autoencoders	A type of machine learning model that can generate new data similar to the data it was trained on.
Confabulations	The generation of inaccurate or misleading information by an AI system, when the AI "fills in the gaps" in a response based on patterns it has learned, rather than providing a fully accurate or factually correct answer.
Hallucinations	AI hallucination is a phenomenon wherein a large language model (LLM) perceives patterns or objects that are nonexistent or imperceptible to human observers, creating outputs that are nonsensical or altogether inaccurate.

Background

There is a growing interest in the utilisation of artificial intelligence (AI) based healthcare technologies to increase the efficiency of healthcare systems and improve patient care.[1] AI driven technologies can improve the accuracy, efficiency, and cost-effectiveness of disease diagnosis, and assist in processes such as patient monitoring, decision-making, guideline establishment, and optimising treatment regimens for individual patients.[2] Generative AI (GenAI) has the potential to rapidly transform the landscape of healthcare by unlocking new possibilities in diagnosis, treatment planning, patient care and drug discovery. Over the last three years, venture capital firms have invested more than 1.7 billion USD into GenAI solutions, with GenAI-enabled drug discovery and software coding getting the most funding.[3] By end 2025, it is expected that over 30% of new drugs and materials will be discovered using GenAI techniques, underscoring its transformative impact on drug discovery and material innovation. Unlike traditional AI which focuses on historical data analysis and pattern recognition, GenAI refers to a subset of AI that uses deep-learning models that can create new content such as text, images, music and code based on patterns and data it has been trained on.[4] The National Institute of Standards and Technology (NIST) defines GenAI as "the class of AI models that emulate the structure and characteristics of input data in order to generate derived synthetic content.[5] Once trained, GenAI models use various sampling techniques to produce new content. These techniques determine how the model selects from the learned patterns to create outputs.[6] Figure 1 below illustrates how new content can be created by GenAI.

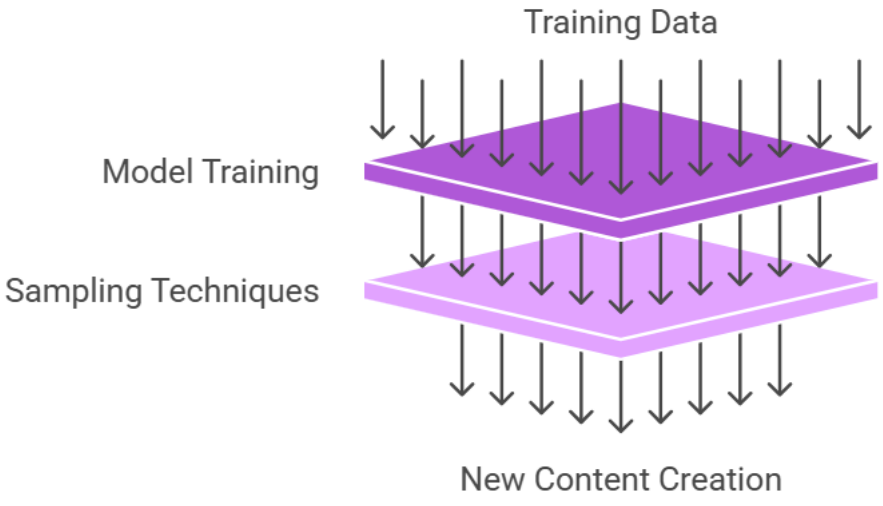


Figure 1: Overview of how generative AI creates new content.

As shown in Fig 1, the GenAI model takes an initial input, processes it through multiple layers, and generates a unique output. This process allows for the creation of diverse and innovative content, ranging from text to images, based on the input data and the model's training. The training data plays a crucial role in shaping the AI model's ability to generate content. It provides the foundational knowledge and patterns the AI uses to understand and transform input into new creations. The quality and diversity of

the training data directly impact the model's performance and the creativity of the output it produces.

Adoption of GenAI-enabled technologies into clinical care pathways has the potential to deliver game-changing advances in healthcare. "GenAI-enabled technology" refers to a technology (device) in which GenAI methods or models are integral to the technology's output or functionality. These technologies can provide benefits such as producing high-quality images for improving diagnostic accuracy, predicting molecular structures and simulating interactions between biological targets and drugs to accelerate drug discovery, producing personalised treatment plans for individual patients, generating insights that support clinical decision-making, and analysing patient data to generate new insights that support clinical decision-making.[7] Moreover, GenAI models can be utilised to facilitate medical training and simulations, as these models are capable of creating realistic patient simulations that adapt in real-time, allowing medical professionals to practice in a controlled environment.[8] Unlike traditional AI which may be used to analyse historical data, GenAI can create novel solutions or responses in real-time, leading to faster and more accurate decision-making and potentially improving patient outcomes in a clinical care setting. However, the innovative nature of these technologies are also associated with significant challenges and risks in the context of healthcare.[9] For instance, a GenAI model is commonly developed on datasets that are so large that it is almost impossible for a human developer to understand all of the dataset's contents during the development phase. As opposed to datasets that are used to develop other AI/ML models, datasets for GenAI model development can be intentionally broad and may not be initially tailored to a specific task. A significant issue with AI models, including some GenAI models such as LLMs, Generative Adversarial Networks (GANs), Variational Autoencoders (VAEs), and Recurrent Neural Networks is that they work as 'black boxes' so that it is not easy to interpret their decision making or how the model arrives at a specific output. [9]

AI models, including GenAI models, may exhibit harmful biases, which may stem from their training data. Such biases may lead to representational harm or perpetuate or exacerbate bias based upon race, gender, disability or other protected categories. Another cybersecurity risk to GenAI is data poisoning, in which an adversary compromises a training dataset used by a model to manipulate its outputs or operation. Malicious tampering with data or parts of the model could exacerbate risks associated with GenAI system outputs.[5] Additionally, AI models, including GenAI models, can also exhibit hallucinations, i.e., the generation of erroneous or false content to meet the programmed objective of fulfilling a user's prompt. In the context of GenAI models this can result in chatbots providing false or misleading medical advice or radiology technologies using GenAI models adding artificial anomalies to an image.[10]

Despite the possibility that GenAI-enabled products will benefit patients and public health in unique ways, they will also present new regulatory challenges to competent authorities. Regulators may find it difficult to apply a risk-based approach to classifying GenAI-enabled devices and determining their applicable regulatory requirements if their intended uses are not well defined[11]. Large language models pose specific challenges to regulators as they do not fit neatly within existing regulatory frameworks. The FDA and EU regulatory bodies classify software as a medical device or software that provides medical recommendations/diagnosis

respectively. LLMs are often general-purpose built models making it unclear whether they fall into the software as a medical device regulatory category. Further, LLMs are dynamic and continuously evolving unlike traditional static software which is easier to regulate and continually monitor.

GenAI-enabled healthcare technologies have therefore been identified as a high priority area for proactive horizon scanning given their potential to have both significant beneficial and disruptive impacts on healthcare systems. The growing investment and development in AI technologies, including generative AI, necessitates the early identification and assessment of these technologies to understand their potential impact on healthcare providers and patients. Horizon scanning is defined as the systematic examination of potential threats, opportunities and likely future developments including but not restricted to those at the margins of current thinking and planning.[12] Horizon scanning can assist in identifying these technologies, understanding key development trends, highlighting the unique challenges and opportunities they present, and identifying areas of unmet need where there is currently a lack of development.[13] This can help decision makers to identify implementation challenges early and therefore develop informed strategies to address any infrastructure, regulatory, or reimbursement barriers to fast implementation of GenAI-enabled health technologies.[9]

Aims

Overall, this project aims to provide a comprehensive overview of the current AI-based health technology landscape. The project deliverables include quarterly horizon scanning outputs, delivered in the form of interactive dashboards, technology radars and reports focusing on a particular subset of AI technologies, or a specific priority area identified in consultation with our critical stakeholders. These outputs will identify AI-based technologies currently in development. Additional data will highlight emerging development trends, the potential implications the technologies may have on the healthcare system, and any areas of unmet clinical need. The analysis undertaken will also consider the key factors driving the rapid development of AI-based healthcare technologies. Each quarter, the Innovation Observatory will focus on a specific area of AI technologies; this current report focuses on Gen-AI.

Objectives

- Identify and describe GenAI-based healthcare technologies currently in development.
- Provide an in-depth analysis of the GenAI-based healthcare landscape, including: disease areas, Gen-AI model utilised, algorithm type, geographical location and sponsor type.
- Identify emerging trends and their potential implications for healthcare provider and patients.
- Produce quarterly reports, dashboards and technology radars to disseminate up-to-date insights from this ongoing horizon scan.

Methods

We used horizon scanning methods as follows:

Scope

The scope of this horizon scan was to identify signals from clinical trials and funding sources relating to emerging GenAI-based technologies being developed for use in a clinical setting to improve patient care. GenAI-based technologies being developed to support healthcare systems or healthcare professionals rather than directly for patient care were not included in the scope of this scan. Technologies across all TRLs from initial technology concept through to technologies ready to be implemented in a clinical setting were included in the scope of this scan.

Eligibility criteria

Inclusion criteria

Medical device, digital health technology or diagnostic

To be eligible, technologies identified in the clinical trial must meet the criteria for being a regulated medical device, digital technology, or diagnostic as defined by EU Regulation 2017/755/ on medical devices (MDR) and EU Regulation 2017/746 on in vitro diagnostic medical devices (IVDR).

AI technology

The AI technology must also satisfy at least one of the following requirements to be considered AI:

- Clinical trial record uses keyword phrases specific to AI (e.g. “deep learning”, “machine learning”).
- One or more of the collaborators specialise in AI AND a relevant technology.
- A technology is described as “intelligent” AND a secondary source confirms the technology is AI-based.
- Significant computation is involved (big-data) AND a secondary source confirms the technology is AI-based.
- Machine learning methodology is described (e.g. an algorithm is “trained” using (big) data).
- Data is collected with the stated intention to apply machine learning.
- Existing AI-based technology (e.g. smart watch) being evaluated for clinical use.
- A novel AI-based technology is in development for: Risk assessment or risk mitigation at the individual patient level; direct patient health impact; or diagnostic application.

Generative AI

There should be explicit mention of terms related to generative AI within the clinical trial or funding grant (e.g. generative AI, deep-learning, neural networks, generative model, adversarial networks, transformers, GPT, large-language models, reinforcement learning).

The AI component is supplied with data, and uses this data to generate new content, new content could include:

- Automating tasks (e.g. text generation) and creating tools (e.g. ChatGPT, virtual assistants) that help to reduce workload for healthcare professionals.
- Image enhancement or automatic image labelling to improve diagnostic accuracy of clinicians.
- Analyse the 3D structure of proteins implicated in disease (e.g. receptors) and suggest properties of a new drug to treat the disease.
- Suggesting patient-specific rehabilitation plans based on data related to lifestyle, medical history, genetic information and patient preference.

In instances where the trial record indicates, but does not specify, that the technology is GenAI-based, a secondary search for published information on the technology was conducted using the clinical trial reference number, company name, or technology brand name as search terms.

Exclusion criteria

- Clinical trials using GenAI in the training and education of healthcare professionals.
- Clinical trials using GenAI to improve health outcomes at a hospital level or disaster management level rather than at an individual level.
- Primary research conducted in the clinical setting.
- Patient information or educational tools whose primary purpose is not to improve health outcomes.
- "Intelligent" technologies that gather enriched information but do not perform independent analysis
- The AI technology is only making predictions based on input data, it is not generating new content.

Search Strategy

Searches for this horizon scan were conducted in October 2024 across all sources outlined below. A comprehensive list of keywords and concepts were compiled using MeSH/key terms with Boolean operators, crafted by the IO's Information Specialist Team, based on key indicator reports. This list of terms was refined using scoping searches, testing, and analysis of the relevance of each concept. A detailed search strategy is provided in Appendix A.

Sources:

Clinical trial registries

- ClinicalTrials.gov
- EudraCT
- International Clinical Trials Registry Platform (ICTRP)
- Scan Medicine

Funding portals

- National Institute for Health Research (NIHR)
- UK Research and Innovation (UKRI)

Search limits

The following limits were applied to the search:

- English language only
- Clinical trial start date or project funding start date between January 2021 and October 2024

Trial and funding searches were limited to the most recent three years in order to capture technologies currently in active development rather than those where development was no longer being continued.

Record selection

Following completion of the systematic searches on the clinical trial registries and funding portals, all search results were exported and collated into one Excel document. Records were deduplicated and the remaining unique records were manually screened against the eligibility criteria by four reviewers. Each record was screened and tagged as either 'include' or 'exclude'. Ten percent of all records were double screened to ensure consistency amongst the reviewers, the remaining records were distributed across the review team and single-screened. During single screening, any record where the screener had a query was marked as such and discussed amongst the full screening team to reach a consensus.

Data Extraction

Manual data extraction for each of the included records was independently conducted by four reviewers. For 10% of records, data extraction was conducted by all of the review team to ensure accuracy and consistency and the remaining 90% of records were allocated to individual members of the review team. Data was extracted onto an Excel data extraction form. For named technologies where data fields could not be populated from information contained within the clinical trial record or funding portal record, the developer website was searched to identify additional information. Any queries relating to the data extraction were recorded and resolved through discussion amongst the full review team.

A summary of the data extraction fields is shown in Table 1 below.

Field	Description
Algorithm type	The type of algorithm the generative AI technology uses was recorded as one of the following: 'reinforcement learning', 'unsupervised', 'supervised', 'semi supervised learning' 'multiple algorithm types' or 'not specified'.
Device type	This was categorized as either 'digital', 'diagnostic' or 'device'.
Developer type	This was recorded as either 'industry' or 'non-industry' depending on the developers.
Developer location	The location in which the developer of the generative AI technology is based.
Generative AI category	The type of content each technology generated was categorized by either 'audio', 'text', 'data generation and augmentation tools', 'code', '3D model', 'image', 'video' or 'other'.
Generative AI method	The method used by the generative AI algorithm was recorded as one of the following: 'large language model', 'Generative Adversarial Networks (GANs)', 'reinforcement learning', 'transformer-based models', 'Variational Autoencoders (VAEs)' , 'other' or 'not reported'. No assumptions were made regarding GenAI method, and where no information was available from the clinical trial record or funding record this was recorded as NR.
NICE disease category	The indication for which each technology was being developed to treat was recorded. Indications were grouped into the most appropriate of the 19 NICE disease categories.
Source type	This was recorded as the source in which the information was found, e.g., ClinicalTrials.gov
Technology readiness level (TRL)	An assessment of the maturity level of a technology. TRL 1: basic principles observed and reported TRL 2: technology concept or application formulated TRL 3: analytical and experimental critical function or proof-of-concept TRL 4: basic validation in a laboratory environment TRL 5: basic validation in a relevant environment TRL 6: prototype demonstration in a relevant environment TRL 7: prototype demonstration in an operational environment TRL 8: actual technology completed and qualified through test and demonstration TRL 9: actual technology qualified through successful mission operations
N/A	Not applicable
N/R	Not reported

Table 1: Summary of the data extraction fields used in this report

Results

Prisma diagram

We identified a total of 161 novel technologies using GenAI; 96 technologies were identified from trial registries, and 64 were identified from funding sources.

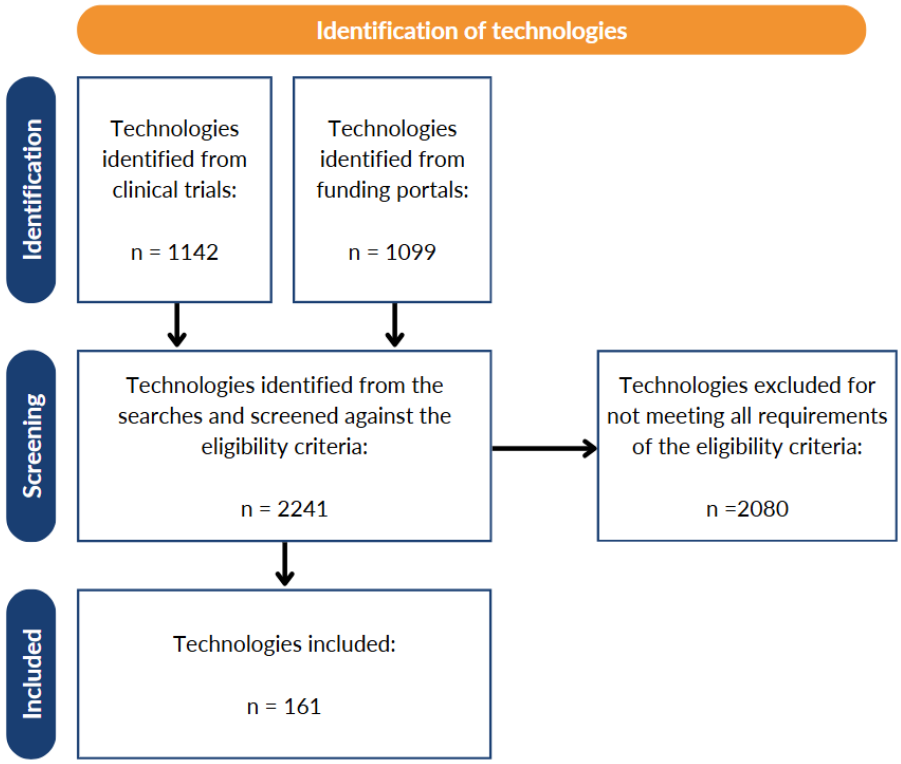


Figure 2: Prisma flowchart showing the number of technologies at each stage of screening

Type of Generative AI models

We gathered data on the type of models used within these technologies to better understand how they generate content as shown in figure 3. Around 51 percent (n=82) of these technologies did not specify the generative model used within either the clinical trial or funding award record, or the developer website. Of the records where the generative model was reported, the most frequently used was large language models (~17%, n=27). Twenty-one technologies used model types that did not fit into the broad categories used for classification; these were classified as "other", examples of these generative AI models include natural language processing and fusion methods. Variational autoencoders was the least used model, with only one technology found to be using this model type.

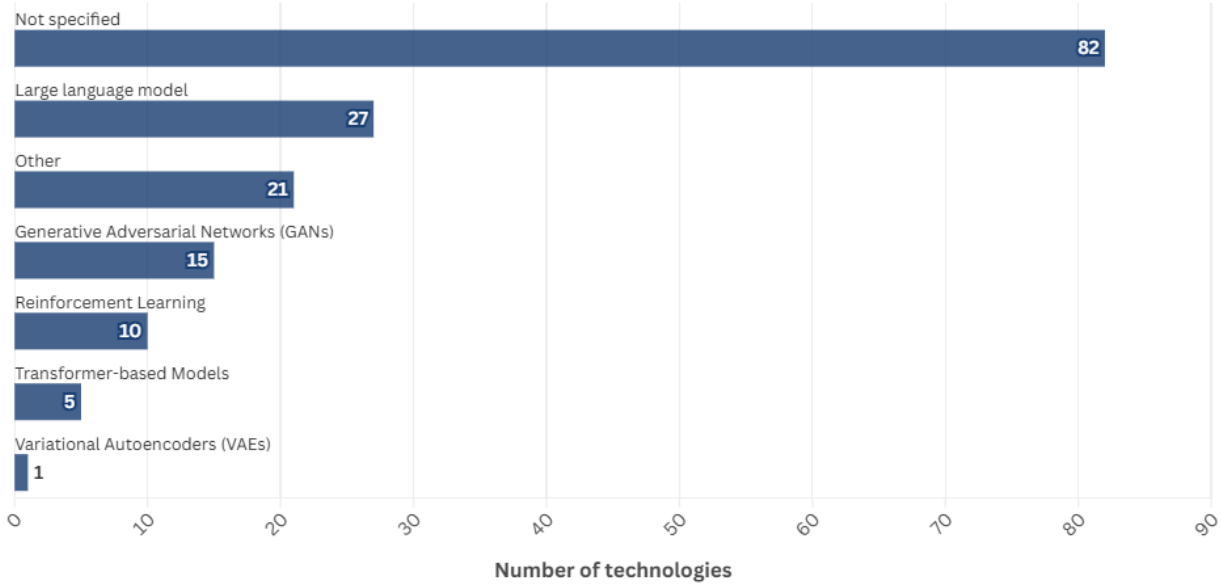


Figure 3: Type of generative AI models used in the identified technologies

Type of generative AI content

Figure 4 depicts the number of content types generated by the included technologies. The most frequently generated content was text (~43%, n=70), followed by data generation and augmentation (~17%, n=27) and image generation (~16%, n=25). Fewer technologies used AI to generate 3D models (n=19) and audio (n=5). Only one technology was found to generate code. Eleven technologies generated content which did not fit into the standardised categories, for example technologies which generated multiple forms of content, the content type generated by these technologies was therefore classified as 'other'.

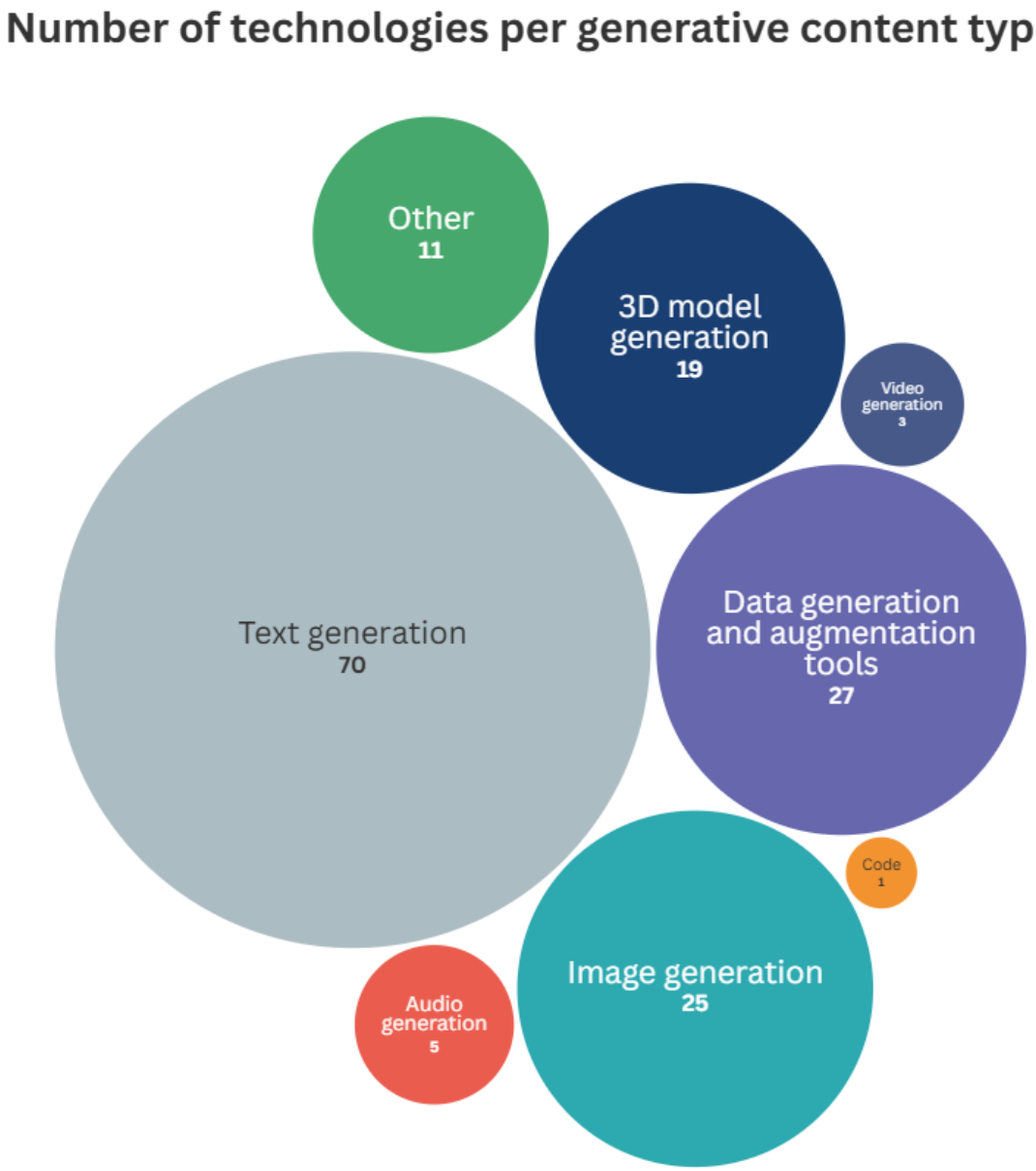
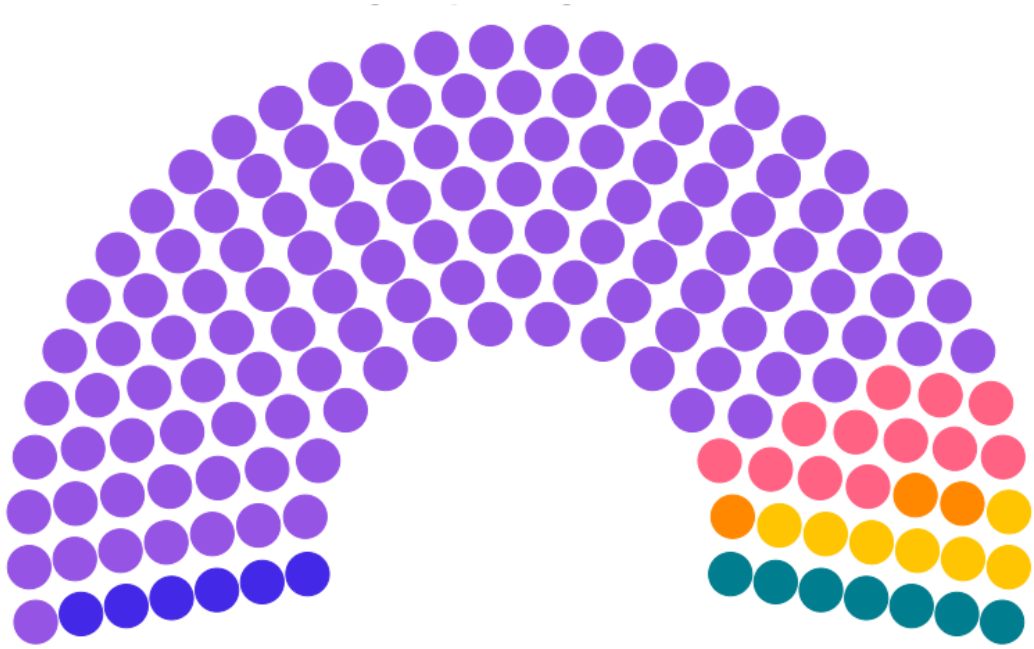


Figure 4: Type of content generated by the generative AI technologies

Type of algorithm used

Figure 5 illustrates the different types of algorithm learning paradigms used within the included technologies. The majority of technologies did not specify the algorithm type used within their description (n=126, ~78%). The most frequently used algorithm type in technologies where this information was reported, was reinforcement learning algorithms (n=12, ~7%). Supervised and unsupervised learning algorithms were used equally as often within our included technologies, representing ~4% (n=7) of algorithm types used respectively. Semi-supervised

learning algorithms was used less frequently in our included technologies, with only three technologies identified using these types of algorithms.



Algorithm category	Number of technologies
Multiple algorithm types	6
Not specified	126
Reinforcement learning	12
Semi supervised learning	3
Supervised	7
Unsupervised	7

Figure 5: Types of algorithms used by the identified generative AI technologies

NICE disease categories

Figure 6 illustrates the distribution of 161 identified technologies by NICE disease category.. Of the 161 technologies identified in this scan over a quarter (n=42, 26%) were not in development for a specific disease category, for example technologies that were used to generate 3D models from medical imaging scans, and were therefore reported as "Not Applicable". The most frequently reported category was mental health, behavioural and neurodevelopmental conditions (n=21, 13%), followed by cancer indications (n=20, 12%). Neurological conditions and disease and other endocrinal, nutritional and metabolic conditions were also frequently reported (n=10, 6% respectively). Several disease categories only had one relevant technology in development; these categories were injuries, accidents and wounds, urological, gynaecological, blood and immune system, liver, skin, kidney and sleep and sleep related conditions.

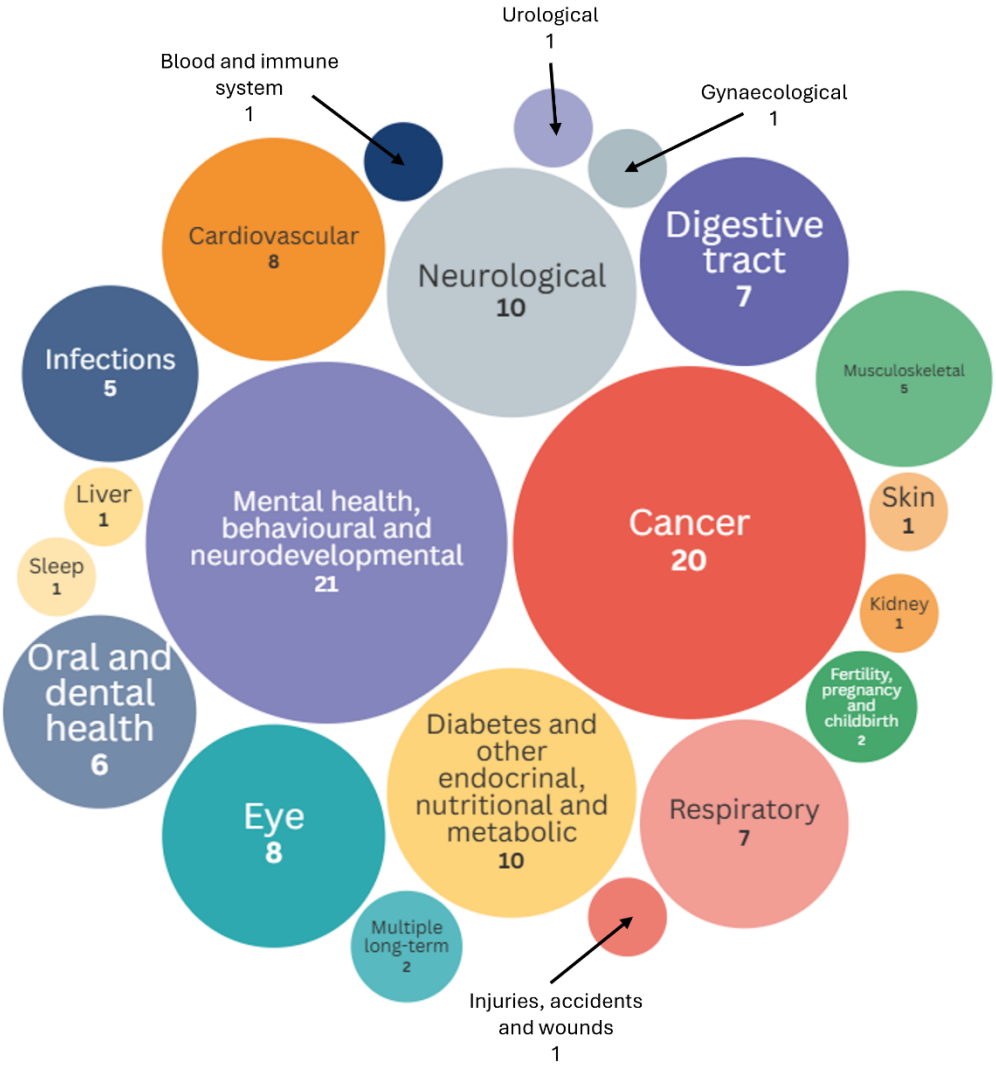


Figure 6: Number of technologies identified, classified by NICE disease category.

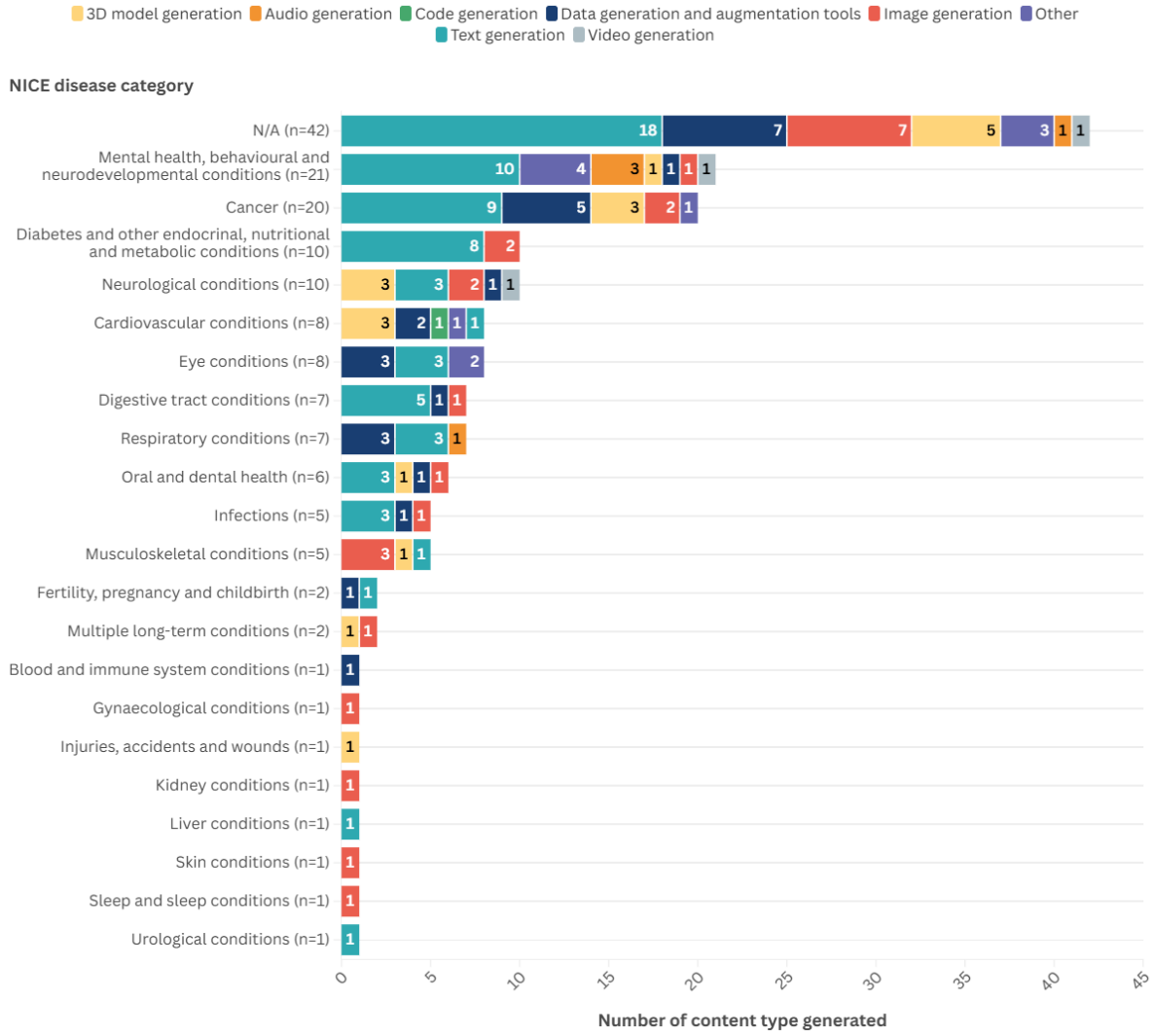


Figure 7: Shows the number of technologies categorised by type of content generated and the NICE disease category.

Figure 7 presents the distribution of content types generated across NICE disease categories. Text generation is the most frequently occurring category, with the highest number of technologies reported in the N/A category (n=18), followed mental health, behavioural, and neurodevelopmental conditions (n=10) and cancer (n=9). Data generation and augmentation tools are also widely represented, particularly technologies reported as N/A (n=7) and cancer (n=5). Image generation appears in multiple categories, including N/A (n=7) and musculoskeletal conditions (n=3). 3D model generation is primarily found in N/A (n=5), cancer (n=3), cardiovascular conditions (n=3) and neurological conditions (n=3). Audio generation is reported in mental health, behavioural, and neurodevelopmental conditions (n=3) and N/A (n=1). Video generation is present in N/A (n=1), mental health (n=1), and neurological conditions (n=1). Code generation was only reported in cardiovascular conditions. The "Other" category appears in mental health, behavioural, and neurodevelopmental conditions (n=4), N/A (n=3), and cancer (n=1).

Developer type

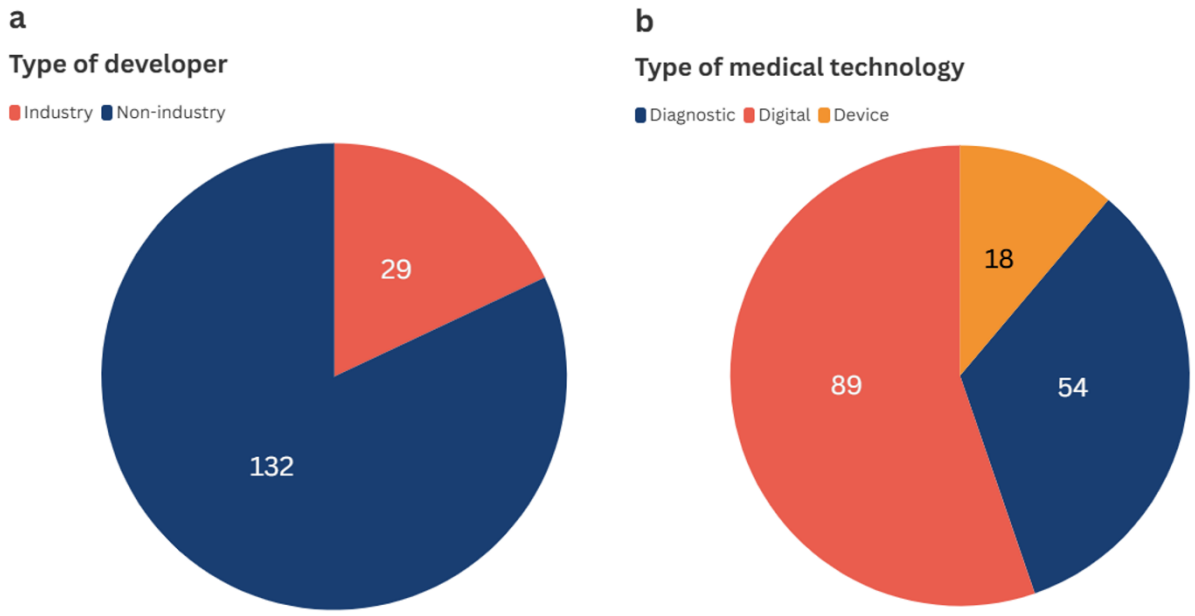


Figure 8: Figure 8a depicts a pie chart of the number of technologies identified by type of developer. Figure 8b shows the number of technologies classified by type of medical technology.

The type of developer was also reported for each technology as shown in Figure 8a. Of all technologies identified, 82% were in development by non-industry developers (n=132), such as universities and hospitals. The remain 18% of technologies were classed as in development by industry (n=29).

Technologies were also classified by the type of medical technology, depicted in Figure 8b. The most common type of medical technology making up 55% of technologies were reported as digital (n=89). This is followed by 34% of technologies classed as diagnostic (n=54). The least reported type of medical technology, accounting for 11% was devices (n=18).

Figure 9 shows the type of content generated categorised into three charts by type of medical technology. Text generation was the most frequent type of content generated in digital technologies, making up 60% of the category (n=53), compared to other types of content. Technologies generating text were used in a wide variety of applications including automatically generating clinical notes, providing recommended treatment plans or dose recommendations, and virtual assistants utilising ChatGPT or other LLMs to provide personalised feedback to patients. The next most common type of content generation was 3D model generation (7%, n=11) followed by data generation and augmentation (6%, n=10). No digital technologies used video or code generation. Technologies classified as devices most commonly used data generation and augmentation tools (n=7), making up 39%. “Other” was the second most reported content type in devices (n=4), accounting for 22%. Audio generation and code generation were not reported to be used in any devices. Image generation was the most frequent type of content reported at 33% in diagnostic technologies (n=18). This was

closely followed by 30% of technologies using text generation (n=16). There were no diagnostic technologies reported to use video generation or audio generation.

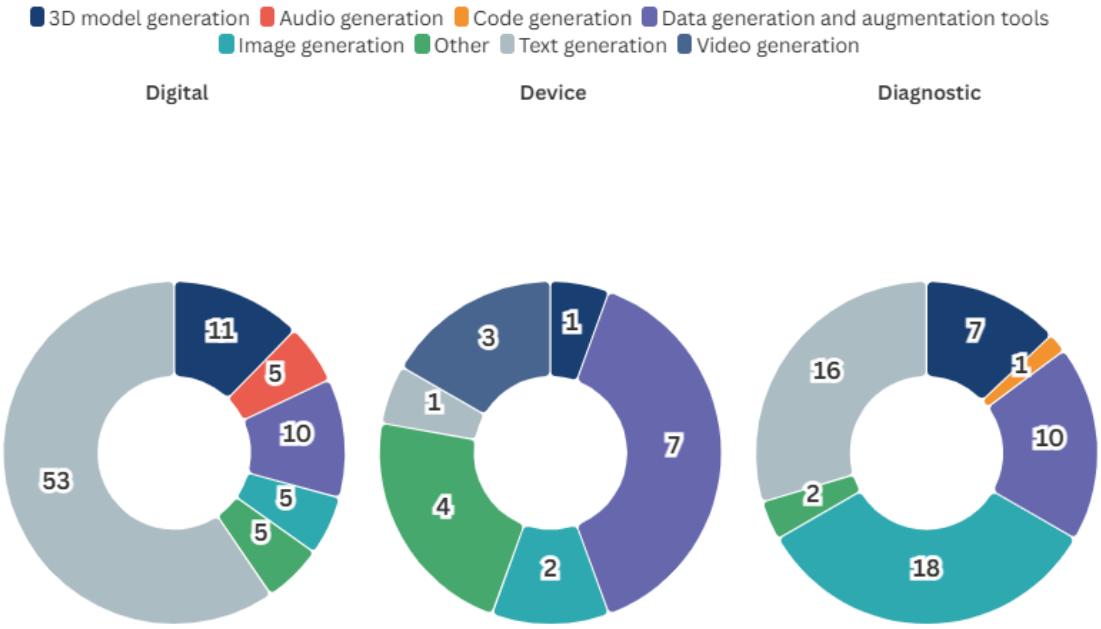


Figure 9: Represents the number of types of content generated, categorised by type of medical technology.

Developer location

Figure 10 shows the number of technologies identified from trial registries. The most common developer location was the United States, which accounted for the majority of the technologies (n=33). This is followed by China (n=15) and Turkey (n=7). In both Canada and UK, five technologies were identified. The names of the UK technologies identified were reported as Limbic Care, AI Clinician XP2, Wysa, spectroscopic probe and BeatClearWalker. Other countries in which technologies were identified ranged between one and three. There were three technologies with location reported as not specified.

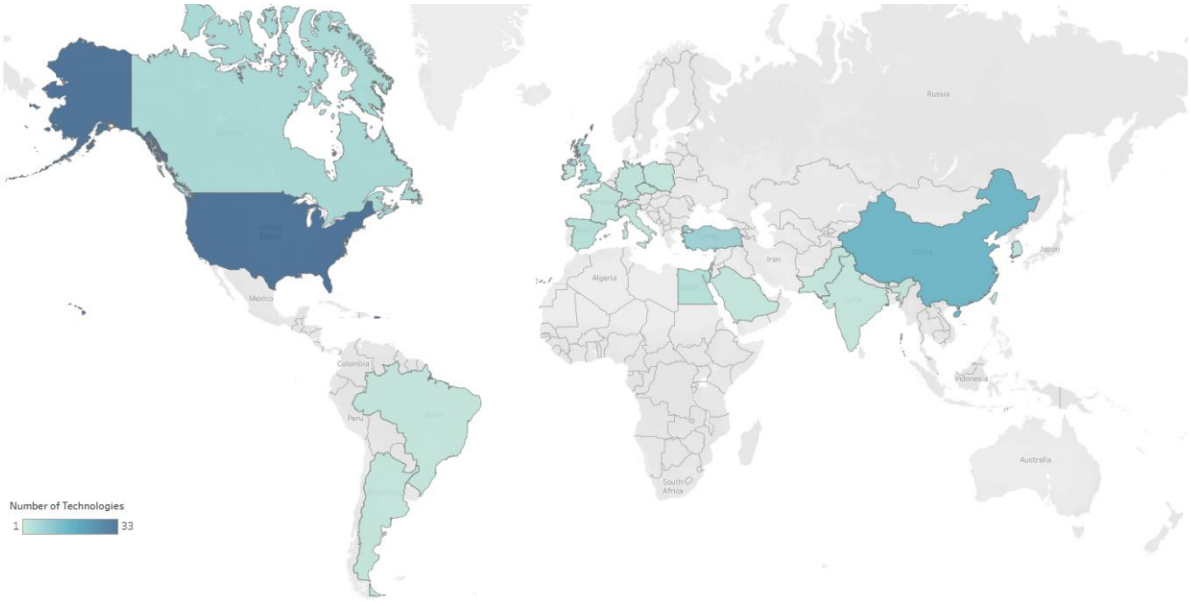


Figure 10: A map depicting developer locations of clinical trials.

Technology readiness level

A total of 75 technologies reported the use of generative AI methods, with large language models (LLMs) being the most common method (n=26), followed by generative adversarial networks (GANs) (n=15). Some algorithm types appeared only once and were found at very early TRLs (TRL 1-3), these included the use of deep convolutional networks, generative modelling, causal generative models, latent diffusion methods, machine/deep learning and neural networks for molecular modelling. Large language models were only found at TRLs 4 and above with 50% of these technologies regarded to be in TRL 6. Figure 11 is a technology radar showing the TRL of identified technologies by generative AI method.

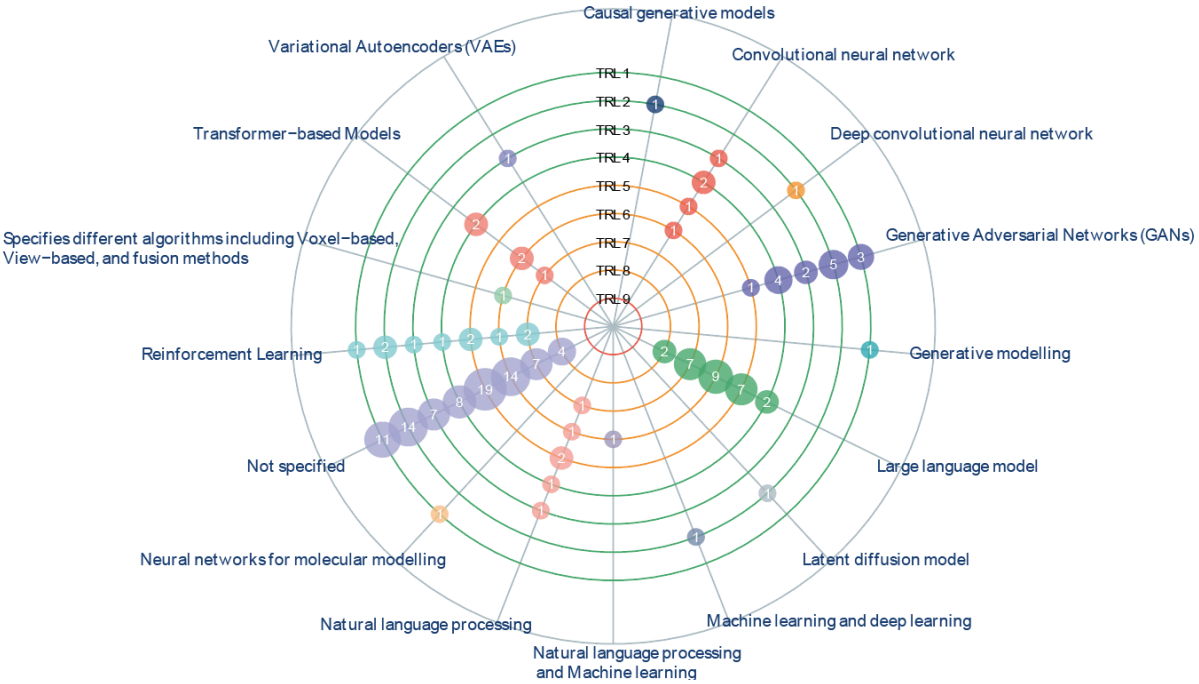


Figure 11: Technology radar to show the generative AI algorithm type by their current technology readiness level (TRLs). Concentric ring colour representative of horizon level; Red: Imminent, Orange: Transitional, Green: Emerging.

Named technologies (UK)

Fifty-four technologies were recorded to have UK locations but 33 of those did not report a clear technology name associated with the technology. Those with names can be seen in Figure 12.

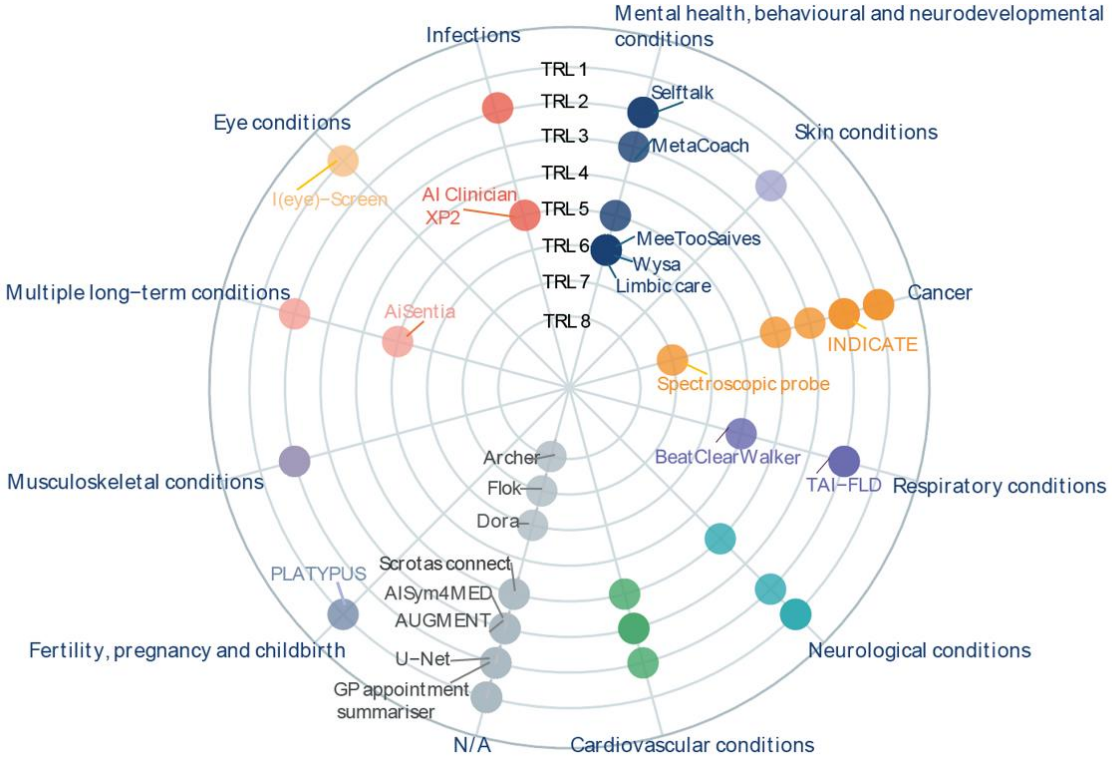


Figure 12: Technology radar representing all UK technologies (n=54) by their technology readiness level (TRL). Technologies without an official name recorded are not labelled and only represented by the plotted points.

Generative AI content type

The type of content generated by GenAI technologies by their TRL is illustrated in figure 13. The most common media type generated by the technologies found was text (n=70) followed by 'data generation and augmentation tools' (n=28) and image generation (n=25). There were eleven technologies that generate data and augmentation tools in the TRL 1, five in 3D model generation, one in the 'other' group and none across any of the other types of media.

Technologies identified in our study closest to market readiness (TRL7-9) included Archer, Flok and spectroscopic probe. For example, Flok is a mobile device which provides individualised physiotherapy, and is being trialled in UK clinical practise to reduce waiting times. The technology generates video and text to provide personal, real-time feedback to patients on managing symptoms of chronic back pain. Whereas another technology currently being used in the UK is Archer, an AI research assistant. The technology also generates text; however, it processes unstructured data from electronic health records for research purposes. Finally, the spectroscopic probe developed by Imperial College London generates images to assist in cancer diagnostics, by creating images with richer colour information to differentiate between normal and cancerous tissues more easily. These examples of technologies either in

deployment or close to market readiness demonstrate the variety of generative technologies becoming available.

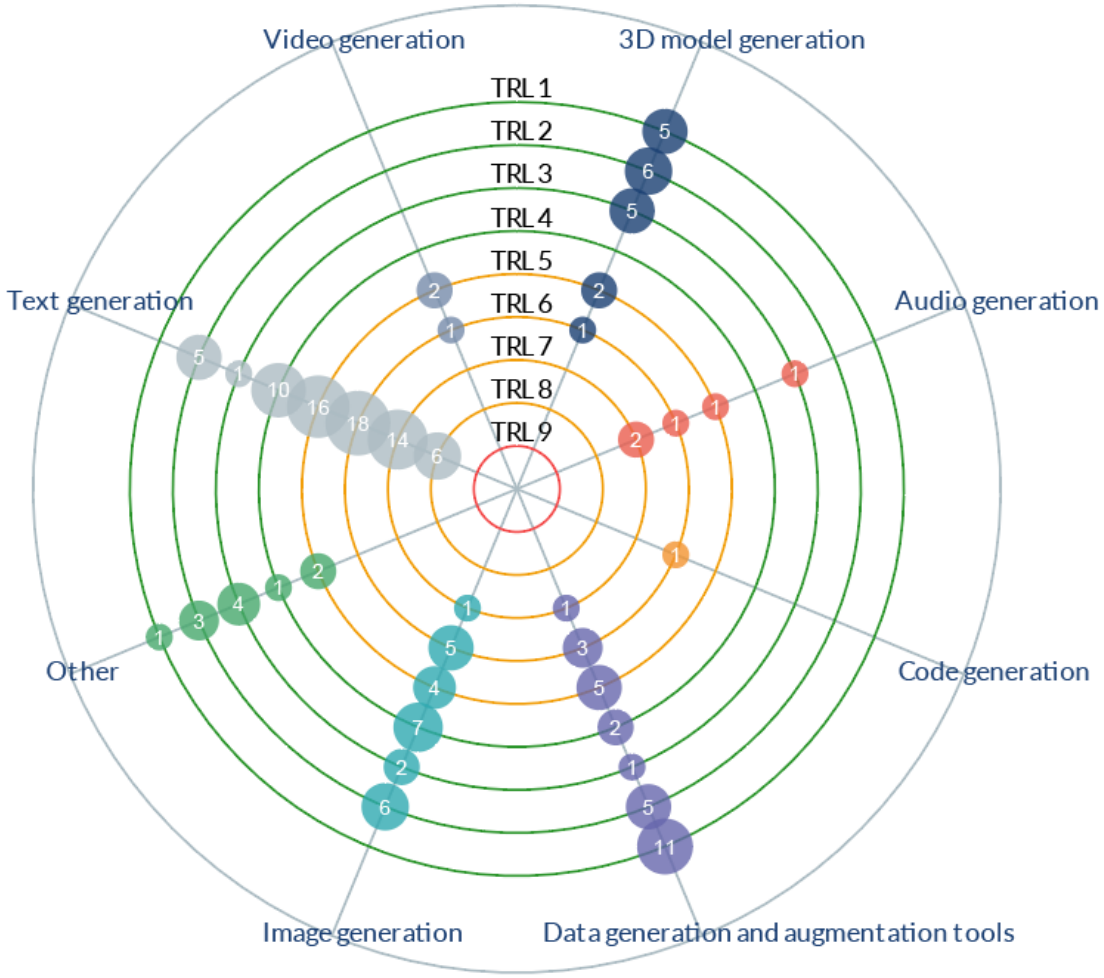


Figure 13: Technology radar to represent technologies by the type of content produced. Concentric ring colour representative of horizon level; Red: Imminent, Orange: Transitional, Green: Emerging.

Implications and Recommendations

Results from our scan revealed a diverse innovation landscape with 161 digital, diagnostic or device technologies using GenAI identified. These technologies leveraged different types of models and algorithms to generate various types of content across a wide range of indications and use cases, highlighting the diverse nature of emerging GenAI innovations in healthcare. Text generation was the most frequent type of content generated by the AI technologies identified in this scan. These technologies were being developed for a range of different applications including: generating clinical notes from unstructured data such as clinical conversations; chatbots and virtual health assistants to generate personalised responses to patient queries and provide specific feedback, education or reminders to the patient; and analysing patient data and automating creation of medical reports from this data. An unexpected finding was that mental health and neurodevelopment conditions was the most common indication for which GenAI technologies were in development for. The majority of these technologies were chatbots or virtual assistants generating personalised text responses. This could have significant implications on the health service in the UK given the extensive waiting times for mental health referral that patients are currently experiencing. GenAI chatbots can provide timely and valuable support for mental health patients, and offers benefits particularly in terms of accessibility, affordability and immediate emotional support. However, these technologies also pose risks such as a lack of human connection which may leave patients feeling more isolated, the technology providing inappropriate responses to patients which could worsen the mental health of vulnerable people, For example, a chatbot might fail to recognize the severity of a patient's condition and provide advice or recommend techniques that are not suitable for someone experiencing a severe mental health crisis.

We mapped the TRL of identified technologies to identify which types of technologies were closer to clinical implementation, enabling stakeholders to identify technologies that are most likely to be immediately disruptive and help to inform decision-makers in the implementation of these technologies into the healthcare system.[14,15] In addition to being the most frequently reported type of content, technologies generating text also had the highest proportion of technologies at later stages of TRL (TRL 7-9), indicating that text-generating technologies are likely to be the first widely implemented into clinical practice. This trend highlights the significant potential for improving efficiency and accuracy in tasks such as health-related documentation and providing personalised patient interaction. In contrast, technologies generating 3D models such as those producing patient-specific anatomical models and surgical simulations from medical image scans, were often at the less mature stages of TRL (TRL 1-3) indicating that these technologies are typically still in the exploratory phase, and it may be some time before they are implemented. Other categories for type of content generated including image generation, and data generation and augmentation tool had a more even distribution across the TRLs. This demonstrates and highlights the diversity and pace of development amongst generative AI technologies with some technologies nearing readiness for clinical implementation and others in experimental phases or proof of concept phases.

A significant challenge is the underreporting of generative AI model types, with 51% of technologies not disclosing this information. This poses challenges for regulators trying to assess these technologies in a timely manner given that different types of models can be associated with different levels of risks and types of mitigation requirements.[16] Reporting the type of model the developers are using at the pre-market development stage helps regulators to: better understand if the technologies are making transparent and accountable decisions; ensure compliance with guidelines from health technology assessment agencies and standardization organisations; establish if there are any ethical concerns such as bias, privacy and individual autonomy associated with the model; and more accurately assess risks associated to a specific type of model to ensure that appropriate measures are in place to mitigate any potential harms.[16,17,18] Among the technologies that reported the type of generative model used, large language models were the most frequently employed, accounting for approximately 17% (n=27) of the total.

The majority of technologies identified in this scan were being developed by non-industry developers, highlighting the critical role of academic and research institutions in driving early-stage innovation in generative AI medical technology. However, academic and research institutions may face difficulties navigating the regulatory landscape, securing funding and scaling-up development which may impact adoption of these technologies, highlighting the need for collaboration between industry and non-industry. [19,20]

Conclusion

This report provides an overview of the current state of GenAI technology development, with a comprehensive summary of global clinical trial activities and a focus on UK funding. Despite the innovative nature of GenAI technologies, there are clear indications of growing investment and development in this area of healthcare. We identified that there was extensive development of GenAI technologies for mental health and neurodevelopment conditions, and for cancer indications. Approximately half of the these technologies were in development for mental health and neurodevelopment conditions, primarily chatbots. The increasing utilisation of GenAI chatbots to provide mental health education and emotional support to patients has the potential to be disruptive to mental health care provisions. Positive impacts of technologies such as the Ryan Companionbot, Wysa and MetaCoach increase accessibility to care by providing 24/7 mental health resources to individuals who may otherwise have difficulty accessing therapy. However, there are concerns regarding the potential risks associated with these technologies. Issues such as hallucinations and misinformation advice that could lead to patients receiving sub-optimal care as a result of GenAI technologies. Additionally, the over-reliance on GenAI tech may result in reduced human interaction between patients and healthcare professionals which can be detrimental to improving mental health. The most common type of content generated by GenAI technologies was text, with significant development in digital assistants. There was also development of other types of devices and diagnostic technologies, however, they were often at much earlier stages of TRL.

Appendix A: Clinical trial and funding search strategy

EudraCT

#	Query/String
	"Generative AI" OR "deepfakes" OR "transformers" OR "chatgpt" OR "gemini" OR "Dall-E" OR "large language models" OR "LLMs" OR "natural language processing" OR "Generative AI models"

ICTRP

#	Query/String
	"Generative AI" OR "deepfakes" OR "transformers" OR "chatgpt" OR "gemini" OR "Dall-E" OR "large language models" OR "LLMs" OR "natural language processing" OR "Generative AI models" OR "GANs and variational autoencoders " OR "Bidirectional Encoder Representations from Transformers" OR "BERT" OR "AlphaFold" OR "GPT" OR "Codex" OR "GitHub Copilot" OR "Tabnine." OR "Descript" OR "Listnr" OR "Synopsys" OR "Cadence" OR "Nvidia" OR "Dialog Applications (LaMDA) transformers" OR "GANs and VAEs." OR "convolutional neural networks" OR "reinforcement learning." OR "Generative models" OR "Transformer models" OR "Creative AI" OR "Synthetic AI" OR "Generative Adversarial Networks" OR "GANs" OR "Variational Autoencoders" OR "VAEs" OR "Autoregressive models" OR "Recurrent Neural Networks" OR "RNNs" OR "Transformer-based models" OR "Text generation" OR "Image generation" OR "Music generation" OR "Video generation" OR "Code generation" OR "Audio generation" OR "Molecular generation" OR "Motion planning" OR "GPT-4" OR "DALL-E" OR "Jukedeck" OR "Synthesia" OR "GitHub Copilot" OR "WaveNet" OR "drug discovery" OR "Neural network-based models" OR "Feedforward Neural Networks" OR "Convolutional Neural Networks" OR "CNNs" OR "Recurrent Neural Networks" OR "RNNs" OR "Large Language Models" OR "LLMs" OR "Probabilistic models" OR "Variational Autoencoders" OR "VAEs" OR "Boltzmann machines" OR "Adversarial models" OR "Generative Adversarial Networks" OR "GANs" OR "Autoregressive models" OR "PixelCNN" OR "WaveNet" OR "Diffusion models"

CT.gov

#	Query/String
	"Generative AI" OR "deepfakes" OR "transformers" OR "chatgpt" OR "gemini" OR "Dall-E" OR "large language models" OR "LLMs" OR "natural language processing" OR "Generative AI models" OR "GANs and variational autoencoders " OR "Bidirectional Encoder Representations from Transformers" OR "BERT" OR "AlphaFold" OR "GPT" OR "Codex" OR "GitHub Copilot" OR "Tabnine." OR "Descript" OR "Listnr" OR "Synopsys" OR "Cadence" OR "Nvidia" OR "Dialog Applications (LaMDA) transformers" OR "GANs and VAEs." OR "convolutional neural networks" OR "reinforcement learning." OR "Generative models" OR "Transformer models" OR "Creative AI" OR "Synthetic AI" OR "Generative Adversarial Networks" OR "GANs" OR "Variational Autoencoders" OR "VAEs" OR "Autoregressive models" OR "Recurrent Neural Networks" OR "RNNs" OR "Transformer-based models" OR "Text generation" OR "Image generation" OR "Music generation" OR "Video generation" OR "Code generation" OR "Audio generation" OR "Molecular generation" OR "Motion planning" OR "GPT-4" OR "DALL-E" OR "Jukedeck" OR "Synthesia" OR "GitHub Copilot" OR "WaveNet" OR "drug discovery" OR "Neural network-based models" OR "Feedforward Neural Networks" OR "Convolutional Neural Networks" OR "CNNs" OR "Recurrent Neural Networks" OR "RNNs" OR "Large Language Models" OR "LLMs" OR "Probabilistic models" OR "Variational Autoencoders" OR "VAEs" OR "Boltzmann machines" OR "Adversarial models" OR "Generative Adversarial Networks" OR "GANs" OR "Autoregressive models" OR "PixelCNN" OR "WaveNet" OR "Diffusion models"

UKRI

#	Query/String
	"Generative AI" OR "deepfakes" OR "transformers" OR "chatgpt" OR "gemini" OR "Dall-E" OR "large language models" OR "LLMs" OR "natural language processing" OR "Generative AI models"

NIHR

#	Query/String
	"Generative AI" OR "deepfakes" OR "transformers" OR "chatgpt" OR "gemini" OR "Dall-E" OR "large language

	models" OR "LLMs" OR "natural language processing" OR "Generative AI models"

References

1 Kwong JCC, Nickel GC, Wang SCY, Kvedar JC. Integrating artificial intelligence into healthcare systems: more than just the algorithm. *npj Digital Medicine*. 2024;7(1):52. Available from: https://doi.org/10.1038/s41746-024-01066-z.

2 Alowais SA, Alghamdi SS, Alsuhebany N, Alqahtani T, Alshaya AI, Almohareb SN, et al. Revolutionizing healthcare: the role of artificial intelligence in clinical practice. *BMC Medical Education*. 2023;23(1):689. Available from: https://doi.org/10.1186/s12909-023-04698-z.

3 Mariani M, Dwivedi YK. Generative artificial intelligence in innovation management: A preview of future research developments. *Journal of Business Research*. 2024;175:114542. Available from: https://doi.org/https://doi.org/10.1016/j.jbusres.2024.114542.

4 Yu P, Xu H, Hu X, Deng C. Leveraging Generative AI and Large Language Models: A Comprehensive Roadmap for Healthcare Integration. *Healthcare (Basel)*. 2023;11(20). Available from: https://doi.org/10.3390/healthcare11202776.

5 National Institute of Standards and Technology (NIST). *Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile (NIST AI 600-1)*. Last Update Date: July 2024. Available from: https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.600-1.pdf [Accessed 13 February 2025].

6 Raza MM, Venkatesh KP, Kvedar JC. Generative AI and large language models in health care: pathways to implementation. *NPJ Digit Med*. 2024;7(1):62. Available from: https://doi.org/10.1038/s41746-023-00988-4.

7 Reddy S. Generative AI in healthcare: an implementation science informed translational path on application, integration and governance. *Implementation Science*. 2024;19(1):27. Available from: https://doi.org/10.1186/s13012-024-01357-9.

8 Sardesai N, Russo P, Martin J, Sardesai A. Utilizing generative conversational artificial intelligence to create simulated patient encounters: a pilot study for anaesthesia training. *Postgraduate Medical Journal*. 2024;100(1182):237-41. Available from: https://doi.org/10.1093/postmj/qgad137.

9 Mennella C, Maniscalco U, De Pietro G, Esposito M. Ethical and regulatory challenges of AI technologies in healthcare: A narrative review. *Heliyon*. 2024;10(4):e26297. Available from: https://doi.org/10.1016/j.heliyon.2024.e26297.

10 Reddy GP, Kumar YVP, Prakash KP. Hallucinations in Large Language Models (LLMs). In: 2024 IEEE Open Conference of Electrical, Electronic and Information Sciences (eStream); 2024 25-25 April 2024; 2024. pp. 1-6.

11 US Food and Drug Administration (FDA). *Total Product Lifecycle Considerations for Generative AI-Enabled Devices*. 2024. Available from: https://www.fda.gov/media/182871/download [Accessed 17 February 2024].

12 GOV.UK. *Futures, Foresight and Horizon Scanning*. 2023. Available from: https://foresightprojects.blog.gov.uk/2023/06/28/update-the-futures-toolkit/ [Accessed 17 January 2024].

13 Hines P, Hiu Yu L, Guy RH, Brand A, Papaluca-Amati M. Scanning the horizon: a systematic literature review of methodologies. *BMJ Open*. 2019;9(5):e026764. Available from: https://doi.org/10.1136/bmjopen-2018-026764.

14 van den Hoed MW, Backhaus R, de Vries E, Hamers JPH, Daniëls R. Factors

contributing to innovation readiness in health care organizations: a scoping review. *BMC Health Serv Res*. 2022;22(1):997. Available from: https://doi.org/10.1186/s12913-022-08185-x.

15 UK Research and Innovation (UKRI). *Technology readiness levels and the role of healthcare funders*. 2022. Available from: https://www.ukri.org/publications/technology-readiness-levels-and-the-role-of-healthcare-funders/ [Accessed 28 January 2025].

16 Warraich HJ, Tazbaz T, Califf RM. FDA Perspective on the Regulation of Artificial Intelligence in Health Care and Biomedicine. *JAMA*. 2025;333(3):241-7. Available from: https://doi.org/10.1001/jama.2024.21451.

17 World Health Organisation (WHO). *WHO outlines considerations for regulation of artificial intelligence for health*. 2023. Available from: https://www.who.int/news/item/19-10-2023-who-outlines-considerations-for-regulation-of-artificial-intelligence-for-health [Accessed 27 January 2025].

18 Shelmerdine SC, Arthurs OJ, Denniston A, Sebire NJ. Review of study reporting guidelines for clinical studies using artificial intelligence in healthcare. *BMJ Health Care Inform*. 2021;28(1). Available from: https://doi.org/10.1136/bmjhci-2021-100385.

19 Kruachottikul P, Tea-makorn P, Dumrongvute P, Hemrungrojn S, Nupairoj N, Junchaya O, Vinayavekhin S. MediGate: a MedTech product innovation development process from university research to successful commercialization within emerging markets. *Journal of Innovation and Entrepreneurship*. 2024;13(1):71. Available from: https://doi.org/10.1186/s13731-024-00439-8.

20 Kings College London. *Case study finds regulatory challenges are affecting MedTech innovation in the UK*. 2023. Available from: https://www.kcl.ac.uk/news/case-study-finds-regulatory-challenges-are-affecting-medtech-innovation-in-the-uk [Accessed 27 January 2025].

NIHR | Innovation Observatory

Newcastle University

*Delivering a **world leading** Health Innovation Observatory that **efficiently captures current trends** in health innovation and can guide on areas for **future promise***

NIHR Innovation Observatory
at Newcastle University

The Catalyst
Room 3.12, 3 Science Square,
Newcastle Helix,
Newcastle upon Tyne NE4 5TG

Copyright © National Institute for Health and Care Research Innovation Observatory (NIHR IO), Newcastle University