

# From AI 1.0 to AI 3.0

## ***A very brief chronicle of AI in medicine and healthcare***

Maged N. Kamel Boulos, MBBCh, PhD, SMIEEE

September 2024

---

### Agenda

- The three epochs of medical AI: key points to note
- AI 1.0 (rule-based AI), 2.0 (deep learning) and 3.0 (generative AI, LLMs and a method called RAG), *a very brief, non-technical introduction with some examples*
- Health digital twins, *in brief*
- AI explainability: is it absolutely necessary?
- Final thought
- Further reading (for those of you who want to dig deeper into the subject of generative AI in medicine and AI in your specific clinical disciplines)

26 more slides to go 😊



## Let us begin by emphasising a few key points:

- AI 1.0, 2.0 and 3.0 are not always meant to replace or supersede one another. They can co-exist and can be combined in the same application.
- More **data and computer power** are needed as we move from AI 1.0 to AI 3.0. This helps explain their chronological order of appearance.
- AI 1.0 and 2.0 applications have narrow scope and focus (task specific). AI 3.0 applications have broader scope (artificial general intelligence).
- They count as **SaMD/AIaMD** (software as a medical device/AI as a medical device) and fall under established MDR (medical device regulation) provisions **when used in patient management**, e.g., to reach a diagnosis or formulate a treatment plan.
- **Regulation** becomes increasingly complex as we move from AI 1.0 to AI 3.0. As of September 2024, the US FDA has approved 950 SaMDs that fall under AI 1.0/AI 2.0 but **none yet** belonging to AI 3.0 (situation is similar in UK, EU & China regarding AI 3.0 SaMDs).

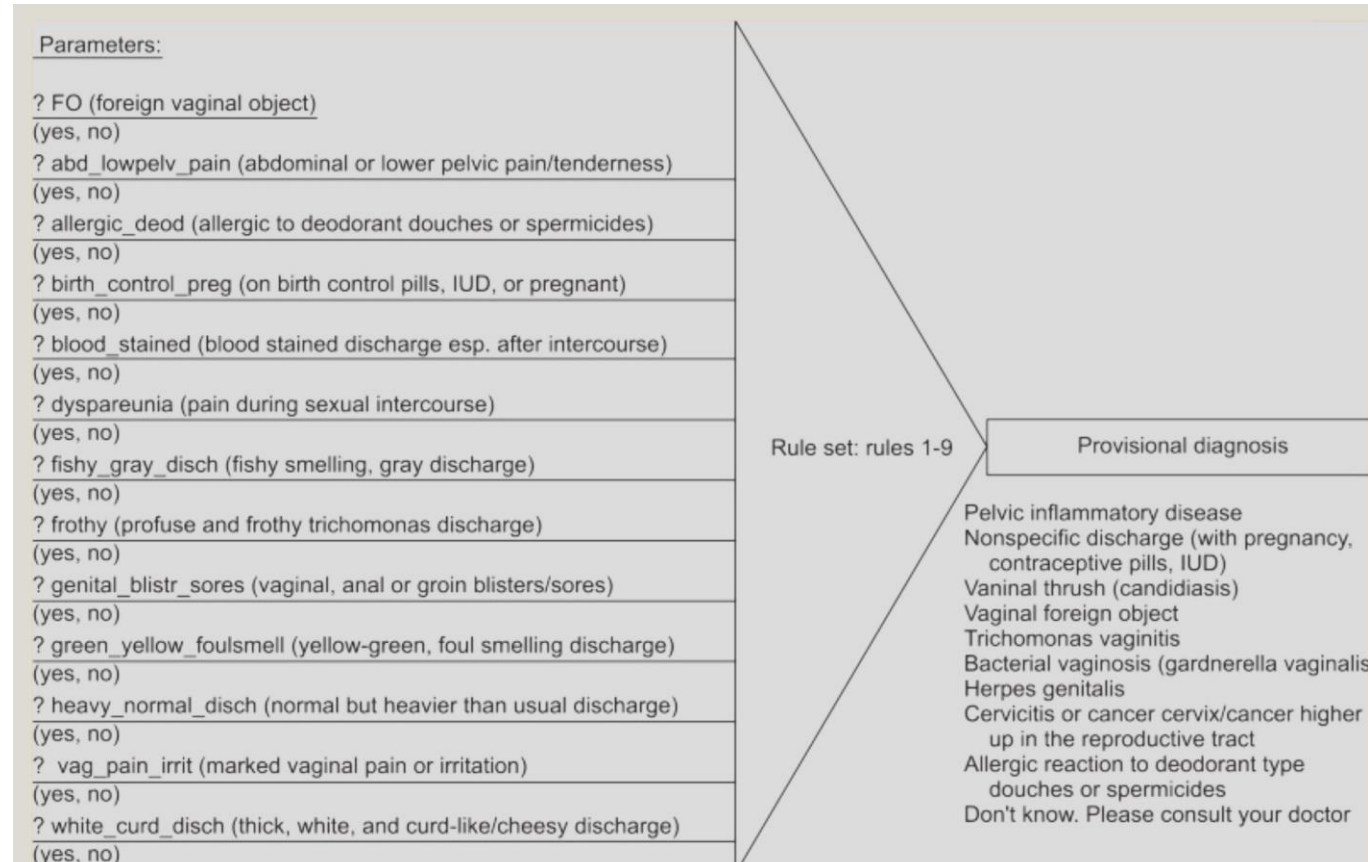
### Three Epochs of Artificial Intelligence in HealthCare

Michael D Howell et al. *JAMA*. 2024;331(3):242-244

<https://pubmed.ncbi.nlm.nih.gov/38227029/>

# AI 1.0: The rule-based beginnings

- Includes symbolic AI, which attempts to encode human knowledge into computational rules, as well as probabilistic models.



[Figure 1](#)

Mockler chart for diagnosis of vaginal discharge. IUD: intrauterine device.

## Additional examples:

Checkout our presentation from 2002 covering some historical AI 1.0 systems at <https://tinyurl.com/historical-AI>

**Expert System Shells for Rapid Clinical Decision Support Module Development: An ESTADemonstration of a Simple Rule-Based System for the Diagnosis of Vaginal Discharge**

Maged N Kamel Boulos. *Healthc Inform Res.*

2012;18(4):252–258

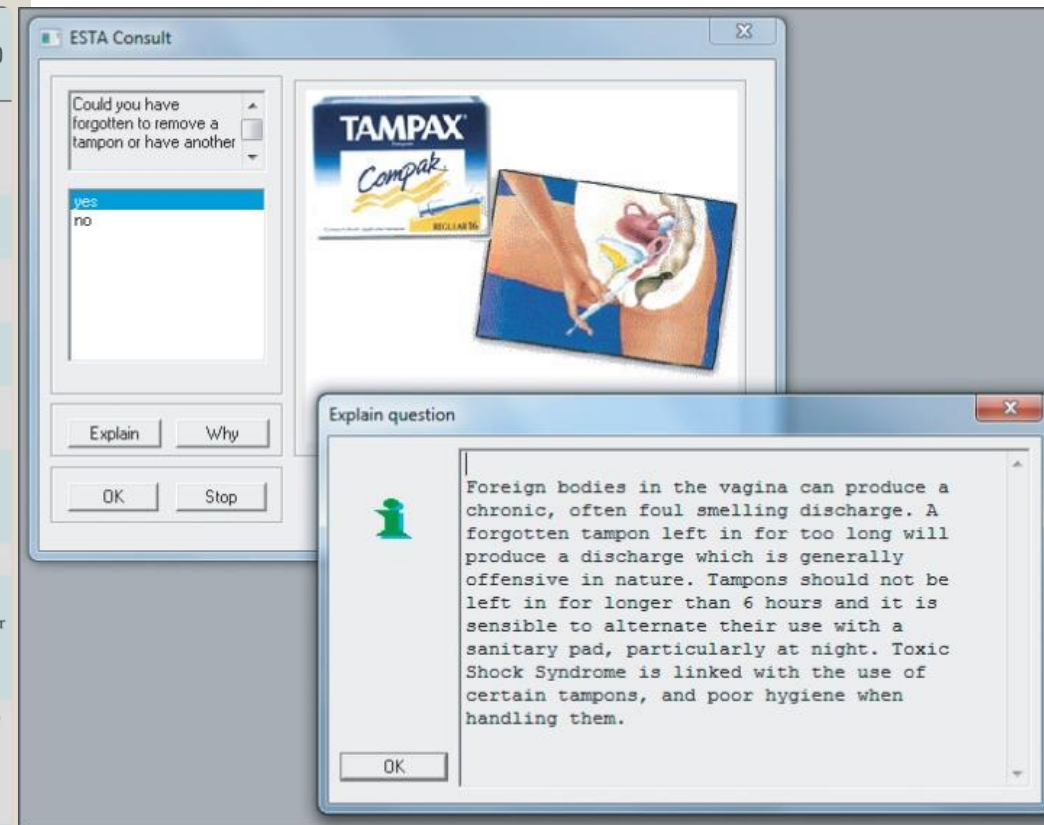
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3548154/>

# Table 1

Decision table for the provisional diagnosis of vaginal discharge (the system's goal)

Rule	FO	abd_lowpelv_pain	allergic_deod	birth_control_preg	blood_stained	dyspareunia	fishy_gray_disch	frothy	genital_blistr_sores	green_yellow_foulsmell	heavy_normal_disch	vag_pain_irrit	white_curd_disch	Diagnosis (advice)
R1	-	Yes	-	-	-	Yes	-	-	-	-	-	-	-	Pelvic inflammatory disease
R2	-	-	-	Yes	-	-	-	-	-	-	Yes	No	-	Non-specific discharge
R3	-	-	-	-	-	Yes	-	-	-	-	-	-	Yes	Vaginal thrush (Candidiasis)
R4	Yes	-	-	-	-	-	-	-	-	Yes	-	-	-	Vaginal foreign object
R5	No	-	-	-	-	No	-	Yes	-	Yes	-	-	-	Trichomonas vaginitis
R6	-	-	-	-	-	-	Yes	-	-	-	-	-	-	Bacterial vaginosis (Gardnerella vaginalis)
R7	-	-	-	-	-	-	-	-	Yes	-	-	Yes	-	Herpes genitalis
R8	-	-	-	-	Yes	-	-	-	-	-	-	-	-	Cervicitis or cancer cervix/cancer higher up in the reproductive tract
R9	-	-	Yes	-	-	-	-	-	-	-	-	-	-	Allergic reaction to deodorant type douches or spermicides
Else	If none	Of the	Above	Rules	Applies >>									Don't know. Please consult your doctor!

FO: foreign vaginal object, abd\_lowpelv\_pain: abdominal or lower pelvic pain/tenderness, allergic\_deod: allergic to deodorant douches or spermicides, birth\_control\_preg: on birth control pills, IUD, or pregnant, blood\_stained: blood stained discharge esp. after intercourse, dyspareunia: pain during sexual intercourse, fishy\_gray\_disch: fishy smelling, gray discharge, frothy: profuse and frothy trichomonas discharge, genital\_blistr\_sores: vaginal, anal or groin blisters/sores, green\_yellow\_foulsmell: yellow-green, foul smelling discharge, heavy\_normal\_disch: normal but heavier than usual discharge, vag\_pain\_irrit: marked vaginal pain or irritation, white\_curd\_disch: thick, white, and curd-like/cheesy discharge.

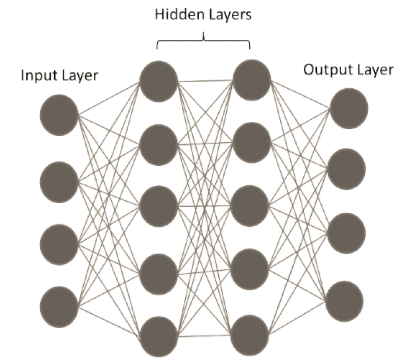


**Expert System Shells for Rapid Clinical Decision Support Module Development: An ESTA Demonstration of a Simple Rule-Based System for the Diagnosis of Vaginal Discharge**  
 Maged N Kamel Boulos. *Healthc Inform Res.* 2012;18(4):252–258  
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3548154/>



# AI 2.0: The deep learning revolution

- Deep learning is a subset of machine learning that uses artificial neural networks **with many layers**. In deep learning, models learn from examples labelled with ground truth.
- Deep learning models are **task-specific**, meaning they do one thing per model as determined by their training. They primarily focus on classification and prediction tasks.
- Deep learning is not a new term or an invention of the last decade (or a product of Google). Deep learning was introduced to the machine learning community in 1986 by computer scientist Rina Dechter and to artificial neural networks in 2000 by computer scientist Igor N Aizenberg et al. Aizenberg and colleagues were in turn building on the output of generations of scientists before them, including the first mathematical model of a neural network developed in **1943** by Walter H Pitts, Jr, a logician, and Warren S McCulloch, a neuroscientist.
- However, current form and applications only became possible with the **very large** amounts of digital data and computer power (esp. the GPUs) we have today, which were not available in those early days.



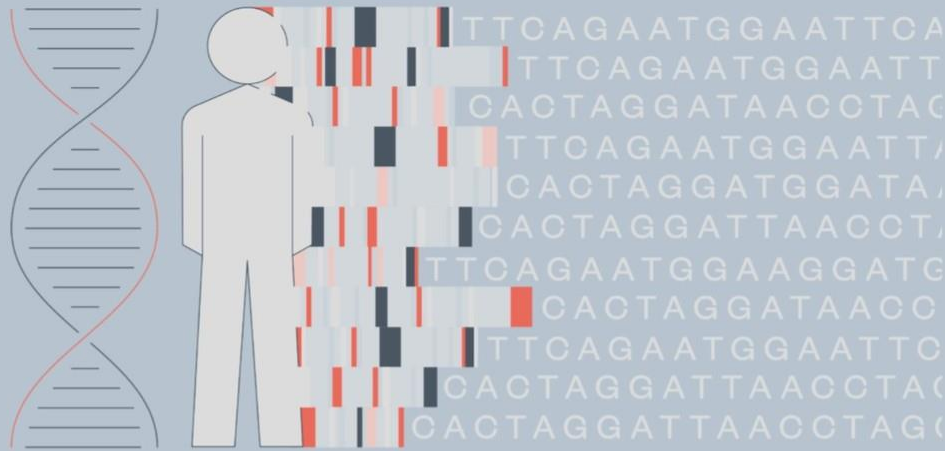
Deep learning is 'standing on the shoulders of giants' (*nanos gigantum humeris insidentes*)



**Disease prediction with multi-omics and biomarkers empowers case-control genetic discoveries in the UK Biobank**

Manik Garg et al. *Nature Genetics*. 2024;56:1821–1831

<https://www.nature.com/articles/s41588-024-01898-1>



- Note the very important **privacy and patient consent** issues mentioned at the end of the news clip.

Report

# Predicting: The future of health?

Assessing the potential, risks and appropriate role of AI-powered genomic health prediction in the UK health system

Harry Farmer

11 September 2024

Reading time: 231 minutes

(PDF 637 KB)



< <https://www.adalovelaceinstitute.org/report/predicting-the-future-of-health/>

- A report from the Ada Lovelace Institute and the Nuffield Council on Bioethics has warned against the mass rollout of AI-powered genomic health prediction (AIGHP) in the NHS
- It says that AIGHP could pose risks around privacy, discrimination and societal problems
- The report recommends policymakers focus on enabling targeted deployments of genomic health prediction within a stronger regulatory framework



# AI 3.0: The age of foundation models, large language models (LLMs) and generative AI

- Most models in AI 3.0 use the **transformer architecture**, a deep learning method introduced by researchers at Google in 2017. OpenAI's ChatGPT, launched in November 2022, is based on this architecture.
- These models have fundamentally new and potentially transformative capabilities, as well as new kinds of risks, such as **hallucinations** (incorrect or misleading outputs presented as facts) and **bias** (biased outputs reflecting and perpetuating human biases and unrepresentative populations in training datasets), among others.
- These models can do **many different kinds of tasks** ('AGI') without having to be retrained on a new dataset. For example, a simple text instruction can change a model's behaviour. Prompts such as *"Write this note for a specialist consultant"* and *"Write this note for the patient's mother"* will produce very different content.
- **Multimodal input and output capabilities:** Some models can take other modalities besides text as input, such as voice or images; for example, to analyse the content of an image or count the instances of some object of interest in it. A number of models can also be prompted to generate multimodal content, such as images, videos or a summary podcast with AI voice actors, in response to user's text input or uploaded documents.



# Doctors using ChatGPT to treat patients

A survey of British family doctors found one in five using AI software to suggest treatments or write letters

Michael Searles  
Health Correspondent

## Related Topics

Artificial Intelligence, ChatGPT, NHS (National Health Service), Doctors, Health news

18 September 2024 7:01am



GPs have been using ChatGPT to treat patients, a Harvard study has warned.

Researchers at the American university found one in five family doctors in the UK had used artificial intelligence tools while treating patients, despite a lack of regulation.

The survey of 1,006 GPs found dozens were using AI to help diagnose conditions and find treatment options.

A quarter of the 205 who admitted using machine-learning tools to help them do their jobs said they had asked the software to suggest treatments.

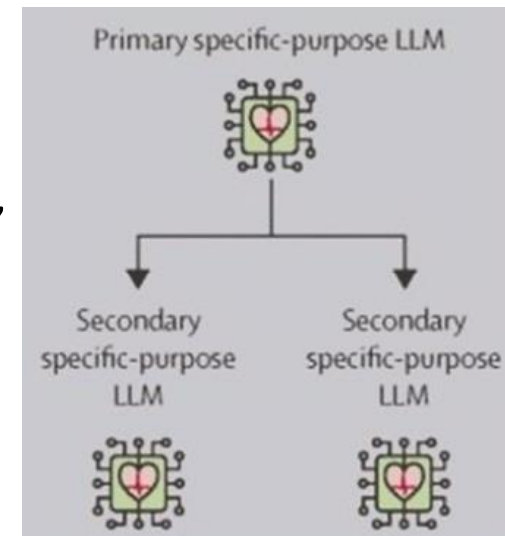
Almost three in 10 said they had used AI to help diagnose a patient. Others admitted they had used it to write letters, generate documents after an appointment with a patient, or create patient summaries and timelines based on past records.

Experts warned that unregulated use of tools such as ChatGPT, Microsoft's Bing AI or Google's Bard, could "risk harm and undermine patient privacy".

# AI 3.0: The age of foundation models, large language models (LLMs) and generative AI

Different models and model versions are optimised for different tasks and applications. It is therefore imperative to choose the right model(s) for the task at hand. There are two broad categories of models:

- **General-purpose models**, such as the different and successive versions of ChatGPT (OpenAI), Claude (Anthropic), Gemini (Google), Llama (Meta), and Mistral, among others; and
- **Specific-purpose/narrow-focus models**, including (in our case):
  - **Medically-trained models**, such as the different and successive versions of Med-PaLM (Medical Pre-trained Language Model, Google), Med-Gemini (Google), PH-LLM (Personal Health Large Language Model, Google; optimised for understanding and reasoning over time-series personal health data, e.g., from mobile and wearable devices), Clinical Camel, MEDITRON-70B, etc.; and
  - **Medical-speciality-specific models.**



Some applications combine more than one model together.

# Medical-speciality-specific models

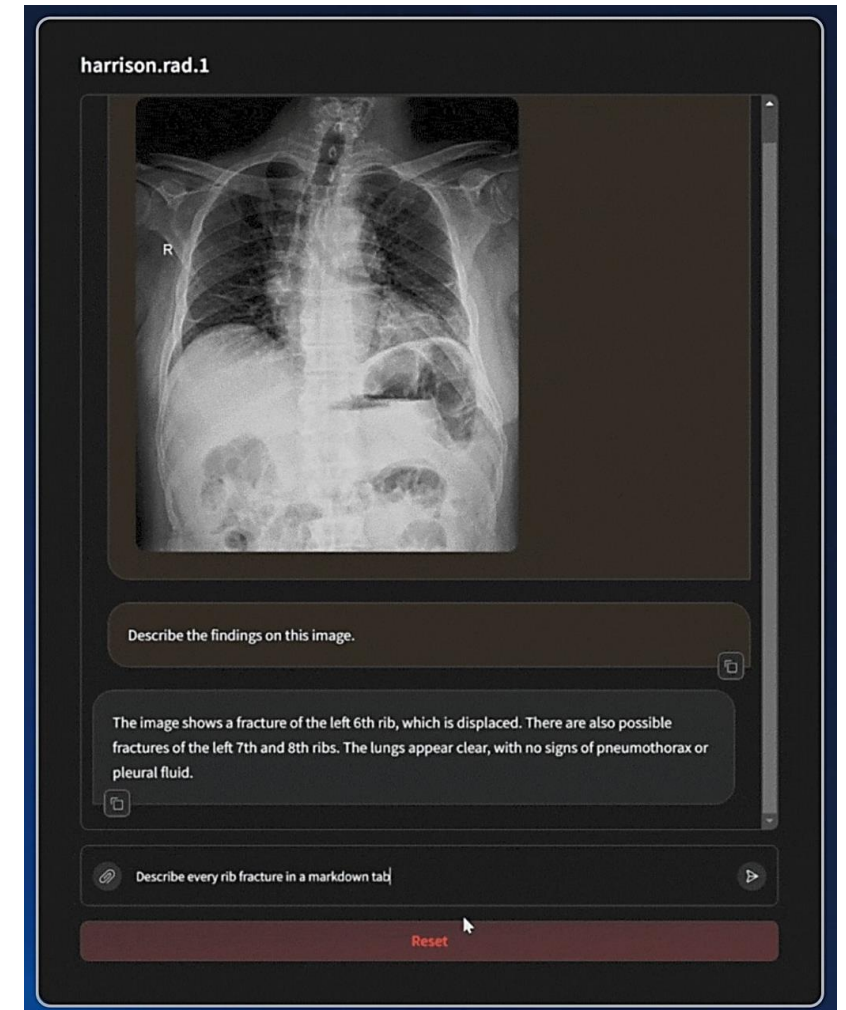
For example:

- **CHIEF (Clinical Histopathology Imaging Evaluation Foundation) model**, a pathology foundation model for cancer diagnosis and prognosis prediction.

**A pathology foundation model for cancer diagnosis and prognosis prediction**

Xiyue Wang et al. *Nature*. 2024 Sep 4. doi:10.1038/s41586-024-07894-z

- **Harrison.rad.1** (<https://harrison.ai/harrison-rad-1/>) a radiology-specific vision language model (VLLM) that recently scored 51.4 out of 60 (85.67%) on the Fellowship of the Royal College of Radiologists (FRCR) Rapid 2B exam, which is twice the score of other frontier models from OpenAI, Google and Anthropic.







# Reducing hallucinations

- A number of methods exist and are being developed to mitigate the issue of LLM hallucinations.
- **Retrieval-Augmented Generation (RAG)** is one of them, whereby **user-supplied material (e.g., research papers, textbooks, etc.)** serves as an **external knowledgebase** to an existing LLM, such as GPT-4 or Mistral 7B, in order to provide more relevant and better cross-linked and referenced outputs in response to user's prompts.
- Here is an example of a RAG-based application we recently developed:

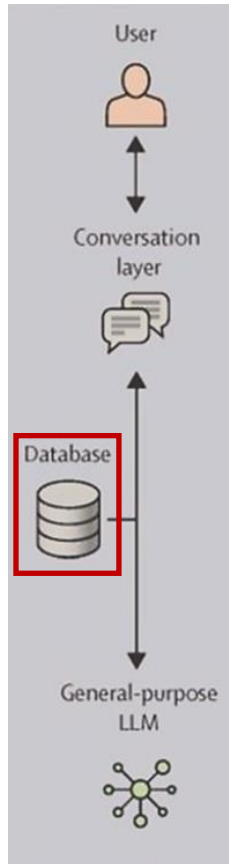
**NVIDIA's "Chat with RTX" Custom Large Language Model and Personalized AI Chatbot Augments the Value of Electronic Dermatology Reference Material**

Maged N Kamel Boulos and Robert Dellavalle. *JMIR Dermatol.* 2024 Jul;7:e58396. doi:10.2196/58396

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC11306952/>

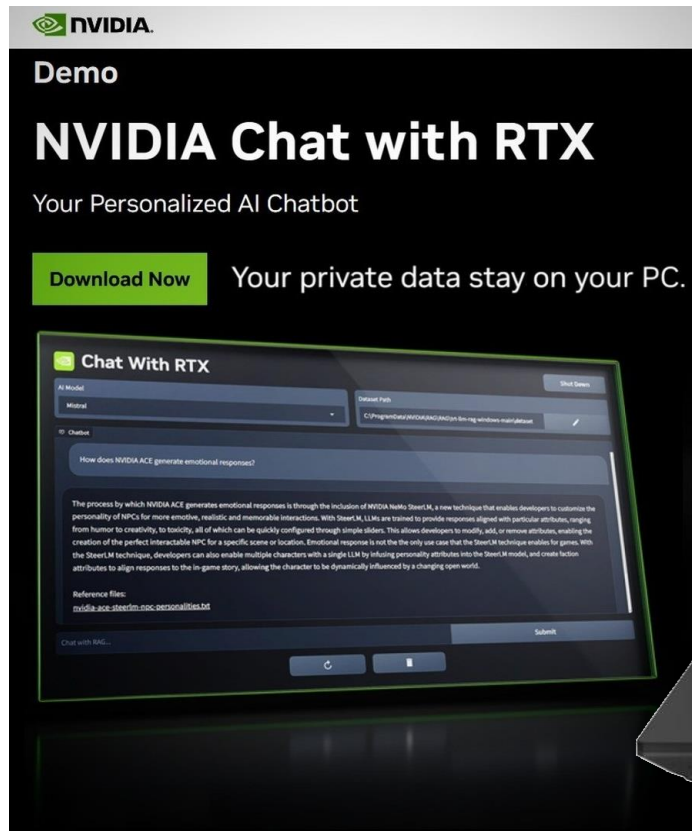
And it is not just for dermatology. The same approach can be applied to other medical specialities or to specific/niche aspects of those specialities: oncology, obstetrics and gynaecology, ovarian cancer; you name it.

We used NVIDIA's ChatRTX a free tool for developing RAG applications with one unique feature: with ChatRTX, small LLMs function entirely **offline** on **consumer-grade hardware**.




# Borrow your child's gaming laptop or desktop (if you can!) *That expensive 'toy' has got a new (clinical) use*

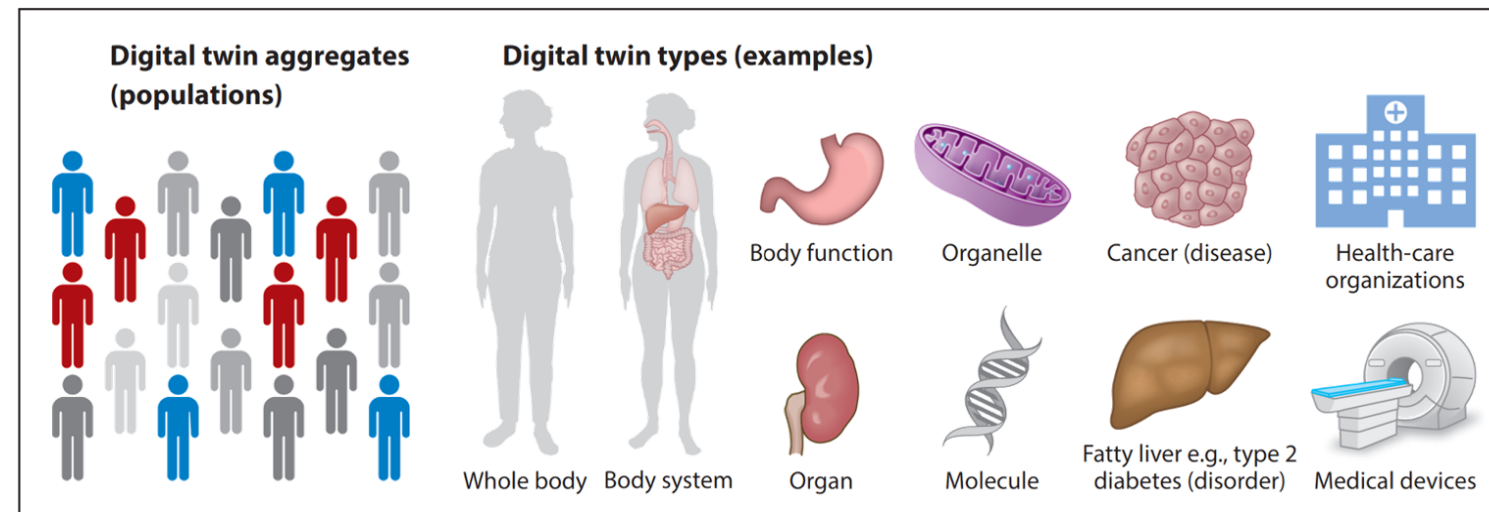
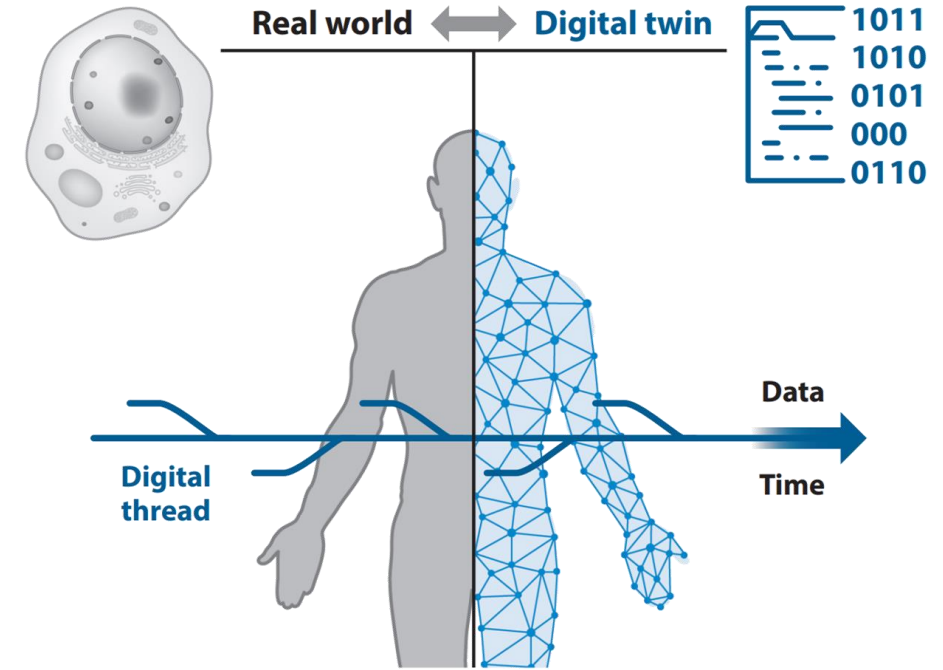
**Benefits:** Delivers additional self-study and research functionality beyond that of the standard electronic search and indexing that is already available in .pdf files, such as the abilities (1) to instantly identify, link, synthesise, and summarise at a single location related information scattered across different book chapters and/or multiple papers, (2) to answer user's questions asked in plain language (free text), and (3) to generate knowledge-testing quizzes with answers.

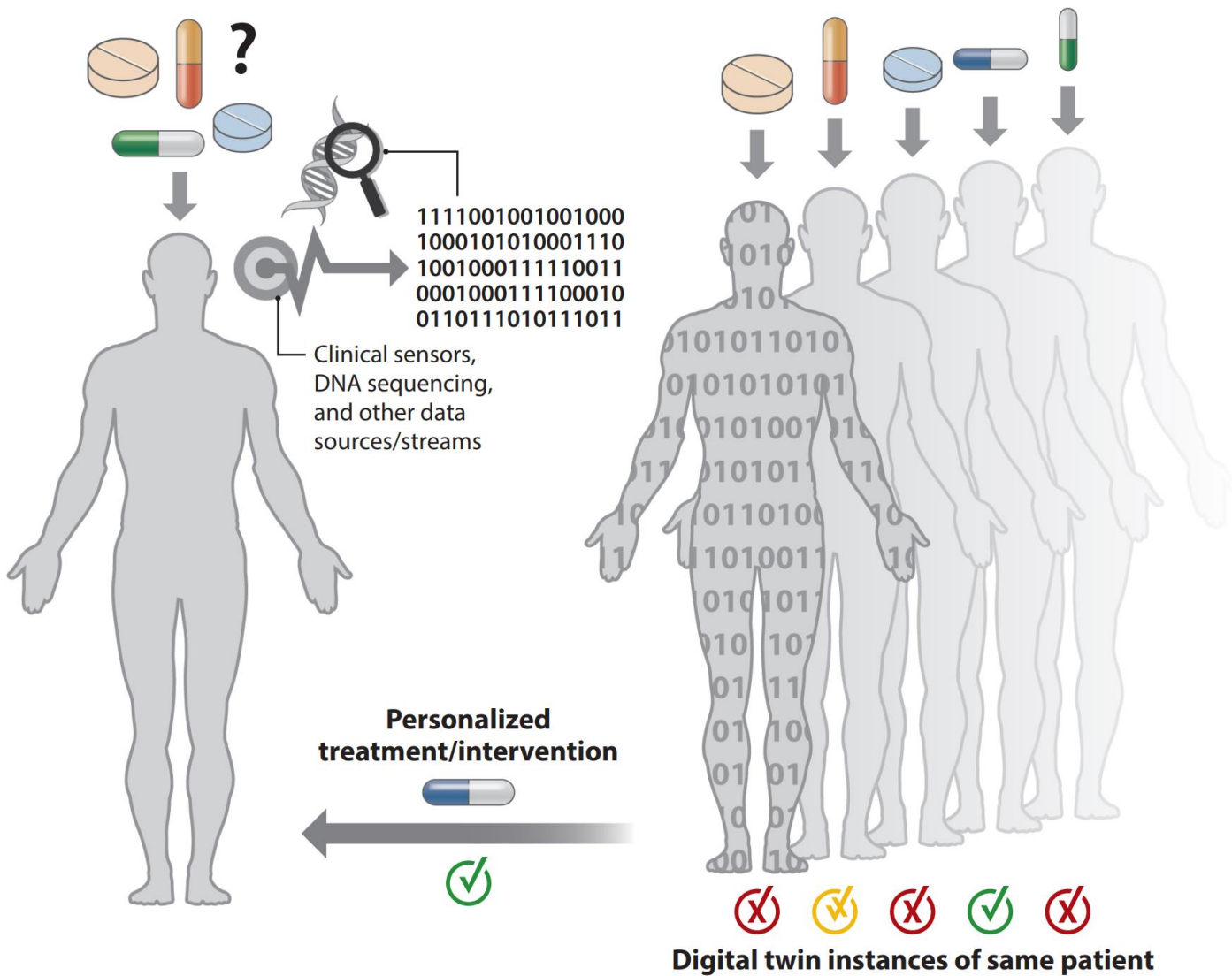




# Health digital twins (HDTs)

- Digital twins are virtual models of physical objects, i.e., the digital twin is a virtual representation of a physical twin.
- Digital twins were introduced in 2002 as a concept in the manufacturing industry for product life-cycle management. In 2010, the concept became known as 'digital twins' when  NASA started creating digital simulations of spacecrafts for testing.
- The technology has only recently been applied to healthcare under the umbrella field of **precision medicine**.
- In health/care, the digital twin becomes a dynamic virtual representation of the patient (or of specific aspects of the patient).





- Various AI/machine learning methods are used alone or combined in HDTs (AI 2.0, A 3.0/generative AI).

<< Here identical copies of a digital twin, all belonging to the same individual, are used for *in silico* testing and comparisons of various treatment(or intervention/procedure) options to answer **individual-specific** questions such as “*which treatment or intervention will be most successful in a given condition*”.

#### Health Digital Twins in Life Science and Health Care Innovation

Kaushik P Venkatesh, Gabriel Brito, Maged N Kamel Boulos.

*Annu Rev Pharmacol Toxicol.* 2024;64:159-170.

doi:10.1146/annurev-pharmtox-022123-022046

<https://www.annualreviews.org/content/journals/10.1146/annurev-pharmtox-022123-022046>

#### Digital Twins: From Personalised Medicine to Precision Public Health

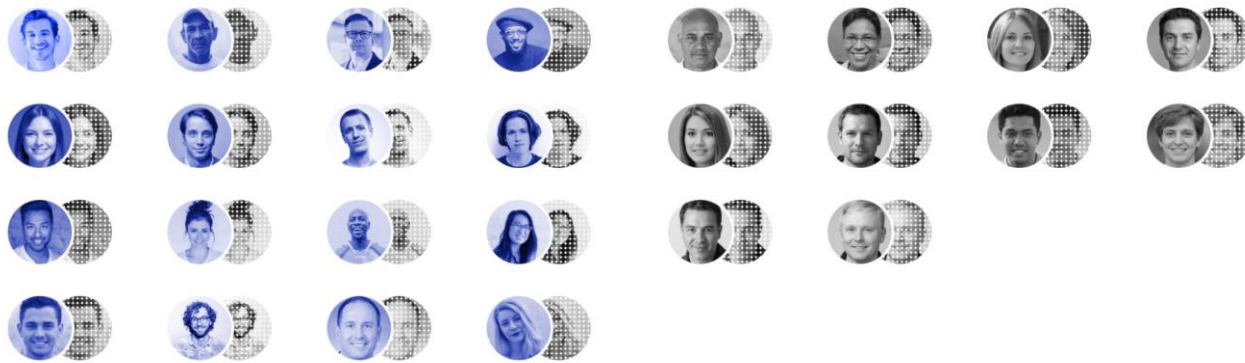
Maged N. Kamel Boulos and Peng Zhang.

*J Pers Med.* 2021;11(8):745. doi:10.3390/jpm11080745

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8401029/>

# HDTs for medical research/clinical trials

- Clinical trials are the last stage of drug development, take 12–15 years on average, and enrol thousands of patients.
- HDTs can be used to **partially virtualise the control arm of clinical trials**, decreasing the number of physical patients needed. The control arm's synthetic data are statistically indiscernible from real collected data.
- By decreasing the number of necessary real-life subjects to recruit, compensate, and administrate, HDTs can deliver **significant cost savings** and **accelerate trial timelines**.



Treatment group

<https://www.unlearn.ai/>

Control group - Significant reduction



Unlearn.AI, a San Francisco-based startup, uses **generative AI** to create HDTs of individual patients at different time points simulating natural disease progression. These HDTs are then used in the control arm of clinical trials.



# AI explainability: An AI workflow to explain AI models

- Explainability is easier with AI 1.0/rule-based systems; more difficult and sometimes even impossible with later systems ('black box').
- A recent study by Lang et al. introduced a novel workflow for **understanding the visual signals in medical images that AI classifiers use for predictions**. The workflow was applied across three different medical imaging modalities.
- The study identified attributes capturing clinically known features and unexpected confounders, as well as new patho-physiologically plausible attributes that could indicate novel research directions (**novel hypotheses for future clinical research**).
- The study concluded by presenting a generative AI-based framework that can generate visual explanations for AI predictions in medical imagery, with the aim of improving trust in AI models, providing a comprehensive understanding of the visual features influencing AI predictions, improving model assessments, aiding in designing better model training datasets, and facilitating novel scientific discoveries.

**Using generative AI to investigate medical imagery models and datasets**

Oran Lang et al. EBioMedicine. 2024;102

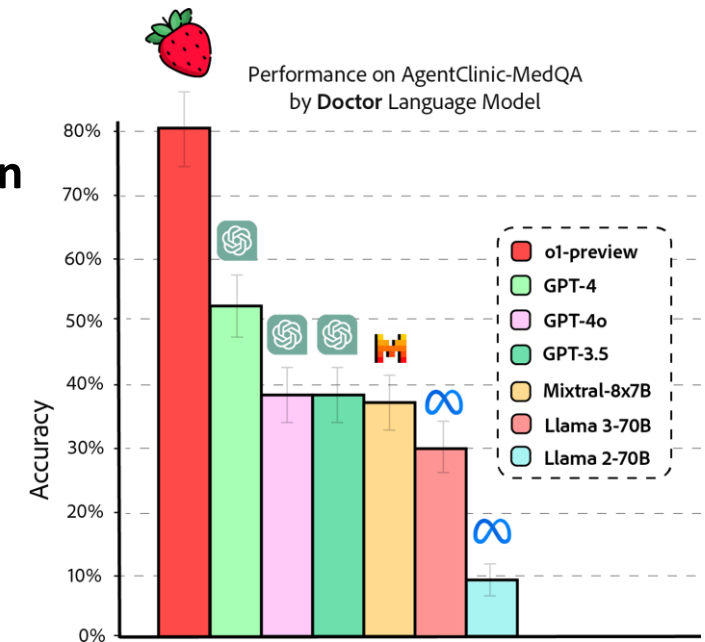
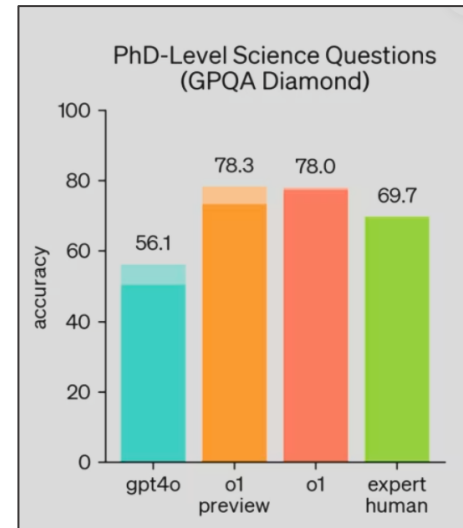
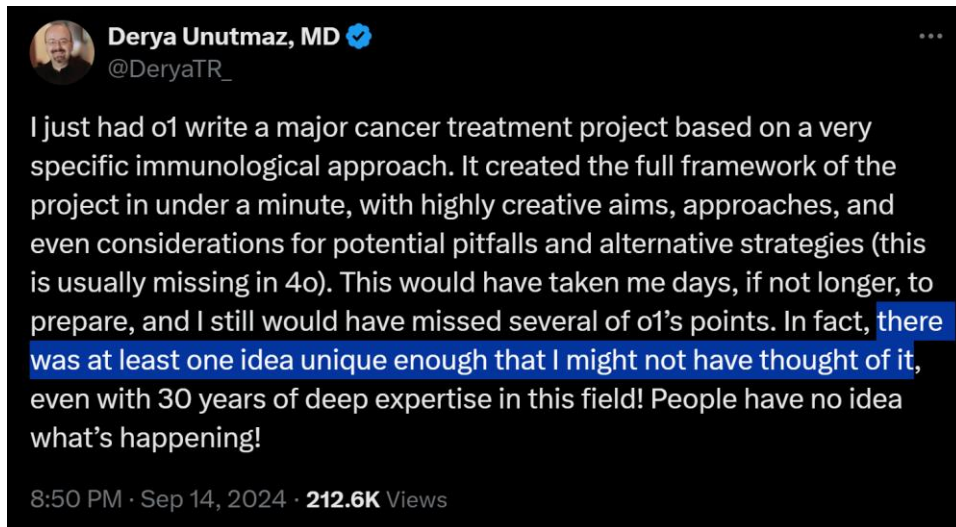
<https://doi.org/10.1016/j.ebiom.2024.105075>



Shedding light on the 'black box'  
(AI art created by DALL-E 3)

# AI explainability: OpenAI o1 'chain of thought'

- OpenAI has been working for some time on a radically new approach and LLM (codenamed 'Strawberry' and 'Orion' respectively) that can do **more complex reasoning** and **context-sensitive solving of harder problems** instead of just the mere pattern recognition and glorified word prediction of previous LLMs.
- On 12 September 2024, a preview of OpenAI o1 (Strawberry - <https://openai.com/o1/>) was launched, heralding the beginning of a new series of AI models designed to **spend more time thinking and producing a long internal chain of thought before responding to the user** (<https://openai.com/index/learning-to-reason-with-llms/>).
- **Although not specifically medically trained, OpenAI o1-preview has shown impressive performance when compared with other LLMs.**



<https://agentclinic.github.io/>

Make a diagnosis based on the following report:

**Phenotypes:**

- Intellectual disability
- Global developmental delay
- Macrodonia
- Triangular face
- Thick eyebrow
- Abnormality of the hand
- Short stature
- Motor delay
- Delayed speech and language
- Feeding difficulties
- Low anterior hairline
- Strabismus
- Anteverted nares
- Bulbous nose

**Excluded phenotypes:**

- Atrioventricular canal defect
- Patent foramen ovale
- Ventricular septal defect
- Pulmonic stenosis
- Bicuspid aortic valve
- Mixed hearing impairment
- Microcephaly
- Synophrys



Based on the reported phenotypes and the exclusion of certain conditions, the most likely diagnosis could be **Cornelia de Lange Syndrome (CdLS)**.

**Rationale:**

**Core Features:** Intellectual disability, global developmental delay, short stature, and distinctive facial features like thick eyebrows, a triangular face, a bulbous nose, and low anterior hairline are characteristic of CdLS.

**Additional Features:** The presence of macrodonia (enlarged teeth), abnormal hand features, motor and speech delays, and feeding difficulties are also commonly associated with this condition.

**Excluded Phenotypes:** The exclusion of heart defects like atrioventricular canal defects, foramen ovale, and septal defects, as well as hearing impairment and microcephaly, helps narrow down the differential diagnosis away from syndromes with these common features, such as some forms of congenital heart disease or other neurodevelopmental disorders with microcephaly.

**Differential Diagnosis Consideration:**

**Smith-Magenis Syndrome:** Also includes intellectual disability, developmental delays, and some facial features but typically involves sleep disturbances and behavioral challenges not mentioned here.

**Rubinstein-Taybi Syndrome:** Presents with intellectual disability, distinctive facial features, and broad thumbs, but the absence of this specific hand abnormality makes it less likely.

Given the overall presentation, **Cornelia de Lange Syndrome** is the most consistent diagnosis based on the provided phenotypic report.



**OpenAI GPT-4o (released in 2024) got it wrong, with no proper reasoning/chain of thought. (Correct answer is KBG Syndrome.)**

AI explainability  
Diagnosing a rare  
genetic disease:  
GPT-4o vs. OpenAI  
o1 ('chain of  
thought')

User

Make a diagnosis based on the following report:

Phenotypes:

- Intellectual disability
- Global developmental delay
- Macrodontia
- Triangular face
- Thick eyebrow
- Abnormality of the hand
- Short stature
- Motor delay
- Delayed speech and language
- Feeding difficulties
- Low anterior hairline
- Strabismus
- Anteverted nares
- Bulbous nose

Excluded phenotypes:

- Atrioventricular canal defect
- Patent foramen ovale
- Ventricular septal defect
- Pulmonic stenosis
- Bicuspid aortic valve
- Mixed hearing impairment
- Microcephaly
- Synophrys



### Chain of thought

So from the user's instructions, I need to "Make a diagnosis based on the following report", which includes:

Phenotypes: (listed ones)

Excluded phenotypes: (another list)

I need to figure out a diagnosis based on the phenotypic features presented, and considering that certain features are excluded.

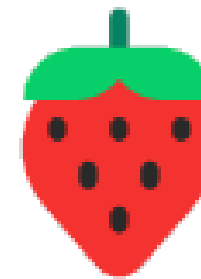
So, the goal is to provide a diagnosis based on the phenotypes.

I should carefully analyze the listed phenotypes and excluded phenotypes, and try to match it to known syndromes or conditions.

First, list the included phenotypes:

**Included Phenotypes:**

1. Intellectual disability
2. Global developmental delay
3. Macrodontia
4. Triangular face
5. Thick eyebrow
6. Abnormality of the hand
7. Short stature
8. Motor delay
9. Delayed speech and language
10. Feeding difficulties
11. Low anterior hairline
12. Strabismus
13. Anteverted nares
14. Bulbous nose



**OpenAI o1 (a preview was released in September 2024) got it right, with proper reasoning and a clear chain of thought explaining why it is KBG Syndrome and why other diagnoses have been excluded.**

Don't perform or facilitate the following activities that may significantly impair the safety, wellbeing, or rights of others, including: Providing tailored legal, [medical/health](#), or financial advice without review by a qualified professional and disclosure of the use of AI assistance and its potential limitations.

OpenAI Usage Policies, 10 Jan 2024



**Excluded phenotypes:**

1. Atrioventricular canal defect
2. Patent foramen ovale
3. Ventricular septal defect
4. Pulmonic stenosis
5. Bicuspid aortic valve
6. Mixed hearing impairment
7. Microcephaly
8. Synophrys

So what syndrome is characterized by the included features and lacks the excluded features?

First, it's intellectual disability and global developmental delay.

Then facial features: triangular face, thick eyebrow, low anterior hairline, strabismus, anteverted nares, bulbous nose.

Then macrodontia (big teeth), abnormality of the hand, short stature, feeding difficulties, delayed speech, motor delay.

I need to think of a genetic syndrome that matches these features.

Notable in the facial features are the triangular face, thick eyebrows, low anterior hairline, bulbous nose, anteverted nares.

Also, macrodontia (big teeth) is a distinctive feature.

Abnormality of the hand, short stature.

Does the combination of triangular face, thick eyebrows, bulbous nose, macrodontia, hand anomalies, short stature, etc., match any known syndromes?

Well, let's consider Cornelia de Lange syndrome.

But Cornelia de Lange tends to have synophrys (uni-brow), which is an exclusion (synophrys is excluded). Also, microcephaly is an exclusion, and it's common in Cornelia de Lange.

What about Rubinstein-Taybi syndrome? But that typically has broad thumbs and big toes, but also facial features including downslanting palpebral fissures, arched eyebrows, long eyelashes, and beaked nose.

Also, macrodontia is big teeth. Is that associated with any condition?

Macrodontia is a rare dental anomaly inclusive of large teeth.

Kabuki syndrome features include typical facial features with arched eyebrows with sparse lateral third (similar), but also includes short stature, intellectual disability.

But they commonly have congenital heart defects, which are excluded here (there are several major heart defects excluded).

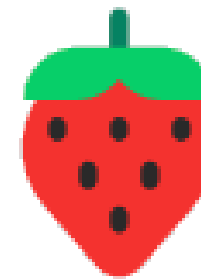
Trisomy 9p: has triangular face, big eyebrows, microcephaly (but microcephaly is excluded).

Is macrodontia associated with KBG syndrome? KBG syndrome is characterized by macrodontia of upper central incisors, short stature, intellectual disability, characteristic facial features including triangular face, bushy eyebrows, bulbous nasal tip, anteverted nares, low anterior hairline.

Yes! KBG syndrome seems to fit.

KBG syndrome is characterized by:

- Macrodontia (especially upper central incisors)
- Short stature
- Intellectual disability
- Triangular face
- Bushy eyebrows
- Anteverted nares



**OpenAI o1 (a preview was released in September 2024) got it right, with proper reasoning and a clear chain of thought explaining why it is KBG Syndrome and why other diagnoses have been excluded.**

Don't perform or facilitate the following activities that may significantly impair the safety, wellbeing, or rights of others, including: Providing tailored legal, [medical/health](#), or financial advice without review by a qualified professional and disclosure of the use of AI assistance and its potential limitations.

OpenAI Usage Policies, 10 Jan 2024

# Now imagine these powerful reasoning capabilities added to future medicine-specific LLMs or RAG!

## OpenAI o1

- Bulbous nasal tip
- Low anterior hairline
- Skeletal anomalies of hands
- Delayed speech and motor development
- Feeding difficulties sometimes
- Strabismus

So this seems to fit.

Moreover, cardiac anomalies in KBG syndrome are less common, and none of the cardiac anomalies are included in the phenotypes.

So the diagnosis is KBG syndrome.

### End of chain of thought

Based on the provided phenotypic features and the exclusion of certain conditions, the most fitting diagnosis is **KBG syndrome**.

### Rationale:

### Included Phenotypes Matching KBG Syndrome:

1. **Intellectual Disability and Global Developmental Delay:** KBG syndrome is characterized by mild to moderate intellectual disability and developmental delays in both motor and speech domains.
2. **Macrodontia:** One of the hallmark features of KBG syndrome is macrodontia, specifically of the upper central incisors.
3. **Triangular Face:** Individuals with KBG often have a triangular-shaped face due to a broad forehead and a pointed chin.
4. **Thick Eyebrows:** Bushy or thick eyebrows are a common facial feature.
5. **Abnormality of the Hand:** Skeletal anomalies, including brachydactyly (short fingers) and clinodactyly (curved fingers), are frequently observed.

## OpenAI o1

6. **Short Stature:** Many individuals exhibit growth delays resulting in short stature.
7. **Motor Delay and Delayed Speech and Language:** Developmental delays are common, affecting both motor skills and speech.
8. **Feeding Difficulties:** Infants may have difficulties with feeding due to hypotonia or other oral-motor issues.
9. **Low Anterior Hairline:** A low hairline at the front of the head is a recognized feature.
10. **Strabismus:** Eye anomalies, including strabismus (crossed eyes), are reported.
11. **Anteverted Nares and Bulbous Nose:** Facial features often include a bulbous nasal tip and anteverted (upturned) nostrils.

### Excluded Phenotypes Aligning with KBG Syndrome:

1. **Cardiac Anomalies:** While some individuals with KBG syndrome may have heart defects, they are not a consistent feature, and the specific heart anomalies listed (e.g., atrioventricular canal defect, ventricular septal defect) are not commonly associated with KBG.
2. **Hearing Impairment and Microcephaly:** These are not primary features of KBG syndrome, and their exclusion aligns with the diagnosis.
3. **Synophrys:** Unibrow or synophrys is not typically associated with KBG syndrome; in contrast, thick but separate eyebrows are characteristic.

### Conclusion:

Given the combination of characteristic facial features (triangular face, thick eyebrows, bulbous nose, anteverted nares, low anterior hairline), macrodontia, skeletal anomalies of the hands, short stature, developmental delays, and the absence of excluded phenotypes, **KBG syndrome** is the most appropriate diagnosis.

### References:

## OpenAI o1

1. **KBG Syndrome Overview:** KBG syndrome is a genetic disorder caused by mutations in the *ANKRD11* gene. It is inherited in an autosomal dominant pattern.
2. **Clinical Features:** The syndrome is characterized by macrodontia of the upper central incisors, distinctive facial features, short stature, skeletal anomalies, and developmental delays.

### Recommendation:

- **Genetic Testing:** Confirmation of the diagnosis can be achieved through molecular genetic testing for mutations in the *ANKRD11* gene.
- **Multidisciplinary Management:** Patients benefit from a team approach, including pediatricians, geneticists, dentists, speech therapists, and occupational therapists, to address the various clinical manifestations.

### Diagnosis: KBG Syndrome



**OpenAI o1 got it right, with proper reasoning and a clear chain of thought explaining why it is KBG Syndrome and why other diagnoses have been excluded.**

However, as of September 2024, OpenAI is reluctant to let users see inside the box: *“We have decided not to show the raw chains of thought to users. We acknowledge this decision has disadvantages. We strive to partially make up for it by teaching the model to reproduce any useful ideas from the chain of thought in the answer.”*



# AI explainability: Is it absolutely necessary?

- While AI explainability is a highly desirable feature, it is not always possible to achieve due to technology limitations, and AI can still be made useful and safe without it.
- We should remember in this context that there are many other things, expert diagnoses based on not-fully-explainable intuition or gut feelings; treatments, and therapies we routinely use in healthcare today but do not fully understand or cannot fully explain, with reliable evidence, how exactly they work (mechanism of action), yet we continue using them without waiting for such explanations of how they work to become available to us before we can successfully harness them for the benefit of our patients.



Future Healthcare Journal  
Volume 11, Issue 3, September 2024, 100171

## Explaining decisions without explainability? Artificial intelligence and medicolegal accountability

Melissa D. McCradden <sup>a b c</sup> ✉, Ian Stedman <sup>d</sup>

Received 26 July 2024, Accepted 6 August 2024, Available online 19 September 2024, Version of Record 19 September 2024.

<p>What makes a 'reasonable decision' for clinicians using AI has not been established. Explainability has been proposed as a key mechanism, but we suggest this is misguided.</p> 	<p>We discuss 2 problems with explainability: reliability and lack of patient-centredness. Instead, we look to how 'reasonableness' has been defined in the law and consider how it applies to AI.</p> 	<p>"Reasonable" decisions involve considering the AI prediction for the patient as an individual, within their unique social and environmental context.</p> 
--	--	---

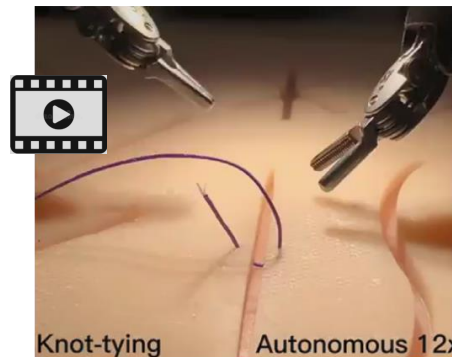
<sup>^</sup> <https://doi.org/10.1016/j.fhj.2024.100171>



< Highly recommended discussion:  
Annabelle Painter et al. **Explaining Explainable AI (for healthcare)**. YouTube, 2024 Aug15. Available at <https://www.youtube.com/watch?v=d5ZMVlgQ0jM>  
(running time: 58 minutes)

# Final thought

- AI holds immense potential in medicine and **is here to stay despite today's challenges.**
- As AI continues evolving and improving, its **ethical, safe and effective integration into clinical practice** will always require rigorous testing, **ongoing monitoring, evaluation and regulation**, and continuous collaboration among stakeholders, including:
  - regulatory bodies (US FDA, UK MHRA, EU EMA, China NMPA, etc.),
  - patients/patient advocates,
  - healthcare professionals and providers,
  - industry representatives,
  - the academia,
  - government partners and lawmakers, and
  - relevant international organisations (e.g., IMDRF, International Medical Device Regulators Forum, and others).
- **AI will undoubtedly reshape, if not disrupt, the medical profession in the coming years, and we should begin now rethinking medical education and curriculum.**



[AI Surgical Robot Transformer \(SRT\) - Automating surgical tasks with imitation learning \(Johns Hopkins & Stanford\)](#)



Derya Unutmaz • 2nd  
Professor at The Jackson Laboratory  
21h • 🌐



This is my final warning for those considering careers as physicians: AI is becoming so advanced that the demand for human doctors will significantly decrease, especially in roles involving standard diagnostics and routine treatments, which will be gradually and increasingly replaced by AI.

This is underscored by the massive performance leap of OpenAI's o-1 model, also known as the "Strawberry" model, which was released as a preview yesterday. The model performs exceptionally well on a specialized medical dataset (AgentClinic-MedQA), greatly outperforming GPT-4o. The rapid advancements in AI's ability to process complex medical information, deliver accurate diagnoses, provide medical advice, and recommend treatments will only accelerate.

Medical tasks like diagnosing illnesses, interpreting medical imaging, and formulating treatment plans will soon be handled by AI systems with greater speed and consistency than human practitioners. While this does not mean medical profession will disappear but as the healthcare landscape evolves in the coming years, the number of doctors needed will drastically shrink, with more reliance on AI-assisted healthcare systems. Thus in a decade or so from now there will be big surplus of physicians. Given it takes at least a decade to train as a physician specialist, people who are considering medicine as a high paying secure job should in future, carefully reconsider!

While human empathy, critical thinking, and decision-making will still play an important role in certain areas of medicine, even these may eventually be supplanted by future iterations of models like voice version of GPT o-1.

Consequently, medicine is becoming a less appealing career path for the next generation of doctors—unless they specialize in intervention-focused areas (such as surgery, emergency medicine, and other interventional specialties), though these, too, may eventually be overtaken by robotic systems...maybe within a decade or two.



# There is very much more to generative AI in medicine and healthcare

- There are many important issues and challenges surrounding the use of generative AI in medicine and healthcare that are beyond the scope of this short presentation.
- We have created a **series of horizon-scanning reviews** and a companion, regularly updated online resource to cover them. Scan the QR code to access them all.

2023:

**Generative AI in Medicine and Healthcare: Promises, Opportunities and Challenges**

Peng Zhang and Maged N Kamel Boulos. *Future Internet*. 2023 Aug;15(9):286

<https://doi.org/10.3390/fi15090286>

2024 incremental update:

**Generative AI in Medicine and Healthcare: Moving Beyond the 'Peak of Inflated Expectations'**

Peng Zhang, Jiayu Shi, Maged N Kamel Boulos. *Preprints*. 2024 Sep;2024090311

<https://doi.org/10.20944/preprints202409.0311.v1>

Companion online resource (where you can also download a PDF of this presentation):

<https://tinyurl.com/genai-med>



# And as always, PubMed® is your friend

- A good start, with many thousands of clinical-discipline-specific AI papers.

[Review](#) > [Strahlenther Onkol.](#) 2024 Aug 30. doi: 10.1007/s00066-024-02281-z.  
Online ahead of print.

## Artificial intelligence for response prediction and personalisation in radiation oncology

[Alex Zwanenburg](#) <sup>1 2 3</sup>, [Gareth Price](#) <sup>4 5</sup>, [Steffen Löck](#) <sup>6 7</sup>

[Affiliations](#) + expand  
PMID: 39212687 DOI: [10.1007/s00066-024-02281-z](#)

> [Lancet Digit Health.](#) 2024 Jul;6(7):e436-e437. doi: 10.1016/S2589-7500(24)00086-4.

## Virtual pregnancies: predicting and preventing pregnancy complications with digital twins

[Adrienne K Scott](#) <sup>1</sup>, [Michelle L Oyen](#) <sup>2</sup>

[Affiliations](#) + expand  
PMID: 38906606 DOI: [10.1016/S2589-7500\(24\)00086-4](#)  
[Free article](#)

> [Nat Rev Urol.](#) 2024 Sep 4. doi: 10.1038/s41585-024-00938-z. Online ahead of print.

## Artificial intelligence in the management of prostate cancer

[Raghav Khanna](#) <sup>1</sup>, [Alejandro Granados Martinez](#) <sup>2</sup>, [Nicholas Raison](#) <sup>3</sup>, [Sebastien Ourselin](#) <sup>2</sup>, [Alberto Briganti](#) <sup>4</sup>, [Francesco Montorsi](#) <sup>4</sup>, [Prokar Dasgupta](#) <sup>3</sup>

[Affiliations](#) + expand  
PMID: 39232058 DOI: [10.1038/s41585-024-00938-z](#)

[Comparative Study](#) > [Gynecol Oncol.](#) 2024 Oct;189:75-79. doi: 10.1016/j.jgyno.2024.07.007.  
Epub 2024 Jul 22.

## ChatGPT compared to national guidelines for management of ovarian cancer: Did ChatGPT get it right? – A Memorial Sloan Kettering Cancer Center Team Ovary study

[Lindsey Finch](#) <sup>1</sup>, [Vance Broach](#) <sup>2</sup>, [Jacqueline Feinberg](#) <sup>2</sup>, [Ahmed Al-Niimi](#) <sup>2</sup>, [Nadeem R Abu-Rustum](#) <sup>2</sup>, [Qin Zhou](#) <sup>3</sup>, [Alexia Iasonos](#) <sup>3</sup>, [Dennis S Chi](#) <sup>4</sup>

[Affiliations](#) + expand  
PMID: 39042956 PMCID: PMC11402584 (available on 2025-10-01)  
DOI: [10.1016/j.jgyno.2024.07.007](#)

[Review](#) > [J Med Internet Res.](#) 2024 Sep 16;26:e54737. doi: 10.2196/54737.

## Artificial Intelligence-Augmented Clinical Decision Support Systems for Pregnancy Care: Systematic Review

[Xinnian Lin](#) <sup># 1</sup>, [Chen Liang](#) <sup># 2 3</sup>, [Jihong Liu](#) <sup>4</sup>, [Tianchu Lyu](#) <sup>3</sup>, [Nadia Ghumman](#) <sup>3</sup>, [Berry Campbell](#) <sup>5</sup>

[Affiliations](#) + expand  
PMID: 39283665 DOI: [10.2196/54737](#)  
[Free article](#)