

# Google AI Search results & the Health Sector

An assessment of impact and concerns

# 1. Executive Summary

Google's 2024 launch of its AI search results panel has not only impacted traffic to support lines and health content across the sector, but there are concerns that this zero-click approach poses a risk to health outcomes.

70 organisations from across the health sector attended a virtual roundtable in March 2025, hosted by Marie Curie and the Patient Information Forum (PIF). These included NHS organisations and a host of national charities who share these concerns and have collaborated on this list of identified risks.

# 2. Working together

Google is a trusted source for access to health information. Research by Ipsos for PIF found 50% of people use Google to search for health information, a similar number to the NHS website. All summaries present a risk to this position of trust. We want to work with Google to identify trusted sources of UK information, improve localisation of summaries and onward signposting. We believe All can support better health outcomes and we make a series of recommendations including:

- Suspending AI summaries on health topics in the UK until accuracy and safety issues are resolved.
- Agreeing a framework for the verification of trusted health information in the UK.
- Ensuring results prioritise verified UK-based sources.
- Ensuring critical health queries are routed to NHS-approved content.
- Providing clearly visible context and support options at every stage.
- Introducing explicit warnings that AI summaries on health topics may not be accurate and should not be taken as authoritative or regulated advice.
- Providing directions in health searches for how to seek a healthcare professional.

# 3. Specific concerns

#### 1. Risks to Health Outcomes

- Health organisations warn that AI-generated information panels may cause users to miss important source content, hindering a nuanced understanding of their condition and limiting access to support. The outcome could be a reduction in quality of life.
- By giving the exact (potentially inaccurate) information for which people search with no context or clear onward user-journey, Google's results risk widening the health inequality gaps these organisations aim to reduce.
- The PIF/Ipsos published research showed only 1 in 10 people are signposted to patient support groups from the NHS. The vast majority reach them through independent search with charities considered the most trusted sources.

We believe it is as critical for Google to maintain its position of trust by providing accurate source data as it is fundamental for health organisations that Google guides people to robust, researched health information and offers of care.

#### 2. Quality and Accuracy of Content

- Charities report their content has been altered in search results. See Appendix for examples.
- Use of authoritative language and references may give invalid source sites the appearance of expertise, making them more likely to be picked by AI.
- Al-generated summaries lack nuance, particularly in health-related topics where circumstances differ; an Al answer for "what are the causes of MASLD" is unlikely to include side links for "what to do if you are worried".
- o It is virtually impossible to obtain the same answer from AI even for questions worded identically, leading to urgent concerns about the consistency and accuracy of information provided, or the ability for someone to revisit advice.

#### 3. Lack of localised source material

- o One of the most repeated concerns is that AI search results favour US-based sources and do not adequately cater to UK or more regional audiences.
- Google has said this is a target for change; we would like to ask for visibility around how source content is being chosen and given weight.

Better still, the sector offers to work with Google to develop a shared understanding of what would constitute a 'trusted publisher' in the Health Information space. PIF runs a certification scheme for health information in the UK under which 150 organisations are certified (the PIF TICK). This forms the basis of the content standard for NHS England and is already used as the basis forof criteria for YouTube's 'Health Shelf'.

#### 4. Health Inequality and Literacy

- Narrowly focused AI responses could widen health inequalities by failing to account for varying literacy levels and broader informational needs.
- People searching for health results may be in a state of panic and in need of additional context. While Google provides links to sources, search results are isolated from the context of other information or support, creating extra obstacles for those seeking help, especially affecting those less digitally included or literate.
- The language used in AI summaries is identified as confusing. For example, wording such as "this is especially the case" indicates information comes from multiple sources; however, users in a state of stress or with low reading age need clear, precise language to guide them to the help they need.

#### 5. Access to support

- Some members have noticed a decline in organic search traffic and have concerns that users are bypassing signposting to source content, which may offer hands-on support and care or contain information they are unaware they need.
- Around 10% of organisations attending the roundtable reported a reduction in calls to their helpline; and 50% reported a fall in traffic. 70% of those reporting a fall attributed this to the introduction of AI overviews rather than other changes in the environment.

### 4. Summary

Representatives from the health sector ask to address the concerns raised in this document directly with Google. We seek reassurance that the identified risks to health outcomes in the UK from the use of Al-generated search panels will be addressed and look forward to a successful collaboration on the maintenance of trust in Google's health-related search results.

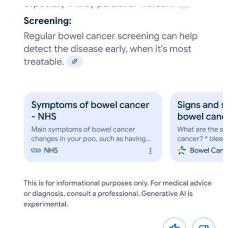
# 5. Appendix 1: Additional questions raised

#### 1. Interpretation of Health-related queries

- Can Google's search engine tell if someone is asking a health-related question? Is it possible to put in special measures to ensure results come from trusted content?
- Does Google have mechanisms able to identify and respond to critical or harmful queries such as health emergencies or medication issues which then direct to relevant sources such as NHS 111 or a charity helpline?

#### 2. Safeguards and Warnings for Health-Related AI Content

- There is a strong wish for clearer, tailored disclaimers on the AI health information, such as urging users to verify data with trusted sources.
- There is a recommendation that Google make it a lot more clear which content has been AI generated. The example to the right shows an AI statement which could be made more prominent, and/or with stronger wording.
- Would Google be willing to introduce an obvious way to turn off AI summaries when searching? Combined with clear statementing around AI-generated content, this may help safeguard users



 There are concerns that harmful searches typically performed by those in fragile mental states could be delivering results that drive them to further harm.

#### 3. Question: Why is Health the biggest sector for AI overviews?

- o Is this a reversal of the decision?
- Concerns were raised that the disproportionate focus on health content in Al summaries could pave the way for health/wellbeing/'wellness' advertising opportunities in the future.

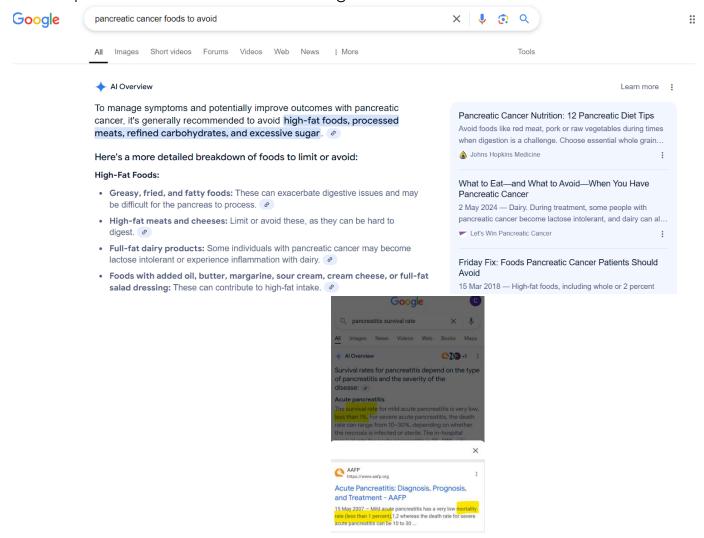
# 6. Appendix 2: Examples of inaccurate Google AI results

#### 1.1. Pancreatic Cancer UK

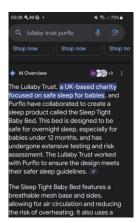
The Google panel below recommends that people with pancreatic cancer avoid highfat and other foods as displayed in a list; however this is the exact opposite of recommendations. Many people with pancreatic cancer will need to eat high fat foods to help them put on or maintain weight, including all the food listed here.

This search result is potentially dangerous to those who need to put on weight so that they can receive cancer treatment. If people follow this, it will lead to worse health outcomes as they potentially won't be able to have treatment for their cancer.

With pancreatic cancer survival rates as low as 1% after 3 years, outcomes can depend on the speed at which treatment can be given.



#### 1.2. The Lullaby Trust



The charity reports this inaccurate Al-generated search result which poses a potential child safety risk as well as a reputational risk.

The result states that the charity has collaborated with Purflo on a baby bed, which is particularly concerning as Purflo sells products that the charity believes are unsafe.

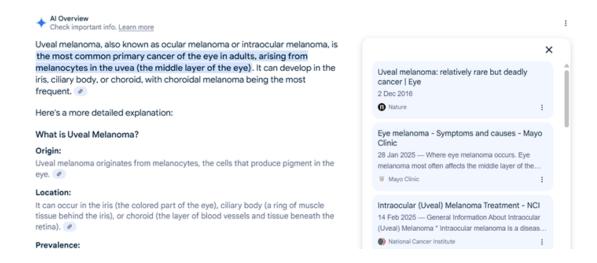
The Lullaby Trust states that it therefore has never and would never work with Purflo. There is no information on their website about Purflo, they cannot find this information on Purflo's site or anywhere online, suggesting that this is an AI hallucination.

#### 1.3. Melanoma Focus

The charity has noticed an increase in US organisations appearing high in Google search results, which is concerning as treatments and medical information differ between the US and the UK.

Below is an example below of AI search results for 'uveal melanoma'. This is a rare type of melanoma for which the charity produces up-to-date UK national guidelines.

When doing a search for 'uveal melanoma' on Google however, the UK national guidelines are not represented. The first link is to an out-of-date 2016 article in a medical journal, the next to Wikipedia and the last to a US organisation.



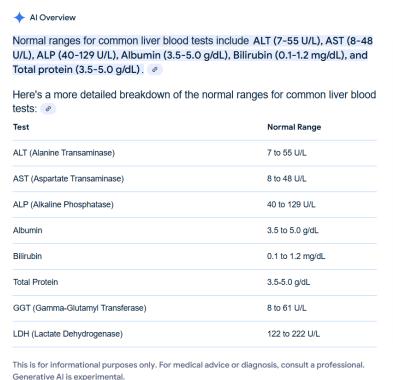
#### 1.4. British Liver Trust

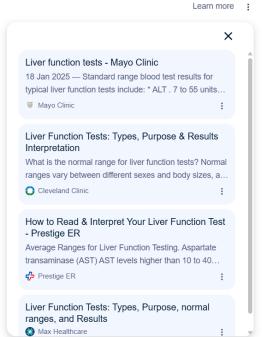
The AI summary results for "What is the normal range for liver blood tests" provides misleading information for Liver Function Tests, which are commonly requested by GPs in the UK to diagnose and monitor liver disease and other conditions. Many people with liver disease show no symptoms until the final stages, making accurate testing crucial.

We know from feedback that these results are often delivered with no context for what the numbers mean, leading people to search online for explanations. The first summary is only accurate for adult men and does not account for wide variances for sex, ethnicity, and age, so a great many will struggle to understand their results.

Three references are from the US and one is from India and both countries have different standards than the UK. As a result, these results could differ from what is considered 'normal' by up to a factor of 10,

Crucially for health outcomes, these result lack critical context. Some people with serious liver disease may have a normal result and could therefore be discouraged from seeking or attending a critical follow-up healthcare meeting. In contrast, 20% of those with results outside the range will have a completely healthy liver and may worry unduly.





#### 1.5. Mind

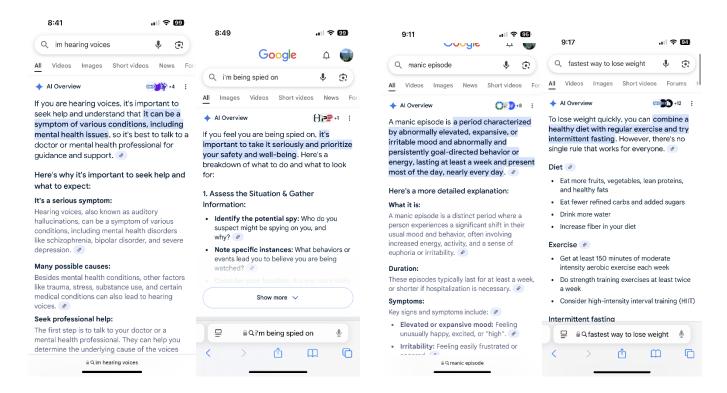
Searches that may be performed by someone experiencing psychosis, in a manic episode, or with an eating disorder, also deliver concerning results.

For 'Hearing voices' and 'psychosis', while the information presented is broadly correct, it signposts people to untrustworthy resources such as WebMD; or to 'Voice Collective', which is a young person only resource – and the link provided goes to ChildLine instead.

Common searches performed by someone experiencing psychosis, such as 'I'm being spied on', generate AI overviews which may dangerously encourage someone's delusions, such as focusing on identifying the spy and the need to take it seriously.

Results for 'manic episode' bring up an AI overview which says that mania lasts for a week, whereas an episode must last a <u>minimum</u> of a week to be considered mania, which may lead people to avoid seeking help.

Al overviews regularly promote harmful calorie-restricted diets to those with eating disorders due to the heavy presence of poor-quality nutrition content online. The example below promotes a recommendation to only eat one 500 calorie meal some days whereas a diet of anything less than 1200 calories a day is considered starvation.



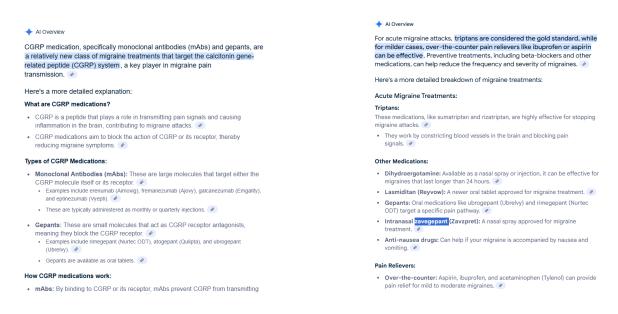
#### 1.6. Migraine Trust

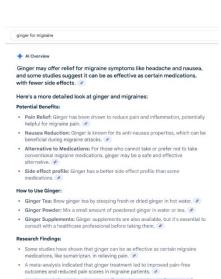
The Migraine Trust reports several examples of concerning results including localisation issues and inaccurate advice.

The first results shown are for a search of CGRP medication which includes medicines that aren't available in the UK

Similarly, a search for the best treatments for migraines also lists non-UK medications.

Both results show US brand names for medication which may be available in the UK but called something else.





 Another study found that ginger was effective in reducing the frequency of monthly migraine episodes. An AI search for migraine treatments also delivered a recommendation to use ginger - then when searching specifically for 'ginger for migraine', it seems to be advocating ginger as a treatment for migraine.

The reality is that the trials it mentions are very small, unreliable and from quite a long time ago; there has never been enough evidence to recommend ginger in migraine treatment.

#### 1.7. MS Society

Below is an example of an AI overview getting something fundamentally wrong about multiple sclerosis (MS).



It may be that the Al logic assumed an order because it knows secondary usually comes after primary, however secondary and primary progressive MS do not follow on from each other. They are different diagnoses with very different treatment options and health outcomes, which makes this a particularly problematic example.

If a user trusts that AI results are correct it can lead to all sorts of problems like pursuing the wrong treatment, or giving up on treatment altogether.

The charity has since retrained the result by using a webpage to answer the specific question. This means the answer (below) is now correct. This was only caught by chance, however. Finding and correcting errors created by AI is not a sustainable approach for charities already pressed for resources.



#### 1.8. **BUPA**

BUPA quote an example in which a random dental site had copy/pasted a Bupa blog onto their own site (https://www.toothtalk.uk/preston/is-vaping-harmful/) and they were cited in the search summary rather than Bupa, despite the content being identical.

This may be because they'd published that information more recently, but that's not necessarily a good gauge of accuracy/relevance.

BUPA has emphasised the concern that AI summaries lack context. They are transparent about how they produce information and clearly outline their stance on this, going to great lengths to remain unbiased and accurate, substantiating each statement with the best possible evidence and getting clinical sign-off.

BUPA are another organisation actively looking for ways to target inaccuracies in results and which has seen a reduction in traffic directly correlated to the launch of the Al search results panel.

# 7. Appendix 3 - Authors

This document has been compiled by Marie Curie, the Patient Information Forum and Macmillan Cancer Support.



Marie Curie is the UK's leading end-of-life charity. We bring 75 years of experience and leading research to the care we give you at home, in our hospices and over the phone. And we push for a better end of life for all by campaigning and sharing research to change the system.



The Patient Information Forum (PIF) is the independent membership body for people working in health information PIF runs the UK quality mark for health information – the PIF TICK.



Because cancer can affect your life in so many ways, we do whatever it takes to give people the support they need.

Additional content has been provided by BUPA, British Liver Trust, The Lullaby Trust, Melanoma Focus, Migraine Trust, MIND, Pancreatic Cancer UK and MS Society, with input from nearly 70 organisations from across the health sector.