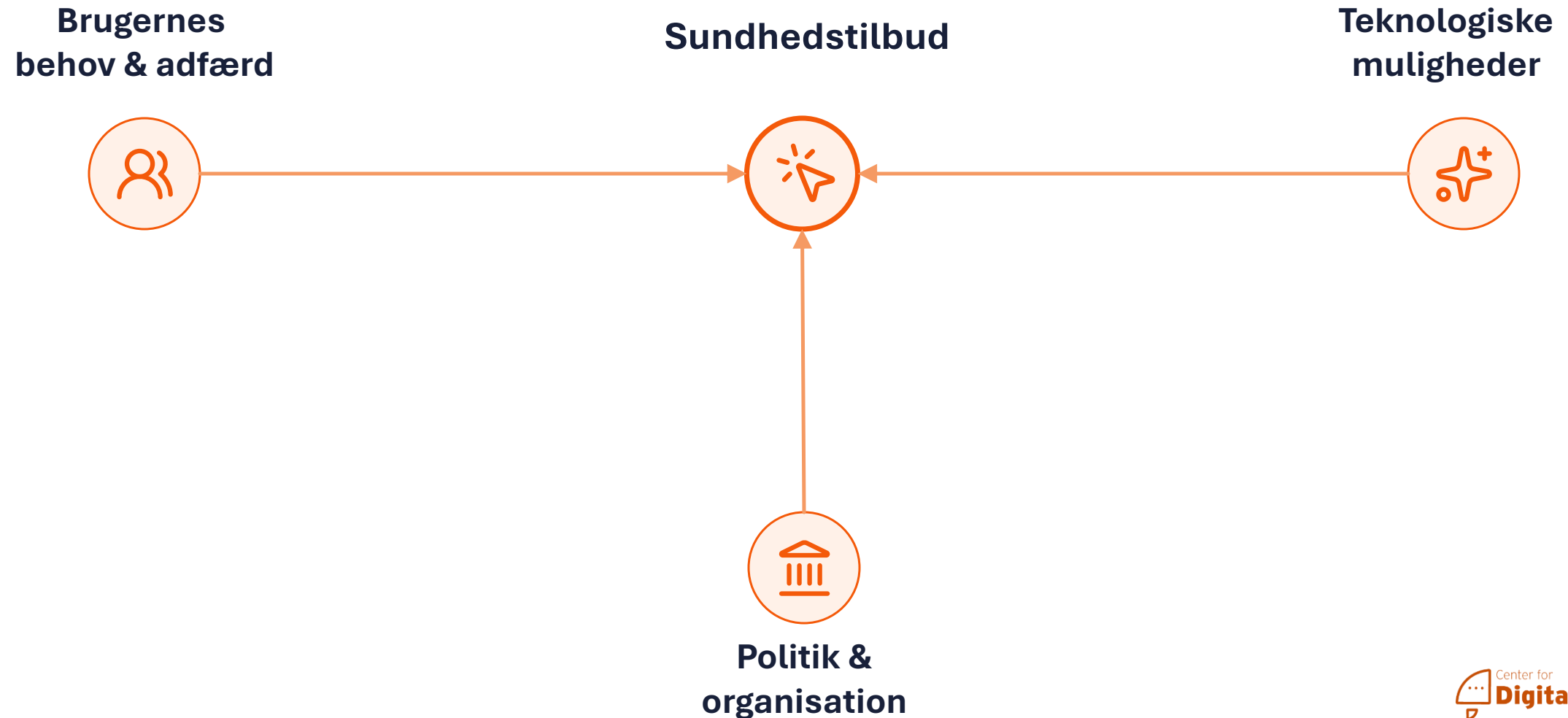


Tor Juul Groth
Digital produktleder
Center for Digital Psykiatri

...og nu også forfatter af AI-rapport

Rollen som digital produktleder



Voksne

Generaliseret angst

Børn

Kvinder

Stress og

Social

Anfald

At miste

✦ AI-oversigt

Angst viser sig gennem **fysiske, psykiske og adfærdsmæssige symptomer**, som kan opdeles i tre hovedgrupper. Du kan læse mere om de forskellige tegn hos [Psykiatrifonden](#). 

Fysiske symptomer

Kroppens alarmsystem aktiveres, hvilket kan føles som:  Internetpsykiatrien

- **Hjertebanken** eller hurtig puls
- **Sveden** og rysten i kroppen
- **Åndenød** eller kvælningfølelse
- **Trykken** eller smerter i brystet
- **Svimmelhed** og uro i maven eller kvalme
- **Mundtørhed** og muskelspændinger  Psykiatrifonden +2

Psykiske og kognitive symptomer

Tankerne og følelserne påvirkes af angsten:  Bedre Psykiatri +1

- **Vedvarende bekymringer** og katastrofetanker
- **Tankemylder** som er svært at styre

Stil et spørgsmål



Generaliseret angst, symptomer og tegn - Sundhed.dk

Angsten gør, at du føler dig konstant belastet, anspændt, ikke kan sove om natten og er træt. Du kan også få fysisk...



Angst - Psykiatrifonden

Symptomer. Typiske tegn på angst. Hjertebanken, sveden og rysten; Svimmelh...



Om angst hos voksne - Sundhed - Region Midtjylland

Krop - symptomer på angst. Man kan opleve: Hjertebanken, sveden, rysten, mundtørhed....



Vis alle

Da Googles AI og ChatGPT ændrede alt



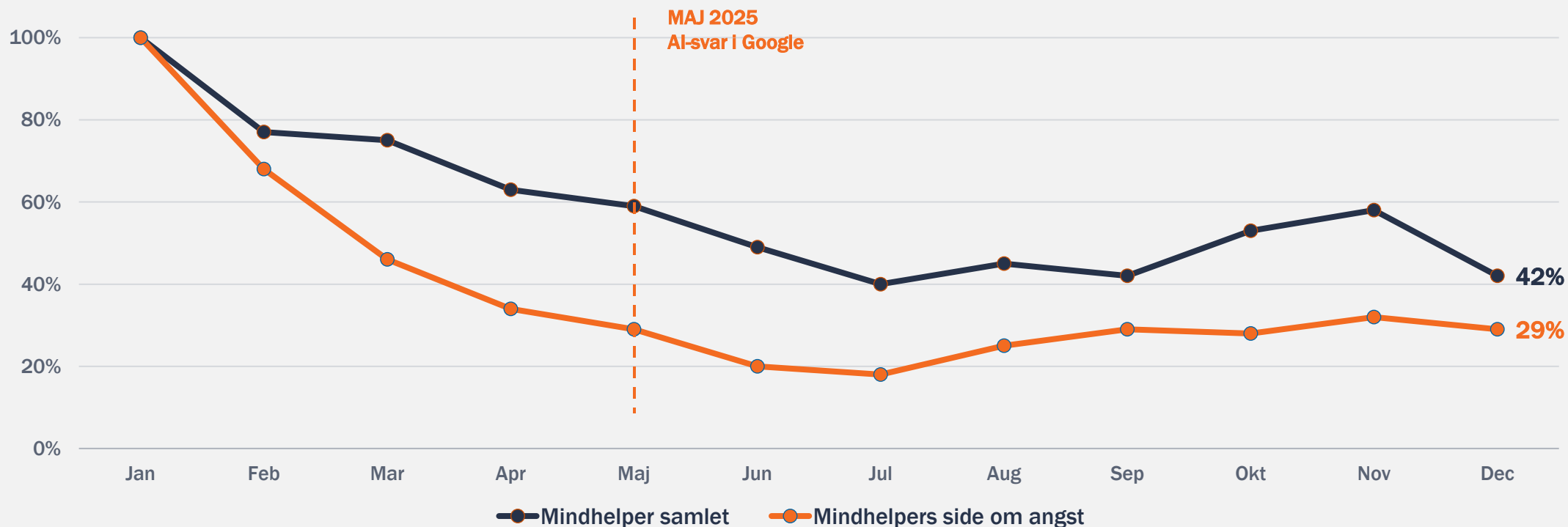
Center for Digital Psykiatri (2026). Mental sundhed: Sådan ændrer AI danskernes søgen efter viden og hjælp.

- Anledning: Googles AI-svar i Danmark (maj 2025)
- Først intern strategiproces, men blev til publikation
- Kortlægger forandringerne i brugeradfærd
- 4 mulige fremtidsscenarier og analyse på tværs
- Udgivet februar 2026, så meget er sket siden da, men analysen holder overordnet set stadig



2025 ændrede mindhelper.dks rolle i samfundet

Andel af trafikken i januar 2025



ChatGPT skabte en ny selvhjælpsadfærd



- ChatGPT bruges til svære tanker, mistrivsel og relationer
- Denne brug af ChatGPT er meget tabubelagt
- Almindelig glidebane fra lektier til støtte og coaching
- Tip: Læs rapportens kvalitative materiale!
- 4 ud af 10 18-29årige har brugt chatbot til at tale om deres mentale helbred eller personlige udfordringer
(Kilde: Psykiatrifonden, 2026)

VIVE: Kessing et. Al. (2026)
+ Psykiatrifonden (2026)



Andel af læste ord om mental sundhed, som er genereret af AI

3 ud af 4

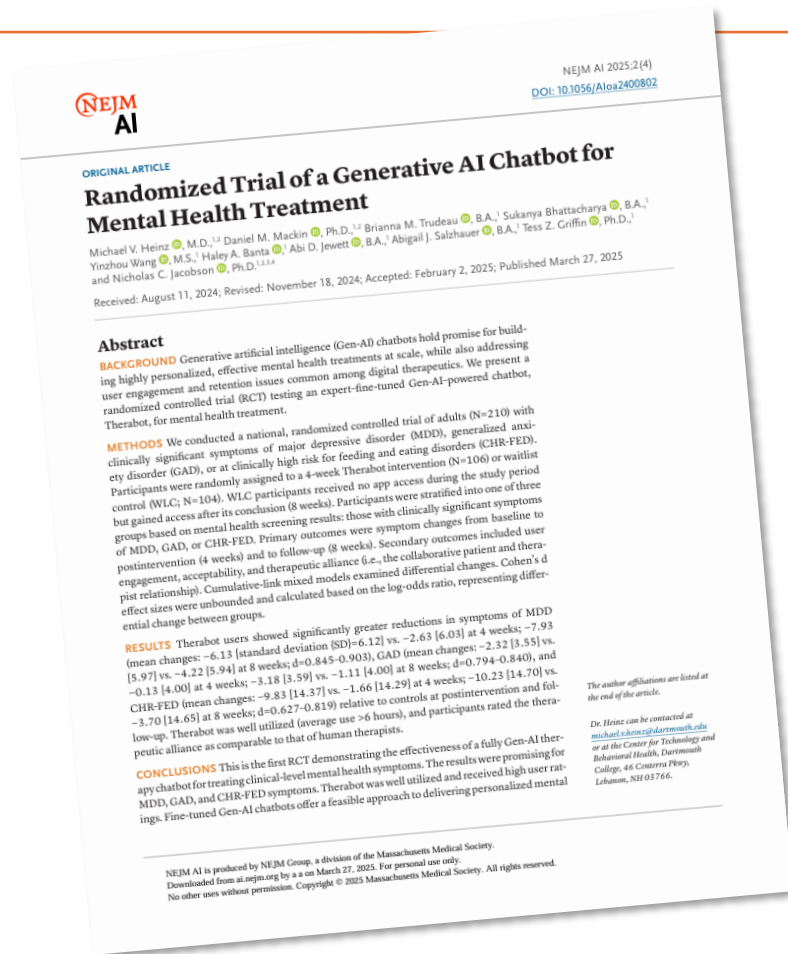
*Tors gæt-stimat på baggrund af:
Cloudflare, SE Rankings audit af Google AI Overviews, Agarwal & Sen (2026)*



Mulighed: 24/7 personlig afklaring, sparring og støtte



Mulighed: Specialiseret chatbot + kliniker



Heinz et. al. (2025)

- RCT med 210 voksne gennemført i foråret 2024
- Ikke generel chatbot: eksperttuning i manualiseret kognitiv adfærdsterapi og human-in-the-loop
- Store symptomfald mod venteliste efter 4 og 8 uger
- 95 % brugte løsningen med høj aktivitet og terapeutisk alliance på niveau med samtaleterapi
- Alle svar overvåget: 28 indgreb – få, men afgørende

Risiko: Fejl leveres i et overbevisende svar



- 49,6 % af AI-svar var problematiske ifølge eksperter
- 19,6% af dem var meget problematiske
- Chatbots afviser nærmest aldrig at besvare et sundhedsspørgsmål
- Ingen chatbot leverede et eneste svar med fuldt korrekte kildehenvisninger

Risiko: Forstærker uhensigtsmæssig adfærd



- Vi har alle haft det på fornemmelsen, nu er det dokumenteret...
- Chatbots forstærker uhensigtsmæssige tanke- og handlemønstre med et “Vulnerability-Amplifying Interaction Loop”
- Mønsteret kan ikke læses af ét svar, men bliver tydeligt, imens chatbotten tilpasser sig brugeren
- 810 samtaler med 9 populære chatbots via 30 simulerede brugerprofiler

Weinhhammer et. al. (2026)

Nej, de bliver ikke nødvendigvis bedre...

MindBench.ai

HealthBench-Psych: A Mental Health Subset of OpenAI's HealthBench

Matthew Flathers¹ Phuong Anh Nguyen¹ Jill Noorily¹ Julian Herpertz²
Meiting Chen¹ Jasreen Multani¹ Samuel Powell¹
Mason Granof¹ Mark Kalinch¹ John Torous¹

¹Division of Digital Psychiatry, Beth Israel Deaconess Medical Center, Boston, MA
²Department of Psychiatry, Beth Israel Deaconess Medical Center, Boston, MA

Abstract

General-purpose health benchmarks increasingly anchor claims about LLM medical performance, but they are not always resolved by clinical specialty, making domain-specific performance hard to isolate. Mental health performance is of acute public-health concern as millions of people turn to LLMs for psychological support, and most existing evaluations are bespoke academic benchmarks that are difficult to integrate into developer workflows. We introduce HealthBench-Psych and HealthBench-Psych-Hard. We screened HealthBench's 5,000 physician-authored rubrics for mental-health relevance with a transparent LLM-applied rubric, then validated the subset through two rounds of blinded clinician review with concealed known-exclude controls, yielding 610 conversations (12.2% of the corpus). Evaluating 20 frontier and open models under a cross-vendor panel of three LLM judges, we find a statistically tied frontier cluster, measurable refusal behavior across judges ($\tau \geq 0.92$). We release the subset, pipeline, model responses, grades, and analysis code as a reusable resource.

1 Introduction

People increasingly bring mental-health concerns to general-purpose LLMs, often in moments of acute distress and outside any clinical pathway (Bodner et al., 2026; Roussmaniere et al., 2026).

Preprint v1.0.0...
Code: <https://github.com/mindbench-ai/healthbench-psych>
Data: <https://huggingface.co/datasets/mindbench-ai/healthbench-psych>
Correspondence: mflathers@bidmc.harvard.edu

Evaluating how these systems behave in mental-health conversations is therefore pressing. Yet the existing mental-health evaluation literature clusters narrowly on suicide-risk detection and diagnostic classification (Flathers et al., 2026), and risk classification and board-style diagnostic reasoning are poor proxies for the open-ended, ambiguous, multi-turn help-seeking that likely constitutes most real usage (Raji et al., 2025; Stade et al., 2024; Luo et al., 2025).

Conversational health benchmarks have begun to close this realism gap. HealthBench (Arora et al., 2025) is the most prominent example, comprising 5,000 multi-turn conversations, each response graded against physician-authored rubrics. Its format is free-text rather than saturated multiple-choice (Singhal et al., 2025), and the benchmark is openly licensed. But it carries no specialty meta-data. Its item tags describe conversational behavior (holding, emergency referral, context-seeking) and physician-uncertainty structure, never specific medical topic. This matters because HealthBench has begun to appear in AI company model cards and technical reports as the core of their health testing, making it one of developers' most widely used mental-health evaluation instruments (OpenAI, 2026; Anthropic, 2026; Meta Superintelligent Labs, 2026). A lab reporting a strong HealthBench number may therefore appear competent in psychiatry and clinical psychology, with no clear way to isolate that domain performance on its own. We address this gap in two stages. We introduce HealthBench-Psych, a clinician-validated subset of 610 mental-health-relevant HealthBench conversations, and use it for the first specialty-resolved comparison of 20 frontier and open models under a cross-vendor panel of three LLM judges. We contribute: (1) a transparent, reproducible pipeline for

- Nyt preprint, der sammenligner 20 sprogmodeller på et helt nyt benchmark for mental sundhed
- Mental sundhed: Claude Fable 5 < Claude Opus 5. Kimi K3 < Kimi K2.6. GPT-5.6-sol < GPT-5.5
- I vanskelige samtaler om mental sundhed falder kvaliteten af svar helt ned til omkring 0,40 (0,0-1,0)
- Kun 13 ud af 12.200 gange afviste en chatbot at svare på et spørgsmål om mental sundhed

Flathers et. al. (2026)

Anbefalinger

1. Forstå hvor markant anderledes adfærd er nu sammenlignet med for 1,5 år siden
2. Sundhedstilbud skal matche den aktuelle brugeradfærd og teknologi
3. Det offentlige skal have få brugerflader og én "digital fordør til sundhedsvæsenet"

