

In Health-Related Stigma AI is Mapping it Faster than it is Fixing it

Researchers explored the role of [artificial intelligence](#) (AI) in health-related stigma.

Health-related stigma is a significant barrier to equitable health, and reducing stigma remains a major objective in clinical practice and [health policy](#). Digital technologies are increasingly used in health communication, prevention, and treatment, which raises concerns about whether and how they alleviate or reproduce stigma, with AI further transforming this landscape.

While natural language and [machine learning](#) (ML) models have been used to detect stigmatizing language and support some automated interventions, they can unintentionally increase discrimination and bias. New ethical issues about data misuse and misrepresentation have also emerged with the advent of large language models. That AI can reinforce or decrease social exclusion highlights the need to analyze how stigma manifests in AI systems across health contexts.



Study and Results

In the present study, researchers reviewed current evidence on the relationship between health-related stigma and AI. As a scoping review, the study was designed to map the field rather than estimate pooled effects or determine clinical effectiveness. First, they searched 10 databases for studies on AI, health, and stigma published from 2012. Eligible studies were those that implemented or discussed AI applications or algorithms, addressed health-related condition(s), and assessed stigma and related constructs (prejudice, [discrimination](#), and stereotypes).

Studies on stigma related to [coronavirus disease 2019](#) (COVID-19) were excluded. Database searches identified 27,552 records, of which 11,769 underwent title/abstract screening after deduplication. Following full-text review and reference searching, 70 studies, published between 2016 and 2025, were included. Early publications were mostly centered on AI quantifying stigma primarily through natural language processing (NLP) analyses.

Studies on AI-related increases in stigma and the [influence](#) of stigma on AI usage began in 2019, whereas those on AI reducing stigma emerged from 2020 onward. Most studies were conducted in the United States (32), followed by the United Kingdom (10) and Singapore (7). ML and NLP were the most commonly used AI approaches, predominantly applied to sentiment detection, text classification, and stigma-related outcome prediction.

Twenty studies evaluated AI services, such as chatbots, diagnostic tools, or virtual agents, while six studies assessed generative systems or language models. Fifty-three studies analyzed stigma related to [mental health](#), representing about 76% of the included literature. Stigma type was most often classified as public stigma, followed by self-stigma, although more than one-quarter of studies did not specify the target stigma. The researchers identified four research themes across publications: AI measuring stigma, stigma influencing AI usage, AI increasing stigma, and AI reducing stigma.

In 42 studies, AI was used to detect, stratify, or measure stigmatizing content. Across these studies, two major types of research questions emerged. One type involved using AI to detect and characterize stigma in [public discourse](#), while the other focused on improving AI's measurement capabilities. These studies used ML and NLP approaches to examine large-scale digital corpora for stigmatizing language.

X (formerly Twitter), Reddit, Weibo, and Facebook were the main sources analyzed for stigmatizing language. The prevalence of stigma within digital corpora was highly variable, ranging from less than 1% to more than 40%. [Obesity](#)-related stigma was usually rare, while schizophrenia-related stigma was more common. Moreover, stigmatizing content often had negative or exclusionary expressions. Fifteen studies examined how stigma affects AI systems' trust, adoption, or acceptance.

The team identified two contrasting patterns across studies on how stigma influences AI use: The anonymity of AI created an [environment](#) that encouraged disclosure and use for some people; in contrast, others were hesitant to use AI because of concerns that it might increase stigma. Notably, there was a greater willingness to use AI for sensitive or stigmatized health conditions. Furthermore, nine studies examined how AI increases stigma.

These studies investigated whether AI itself or people using AI applications could increase stigma toward certain health groups. In these studies, prompts with references to stigmatized conditions, e.g., disability, [mental illness](#), etc., produced more fearful and negative responses than those with neutral terms. Other studies found that language model embeddings and image-generation models could reproduce negative associations or harmful visual stereotypes of illness. In one study, healthcare professionals who were exposed to ML-based predictive assessments reported greater fear and less anger toward patients.

There were only four studies on AI reducing stigma. These studies used conversational agents engaging users in mental health-related dialogue and found a decrease in stigmatizing attitudes. The effects were most pronounced when conversational agents shared first-person narratives, i.e., living with a health condition. Moreover, recent research has focused on adapting such agents as educational tools to decrease stigma among [healthcare providers](#). However, this evidence remains preliminary and largely based on small-scale experimental studies rather than long-term clinical or community evaluations.

Conclusion

In sum, research on health stigma and AI has been growing but remains uneven, dominated by studies on AI use to detect and quantify stigma. As such, AI is currently more akin to an analytical tool than to decreasing stigma and promoting health equity. Notably, studies have disproportionately focused on mental health conditions, with limited representation of conditions recognized in stigma research (e.g., [leprosy](#)).

The review also highlighted inconsistent definitions of stigma, limited cross-cultural perspectives, scarce real-world evaluations, and few multi-modal studies beyond text-based AI. Overall, advancing research in this field requires deeper integration of clinical, social, and [computational perspectives](#).

Source:

<https://www.news-medical.net/news/20260601/AI-is-mapping-health-stigma-faster-than-it-is-fixing-it.aspx>