

Могут ли большие языковые модели служить консультантами при анализе причин смерти в судебной медицине? Многомерная оценка

Источник: Frontiers in AI — Medicine

Оригинал: <https://www.frontiersin.org/articles/10.3389/frai.2026.1752043>

LLM

decision support

галлюцинации ИИ

диагностика

судебная медицина

Введение

Большие языковые модели (Large Language Models, LLMs) были предложены в качестве инструментов поддержки принятия решений в медицине, однако их роль в судебно-медицинском анализе причин смерти остается неизученной.

Методы

В данном исследовании мы использовали 118 реальных случаев, охватывающих различные категории смерти, чтобы систематически оценить эффективность четырех репрезентативных LLM (**GPT-4o**, **OpenAI o3**, **Gemini-2.5pro** и **DeepSeek-R1**) в судебно-медицинском анализе причин смерти. Два старших судебно-медицинских патологоанатома независимо оценивали возможности принятия решений каждой моделью в отношении качества логических выводов и точности заключений. Эти показатели оценивались с использованием экспертной системы баллов по 5-балльной шкале Лайкерта, где исходные аналитические заявления и юридически обоснованные экспертные заключения служили объективными «золотыми стандартами». В рамках дополнительного исследования мы изучили

потенциал применения локально развернутой модели с открытым исходным кодом **DeepSeek-R1:32b**. Кроме того, был проведен целевой ретроспективный анализ для количественной оценки частоты возникновения и типологии галлюцинаций ИИ.

Результаты

DeepSeek-R1 продемонстрировала статистически значимое преимущество в показателях качества логических выводов по сравнению с **GPT-4o** ($p = 0,015$, $rrb = 0,28$) и **Gemini-2.5pro** ($p = 0,000003$, $rrb = 0,46$), в то время как между четырьмя моделями не было обнаружено статистически значимых различий в показателях точности заключений. Локально развернутая модель **DeepSeek-R1:32b** также не показала статистически значимых отличий от **GPT-4o** в точности заключений. Тем не менее, галлюцинации постоянно проявляются в отчетах о результатах всех LLM.

Обсуждение

LLM могут представлять ограниченную вспомогательную ценность в анализе причин смерти, но они не должны заменять окончательное суждение судебно-медицинских экспертов. LLM по-прежнему требуют экспертного контроля для обеспечения целостности доказательств и минимизации таких рисков, как галлюцинации. LLM с открытым исходным кодом могут дополнительно снизить опасения, связанные с конфиденциальностью данных, и обеспечить практическую поддержку при анализе причин смерти.