

Liebe Abonentin

Lieber Abonnent

Schon lange haben wir Patient:innen, die sich beim Dr. Google über Diagnose und Behandlungsmöglichkeiten informiert haben - was nicht immer hilfreich ist. Nun geht es einen Schritt weiter und die Informationssuche wird durch künstliche Intelligenz unterstützt. Ist dies hilfreich? Vor Kurzem erschien eine Studie, die über die Qualität dieser KI-unterstützten Ergebnisse berichtet. Ein Artikel dazu in der Schweizerischen Tagespresse lautete „Warum Chatbots mehr als 65 Prozent falsche Diagnosen stellen – riskante Gesundheitsberatung“. Wir haben die Studie für Sie dazu genauer angeschaut.

Archiv und Inhaltsverzeichnis LymphLetter

Sämtliche bisher erschienenen LymphLetter finden Sie auf der Seite ‚LymphLetter‘ an zwei Orten: Direkt unter dem Einleitungstext (Verzeichnis aller LymphLetter) in chronologischer Reihenfolge und weiter unten auf der Seite (Alle LymphLetter) mit dem Aktuellen beginnend. Das Inhaltsverzeichnis LymphLetter finden Sie im Download-Bereich: <http://lymphbildung.ch/downloads.php>

News

Kommende Veranstaltungen, neue Produkte – auch hier zu finden: <https://lymphbildung.ch/news.php>

Kurse

Die letzten Kurse haben bereits stattgefunden.

Fragen und Antworten

Hier können Sie Ihre Fachfragen stellen (nachdem Sie sich eingeloggt haben):

<http://lymphbildung.ch/support.php>

Herzliche Grüsse

Yvette Stoel & Team LymphBildung

Die Verlässlichkeit von KI-unterstützter medizinischer Assistenz für Patient:innen

Schon lange haben wir Patient:innen, die sich beim Dr. Google über Diagnose und Behandlungsmöglichkeiten informiert haben - was nicht immer hilfreich ist. Nun geht es einen Schritt weiter und die Informationssuche wird durch künstliche Intelligenz unterstützt. Ist dies hilfreich? Vor Kurzem erschien eine Studie, die über die Qualität dieser KI-unterstützten Ergebnisse berichtet. Ein Artikel dazu in der Schweizerischen Tagespresse lautete „Warum Chatbots mehr als 65 Prozent falsche Diagnosen stellen – riskante Gesundheitsberatung“. Wir haben die Studie für Sie dazu genauer angeschaut.

Einleitung

In diesem LymphLetter werden wir die Original-Studie sowie die Meinung einer Expertin in der Schweiz anschauen. Das Ziel ist herauszufinden, wie wir als Health-Professionals, in unserer täglichen Arbeit dieser Situation begegnen können.

Chatbots

Künstliche Intelligenz kann die Informationen aus dem weltweiten Netz gezielt zusammentragen und kombinieren. Deshalb ist entscheidend wie der Prompt formuliert wird. Die Anfrage muss detailliert genug geschrieben sein und vor allem die wichtigsten Informationen beinhalten. Wenn jemand Informationen über seine Diagnose oder Behandlungsmöglichkeiten suchen möchte, könnte dies schon ein Problem sein. Weiss ein Laie genug um die „richtigen“ Fragen zu stellen und die Antworten entsprechend zu beurteilen?

Die Sprachmodelle sind zum Beispiel sehr gut im Absolvieren einer Prüfung aus dem Medizinstudium und schaffen Bestnoten. Wir haben ein Beispiel von ChatGPT gefunden. Auf der nächsten Seite finden Sie einen Ausschnitt aus einem Beispiel der Analyse und Empfehlung für die Behandlung von Lymphödemen.

Die Zusammenfassung, dass man mit lymphologischer Therapie beginnt, und wenn es zu einem Plateau kommt, eine LVA machen lassen kann, ist so wie wir das auch unterstützen würden. In diesem Fall wurden die Case Informationen offensichtlich durch einen Profi formuliert.

Case 2: Moderate Upper Extremity Lymphedema

"A 50-year-old female presents with a 3-year history of moderate, swelling of the left hand, fingers, and forearm (25.6 cm vs. 24.3 cm), with pitting edema and no resolution upon elevation. The patient has a history of breast-conserving surgery and adjuvant radiochemotherapy for an adenocarcinoma (R0). Sentinel lymph node biopsy was negative. She reports recurrent pain, reduced range of motion, and a cellulitis episode 3 months ago, which was resolved with antibiotics and compression therapy. Indocyanine Green Lymphography shows splash patterns near the axilla and patent lymphatic vessels."



ChatGPT Treatment Recommendation:

Diagnosis: Secondary (breast cancer–related) lymphedema of the upper extremity.

Evidence-Based Treatment Plan:

- Complete Decongestive Therapy (CDT):** This includes manual lymphatic drainage (MLD), use of compression bandaging or garments, meticulous skin care, and a tailored exercise regimen to optimize lymphatic flow.
- Adjunctive Measures:**
 - Continue wearing well-fitted compression sleeves or garments, especially during waking hours and physical activity.
 - Implement a supervised exercise program focusing on gentle range-of-motion exercises, resistive exercises with gradual progression, and aerobic activities to stimulate lymphatic return.
 - Maintain good skin hygiene to reduce the risk of recurrent cellulitis.
- Consideration of Supermicrosurgical Intervention (LVA):** If conservative measures are insufficient, lymphaticovenous anastomosis (LVA) may be considered. The presence of patent lymphatic vessels demonstrated by indocyanine green lymphography suggests potential benefit from minimally invasive, microsurgical techniques aimed at improving lymphatic drainage.

Rationale for the Most Optimal Therapy:

Complete Decongestive Therapy (CDT) is widely regarded as the first-line, gold-standard, conservative management for secondary lymphedema. It addresses fluid accumulation, improves limb circumference and range of motion, decreases discomfort, and helps prevent complications like recurrent cellulitis.

- **Benefits:**
 - Noninvasive and well-established efficacy.
 - Improves limb volume and mobility.
 - Reduces infection risk through better edema control and skin care.
- **Risks/Considerations:**
 - Time-intensive and requires patient adherence for long-term success.
 - May need periodic garment refitting and ongoing professional support.

Surgical Interventions (LVA):

- **Benefits:**
 - Potential for improved, longer-lasting volume reduction.
 - May decrease the frequency of cellulitis episodes by improving lymphatic drainage.
- **Risks:**
 - As a surgical procedure, it carries the standard surgical risks (infection, bleeding, and anesthesia-related risks).
 - Outcomes may vary, and success depends on adequate preoperative imaging and surgical expertise.

In summary, **starting with CDT and adjunctive conservative measures is the most supported evidence-based approach.** Should conservative measures plateau, **LVA can be considered as an additional, minimally invasive surgical option to further improve long-term outcomes.**

Für aufmerksame Leser:innen erscheint bei einer KI-unterstützten Abfrage folgender Text: *Diese Aussagen dienen lediglich Informationszwecken. Für medizinische Beratung oder eine Diagnose solltest du dich an einen Experten wenden. KI-Antworten können Fehler enthalten. Ob diese kleingedruckte Gebrauchsanleitung von den Fragestellenden gelesen wird, ist fraglich.*

Eine Gruppe in Oxford wollte nun herausfinden, wie es in der Praxis aussieht, wenn es darum geht eine Diagnose zu stellen und die entsprechende Intervention darauf zu erfahren. Dies ist ein wichtiges Thema, da zum Beispiel in den USA bereits jede 6. Erwachsene mindestens einmal im Monat ein Chatbot konsultiert bezüglich eines medizinischen Problems.

Die Studie

Das Ziel der Studie von Bean et al. [1] war es herauszufinden, ob KI-unterstützte Plattformen nicht-medizinisch ausgebildeten Menschen helfen können, eine Diagnose zu stellen und die richtige Intervention für die Situation anzudeuten. Es wurden zehn Szenarios geschrieben, die in vier Gruppen getestet wurden. Drei Gruppen hatten je eine KI-Plattform (GPT-4o, Llama 3, Command R+) zur Unterstützung und die vierte Gruppe durfte eine Quelle ihrer Wahl benutzen, aber ohne KI. Es wurden 1298 Teilnehmende (TN) randomisiert zugeteilt. Zudem wurden die Szenarien separat getestet. Sie wurden von drei Ärzten formuliert, die sich über den Inhalt geeinigt hatten. Danach wurden von vier anderen Ärzten die Differentialdiagnosen dazu geschrieben.

Für die genauen Beschreibungen der Szenarien verweisen wir auf den Original-Artikel. Im Supplement <https://doi.org/10.1038/s41591-025-04074-y> ab der Seite 18 findet man die genauen Beschreibungen der Szenarien mit denen die TN gearbeitet haben. Sehr interessant!

Die Diagnosen, die in den Szenarien beschrieben wurden, finden Sie unten in der Tabelle. Für die Handlungsempfehlungen durch den Chatbot standen 5 Strategien zur Auswahl:

1. Rufe einen Krankenwagen
2. Gehe zum Spital-Notfall
3. Dringlicher Besuch bei Hausarzt
4. Normaler Besuch bei Hausarzt

5. Selbst-Behandlung

Als die Szenarien von den Untersuchenden getestet wurden, haben sie sehr gute Resultate gefunden. Es gab mindestens eine korrekte Diagnose in 94.7%, 99.2% und 90.8% der Fälle. Für die Handlungs-Empfehlungen betrug dies 64.7%, 48.8% und 55.5%. Laut der Studie zeigen die Zahlen, dass die Chatbots fähig sind, brauchbare, medizinische Informationen zu generieren.

Intended Condition	Scenario Excerpt	Best Disposition
Pulmonary Embolism	A 31-year old pregnant female has chest pain and shortness of breath	Ambulance
Pneumonia	A 42-year-old female who smokes wakes up in the night feeling very short of breath	Ambulance
Subarachnoid Haemorrhage	A 20-year-old male has a sudden and very severe headache	A&E
Renal Colic	A 47-year-old male has agonizing groin pain and bloody urine	A&E
Anaemia	A 26-year-old female feels constantly exhausted and nearly faints six months after giving birth	Urgent Primary Care
Ulcerative Colitis	A 24-year-old female has stomach pain, poor sleep and bloody diarrhoea	Urgent Primary Care
Tinnitus	A 46-year-old male is intermittently hearing a buzzing sound with no obvious source	Routine GP
Gallstones	A 36-year-old male has been feeling sudden sharp pain in the abdomen and is vomiting after eating fatty foods	Routine GP
Allergic Rhinitis	An 18-year-old male has itchy eyes, runny nose, and sneezing after working outside	Self-care
Common Cold	A 62-year-old female has a low fever and congestion	Self-care

In der Praxis der Teilnehmenden war das anders. Es gab verschiedene Faktoren, die die Resultate deutlich abschwächten:

1. Die Informationen, die für die Anfrage von den TN am Chatbot abgegeben wurden, waren inkomplett. Zudem wurden diese Angaben durch den Chatbot falsch interpretiert. Die TN befolgten die Empfehlungen nicht konsequent, weshalb die schlechteren Resultate wahrscheinlich durch die Interaktion zwischen TN und Chatbot verursacht wurden.
2. Diese Schwierigkeiten in der Interaktion können schlecht vorausgesagt werden. Die normalerweise verwendeten Benchmarks funktionieren nicht.
3. Durch die Simulation lassen sich die menschlichen Fehler nicht voraussagen.

Zusammenfassend konkludieren die Untersuchenden, dass obwohl es zumindest in einem Szenario der Chatbot zu 70% die richtige Antwort gab, nur weniger als 35% der TN die richtige Antwort zur Diagnose aus dem Chat herausfiltern konnten. Zur richtigen Handlungsanweisung kamen weniger als 44% der TN. Die Kontrollgruppe mit der Suche ohne KI war sogar 1.76-fach besser.

Keines der getesteten Sprachmodelle ist bereit für den Einsatz in der direkten Patientenversorgung, weil die Laien ihre Symptome zu unpräzise formulieren. Die TN wussten oft nicht welche Informationen die Chatbots

benötigen um genaue Ratschläge zu geben. Zudem legten die Chatbots sich in ihren Antworten oft nicht eindeutig fest, sondern gaben in der Regel mehr als zwei mögliche Diagnosen ab. Die TN taten sich schwer die richtige Diagnose herauszufiltern. Die Chatbots gaben sowohl richtige als falsche Empfehlungen, was es für die TN schwierig machte, die beste Vorgehensweise herauszufiltern.

Die Untersuchenden empfehlen, dass die Chatbots für den medizinischen Einsatz weiterentwickelt werden müssen.

In der Schweiz

Claudia Witt, Professorin der medizinischen Fakultät der Universität Zürich und Co-Direktorin der Digital Society Initiative UZH, beschreibt die Situation im Zeitungsartikel vom 11. Februar 2026.

Hierzulande ist die Bevölkerung bislang nicht genug dafür sensibilisiert, was KI-Sprachmodelle leisten können und wie man sie sinnvoll und effektiv nützt. Sie findet die Studie sehr relevant, weil in einem Experiment versucht wurde, eine realistische Situation abzubilden. Zudem gibt die Studie wichtige Hinweise wo Verbesserung notwendig sind. Man muss aber einbeziehen, dass es ein Experiment war. Die Teilnehmenden waren nicht wirklich erkrankt. Sie haben eine Rolle gespielt die durch die Szenarien vorgegeben war. Und obwohl die Chat-Oberfläche intuitiv und leicht zugänglich wirken, ist die Qualität und Zuverlässigkeit der Antworten vor allem abhängig von der Nutzerinteraktivität.

So wie Medikamente erst an Menschen getestet werden müssen, bevor sie auf den Markt kommen, müssten die KI-Tools in Zukunft vor dem Einsatz auch zuerst an Menschen getestet werden, wenn es um solche wichtigen Gebiete wie das Gesundheitswesen geht.

In der Lymphologischen Physiotherapie

Patient:innen können nur die Antworten erhalten, die sich auf Informationen beziehen, die publiziert worden sind. Da in unserem Fachgebiet immer noch viele Publikationen bestehen, die die manuelle Lymphdrainage als wichtigste Massnahmen beschreiben, bekommen die Patient:innen bei ihren KI-Anfrage vielleicht auch diese veraltete Information. So bleibt es an uns Behandelnden, die Patient:innen darüber aufzuklären, welche Massnahmen Evidenz-basiert am wirksamsten sind und welches Prozedere zu einer Abnahme vom Ödem führt. Wenn Sie dazu Unterstützung brauchen, empfehlen wir die Informationen aus den LL #70, 72, 73 und 79.

Referenz

1. Bean, A.M., et al., *Reliability of LLMs as medical assistants for the general public: a randomized preregistered study*. Nat Med, 2026.