

# **ChatGPT – SEIN oder nicht SEIN**

**„ChatGPT vielen Neurologen überlegen – KI stärkt neurologische Diagnosen“**

Die erste Überschrift („ChatGPT vielen Neurologen überlegen“) kann bei Fachleuten leicht den Eindruck erwecken, die KI solle den Arzt ersetzen und gefährde damit Arbeitsplätze oder die ärztliche Verantwortung. Die zweite Formulierung („KI stärkt neurologische Diagnosen“) betont dagegen die Rolle der KI als unterstützendes Werkzeug, das auf ein riesiges, ständig aktualisiertes Wissen zurückgreift und gerade in schwierigen Fällen neue Perspektiven eröffnen kann. Berücksichtigt man beide Sichtweisen, zeigt sich: Die KI ist dann ein wertvolles Hilfsmittel, wenn sie verantwortungsvoll genutzt wird und der Arzt ihre Vorschläge kritisch prüft. Wenn sie jedoch als Rechtfertigung für Diagnosen dient, ohne ärztliches Urteil einzubeziehen, kann das sowohl für Patienten als auch für Behandler und letztlich für das Vertrauen in das Gesundheitssystem zu einem ernsthaften Problem werden.

**„KI stärkt Diagnosen bei Polyneuropathie“**



Kann künstliche Intelligenz bei der Diagnose komplexer Nervenerkrankungen mithalten? Eine internationale Studie unter der Leitung von Dr. Alberto De Lorenzo vom Klinikum Humanitas in Mailand hat genau das untersucht – mit erstaunlichen Ergebnissen.

Im Test standen 100 echte, sorgfältig dokumentierte Fälle von Patientinnen und Patienten mit Polyneuropathien, darunter chronisch inflammatorische demyelinisierende Polyneuropathie (CIDP), das POEMS-Syndrom und die Charcot-Marie-Tooth-Krankheit. ChatGPT in der Version 4o analysierte die standardisierten Fallberichte und sollte jeweils drei differenzialdiagnostische Vorschläge inklusive Begründung machen.

Parallel beurteilten 36 Neurologinnen und Neurologen aus zehn Ländern die gleichen Fälle. Knapp die Hälfte der Ärztinnen und Ärzte war auf periphere Nervenerkrankungen spezialisiert, die andere Hälfte nicht.

Das Ergebnis: ChatGPT lag in 66 Prozent der Fälle mit der Hauptdiagnose richtig – deutlich besser als die nicht spezialisierten Fachkräfte (54 Prozent). Lediglich die erfahrenen Spezialistinnen und Spezialisten schnitten noch besser ab und erreichten eine Trefferquote von 74 Prozent.

Betrachtete man alle drei Vorschläge, enthielten 82 Prozent der KI-Antworten die korrekte Diagnose. Zum Vergleich: Die nicht spezialisierten Ärzte kamen auf 69 Prozent, die Experten auf etwa 90 Prozent.

Auch bei der Wahl der wichtigsten konfirmatorischen Tests überzeugte die KI: In 68 Prozent der Fälle lag sie richtig – ein Wert, der fast exakt dem der Spezialisten entsprach.

Allerdings zeigte sich auch, wo die Grenzen liegen. Die KI neigte dazu, Laborwerte zu überbewerten und klinische Details zu vernachlässigen. Immerhin: Nur in fünf Prozent der Antworten wurden Informationen frei erfunden – deutlich weniger als bei früheren Modellen.

Nach dem Vergleich mit den KI-Vorschlägen änderte etwa jeder fünfte nicht spezialisierte Neurologe seine Diagnose. Unter den Spezialisten war es nur jeder Zehnte.

Fachleute sehen in solchen Systemen vor allem eine Chance, wo spezialisierte Expertise fehlt. Richtig eingesetzt können KI-Modelle Diagnosen strukturieren, Tests gezielter auswählen und wertvolle Denkanstöße geben.

*„Es kommt auf die Prompts, das kritische Hinterfragen und die ärztliche Verantwortung an.“ so Siegels Fazit.*

Quelle: SpringeMedizin 26.06.2025 | [Periphere Neuritis](#) | Kongressbericht