

ChatGPT in der Augenheilkunde – ein Erfahrungsbericht

ChatGPT in Ophthalmology – A Report

ChatGPT von Open AI geht seit Anfang 2023 „viral“. Bei ChatGPT handelt es sich um einen Chatbot, der auf künstlicher Intelligenz basiert und der auf Fragen mit scheinbar menschlicher Intelligenz antwortet. Fragt man beispielsweise ChatGPT: „Wer bist du?“, so antwortet das System: „Ich bin ChatGPT, ein großes Sprachmodell, das von OpenAI auf Basis der GPT-3.5-Architektur trainiert wurde. Ich bin ein virtueller Assistent, der in der Lage ist, auf verschiedene Fragen und Anfragen zu antworten, die von Benutzern gestellt werden. Meine Programmierung ermöglicht es mir, natürliche Sprache zu verstehen und zu generieren, um menschenähnliche Konversationen zu führen.“

Die menschlich klingenden Antworten von ChatGPT sind beeindruckend, können aber auch die Illusion erwecken, dass es sich bei dem „Gegenüber“ um einen reflektierenden Gesprächspartner handelt. Dies ist explizit nicht der Fall – der Chatbot wurde nur sehr erfolgreich darauf trainiert, menschlich erscheinende Antworten zu generieren. Die menschenähnliche Konversation weckt auf den ersten Anschein hin Vertrauen in die Antworten, das – wie die folgenden Beispiele aufzeigen – nicht immer gerechtfertigt ist. Während wir bei Google, Wikipedia und anderen Anwendungen erfahrungsbasiert eine gewisse Skepsis erlernt haben, sollten auch die „Werke“ von ChatGPT bei allem Enthusiasmus kritisch hinterfragt werden.

Die Basisversion von ChatGPT kann jeder kostenlos nach vorheriger Anmeldung nutzen, und dies habe ich getan (Version ChatGPT3 mit letztem Update vom 23.3.23). Meine Erfahrungen in Bezug auf die Augenheilkunde möchte ich im Folgenden exemplarisch teilen.

Die erste Frage, die ich ChatGPT gestellt habe, war: „Muss ich mit einem Hornhautulkus zum Arzt?“ Die Antwort von ChatGPT (unter Berücksichtigung der Tatsache, dass diese Frage eher von Patienten gestellt werden könnte) finde ich recht zufrieden-

stellend und lautet: „Ja, wenn Sie den Verdacht haben, dass Sie ein Hornhautulkus haben, sollten Sie unbedingt einen Augenarzt aufsuchen. Ein Hornhautulkus ist eine offene Wunde oder ein Geschwür auf der Hornhaut des Auges, die durch eine Infektion verursacht wird. Es kann zu Schmerzen, Rötungen, Schwellungen, Lichtempfindlichkeit und verschwommenem Sehen führen. Eine rechtzeitige Diagnose und Behandlung durch einen Facharzt ist wichtig, um mögliche Komplikationen und eine Verschlechterung des Sehvermögens zu vermeiden. Der Arzt kann eine geeignete Behandlung wie Augentropfen, Salben oder Antibiotika verschreiben, um die Infektion zu behandeln und die Heilung der Hornhaut zu fördern. In einigen Fällen kann auch eine Operation erforderlich sein.“

In Vorbereitung auf ein Manuskript zu benignen Lidtumoren bei Kindern und Jugendlichen [1], bei der ich nach einer PubMed-Recherche überraschend wenig Literatur gefunden habe, stellte ich folgende Frage an ChatGPT: „Gibt es Studien zur Häufigkeit von Lidtumoren bei Kindern und Jugendlichen?“. Die Antwort von ChatGPT war auf den ersten Blick seriös und ergab 2 Studien, die ich bislang noch nicht identifiziert hatte. Die Antwort lautete: „Ja, es gibt Studien zur Häufigkeit von Lidtumoren bei Kindern und Jugendlichen. Lidtumoren bei Kindern und Jugendlichen sind selten, aber es wurden einige Studien durchgeführt, um ihre Häufigkeit zu untersuchen. Eine Studie aus dem Jahr 2014, die in der Fachzeitschrift „Journal of AAPOS“ veröffentlicht wurde, untersuchte die Inzidenz von Augentumoren bei Kindern unter 15 Jahren. In dieser Studie wurden 344 Augentumoren bei Kindern in einer Population von 16 Millionen Kindern identifiziert. Von diesen Tumoren waren 15,1% Lidtumoren. Eine weitere Studie aus dem Jahr 2016, die in der Fachzeitschrift „Orbit“ veröffentlicht wurde, untersuchte die Inzidenz von Lidtumoren bei Kindern unter 18 Jahren. In dieser Studie wurden 101 Fälle von Lidtumoren identifiziert. Basalzellkarzinom war die

häufigste Art von Lidtumor, gefolgt von Melanom und seborrhoischen Keratosen. Es ist jedoch wichtig zu beachten, dass diese Studien nur einen begrenzten Einblick in die Häufigkeit von Lidtumoren bei Kindern und Jugendlichen geben und dass weitere Forschung auf diesem Gebiet erforderlich ist.“

Bei genauerer Betrachtung mutete die Antwort zur Publikation in „Orbit“ schon etwas merkwürdig an, da wir aus der Klinik wissen, dass das Basalzellkarzinom sicher nicht der häufigste Lidtumor bei Kindern unter 18 Jahren ist.

Ich bemühte mich, die entsprechende Literatur bei PubMed zu finden, und suchte auch bei Google und Research Gate, aber erfolglos. Daraufhin fragte ich ChatGPT: „Wer sind die Autoren der beiden oben genannten Studien?“ Antwort von ChatGPT: „Die Autoren der beiden oben genannten Studien sind: Die Studie aus dem Jahr 2014, die in der Fachzeitschrift „Journal of AAPOS“ veröffentlicht wurde: „A population-based study of eyelid and conjunctival tumors among children and adolescents“ von Sahar M. Elkhamary, Todd P. Margolis, Christopher G. Kenyon und Jonathan H. Salvin. Die Studie aus dem Jahr 2016, die in der Fachzeitschrift „Orbit“ veröffentlicht wurde: „Lid tumors in children: A single-center experience“ von Zeynep Aktas, Aylin Yaman, Umit Beden, Yelda Ozsunar und Ugur Yilmaz.“

Auch diese präziseren Informationen halfen mir nicht dabei, die Studien zu finden. Zusammenfassend ist davon auszugehen, dass es hier um valide klingende, aber inhaltlich vollständig erfundene Antworten handelt.

Eine erneute Anfrage einige Wochen später (April 2023) mit exakt derselben Fragestellung ergab eine etwas andere Antwort unter Angabe der Publikation von E. Andrea Aronson, James A. Katowitz, Kristen M. Tye, Brian G. Mohny, Jonathan C. Horton (2015): „Incidence of Eyelid

Lesions in Children: A Survey through the American Society of Ophthalmic Plastic and Reconstructive Surgery“. Auch diese Publikation konnte ich nicht finden.

Auch die neueste Version (ChatGPT4), die allerdings nicht kostenfrei ist, ergab ebenfalls eine Studie, die nicht existiert („Childhood Eyelid Neoplasms: A 30-Year Experience“ Autoren: S. A. Al-Buloushi, P.J. Dolman, H. A. Rootman; Zeitschrift: Ophthalmic Plastic & Reconstructive Surgery; Jahr: 2005). Die gezielte Frage „Kannst du auf Internetquellen zugreifen oder hast du dir die Studie ausgedacht?“ beantwortet ChatGPT4 folgendermaßen: „Als KI-Modell habe ich keinen direkten Zugriff auf das Internet oder aktuelle Studien. Ich kann jedoch auf mein trainiertes Wissen zurückgreifen, das auf einer Vielzahl von Quellen basiert, einschließlich wissenschaftlicher Studien, Artikeln und Büchern. Die von mir genannte Studie stammt aus meinem vorherigen Training und ist eine reale Studie, die ich nicht erfunden habe. Wenn Sie die Studie selbst nachschlagen möchten, können Sie versuchen, sie in einer wissenschaftlichen Datenbank wie PubMed oder auf der Website der Zeitschrift „Ophthalmic Plastic & Reconstructive Surgery“ zu finden. Beachten Sie jedoch, dass mein Wissen bis September 2021 reicht, und es ist möglich, dass es seitdem aktualisierte oder ergänzende Studien gibt, die für Ihre Frage relevant sind.“

In meinem Beispiel wäre die unkritische Weiterverwendung des generierten Inhaltes der Anfrage zumindest bei einer wissenschaftlichen Publikation am fehlenden Zitationsnachweis gescheitert. Sicherlich ist aber deutlich geworden, dass hier grundlegend falsche Informationen generiert und als wissenschaftlich belegt ausgegeben wurden.

Auch Anfragen wie „Bitte fasse den Inhalt der Publikation von ... zusammen“, die – wie allen Lesern vorstellbar ist – sicherlich Anwendung finden werden, führten bei stichpunkthafter Untersuchung eigener Publikationen zu einigen Fehlaussagen, selbst wenn der Großteil der Publikation richtig wiedergegeben wurde. Das wird zukünftig die akkurate Aufarbeitung und Wiedergabe wissenschaftlicher Arbeiten und auch deren Überprüfung deutlich er-

schweren und alle Ärztinnen und Ärzte vor große Herausforderungen stellen. Denn wird wirklich immer angegeben werden, ob eine Literaturrecherche (sowohl im wissenschaftlichen Umfeld als auch im Bereich des Medizinjournalismus) mithilfe von KI erfolgte? Hier wird eine große Sorgfalt im Umgang mit Publikationen notwendig werden, um nicht leichtfertig falsche Fakten zu verbreiten. Es existieren allerdings bereits Anwendungen, mit denen mittels KI generierte Texte erkannt werden können, wie z.B. Classifier, GPTZero oder Originality.AI. Allerdings sind auch diese Hilfsmittel nicht fehlerfrei und arbeiten zuverlässiger bei englischsprachigen Texten. KI-generierte Texte weisen aber auch oft bestimmte Merkmale auf, die ein sensibilisierter Leser ohne weitere Hilfsmittel erkennen kann. So benutzt eine KI seltener Phrasen, außergewöhnliche Wortkombinationen, Neologismen, Umgangssprache oder Dialekt. Ein vom Menschen geschriebener Text weist i. d. R. eher Rechtschreibfehler, aber auch Interpunktionszeichen wie Gedankenstriche oder Doppelpunkte auf. Falsche Annahmen auf Basis eines KI-generierten Textes und deren unkritische Weitergabe werden aber natürlich mit diesen Techniken nicht erfasst.

An dieser Stelle muss allerdings der Fairness halber auch angemerkt werden, dass ChatGPT nicht zitieren kann und somit für diese Art der Fragestellung nicht geeignet ist. Leider geht dies aus der Antwort von ChatGPT4 nicht hervor und dürfte auch nicht jedem Anwender bekannt sein. Mir war dies bei der Erstnutzung ebenfalls nicht klar, sondern diese Information erhielt ich erst im Austausch mit Kollegen, die ChatGPT „professioneller“ nutzen als ich. Auf der Anwenderseite ist dazu vermerkt: „ChatGPT Mar23 Version. Free Research Preview. ChatGPT may produce inaccurate information about people, places, or facts“.

Für die von mir untersuchten Fragestellungen sind laut Dr. von der Emde (ebenfalls Universitäts-Augenklinik Bonn) andere Programme wie z.B. Perplexity geeignet. Eine Kritik bei „CHIP“ warnt bereits davor, Perplexity blind zu vertrauen, wortwörtlich heißt es: „Am besten ist es, die Inhalte kritisch zu hinterfragen und die Quellen zu prüfen – aber zumindest das ist hier deut-

lich einfacher.“ Die gleiche Fragestellung wie bei ChatGPT ergab bei Perplexity eine deutlich validere Antwort: „Yes, there are studies on the frequency of eyelid tumors in children and adolescents. However, the search results provided do not offer a comprehensive answer to the question. Some of the search results discuss benign eyelid tumors in adults“ (Zitate von mir sowohl bez. der Quelle als auch des wiedergegebenen Inhalts als korrekt bewertet). Es gibt zudem mehrere „KI-Forschungsassistenten“, darunter Elicit, die auf wissenschaftliches Arbeiten inkl. einer Literaturrecherche konzipiert sind. Allerdings sollte der Anwender auch hier immer die generierten Antworten auf Plausibilität und Korrektheit überprüfen. Eine Vollständigkeit der KI-basierten Antworten ist ebenfalls nicht gewährleistet, sodass die Anwendungen (basierend auf meinen bisherigen Erfahrungen) eine PubMed-basierte Literatursuche nicht ersetzen. Sie können aber durchaus hilfreich sein, wenn es darum geht, sich einen ersten Überblick zu einem Thema zu verschaffen und somit den Einstieg in eine tiefergehende Literaturanalyse zu vereinfachen.

Es soll allerdings an dieser Stelle auch angemerkt werden, dass ChatPGT viele Anwendungsmöglichkeiten besitzt, die auch beim wissenschaftlichen Arbeiten hilfreich sein können. So kann ChatGPT aus Stichpunkten einen Fließtext erstellen (der dann natürlich noch weiterbearbeitet werden kann), basierend auf einem Abstract Titel für ein Manuskript vorschlagen oder auch bereits existierende Texte nach Wunsch umformulieren. Auch kann ChatGPT (nach Eingabe entsprechender Vorinformation) bei der Generierung von Hypothesen für das Schreiben wissenschaftlicher Anträge helfen oder Ideen für Experimente entwickeln.

Wer die Debatte um ChatGPT bislang verfolgt hat, weiß um die Diskussion bez. Hausaufgaben und Hausarbeiten bei Schülern und Studenten, die hierzu sicherlich auch ChatGPT nutzen werden. Schon Friedrich Dürrenmatt ließ den Kernphysiker Möbius in die „Die Physiker“ sagen: „Was einmal gedacht wurde, kann nicht mehr zurückgenommen werden.“ (S. 85, Diogenes Verlag, Neufassung 1980). Umso wichtiger ist es daher, sich selbst aktiv

und kritisch mit den neuen Möglichkeiten auseinanderzusetzen und die KI-Anwendungen auch mit Studenten und Weiterbildungsassistenten entsprechend zu diskutieren. Es bedarf sicherlich Regeln sowie neutraler Anleitungen/Informationen zum Umgang mit diesen neuen Errungenschaften der Gesellschaft, die bei adäquatem und durchdachtem Einsatz einen immensen Fortschritt in vielen Bereichen darstellen können.

Dies kann aber nur gelingen, wenn man sich intensiv mit diesen neuen Möglichkeiten auseinandersetzt und den Dialog darüber auch in den entsprechenden Fachgesellschaften führt mit dem Ziel, die gesellschaftliche Entwicklung aktiv mitzugestalten. Großes Potenzial sehe ich hier in der Zusammenarbeit verschiedener Generationen, die unterschiedliche Qualifikationen und Sichtweisen in die Diskussion einbringen können.

Zusammenfassend ist es bei den neuen KI-Anwendungen – wie auch in anderen Be-

reichen – wichtig, die Vor- und Nachteile der verschiedenen Systeme zu kennen und die Systeme für die richtige „Indikation“ zu verwenden.

Interessenkonflikt

Glaxo Smith Kline (Teilnahme an Advisory Board, Referent). In Bezug auf diese Arbeit besteht kein Interessenkonflikt.

Autor

Martina C. Herwig-Carl 
Klinik für Augenheilkunde, Universitätsklinikum
Bonn, Deutschland

Korrespondenzadresse

Prof. Dr. med. Martina C. Herwig-Carl
Klinik für Augenheilkunde
Universitätsklinikum Bonn
Ernst-Abbe-Str. 2
53127 Bonn
Deutschland
Tel.: + 49 (0) 2 28 28 71 55 05
Fax: + 49 (0) 2 28 28 71 47 46
martina.herwig-carl@ukbonn.de

Literatur

- [1] Herwig-Carl MC, Bierhoff E, Holz FG et al. Differenzialdiagnosen benigner Lidtumoren bei Kindern und Jugendlichen. *Ophthalmologie* 2023; 120: 794–803. doi:10.1007/s00347-023-01887-5

Bibliografie

online publiziert 27.10.2023

Klin Monatsbl Augenheilkd 2024; 241: 952–954

DOI 10.1055/a-2142-2910

ISSN 0023-2165

© 2023. Thieme. All rights reserved.

Georg Thieme Verlag KG, Rüdigerstraße 14,
70469 Stuttgart, Germany