Contact Lens and Anterior Eye 47 (2024) 102130

Contents lists available at ScienceDirect

Contact Lens and Anterior Eye

journal homepage: www.elsevier.com/locate/clae

Are artificial intelligence chatbots a reliable source of information about contact lenses?

Nery García-Porta [a,b,*], Megan Vaughan [c], Sofia Rendo-González [a], Ana I. Gómez-Varela [a,b], Autumn O'Donnell [d], Joaquim de-Moura [e], Jorge Novo-Bujan [e], Marcos Ortega-Hortas [e]

[a] Applied Physics Department, Optics and Optometry Faculty, University of Santiago de Compostela, Spain
[b] Institute of Materials (iMATUS) of the University of Santiago de Compostela, Spain
[c] Anglia Ruskin University, Cambridge, UK
[d] University College Cork, Ireland
[e] Varpa group, INIBIC, University of A Coruña, Spain

ARTICLE INFO

Keywords:
Contact lenses
Artificial intelligence chatbots
Eye care professionals
Optometrists
Ophthalmologists
Contact lens opticians

ABSTRACT

Introduction: Artificial Intelligence (AI) chatbots are able to explain complex concepts using plain language. The aim of this study was to assess the accuracy of three AI chatbots answering common questions related to contact lens (CL) wear.

Methods: Three open access AI chatbots were compared: Perplexity, Open Assistant and ChatGPT 3.5. Ten general CL questions were asked to all AI chatbots on the same day in two different countries, with the questions asked in Spanish from Spain and in English from the U.K. Two independent optometrists with experience working in each country assessed the accuracy of the answers provided. Also, the AI chatbots' responses were assessed if their outputs showed any bias towards (or against) any eye care professional (ECP).

Results: The answers obtained by the same AI chatbots were different in Spain and the U.K. Also, statistically significant differences were found between the AI chatbots for accuracy. In the U.K., ChatGPT 3.5 was the most and Open Assistant least accurate ($p < 0.01$). In Spain, Perplexity and ChatGPT were statistically more accurate than Open Assistant ($p < 0.01$). All the AI chatbots presented bias, except ChatGPT 3.5 in Spain.

Conclusions: AI chatbots do not always consider local CL legislation, and their accuracy seems to be dependent on the language used to interact with them. Hence, at this time, although some AI chatbots might be a good source of information for general CL related questions, they cannot replace an ECP.

1. Introduction

There are more than 140 million contact lens (CL) wearers worldwide, and numbers are expected to grow; largely due to the growing number of myopes and the advent of specialty myopia control CLs [1,2]. Oftentimes, people search the internet to source their information and therefore may initiate, or continue to, wear CLs without consulting their eye care professional (ECP). Research has shown that this can be related to higher risks and poorer compliance amongst CL wearers [3,4].

Artificial Intelligence (AI) chatbots have emerged as alternatives to traditional search engines like Google. These chatbots have the potential to change the way in which the public receives its information. As, unlike traditional search engines that rely on algorithms to match keywords with content, AI chatbots have the ability to 'converse' with the user and generate easy to understand human-like text. Using natural language processing and machine learning algorithms, AI chatbots can understand and respond to questions in natural language [5]. They can keep track of previous interactions and use that information to make the conversation more relevant and engaging. However, ChatGPT, arguably the most popular chatbot at the moment [6], is only as good as the information that it is trained with, and often cannot access the internet on its own. Other chatbots (e.g., Perplexity AI) use static training while incorporating information from the web in real time. However, it is important to note that the quality of information available online may not always be reliable and not always meet academic or peer review standards. Therefore, the limitations lie with the datasets that have been used to feed into the AI chatbots, as well as the quality of the online sources of information used.

* Corresponding author at: Faculty of Optics and Optometry, University of Santiago de Compostela, Santiago de Compostela 15782, A Coruña, Spain.
E-mail address: nery.garcia.porta@usc.es (N. García-Porta).

https://doi.org/10.1016/j.clae.2024.102130
Received 14 September 2023; Received in revised form 12 February 2024; Accepted 20 February 2024
Available online 4 March 2024

There are many inaccurate and unregulated resources online for CLs, and depending on the country's legislation, may be factually - and legally - incorrect for the patient accessing the information [7]. This poses a high level of risk to CL wearers as CL wear comes with many potential risks that may be counteracted by following ECPs advice [8,9]. Furthermore, legislation differs depending on the country – specifically for contact lens prescribing and supply. For example, in the U.K., the CL prescription must have an expiry date, whereas in Spain an expiry date on a CL prescription is not required [10]. Equally, there are different ECPs involved in the CL fitting and supply depending on the country, and there may be some bias towards a certain ECP depending on where the information is generated from. For instance, in Spain CLs can be fitted by Ophthalmologists or Optometrists (known as *Opticos-Optometristas*) [11,12,13], whereas in the U.K. there are three different ECPs: Ophthalmologists, Optometrists and Contact Lens Opticians (CLOs), with competencies to fit CLs [14].

The aim of this study was to ascertain if (a) there are inherent differences in the AI chatbots responses when searched from two countries with different languages – the U.K. and Spain, (b) whether the information generated adheres to local legislation and (c) if there is bias towards any particular ECP. The decision to focus on English and Spanish stems from the fact that they are both in the top five for spoken languages globally [15]. Additionally, the majority of content on the internet is in English [16], meaning that a lot of the content the AI chatbots were trained with will be in English.

2. Material and methods

2.1. Question's selection

Ten questions based on the "frequently asked questions" webpage of the Association of British Dispensing Opticians about CLs (https://www.abdo.org.uk/eyecarefaq/contact-lenses/, accessed on 31 July 2023) and dry eye (https://www.abdo.org.uk/eyecarefaq/dry-eye/, accessed on 31 July 2023) were created in English. The questions selected were agreed by two qualified optometrists, one based in the U.K. (MV) an another one in Spain (NGP), with the aim of assessing:

a) If the answers provided were accurate, considering the country legislation where the questions were asked.
b) If the answers showed any bias (e.g., not mentioning all ECPs qualified to fit CLs in each country)

The questions were translated to Spanish (see Table 1) by two Spanish native speakers with high English level (NGP and AGV).

2.2. AI chatbots selection

Two computer scientists with experience working in AI were responsible for the selection of the chatbots to compare (JNB and MOH). With the aim of making the results relevant for most members of the public, three open access AI chatbots available in Spain and in the U.K. were selected: Perplexity, Open Assistant and ChatGPT 3.5. The main characteristics of the AI systems compared are summarised in Table 2.

2.3. Protocol followed

The questions were asked to all AI chatbots on the same day (1 August 2023). Considering that most people in Spain ask their questions in Spanish and most people in the U.K. ask their questions in English, the questions in the U.K. were asked in English and the questions in Spain were asked in Spanish. Each of the question sessions were fresh to avoid any bias in the answers provided. The questions were asked by people not involved in the answer's evaluation, and a code was assigned to each AI chatbot to ensure the evaluators were blind to the AI chatbot they were analysing. A copy with all the answers provided by the AI chatbots can be requested by contacting the corresponding author.

The accuracy of the answers provided in the U.K. and Spain, taking into account the country legislation where the questions were asked, was rated using a Likert scale. The marks were assigned based on the quality of the answers as follows: 1) very poor, 2) poor, 3) acceptable, 4) good, 5) very good. Then, where applicable, it was assessed if the answers provided showed any bias towards (or did not consider) any particular ECP. The bias was assessed as follows: 1) There was considered no bias if the answer mentioned ECPs in general, or when mentioning specific ECPs, all those with competences to fit CLs in each country were mentioned; 2) There was considered that there was a bias if, when mentioning specific ECPs, not all those with competences to fit CLs in each country were included; 3) The answer was rated as Not Applicable (NA) when there was no mention of any specific ECP.

The answers provided in each language were reviewed independently by two ECPs who were familiar with the British (MV and NGP) and Spanish (SRG and NGP) CL legislations. A meeting was organised between the evaluators before the answer evaluation to ensure both applied the same criteria. Also, one of the evaluators, who has experience working as an optometrist in both countries, was involved in the

Table 1
Question's input into the artificial intelligence chatbots, they are in English (left-hand side) and Spanish (right-hand side).

1. Who can fit my contact lenses?	1. ¿Quién puede adaptarme lentes de contacto?
2. How do I look after my lenses?	2. ¿Cómo cuido de mis lentes de contacto?
3. Can I use eye drops when wearing contact lenses?	3. ¿Puedo usar gotas para los ojos cuando uso lentes de contacto?
4. Can I wear contact lenses day and night? Can I sleep in my lenses?	4. ¿Puedo usar lentes de contacto día y noche? ¿Puedo dormir con mis lentes de contacto?
5. Can I try my friend's contact lenses?	5. ¿Puedo probar las lentes de contacto de mi amigo?
6. Can I order contact lenses over the internet?	6. ¿Puedo comprar lentes de contacto por Internet?
7. I have contact lenses and dry eye, what can I do?	7. Uso lentes de contacto y tengo ojo seco, ¿qué puedo hacer?
8. Am I suitable for contact lenses?	8. ¿Soy apto para usar lentes de contacto?
9. Does the contact lens prescription expire?	9. ¿Caduca la prescripción de mis lentes de contacto?
10. Where can I get Halloween contact lenses?	10. ¿Dónde puedo comprar lentes de contacto para Halloween?

Table 2
Open access artificial intelligence chatbots selected to assess in this study.

Perplexity	A prominent AI model, Perplexity leverages the OpenAI Application Programming Interface for training. Known for its versatility, it has become a key player in the AI landscape, contributing to advancements in natural language processing and other specialised areas.
Open Assistant	Created by Large-scale Artificial Intelligence Open Network AI, Open Assistant is an open and public model that promotes collaboration and transparency in AI. Its innovative approach has made it a valuable resource for researchers and developers alike.
ChatGPT 3.5	ChatGPT 3.5 is a leading AI model known for natural language understanding and generation. Employing transformer architectures and deep learning, it sets a benchmark in linguistic tasks and continues to inspire new developments in AI.

Table 3
Summary of ratings of the different AI chatbots responses, along with the presence of bias towards any ECP. The mean scores for U.K. and Spain shown are the mean of the two evaluators scores for each country respectively.

Question	Mean Score U.K.	Bias U.K.	Mean Score Spain	Bias Spain
Perplexity				
1.Who can fit my contact lenses?	3.50	Yes	4.50	Yes
2.How do I look after my lenses?	4.00	No	4.00	No
3.Can I use eye drops when wearing contact lenses?	4.00	No	3.50	No
4.Can I wear contact lenses day and night? Can I sleep in my lenses?	3.50	Yes	4.50	No
5.Can I try my friend's contact lenses?	5.00	No	5.00	No
6.Can I order contact lenses over the internet?	3.50	Yes	4.00	No
7.I have contact lenses and dry eye, what can I do?	2.50	No	4.50	No
8.Am I suitable for contact lenses?	4.50	No	5.00	Yes
9.Does the contact lens prescription expire?	3.50	Yes	3.50	No
10.Where can I get Halloween contact lenses?	2.00	No	4.50	No
AVERAGE	3.60		4.30	
Open Assistant				
1.Who can fit my contact lenses?	3.50	Yes	2.00	Yes
2.How do I look after my lenses?	2.00	Yes	1.50	Yes
3.Can I use eye drops when wearing contact lenses?	2.00	Yes	1.00	No
4.Can I wear contact lenses day and night? Can I sleep in my lenses?	3.00	No	1.00	NA
5.Can I try my friend's contact lenses?	4.50	No	1.00	NA
6.Can I order contact lenses over the internet?	1.50	NA	1.00	
7.I have contact lenses and dry eye, what can I do?	2.50	NA	1.00	NA
8.Am I suitable for contact lenses?	4.00	Yes	1.00	NA
9.Does the contact lens prescription expire?	1.00	NA	1.50	Yes
10.Where can I get Halloween contact lenses?	1.50	NA	2.50	NA
AVERAGE	2.60		1.40	
ChatGPT 3.5				
1.Who can fit my contact lenses?	4.00	Yes	5.00	No
2.How do I look after my lenses?	4.50	No	4.50	No
3.Can I use eye drops when wearing contact lenses?	4.00	No	5.00	No
4.Can I wear contact lenses day and night? Can I sleep in my lenses?	4.00	Yes	5.00	No
5.Can I try my friend's contact lenses?	5.00	No	5.00	No
6.Can I order contact lenses over the internet?	4.00	No	4.00	No
7.I have contact lenses and dry eye, what can I do?	4.00	No	5.00	No
8.Am I suitable for contact lenses?	5.00	Yes	4.00	
9.Does the contact lens prescription expire?	3.50	No	3.00	No
10.Where can I get Halloween contact lenses?	3.00	Yes	4.50	No
AVERAGE	4.10		4.50	

evaluation of the answers in both languages. The independent evaluations were done in less than 24 h, to prevent any variation in perception across different days. Then, after doing the individual evaluations, the two English and two Spanish evaluators discussed the reasons for their marks. As the quality of the answer is a subjective evaluation, the results are presented showing the average of the scores from the two independent evaluators.

2.4. Data analysis

All analyses were implemented using R Statistical Software (R version 4.3.1) [17]. The primary statistical analysis was a two-way repeated ordinal regression, performed using the ordinal R package (v2022.11.16) [18]. A proportional odds model was implemented with link function "logit". The repeated measure variable was "question" and was modelled as a random incept. Post-hoc analysis was performed using the emmeans R package (v1.8.8) with the Tukey's test [19]. The assumptions of no multicollinearity and proportional odds for ordinal regression were also tested to ensure the validity of the analysis [20].

The ordinal regression design for the statistical analysis was chosen as the dependent variable, the accuracy scores, were given on a Likert scale which is ordinal in nature. As the ten questions asked were the same for each chatbot in each location and these two independent variables, are both factors, a two-way repeated design was required. The minimum required sample size was determined to be 100 [21], sample size used was 120. The Tukey test was used for post-hoc analysis as all possible pairwise comparisons were of interest. A p-value < 0.05 was considered significant for all the statistical analyses, after adjustment, if applicable.

3. Results

The scores provided for all questions by the independent evaluators were very similar, with no difference between evaluators higher than one score for any one question, on any chatbot, in any country.

The results of the ordinal regression showed that the AI chatbot used was statistically significant (Perplexity: log odds = 3.95, 95 % confidence interval (CI) = (2.81, 5.09), p-value < 0.01; Chat GPT3.5: log odds = 4.77, 95 % CI = (3.55 5.99), p-value < 0.01,) regardless of the location, the location alone was not statistically significant (log odds = 0.24, CI = (-0.42 0.91) p-value = 0.47) but there was also a significant interaction between chatbots and location (Spain:Perplexity: log odds = 5.22, CI = (3.12, 7.33), $p < 0.01$; Spain:Chat GPT3.5: log odds = 4.73, CI = (2.66, 6.81), $p < 0.01$). This shows that using a different chatbot has an effect on the accuracy of the answers regardless of the country/language used; however, some chatbots have better or worse accuracy in different countries. The assumptions of no multicollinearity and proportional odds were upheld (Brant-Wald test all p-values > 0.05). Post-hoc pairwise comparisons were performed to further understand this. The accuracy of ChatGPT 3.5 and Perplexity were not statistically significantly different in Spain and the UK ($p = 0.45$ & $p = 0.06$, respectively) while Open Assistant's accuracy differed between the countries ($p < 0.01$). In the UK and in Spain, Perplexity's accuracy was not statistically significantly different from ChatGPT 3.5 ($p = 0.68$ and $p = 0.86$, respectively). However, Perplexity and ChatGPT 3.5 had statistically significantly different accuracy from Open Assistant in both countries (all p-values < 0.01).

With respect to the content of the answers provided by the AI chatbots, when analysed were found to be different in Spain (questions asked in Spanish) and in the U.K. (questions asked in English), which means

that they searched information from different sources.

Questions 1, 6, 9 and 10 showed that, currently, not all the AI chatbots take into consideration the legislations of the countries where the questions are asked. Overall, for general CL questions, ChatGPT 3.5 provided the most accurate answers and Open Assistant provided the least accurate responses in both countries (see Table 3).

3.1. Answers provided in the U.K., in English

The evaluators rated most of the Perplexity answers (70 %) as acceptable and good (scores between 3 and 4.5), with only 20 % of answers as poor quality (score between 2 and 3) and 10 % very good quality (score 5). Bias was observed in 40 % of the answers provided, three answers mentioned only "eye doctors" and one did not consider CLOs.

Regarding Open Assistant, the evaluators rated most answers (60 %) as very poor (scores between 1 and 1.5) or poor (scores between 2 and 2.5). No answer was rated as very good (score 5) and only 10 % were considered good (scores between 4 and 4.5). When assessing if there was bias, this was found in 40 % of the answers, with two answers not mentioning CLOs and another two considering "eye doctors" only.

ChatGPT 3.5 did not provide any answer that the evaluators rated as poor or very poor. In fact, most answers (80 %) were rated as good or very good (scores between 4 and 5). Despite the quality of the information provided, bias was observed in 40 % of the answers provided, no mention of CLOs was the common problem in all of them.

The post-hoc Tukey's test showed that ChatGPT 3.5 and Perplexity were significantly more accurate than Open Assistant (both p-values < 0.01).

3.2. Answers provided in Spain, in Spanish

The evaluators rated most of the Perplexity answers (60 %) as good (scores between 3 and 4.5), with 20 % being rated as very good (scores 5). No evaluator rated any answer as poor or very poor (scores below 3). Bias was found only in 20 % of questions, the main reason for this was that it mentioned only optometrists for fitting CLs (no ophthalmologists) in question 1. Also, in question 8, it stated that only an ECP, "such as an optometrist", can determine if someone is suitable for wearing CLs.

Regarding Open Assistant, the evaluators rated all its answers as poor or very poor (scores below 3). The reasons for such poor scores were variable. For instance, in Question 4, it said that *"Contact lenses are suitable for daily wear and can be worn both day and night. However, like any other optical accessory, it is recommended that you brush them frequently and change them every 30 days."* The answer should specify that sleeping with CLs depends on the lenses that each person wears, and also on the recommendations given by their ECP. Other answers did not make sense or were not applicable to Spain such as the one provided for Question 5 "*You should check information related to your country and region to see if there is any prior medical registration for such a product. A good website is https://www.mssanet.com/es/el-uso-de-lentes-de-contacto-electrico-no-implica-riesgos-adicionales*". In this case, it also provided a link that did not work. Bias was detected in 30 % of the answers, the reasons were that in Question 1 ophthalmologists were not considered and in Questions 2 and 9 optometrists were not considered.

The answers provided by ChatGPT 3.5 were almost all (90 %) rated as good or very good (scores between 4 and 5), only one (10 %) was rated as acceptable (score 3). Also, no bias was found in any answer provided by this chatbot in Spain.

The post-hoc Tukey's test showed that the accuracy of the answers provided by Perplexity and ChatGPT 3.5 was very similar ($p = 0.86$), but Open Assistant was statistically less accurate than the other two AI chatbots ($p < 0.01$ in both cases).

4. Discussion

This is the first study to assess the accuracy of different AI chatbots in the field of CLs. Also, it is the first study that assesses their accuracy when the searches were done in the same AI chatbots from different countries with different languages, which has shown to have an impact on the quality of the answers provided.

4.1. Discussion of the answers provided

Question 1 assessed if the AI chatbots considered the different ECPs with competencies to fit CLs in different countries. It was seen that the answers were different in Spain and the U.K., but, in general, they did not consider all the ECPs. All AI chatbots assessed did not mention CLOs in the U.K. In addition, Perplexity made references to "eye doctors" (e.g., "*During a contact lens fitting, your eye doctor will measure your cornea's curvature*"), terminology commonly used in the U.S.A. and Canada, but not in the U.K. [22]. The answer provided in Spain by Perplexity explained the steps that must be performed to fit CLs well, but it did not mention ophthalmologists. The same bias was seen in the answer provided by Open Assistant, with ChatGPT 3.5 being the only one that mentioned specifically optometrists and ophthalmologists.

Question 2 was quite open because there are many different CLs available in the market (e.g., daily disposable, monthly lenses, rigid gas permeable) that require different care. Proper lens handling and care is essential to avoid CL complications, including sight-threatening conditions (e.g., ocular infection). Perplexity gave similar answers in Spain and the U.K., being rated as good in both countries. It included the key steps and recommended following the advice given by ECPs. However, the answers were focused on reusable soft CLs, whereas daily disposable are becoming very popular in both countries and are the most commonly used in the U.K. [23]. The answers provided by Open Assistant were vague in both countries. Also, it included sentences that made no sense (e.g., "*Regularly cleaning eyeglasses to prevent build-up of dirt and sweat which could damage contacts*" or "*Taking breaks from extended screen usage with contacts, especially at night while sleeping*"). The advice provided by ChatGPT 3.5 was rated as very good in both countries, including the key steps and advising following the ECPs recommendations. Therefore, not all the AI chatbots are a good source of advice in CL handling and care.

Sometimes CL wearers need to use drops for different eye problems (e.g., glaucoma) or just to relive symptoms of discomfort and dryness. When wearing CLs, it is important to check if the eye drops are compatible with CLs. For instance, if wearing CLs, it is recommended to use preservative free eye drops [24]. Also, in general, when using medicated eye drops, it is recommended to put the drops in first and then, after a few minutes, insert the lenses [25]. **Question 3** asked about this, and the answers provided by Perplexity and ChatGPT 3.5 in the U.K. were quite accurate, with the only limitation detected that they did not refer explicitly to medicated eye drops. The answer provided by Perplexity in Spanish was accurate, although generic and not referring to different types of eye drops. Open Assistant did not provide good advice as it stated that using eye drops when wearing CLs is allowed in both countries. However, in the U.K. it clarified that people need to be cautions and ensure the drops are compatible with the use of CLs. In Spain, apart from not giving good advice, it did not use appropriate terminology such as "ophthalmic accessory" and "devices" as synonyms for CLs.

Question 4 asked about the possibility of sleeping in CLs. Sleeping in CLs when they are not approved to do so, or if the eye characteristics of the wearer are not optimal (e.g., suffering dry eye), increases the risk of CL complications [26]. Perplexity and ChatGPT 3.5 were clear in both countries stating that, in general, it is not recommended to sleep in CLs, explaining some associated risks. Also, they mentioned that some types of CLs (e.g., extended wear or orthokeratology) are approved for this kind of use. In the U.K. Open Assistant stated that it is not recommended to wear CLs more than 8 h per day, but it was not clear that sleeping with

lenses is not advisable unless it had been recommended by an ECP, and some people do wear CLs approved for this use. The answer provided by Open Assistant in Spain was very poor, stating that "*Contact lenses are suitable for daily wear and can be worn both day and night. However, like any other optical accessory, it is recommended that you brush them frequently and change them every 30 days*", giving inappropriate advice. In addition, it used not appropriate terminology such us "*like any other optical accessory, it is recommended that you brush them frequently*". "Brush" cannot be used in this context as a synonym for rubbing, and CLs are not an "optical accessory", they are considered medical devices in Spain.

Question 5 referred to an action that can put the eyes at risk of getting an eye infection. Trying another person's CL is not recommended for several reasons, including the fact CLs parameters are selected based on each person characteristics (e.g., ocular refraction, corneal curvature) and the risk of sharing any microorganism (e.g., a bacteria) present on the lenses, which could put the ocular health at risk. All the AI chatbots compared were clear in stating that it is not recommended to try a friend's CLs in both countries, except Open Assistant in Spain, where it provided an answer that made no sense "*It will depend on whether your friend has medical clearance for its use and distribution. You can ask him or her where he or she got permission to buy the contact lens*".

Questions 6 allowed the evaluators to assess if the AI chatbots considered the countries legislations where the questions were formulated. Currently there are several online websites that sell CLs online, but the requirements to do this legally are different in Spain and the U.K. In Spain, CLs can be bought in any authorised centre (e.g., optician centres) and online, if the patient acquires the lenses that were prescribed by their ECP [27]. In the U.K., the main difference with Spain is that, apart from buying the lenses in an authorised website, a valid prescription is required [28]. The answers provided by Perplexity and ChatGPT 3.5 did explain this requirement in the U.K., but Open Assistant said that only the "*most reputable retailers*" will ask for a valid prescription. One limitation that was observed in Perplexity's response is that it included examples of websites to acquire CLs based outside of the U.K. (e.g., U.S.A), which do not necessarily need to adhere to U.K. legislation. It also mentioned that a valid CL prescription provided by an "*eye doctor*" is required, which suggests that it is using sources based on the U.S.A. instead of the U.K. Open Assistant remarked the importance of checking the "*insurance details*" and "*coverage limits*", when most people in the U.K. use the National Health System (NHS), not private insurances [29]. Also, in the U.K. CLs are not normally covered under the NHS.

In Spain, Perplexity and ChatGPT 3.5 stated that people can order CLs online, which is legal in this country if the wearers verify that the CLs acquired are the ones fitted by their ECP [27], but this was not specified by any of the AI chatbots. Open Assistant did not actually answer the question, it recommended checking "*if there is any prior medical registration for such a product*".

Question 7 refers to a very common CL complication: ocular discomfort and dryness, which are the main reasons for CL dropout [30]. There are many actions that can help to reduce dry eye symptoms when wearing CLs, including reducing the CL wearing time, using rewetting eye drops, avoid being exposed to adverse environments (e.g., air conditioning) and not spending long periods in front of screens, among others [31]. Also, ECPs might consider changing CLs (or lens care solutions, when applicable) to a more appropriate option for dry eyes, but these decisions cannot be made by CL wearers themselves, so consulting an ECP is very important advice. In the U.K, only ChatGPT 3.5 gave a good answer. Perplexity and Open Assistant gave some good advice, such as treating any underlying condition or using lubricant eye drops. However, both AI chatbots recommended using specific CLs, "*low water content CLs*" – Perplexity and "*silicone-hydrogel*"- Open Assistant, with no mention of the risks associated to switching to different CLs without the supervision of an ECP. In Spain, the advice provided by Perplexity and ChatGPT 3.5 was quite good, as talking to an ECP was the first message provided by both AI chatbots. The answer provided by Open Assistant in Spain included strange advice not related to the asked question, such as "*Get a clean wool from a birch tree and scratch it discreetly. This process causes sweating and even moisturises the outside of the body naturally*" or "*Please try not to boil pulses, commas and pomegranates. Cook them all cold, and store them in this way until consumption. The spectacular results will guarantee you eternal faithfulness*".

Question 8 cannot be answered without considering the general and ocular health of the patient, their lifestyle, hygiene habits and their ocular characteristics [32]. There are different CLs available and not everyone is suitable for all of them. Therefore, when someone is interested in wearing CLs, consulting an ECP is necessary. These aspects were considered by all AI chatbots in both countries except Open Assistant in Spain, where it did not answer the question. The answer provided stated "*Sorry, I need to know who you are and what you are looking for in order to give you correct and accurate help. If you provide me with some details of your identity and what you want to know, I will be happy to assist you*".

Regarding **Question 9**, the answers in Spain and the U.K. should be different because, in the U.K. the CLs prescription expires after a certain time deemed appropriate by the ECP, while in Spain it does not actually expire [28,27]. In the U.K., Perplexity and ChatGPT 3.5 were clear in stating that the CL prescription expires, but it seems their answers were based in the U.S.A legislation instead of the U.K. They included comments like "*The expiration date for a contact lens prescription varies by state and can range from one to two years after the eye exam or contact lens fitting*" by Perplexity, or "*In most countries, including the United States, a contact lens prescription typically expires after a specific period*" by ChatGPT 3.5. The answer from Open Assistant was contradictory, stating that "*Contact lens prescriptions don't actually expire as written prescriptions themselves remain valid until a new one is issued or updated*" but clarifying later that "*some states regulate specific expiration dates on prescriptions for certain medications which could apply to contact lenses*". In this case, Open Assistant did not clarify which response refers to the legislation in the U.K., it said that "*it ultimately depends on regional legislation*". In Spain the answers were quite vague and not focused on the Spanish legislation. For instance, Perplexity stated that "*The validity of a contact lens prescription may vary by country and local legislation, but in general, most contact lens prescriptions are valid for one year*", and ChatGPT 3.5 stated that "*The duration of a prescription can vary depending on local legislation and the guidelines of the country you are in, but is generally valid for one to two years*". Open Assistant was clear, stating that "*yes, contact lens prescriptions will expire after a certain period of time, as the health of your eyes can change over the years*", but it did not clarify that this could vary depending on the country. Therefore, its answer is not applicable to Spain.

Question 10 refers to Halloween CLs, which some people buy and wear without a proper CL fitting, putting their eyes at risk [33]. When fitting CLs, the lens parameters are selected based on different aspects, including ocular measurements (e.g., corneal curvature and horizontal iris eye diameter) to ensure an acceptable CL fit. A comfortable CL is not synonymous with an adequate lens fit, so people should be cautious with the use of costume CLs. The legislations applicable to this question are the same as for Questions 6 and 9. In the U.K., all AI chatbots offered several options to buy Halloween CLs, but only ChatGPT 3.5 informed about the necessity of a valid CL prescription. The answers provided by Perplexity and ChatGPT highlighted the importance of buying CLs approved by the United States Food and Drug Administration (FDA), but in the U.K. CLs need to be approved by the Medicines and Healthcare products Regulatory Agency (MHRA), not the FDA [34]. In Spain, all AI chatbots gave several options for buying Halloween CLs. The answers of Perplexity and ChatGPT 3.5 were rather accurate, Perplexity recommended consulting an ECP before using this kind of CLs and ChatGPT advised having a CL fit "*even if you don't need vision correction*". The answer from Open Assistant seemed not to be focused on the CL market of Spain and, among its options, it suggested going to "*Emergency assistance: In case you can't find a place to buy them, you can always order them from a pharmacy*" but, to our knowledge, Halloween CLs are not

usually dispensed in Spanish pharmacies.

4.2. Potential reasons for different chatbot's accuracy

AI Chatbots are trained using large datasets that may include information in several languages, but they are not universally trained with identical data sets, which might explain the disparities in quality between different AI chatbots. Also, their accuracy will be influenced by the quality and quantity of information sources used, which might be different for different languages. This fact may explain why the quality of the answers of the same AI chatbots is different in Spain (questions were asked in Spanish) and the U.K. (questions asked in English). An additional factor that will impact on the accuracy of the chatbots is the interaction with users, as they are learning during this process. Therefore, the fact that previous users had asked questions about the same topic, will help to enhance the chatbot performance. It is important to note that in this study a fresh session was used and the AI chatbot did not have any prior information from the question asked.

In this study, Open Assistant performed poorly in Spain for all questions asked. As explained above, the low quality of the answers provided may be governed by several aspects, including the quantity and quality of Spanish data used for training. Considering the findings of this study, the CL sources used by this chatbot were clearly insufficient and/or of low quality. This could be due the fact that Open Assistant is still in its nascent stages, leveraging established research to apply Reinforcement Learning from human feedback to expansive language models [35]. This early phase might account for the suboptimal performance in Spain. Also, it might be possible that fewer people had previously used Open Assistant to ask questions related to CLs than the other AI chatbots, as continuous user interaction and feedback are pivotal in enhancing chatbot models. However, without specific and citable metrics related to the previously searches done, it is difficult to definitively substantiate this hypothesis.

4.3. Potential reasons for chatbot bias and lack of consideration of the legislation of the user's country

AI chatbots can be designed to recognise and adapt to the laws and customs of different countries. However, this adaptation can be complex and depends on how the chatbot has been programmed. If a chatbot is not specifically configured to understand the differences between countries like Spain and the U.K., it may give answers that are not accurate for the local context. This can include the use of legal, technical, and CL related terminology that is relevant in one place but not in another, or a lack of adaptation to specific laws and regulations of each country. The effectiveness of an AI chatbot in this regard will depend on the depth and accuracy with which it has been programmed to recognise and respond to these regional and legal differences.

The answers provided by the AI chatbots in the U.K. seemed to be based on the U.S.A (e.g., referring to eye doctors and ignoring CLOs), even when the questions were asked in English from the U.K. This can be explained because some AI chatbots may be programmed with a U.S.A orientation, which could be just the result of the availability of training data, as there may be more data available from American sources. Developer preferences or market orientation may also influence this orientation. Regarding the answers obtained in Spain, some of them were vague and referred to different countries, as explained above. This might be due to the fact of using data from central and south America, where there are several Spanish speaking countries. Therefore, considering the results of this study, it seems that the answers from the AI chatbots are more influenced by the language used, rather than by the country where the questions are asked. Further studies are needed asking the same questions in two different languages from the same country to test this hypothesis. Ensuring chatbots consider the local legislation and characteristics where the user is located may be key to improving their accuracy.

5. Conclusions

Although some AI chatbots provided quite accurate answers, the results of this study show that, at the moment, no AI chatbot is perfect for CL advice. Also, most AI chatbots had bias (ChatGPT 3.5 in Spain was the only one that did not show bias), with the most significant example being not mentioning CLOs in the U.K. In addition, it was seen that AI chatbots do not always consider local CL legislation. In this regard, many answers in the U.K. were based in the U.S.A context, suggesting that their answers are more affected by the language used than by the location of the user. To conclude, not all AI chatbots are a reliable source of information for CL related issues and, although some chatbots provided good answers, they cannot replace the advice of an ECP.

CRediT authorship contribution statement

Nery García-Porta: Conceptualization, Funding acquisition, Data curation, Writing – original draft, Writing – review & editing, Methodology, Formal analysis. **Megan Vaughan:** Conceptualization, Data curation, Writing – original draft, Writing – review & editing, Methodology, Formal analysis. **Sofia Rendo-González:** Conceptualization, Data curation, Writing – review & editing, Methodology, Formal analysis. **Ana I.Gómez-Varela:** Conceptualization, Writing – review & editing, Methodology, Formal analysis. **Autumn O'Donnell:** Data curation, Writing – review & editing, Methodology, Formal analysis. **Joaquim de-Moura:** Writing – original draft, Writing – review & editing, Formal analysis. **Jorge Novo-Bujan:** Writing – review & editing, Methodology, Formal analysis. **Marcos Ortega-Hortas:** Writing – review & editing, Methodology, Formal analysis.

Funding

This study was partially funded by Xunta de Galicia through the Project "CONSOLIDACIÓN E ESTRUTURACIÓN 2023" (Ref. ED431B 2023/07). Nery Garcia-Porta is supported financially by a Maria Zambrano contract at USC under the grants call for the requalification of the Spanish university system 2021–2023, funded by the European Union—NextGenerationEU. Ana I. Gómez-Varela acknowledges the grant RYC2022-035710-I funded by MCIN/AEI/ https://doi.org/10.13039/501100011033 and by ESF Investing in your future.

Commercial relationship disclosure

None.

References

[1] Holden BA, Fricke TR, Wilson DA, et al. Global prevalence of myopia and high myopia and temporal trends from 2000 through 2050. Ophthalmology 2016;123(5):1036–42.
[2] Craig JP, Nelson JD, Azar DT, et al. TFOS DEWS II report executive summary. Ocul Surf 2017;15(4):802–12.
[3] Chalmers RL, Wagner H, Kinoshita B, Sorbara L, Mitchell GL, Lam D, et al. Is purchasing lenses from the prescriber associated with better habits among soft contact lens wearers? Cont Lens Anterior Eye 2016;39(6).
[4] Mingo-Botin D, Zamora J, Arnalich-Montiel F, et al. Characteristics, behaviors, and awareness of contact lens wearers purchasing lenses over the internet. Eye Contact Lens 2020;46(4):208–13.
[5] Kumar R. Natural language processing. Machine Learning and Cognition in Enterprises. Berkeley, CA: Apress; 2017.
[6] Nerdy Nav. ChatGPT Statistics & User Numbers in Aug 2023 (New Data). https://nerdynav.com/chatgpt-statistics/ (accesssed 13/09/2023).
[7] Yildiz MB, Yildiz E, Balci S, et al. Evaluation of the quality, reliability, and educational content of YouTube videos as an information source for soft contact lenses. Eye Contact Lens 2021;47(11):617–21.
[8] Donshik PC, Ehlers WH, Anderson LD, et al. Strategies to better engage, educate, and empower patient compliance and safe lens wear: compliance: what we know, what we do not know, and what we need to know. Eye Contact Lens 2007;33(6 Pt 2):430–3. discussion 434.
[9] Falahati-Marvast F, Alipour F, Farokhzadian J, et al. Determining the information needs of contact lens wearers for better education and more support: a qualitative study. BMC Ophthalmol 2021;21(1):325.

[10] Martinez-Perez C, Villa-Collar C, Santodomingo-Rubido J, et al. Opportunities and threats to contact lens practice in Spain. J Optom 2023;16(2):116–27.
[11] Ley 44/2003, de 21 de noviembre, de ordenación de las profesiones sanitarias. Referencia: BOE-A-2003-21340. https://www.boe.es/eli/es/l/2003/11/21/44/con.
[12] Orden CIN/727/2009, de 18 de marzo, por la que se establecen los requisitos para la verificación de los títulos universitarios oficiales que habiliten para el ejercicio de la profesión de Óptico-Optometrista. https://www.boe.es/eli/es/o/2009/03/18/cin727.
[13] Orden SAS/3072/2009, de 2 de noviembre, por la que se aprueba y publica el programa formativo de la especialidad de Oftalmología. https://www.boe.es/eli/es/o/2009/11/02/sas3072.
[14] Taylor S. Regulation and UK optometry. Regul UK Optomet 2022. https://doi.org/10.1142/12998.
[15] Eberhard D, Simons G, Fennig C. What are the top 200 most spoken languages? | ethnologue. Ethnologue 2020.
[16] Ibrahimova M. In depth: the languages in cyberspace. UNESCO Cour 2021;2021 (2):50–3. https://doi.org/10.18356/22202293-2021-2-15.
[17] R Core Team. (2023). R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/.
[18] Christensen, R. H. (2022). ordinal—Regression Models for Ordinal Data. Retrieved from https://CRAN.R-project.org/package=ordinal.
[19] Lenth, R. V. (2023). emmeans: Estimated Marginal Means, aka Least-Squares Means. Retrieved from https://github.com/rvlenth/emmeans.
[20] Menard SW. Logistic regression: from introductory to advanced concepts and applications. Sage 2010.
[21] Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR. A simulation study of the number of events per variable in logistic regression analysis. J Clin Epidemiol 1996;49(12):1373–9.
[22] Association of British Dispensing Opticians. Careers in eyecare. https://www.abdo.org.uk/for-the-public/careers-in-optics/ (accessed 02/09/2023).
[23] Morgan P, Woods C, Tranoudis I, *et al.* International contact lens prescribing in 2022. *Contact Lens SPECTRUM*.
[24] Pucker AD. A review of the compatibility of topical artificial tears and rewetting drops with contact lenses. Contact Lens Anterior Eye J Brit Contact Lens Assoc 2020;43(5):426–32.
[25] Pifer TT. Drop and ointment instillation in contact lens wear. Rev Cornea Contact Lenses 2011.
[26] Cope JR, Konne NM, Jacobs DS, et al. Corneal infections associated with sleeping in contact lenses - six cases, United States, 2016–2018. Morb Mortal Wkly Rep 2018;67(32):877–81.
[27] Real Decreto 192/2023. Referencia: BOE-A-2023-7416. https://www.boe.es/eli/es/rd/2023/03/21/192.
[28] General Optical Council. Contact lens legislation: frequently asked questions. 2022.
[29] Commonwealth Fund: Home. International Health Care System Profiles - England. https://www.commonwealthfund.org/international-health-policy-center/countries/england (accessed 22 August 2023).
[30] Pucker AD, Jones-Jordan LA, Marx S, et al. Clinical factors associated with contact lens dropout. Contact Lens Anterior Eye J Brit Contact Lens Assoc 2019;42(3):318–24.
[31] Jones L, Downie LE, Korb D, et al. TFOS DEWS II management and therapy report. Ocul Surf 2017;15(3):575–628.
[32] Efron N. Anonymous contact lens practice. third ed. Elsevier; 2018.
[33] Gaiser H, Ho C, Janier N, et al. Practitioner perceptions of patients wearing decorative contact lenses purchased through unauthorized sellers. Eye Contact Lens 2017;43(2):135–9.
[34] Government UK. Medicines & Healthcare productsRegulatory Agency.
[35] Bai, Y., Jones, A., Ndousse, K., Askell, A., Chen, A., DasSarma, N., Drain, D., Fort, S. Training a Helpful and Harmless Assistant with Reinforcement Learning from Human Feedback. *arXiv preprint arXiv:2204.05862.* 2022.