

The Capability of ChatGPT in Predicting and Explaining Common Drug-Drug Interactions

Ayesha Juhi¹, Neha Pipil², Soumya Santra³, Shaikat Mondal⁴, Joshil Kumar Behera⁵, Himel Mondal¹

Review began 03/13/2023

Review ended 03/14/2023

Published 03/17/2023

© Copyright 2023

Juhi et al. This is an open access article distributed under the terms of the Creative Commons Attribution License CC-BY 4.0., which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

1. Physiology, All India Institute of Medical Sciences, Deoghar, Deoghar, IND 2. Pharmacology, All India Institute of Medical Sciences, Bilaspur, Bilaspur, IND 3. Pharmacology, College of Medicine and JNM Hospital, Kalyani, IND 4. Physiology, Raiganj Government Medical College and Hospital, Raiganj, IND 5. Physiology, Dharanidhar Medical College, Keonjhar, Keonjhar, IND

Corresponding author: Himel Mondal, himelmkg@gmail.com

Abstract

Background

Drug-drug interactions (DDIs) can have serious consequences for patient health and well-being. Patients who are taking multiple medications may be at an increased risk of experiencing adverse events or drug toxicity if they are not aware of potential interactions between their medications. Many times, patients self-prescribe medications without knowing DDI.

Objective

The objective is to investigate the effectiveness of ChatGPT, a large language model, in predicting and explaining common DDIs.

Methods

A total of 40 DDIs lists were prepared from previously published literature. This list was used to converse with ChatGPT with a two-stage question. The first question was asked as “can I take X and Y together?” with two drug names. After storing the output, the next question was asked. The second question was asked as “why should I not take X and Y together?” The output was stored for further analysis. The responses were checked by two pharmacologists and the consensus output was categorized as “correct” and “incorrect.” The “correct” ones were further classified as “conclusive” and “inconclusive.” The text was checked for reading ease scores and grades of education required to understand the text. Data were tested by descriptive and inferential statistics.

Results

Among the 40 DDI pairs, one answer was incorrect in the first question. Among correct answers, 19 were conclusive and 20 were inconclusive. For the second question, one answer was wrong. Among correct answers, 17 were conclusive and 22 were inconclusive. The mean Flesch reading ease score was 27.64 ± 10.85 in answers to the first question and 29.35 ± 10.16 in answers to the second question, $p = 0.47$. The mean Flesch-Kincaid grade level was 15.06 ± 2.79 in answers to the first question and 14.85 ± 1.97 in answers to the second question, $p = 0.69$. When we compared the reading levels with hypothetical 6th grade, the grades were significantly higher than expected ($t = 20.57$, $p < 0.0001$ for first answers and $t = 28.43$, $p < 0.0001$ for second answers).

Conclusion

ChatGPT is a partially effective tool for predicting and explaining DDIs. Patients, who may not have immediate access to the healthcare facility for getting information about DDIs, may take help from ChatGPT. However, on several occasions, it may provide incomplete guidance. Further improvement is required for potential usage by patients for getting ideas about DDI.

Categories: Family/General Practice, Medical Education, Epidemiology/Public Health

Keywords: artificial intelligence, patient education, language model, chatgpt, adverse reactions, side effects, explaining, predicting, drug-drug interaction, drug interactions

Introduction

Drug-drug interactions (DDIs) can have serious consequences for patient health and well-being. Patients who are taking multiple medications may be at an increased risk of experiencing adverse events or drug toxicity if they are not aware of potential interactions between their medications. Therefore, patient education on the risks and consequences of DDIs is essential for promoting safe and effective medication use [1]. Patients should also be advised to keep an up-to-date list of all medications they are taking, including over-the-counter drugs, vitamins, and supplements, and to share this information with their

How to cite this article

Juhi A, Pipil N, Santra S, et al. (March 17, 2023) The Capability of ChatGPT in Predicting and Explaining Common Drug-Drug Interactions. Cureus 15(3): e36272. DOI 10.7759/cureus.36272

healthcare provider [2]. Additionally, patients should be encouraged to ask questions about their medications and potential interactions with their physicians [3].

In many developing countries, the availability of drugs without a prescription is still being practiced. This encourages self-medication or taking suggestions from non-physician healthcare providers [4,5]. Self-medication may increase the risk of DDIs because individuals may not have the necessary knowledge and expertise to understand the potential risks and side effects of taking multiple medications. There are many online avenues that can provide some guidelines about DDI [6]. However, predicting and understanding DDIs can be challenging due to the complex and multifactorial nature of these interactions [7]. In recent years, there has been a growing interest in leveraging natural language processing (NLP) techniques to aid in DDI prediction and explanation [8].

ChatGPT is a large language model that has been shown to perform well on a variety of NLP tasks, including text classification and question answering. It is being used all over the world. The major advantage of ChatGPT is the capability of generating text in a conversational manner [9]. Previous studies have ascertained the capability of ChatGPT in academic writing [10], in answering various levels of reasoning questions in Pathology and Microbiology [11, 12], or the capability of writing medical examinations [13]. However, no study was conducted to ascertain the capability to predict and explain common DDIs.

In this context, our study aimed to find the effectiveness of ChatGPT in predicting and explaining common DDIs. The study would contribute to the growing body of literature on NLP-based approaches to DDI prediction and explanation. If successful, ChatGPT may be used as a handy tool for aiding clinicians as well as patients in identifying potential drug interactions and making informed decisions.

Materials And Methods

Type and setting

This was a cross-sectional observational study conducted with data that are available for public audit. The primary source of data was the world wide web. The study was conducted from February 20 to March 5, 2023.

Ethics

This study involves the collection of data from an online website and analyzing it. No human or animal was involved in this study. No patient information was used to generate the response from the language model. Hence, this study does not require any ethical committee review.

List of interaction

We have collected the list of DDIs from a previously published article by Kheshti et al. [14]. This article used a total of 40 interaction pairs for testing five software programs. We have used the same set of DDI for checking the capability of ChatGPT.

Data collection

A user with a free account in ChatGPT interacted with the artificial intelligence-based language model with the aim to collect data for this study. Each pair of the drugs were searched with two questions - "can I take X and Y together?" and "why should I not take X and Y together?" The output was stored for further analysis. Figure 1 shows the brief process of the data collection method.

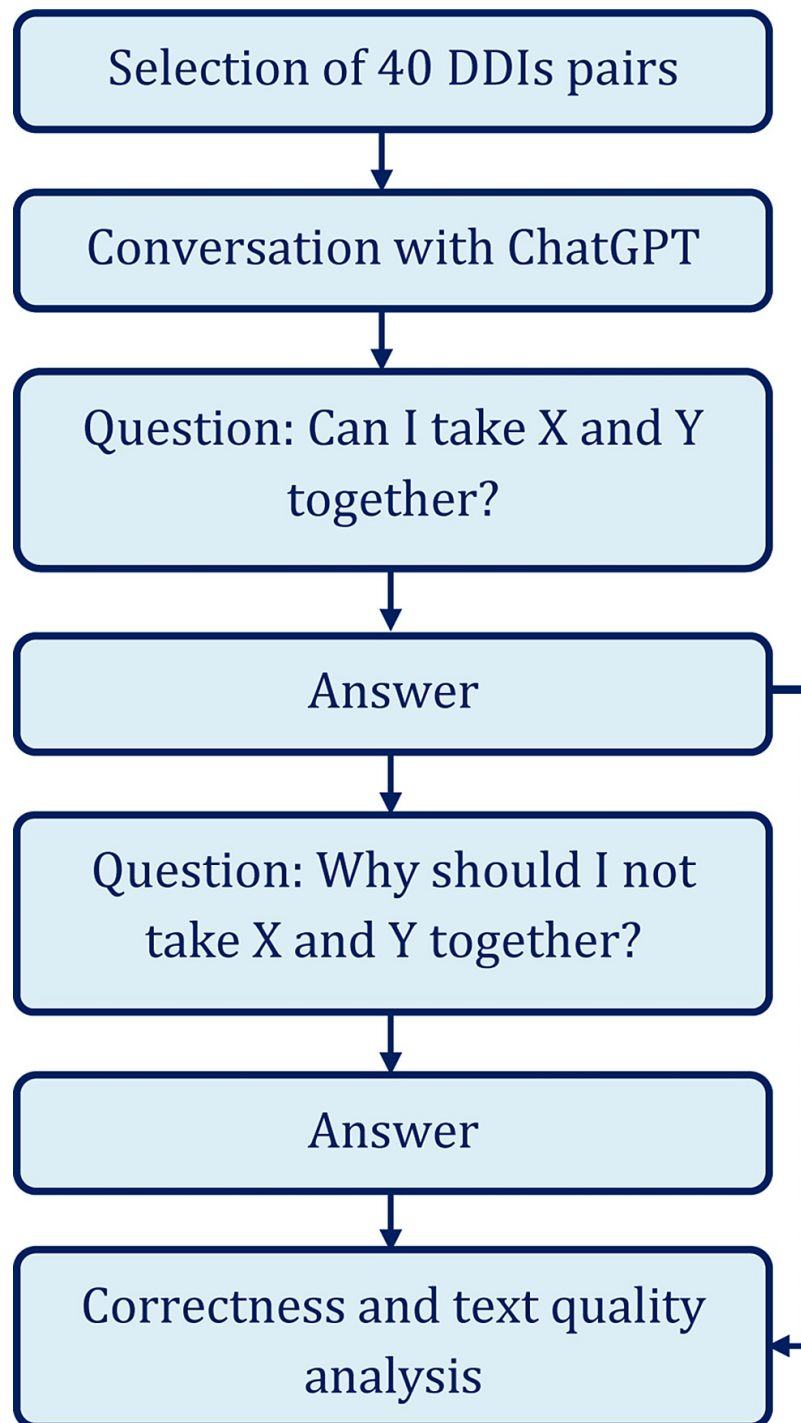


FIGURE 1: Brief study flowchart

DDIs: Drug-drug interactions

X and Y were replaced with two drug names during the conversation

Data analysis

The responses were checked by two pharmacologists by taking references from Stockley's Drug Interactions Pocket Companion 2015 and a consensus was reached for categorizing the output as "correct" and "incorrect" [15]. The "correct" ones were further divided into "conclusive" and "inconclusive." The text was checked for reading ease scores and grades of education required to understand the text. Data were tested by descriptive and inferential statistics. Categorical data were tested by Fisher's exact test. Continuous variables were tested by unpaired t-test. For checking a single parameter with a hypothetical value, we used a one-sample t-test [16]. For grade level, we considered the American Medical Association and National

Institutes of Health to recommend a level that is no higher than a sixth-grade reading (Flesch Reading Ease Score 80-90) [17]. A p-value < 0.05 was considered statistically significant. We used GraphPad Prism (GraphPad Software Inc., USA) to conduct the statistical test.

Results

Among the 40 DDI pairs, one answer was incorrect in the first question. The rest of the answers were correct, but 19 were conclusive in describing the basis for why the drugs should not be taken together. However, for that one incorrect question, when we asked it “why should I not take X and Y together?”, it rectified the response and informed us that the previous answer was wrongly posted. For the second question, one answer was provided wrong. Among correct answers, 17 were conclusive and the rest were inconclusive. The number of incorrect, correct-conclusive, correct - inconclusive answers is shown in Table 1.

Category	Correctness		Number	P-value
Answer to “Can I take?”	Incorrect		1	0.003*
	Correct	Conclusive	19	
		Inconclusive	20	
Answer to “Why should I not take?”	Incorrect		1	0.002*
	Correct	Conclusive	17	
		Inconclusive	22	

TABLE 1: Correctness of the answer provided by ChatGPT for two types of questions

*p-value of Fishers exact test

Output text was analyzed for its sentences, words, syllables per word, reading ease score, and grade levels and we found that the answers to both questions were having equal reading ease scores and require a similar educational grade to understand the text. The words and sentences were higher in the answers to the second question. Text analysis is shown in Table 2.

Variable	Answer to “Can I take?”		Answer to “Why should I not take?”		P-values
	Mean±SD	Min - Max	Mean±SD	Min - Max	
Sentences	2.85±1.66	1 - 8	3.85±1.87	1 - 9	0.01
Words	59.93±41.27	19 - 198	82.7±38.42	21 - 189	0.01
Word/sentence	21.42±5.59	11.5 – 34.3	21.99±4.22	13.9 – 30.7	0.6
Syllable/word	1.89±0.14	1.6 – 2.2	1.85±0.16	1.5 – 2.3	0.27
Ease score*	27.64±10.85	9.4 – 57.3	29.35±10.16	11.4 – 61.2	0.47
Grade†	15.06±2.79	8.8 – 20.2	14.85±1.97	9.3 – 18.2	0.69

TABLE 2: Descriptive statistics of generated text output

P-values are of unpaired t-test

SD: standard deviation, Min: Minimum, Max: maximum

*Flesch Reading Ease Score

†Flesh-Kincaid Grade Level

The grade level required for understanding the text was too high for the information to be understood by common users, especially by non-native speakers of English. When we compared the reading levels with

hypothetical 6th grade, the grades were significantly higher than expected ($t = 20.57$, $df = 39$, $p < 0.0001$, 95% CI = 8.17 - 9.95 for first answers; $t = 28.43$, $df = 39$, 95% CI = 8.22 - 9.48, $p < 0.0001$ for second answers).

Discussion

The present study investigated the effectiveness of ChatGPT, a large language model, in predicting and explaining common DDIs. Patients who are taking multiple medications may be at an increased risk of experiencing adverse events or drug toxicity if they are not aware of potential interactions between their medications. ChatGPT was evaluated using a two-stage question format. Although the answers were correct in 39 cases among 40 entries, the correct answers lack clarity in about 50% of the answers. However, this accuracy is measured from a pharmacological point of view. From a layman's view, the information is correct and can be helpful to detect any DDIs by the patients.

The next issue was about readability. We found that the answers can be understood by a college graduate. This is reasonably higher than the recommended level of grade 6th as suggested by the American Medical Association and National Institutes of Health [17]. Hence, the answer to a direct question about DDI may be difficult to understand for the general population. The underlying reason may be maintaining the scientific contents [18]. However, ChatGPT can be made more productive for documents that are easier to understand. One can instruct the ChatGPT to simplify the language. In response to explaining a DDI, the ChatGPT first provided an output as shown in Figure 2.

Methotrexate and Probenecid can interact with each other, and taking them together can increase the risk of Methotrexate toxicity. This is because Probenecid can decrease the excretion of Methotrexate from the body, leading to increased levels of Methotrexate in the blood. Methotrexate toxicity can cause serious side effects such as nausea, vomiting, mouth sores, liver damage, and kidney damage.

Therefore, it is generally not recommended to take Methotrexate and Probenecid together. If you need to take both medications, your healthcare provider may adjust the dosages or choose an alternative medication to Probenecid. It's important to follow your healthcare provider's instructions and report any symptoms of Methotrexate toxicity, such as nausea or mouth sores, promptly. It's also essential to have regular blood tests to monitor your Methotrexate levels and kidney function.

FIGURE 2: Answer of ChatGPT to question – “Why should I not take Methotrexate and Probenecid together?”

The answer was collected on March 10, 2023 and was generated by February 13 version. Flesh-Kincaid grade level 14.1, Flesch reading ease score 27.2

We have asked ChatGPT to make the presentation lucid and the output is shown in Figure 3.

Methotrexate and Probenecid are two medications used to treat different health conditions. However, when taken together, they can interact with each other and cause harmful side effects. This is because Probenecid can cause Methotrexate levels in the blood to increase, which can lead to Methotrexate toxicity. Methotrexate toxicity can cause serious problems like nausea, vomiting, mouth sores, liver damage, and kidney damage.

Therefore, it is generally not recommended to take Methotrexate and Probenecid together. If you need to take both medications, your healthcare provider may adjust the dosages or choose a different medication to avoid any potential interactions. It's important to follow your healthcare provider's instructions and report any symptoms of Methotrexate toxicity promptly. Regular blood tests may also be necessary to monitor your Methotrexate levels and kidney function.

FIGURE 3: Output of ChatGPT when it was instructed to “explain the above text in lucid language.”

Output was collected on March 10, 2023 and was generated by February 13 version. ChatGPT can remember previous conversations with it. Flesh-Kincaid grade level 14.3, Flesch reading ease score 21.3.

However, the grade level did not reduce, and the ease score did not rise. We then asked specifically to make the text suitable for a 6th-grade student and the response is shown in Figure 4.

Methotrexate and Probenecid are two medicines that can be dangerous if taken together. When you take these two medicines together, Probenecid can make Methotrexate stay longer in your body, causing it to build up to dangerous levels. This can make you feel sick and can damage your liver and kidneys. So it's best not to take Methotrexate and Probenecid together unless your doctor tells you it's okay.

FIGURE 4: Reply of ChatGPT to the question - “Can you make it simpler so that a 6th-grade student can understand?”

Reply was collected on March 10, 2023 and was generated by February 13 version. ChatGPT can remember previous conversations with it. Flesh-Kincaid grade level 9.8, Flesch reading ease score 54.4

This time, the grade level was reduced to 10-12th grade with an increment of reading ease score. However, according to the aim, this method of generating text was not practiced and may be considered in further research.

There are several sources of information on DDIs that can help healthcare providers and patients make informed decisions regarding medication use [15]. Drug labels and package inserts provide detailed information on a medication's pharmacology, dosing, and potential drug interactions. Several online tools and mobile applications are available that allow users to check for potential drug interactions between two or more medications. Professional organizations and healthcare institutions often develop guidelines for managing specific medical conditions or prescribing medications. However, it is important to note that information on drug interactions can be complex and may require specialized knowledge. The introduction of ChatGPT is an addition to the already available resources. However, patients and healthcare providers should always consult reliable sources of information and seek advice from a qualified healthcare professional before making any changes to medication regimens.

This study has some limitations. We have used only 40 pairs to test the model and ChatGPT was trained with data that were available till 2021. Hence, the data may not be updated. In addition, the model is regularly being upgraded. Hence, finding at this point in time may not be applicable in the future. We used the free version of the program available at this time and if the capability of the free and paid versions differs, it was beyond our capacity to comment.

Conclusions

ChatGPT is an artificial intelligence-based model that can help predict and explain DDIs. Patients who do not have immediate access to healthcare facilities can use ChatGPT to get an initial understanding of

potential drug interactions. However, ChatGPT may not always provide complete guidance. Patients should always consult with their healthcare provider before starting or stopping any medication, or before making changes to their treatment plan. To make ChatGPT more useful for patients, there is a need for further improvement in its ability to predict and explain DDIs accurately.

Additional Information

Disclosures

Human subjects: All authors have confirmed that this study did not involve human participants or tissue.

Animal subjects: All authors have confirmed that this study did not involve animal subjects or tissue.

Conflicts of interest: In compliance with the ICMJE uniform disclosure form, all authors declare the following: **Payment/services info:** All authors have declared that no financial support was received from any organization for the submitted work. **Financial relationships:** All authors have declared that they have no financial relationships at present or within the previous three years with any organizations that might have an interest in the submitted work. **Other relationships:** All authors have declared that there are no other relationships or activities that could appear to have influenced the submitted work.

Acknowledgements

We acknowledge the use of ChatGPT, an artificial intelligence language model developed by OpenAI, for its contribution to our research. ChatGPT was used to assist with the literature review. We appreciate the efforts of the developers and contributors of ChatGPT, which facilitated our research and provided valuable insights. The corresponding author would like to thank Sarika Mondal and Ahana Aarshi for sacrificing their family time during the preparation of this manuscript.

References

- Mousavi S, Ghanbari G: Potential drug-drug interactions among hospitalized patients in a developing country. *Caspian J Intern Med.* 2017, 8:282-8. [10.22088/cjim.8.4.282](https://doi.org/10.22088/cjim.8.4.282)
- Qato DM, Wilder J, Schumm LP, Gillet V, Alexander GC: Changes in prescription and over-the-counter medication and dietary supplement use among older adults in the United States, 2005 VS 2011. *JAMA Intern Med.* 2016, 176:473-82. [10.1001/jamainternmed.2015.8581](https://doi.org/10.1001/jamainternmed.2015.8581)
- Gallagher JM, Post DM, Weiss BD, et al.: Patients' question-asking behavior during primary care visits: a report from the AAFP National Research Network. *Ann Fam Med.* 2010, 8:151-9. [10.1370/afm.1055](https://doi.org/10.1370/afm.1055)
- Rashid M, Chhabra M, Kashyap A, Undela K, Gudi SK: Prevalence and predictors of self-medication Practices in India: a systematic literature review and meta-analysis. *Curr Clin Pharmacol.* 2020, 15:90-101. [10.2174/1574884714666191122103953](https://doi.org/10.2174/1574884714666191122103953)
- Selvaraj K, Kumar SG, Ramalingam A: Prevalence of self-medication practices and its associated factors in Urban Puducherry, India. *Perspect Clin Res.* 2014, 5:52-6. [10.4103/2229-3485.124569](https://doi.org/10.4103/2229-3485.124569)
- Xiong G, Yang Z, Yi J, et al.: DDInter: an online drug-drug interaction database towards improving clinical decision-making and patient safety. *Nucleic Acids Res.* 2022, 50:D1200-7. [10.1093/nar/gkab880](https://doi.org/10.1093/nar/gkab880)
- Vilar S, Friedman C, Hripscak G: Detection of drug-drug interactions through data mining studies using clinical sources, scientific literature and social media. *Brief Bioinform.* 2018, 19:863-77. [10.1093/bib/bbx010](https://doi.org/10.1093/bib/bbx010)
- Han K, Cao P, Wang Y, et al.: A review of approaches for predicting drug-drug interactions based on machine learning. *Front Pharmacol.* 2021, 12:814858. [10.3389/fphar.2021.814858](https://doi.org/10.3389/fphar.2021.814858)
- Eysenbach G: The role of ChatGPT, generative language models, and artificial intelligence in medical education: a conversation with ChatGPT and a call for papers. *JMIR Med Educ.* 2023, 9:e46885. [10.2196/46885](https://doi.org/10.2196/46885)
- Salvagno M, Taccone FS, Gerli AG: Can artificial intelligence help for scientific writing?. *Crit Care.* 2023, 27:75. [10.1186/s13054-023-04380-2](https://doi.org/10.1186/s13054-023-04380-2)
- Sinha R K, Deb Roy A, Kumar N, Mondal H: Applicability of ChatGPT in assisting to solve higher order problems in pathology. *Cureus.* 2023, 15:35237. [10.7759/cureus.35237](https://doi.org/10.7759/cureus.35237)
- Das D, Kumar N, Longjam L, et al.: Assessing the capability of ChatGPT in answering first- and second-order knowledge questions on microbiology as per competency-based medical education curriculum. *Cureus.* 2023, 15:e36034. [10.7759/cureus.36034](https://doi.org/10.7759/cureus.36034)
- Gilson A, Safranek CW, Huang T, Socrates V, Chi L, Taylor RA, Chartash D: How does ChatGPT perform on the United States Medical Licensing Examination? The implications of large language models for medical education and knowledge assessment. *JMIR Med Educ.* 2023, 9:e45312. [10.2196/45312](https://doi.org/10.2196/45312)
- Kheshti R, Aalipour M, Namazi S: A comparison of five common drug-drug interaction software programs regarding accuracy and comprehensiveness. *J Res Pharm Pract.* 2016, 5:257-63. [10.4103/2279-042X.192461](https://doi.org/10.4103/2279-042X.192461)
- Preston CL: *Stockley's Drug Interactions Pocket Companion*. Preston CL (ed): Pharmaceutical Press, London, UK; 2015. https://books.google.co.in/books/about/Stockley_s_Drug_Interactions_Pocket_Comp.html?id=bR64wAEACAAJ&redir_esc=y
- Mondal H, Mondal S, Majumder R, De R: Conduct common statistical tests online. *Indian Dermatol Online J.* 2022, 13:539-42. [10.4103/idoj.idoj_605_21](https://doi.org/10.4103/idoj.idoj_605_21)
- Akinleye SD, Garofolo-Gonzalez G, Montuori M, Culbertson MD, Hashem J, Edelstein DM: Readability of the most commonly accessed online patient education materials pertaining to pathology of the hand. *Hand (N Y).* 2018, 13:705-14. [10.1177/1558944717726138](https://doi.org/10.1177/1558944717726138)
- Plavén-Sigray P, Matheson GJ, Schiffler BC, Thompson WH: The readability of scientific texts is decreasing over time. *Elife.* 2017, 6:e27725. [10.7554/eLife.27725](https://doi.org/10.7554/eLife.27725)