# Assessing the Accuracy of ChatGPT in Generating Patient-Facing Information on Pediatric Cataracts

**Deniz Goodman,** Ta Chen Peter Chang, on behalf of the Pediatric Cataract Artificial Intelligence Research Study Group

# Financial Disclosure

- Presenter

- Deniz Goodman: None


- Co-authors

- Ta Chen Peter Chang: None

# Background

- ChatGPT

  - Human-like responses to questions

  - Search engine functionality[1]

  - Problem: Unable to assess reliability of sources[1]

  - GPT-3.5 knowledge cutoff date: September 2021[2]

1.  Dave T, Athaluri SA, Singh S. ChatGPT in medicine: an overview of its applications, advantages, limitations, future prospects, and ethical considerations. Front Artif Intell. 2023 May 4;6:1169595. doi: 10.3389/frai.2023.1169595.
2.  https://twitter.com/OpenAI/status/1707077710047216095?s=20
https://blog.logomyway.com/chatgpt-logo-history/

# Objective

- Evaluate the accuracy of ChatGPT responses to parents' frequently asked questions about pediatric cataract.

Bascom Palmer
Eye Institute
UNIVERSITY OF MIAMI HEALTH SYSTEM

SAMUEL & ETHEL BALKAN
INTERNATIONAL PEDIATRIC
GLAUCOMA CENTER

# Compiling Pediatric Cataract FAQ's

| Frequently Asked Question (FAQ) Number | FAQ |
| --- | --- |
| 1 | What is a pediatric cataract? |
| 2 | How can I prevent cataracts in my child? |
| 3 | Which children are at risk for cataracts? |
| 4 | What are the symptoms of cataracts in a child? |
| 5 | How are cataracts diagnosed in a child? |
| 6 | What are possible complications of cataracts in a child? |
| 7 | What are the treatment options for pediatric cataracts? |
| 8 | How is a cataract removed in infants and children? |
| 9 | What causes pediatric cataracts? |
| 10 | How does a pediatric cataract cause vision loss? |

# Question Websites and Categories

- Question Categories
  - General Information
  - Diagnosis, Signs, and Symptoms
  - Pathophysiology
  - Prognosis and Treatment
- Website Sources
  - American Association for Pediatric Ophthalmology and Strabismus
  - Stanford Medicine Children's Health
  - Boston Children's Hospital
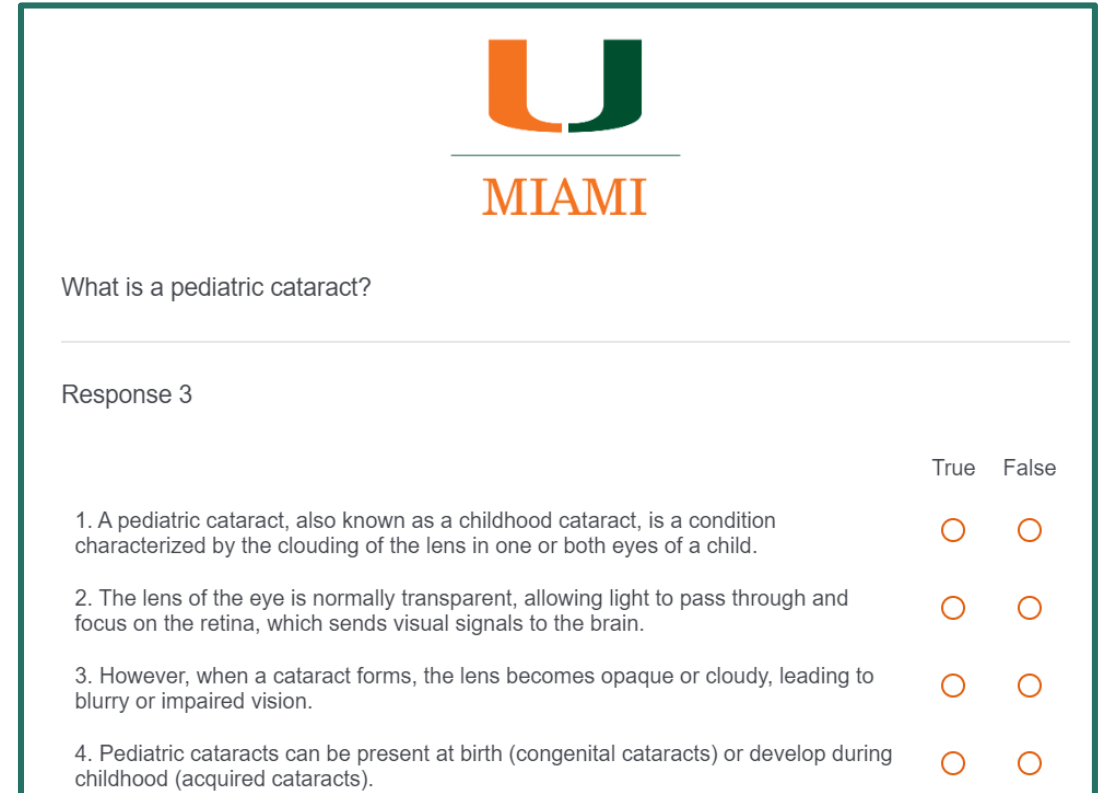  - Children's Health (Dallas)
  - March of Dimes

# ChatGPT Response Generation

| Unprompted ChatGPT | Prompted ChatGPT | Website Response |
|---|---|---|
| FAQ alone | FAQ + "Be specific and incorporate any applicable medical guidelines."[2] | Gold standard reference response |

2. Johnson D, Goodman R, Patrinely J, Stone C, Zimmerman E, Donald R, Chang S, Berkowitz S, Finn A, Jahangir E, Scoville E, Reese T, Friedman D, Bastarache J, van der Heijden Y, Wright J, Carter N, Alexander M, Choe J, Chastain C, Zic J, Horst S, Turker I, Agarwal R, Osmundson E, Idrees K, Kieman C, Padmanabhan C, Bailey C, Schlegel C, Chambless L, Gibson M, Osterman T, Wheless L. Assessing the Accuracy and Reliability of AI-Generated Medical Responses: An Evaluation of the Chat-GPT Model. Res Sq [Preprint]. 2023 Feb 28:rs.3.rs-2566942. doi: 10.21203/rs.3.rs-2566942/v1.

# Exercise Development

- Exercise created on Qualtrics

- Responses split into statements
  - As defined by individual sentences or items in a list

- Added 1 quality control statement per question

- Removed the chatbot's references to itself to avoid biasing graders who were masked to the source of the response

What is a pediatric cataract?

Response 3

|  | True | False |
|---|:---:|:---:|
| 1. A pediatric cataract, also known as a childhood cataract, is a condition characterized by the clouding of the lens in one or both eyes of a child. | ○ | ○ |
| 2. The lens of the eye is normally transparent, allowing light to pass through and focus on the retina, which sends visual signals to the brain. | ○ | ○ |
| 3. However, when a cataract forms, the lens becomes opaque or cloudy, leading to blurry or impaired vision. | ○ | ○ |
| 4. Pediatric cataracts can be present at birth (congenital cataracts) or develop during childhood (acquired cataracts). | ○ | ○ |

# Analysis

- Comparison: proportion of statements graded as correct between sources
- Chi-square tests

| Unprompted ChatGPT | Prompted ChatGPT | Website Response |
|---|---|---|
| FAQ alone | FAQ + "Be specific and incorporate any applicable medical guidelines."[2] | Gold standard reference response |

# Results

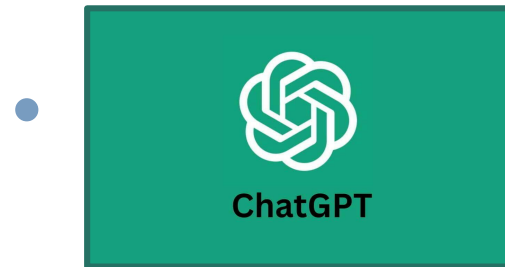## Pediatric Cataract (n=15 reviewers)

| Category | N | Mean Proportion of Statements Graded as Correct | | |
|---|---|---|---|---|
| | | Unprompted | Prompted | Website |
| General Cataract | 3 | 0.90 | 0.87 | 0.97 |
| Diagnosis, Signs, and Symptoms | 2 | 0.96 | 0.95 | 0.84 |
| Pathophysiology | 2 | 0.95 | 0.96 | 0.87 |
| Prognosis and Treatment | 3 | 0.91 | 0.89 | 0.95 |
| All FAQs* | 10 | 0.93 | 0.91 | 0.90 |

*P=0.0103

# Discussion

- Accuracy
  - Prompted ChatGPT ≈ Unprompted ChatGPT
  - ChatGPT > Websites
- Websites may contain outdated information
  - Unknown author and publication/date last updated
- ChatGPT continually updated over time
  - Results may vary with future updates

# Conclusions



- ChatGPT ≠ [doctor]

- Looking ahead: medically-validated chatbots answer remaining questions after appointments

- True statement ≠ useful information

- Future directions: evaluate relevance and comprehensiveness