**The CHART Abstract Checklist**

|  |  |  |  |
| --- | --- | --- | --- |
| HEADING | CHART Checklist # | ITEM | Page # |
| Background | **2a** | State the scientific background, rationale, and healthcare context for evaluating the generative AI-driven chatbot(s), referencing relevant literature when applicable. |  |
|  | **2b** | State the aims and research questions including the target audience, intervention, comparator(s), and outcome(s). |  |
| Methods |  |  |  |
| Model Identifiers | **3a** | State the name and version identifier(s) of the generative AI model(s) and chatbot(s) under evaluation, as well as their date of release or last update. |  |
|  | **3b** | State whether generative AI model(s) and chatbot(s) are open-source versus closed-source/proprietary. |  |
| Model Details | **4a** | State whether the generative AI model was a base model or a novel base model, tuned model, or fine-tuned model. |  |
| Prompt Engineering | **5a** | Describe the evolution of study prompt development. |  |
|  | **5ai** | Describe the sources of prompts. |  |
|  | **5aii** | State the number and characteristics of the individual(s) involved in prompt engineering. |  |
|  | **5aiii** | Provide details of any patient and public involvement during prompt engineering. |  |
| Query Strategy | **6a** | State route of access to generative AI model. |  |
|  | **6b** | State the date(s) and location(s) of queries for the generative AI-driven chatbot(s) including the day, month, and year as well as city and country. |  |
| Performance Evaluation | **7a** | Define the ground truth or reference standard used to define successful generative AI-driven chatbot performance. |  |
|  | **7b** | Describe the process undertaken for the performance evaluation of the generative AI-driven chatbot(s). |  |
| Sample Size | **8** | Report how the sample size was determined. |  |
| Data Analysis | **9a** | Describe statistical analysis methods, including any evaluation of reproducibility of generative AI-driven chatbot responses. |  |
| Results |  |  |  |
|  | **10a** | Report the performance evaluation undertaken including the alignment between generative AI-driven chatbot output and ground truth or reference standard using quantitative or mixed methods approaches as applicable. |  |