Our original submission number was 1720. We would like to thank all reviewers for their very valuable and constructive comments, which we have leveraged to improve this paper. We feel highly encouraged that this work is conditionally accepted. In this letter, we answer ethics-related questions and summarize the revisions we made requested by the EAC meta-reviews.

1 Requested Revisions

**Requested Revision 1:** Reviewers expressed concern about how privacy was protected and consent was obtained to use medical records. In your camera-ready, you should share how your team ensured that patient privacy was protected in this dataset and through what mechanism the data was made available to share that ensures no breach of patient confidentiality.

**Our Revision:** We added the following information in the camera ready (Section 3): For both datasets, the dialogs are crawled from openly accessible medical websites whose owners make these dialogs visible to the public. The patients’ personal information is de-identified by owners of these websites. We further checked the crawled dialogs to ensure they do not contain private information of patients. Besides, we also manually removed borderline sensitive information, such as specific dates and destinations in patients’ travel histories. Experts in privacy and security domains helped to check the final version of shared data and ensured there is no breach of patient privacy or confidentiality.

2 Response to Questions Related to Ethics

**Question 1:** How many persons were consulted for creating the dialogues?

**Response:** In the English dataset, the conversations are from 582 patients and 117 doctors. In the Chinese dataset, the conversations are from 935 patients and 352 doctors.

**Question 2:** How did you anonymize the patients information for preserving the privacy related information?

**Response:** The patients’ personal information was de-identified by owners of medical websites from which the dialogs are crawled. We further checked the crawled dialogs to ensure they do not contain private information of patients. Besides, we also manually removed borderline sensitive information, such as specific dates and destinations in patients’ travel histories. Experts in privacy and security domains helped to check the final version of shared data and ensured there is no breach of patient privacy or confidentiality.

**Question 3:** Did you consult any medical practitioners for judging the authenticity of the responses generated?

**Response:** We performed human evaluation of the generated responses. Five medical professionals were asked to give ratings (from 1 to 5, higher is better) to generated responses in four aspects: 1) Correctness: whether the response is clinically correct; 2) Relevance: how relevant the response is to the conversation history; 3) Informativeness: how much medical information and suggestions are given in the response; and 4)
Doctor-like: how the response sounds like a real doctor. The human evaluation results in Table 3 and 8 show that the generated responses are promising in being doctor-like, relevant to conversation history, clinically informative and correct.

**Question 4:** What procedures were followed to acquire these data?

**Response:** For both datasets, the dialogs are crawled from openly accessible medical websites whose owners make these dialogs visible to the public. The patients’ personal information is de-identified by owners of these websites. We further checked the crawled dialogs to ensure they do not contain private information of patients. Besides, we also manually removed borderline sensitive information, such as specific dates and destinations in patients’ travel histories. Experts in privacy and security domains helped to check the final version of shared data and ensured there is no breach of patient privacy or confidentiality.