

Impact de l'association d'une base de connaissance sur la qualité des réponses d'un agent conversationnel accompagnant les patients dans la gestion de leurs médicaments à la sortie d'hôpital

Krifa S¹², Charveys A¹², Chalot E³, Singer F³, Houeix J³, Kaestli L-Z¹, Bonnabry P¹²

¹ Pharmacie, Hôpitaux Universitaires de Genève, Suisse • ² ISPSO, Université de Genève, Suisse • ³ Direction de la transformation numérique-Intelligence augmentée, Hôpitaux Universitaires de Genève, Suisse

25 médicaments de sortie d'hôpital testés

6 questions types de patients posées au *chatbot* LLM

2 versions GPT-4.1 comparées

11 évaluateurs finaux (patients et professionnels de santé)

5 Dimensions qualité évaluées

Contexte

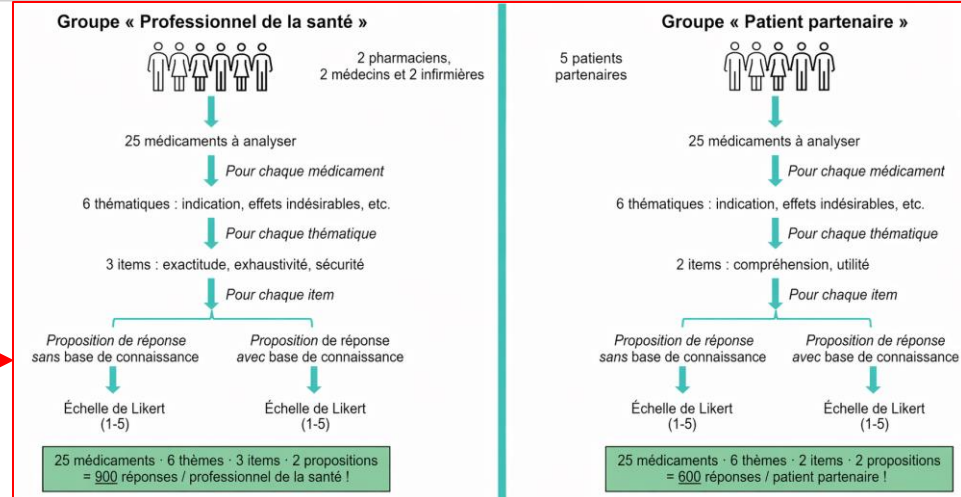
- **Sortie d'hôpital :**
 - Période à risque d'erreurs médicamenteuses et de réadmissions.
- **Chatbots LLM (Large Language Model) :**
 - Capacité à répondre aux questions des patients.
 - Risque d'hallucinations.
- **Notre hypothèse:**
 - Une **base de connaissance (BC)** peut **sécuriser** les réponses d'un *chatbot* et en améliorer la **qualité**.

Objectif

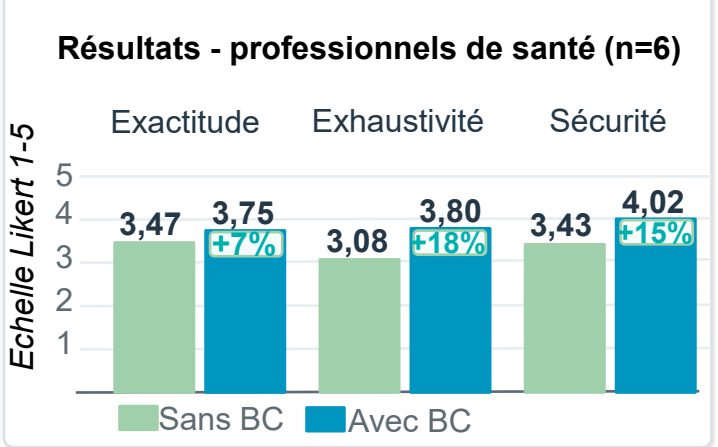
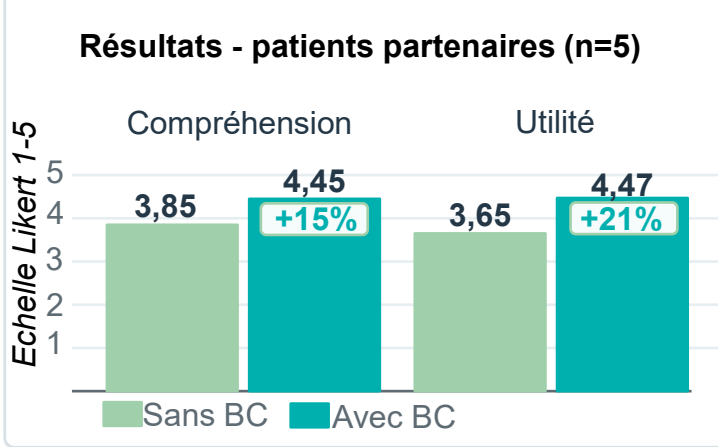
Évaluer l'**impact** de l'association d'une **base de connaissance** clinique validée à un même **LLM** sur la **qualité** des **réponses** aux questions des **patients** sur leur **médication** de **sortie**.

Méthode

- 1• Construction **base de connaissance** à partir de 4 sources d'informations destinées aux **patients** : *Compendium suisse, Mayo Clinic, MedlinePlus, Drugs.com.*
 - 2• **Evaluation** des réponses (GPT4.1 VS [GPT4.1 + BC]):
 - a-Axe **quantitatif** :
 - b-Axe **qualitatif** :
- Focus group** de discussion sur thématiques spécifiques.



Résultats



Conclusion

0 hallucination identifiée dans les réponses évaluées.
TOUTES les dimensions qualité ont été améliorées par la BC.
LLM seul globalement satisfaisant, mais moins complet.

Perspectives

Evaluation d'un **prototype** de *chatbot* [LLM + BC] en conditions réelles, avec des **patients** sortant d'hôpital.