

# DEVELOPMENT OF A PYTHON SCRIPT FOR AUTOMATED DATA EXTRACTION FROM PHARMACEUTICAL ARTICLES AND COMPARISON OF CHATGPT 4.5 AND 5.0 MODEL PERFORMANCE

Juliette Mutin<sup>1</sup>, Louise Lardeux<sup>1</sup>, Missipsa Annane<sup>2</sup>, Clarisse Martin<sup>1</sup>, Jean-François Bussières<sup>1,3</sup>  
<sup>1</sup> Unité de recherche en pratique pharmaceutique, CHU Sainte-Justine, Montréal (Québec) Canada;  
<sup>2</sup> Département de génie informatique et génie logiciel Polytechnique Montréal  
<sup>3</sup> Faculté de pharmacie, Université de Montréal, Montréal (Québec) Canada

6ER-040

## INTRODUCTION

2010

The emergence of artificial intelligence (AI), including conversational agents such as ChatGPT.

2024

After optimization of the script and the prompts, the mean compliance increased to 85 ± 13%, and only 10% of articles showed hallucinations. However, human validation remained necessary.

Impact Pharmacie is updated once a year by pharmacy students. Each article requires **1 to 4 hours** of work.

2022

A first study developed a custom GPT to extract data, comparing the results with those entered into Impact Pharmacie. Mean compliance was 77 ± 19%, with 23% of articles containing hallucinations.

2025

## OBJECTIVES

- To develop a Python script leveraging the ChatGPT API to automate the extraction of targeted data from pharmaceutical literature.

## METHODS

**Study design :** Retrospective descriptive study

**Team:**

- Five pharmacy students trained in the reading and analysis of scientific articles using a standardized operating procedure
- One engineering student from Polytechnique Montréal (computer engineering), trained in Python coding
- One supervising pharmacist

**Sample:**

- 26 pharmaceutical interventional studies in English addressing pharmacists' roles
- Studies whose data were extracted and integrated into the Impact Pharmacie platform in summer 2025, using a methodology based on a comparison between ChatGPT versions 4.5 and 5.0, with human data verification

**Study variables:**

- 22 criteria for ChatGPT 4.5 and 23 criteria for ChatGPT 5.0, corresponding to the 23 questions used to query the AI

**Process:**

- Reuse of the prompt and the 22 questions developed for data extraction with ChatGPT 4.5 in summer 2025, adapted into a Python-compatible code format
- Development of an English version in addition to the French version
- Coding performed by a student
- Software use
- Code execution on the 26 studies
- Assessment of response compliance compared with pre-identified elements by five students, using three levels (Compliant, Non-compliant, Incomplete)
- Code optimization and execution on five additional articles

**Analysis plan - Variables:**

- Calculation of compliance rates (per variable and overall mean) (%)
- Analysis time per article (minutes)

## RESULTS

- Compliance of AI-generated responses (Table 1):
  - In French: 72 ± 9% for ChatGPT 4.5 and 86 ± 8% for ChatGPT 5.0.
  - In English: 77 ± 8% for ChatGPT 4.5 and 87 ± 9% for ChatGPT 5.0.
- Missing entries accounted for 2% (25/1,334).
- Five articles contained some responses formulated in English instead of French.

Variable	ChatGPT 4.5 – English	ChatGPT 4.5 – French	ChatGPT 5.0 – English	ChatGPT 5.0 – French
	Total compliance (%) (n/N)			
Year of publication	20/26 = 77%	21/26 = 81%	25/26 = 96%	26/26 = 100%
Country	23/26 = 88%	23/26 = 88%	26/26 = 100%	26/26 = 100%
Type of healthcare setting	24/26 = 92%	22/26 = 85%	23/26 = 88%	22/26 = 85%
Monocentric study	/	/	22/26 = 88%	22/26 = 92%
Study dates and duration	19/26 = 73%	14/26 = 54%	21/26 = 81%	21/26 = 81%
Study design	16/26 = 62%	17/26 = 65%	22/26 = 85%	20/26 = 77%
Type of study (temporality)	26/26 = 100%	26/26 = 100%	26/26 = 100%	26/26 = 100%
Primary objective	24/26 = 92%	21/26 = 81%	24/26 = 96%	22/26 = 88%
Secondary objectives	13/26 = 50%	14/26 = 54%	13/26 = 50%	13/26 = 50%
Number of groups	23/26 = 88%	23/26 = 88%	23/26 = 88%	23/26 = 88%
Control group	23/26 = 88%	22/26 = 85%	24/26 = 92%	22/26 = 85%
Allocation method (participant assignment)	26/26 = 100%	25/26 = 96%	25/26 = 96%	25/26 = 96%
Inclusion criteria	20/26 = 77%	21/26 = 81%	25/26 = 96%	25/26 = 96%
Exclusion criteria	24/26 = 92%	23/26 = 88%	25/26 = 96%	24/26 = 92%
Blinding (masking)	23/26 = 88%	25/26 = 96%	25/26 = 96%	26/26 = 100%
Primary outcome – description	22/26 = 85%	12/26 = 46%	19/26 = 73%	21/26 = 81%
Primary outcome – results and significance	19/26 = 73%	13/26 = 50%	19/26 = 73%	20/26 = 77%
Secondary outcomes	17/26 = 65%	9/26 = 35%	19/26 = 73%	18/26 = 69%
Intervention description	12/26 = 46%	11/26 = 42%	24/26 = 92%	25/26 = 96%
Intervention delivery	19/26 = 73%	21/26 = 81%	23/26 = 88%	24/26 = 92%
Intervention providers	21/26 = 81%	18/26 = 69%	23/26 = 88%	22/26 = 85%
Intervention duration	11/26 = 42%	12/26 = 46%	18/26 = 69%	14/26 = 54%
Study limitations	18/26 = 69%	18/26 = 69%	24/26 = 92%	23/26 = 92%

Table 1: Variable compliance profile in French and English for ChatGPT versions 4.5 and 5.0

- Extraction time: 380.03 minutes.
- Number of tokens per article: 122,227.29.
- Cost per article: \$0.65.

## DISCUSSION

- Python script via API outputs Excel → SQL → optional website DB integration
- English extraction shows higher compliance than French
- API extraction performs worse than conversational chat, possibly due to:
  - No memory in the API
  - Reduced script for coding constraints
  - JSON responses more concise than natural language
- Perspectives
  - Code optimization
  - Prompt and question refinement
  - Automated Impact Pharmacie updates via Python extraction
  - Large-scale AI-driven data generation

## LIMITATIONS

- Assessment performed by five students
- Small sample size (n = 26)

## CONCLUSION

- Python script integrated into the ChatGPT API for automated extraction of targeted data from pharmaceutical articles
- Compliance rates suggest
  - Automation is feasible
  - Lower performance via the API compared with the ChatGPT conversational interface
  - Human validation remains necessary
- Extraction of 26 articles in 380.03 minutes, representing a significant time gain for updating the Impact Pharmacie platform

Contact: jean-francois.bussieres.hs@ssss.gouv.qc.ca

Conflict of interest: None Funding: None // Poster presented at EAHP 2026

