

Brevity Without Clarity? Readability of AI-Generated PIL in Orthopaedics

Nasir Kharmal¹, Saran Gill², Chayan Shanmugaratnam¹
¹Department of Surgery and Cancer, Faculty of Medicine, Imperial College London, United Kingdom
²Faculty of Medicine, Imperial College London, United Kingdom

Introduction

Effective patient education is critical in orthopaedic care. Clear, accessible patient information leaflets (PILs) can enhance satisfaction, adherence to treatment, and outcomes¹. However, almost half of UK adults struggle to understand written health information, especially those from socioeconomically deprived groups².

Generative AI, such as ChatGPT, offers a novel, rapid method for producing PILs. While it holds promise for scalability and efficiency, its impact on readability and accessibility remains uncertain, especially when used without clinical oversight.

Aim

To evaluate the readability of AI-generated orthopaedic PILs with those produced by UK Orthopaedic societies, using validated objective metrics.

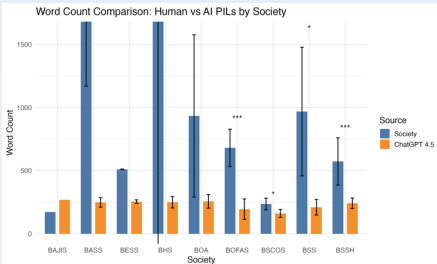
Method

A retrospective quantitative study was conducted comparing PILs from nine UK orthopaedic subspecialty societies with matched AI-generated counterparts created using ChatGPT 4.5. AI responses were generated using simple, single lined patient-style prompts to simulate real-world queries. PILs were categorised as either condition-based, procedure-based, or general information leaflets.

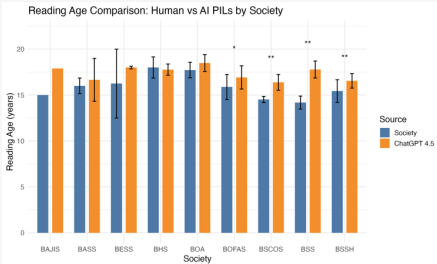
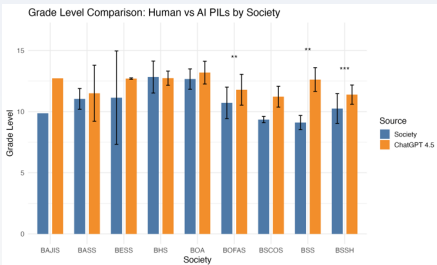
Readability was assessed using validated metrics including Flesch-Kincaid Grade Level (FKGL) and Reading Age, FORCAST, New Dale-Chall, SMOG, Gunning Fog Index, and Flesch Reading Ease (FRE).

Word counts were also analysed. Grade levels were interpreted according to U.S. educational standards. Statistical comparisons between AI and human-generated materials were performed using appropriate parametric and non-parametric tests, with statistical significance set at $p < 0.05$.

Word Count Comparison

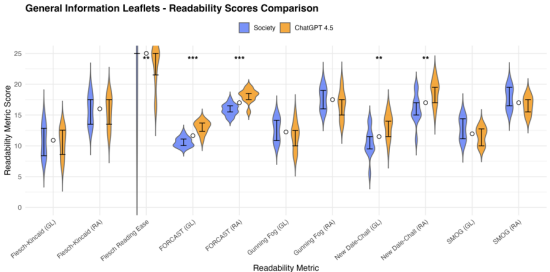
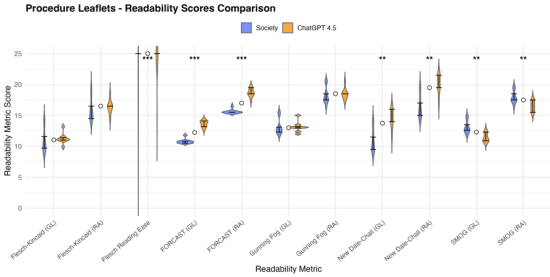
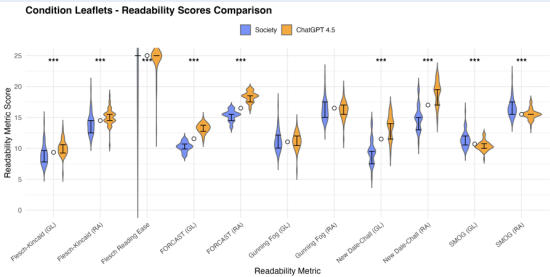


Average Grade Level and Reading Age Comparison



Results

Readability Metric Comparison by PIL Category



Reading & Grade Level Guidance

Literacy Level Related to Age & Understanding

Literacy national standard	Age and school level	Adults will be able to:	In a health setting adults will be able to:
Entry 1	Typical 7 year old	- read short texts with repeated language patterns on familiar topics - read signs and symbols, produce limited writing – only short sentences - engage in simple exchanges of information	- understand pictures on a health promotion poster - read and understand 'way in' and 'way out' signs but not 'warning' and 'next' signs - tell a clinician that they are not feeling well but not describe degrees or type of pain
Entry 2	Typical 9 year old	- read short straightforward texts on familiar topics - obtain information from familiar sources (e.g. a website, a short letter) - show some awareness of audience when writing (e.g. a short informal letter or report) - engage in discussions with familiar people	- understand the words on a simple poster such as 'Smoking is bad for you' - understand the words but not necessarily the numbers of a routine appointment letter - tell a clinician in simple language the degrees and type of pain they have
Entry 3	Typical 11 year old	- obtain information from everyday sources e.g. newspapers - communicate orally and in writing information and opinions with some adaptation to the intended audience	- understand the words or more complex posters and simple medical leaflets - understand short formal letters, notes or forms, telephone call to 111 - describe a new oral medicine and type of pain and understand what they have been prescribed by a clinician
Level 1	GCSE grade D-F	- read texts of varying lengths on a variety of topics - obtain information from different sources (simple reports, text books, work manuals) - in written communication, demonstrate an ability to express ideas and opinions clearly using length, format and style appropriate to audience and purpose (formal letter, memo, brief report etc) - be confident in oral communication - make contributions to discussions that demonstrate awareness of other views	- understand more complex information on a variety of different health related matters e.g. health promotion information on the importance of diet and simple clinical information, mediated and interpreted in a supported manner - can describe confidently degree and severity of pain - understand simple and instructions but not the importance of compliance or can clarify what they have been told
Level 2	GCSE grade A*-C	- read from texts of varying complexity accurately and independently - write to communicate information, ideas and opinions clearly and effectively using length, format and style appropriate to purpose, context and audience (e.g. complete letter, essay, report) - make a brief presentation or speak in a meeting	- understand and interpret more complex information on a variety of different health related matters e.g. health promotion information on the importance of diet and simple clinical information, mediated and interpreted in a supported manner - volunteer considered information about degree and severity of pain - ask simple questions in order to gain an understanding of the rationale behind a prescribed course of treatment

US Grade Levels & Flesch Reading Ease Interpretation

Table 21.1: United States of America

Grade Levels ^{a,b}					
Kindergarten (grade 0)					
Grades 1–8					
High School (grades 9–12)					
College (grades 13–16)					
Graduate School (grades 17–18)					
Doctorate (grade 19+)					
Reading Ease	Grade	Description of Style	Syllables per 100 Words	Av. Sentence Length	
90–100	5	Very easy	123	8	
80–90	6	Easy	131	11	
70–80	7	Fairly easy	139	14	
60–70	8–9	Standard	147	17	
50–60	10–12	Fairly difficult	155	21	
30–50	College	Difficult	167	25	
0–30	College graduate	Very difficult	192	29	

Conclusion

AI-generated PILs offer brevity but do not consistently improve readability, with some indices suggesting increased complexity, failing to meet NHS guidance on readability. While AI holds promise, clinician oversight and further validation are essential to ensure AI-generated materials enhance, rather than hinder, patient understanding and engagement.

References

- Martin LR, Williams SL, Haskard KB, Dimatteo MR. The challenge of patient adherence. Ther Clin Risk Manag. 2005;1(3):189-99.
- Rowlands G, Protheroe J, Winkley J, Richardson M, Seed PT, Rudd R. A mismatch between population health literacy and the complexity of health information: an observational study. British Journal of General Practice. 2015;65(635):e379.