

Prevalence and characteristics of plain language summaries indexed in PubMed

William Gattrell^a, Kim Wager^b, Nisha Sheikh^a, Alison Chisholm^b

^aIpsen, Abingdon, United Kingdom; ^bOxford PharmaGenesis, Oxford, United Kingdom



For further information, please send your question(s) to will.gattrell@ipsen.com
For audio commentary, please click the icon.
To download the poster, please click the icon.

Copies of this eposter are for personal use only and may not be reproduced without written permission from the authors.

Objective

- In 2021, Open Pharma issued new recommendations for plain language summaries (PLSs) for peer-reviewed journal articles.¹
- The standardized text format recommended by Open Pharma has the potential to facilitate indexing in directories, such as PubMed.¹
- We evaluated the prevalence and characteristics of PLSs among PubMed-listed articles to provide a benchmark for monitoring future PLS uptake.

Research design and methods

- We searched PubMed-listed publications (1 January 1781–20 August 2021) for text-only PLSs indexed with an XML <plain-language-summary> tag.
- Identified records were de-duplicated and screened programmatically for eligibility.
- Prevalence of PLSs identified by XML tag was compared with that of PLSs identified by PubMed search engine query. PLSs were categorized by journal, publication year, and publication-type tag (Panel 1).

Results

- Of 32,939,900 unique PubMed-listed records, 2,050 publications had XML <plain-language-summary> tags.
- Following data cleaning (the exclusion of absent PLSs and technical abstracts), 2,010 PLSs remained, giving a prevalence of 61.0/1,000,000 articles. In comparison, a PubMed search engine query indicated a PLS prevalence of 17.4/1,000,000 articles (Panel 2).
- Most indexed PLSs (98.9%) were published in the past 3 years (9.3% in 2019; 38.2% in 2020; 51.3% in 2021 (Panel 2)).
- In total, 91 journals published PubMed-indexed PLSs. Ten journals contributed 73.5% of indexed PLSs; the most frequent source was *eLife* (32.9% of all PLSs) (Panel 3).
- Of indexed PLSs, 96.8% were simply tagged as 'journal article', 10.5% had tags for clinical trials, 4.9% were tagged as 'review', and 2.4% were tagged as 'systematic review' (Panel 4).

Conclusions

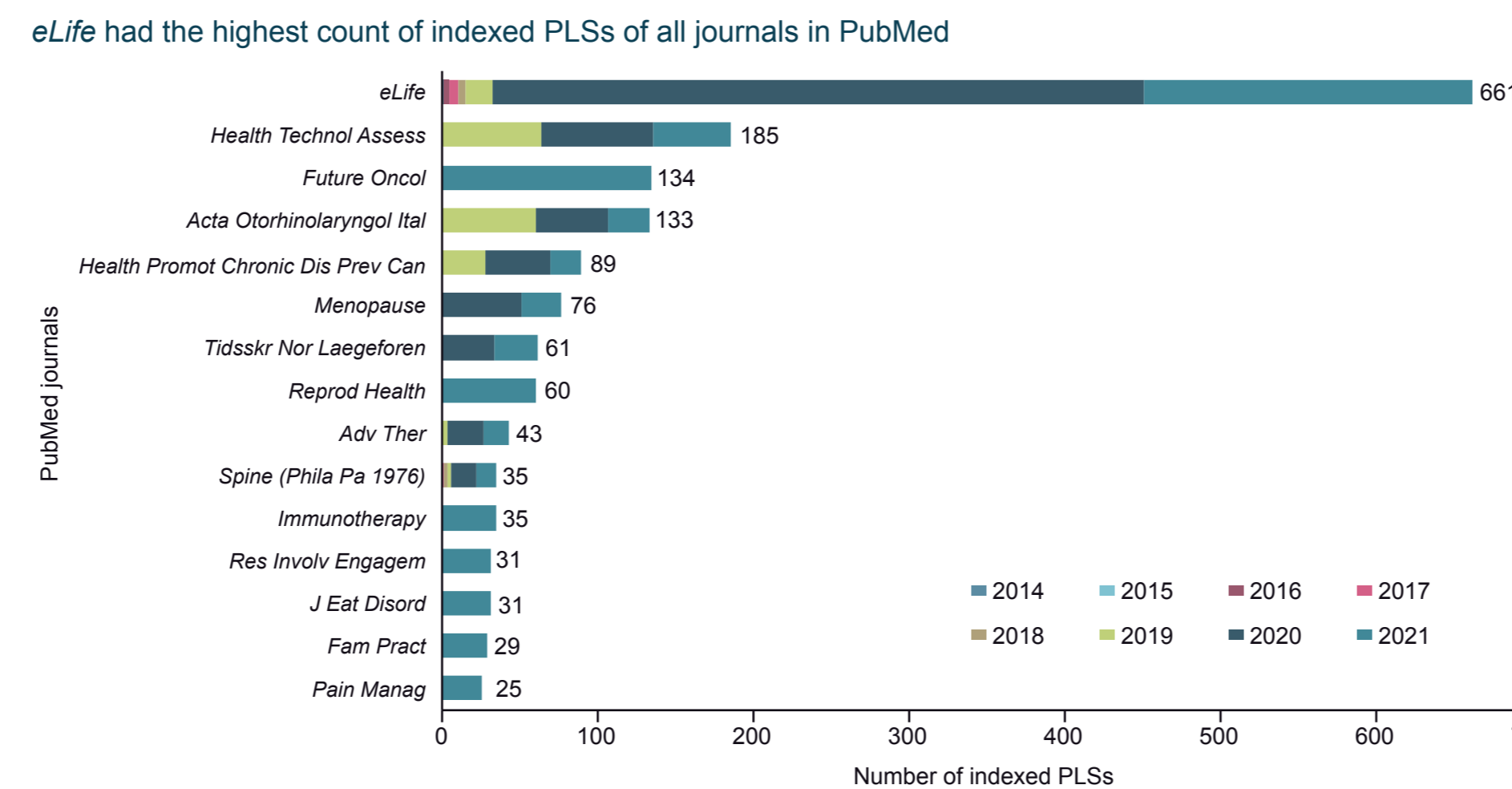
- A minority of PubMed-listed articles featured indexed PLSs. Of these, most were published between 2019 and 2021.
- Using XML tags was more than 3.5-fold more effective at identifying PLSs than interrogating the PubMed search engine (Panel 2).
- The 'journal article' tag offers limited insight into the types of research currently accompanied by indexed PLSs in PubMed.
- Collaborative efforts are required to standardize the implementation and indexing of PLSs by journals to help to align practices for PLSs and to improve their discoverability and accessibility.
- This method could be used to measure future uptake of PLSs in the literature.

TAKE-HOME MESSAGE Despite increased indexing of PLSs among PubMed articles since 2019, prevalence of indexed PLSs remains low. Collaborative efforts are required to align and optimize PLS practices to help to improve the discoverability and accessibility of articles.

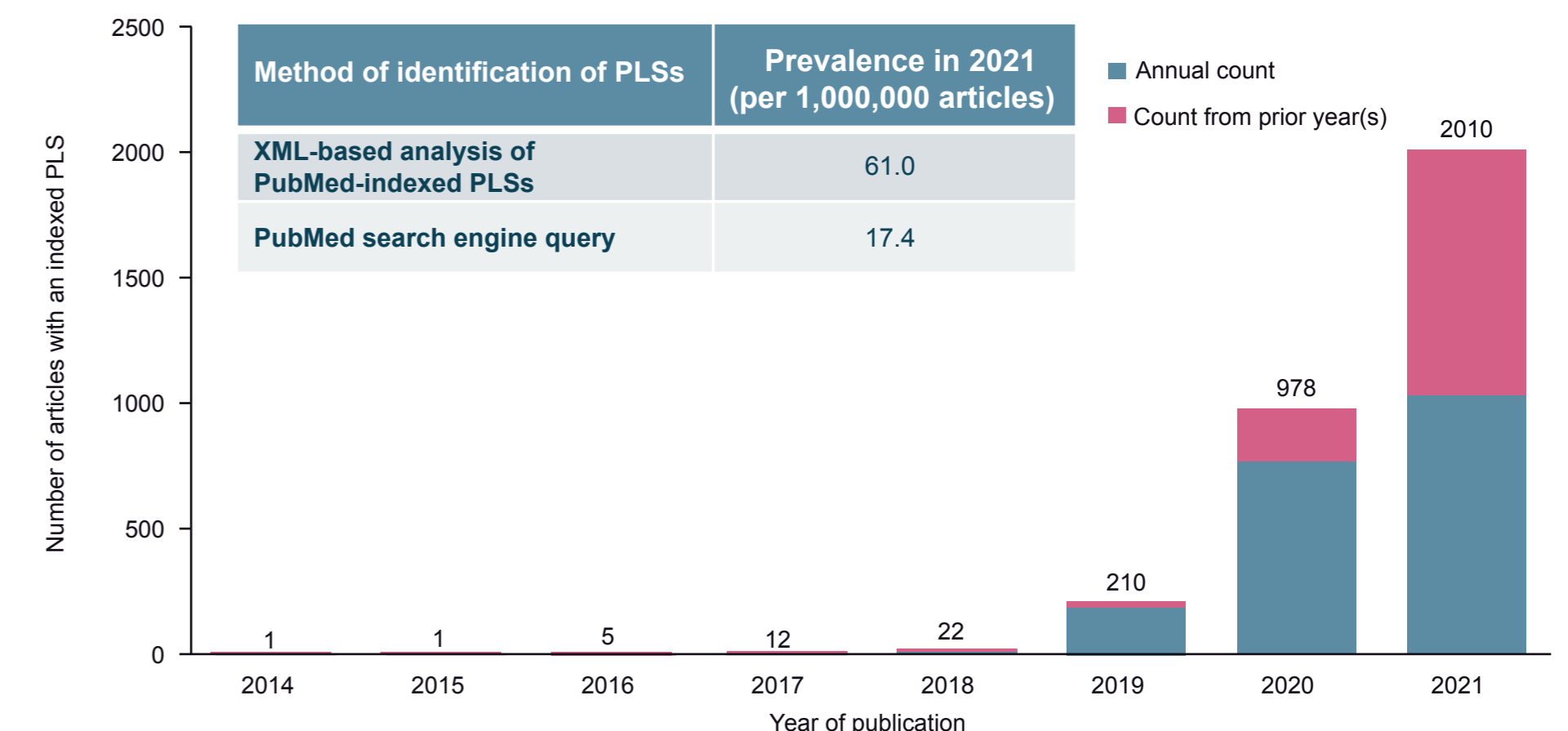
Panel 1. Methods of data analysis

- Data source:** annual baseline and daily update files were downloaded from the PubMed FTP site
- Identification of PLSs:** XML parsing, conversion to JSON, and identification of correct metatags were conducted using PHP
- Identification of unique records:** de-duplication was carried out using SQL
- Analysis:** unique files with PubMed-indexed PLSs were analyzed and the data visualized using R

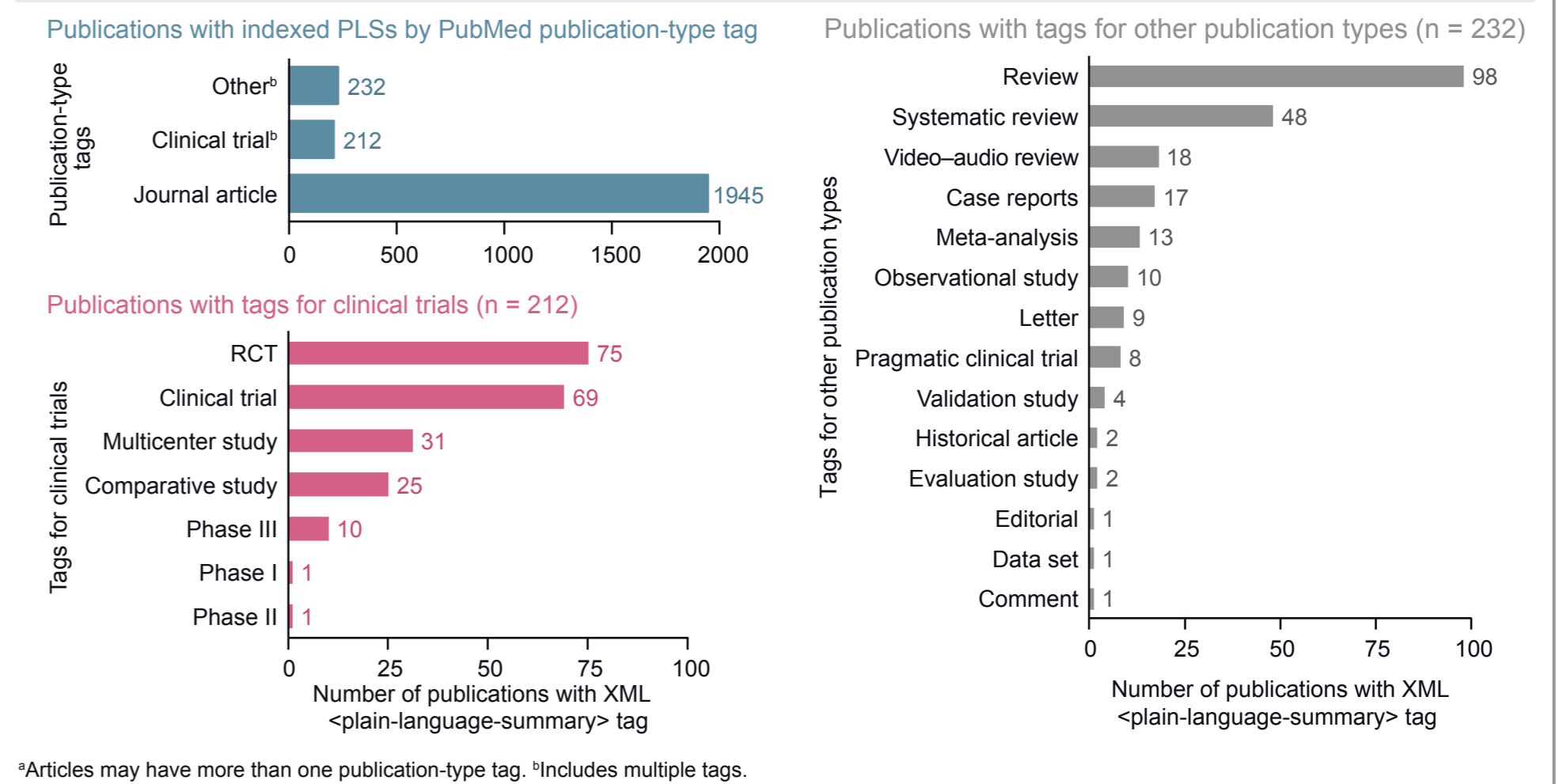
Panel 3. Journals most frequently indexing PLSs



Panel 2. Cumulative count of PubMed-indexed PLSs and prevalence in 2021



Panel 4. Breakdown of publication types with indexed PLSs^a



Abbreviations FTP: file transfer protocol; JSON: Java Script Object Notation; PHP: hypertext preprocessor; PLS: plain language summary; RCT: randomized controlled trial; SQL: structured query language; XML: extensible markup language.

References 1. Rosenberg A et al. *Curr Med Res Opin* 2021;37:2015–6.

Author Contributions Substantial contributions to study conception/design, or acquisition/analysis/interpretation of data: WG, KW, NS, AC; Drafting of the publication, or revising it critically for important intellectual content: WG, KW, NS, AC; Final approval of the publication: WG, KW, NS, AC.

Disclosures WG: Employee and shareholder of Ipsen; NS: Employee of Ipsen; KW, AC: Employees of Oxford PharmaGenesis.

Acknowledgments The authors thank Oxford PharmaGenesis who provided editorial and layout support, which was sponsored by Ipsen in accordance with Good Publication Practice guidelines.