



## Articles

# Changing roles of patient information leaflets in the UK: A corpus-assisted discourse analysis

Nicola Pelizzari <sup>\*</sup> 

University of Brescia, Department of Medical and Surgical Specialties, Radiological Sciences, and Public Health, Section of Hygiene, Epidemiology and Public Health, Viale Europa, 11, Brescia, 25121, Italy

## ARTICLE INFO

## Keywords:

Patient information leaflets  
Corpus-assisted discourse analysis  
Health communication

## ABSTRACT

Effective health communication with non-expert audiences is crucial to empower patients to make informed decisions about their medications. Patient information leaflets (PILs), which accompany medicines, aim to provide essential information about the medication, its administration, precautions, and potential side effects. The contents and objectives of PILs have changed over time, following the shift towards a patient-centered healthcare. With the aim of capturing this shift, this study uses corpus-assisted discourse analysis to analyse two corpora of British PILs across two distinct periods, 1900–1930 and 2010–2020. This paper combines quantitative and qualitative methods to investigate linguistic and pragmatic trends and the change in PILs' communicative roles. The findings reveal that early 20th-century PILs frequently relied on persuasive, promotional rhetoric with limited scientific grounding, whereas contemporary PILs prioritise accuracy, patient education, and adherence to regulatory standards. This shift reflects changes in regulatory frameworks and a move towards fostering patient autonomy.

## 1. Introduction

Patient information leaflets (PILs) are technical documents included in medicine packages, designed to deliver essential written information about medications (Herber et al., 2014). PILs serve as a primary source of information for patients regarding the safe and effective use of medications. Properly designed PILs can enhance patient understanding, ensure correct medication usage, and improve health outcomes. By providing accessible and reliable information, PILs have the potential to empower patients, fostering informed decision-making and supporting medication adherence (Askehave and Zethsen, 2002). However, when they first emerged in the early 20th century, PILs primarily functioned as marketing tools for pharmaceutical companies rather than comprehensive informational resources. This initial role was shaped by the lack of regulatory frameworks in pharmaceutical sales, resulting in an unregulated market where companies could distribute drugs without disclosing active ingredients, efficacy, or potential risks to users (Mcmanus, 2009). This often resulted in PILs that used promotional language, with little emphasis on patient safety, frequently lacking detailed guidance on dosages, contraindications, or potential side effects, reflecting an era when the focus was on promoting products rather

than fostering informed decision-making and patient safety (Shrank et al., 2007). Today, PILs have evolved into standardised, informative documents that prioritise patient safety and aim to instruct users on the correct and effective use of medications through comprehensive information. This transformation required decades of adjustments, driven by legislative developments, regulatory frameworks, and a broader shift in healthcare from a paternalistic model to one emphasising patient-centred care (Sandman and Munthe, 2010).

PILs have the potential to become an indispensable element within the healthcare continuum, acting as a critical nexus between healthcare professionals and patients. Their role extends beyond the mere provision of medication instructions; they are designed to ensure comprehensive patient understanding regarding the use, potential side effects, contraindications, and administration guidelines of prescribed medications (Wongtaweepkij et al., 2021). This level of informed awareness is essential, fostering a conducive environment for informed decision-making, thereby significantly enhancing patient adherence to treatment plans. Shrank et al. (2007) underscore the criticality of PILs, noting their direct impact on patient outcomes through improved treatment adherence and patient safety.

This multifunctionality reflects a significant paradigm shift towards a

\* Corresponding authors:

E-mail address: [nicola.pelizzari@unibs.it](mailto:nicola.pelizzari@unibs.it).

<https://doi.org/10.1016/j.acorp.2025.100129>

Received 12 October 2024; Received in revised form 28 February 2025; Accepted 17 March 2025

Available online 18 March 2025

2666-7991/© 2025 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

more patient-centered approach in medicine, a movement that places patient needs, experiences, and values at the forefront of healthcare delivery (Page-Butler, 2013). The evolution of PILs into informative documents signifies a profound change in how information is shared and utilised in healthcare settings, highlighting the shift from a passive receipt of healthcare to an active, engaged process of health management. The transformation of PILs from optional inserts to mandatory components of medication packaging underlines the growing recognition of the patient's right to be fully informed about their treatment (Oxman et al., 2020). This development is indicative of a broader societal and regulatory shift towards ensuring transparency in healthcare, safeguarding patient autonomy, and promoting informed consent as a fundamental patient right (Pathberiya, 2016; Zanola, 2019).

### 1.1. Regulatory frameworks and functions of patient information leaflets

The advent and subsequent evolution of PILs within the healthcare domain are significantly influenced by regulatory measures aimed at ensuring patient safety, promoting health literacy, and facilitating informed decision-making. These regulations, often formulated by health authorities and legislative bodies, underscore the criticality of providing patients with accessible, accurate, and comprehensive information regarding their medications. From a regulatory perspective, PILs serve as a direct response to the need for transparency in patient-medication information. The Medicines and Healthcare products Regulatory Agency (MHRA) in the United Kingdom mandates the inclusion of PILs within all medication packaging (MHRA, 2020), a requirement echoed by similar bodies worldwide (e.g. European Medicines Agency, 2023; Food and Drug Administration of the United States of America, 2023; Therapeutic Goods Administration, 2023; National Medical Products Administration, 2022). These regulatory directives are rooted in the principle of informed consent, recognising the patient's right to understand the medications they are prescribed fully. This includes not only the benefits and potential outcomes but also the risks and side effects associated with medication usage (O'Donovan et al., 2019).

The functions of PILs, delineated by these regulatory frameworks, extend beyond mere compliance. They embody a multifunctional approach to patient care, intertwining information dissemination with patient empowerment and engagement strategies. PILs aim to facilitate a dialogue between patients and healthcare providers, encouraging patients to ask questions and seek clarifications, thereby enhancing the therapeutic alliance (Garner et al., 2012).

In the UK, regulatory standards governing PILs emphasize clarity, comprehensibility, and non-ambiguity, aiming to make these documents accessible to a broad patient demographic, including those with varying levels of health literacy (MHRA, 2020). The development and review process of PILs involves rigorous testing, including patient feedback mechanisms, to ensure that the information is effectively communicated and understood. This iterative process underscores the dynamic nature of PILs, reflecting ongoing advancements in medical science, changes in healthcare delivery models, and evolving patient needs (Madison, 2010). Moreover, the regulatory standards governing PILs in the UK, emphasise the importance of clarity, comprehensibility, and non-ambiguity, aiming to make these documents accessible to a broad patient demographic, including those with varying levels of health literacy (MHRA, 2020).

Despite regulatory frameworks designed to ensure accessibility and clarity, PILs remain challenging for many readers due to persistent linguistic complexities. Research highlights significant issues in their lexical, syntactic, and discourse-level features, which often undermine their intended purpose of providing clear and actionable medical information (Bouayad-Agha et al., 2000; Maglie, 2015; Paiva, 2000). Stylistic analyses reveal that PILs often oscillate between engaging, reader-oriented styles and more abstract, impersonal tones. Paiva (2000) identified significant stylistic variability, driven by factors such as the type of medication, its intended use, and the assumptions about the reader's

role in administering the treatment. These stylistic shifts, while reflecting the intended purpose of the document, can result in inconsistencies that undermine comprehension and engagement. For instance, PILs designed for lay readers frequently incorporate features characteristic of technical or regulatory genres, creating a gap between the intended audience and the language used.

From a syntactic perspective, PILs often feature long sentences and dense constructions that increase cognitive load, making it difficult for readers to process and act on the information provided. These linguistic features, while aligning with the legal and scientific requirements for accuracy, contrast sharply with best practices for reader-centred communication (Bouayad-Agha et al., 2000; Gotti et al., 2015). Furthermore, the use of specialised terminology (Lonati, 2019) further compounds these issues. Studies highlight the overuse of technical terminology, legal expressions, and formulaic language, which prioritise regulatory compliance but reduce readability (Page-Butler, 2013; Grabowski, 2015; Gyasi, 2023; Maglie, 2015). The reliance on fixed phrase frames, as noted in corpus-based research, underscores the limited adaptability of PILs to accommodate diverse literacy levels and user needs (Pelizzari, 2024). Such rigid structures often fail to align with principles of audience-centred communication, which advocate for dynamic and context-sensitive language choices (Hegedűs, 2008).

At the discourse level, PILs often fail to fully address the needs of their target audience. Research highlights a mismatch between the formal tone of PILs and the expectations of lay readers, who may find such documents unengaging or intimidating. This disconnect is exacerbated by inconsistent efforts to integrate lay-friendly language and clear organisational structures, resulting in documents that neither fully engage readers nor enhance comprehension (Bouayad-Agha et al., 2000; Gotti et al., 2015). These linguistic features not only affect individual comprehension but also have broader implications for the usability and inclusivity of PILs. Literature therefore suggests that improving PILs requires a holistic approach that integrates stylistic adaptability, lexical simplification, and iterative user feedback (Buts et al., 2021). While progress has been made, the evolution of PILs in the UK remains an ongoing process that must find the right balance between regulatory requirements, linguistic accessibility, and the communicative needs of diverse patient populations.

### 1.2. Tracing the evolution of patient information leaflets

The current regulatory framework governing PILs in the UK reflects significant changes within healthcare, particularly in the realms of patient empowerment and access to information. The regulated state of PILs observed today represents a marked shift from earlier practices, moving from rudimentary and often perfunctory documents to detailed and regulated sources of medication information (Gotti, 2015). This transformation aligns with broader societal, regulatory, and healthcare shifts, particularly the growing emphasis on patient autonomy and the democratisation of medical knowledge (Dent, 2006).

This study investigates a specific linguistic and communicative shift by comparing two distinct points in time over the course of about a century. By analysing British PILs from the early 20th century (1900–1930) and the early 21st century (2010–2020), this study employs a mixed-methods approach, integrating quantitative corpus linguistics and qualitative discourse analysis to identify and characterise the changes that have occurred in response to differing societal, medical, and regulatory contexts. Furthermore, standardized type-token ratio (STTR) (McCarthy and Jarvis, 2010) was used to quantify shifts in lexical diversity in the respective corpora. Descriptive and inferential statistics (e.g., Cohen's D, Cliff's Delta) further validate the significance of observed patterns (Kyle et al., 2018; Gries, 2021).

This approach acknowledges that while the time points may not capture a full trajectory, they provide valuable insight into how the communicative goals and structural features of PILs have adjusted to meet new demands. From an applied linguistic perspective, examining

this shift involves exploring changes in the linguistic, pragmatic, and discourse-level features of PILs.

These analyses allow for a deeper understanding of how language use reflects broader changes in healthcare communication practices. Such a comparative approach is well-suited to reveal patterns in linguistic and communicative practices that illuminate how PILs have been adapted to address patient engagement and regulatory priorities over time (Gotti, 2015; Madison, 2010).

This study aims to determine how the communicative aims of PILs have shifted across these two periods, focusing on the implications of these changes for healthcare communication and patient engagement strategies.

## 2. Material and methods

The methodology of this study is structured around a corpus-assisted discourse analysis approach (Partington et al., 2004) that integrates the principles of corpus linguistics and discourse analysis, particularly suitable for examining linguistic and communicative shifts within a specified genre or domain, as it combines quantitative precision with contextual depth, enabling a detailed exploration of changes in language use and communicative practices within PILs. The corpus-based component provides a robust empirical foundation by quantifying linguistic features and identifying patterns and trends (McEnery and Hardie, 2012). This quantification is complemented by the discourse analysis component, which allows for an in-depth investigation of how linguistic features function within their specific social and regulatory contexts (Schiffrin et al., 2018). This dual focus on form and function enables a richer interpretation of the pragmatic and functional aspects of PILs (Gee, 2014, pp. 32–41). Additionally, critical discourse analysis (CDA), a subset of discourse analysis, is employed for its focus on power dynamics, ideology, and social change as reflected in language use (Fairclough, 2013). This perspective is particularly relevant for analysing texts that operate within institutional and regulatory frameworks, as it reveals how language choices reflect broader social and regulatory priorities. While this study does not examine continuous language evolution, the focus on two key temporal snapshots highlights significant shifts in PILs, offering insights into how these documents have been adapted to meet changing societal, medical, and regulatory demands.

### 2.1. Corpus compilation

To explore the shift of healthcare communication and societal expectations, this study compares PILs from two distinct historical periods. To establish an ideal timeframe, we identified the earliest available English-language PILs from the United Kingdom. This step allowed to provide a historical reference point for comparison, enabling an analysis of how PILs have evolved in response to changes in healthcare communication practices and societal expectations over time. This required an initial enquiry to the Royal Pharmaceutical Society of Great Britain in London, followed by a formal archival research application, to access their collection of PILs.

The Royal Pharmaceutical Society Museum (RPSM), established in 1842, preserves and showcases the history of British pharmacy. In July 2024, access was granted to the RPSM archive, which houses a collection of historical PILs (Access identification RPSM 104/2024). To identify relevant documents, we consulted the archive's metadata records cataloging its physical holdings. However, as the PILs are not digitized, manual verification of the physical documents was required. The search revealed that the earliest available PILs were very limited in number, largely due to the non-compulsory nature of these documents in the early 20th century. At that time, medication information was typically provided by pharmacists or doctors directly, or printed on the packaging rather than in a dedicated leaflet (McManus, 2009). Additionally, preservation challenges further reduced the availability of these documents, as only a fraction of printed materials from that period have

survived intact to the present day.

The selection of PILs for the early corpus was guided by predefined inclusion criteria to ensure consistency and comparability across the dataset. The inclusion criteria ensured that only documents meeting key requirements were selected for the corpus. These included the presence of text explicitly providing therapeutic indications, clear identification of the country of origin (the UK), and a general explanation of the medication's purpose or function. Documents that lacked these essential features, such as those containing only images or visual elements without accompanying textual information, were excluded. Using these selection criteria, we identified approximately 30 PILs from 1900 to 1930 that met the requirements for inclusion in the study. The selection of 1900 as the starting point for Corpus A was determined based on the earliest available PILs that met the predefined inclusion criteria. PILs produced after 1930 exhibited notable differences, including greater standardization of their layout and a broader range of information addressing dosage, side effects, and contraindications. To maintain focus on the earliest available examples and ensure comparability, the selection was limited to PILs that exemplified the genre in its formative stage.

To represent the contemporary landscape, the second corpus consists of PILs published between 2010 and 2020. This timeframe was chosen to ensure the inclusion of the most up-to-date documents. The period considered for the twentieth-century corpus is broader than that of the twenty-first century due to the difficulty of finding early versions of these documents. This challenge, however, does not apply to the more recent corpus, which consists of readily available and contemporary documents. The documents included were obtained through the Electronic Medicines Compendium (EMC), a publicly accessible and free database providing comprehensive access to PILs of medicines currently on the market (Datapharm, 2024). The availability of these documents through the EMC, reflects the significant regulatory evolution over the past century. This regulatory transformation underscores the importance of analyzing PILs across different time periods to understand how linguistic and structural changes have shaped patient communication.

Once the time-lags in the two corpora were defined, a set of inclusion criteria were established to ensure high comparability of the corpora, which is essential for valid linguistic analysis, adhering to the methodological practices of corpus linguistics that advocate comparability to accurately assess language use in different contexts (Baker, 2006; Brezina, McEnery, and Gablasova, 2021). The first criterion focuses on the geographic and marketing context, requiring the inclusion of PILs for drugs produced and marketed in the United Kingdom. This condition is met only if the leaflet explicitly states that the drug manufacturer is based in the UK and that the drug is distributed by a UK pharmacy or retailer. Additionally, practical criteria were applied to ensure the documents' usability: each PIL had to be intact, clearly legible, and consultable in its entirety. In order to achieve maximum uniformity, the documents had to reflect the macro-function of the PIL, that is, to provide information about the drug and be directed to a patient audience. Addressing the historical absence of a clear distinction between prescription-only and over-the-counter medications, this study's methodology pivots to the selection of medications based on therapeutic aims. Specifically, the chosen PILs represent drugs that, across both the historical and contemporary corpora, target analogous diseases or ailments, therefore focusing on areas with comparable therapeutic aims, such as pain relief and inflammation management (e.g., analgesics), gastrointestinal health (e.g., laxatives and antispasmodics), neurological disorders (e.g., anticonvulsants), and respiratory health (e.g., antitussives in the early corpus and mucolytics or bronchodilators in the contemporary corpus).

These criteria enabled the composition of two corpora comprising 30 PILs each: the early 20th-century corpus and the early 21st-century corpus. The documents from the former have been duly transcribed and proofread. Documents from the latter, available in digital format, were converted to cleaned-up PDF format. Once ready, the two corpora were entered into AntConc - Version 4.2.4 (Anthony, 2023) for the

quanti-qualitative analyses object of this study.

### 3. Results and discussion

#### 3.1. Quantitative analysis

As mentioned in the methodological section (2.2), this study analyses two corpora, each comprising a collection of 30 patient information leaflets grouped according to the following criteria:

- 1). The first corpus contains 7,164 words from 30 British PILs of medicines sold in the UK between 1900 and 1930 (henceforth, Corpus A).
- 2). The second corpus contains 71,976 words from 30 British PILs medicines sold in the UK between 2010 and 2020 (henceforth, Corpus B).

Data concerning the PILs constituting the respective corpora, indicating their temporal distribution and the type of medication they describe can be found in the supplementary materials.

Table 1 presents the comparative analysis of lexical diversity, measured using the standardized type-token ratio (STTR). The type-token ratio (TTR), which expresses the ratio of unique words (types) to the total number of words (tokens) in a text, is a widely used measure of lexical diversity (Kettunen, 2014; Malvern et al., 2004; Richards, 1987). However, TTR is sensitive to text length, as shorter texts tend to yield inflated values (Covington and McFall, 2010). To address this, STTR normalizes text length by calculating the average TTR across fixed-length segments, typically 100 tokens, enabling comparisons across texts of different sizes (McCarthy and Jarvis, 2010). In this study, each text was divided into non-overlapping segments of 100 tokens with WordSmith tool 9 (Scott, 2020). The TTR was calculated for each segment, and the average TTR across all segments was computed to obtain the STTR value for each text.

A *t*-test was conducted to assess statistical significance, and Cohen's *d* was calculated to determine the practical significance of the differences (Brezina, 2018). The results show that Corpus A uses a much wider variety of words (38.12 %) compared to Corpus B (28.04 %). This difference is not only very unlikely to be due to chance (indicated by the extremely low *p*-value, < 0.0001) but also large enough to be meaningful in practical terms (Cohen's *d* = 4.36, which indicates a very strong effect). The magnitude of Cohen's *d* (4.36) far exceeds conventional benchmarks (Cohen, 1988), yet it is consistent with corpus linguistic studies of bounded metrics like STTR. A value greater than 1.0 simply means that the difference between the two group means is larger than the pooled standard deviation.

Two interrelated factors explain this inflation: (1) sample sizes reduce variability, amplifying standardized effects even for modest raw differences (Larson-Hall and Herrington, 2010), and (2) STTR's bounded range (0–100 %) inherently constrains variability, shrinking the pooled standard deviation and inflating the ratio of mean difference to variability (McCarthy and Jarvis, 2010). Similar patterns are observed in studies of syntactic complexity and readability, where bounded metrics in corpora yield inflated effect sizes (Kyle, 2021; Plonsky and Oswald, 2014). For instance, Gries (2021) notes Cohen's *d* > 2.0 is common in diachronic corpus comparisons, while Egbert et al. (2023) report *d* = 4.1

**Table 1**  
Descriptive statistics of lexical diversity between the corpora.

Corpus	Mean STTR (%), Mean Standard Deviation	Token (Word count)	Types (Unique words)
Corpus A (1900–1930)	38.12 ± 2.50	7,164	2,731
Corpus B (2010–2020)	28.04 ± 1.80	72,257	20,262

in 19th-century newspaper analyses.

To contextualize practical significance, we emphasize the raw mean difference of 10.08 % (26.4 % relative increase) and validate it with Cliff's Delta ( $\Delta = 0.55$ ), a non-parametric measure robust to distributional assumptions (Pearson et al., 2010), which indicates a large practical effect. This supports the raw mean difference by confirming that the observed STTR divergence is not an artifact of parametric assumptions (e.g., normality) or outliers. It reinforces that the effect is robust and practically meaningful. Table 2 presents the inferential analysis of lexical diversity between Corpus A (1900–1930) and Corpus B (2010–2020), highlighting statistical significance.

This suggests that the early 20th-century PILs were much more diverse in their choice of words, while the modern ones are more standardized and repetitive, likely reflecting changes in communication goals.

As frequent words are often indicative of the functional and thematic focus of a text, reflecting core linguistic and communicative practices within a specific domain (Brezina, 2018; Biber et al., 1999). The most recurrent lexical words in both corpora are listed in Table 3. We decided to focus on the 15 most frequent words to ensure a balance between representativeness and analytical clarity (Brezina, 2018). The raw word frequencies for both corpora were statistically normalized to frequencies per 1000 words to account for differences in corpus size, ensuring comparability across datasets. Normalization was achieved by dividing each word's frequency by the total word count of its respective corpus and multiplying the result by 1000 (Brezina, McEney, and Gablasova, 2021) Even after the normalization process, the word frequency did not change.

Table 3 shows raw and normalised frequencies of the most recurrent lexical terms in both corpora.

Limiting the analysis to 15 items allowed for an in-depth examination of these central features without diluting the analysis with less significant or less frequent words.

The most frequent terms in Corpus A are *pills* (*nf* = 5.58) and *remedy* (*nf* = 4.05). *Pills* explicitly denotes the physical form of medication. In contrast, *remedy* functioned as a more ambiguous term, often encompassing a broader range of therapeutic substances. The terms *teaspoonful* (2.51) and *bottle* (2.51) reflect the use of household items for measurement, which contrasts with the precise and standardised units seen in modern practices. Additionally, the inclusion of *bishop* (2.37), as part of a brand name, underscores the practice of using recognizable or culturally resonant branding to promote trust in products. This language reflects a period when medical communication was less formal and standardised, relying on familiar and relatable terms to connect with patients.

In contrast, Corpus B demonstrates a clear shift toward regulatory compliance. Terms like *doctor* (10.56) and *pharmacist* (4.92) emphasise the central role of healthcare professionals in guiding patient care. The frequent appearance of *mg* (5.56) and *tablets* (8.59) points to the standardisation of medical language, particularly in describing medication forms and dosages. These changes reflect advancements in pharmacology and a growing emphasis on clear and accurate instructions, driven by regulatory requirements and the need for patient safety (EMA, 2023; MHRA, 2020).

Both corpora highlight the importance of instructions, with the term *take* appearing prominently in each, though it occurs more frequently in

**Table 2**  
Inferential statistics of lexical diversity between the corpora.

	Mean Difference ( $\Delta M$ )	<i>t</i> -test	<i>p</i> -Value	Cohen's <i>d</i>	Cliff's $\Delta$
Comparison (Corpus A and B)	10.08 % (26.4 % relative increase)	<i>t</i> (58) = 16.88	<i>p</i> < 0.001	4.36 (Large effect)	0.55 (Large effect)

**Table 3**  
Raw and normalized frequencies of most recurrent terms in the corpora.

CORPUS A			CORPUS B		
Word	Raw Frequency	Normalized Frequency	Word	Raw Frequency	Normalized Frequency
Pills	40	5.58	doctor	760	10.56
Remedy	29	4.05	tablets	618	8.59
One	28	3.91	have	590	8.19
years	27	3.77	medicine	549	7.63
take	25	3.49	side	507	7.05
water	21	2.93	take	478	6.64
cold	19	2.65	effects	469	6.51
two	19	2.65	blood	420	5.54
bottle	18	2.51	mg	400	5.56
teaspoonful	18	2.51	pharmacist	354	4.92
bishop	17	2.37	dose	344	4.78
rheumatism	17	2.37	used	327	4.54
throat	17	2.37	do	325	4.51
children	16	2.23	can	301	4.18
dose	16	2.23	use	292	4.05

Corpus B (nf = 6.64) compared to Corpus A (nf = 3.49). In addition to the prominence of *take*, Corpus B demonstrates a broader range of verbs with significant frequencies, such as *have* (8.19), *use* (4.05), and *do* (4.51). These verbs expand the communicative scope of modern PILs, addressing diverse aspects of patient engagement: *take* remains central for medication administration, while *have* and *do* suggest potential queries, actions, or recommendations, and *use* conveys broader application or contextual instructions. This variety reflects the evolved role of PILs in facilitating patient interaction and understanding.

By contrast, Corpus A centers its communication heavily around *take*, reflecting a singular focus on adherence to medication instructions. The limited range of verbs suggests that other aspects of patient engagement, such as broader guidance or interaction, were secondary considerations. This narrower lexical focus aligns with the priorities of the period, where PILs aimed to deliver instructions on administration, leaving little room for the more expansive patient education seen in Corpus B. Similarly, the term *dose* appears more often in Corpus B (4.78) than in Corpus A (2.23), reflecting the shift from informal language, such as *teaspoonful*, to standardised and professionalized descriptions of dosage (MHRA, 2020).

One of the most striking differences between the two corpora is the focus on safety and risks. In Corpus B, terms like *side* (7.05) and *effects* (6.51) emphasise the importance of disclosing potential adverse outcomes, which aligns with modern regulatory frameworks that mandate transparency in healthcare communication. These terms are absent in Corpus A, where the language tends to prioritise therapeutic benefits without addressing potential risks. This difference reflects a significant change in the expectations placed on pharmaceutical communication over time.

These lexical differences also reveal broader implications for health literacy. Corpus A reflects an era when PILs served primarily to provide basic instructions, catering to a population with limited access to detailed medical information. The language is simple and accessible, emphasizing ease of understanding. In comparison, Corpus B reflects a shift toward empowering patients with comprehensive, detailed, and regulated communication. Modern PILs aim to support active participation in healthcare decisions, reflecting higher expectations for health literacy and patient engagement (Raynor et al., 2007; Rudd et al., 2023).

The increase in word count from 7,164 in Corpus A to 71,976 in Corpus B reflects the modern necessity to provide extensive information covering all aspects of medication use, from dosage and administration to side effects and contraindications (Koo et al., 2007). This comprehensive approach ensures that patients are well-informed about their treatments, enhancing safety and efficacy.

### 3.2. Qualitative analysis

In this qualitative analysis, we examine the evolving linguistic and communicative attributes of PILs across the two corpora. The focus is on high-frequency lexical items and their surrounding contexts, exploring how these terms function pragmatically within the broader discourse of each period. To achieve this, the investigation employs a combination of collocation analysis, concordance analysis, and contextual reading, enabling the systematic exploration of recurring word pairings, thematic patterns, and rhetorical strategies. These methods shed light on how PILs reflect their respective socio-cultural and regulatory environments. Concordance analysis, in particular, further contextualised the collocations by systematically examining every occurrence of high-frequency terms within the corpus to identify recurring patterns and pragmatic functions. This approach facilitated the identification of broader thematic trends, rather than isolated or anecdotal examples, ensuring that the findings are representative of the corpus as a whole.

#### 3.2.1. CORPUS a – early 20th century PILs (1900–1930)

**3.2.1.1. Pills.** The analysis of the contexts in which the term *pills* (nf = 5.58) appears, often reveals a recurrent association with hyperbolic claims and psychologically manipulative marketing strategies in early 20th-century PILs. Products were described in hyperbolic terms to captivate consumers and emphasise their supposed efficacy.

*These pills are the greatest blessing to humanity.*

*These pills are the most wonderful and valuable remedy ever discovered.*

*Bishop's pills are a distinct pharmaceutical triumph.*

*These pills dissolve and expel tumours.*

Such claims reflect a time when pharmaceutical marketing prioritised attention-grabbing language over scientific rigour. The dramatic assertion that “These pills dissolve and expel tumours” exemplifies this trend, despite the modern understanding that treating cancer is a highly complex and challenging endeavour (American Cancer Society, 2020; Siegel et al., 2023). Furthermore, PILs frequently incorporated refund policies designed to instil confidence in the consumer.

*Every druggist is instructed to refund the price while you wait at the counter if you don't feel relief coming in two minutes from taking the pills.*

*After taking the pills, a refund for unsatisfied customers.*

These strategies were economically and psychologically impactful, driving sales while fostering a sense of security among consumers (Courtwright, 2001). The promise of a money-back guarantee added a layer of reassurance, positioning the products as both effective and

risk-free. Such marketing creativity highlights the unregulated environment of the period, contrasting sharply with the requirements of modern pharmaceutical communication, where claims must be scientifically validated (Cialdini, 2007).

3.2.1.2. *Remedy*. The qualitative analysis of the term *remedy* (nf = 4.05) builds upon its prominence in the quantitative findings, which highlighted its frequent use in Corpus A. Collocation analysis showed that *remedy* co-occurred with positive evaluative adjectives such as *safe* (nf = 0.4), *effective* (nf = 0.3), and *perfect* (nf = 0.3). These modifiers framed the term as reliable and universally beneficial, aligning with the promotional tone characteristic of early 20th-century PILs. For instance, typical collocations included:

*A safe remedy for all ages.*

*The most effective remedy for coughs and colds.*

*The perfect remedy for your family.*

Such patterns emphasised emotional reassurance and consumer appeal, portraying the product as a trusted solution rather than a specific treatment backed by evidence. Concordance analysis further contextualised these collocations, revealing that *remedy* was often framed as a universal cure endorsed by vague authority figures or institutions. For example, concordance lines included:

*This remedy has been highly commended by leading physicians.*

*No family should be without this remedy.*

These instances underscore the reliance on vague endorsements to enhance credibility without providing detailed or verifiable scientific data. A closer reading of representative texts confirmed this pattern:

*Had this lady only known it, one delicious little glass of this remedy, greatest discovery of modern times, highly commended by the Lancet and British Medical Journal, and which has the power of instantly curing headache and invigorating the tired body or the overworked brain, Bishop's Citrate, she would have instantly cured her headache.*

This hyperbolic description demonstrates how the term *remedy* was embedded in broader rhetorical strategies aimed at persuasion, leveraging endorsements and exaggerated claims to instill trust. The text does not merely inform but seeks to persuade by appealing to the reader's emotions - specifically, hope, trust, and a desire for relief from discomfort. This is achieved through the use of language that is evocative, exaggerated, and relatable, constructing a compelling narrative around the product's purported benefits. This approach aligns with Hyland's (2010) exploration of metadiscourse, which emphasises how writers use emotionally charged language to engage and persuade their audience.

The use of testimonials is another prominent feature in Corpus A, further enhancing the persuasive appeal of the term *remedy*. These testimonials presented unverifiable accounts from anonymous users, claiming substantial benefits without offering any context regarding the credibility of the individuals quoted (Hovland et al., 1951).

*I have found much benefit from taking Brown's Universal Remedy for Rheumatism.*

*I have been a great sufferer for years, but since using this remedy, I have always felt relief after a dose or two and feel it is my duty to make it known.*

*I have received so much benefit from this remedy that if I ever have pain again, this will be the first medicine I turn to.*

Such testimonials exemplify how PILs used anecdotal evidence to position the products as universally effective, leveraging personal narratives to instill trust and appeal to the reader's emotions. However, these accounts lacked substantiation or scientific validation, relying

instead on their relatability and simplicity to persuade consumers.

Endorsements from prominent figures, such as royalty, politicians, and respected professionals, further bolstered the credibility. This technique, known as the appeal to authority, was a common strategy to persuade consumers by associating products with influential individuals or institutions.

*This remedy has been highly commended by members of the Royal Household.*

*This remedy was supplied to The Archbishop of Canterbury, H.R.H. The Duke of Kent, and His Eminence Cardinal Manning.*

These endorsements elevated the perceived reliability and desirability of the remedy, embedding it within a broader narrative of trustworthiness. Rowe & Frewer (2005) discuss how vague references can be used to enhance the perceived credibility of a text without providing substantial evidence. This approach creates an impression of reliability and trustworthiness, leveraging the reputation of the cited sources to persuade the reader. The appeal to authority was a well-known persuasive technique during this period, where communicators relied on the reputation of authoritative sources to support their claims (Cummings, 2015).

3.2.1.3. *Teaspoonful*. Collocation analysis showed that *teaspoonful* appeared alongside vague directives and quantities, such as a scant *teaspoonful* (nf = 0.4), one *teaspoonful* (nf = 0.53), and two *teaspoonfuls* (nf = 0.43). These collocations suggest that early PILs relied heavily on familiar household measures, likely aiming to make dosage instructions accessible to a general audience. However, this approach also highlights a lack of precision in conveying medically accurate information.

*Take one scant teaspoonful in a glass of water.*

*Administer two teaspoonfuls after each meal.*

*Take one teaspoonful of the remedy, following the maxim 'little but often'.*

Concordance analysis further contextualised the term, illustrating its use in broader instructions that often lacked specificity or scientific grounding.

*Take one teaspoonful as needed.*

*One teaspoonful every few hours.*

These examples reflect the absence of standardised measurement practices that are now central to modern pharmaceutical communication. The reliance on vague and subjective language underscores the era's prioritisation of simplicity over accuracy. A representative example from Dr. Peebles' epilepsy remedy PIL illustrates this trend.

*Take one scant teaspoonful in one-fourth of a tumbler of water a short time after each meal.*

This imprecision, combined with the lack of specific timing or quantified intervals, highlights the broader risks associated with informal dosing practices in early 20th-century PILs.

3.2.1.4. *Bottle*. *Bottle* appeared alongside phrases that sufficiency, reliability, and authenticity.

*This bottle contains enough remedy for two weeks.*

*A bottle a day keeps the symptoms at bay.*

*Insist on the original bottle.*

Concordance analysis contextualised these collocations, revealing a strong emphasis on the term *bottle* as both a practical container and a symbol of product authenticity. For instance, some PILs stressed the

need to avoid counterfeit products by recognising the bottle.

*Beware of imitation bottles. Look for our name embossed on the glass.*

*Beware of spurious bottles.*

This recurring theme suggests that they were marketed not only as a functional object but also as a visual guarantee of quality, helping to build trust in the product. A closer reading of excerpts illustrates how the term bottle functioned within promotional narratives.

*Each bottle of Dr. Smith's Elixir is sufficient for 14 days of treatment, ensuring results with every dose.*

*One bottle is all you need to experience noticeable improvement and regain your vitality.*

Here, the emphasis on sufficiency and guaranteed outcomes positions the bottle as an integral part of the remedy's perceived efficacy.

**3.2.1.5. Dose.** *Dose* (nf = 2.23) is found associated with vague or subjective language, highlighting the imprecision of early dosing instructions and the reliance on patient discretion.

*A small dose as needed.*

*Increase the dose gradually until relief is achieved.*

Additionally, close reading highlights that the lack of standardisation in dosing was compounded by the limited information on the composition of these medications. Literature confirms that formulations were often kept confidential, and when ingredient information was disclosed, it revealed that many medicines contained highly addictive substances such as cocaine, opioids, and alcohol (Porter, 1989). These ingredients posed significant health risks, especially to children. At the time, they were neither illegal nor regulated, leading to widespread abuse and addiction (Lomax, 1973). This is corroborated by evidence found in the corpus.

*A combination of soluble ingredients, which are put up in a small ball about two inches in diameter. These ingredients are all of a very healing character, and are so light that they readily dissolve in the air.*

*Vinacola is merely a solution of cocaine, which cannot be too severely condemned.*

To sum up, the analysis of early 20th-century PILs reveals a landscape dominated by persuasive and promotional language, strategic use of authority and testimonials, exaggerated claims, aggressive marketing strategies, and significant omissions of crucial safety information, highlighting the challenges and risks associated with the largely unregulated pharmaceutical market of that era, and emphasizing the need for rigorous scientific validation and transparent communication in the promotion of medical products.

### 3.2.2. Corpus b – early 21st century (2010 and 2020)

The analysis of Corpus B reveals significant advancements and shifts from the early 20th-century counterparts, particularly in the use of more precise and domain-specific medical vocabulary. High-frequency words such as *doctor* (nf = 10.56), *tablets* (nf = 8.59), *medicine* (nf = 7.63), and *dose* (nf = 4.78) reflect a broader emphasis on clarity, patient understanding, and effective communication in healthcare contexts. These terms underscore the modern priorities of regulatory compliance and the need for detailed, patient-centered communication.

**3.2.2.1. Doctor and pharmacist.** The prominence of *doctor* and *pharmacist* in the corpus highlights the central role of healthcare professionals in guiding patients.

*Consult your doctor or pharmacist if you have any questions about this medication.*

*Your doctor will determine the appropriate dose for your condition.*

*Ask your pharmacist for advice if you are unsure.*

This directives emphasise the importance of professional oversight, aligning with contemporary healthcare models that promote shared decision-making and patient engagement (Koo et al., 2007).

**3.2.2.2. Tablets.** Similarly, the frequent use of *tablets* (nf = 8.5) reflects the shift toward standardised medication forms and dosages.

*Take two 5 mg tablets once daily.*

*Swallow the tablets whole with water. Do not crush or chew the tablets.*

The structured and systematic presentation of this information is intended to help patients adhere to dosing protocols more safely and effectively, aligning with the modern emphasis on patient safety and adherence.

**3.2.2.3. Medicine.** The term *medicine* appears frequently in the corpus (nf = 7.6), suggesting its importance within the language of modern PILs. Unlike the promotional and vague usage observed in early 20th-century texts, *medicine* in contemporary PILs seems to be framed as a vital and manageable aspect of everyday healthcare. Instructions are often presented in patient-friendly contexts, with an apparent focus on using clear and accessible language to help patients engage with their treatment more effectively.

*This medicine helps to lower your blood pressure and reduce your risk of heart attack or stroke.*

*If you experience symptoms such as a rash or difficulty breathing, stop taking this medicine and seek medical advice immediately.*

This straightforward guidance appears to aim at simplifying adherence while reflecting an awareness of varying levels of health literacy among patients described in literature (Rudd et al., 2023). The language used in modern PILs seems to emphasise the importance of prompt action and professional consultation when side effects occur. By presenting *medicine* as an accessible and supportive tool, contemporary PILs appear to prioritise patient understanding and engagement.

**3.2.2.4. Side and effects.** The frequent appearance of *side* (nf = 7) and *effects* (nf = 6.5) in modern PILs suggests an increased emphasis on transparency and patient safety. This shift is evident in the structured presentation of potential risks, with many PILs clustering side effects into categories such as “Possible side effects” and “Serious side effects.”

*Possible side effects may include drowsiness or dry mouth.*

*This medicine may cause side effects, but not everybody gets them.*

This detailed approach reflects the influence of modern regulatory frameworks, which require the disclosure of potential adverse outcomes in a clear and systematic manner. By addressing side effects explicitly, these PILs aim to provide patients with information about risks and encourage timely action when necessary (Edwards, 2004; Vander Stichele, 2004).

**3.2.2.5. Take, use, can.** Words such as *take* (nf = 6.6), *use* (nf = 4), and *can* (nf = 4.1) exemplify the directive tone characteristic of modern PILs. These verbs function pragmatically to deliver actionable guidance, enabling patients to manage their treatments effectively while supporting a sense of autonomy in healthcare decisions.

*Take one tablet with a glass of water every 6 hours as needed.*

*You can crush the tablet if you have difficulty swallowing.*

*Use this medication exactly as prescribed by your doctor.*

These examples highlight the interplay between directive language

and patient empowerment. By blending precise instructions with an emphasis on safety and responsibility, modern PILs align with contemporary medical discourse, which prioritises clarity, shared accountability, and informed decision-making in healthcare practices.

The increase in word count from early 20th-century PILs to those of the 21st century reflects the modern necessity to provide extensive information covering all aspects of medication use, from dosage and administration to side effects and contraindications. This comprehensive approach ensures that patients are well-informed about their treatments, enhancing safety and efficacy. The language characteristics of the PILs contained in corpus B demonstrate a significant shift towards a more technical, precise, and patient-centered approach. The use of structured formats, detailed safety information, clear dosage instructions, and patient-friendly language reflects contemporary priorities in healthcare communication, emphasizing patient education, safety, and engagement. These characteristics highlight the evolution of PILs into comprehensive tools designed to support informed decision-making and active participation in healthcare, aligning with the broader trends in modern medical practice and regulatory standards (European Medicines Agency, 2023; Raynor et al., 2007).

#### 4. Conclusions

The corpus-assisted discourse analysis of PILs from the early 20th and early 21st centuries illuminates significant transformations in healthcare communication practices, reflecting broader societal, regulatory, and technological changes. This study highlights the evolution from rudimentary, often persuasive, and promotional tools to more standardized, informative, and patient-centered documentation. This transition is critical for understanding contemporary health communication's role in promoting patient autonomy and safety. Early 20th-century PILs primarily employed persuasive language, with heavy reliance on testimonials, endorsements from authoritative figures, and exaggerated claims about the efficacy of medications. This language was designed to persuade patients to purchase and use medications, often without providing comprehensive information on dosage, side effects, or contraindications. The frequent use of vague references to esteemed publications and the strategic deployment of testimonials underscored the promotional nature of these documents, as they sought to build credibility through association rather than through detailed scientific evidence (Coulter, 1998). In contrast, PILs from the early 21st century exhibit a marked shift towards clarity, precision, and comprehensiveness. The language used in these documents is more technical and standardized, reflecting regulatory requirements aimed at ensuring that patients receive accurate and complete information. Modern PILs include detailed sections on drug mechanisms, dosage instructions, potential side effects, interactions with other medications, and warnings for specific patient populations (Raynor et al., 2009). This shift towards informative language aligns with the increasing emphasis on patient education and autonomy in healthcare, supporting informed decision-making and adherence to treatment plans (Atal et al., 2018). The transformation in the linguistic and pragmatic features of PILs can be largely attributed to evolving regulatory frameworks. Early 20th-century PILs were produced in an era with minimal regulatory oversight, allowing for a wide range of rhetorical strategies aimed at persuasion rather than information dissemination (Herber et al., 2014). The lack of standardised requirements for the content and presentation of these documents meant that patients often received incomplete or misleading information, potentially compromising their safety. In contrast, the modern regulatory environment mandates the inclusion of comprehensive and clear information in PILs. Agencies such as the Medicines and Healthcare products Regulatory Agency (MHRA) in the UK have established guidelines to ensure that PILs are accessible, accurate, and comprehensive, promoting transparency and patient safety (MHRA, 2020). These regulatory measures are rooted in the principles of informed consent and patient rights, reflecting a broader societal shift

towards greater patient involvement in healthcare decisions.

The changes observed in PILs also mirror broader socio-cultural and ideological shifts in society. The early 20th century was characterised by a paternalistic approach to healthcare, where medical professionals were viewed as the primary authority figures, and patients were expected to follow their advice without question. This power asymmetry is evident in the promotional and authoritative tone of early PILs, which often omitted critical safety information and relied on endorsements and testimonials to build trust (Coulter, 1998). The early 21st century, however, has seen a move towards a more patient-centered approach in healthcare. This shift is reflected in the design and content of modern PILs, which prioritise patient understanding and engagement. The detailed and structured presentation of information in contemporary PILs supports the goal of empowering patients to make informed decisions about their treatment. This approach aligns with the concordance model of medicine taking, which emphasises the importance of involving patients in decisions about their care to improve adherence and therapeutic outcomes (Atal et al., 2018). Moreover, the growing focus on health literacy has influenced the design of modern PILs, thus reflecting a commitment to reducing health disparities and promoting equitable access to healthcare information.

To our knowledge, this is the first study to perform a comparative analysis of PILs in the UK using a corpus-assisted discourse approach in two different centuries. This contribution highlights how PILs have transformed from persuasive and promotional devices into detailed, patient-focused documents. Future research should extend this analysis to other regions, examining how these documents have evolved and mirrored societal changes over time. Finally, this study underscores the importance of health literacy in the design of PILs highlighting the need to develop initiatives to ensure equitable access to healthcare information and foster patient safety and empowerment, with the aim of serving the evolving needs of healthcare.

#### 5. Limitations of the study

This study's findings should be considered within the limitations of its corpus size and methodological scope. While the two corpora consist of 30 PILs each, with 7,164 tokens in Corpus A and 71,976 in Corpus B, the smaller size of Corpus A reflects the scarcity of preserved early 20th-century PILs. This limitation is mitigated by the application of stringent inclusion criteria, ensuring comparability across the two datasets and adherence to established methodological standards in corpus linguistics (Brezina et al., 2021). The difference in corpus size also highlights the transformation in the genre itself, with contemporary PILs containing significantly more detailed information. The qualitative analysis began with keywords identified through quantitative methods, providing a robust foundation for linguistic exploration. However, the analysis evolved by employing close reading of concordance lines, collocations, and clusters, enabling a more nuanced examination of how high-frequency terms function pragmatically and rhetorically within the broader discourse. This layered approach ensured that the findings extend beyond surface-level observations, offering deeper insights into the socio-communicative shifts represented in the PILs.

#### Ethics and integrity policies

This study involved the analysis of PILs the early 20th and early 21st centuries. As the research did not involve human participants or sensitive personal data, formal ethical approval was not required. All data sources were publicly accessible, including archives from the Royal Pharmaceutical Society of Great Britain and the Electronic Medicines Compendium. Efforts were made to ensure the accuracy and authenticity of the data. The authors have properly cited all referenced works to maintain academic integrity.

## CRedit authorship contribution statement

**Nicola Pelizzari:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Funding

The author declares that there are no conflicts of interest regarding the publication of this paper. No funding or financial support was received from any entities that could have influenced the research findings. The authors are solely responsible for the content and have no affiliations with any organizations that could be perceived as influencing the research.

## Acknowledgements

Acknowledgment is extended to the Royal Pharmaceutical Society of Great Britain for providing access to their archival collection of Patient Information Leaflets, which was crucial for the completion of this study

## Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.acorp.2025.100129](https://doi.org/10.1016/j.acorp.2025.100129).

## References

- American Cancer Society. (2020). *Cancer Treatment & Survivorship Facts & Figures 2019–2021*. American Cancer Society. Retrieved from <https://www.cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/cancer-treatment-and-survivorship-facts-and-figures/cancer-treatment-and-survivorship-facts-and-figures-2019-2021.pdf>. (Last accessed 29/07/2024).
- Anthony, L., 2023. *AntConc* (Version 4.2.4) [Computer software]. Waseda University. Retrieved from <https://www.laurenceanthony.net/> (Last accessed 29/07/2024).
- Askehave, I., Zethsen, K.K., 2002. Translating for laymen. Perspectives: Studies in Translatology 10 (1), 15–29. <https://doi.org/10.1080/0907676X.2002.9961431>.
- Atal, I., Trinquart, L., Ravaud, P., Porcher, R., 2018. A mapping of 115,000 randomized trials revealed a mismatch between research effort and health needs in non-high-income regions. J Clin Epidemiol 98, 123–132. <https://doi.org/10.1016/j.jclinepi.2018.01.006>.
- Baker, P. (2006). *Using Corpora in Discourse Analysis*. Continuum. <https://doi.org/10.5040/9781350933996>.
- Biber, D., Conrad, S., Reppen, R., Leech, G., 1999. Corpus linguistics: investigating language structure and use. International journal of corpus linguistics 4 (1), 185–188. <https://doi.org/10.1075/ijcl.4.1.11lee>.
- Bouayad-Agha, N., Scott, D., Power, R., 2000. Integrating content and style in documents: a case study of patient information leaflets. Information Design Journal 9 (2–3), 161–176. <https://doi.org/10.1075/idj.9.2-3.05bou>.
- Brezina, V. (2018). *Statistics in Corpus linguistics: A practical Guide*. Cambridge University Press. <https://doi.org/10.1017/9781316410899>.
- Brezina, V., McEnery, T., Gablasova, D., 2021. The Written British National Corpus 2014 – design and comparability. Text & Talk 41 (5–6), 595–612. <https://doi.org/10.1515/text-2020-0052>.
- Buts, J., Baker, M., Luz, S., Engebretsen, E., 2021. Epistemologies of evidence-based medicine: a plea for corpus-based conceptual research in the medical humanities. Medicine, Health Care and Philosophy 24 (4), 621–632. <https://doi.org/10.1007/s11019-021-10027-2>.
- Cialdini, R.B. (2007). *Influence: The psychology of Persuasion*. New York, NY: HarperCollins.
- Cohen, J. (1988). *Statistical Power Analysis For the Behavioral Sciences* (2nd ed.). Routledge.
- Coulter, A. (1998). *Informing patients: An assessment of the Quality of Patient Information Materials*. King's Fund Publishing. <https://doi.org/10.1046/j.1369-6513.1999.0040c.x>.
- Courtwright, D.T. (2001). *Forces of habit: Drugs and the Making of the Modern World*. Harvard University Press. <https://doi.org/10.2307/j.ctvjsf4p9>.
- Covington, M.A., McFall, J.D., 2010. Cutting the gordian knot: the moving-average type-token ratio (MATTR). J Quant Linguist 17 (2), 94–100. <https://doi.org/10.1080/09296171003643098>.
- Cummings, L. (2015). Argument from authority. In *Reasoning and Public health: New ways of Coping With Uncertainty* (pp. 67–92). Springer. <https://doi.org/10.1007/978-3-319-15013-0>.
- Datapharm. (2024). *Electronic Medicines Compendium (eMC)*. Retrieved from <https://www.medicines.org.uk/emc>. (Last accessed 29/07/2024).
- Dent, M., 2006. Patient choice and medicine in health care: responsabilization, governance and proto-professionalization. Public Management Review 8 (3), 449–462. <https://doi.org/10.1080/14719030600853360>.
- Edwards, A., 2004. Communicating risks through patient information leaflets. Br J Clin Pharmacol 58 (4), 351–356. <https://doi.org/10.1136/bmj.327.7417.691>.
- Egbert, J., Burch, B., 2023. Which words matter most? Operationalizing lexical prevalence for rank-ordered word lists. Applied linguistics 44 (1), 103–126. <https://doi.org/10.1093/applin/amac030>.
- European Medicines Agency. (2023). *Guidelines On the Readability of the Label and Package Leaflet of Medicinal Products For Human Use*. European Commission. Retrieved from [https://health.ec.europa.eu/system/files/2023-09/2018\\_packaging\\_guidelines\\_en\\_1.pdf](https://health.ec.europa.eu/system/files/2023-09/2018_packaging_guidelines_en_1.pdf). (Last accessed 29/07/2024).
- Page-Butler, A., 2013. Including patients' perspectives in patient information leaflets: a polyocular approach. Fachsprache 35 (3–4), 140–154. <https://doi.org/10.24989/fs.v35i3-4.1326>.
- Fairclough, N. (2013). *Critical Discourse analysis: The critical Study of Language* (2nd ed.). Routledge.
- Food and Drug Administration of the United States of America. (2023). *Patient Medication Information: A 2023 Proposed Rule to Help Patients Understand Their Prescription Medication Information*. FDA. Retrieved from <https://www.fda.gov/drugs/idas-labeling-resources-human-prescription-drugs/patient-labeling-resources> (Last accessed 29/07/2024).
- Garner, M., Ning, Z., Francis, J., 2012. A framework for the evaluation of patient information leaflets. Health Expectations 15 (3), 283–294. <https://doi.org/10.1111/j.1369-7625.2011.00665.x>.
- Gee, J.P. (2014). *An Introduction to Discourse analysis: Theory and Method* (4th ed.). Routledge. <https://doi.org/10.4324/9781315819679>.
- Gotti, M., 2015. Insights into medical discourse: diachronic and synchronic perspectives. ESP Today 3 (1), 5–24.
- Grabowski, L., 2015. Phrase frames in english pharmaceutical discourse: a corpus-driven study of intra-disciplinary register variation. Research in Language 13 (3), 266–284. <https://doi.org/10.1515/rela-2015-0025>.
- Gries, S. (2021). *Statistics For Linguistics With R: A Practical Introduction*. Berlin, Boston: De Gruyter Mouton. <https://doi.org/10.1515/9783110718256>.
- Gyasi, W.K., 2023. Linguistic complexity of patient information leaflets: a Coh-Metrix analysis. Athens J. Health Med. Sci. 10 (4), 269–286. <https://doi.org/10.30958/ajhms.10-4-4>.
- Hegedűs, A. (2008). *Modality in Drug Information Leaflets: A Corpus-Based Analysis* (Doctoral dissertation, University of Pécs (Hungary)).
- Herber, O.R., Gies, V., Schwappach, D., Thürmann, P., Wilm, S., 2014. Patient information leaflets: informing or frightening? A cross-sectional study on risk communication in cancer care. Psychooncology 23 (10), 1173–1179. <https://doi.org/10.1186/1471-2296-15-163>.
- Hovland, C.I., Weiss, W., 1951. The influence of source credibility on communication effectiveness. Public Opin Q. <https://doi.org/10.1086/266350>.
- Hyland, K.L., 2010. Metadiscourse: mapping interactions in academic writing. Nordic Journal of English Studies. <https://doi.org/10.35360/njes.220>.
- Kettunen, K., 2014. Can type-token ratio be used to show morphological complexity of languages? J Quant Linguist 21 (3), 223–245. <https://doi.org/10.1080/09296174.2014.911506>.
- Kyle, K., 2021. Natural language processing for learner corpus research. Int. J. Learn. Corpus Res. 7 (1), 1–16. <https://doi.org/10.1075/ijlcr.00019.int>.
- Kyle, K., Crossley, S.A., 2018. Measuring syntactic complexity in L2 writing using fine-grained clausal and phrasal indices. The Modern Language Journal 102 (2), 333–349. <https://doi.org/10.1111/modl.12468>.
- Koo, M.M., Krass, I., Aslani, P., 2007. Evaluation of written medicine information: validation of the consumer information rating form. Annals of Pharmacotherapy 41 (6), 951–956. <https://doi.org/10.1345/aph.1K083>.
- Larson-Hall, J., Herrington, R., 2010. Improving data analysis in second language acquisition by utilizing modern developments in applied statistics. Applied Linguistics 31 (3), 368–390. <https://doi.org/10.1093/applin/amp038>.
- Lomax, E., 1973. The uses and abuses of opiates in nineteenth-century England. Bull Hist Med 47 (2), 167–176.
- Lonati, E., 2019. Stabilising the scientific lexicon in eighteenth-century British encyclopaedias and specialised dictionaries: a focus on medical terminology. Token 8, 107–142. <https://doi.org/10.25951/4063>.
- Madison, K., 2010. Patients as “regulators”? Patients' evolving influence over health care delivery. Journal of Legal Medicine 31 (1), 9–34. <https://doi.org/10.1080/01947641003598195>.
- Maglie, R. (2015). Linguistic features of patient information leaflets: a corpus-based study. In Gotti, M., Maci, S., & Sala, M. (Eds.), *Insights Into Medical Communication* (pp. 121–139). Peter Lang. <https://doi.org/10.4000/asp.4739>.
- Malvern, D., Richards, B., Chipere, N., & Durán, P. (2004). Traditional approaches to measuring lexical diversity. In *Lexical Diversity and Language Development* (pp. 23–35). Palgrave Macmillan. [https://doi.org/10.1057/9780230511804\\_2](https://doi.org/10.1057/9780230511804_2).
- McCarthy, P.M., Jarvis, S., 2010. MTL, vocd-D, and HD-D: a validation study of sophisticated approaches to lexical diversity assessment. Behav Res Methods 42 (2), 381–392. <https://doi.org/10.3758/BRM.42.2.381>.

- McManus, J., 2009. The ideology of patient information leaflets: a diachronic study. *Discourse & Communication* 3 (1), 27–56. <https://doi.org/10.1177/1750481308098763>.
- McEnery, T., & Hardie, A. (2012). *Corpus linguistics: Method, Theory and Practice*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511981395>.
- MHRA. (2020). *Best practice guidance on patient information leaflets*. Retrieved from [https://assets.publishing.service.gov.uk/media/5fe086c18fa8f5149718d66a/Best\\_practice\\_guidance\\_on\\_patient\\_information\\_leaflets.pdf](https://assets.publishing.service.gov.uk/media/5fe086c18fa8f5149718d66a/Best_practice_guidance_on_patient_information_leaflets.pdf). (Last accessed 29/07/2024).
- National Medical Products Administration (NMPA). (2022). *Patient information leaflets: regulatory guidelines in China*. Retrieved from <https://www.nmpa.gov.cn> (Last accessed 29/07/2024).
- O'Donovan, B., Rodgers, R.M., Cox, A.R., Krska, J., 2019. Use of information sources regarding medicine side effects among the general population: a cross-sectional survey. *Prim Health Care Res Dev* 20, e153. <https://doi.org/10.1017/S1463423619000574>.
- Oxman, A.D., Lavis, J.N., Lewin, S., Fretheim, A., 2020. SUPPORT tools for evidence-informed health policymaking (STP) 1: what is evidence-informed policymaking? *Health Research Policy and Systems* 7 (1), S1. <https://doi.org/10.1186/1478-4505-7-S1-S1>. Suppl.
- Paiva, D.S. (2000). Investigating style in a corpus of pharmaceutical leaflets: results of a factor analysis. In *Proceedings of the 38th Annual Meeting of the Association for Computational Linguistics* (pp. 237–240).
- Partington, A., Duguid, A., & Taylor, C. (2004). *Patterns and Meanings in discourse: Theory and Practice in Corpus-Assisted Discourse Studies (CADS)*. John Benjamins. <https://doi.org/10.1075/dapsac.10>.
- Pathberiya, S.D., 2016. Patient autonomy and decision making. *BMJ* 10. <https://doi.org/10.1186/s12978-016-0264-3>.
- Pearson, R.H., Mundform, D.J., 2010. Recommended sample size for conducting exploratory factor analysis on dichotomous data. *Journal of Modern Applied Statistical Methods* 9 (2), 359–368. <https://doi.org/10.22237/jmasm/1288584240>.
- Pelizzari, N., 2024. The challenges for EU user testing policies for patient information leaflets. *Int J Environ Res Public Health* 21 (10), 1301. <https://doi.org/10.3390/ijerph21101301>.
- Plonsky, L., Oswald, F.L., 2014. How big is “big”? Interpreting effect sizes in L2 research. *Lang Learn* 64 (4), 878–912. <https://doi.org/10.1111/lang.12079>.
- Porter, R. (1989). *Health For sale: Quackery in England, 1660–1850*. Manchester University Press. <https://doi.org/10.1017/S0025727300054673>.
- Raynor, D.K., Dickinson, R., Knapp, P., Long, A.F., Nicolson, D.J., 2007. Consumer medication information in Europe: an independent, expert review and recommendations. *Drug Saf* 30 (10), 837–850. <https://doi.org/10.1331/JAPhA.2007.06141>.
- Raynor, D.K., Knapp, P., Silcock, J., Parkinson, B., 2009. Performance-based readability testing of participant materials for a phase I trial: TGN1412. *J Med Ethics* 35 (9), 573–578. <https://doi.org/10.1136/jme.2008.026708>.
- Richards, B., 1987. Type/token ratios: what do they really tell us? *J Child Lang* 14 (2), 201–209. <https://doi.org/10.1017/S030500900012885>.
- Rowe, G., Frewer, L.J., 2005. A typology of public engagement mechanisms. *Science, Technology, & Human Values* 30 (2), 251–290. <https://doi.org/10.1177/016224390427172>.
- Rudd, R.E., Anderson, J.E., Oppenheimer, S., & Nath, C. (2023). Health literacy: an update of medical and public health literature. In *Review of Adult Learning and Literacy, Volume 7* (pp. 175–204). Routledge.
- Sandman, L., Munthe, C., 2010. Shared decision making, paternalism and patient choice. *Health Care Analysis* 18 (1), 60–84. <https://doi.org/10.1007/s10728-008-0108-6>.
- Schiffrin, D., Tannen, D., & Hamilton, H. E. (Eds.). (2018). *The Handbook of Discourse Analysis* (2nd ed.). Wiley-Blackwell. <https://doi.org/10.1002/9781118584194>.
- Scott, M. (2020). *WordSmith Tools* (Version 9) [Computer Software]. Lexical Analysis Software. Retrieved from <https://www.lexically.net/wordsmith/>.
- Shrank, W., Avorn, J., Rolon, C., Shekelle, P., 2007. Effect of content and format of prescription drug labels on readability, understanding, and medication use: a systematic review. *Ann Pharmacother* 41 (5), 783–801. <https://doi.org/10.1345/aph.1H582>.
- Siegel, R.L., Miller, K.D., Wagle, N.S., Jemal, A., 2023. Cancer statistics, 2023. *CA Cancer J Clin* 73 (1). <https://doi.org/10.3322/caac.21763>.
- Therapeutic Goods Administration (TGA). (2023). *Patient information leaflets (PILs): guidance for medicines in Australia*. Retrieved from <https://www.tga.gov.au/resources/guidance/providing-patient-information-leaflets-and-implant-cards-medical-devices> (Last accessed 29/07/2024).
- Vander Stichele, R.H. (2004). *Impact of Written Drug Information in Patient Package inserts: Acceptance and Benefit/Risk Perception*. Academia Press.
- Wongtaweepkij, K., Corlett, S., Krska, J., Pongwecharak, J., Jarernsripornkul, N., 2021. Patients' experiences and perspectives of receiving written medicine information about medicines: a qualitative study. *Patient Prefer Adherence* 569–580. <https://doi.org/10.2147/PPA.S298563>.
- Zanola, A., 2019. *Informed Consent' (IC) in Health Literacy (HL): State-of-the-art of an elaborate legal-lay communication process*. *Anglistica AION* 23 (1), 95–105. <https://doi.org/10.19231/angl-aion.201916>.