

CORPUS-BASED ANALYSIS OF BIOMEDICAL ABSTRACTS FOR MEDICAL ENGLISH PEDAGOGY

Anișoara POP¹

Abstract

Corpus linguistics provides empirically grounded insights into the pedagogy of abstract writing in Medical English, offering a more evidence-based and rhetorically informed alternative to mechanical, prescriptive approaches. This study applies a speciality-specific corpus methodology to examine key linguistic features of biomedical abstracts in a corpus of 48 original research article abstracts published in the Intensive Care Medicine journal (ICM) between January and October 2024. The analysis focuses on the use of passive voice and nominalisations, with a view to their exploitation in Medical English teaching, given the conflicting recommendations journals and publishers provide regarding these features. Although style manuals advise against passive constructions and nominalisations, the present findings demonstrate that these features are frequently employed in this specialised corpus of intensive care medicine, often preferred over alternative constructions. Current tendencies should inform Medical English instruction, where pedagogical recommendations benefit from actual language use and academic writing practice.

Keywords: biomedical abstracts, corpus linguistics, specialty-specific corpus, passive voice, nominalisations, Medical English.

DOI: 10.24818/SYN/2026/22/1.14

1. Introduction

Medical abstracts serve a crucial role in disseminating scientific knowledge and communicating research findings efficiently (Kramar and Ilchenko, 2023). They are highly condensed summaries that require clarity, precision, and rhetorical effectiveness. As a key academic skill, the ability to write coherent and well-structured abstracts is increasingly necessary for medical students in scientific communication, as part of publishing research articles, proposal funding, thesis/dissertation, and research report writing. Like any other piece of biomedical scientific writing, and especially because of their brevity, good abstracts require strategic planning, drafting, and revision, informed by a solid understanding of the genre's functional, structural, and linguistic expectations.

¹ Anișoara, Pop, *George Emil Palade* University of Medicine, Pharmacy, Science, and Technology of Targu Mureș, *Petru Maior* Faculty of Sciences and Letters, Romania, anișoara.pop@umfst.ro.

Functionally, abstracts are condensed versions of full research articles, with a word limit imposed (between 100 and 500 words). They help to filter and understand the main arguments of the research in a short time while selling it to potential readers and editors, and making its indexing and retrieval possible by search engines.

Content-wise, the genre-specific abstract structure follows the IMRaD structure, which includes the basic moves of Introduction, Methods, Results and Discussion, as well as Conclusion (Moskovitz et al., 2024). From a stylistic perspective, biomedical research journals vary widely in their recommendations regarding formality level, where the use of passive versus active voice and the extent of nominalisations are concerned. While formal scientific writing traditionally favours a sense of distancing of the author from the content and a more objective and impersonal tone achieved through employment, among others, of passive voice and nominalisations, many contemporary journal guidelines promote active voice, claiming the latter fosters conciseness, clarity, and transparency. Table 1 below illustrates examples of such contradictory recommendations that legitimate the ongoing debates surrounding desirable stylistic conventions in biomedical abstract writing.

Table 1. For and against the use of Passive Voice and Nominalisations in biomedical abstracts

Feature	Against	For
Passive voice	- Considered a fundamental principle of scientific writing, but increasingly criticised for making texts “ugly and unwieldy.” According to Cambridge University Press: <i>“When the passive voice is used, the writers take a back seat, as if they had watched the study from a distance... [Conversely, the use of active voice] lends a paper a directness that gives readers the sense that the writers actually did the work, or hold the opinions, that they describe.”</i> (Goodman and Edwards, 2006:221)	- The main factor in voice selection is sentence information structure: English tends to place “old” (given) information in subject position and introduce “new” information towards sentence rear. Thus, passive voice can be an effective resource when aligning with natural information flow (Minton, 2015).
Nominalisations	- Writers are advised to avoid nominalisations and instead, use precise, clear language (Arduengo, 2012)	- Nominalisations are widespread in academic writing because they promote brevity and support an impersonal tone. (Kwasi, 2015:56)

Source: Author’s compilation

Apart from the confusing nature of such instructions, different studies in the literature (Connaster, 2004; Millar et al., 2013) encourage a critical examination of the authority of style guides, putative rules of grammar and mechanics, pleading for a deeper understanding of the descriptive grammar of academic texts based on research. The present study contributes to the descriptive perspective by examining the organic, evidence-based use of language or “grammar” of biomedical abstracts, specifically passive versus active voice and authorial presence, as well as nominalisation, in a speciality-specific corpus of intensive care medicine (ICM) abstracts. The findings aim to inform pedagogical decisions in the teaching of Medical English abstract writing with students, where instructional practices can benefit from being aligned with authentic professional usage.

2. Corpus linguistics

According to McEnery and Hardie, “a corpus is a collection of electronic naturally-occurring texts, written or spoken, which are selected to be representative of a particular language or language variety.” (2011:1) Corpus linguistics adopts a frequency-based approach to language analysis, enabling researchers to identify recurring patterns across large collections of naturally occurring or “real” language datasets. Large amounts of data reveal tendencies and what is normal or typical in real-life language use. Compared with intuition-based or manually conducted analysis, which are slow and prone to mistakes, corpus-based methods offer higher accuracy, efficiency, and reliability.

Lancsbox software (<https://lancsbox.lancs.ac.uk/>), developed at Lancaster University, was used to analyse and visualise linguistic patterns within the ICM corpus. Corpus linguistics quantitative methods (frequency counts and collocation analysis) were integrated with qualitative analysis (key-word in context and manual verification) that allowed for a detailed examination of linguistic devices specific to the ICM corpus. The analysis focused on the categories of voice (passive vs active constructions), nominalisation, but also markers of authorial voice (“I”/“we” vs “the study”/“the results”). Authorial voice was included in the analysis as it signals the presence of the author in the text and is, as such, associated with active voice.

2.1 Description of the ICM Speciality-Specific Corpus

The compiled corpus consists of abstracts from original articles published in the Intensive Care Medicine journal (ICM) (<https://link.springer.com/journal/134>), the official journal of the European Society of Intensive Care Medicine and a high-impact Springer Nature journal (IF of 29.6 for 2023). This international intensive care medicine journal was initially selected as part of a larger research project involving comparison with a local Romanian journal of intensive care medicine.

However, those comparative findings, whether with a Romanian or a larger international corpus, are beyond the scope of the present paper.

Only abstracts of original research articles were included in the corpus as they conformed to the IMRaD structure, while all the other types of articles/abstracts were excluded, e.g., From the inside, What's new, Review, Systematic review, Correspondence, Narrative review, Editorial, Imaging, Retraction note, Correction, and Conference reports. The resulting dataset comprises 48 abstracts (N=48) published for a period of 10 months, i.e., between January and October 2024 (the month when the corpus was compiled retroactively for the year). After lemmatisation and part-of-speech tagging (word, pos, class, hw), the corpus contained 2,945 terms and 16,254 grammatical tokens. This dataset served as the basis for the linguistic analyses presented in the subsequent sections.

3. General frequencies in the ICM corpus

The first analysis of the ICM corpus referred to term frequency by using the Word (W) function – the number of times a search query matched the text in the corpus (absolute and relative frequency, the latter referring to the hits/number of words). Each linguistic aspect was first quantitatively analysed, then qualitatively checked for compliance with or exclusion from the pre-set description in the key-word-in-context function (KWIC), with an equal context size of 5 words to the left and right. In the second stage, analysis was performed manually, with the Text (T) tool.

3.1 Nouns

Apart from conjunctions ("and" - 525), prepositions ("of"- 470, "in" - 360, "to" - 295, "with" - 281), and articles ("the" - 415; "a" - 205), which are expected to rank first in analytical languages, top class frequency (relative normalised frequency to a base of one million) was held by the NOUN with 6,424 hits, which characterises the medical ICM abstract as a nominalised discourse. Of the nouns, "patient/s" was the term with the highest frequency (276 hits) and measure of textual lexical diversity (MTLD) of 90.98, followed by: a) the acronym "ICU" - 119 hits; b) "mortality" – 98; and "care" - 96 hits. "ICU" has the highest Log Dice score, 10.2, i.e., "patient" collocates the most with "ICU". Only this graph coll (collocations), called formulaic building blocks, or Lexical Frames (LFs), was calculated (Image 1).

Collocate	Distribution	Freq. (collocation)	Freq. (corpus)	Log Dice	MI	Delta P1	Delta P2
icu		5	119	10.2	5.2	0.3	0.04
the		8	525	8.9	3.8	0.4	0.01
in		5	395	8.6	3.5	0.3	0.01
of		5	481	8.4	3.2	0.2	0.01
and		5	525	8.2	3.1	0.2	0.009

Image 1. Most frequent collocation: "patient – ICU"

Source: Author's compilation

As expected, the ICU discourse proved to be built around the patient, patient care, and the extreme challenge of this speciality, i.e., mortality.

3.2 Verbs

Within the VERB category (1,843 hits), top frequencies were held by the auxiliaries: a) "Was" - 183; b) "Were" -160, and "Had" - 37 hits, a result also probably related to passive voice frequency.

4. Patterns of Passive Voice Use in Context

After frequency calculation, the study of the verb was expanded to include the type of voice. The corpus revealed a strong preference for passive constructions. Quantitatively, passive had 240 hits (14.827 tokens) in the ICM corpus (Image 2).

File	Left	Node	Right
INT C Med,	In this analysis, family members	were assessed	12 months after patient discharge
INT C Med,	in four ICUs (Melbourne, Australia)	were randomised	1:1 to penicillin (250 mg
INT C Med,	A total of 622 patients	were included:	287 (46%) had CRRT reinstated
INT C Med,	total of 519 family members	were analyzed (288 in the flexible visitation
INT C Med,	antibiotic-associated serious adverse events	were identified.	32/40 (80%) received a repeat
INT C Med,	mean age 66, 42% female)	were analyzed:	387 (44.4%) in the intervention
INT C Med,	competing outcomes, and administrative censoring	was imposed	72 h after randomization if
INT C Med,	toward appropriate management. Methods PTSD	was measured	90 days after ICU discharge

Image 2. Frequency of Passive constructions in ICM corpus

Source: Author's compilation

Qualitative analysis was done manually on the already identified entries/tokens. The role of manual analysis of results was to exclude wrongly selected tokens of the Passive involving co-occurrence of be+Vb-ed (nominal predicate), have+Vb-ed (Present/Past Perfect active voice), Vb-ed+Vb-ed (verbal adjective pre-modification). These constructions were excluded from the overall Passive frequency (Image 3).

The screenshot shows the LANCESBOX search interface. At the top, it says 'Intensive Care Med Abstracts' with a search filter set to 'whole corpus' and 16K results. Below that, a search for 'PASSIVE' is shown with 240 hits (14,765,60) and 1/1 text. A sample text snippet is displayed, highlighting several passive constructions: 'were obtained', 'We used mixed-effect Cox models to investigate the association between day one PaO2 or excess oxygen use (inspired fraction of oxygen ≥ 0.6 with PaO2 > 100 mmHg) on day-28 mortality.', 'Results 11,249 patients were included.', 'On day one, 5716 patients (50.8%) had normoxemia (60 ≤ PaO2 ≤ 100 mmHg), 1454 (12.9%) hypoxemia (PaO2 < 60 mmHg), and 4079 patients (36.3%) hyperoxemia (PaO2 > 100 mmHg). Excess oxygen was used in 2201 patients (20%). Crude day-28 mortality rate was 40.6%. There was a significant association between PaO2 and day-28 mortality with a U-shaped relationship (p < 0.001). Higher PaO2 levels (> 100 mmHg) were associated with day-28 mortality with a dose-effect relationship. Subgroup analyses showed an association between hyperoxemia and mortality in patients admitted with neurological disorders; however, the opposite relationship was seen across those admitted with sepsis'.

Image 3. Example of correct versus wrong selection of Passive Voice in LANCESBOX, followed by qualitative analysis.

Source: Author's compilation

To illustrate the process, several instances of passive hits incorrectly selected by LANCESBOX and excluded during the qualitative manual analysis from the overall Passive voice frequency are detailed below:

- The passive construction “Patients were randomly assigned to receive either....” occurred in parallel with frequent instances of false passive constructions with a role of adjectival pre-modification of the NP “trial” such as: “...international double-blinded *randomized-controlled* trial, SA-AKI patients were enrolled < 72 h on vasopressor and < 24 h of AKI.” The false passive construction “randomized controlled trial” occurred both in conjunction with:
 - the active voice: “In this prospective follow-up study of the multicenter EPaNIC randomized controlled trial, *we investigated* the impact...”, and
 - an elliptic passive: “We performed an individual patient-data analysis including 10 multicenter randomized controlled trials [that were] *conducted* over 20 years.”
- Other adjectival constructions with past participle form: “PTSD was measured 90 days after ICU discharge using validated instruments; For the *analysis*

conducted using weakly informed priors, and as compared to the reference category of age...”

- Past perfect/Present Perfect constructions, active voice:
 - “At one year, 117/348 (33.6%) of patients in the lower-oxygenation group *had died* compared to....”; “51% had acute leukemia, 38% lymphoma/multiple myeloma, and 40% *had received* a hematopoietic stem cell transplant (HCT)”
 - “Severe *Pneumocystis jirovecii* pneumonia (PJP) requiring intensive care *has been* the subject of few prospective studies”
- Nominal predicate constructions (to be + adjective): “Alveolar inflammation *is elevated* after HSV reactivation.”

The most frequent tenses associated with the Passive voice were the simple ones: Simple past tense, which occurred sometimes even in series, (e.g., “Interviews were conducted by two social science researchers, audio recorded, transcribed, and analyzed using thematic content analysis.”) was followed by Simple present, with the functions of defining, updating, and expressing limitations and future research opportunities:

- “Continuous renal replacement therapy (CRRT) *is used* for supportive management of acute kidney injury (AKI) and disorders of fluid balance (FB). Little is *known* about the predictors of successful liberation in children and young adults.”
- “Optimal oxygenation targets for patients with acute hypoxemic respiratory failure in the intensive care unit (ICU) *are not* clearly *defined* due to substantial variability in design of previous trials.”
- “HSV viral load dynamics prior to or without antiviral treatment *are associated* with mortality. Cardiogenic shock secondary to acute myocardial infarction (AMI-CS) *is associated* with substantial short- and long-term morbidity and mortality.”
- “A relatively low recruitment to screened ratio suggests that more inclusive eligibility criteria and integration of allergy assessment into routine ICU processes *are needed* to optimise allergy delabelling in critical illness.”
- “Echocardiography *is recommended* as a first-line tool in the assessment of patients with shock. The current provision of echocardiography in critical care *is poorly defined*.”

In decreasing order of frequency, other tenses associated with passive voice in the ICM corpus were:

- Present Perfect Passive affirmative or negative - “Herpesvirus reactivation *has been documented* among patients in the intensive care unit (ICU) and is associated with increased morbidity and mortality, particularly for cytomegalovirus (CMV). Epstein–Barr virus (EBV) *has been poorly studied* despite.....”; “Patients with hematologic malignancy (HM) commonly

develop critical illness. Their long-term survival and functional outcomes *have not been well described.*”;

- Past Perfect Passive: “Adult ICU patients with septic shock *had been randomized* to restrictive (n = 385) or standard (n = 392) intravenous fluid therapy.”
- Non-finite verb forms (passive gerund): “...with only severe hypocapnia and hypercapnia *being associated* with ...”; “We conducted a 1:1 case-cohort study matched on the PS score of *being adequately treated.*”

4.1 Discussion on Passive Voice Results

Employment of the passive voice in scientific writing has garnered the most attention from researchers. As a staple of formal academic writing, passive voice contributes objectivity, credibility, and informativeness (Pop, 2019:25). However, different academic writing guides discourage the use of passive voice due to its lack of brevity and clarity of accountability (i.e., it obscures who did the research). To exemplify, The Journal of Neuroscience advises its authors that: “Overuse of the passive voice is a common problem in writing. Although the passive has its place—for example, in the Methods section—in many instances it makes the manuscript dull by failing to identify the author's role in the research.... Use direct, active-voice sentences.” (Westbrook and Cooper, 2009)

Passive voice is also claimed to be wordy and boring for the reader as it reverses the natural order of sentence elements (Pop and Zdrenghea, 2007). Consequently, according to Jiang and Hyland (2023), the abstract style has undergone substantial changes over the last few years. One such change refers to the tendency to discard the use of the passive in favour of the active voice, especially in the field of social sciences, and less in the natural sciences (Zhang and Cao, 2025).

Concurrent with previous studies (Ferreira, 2021:145) and in contrast with recent recommendations (Westbrook and Cooper, 2009; Behavioral Ecology², British Medical Journal³, The Journal of Trauma and Dissociation⁴, Nature⁵, Ophthalmology⁶), the speciality-specific ICM corpus displayed high frequencies of passive voice. This preference aligns with the conventions of biomedical writing,

² “The first-person active voice is preferable to the impersonal passive voice.”

³ “Please write in a clear, direct, and active style... Write in the active [voice] and use the first person where necessary.”

⁴ “Use the active voice whenever possible: We will ask authors that rely heavily on use of the passive voice to re-write manuscripts in the active voice.”

⁵ “Nature journals like authors to write in the active voice (‘we performed the experiment...’) as experience has shown that readers find concepts and results to be conveyed more clearly if written directly.”

⁶ “Active voice is much preferred to passive voice, which should be used sparingly...”

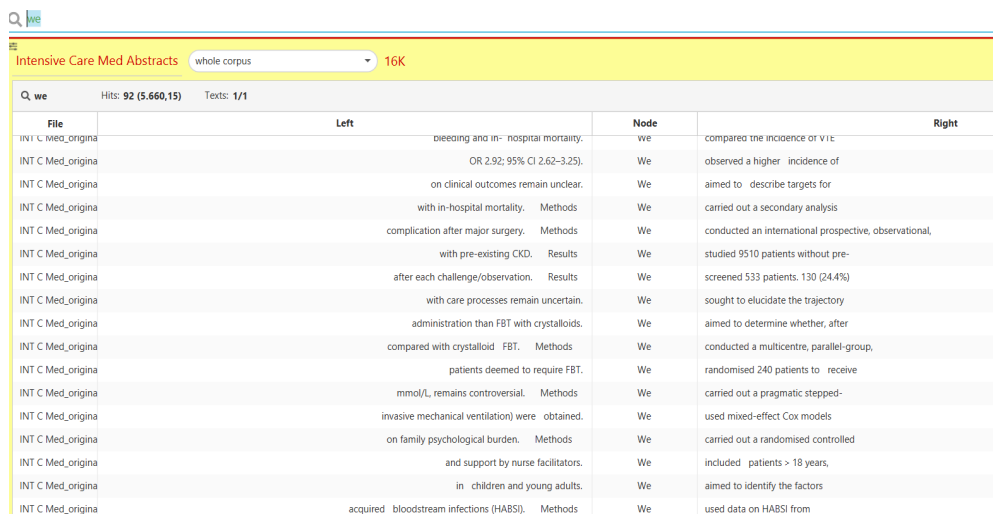
where the passive voice contributes to objectivity and rhetorical distance (Zeiger, 2000; Minton, 2015). As demonstrated in the qualitative analysis, many sentences employed the passive voice to describe methodological processes, instead of using an active voice that binds authors to the claims, often implying that the identity of the agent is either irrelevant, obvious, or intentionally backgrounded. Examples include statements such as:

“Similar results *were found* between increasing age and post-ECMO complications.”

“A three-step Delphi process based on an iterative approach *was used* to obtain the final consensus.”

5. Personal versus Non-Personal Active Voice

The active voice analysis distinguished personal active voice (first-person pronoun) from non-personal active structures. Personal active voice is a construction that foregrounds the author, and is linked with the use of a singular or plural first-person pronoun (“I”, “we”). Conversely, non-personal active constructions avoid the first-person pronouns and attribute the agency to “the paper” or “the study”. In terms of frequencies, the corpus contained no instances of “I” (0 hits), while “We” occurred 92 times, predominantly in the Methods and Results sections of the abstracts, where authors typically foreground their research actions (e.g., “We conducted”, “We performed”, etc.). (Image 4)



Intensive Care Med Abstracts whole corpus 16K

Q: we Hits: 92 (5,660,15) Texts: 1/1

File	Left	Node	Right
INT C Med_origina	diecing and in-hospital mortality.	we	compared the incidence of vte
INT C Med_origina	OR 2.92; 95% CI 2.62–3.25).	We	observed a higher incidence of
INT C Med_origina	on clinical outcomes remain unclear.	We	aimed to describe targets for
INT C Med_origina	with in-hospital mortality. Methods	We	carried out a secondary analysis
INT C Med_origina	complication after major surgery. Methods	We	conducted an international prospective, observational,
INT C Med_origina	with pre-existing CKD. Results	We	studied 9510 patients without pre-
INT C Med_origina	after each challenge/observation. Results	We	screened 533 patients. 130 (24.4%)
INT C Med_origina	with care processes remain uncertain.	We	sought to elucidate the trajectory
INT C Med_origina	administration than FBT with crystalloids.	We	aimed to determine whether, after
INT C Med_origina	compared with crystalloid FBT. Methods	We	conducted a multicentre, parallel-group,
INT C Med_origina	patients deemed to require FBT.	We	randomised 240 patients to receive
INT C Med_origina	mmol/L, remains controversial. Methods	We	carried out a pragmatic stepped-
INT C Med_origina	invasive mechanical ventilation) were obtained.	We	used mixed-effect Cox models
INT C Med_origina	on family psychological burden. Methods	We	carried out a randomised controlled
INT C Med_origina	and support by nurse facilitators.	We	included patients > 18 years,
INT C Med_origina	in children and young adults.	We	aimed to identify the factors
INT C Med_origina	acquired bloodstream infections (HABS). Methods	We	used data on HABS from

Image 4. Personal first-person plural active voice in the ICM corpus.

Source: Author compilation

5.1. Non-personal Active Voice

Far more common than the personal active voice structures were the non-personal active constructions ("There are/is/were") that maintain authorial distance, specific to objective writing. Such structures included:

- *Existential constructions* (24 hits = absolute frequencies), e.g., "There was no difference observed in in-hospital mortality..."
- *Evidence-based subjects*, for example: "These findings support the safety validity..."; "These results may inform the design of future clinical trials..."; "These results support..."
- *Self-referential nominals*, where "the writing", "the study" or another non-human entity, etc., was personified and made the grammatical subject of the sentence, e.g., instead of using: "We included.....patients in the study", the authors used: "The study included 129 patients".

5.2. Linguistic Impersonality Through Self-Referential Nominals

Marina Bondi (2014) designates the nouns "study", "paper", "analysis", "results", etc., used with personified function as «self-referential nominals» (p.251). Self-referential nominals that background the authors and foreground the study or its results as the subject of the sentence were widespread in the corpus: 200 occurrences of non-personal active voice, underlining that authors did not choose to position themselves as the authors, trying to minimise personal involvement and responsibility for the scientific results:

- "This study showed that few patients with PJP admitted to intensive care received prophylactic antibiotic therapy";
- "The purpose of this study was to profile genetic causal factors of acute respiratory distress syndrome...";
- "This study aimed to determine the best strategy to achieve fast and safe extubation";
- "This study identifies a wide range of ARDS-relevant factors....The aim of this study was to determine whether selective....."

This self-referential nominals strategy reinforces the impersonal and objective tone common to the sciences and valued in the medical discourse, thus prioritizing the product over the researcher.

6. Nominalisations and Lexical Density

Nominalisations constituted another salient feature of the ICM corpus. Defined as the conversion of verbs or adjectives into nouns, nominalisations are widely acknowledged as markers of the academic and scientific register (Jalilifar and

Memari, 2017:131; Yuan, 2022). They convert processes into single events, thus shifting the focus of the sentence away from the action (verb) onto an object or concept (noun) and contributing, thus, to lexical density (Pop, 2007:396). Through lexical density, nominalisations in turn add conciseness but also abstraction, and thus complexity, and formality to the medical abstracts.

In the ICM corpus, 758 nominalisation hits were identified, with the most frequently used suffixes being *-sion/-tion*, followed by *-ment* and *-ity* (Image 5).

Q NOMINALIZATION Hits: 758 (46.634,67) Texts: 1/1

File	Left	Node
INT C Med_origina	AKI. However, early postoperative-AKI	severity
INT C Med_origina	early postoperative-AKI severity and	duration
INT C Med_origina	confirmatory test on time to	extubation
INT C Med_origina	to achieve fast and safe	extubation.
INT C Med_origina	low-to-intermediate risk for	extubation.
INT C Med_origina	of arterial oxygen (PaO ₂) to	fraction
INT C Med_origina	of arterial oxygen (PaO ₂) to	fraction
INT C Med_origina	Primary outcomes were time-to-	extubation
INT C Med_origina	weaning rate. Secondary outcomes included	reintubation

Q NOMINALIZATION Hits: 758 (46.634,67) Texts: 1/1

INT C Med_origina articles whole abstracts.pdf 758 (46.634,67)

Tokens: 16,254 MATTR₉₀: 0,82 MTLT: 90,98

Identification of genetic profile and biomarkers involved in acute respiratory distress syndrome Purpose The purpose of this study was to profile genetic causal factors of acute respiratory distress syndrome (ARDS) and early predict patients at high ARDS risk. Methods We performed a phenome-wide Mendelian Randomization analysis through summary statistics of an ARDS genome-wide association study (1250 cases and 1583 controls of European ancestry) and 33,150 traits. Transcriptomic data from human blood and lung tissues of a preclinical mouse model were used to validate biomarkers, which were further used to construct a prediction model and nomogram. Results A total of 1736 traits, including 1223 blood RNA, 159 plasma proteins, and 354 non-gene phenotypes (classified by Biochemistry, Anthropometry, Disease, Nutrition and Habit, Immunology, and Treatment), exhibited a potentially causal relationship with ARDS development, which were accessible through a user-friendly

Image 5. Examples of nominalisations.

Source: Author compilation

All nominalisations contributed to:

- information condensation, allowing complex processes to be expressed compactly,
- lexical cohesion, linking clauses and sentences through abstract nouns,
- increased abstraction and technicality, consistent with the scientific discourse, and
- "subject weakening", a function that reduces the need to express the doer or agent of the action, thereby adding further impersonality.

The high frequency of nominalisations made the abstracts more information-dense and, together with the passive voice and impersonal active constructions,

nominalisations contributed substantially to the high lexical density and formality of the analysed abstracts.

7. Overview of Linguistic Trends in ICM Abstracts

Across the linguistic categories examined – voice, authorial presence, and nominalisations, several tendencies emerged. First, passive voice remains widely used, despite current calls for more active constructions. Recent research (Zhang and Cao, 2025) also suggests that passive voice frequency has remained relatively stable in recent years. Moreover, authorial presence is limited: the corpus contains no instances of "I", while "we" appears only moderately. Instead, authors of the examined ICM abstracts frequently employ non-personal active structures that foreground results of the study itself rather than the researchers.

Nominalisations occur at high rates, contributing to conciseness, abstraction, and an impersonal tone. Collectively, these features align with established conventions in biomedical and scientific discourse mentioned above (see section 4.1), confirming patterns reported in previous findings in biomedical sciences.

8. Conclusion

The analysis conducted in this study demonstrates that the stylistic profile of biomedical abstracts in Intensive Medicine continues to be shaped by linguistic features that have been traditionally associated with scientific discourse, namely the predominant use of the passive voice, extensive nominalisation, and an impersonal stance. Despite contemporary recommendations from editors and style guides to prioritise active constructions and explicit agency, the evidence from this specialty-specific corpus indicates that such prescriptive guidelines fail to reflect current biomedical practice.

Instead, examination of the ICM corpus within the corpus linguistic framework demonstrates that the abstracts systematically employed passive forms, self-referential nominals, and nominalisations in order to foreground processes, results, and outcomes, rather than researchers. Such findings confirm the rhetorical value of impersonality and objectivity in medical scientific discourse, aligning with other corpus-based studies that suggest limited change in passive voice frequency across biomedical science.

By adopting a corpus-based approach, the study highlights how empirical evidence can inform pedagogical decisions in Medical English teaching and learning. The linguistic patterns identified through this analysis – especially the high lexical density generated through nominalisations and the strategic use of both passive voice and non-personal active constructions – offer insights that can be

incorporated into instructional materials, writing scaffolding tools, and genre-based teaching for students.

To conclude, abstracts function rhetorically not only to condense information but also to negotiate credibility, stance, and disciplinary identity, balancing impersonality with the need to signal the researcher's contribution within the medical research community. Whenever style guidelines are missing, as in the case of the ICM, it is up to the author to negotiate such rhetorical functions either towards the epistemic culture of medicine, which values objectivity, replicability, and methodological transparency inherent among others in passive constructions, or towards a more controlled authorial presence characterised by employment of first-person pronouns and active constructions.

This study needs to acknowledge limitations related to corpus size and journal scope. Further research based on larger, more diverse datasets across multiple medical specialties and including a comparative medical corpus would enable a more comprehensive exploration. Despite limitations, the analysis illustrates the value of specialty-specific corpora in describing authentic language use, underscoring the role of corpus linguistics as a reliable methodology in Medical English pedagogy and technical editing.

References and bibliography

- Arduengo, M.** 2012. 'Getting what you want from your scientific writing: tips for writing clearly', in *Medical Writing*, 21(3):197-200. DOI: 10.1179/2047480612Z.0000000004
- Instructions to authors. Oxford Journals Web site: Behavioral Ecology. http://www.oxfordjournals.org/our_journals/beheco/for_authors/general.html
- British Medical Journal The essentials of BMJ style. BMJ Publishing Group Web site: British Medical Journal. <http://resources.bmj.com/bmj/authors/bmj-house-style>.
- Bondi M.** 2014. 'Changing voices: Authorial voice in abstracts' in Bondi M. and R. Lorés Sanz (Eds.). *Abstracts in Academic Discourse: Variation and Change*. Bern: Peter Lang, 243–270.
- Connaster, B.R.** 2004. *Reconsidering Some Prescriptive Rules of Grammar and Composition*.
- Ferreira, F.** 2021. 'In defense of the passive voice'. *Am Psychol*. 76(1):145-153. doi: 10.1037/amp0000620.
- Goodman, N.W., and M.B. Edwards.** 2006. *Medical Writing. A Prescription for Clarity*, 4th edition, Cambridge University Press.
- Guide for authors.** Elsevier Web site: *Ophthalmology: Journal of the American Academy of Ophthalmology*. Retrieved from <http://www.elsevier.com/>

wps/find/journaldescription.cws_home/620418/authorinstructions. Accessed on 1 February 2026.

- Guidelines for authors.** International Society for the Study of Trauma Web site: The Journal of Trauma & Dissociation. Retrieved from <http://www.isstd.org/jtd/journal-trauma-dissociation-info-for-authors.htm>. Accessed on 1 February 2026.
- Jalilifar, A. and M. Memari.** 2017. 'Exploring Nominalization in Physics and Applied Linguistics Textbooks with Different Levels of Difficulty: Implications for English for Specific Purposes' in *Teaching English Language*, 11(2):131-159.
- Jiang, F.K. and K. Hyland.** 2023. 'Changes in Research Abstracts: Past Tense, Third Person, Passive, and Negatives', in *Written Communication* 40 (1): 210-237.
- Kramar, N. and O. Ilchenko.** 2023. 'Stylistic features of English language linguistic abstracts: A corpus-based study', in *Movosnavsto*. 331:41-53. 10.33190/0027-2833-331-2023-4-003.
- Kwasi, S.A.** 2015. 'Nominalizations in research article abstracts: A comparative study', in *European Journal Language Studies*, 2(1):56-66.
- McEnergy, T. and A. Hardie.** 2011. 'What is corpus linguistics?', in *Corpus Linguistics: Method, Theory and Practice*. Cambridge Textbooks in Linguistics. Cambridge University Press, 1-24.
- Millar, N., B. Budgell, and K. Fuller.** 2013. 'Use the active voice whenever possible': The Impact of Style Guidelines in Medical Journals. *Applied Linguistics*. 34(4): 393-414. doi:10.1093/applin/ams059.
- Minton, T.D.** 2015. 'In defense of the passive voice in medical writing', in *Keio J Med*. 64(1):1-10. doi: 10.2302/kjm.2014-0009-RE.
- Moskovitz, C., B. Harmon, and S. Saha.** 2023. The Structure of Scientific Writing: An Empirical Analysis of Recent Research Articles in STEM. *Journal of Technical Writing and Communication*, 54(3): 265-281. <https://doi.org/10.1177/00472816231171851>.
- Nature.** How to write a paper: writing for a Nature journal. Nature Publishing Group Web site: Nature. Retrieved from http://www.nature.com/authors/author_services/how_write.html. Accessed on 15 January 2026.
- Pop, A.** 2007. 'Information Density in Classified Advertising', in *European Integration Between Tradition and Modernity*, Ed. „Petru Maior” Țirgu-Mureș, 396-403.
- Pop, A. and M. Zdrenghea.** 2007. 'The Role of Thematic Information in Advertising Headlines', *Studia Universitatis Babeș-Bolyai*, 4:201-207.
- Pop, A.** 2019. *Basic English Grammar for Medical Professionals*, Lambert Academic Publishing.
- Yuan, Y.** 2022. 'Examining the application of nominalization in segmented abstracts: a corpus-based study' *Journal of China Computer-Assisted Language Learning*, 2(2):281-298. <https://doi.org/10.1515/jccall-2022-0026>

- Westbrook G, L. Cooper.** 2009. 'Writing tips: Techniques for clear scientific writing and editing'. The Society for Neuroscience Web site: The Journal of Neuroscience. <http://www.jneurosci.org>. **Zeiger, M.** 2000. *Essentials of Writing Biomedical Research Papers*. 2nd ed. New York, NY: McGraw-Hill.
- Zhang, F. and R. Cao.** 2025. 'The Passive Voice in Abstracts in Research Articles: A Diachronic Variation across Disciplines', in *International Journal of Education and Humanities*, 18(1):57-61. <https://doi.org/10.54097/49q16a46>.

The author

Anișoara Pop is Professor and PhD supervisor (habilitation) in the field of English language and linguistics at *GE Palade* University of Medicine, Pharmacy, Science and Technology. Her interests include discourse analysis, specialised medical terminology and translation, Medical English, and, of late, narrative medicine.

ORCID profile: <https://orcid.org/0000-0001-6990-9990>, Google scholar:

https://scholar.google.com/citations?user=PO_VAzAAAAAJ&hl=en, Brain map:

<https://www.brainmap.ro/anisoara-pop>.