Candel-Mora, Miguel Ángel (2024): Denominative variation in the terminological representation of Women's Health. *Cultura, Lenguaje y Representación*, Vol. XXXIV, 131-150
ISSN 1697-7750 · E-ISSN 2340-4981
DOI: https://doi.org/10.6035/clr.7888

Denominative variation in the terminological representation of Women's Health

Variación denominativa en la representación terminológica de la salud de la mujer

MIGUEL ÁNGEL CANDEL-MORA
UNIVERSITAT POLITÈCNICA DE VALÈNCIA[1]
https://orcid.org/0000-0001-8754-6046

Artículo recibido el / *Article received*: 2024-01-19
Artículo aceptado el / *Article accepted*: 2024-04-05

ABSTRACT: Medical language is characterized by its veracity, precision, and clarity (Navarro, 2009). However, due to the different communicative situations and contexts in which it is used, it is one of the special languages with more terminological variation (Bowker and Hawkins, 2006). From the point of view of terminology work, in any of its applications: language planning, standardization or translation, the first steps consist of structuring the subject area and accurately define the conceptual field (Cabré, 2005; ISO, 2022; Wright, 1997), and variation is usually an obstacle during this stage. This paper presents the findings of a study for the elaboration of terminological resources on Women's Health from a corpus of specialized academic articles in English. Preliminary results reveal a lack of uniformity in the identification of the most representative lexical units regarding issues that specifically affect Women's Health. This analysis offers a typology of denominative variation in the subject field of Women's Health in academic journals in English prior to initiate the delimitation of the conceptual field in Spanish and standardize terminology equivalence in order to ensure efficient communication.

Key words: women's health; denominative variation; terminology management, medical language

[1] This publication is part of the project «Automatic terminology extraction and corpora of specific domains for the visibilisation of women-related health problems» (TED2021-130040B-C22), funded by the Spanish Ministry of Science and Innovation, the State Research Agency (10.13039/501100011033) and by the European Union «NextGenerationEU»/ Recovery, Transformation and Resilience Plan.

DOI: https://doi.org/10.6035/clr.7888

RESUMEN: El lenguaje médico se caracteriza por su veracidad, precisión y claridad (Navarro, 2009). Sin embargo, debido a las diferentes situaciones y contextos comunicativos en los que se utiliza, es una de las lenguas de especialidad con mayor variación terminológica (Bowker and Hawkins, 2006). Desde el punto de vista del trabajo terminológico, en cualquiera de sus aplicaciones: planificación lingüística, normalización o traducción, los primeros pasos consisten en estructurar el área temática y definir con precisión el campo conceptual (Cabré, 2005; ISO, 2022; Wright, 1997), y la variación suele ser un obstáculo durante esta etapa. Este artículo presenta los resultados de un estudio para la elaboración de recursos terminológicos sobre la salud de la mujer a partir de un corpus de artículos académicos especializados en inglés. Los resultados preliminares revelan una falta de uniformidad en la identificación de las unidades léxicas más representativas en relación con los temas que afectan específicamente a la salud de la mujer. Este análisis ofrece una tipología de la variación denominativa en el campo temático de la salud de la mujer en revistas académicas en inglés antes de iniciar la delimitación del campo conceptual en español y estandarizar la equivalencia terminológica con el fin de garantizar una comunicación eficaz.

Palabras clave: salud de la mujer; variación denominativa; gestión terminológica, lenguaje médico.

1. INTRODUCTION

The most established currents in the study of terminology, language planning (Auger and Rousseau, 1987; TERMCAT, 1990; Wüster, 1998; Arntz and Picht, 1995) and standardization (2022), agree on the key phases of terminology work: preparation of the work, organization and presentation of the work, and revision. The first phase focuses on the preparation of the work, the choice, structuring, and conceptual delimitation of the area, as well as the compilation of documentary resources available. It is in the second phase, term extraction and delimitation of terminological units, that variation plays a key role. According to Sager:

> The lexicon of a special subject language reflects the organisational characteristics of the discipline by tending to provide as many lexical units as there are concepts conventionally established in the subspace and by restricting the reference of each such lexical unit to a well-defined region. Beside containing a large number of items which are endowed with the property of special reference the lexicon of a special language also contains items of general reference which do not usually seem to be specific to any discipline or disciplines and whose referential properties are uniformly vague or generalised. (Sager, 1990, p. 19)

These preliminary stages of terminology work are foundational to the development and refinement of a specialized language within a particular domain. This crucial phase involves systematic exploration, organization, and analysis of terminology

to establish a robust foundation for effective communication. Terminology, which comprises specialized terms unique to a particular domain, holds a significant role in guaranteeing precision, clarity, and consistency in communication among professionals, researchers, and stakeholders. This is in line with Navarro (2009: 90) who notes that, in the case of medical language the three main features of scientific language in general, and of medical writing in particular, are truthfulness, precision and clarity; in other words, what is expressed in a scientific text should not be false, ambiguous, incomprehensible, shocking, or cumbersome to read.

One of the fundamental distinguishing features of specialized languages is their purpose for exchanging technical or specialized knowledge: definitions of specialized language include aspects that differentiate it from the general language such as its use in specialized communication, the type of texts in which it appears and the use of specific terminology (Alcaraz, 2000; Dubuc and Lauriston, 1997; Resche, 2000; Picht, 1987; Wright and Wright, 1997).

Lerat (1997: 18) highlights that a specialized language should not be limited to terminology alone. This is because, apart from employing specialized terms, it also encompasses non-linguistic symbols within sentences that include the standard linguistic elements of a particular language. Furthermore, Lerat notes that the level of specialization contained within a text varies depending on the specific communication requirements. Cabré (1993: 137) further rationalizes the significance of terminology in specialized languages from a communicative perspective. She emphasizes that specialized discourse deviates from general language norms due to various factors, including subject matter, interlocutors' specific traits, communicative situation, or transmission channel.

For Sager (1990: 215), the use of variation revolves around different hypotheses: the differences between text types; the higher density of alternative forms in special language discourse; alternative designations are realized by means of absolute synonyms, the use of contextual synonyms and the choice of the immediate hyperonym; variation may be created by a process of abbreviation; in compound terms the preferred mode of abbreviation is that of omitting one element; and, contextual abbreviation or reduction.

In any case, it is traditionally emphasized that, in order to facilitate specialized communication and knowledge transfer, terms are relatively fixed elements and should not be subject to variation (Picht, 1987; Sager, 1990; Wüster, 1998). However, as Bowker and Hawkins (2006: 80) point out, different studies reveal that, even within the boundaries of specialized communication, terminological variation exists, and they highlight that one specialized field in which terminological variation seems particularly prevalent is medicine, the subject matter of this study. According to Bowker and Hawkins (2006: 100), «Language, and particularly specialized language, cannot be completely random or people would not understand one another; however, it does admit a greater degree of variability than previously thought.»

This variation is due to conceptual motivational factors, linguistically motivated and socially motivated factors (Bowker and Hawkins, 2006, p. 82). According to these authors, it is complex to deduce both conceptually motivated and socially motivated factors simply by studying a corpus, so although «Linguistically motivated term choices are not as important as conceptually motivated term choices with regard to the potential for meaning distortion, but they are important nonetheless because they will affect the naturalness or idiomaticity of the text.» (Bowker and Hawkins, 2006, p. 92).

As a lexical-semantic phenomenon, denominative variation refers, according to Tercedor-Sánchez (2011), to linguistic representations that share with the main term

 DOI: https://doi.org/10.6035/clr.7888

certain semantic and conceptual features that are activated in certain contexts and situations. Likewise, Freixa (2006: 52) presents a typology of five causes for terminological variation: dialectal, due to the origin of the authors; functional, due to different communicative registers; discursive, due to the stylistic and expressive needs of the authors; interlinguistic, due to contact between languages; and finally, cognitive, due to different conceptualizations and motivations.

In a study on denominative variation in the language of mathematics, Freixa and Montané (2006: 212) conclude that the causes for the appearance of synonymous forms in mathematical terminology coincide with those of other fields of knowledge and that the most frequent reason for variation is the stylistic need to avoid repetition, to vary the expression so that the text does not sound too repetitive and that the degree of specialization of the texts would be the least convincing cause of denominative variation. According to Freixa and Montané (2006: 212), among the causes of denominative variation are lexical changes when substituting an element of the terminological syntagm for a synonym; reductions of the extension and the base of the syntagm; lengthening, in which a semantic aspect of the concept is introduced in the denomination; morphosyntactic changes, such as the alternation between the presence and absence of the definite article; and graphic changes, common in the language of mathematics:

> It is possible that the particular characteristics of specialised texts determine denominative alternatives, and therefore it may be necessary to add a type of textual causes (and maybe even other types). It is very probable that, for each of the identified types, the different sub-cause identified may be described a little further, as a result of the analysis of real texts from different domains of specialty, from different levels of specialisation and different languages. (Freixa, 2006: 70-71)

From the perspective of translation studies, Alarcón-Navío et al. (2016: 118) point out that one of the consequences of variation is that the translator must face a high degree of uncertainty both at the cognitive level (since it affects comprehension) and at the level of equivalences. According to (Tercedor-Sánchez and López-Rodríguez, 2012), in medicine, medical concepts can be lexicalized in diverse ways depending on their appropriateness to a specific communicative situation or the facet or dimension of the concept in question.

For this work, a corpus of journals specialized in Women's Health composed of articles from three journals: *Journal of Women's Health Care*, *Women's health issues and Women's Health*, was analyzed over a period of three years (2021-2023) with a total of 405 articles and 1,957,385 words to empirically observe the type of terminological variations carried out in issues affecting Women's Health and how to address them. The study was carried out using corpus linguistics techniques, with Sketch Engine. The corpus-based approach to study combining forms in context seems an adequate option as supported by previous research (Prieto-Velasco et al., 2012; Freixa, 2006; Sager, 1990; Peters et al., 2018; Wiese, 2018). Following Bowker and Hawkins (2006: 101) «By studying these terms in context, we were able to uncover a number of regular patterns of variation, which allowed us to deduce various possible motivations behind term choice, including conceptual, linguistic and social motivations».

It should be noted that this is a preliminary study, prior to the development of terminology databases on Women's Health in English and Spanish. In a first phase, the

aim is to identify the type of terminological variations in English in order to, later on, contrast the results with those obtained in a similar study with a corpus of scientific articles in the field of Women's Health written in Spanish.

2. METHODOLOGY

This work is based on the analysis of a corpus of research papers specialized in Women's Health from 3 journals: *Journal of Women's Health Care*, *Women's health issues* and *Women's Health*, during a period of three years (January 2021-April 2023) with a total of 405 articles and 1,957,385 words to observe empirically the type of terminological variations in medical language on Women's Health and identify the trends and patterns, and linguistic variations within this domain.

Table 1. Corpus data

Journal title	Publisher	Number of articles	Country of publication	Journal impact factor (JCR)
Journal of Women's Health Care	Longdom Group	96	Belgium	5.14
Women's health issues	Elsevier	103	USA	3.053
Women's Health	Sage	206	UK	2.4

These publications were chosen because of their academic rigor, because they are peer reviewed, international, and indexed in prestigious databases, they focus on Women's Health, and each publication provides different approaches that will be useful to observe terminological variation, object of this study. Thus, *Journal of Women's Health Care* (JWHC) contains research based, clinical and non-clinical, diagnostic, and social aspects in the field of medical sciences in the form of articles, review articles, case reports, and short communications. *Women's Health Issues* (WHI) is the official journal of the *Jacobs Institute of Women's Health* and is dedicated to improving the health of women in the context of the U.S., health care delivery system and policymaking processes. Finally, *Women's Health* (WHE) focuses on all aspects of women's healthcare, from childhood/adolescence to menopause and beyond, with primary research, systematic reviews, meta-analyses, and reviews from both low- and high-resource countries.

The Sketch Engine tool was used for the analysis of the corpus, and in a first phase, the basic functions of keyword identification and word lists were applied to gain preliminary insights into language structure, usage, and variability. In this application, keywords are defined as typical words and phrases of the corpus because they appear more frequently than in the general language when compared to the reference corpus *English Web Corpus* (enTenTen) composed of 52 billion words.

With the data obtained in this first phase and the evaluation of previous studies on variation in terminology, we proceeded to design an ad hoc framework of analysis to organize and structure the findings obtained during the corpus analysis. To this end, a first approach to linguistic variation (Sager, 1990; Freixa, 2006; Freixa and Montané, 2006; Arntz and Picht, 1995) was proposed, focused on aspects such as lexical changes, synonyms, abbreviations, omissions and morphosyntactic changes.

 DOI: https://doi.org/10.6035/clr.7888

Other useful features of Sketch engine in this research included Word Sketches, which offer comprehensive information about a word, including its collocations, grammatical relationships, and common contexts, as well as the Thesaurus and Synonymy, which allow for the examination of word relationships, the discovery of synonyms, and the identification of related terms.

As a starting point for the description of the conceptual field, the definition of Women's Health of the National Institute of Child Health and Human Development was used: «Women's health is a broad category that includes health issues that are unique to women, such as menstruation and pregnancy, as well as conditions that affect both men and women, but that may affect women differently, such as heart disease and diabetes.»[2] And the definition from the National Library of Medicine of the National Institutes of Health of the United States: «Women's health refers to the branch of medicine that focuses on the treatment and diagnosis of diseases and conditions that affect a woman's physical and emotional well-being»[3].

The Medical Literature Analysis and Retrieval System Online (MEDLINE), a bibliographic database of life sciences and biomedical literature, part of the larger PubMed database, maintained by the United States National Library of Medicine (NLM), also provides the range of specialties and focus areas of Women's Health, as shown in Table 2.

Table 2. Range of specialties and focus areas in Women's Health

1.	Birth control, sexually transmitted infections (STIs), and gynecology
2.	Breast cancer, ovarian cancer, and other female cancers
3.	Mammography
4.	Menopause and hormone therapy
5.	Osteoporosis
6.	Pregnancy and childbirth
7.	Sexual health
8.	Women and heart disease
9.	Benign conditions affecting the function of the female reproductive organs

These working definitions, together with the categories proposed by the National Library of Medicine (Table 2), were used as a starting point for locating denominative variations in the initial structure of the conceptual field of Women's Health. For this purpose, after the identification of keywords, the Thesaurus functionality of Sketch Engine was used to generate lists of synonyms belonging to the semantic field under study. However, since the lists are produced based on the context in which the words appear in the selected corpus, the results are not always precise. In order to address any inaccuracies, the author implemented a manual revision process to identify synonyms

[2] https://www.nichd.nih.gov/health/topics/womenshealth
[3] https://medlineplus.gov/ency/article/007458.htm

with a greater similarity in meaning than what was provided by the automatic processing of Sketch engine.

3. DENOMINATIVE VARIATION IN WOMEN'S HEALTH LANGUAGE

This part of the analysis and discussion of the findings is divided into four sections where the results of the initial corpus analysis method are examined. After that, it concentrates on the particular discoveries regarding changes in vocabulary and grammar structure. This section thoroughly investigates the patterns, trends, and factors that impact the selection of denominations, revealing the ever-changing nature of language within the studied context. Through a detailed exploration of denominative variation, this analysis seeks to unravel the complexities inherent in the selection and usage of terms, offering insights into the factors that contribute to the diversity of denominations observed. This discussion is not only an exploration of linguistic nuances, but also a reflection on the broader implications of denominative choices on communication within the domain of medical articles on Women's Health.

3.1. PRELIMINARY APPROACH TO CORPUS: FREQUENCY LISTS

An initial analysis of the wordlist of the first 100 most frequent nouns in the corpus reveals that the most common medical specialties correspond to the nine categories proposed by the National Library of Medicine (Table 2).

Table 3. Most frequent nouns in corpus

Item / Frequency in corpus distribution			
woman	19,889	breast	1,427
health	19,508	Sex	1,380
pregnancy	5,259	depression	1,271
care	7,748	infection	1,213
risk	3,629	anxiety	1,087
birth	3,186	contraception	931
HIV	3,116	menopause	732
cancer	2,958	vaccine	663
child	2,722	syndrome	643
mother	2,464	obstetrics	629
covid-19	2,271	fertility	624
screening	1,754	cycle	617
abortion	1,494		

DOI: https://doi.org/10.6035/clr.7888

From this selection of the 25 most frequent nouns, terms such as *health*, *care*, *risk*, and *screening*, may be adequate for several of the specialties; however, other terms such as *pregnancy*, *birth*, *cancer*, *breast*, and *sex*, correspond to the specialties indicated by the National Library of Medicine (Table 2).

Next, a second analysis was carried out, this time on keywords, to find the frequency of the keywords in the focus corpus and reference corpus, which yields more significant data (Table 4) that will be useful for the detailed analysis in the second part of this analysis and discussion section.

Table 4. Frequency of keywords in focus and reference corpus

Item	Frequency (focus)	Frequency (reference)	Score
postpartum	1,880	90,608	272.3
antenatal	772	38,583	170.26
maternal	3,503	464,254	146.75
contraceptive	1,421	152,735	146.23
obstetric	623	39,101	136.81
menstrual	1,183	136,259	131.93
perinatal	659	57,364	122.46
cervical	1,501	223,325	116.18
contraception	931	141,071	101.42
gestational	566	63,981	99.70
cesarean	410	35,601	93.49
gynecology	477	53,483	91.77
ANC	817	137,519	90.62
HPV	757	124,127	90.05
postnatal	433	46,188	89.01
morbidity	739	129,885	85.27
menopause	732	133,100	83.07
pregnancy	5,259	1,356,464	81.67
IPV	302	22,429	79.85
depressive	624	111,815	79.56
endometriosis	455	72,003	75.43
childbirth	656	141,549	71.38
vaginal	782	180,815	71.26
prenatal	612	130,567	70.425
reproductive	1764	495,270	69.837

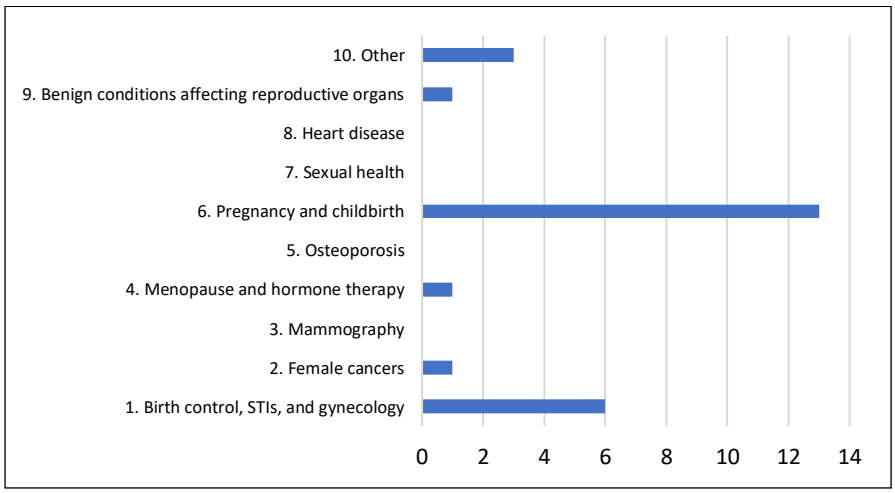
As in the results shown in Table 3, along with generic terms that apply to general Women's Health and clinical studies such as *cross-sectional*, *obstetric*, *morbidity* and *gynecology*, other terms highlight the most common specialties in Women's Health such as *postpartum*, *antenatal*, *contraceptive*, *menstrual*, *cervical*, *gestational.*

Although most specialties are represented in this list (Table 4), a more detailed analysis reveals that the focus of the publications is on specialty 6, pregnancy and childbirth, with almost 50% of the most frequent words in the corpus.

Also noteworthy is the appearance of abbreviations in this list of the most frequent words, ANC (antenatal care), HPV (human papilloma virus) and IPV (intimate partner violence). The latter, IPV, acquires special relevance since, although it is a term more commonly used in social and legal fields, its frequency of appearance (302 cases) in studies on Women's Health requires attention since it is usually included in studies related to mental health. Similarly, the word *depressive* with 624 occurrences or terms such as *depression*, *anxiety*, and *trauma*, which although do not appear in the list of the 25 most frequent words are very common in the corpus, support the need to include the section on mental health in studies on Women's Health.

Figure 1 shows the distribution of terms according to the initial Women's Health classification. In addition to the relevance of Category 6 (pregnancy and childbirth) it is also noteworthy that other specialties such as *heart disease*, *sexual health*, *osteoporosis*, and *mammography*, are not represented during this initial stage of corpus observation.

Figure 1. Distribution of terms and Women's Health specialties



3.2. FREQUENT MULTI-WORD TERMS

Finally, the extraction of multi-word terms was performed, which ultimately will illustrate more precisely the denominative variation object of this study. Using the Keywords function, Sketch Engine identifies what is unique in the focus corpus (medical articles) compared to the reference corpus (English Web Corpus 2021) and provides a list of multi-word units which are typical of a corpus or which define its content or topic and which will constitute the beginning of the detailed research on denominative variation in

 DOI: https://doi.org/10.6035/clr.7888

the field of Women's Health. Terms related to generic clinical research were removed from the list, such as *associated factor*, *formal analysis*, *odds ratio*, *systematic review*, or *cross-sectional study*, which appear frequently due to the textual characteristics of the genre studied. Only the terms related to Women's Health were selected for further study as shown in Table 5.

In line with the analyses of the previous lists (Tables 3 and 4) the specialties of *pregnancy and childbirth*, *cancer*, and *birth control* occupy the most representative positions. However, a large number of occurrences of terms such as *depressive symptom* and *intimate partner violence* stands out. A detailed study of these multiword terms reveals that of the 460 occurrences of *intimate partner*, the collocates found are limited only to words such as *abuse*, *violence*, and *homicide*.

Table 5. Most frequent multi-word terms from corpus

	Multi-word term	**Frequency (focus)**	**Score**
1	women's health issues	605	214.77
2	cervical cancer	1,060	169.47
3	women's health	1,076	168.64
4	maternal health	543	146.62
5	antenatal care	448	137.88
6	maternal morbidity	358	122.03
7	depressive symptom	526	121.26
8	maternal mortality	458	112.69
9	pregnant woman	1,523	104.3
10	intimate partner	447	101.4
11	reproductive health	595	91.09
12	unintended pregnancy	297	90.9
13	reproductive age	288	87.85
14	substance use	729	87.20
15	cervical screening	241	78.49
16	contraceptive method	248	78.12
17	severe maternal morbidity	208	74.40
18	postpartum period	225	73.88
19	contraceptive use	234	72.75
20	cervical cancer screening	218	71.44
21	menstrual cup	217	71.09
22	birth weight	319	67.75
23	maternal death	234	67.66
24	family planning	413	65.84
25	cancer screening	317	63.49

3.3. PATTERNS OF LINGUISTIC VARIATION: LEXICAL CHANGES

This section presents a closer view of the patterns of lexical change found during the analysis of the corpus with special attention to the cases where denominative variation is remarkable, as compared with the other instances of variation.

3.3.1. References to the subjects of the study: women

Although in the language of academic medical research reference to subjects of study is usually carried out using a gender-neutral language, in the corpus analyzed, the terminology used to refer to these subjects presents the highest degree of variation found during the study. As shown in Table 6, the method to refer to the clinical study subjects and the recipients of the medical treatment or research presents up to nine different forms (Table 6).

Table 6. References to subjects of study

Lemma	Occurrences in corpus
woman	19,890
participant	5,047
patient	3,631
mother	2,421
people	1,520
individual	970
female	391
lady	344
subject	214

As it can be seen in Table 6, the references to the subjects of study are mostly made by means of the lemma *woman*, with 50% of the cases. The occurrences of common terms such as individual and subject, are almost negligible, barely reaching 1%. In order to determine whether these variations correspond to a specific pattern, collocation analyses were carried out for each of the items in Table 6 and it was found that, as pointed out by different authors (Alarcón-Navío et al., 2016; Bowker & Hawkins, 2006; Freixa, 2006; Tercedor, 2011), it was due to stylistic variation in most cases, and in others, such as *participant*, *patient* and *subject*, it was a contextual variation as this is the most accepted terminology in the performance of surveys and clinical studies. However, Table 7 shows some examples in which many of the references to study subjects are made indiscriminately.

 DOI: https://doi.org/10.6035/clr.7888

Table 7. Examples of denominative variation in the notion of the subject of study.

lemma	pregnant	race/origin	age
participant	participants who were pregnant	English-speaking participants	participants who are aged 30 years
patient	Pregnant patient	Spanish-speaking patient	patient age <20
subject	nonpregnant subjects	subjects were Japanese women	elderly subjects
individual	pregnant individuals	White individuals	individuals aged 18
woman	pregnant women	black women	young women
female	pregnant female	Kosovan Albanian females	Females aged 9-45 years old
mother	pregnant mothers	Canadian mother	31 years old mother
lady	pregnant ladies	Ghanaian ladies	middle-aged ladies
people	pregnant people	Black people	people aged 50 years

As shown in Table 7, all the variants collocate with contexts referring to pregnancy, race/origin, and age. However, the variants *subject* and *participant* affect the naturalness or idiomaticity of the text (Bowker and Hawkins, 2006, p. 92) as they require further syntactic changes: *participants who are aged 30 years*, *subjects were Japanese women*… This mechanism to include pronouns, prepositions or copular verbs will be detailed later as it is probably one of the most frequent methods of variation.

Gender-inclusive language aims at including all individuals regardless of their gender identity or expression in order to promote gender equality and respecting diverse identities and experiences and seeks to avoid making assumptions about a person's gender and to be sensitive to the fact that not everyone identifies as strictly male or female. However, recent gender-inclusive language approaches in healthcare and medical settings promote linguistic changes to make biological sex less visible that have sparked the discussion between patients, clinicians, and academics. Linguistic changes aimed at gender inclusivity might inadvertently lead to a reduced visibility of the concept of biological sex, making it more challenging to articulate clearly in healthcare and medical education. With regards to gender-inclusive terminology, collocations with related terms such as *TNB*, *transgender*, *non binary*, *cisgender* or *agender*, mostly collocate with the gender-neutral language used in medicine as in the following examples: *genderfluid participant, Non binary participant, transgender and nonbinary (TNB) individuals*. However, in other cases such as *transgender*, the corpus analysis reveals a wider variety of options: as *transgender* collocates with nouns such as *women*, *population*, *persons*, *individuals*, *patients*, *young adults*, and *adolescents*.

To illustrate variation in context, different keywords have been chosen from the nine specialties selected for this study to observe recurring patterns that would help illustrate the cause for variation. From the results shown in Table 7 it can be concluded that the reason for the variation in the form of address to the subject of medical studies is mainly due to stylistic and contextual reasons, since the analysis of the corpus reveals that

all the variants seem interchangeable, even within the same research article. To refer to the subjects of clinical studies, *woman* is the most frequent word in the corpus, although in collocations with *pregnant*, to express origin or race, or age, are practically interchangeable with the other alternatives.

What seems clear is that *participant*, *patient*, and *subject* are more naturally used in contexts related to the description of studies, surveys, and experiments. Proof of this is that a very recurrent pattern of *participant/patient/subject* + verb is observed, as in the examples «*One participant described...*, *another patient remarked...*, *one subject shared...*», which is not observed in the rest of the variations. In addition, these three terms are frequently found in the methodology section.

Participant, with 2486 occurrences, appears as the subject of verbs such as *report*, *describe*, *express*, *share*, *state*, *experience*, *mention*, which reveals a clear interaction with the subjects of study and the subsequent analysis of the data obtained. With the term *patient* (896 occurrences), something similar happens, although instead of reporting verbs, the most common collocations are with verbs indicating reception of treatment, experiential verbs, or endurance verbs such as: *receive*, *undergo*, *experience*, *present*, *feel*, or *suffer*. Finally, in the case of *individuals* (970 occurrences), common in the medical scientific literature, it is located mainly in the introduction of the studies, indicating that they have not yet undergone the study, or in the analysis and discussion section, after the surveys and experiments, and is usually placed with verbs such as *be*, *seek*, *experience*, *suffer*. Some examples are shown below:

- A cystic fibrosis study found that 50% of **individuals** who had screened negatively...
- Many **individuals** choose not to disclose sexual victimization across their life span.
- **Individuals** were eligible for this study if they...
- 10 **individuals** did not sign the consent form to participate in the study....
- Black and Native American **individuals** are particularly burdened by...

Although with less intensity, only in 52 cases, the same happens with *subject* as subject of the verb: *participate*, *exhibit*, *undergo*, *report*, *experience:*

- **Subjects** reported personal barriers…
- Male **subjects** experience a stronger «inflamm-aging» syndrome…
- …the female **subjects** exhibited a higher antibody response…
- 24 **subjects** underwent vaginal delivery…

Finally, collocations with the word pregnant, confirm the preference in this type of publications for the use of *woman*. There exists variation, however negligible if we consider the number of occurrences of this collocation, 1523.

DOI: https://doi.org/10.6035/clr.7888

Table 8. Collocations with pregnant

Collocations with pregnant	Occurrences in corpus
pregnant + woman	1523
pregnant + mother	111
pregnant + lady	26
pregnant + person	16
pregnant + people	40
pregnant + patient	20
pregnant + participant	4

3.3.2. Synonyms

In medical language, it is common to use synonyms to make terminology more understandable to less specialized users. However, even though the corpus analyzed is aimed at a specialized audience, the use of synonyms is a frequent artifact for lexical variation.

Table 9. Lexical variations

Reference term	Occurrences in corpus	Lexical variant	Occurrences in corpus
menstrual cycle	190	individuals' cycles	1
		feminine cycle	3
		28-day cycle	24
		cycle	381
cervical cancer	650	cervix cancer	4
		cancers of cervix	1
		cancer of the uterine cervix	1
ovarian cancer	75	cancer of ovary	5
antenatal	772	prenatal	742
unintended pregnancy	297	unwanted pregnancy	170
		unplanned pregnancy	166
uterus	71	womb	7
colon cancer	4	colonic cancer	4
		cancer of colon	1
colorectal cancer	11	colon and rectal cancer	2
breast cancer	762	cancer of the breast	1

Table 9 reveals that, apart from *menstrual cycle/cycle* and *antenatal/prenatal*, most lexical variations lack significance in their occurrence frequency within the corpus. This observation aligns with the principle of monosemy and mononymy, which asserts that a term should denote a singular concept, and conversely, a concept should be denoted by a singular term. In fact, in most of the cases the alternative term is formed by lexical patterns (*cancer of ovary*) or ellipsis of the main term (*cycle*).

Other cases such as *mammography screening* run parallel in occurrences with *breast screening*, and as shown in the examples below:

- Adjusting the frequency of mammography screening…
- Periodicity of mammography screening…
- …have a limited understanding of the harms of breast screening…
- …the benefits of breast screening are less clear…
- …factors involved in breast cancer screening…

Although the corpus reveals no difference between the collocations of *unwanted*, *unplanned* and *unintended* with *pregnancy*, there seems to be a slight difference between them, as some concordances provide further explanation of the use of *unwanted* as mistimed, or something done at the wrong moment: «…associated with increased risk of having an unwanted or mistimed pregnancy…», «…and pregnancy intentions (wanted, mistimed, or unwanted)…», …«wanting to get pregnant but not at this time [mistimed pregnancy]…», «used to calculate the odds of having an unwanted or mistimed pregnancy versus wanted pregnancy…».

This mechanism is also frequent in the corpus, thus revealing that authors are aware of their use of denominative variations and at the same time express their concern for their use of medical terminology as self-evident and reflect the essential characteristics of the concept it designates.

Finally, it should be noted that *unplanned* appears exclusively in collocations with *pregnancy*, while *unwanted* appears with *sexual activity*, *sex*, and *unintended* also collocates with *consequences*, *outcome*, *births*, and *conceptions.*

3.4. PATTERNS OF LINGUISTIC VARIATION: MORPHOSYNTACTIC CHANGES

Table 10. Frequent patterns of morphosyntactic changes

Type of variation	Reference item	Occurrences in corpus	Alternative	Occurrences in corpus
morphological variants	gynecological	70	gynecologic	51
orthographic variants	gynaecological	100	gynecological	70
			gynecologic	51
	fetus	81	foetus	28
	anemia	297	anaemia	195
	cesarean birth	410	Caesarean birth	175
	dysmenorrhea	41	dysmenorrhoea	8
	edema	10	oedema	2
	hemorrhage	83	haemorrhage	6

 DOI: https://doi.org/10.6035/clr.7888

Type of variation	Reference item	Occurrences in corpus	Alternative	Occurrences in corpus
ellipted forms (pre-paid telephone card vs. phone card	menstrual cycle	190	cycle	381
graphical variation (online vs. on-line)	nonpregnant	26	non-pregnant	121
	prevention of mother to child transmission	16	Prevention of mother-to-child transmission	6
	Non-medical	15	nonmedical	9
	non-use	19	nonuse	5
permutation	cervical cancer	650	cancers of cervix	1
	ovarian cancer	75	cancer of ovary	5

In medical terminology, many terms are derived from Latin and Greek roots, and Latin spelling is often used to form the basis of these terms. However, modern medical terminology has undergone modifications over time to form comprehensive and precise terms. As illustrated with the examples from Table 10, in the case of *cycle*, it seems that the prevailing method is the omission and the use of the ellipted forms, although the reference term is *menstrual cycle*.

3.4.4. Abbreviations

The use of abbreviations in medical research articles is a widespread practice that aims to enhance clarity, conciseness, and readability. Abbreviations are used throughout the main body of medical articles to represent medical terms, study variables, and statistical measures to streamline the text and make it more accessible to readers. In our corpus, abbreviations deserve special attention since there is no specific pattern in the way they are integrated into the text. As in most specialized languages, in order to synthesize information, abbreviations are commonly used to present and discuss a wide range of concepts, conditions, and treatments. In our corpus, abbreviations appear mostly without the full form, as in the following examples:

- This study explored the postresidency provision of **EPL** management...
- Expanding **EPL** management in family medicine office-based settings...
- We defined medication management of **EPL** as using…

While in other cases it appears in parentheses after the complete form:

- One in five women will experience early pregnancy loss (EPL), or miscarriage,
- Early pregnancy loss (EPL) is a common experience.

Table 11. Use of acronyms vs. full forms in corpus

Acronym	Occurrences in corpus	Full form	Occurrences in corpus
HPV	757	Human papillomavirus	171
STI	110	Sexually transmitted infection	76
SMM	438	severe maternal morbidity	231
HCV	125	Hepatitis C virus	35
IPV	302	intimate partner violence	358
EPL	116	Early pregnancy loss	22
IUD	147	intrauterine device	76
SHR	118	sexual and reproductive health	141
PMTCT	51	prevention of mother-to-child transmission	22
PCMH	198	patient-centered medical home	61
ANC	817	antenatal care	478
AIDS	496	Acquired Immune Deficiency Syndrome	27
BSE	129	Breast Self Examination	77
CBE	14	Clinical Breast Examination	6
BC	94	birth control	139
ER	45	emergency room	33
PRAMS	98	Pregnancy Risk Assessment Monitoring System	43
LBW	213	low birthweight	32

One of the reasons for including the full form without context is due to the possibility of ambiguities as in the case of the acronym: CDC, which may represent *Cancer Prevention and Control*, *Communicable Disease Control*, and *Centers for Disease Control and Prevention*, and which was not possible to process accurate frequencies as in most cases it did not appear in the text with enough context to distinguish the correct one.

As an illustration, of the 757 occurrences of the acronym *HPV*, only in 31 instances it appeared in brackets after the full form, which is an indication of the consolidation of this acronym in the field.

- Human papillomavirus (HPV), a sexually transmitted disease,
- Awareness and attitude towards human papillomavirus (HPV) vaccine among medical students
- Self-sampling forhuman papillomavirus (HPV) testing

 DOI: https://doi.org/10.6035/clr.7888

The case of *Intimate partner violence* (IPV), or the pattern of abusive behavior within an intimate relationship where one partner seeks to assert power and control over the other, was mentioned before as it involves a significant public health concern but is not included in Medline indices. This is the only term that is more frequently used in its full form than its acronym.

4. CONCLUSIONS

For this research paper we conducted a comprehensive analysis of a corpus of medical articles to explore patterns, trends, and linguistic variations within the domain of Women's Health. By means of a corpus-based approach, the study focused on understanding the language used in medical literature, aiming to uncover insights into terminology denominative variation and key themes prevalent in the corpus analyzed. With this aim in mind, we employed Sketch Engine as linguistic tool and design an ad hoc methodology based on previous studies to extract meaningful information from the medical corpus and examine frequency patterns of specific terms and identify common collocations.

Overall, this research work on denominative variation of the terminological representation of Women's Health has revealed a nuanced landscape shaped by linguistic choices and identified several factors influencing the selection of terms, focused on linguistic motivation and the presence of accepted shortened forms such as abbreviations or acronyms.

Based on the evidence found in the corpus, we can conclude that in the field of Women's Health, lexical changes are not as frequent as expected in medical language, apart from the references to the subject of study - women - which as shown in the analysis and discussion section, is by far the most frequent lexical change in the corpus.

On the other hand, morphosyntactic changes correspond with the results found in the literature consulted (Bowker and Hawkins, 2006; Daille et al., 1996; Freixa, 2006; Velasco et al., 2013) and include morphological variants, orthographic variants, ellipted forms, graphical variation, and permutation.

Special attention was given to the use of abbreviations, which are prevalent in medical language, as they facilitate concise and efficient communication among healthcare professionals and are widely accepted within the medical community, although in the case of the corpus analyzed revealed a quite irregular usage.

This study highlights the significance of clear and consistent language in the domain of Women's Health, where effective communication is paramount. The observed denominative variations emphasize the need for a thoughtful approach to terminology selection, to ensure precision and comprehension in the communication between medical professionals, researchers, and the broader audience. The findings offer a foundation for future studies to explore deeper into the implications of denominative choices on communication and the overall advancement of healthcare practices in the specific context of Women's Health. Further research should consider the impact of linguistic choices on readability and comprehension, with implications for both healthcare professionals and the wider audience.

The shortcomings of variation lie in the incorporation of modern technologies and information retrieval systems in which terminology plays a crucial role in facilitating accurate and efficient access to information and where the use of precise and standardized terms is key. Thus, standardization, consistency, indexing, and synonym management are

among the aspects required for a smooth information retrieval workflow and consequently for a proper preparation of terminology management work.

BIBLIOGRAPHIC REFERENCES

Auger, Pierre and Rousseau, Louis-Jean (1987). *Metodologia de la recerca terminológica*. (M.T. Cabré, Trad.). Departament de Cultura de la Generalitat de Catalunya.

Alarcón-Navío, Esperanza, López-Rodríguez, Clara Inés, and Tercedor-Sánchez, Maribel. (2016). Variation dénominative et familiarité en tant que source d'incertitude en traduction médicale. *Meta*, 61(1), 117-144. https://doi.org/10.7202/1036986ar

Alcaraz, Enrique (2000). *El inglés profesional y académico*. Alianza Editorial.

Arntz, Reiner and Picht, Heribert (1995). *Introducción a la terminología.* (A. de Irazazábal, M. J. Jiménez, E. Schwarz, y S. Yunquera, Trads.). Pirámide.

Bowker, Lynne, and Hawkins, Shane (2006). Variation in the organization of medical terms: Exploring some motivations for term choice. *Terminology*, 12, 79-110. https://doi.org/10.1075/term.12.1.05bow

Cabré, Teresa (1993). La terminología. Teoría, metodología, aplicaciones. (C. Tebé, Trad.). Antàrtida/Empúries.

Daille, Béatrice, Habert, Benoît, Jacquemin, Christian, and Royauté, Jean. (1996). Empirical observation of term variations and principles for their description. Terminology. International Journal of Theoretical and Applied Issues in Specialized Communication, 3(2), 197-257. https://doi.org/10.1075/term.3.2.02dai

Dubuc, Robert y Lauriston, Andy (1997). Terms and Contexts. In S. E. Wright and G. Budin (Eds.), *Handbook of Terminology Management*. (pp. 80-88) John Benjamins Publishing.

Freixa, Judit, and Montané March, M. Amor (2006). Variación denominativa y biunivocidad en el lenguaje de las matemáticas. *Revista española de lingüística*. 36: 189-215.

Freixa, Judit (2006). Causes of denominative variation in terminology: A typology proposal. *Terminology*, 12(1), 51 – 77. https://doi.org/10.1075/term.12.1.04fre

International Organization for Standardization. (2022). *Terminology work — Principles and methods* (ISO 704).

Lerat, Pierre (1997). *Las lenguas especializadas.* Ariel.

Navarro, Fernando (2009). La precisión del lenguaje en la redacción médica. Quaderns de la Fundació Dr. Antoni Esteve, 89-104.

Peters, Pam, Qian, Yan, and Ding, Jun. (2018). Translating medical terminology and bilingual terminography. Lexicography, 3(2), 99–113. https://doi.org/10.1007/s40607-018-0037-y

Picht, Heribert (1987). Terms and their LSP environment-LSP Phraseology. *Meta*, 32, 2, 149-155. https://doi.org/10.7202/003836ar

Resche, Catherine (2000). An Approach to Interface Terminology: The Example of Environmental Economics in English as a Foreign Language. *Meta* XLV 4, 628:645. https://doi.org/10.7202/003941ar

Sager, Juan Carlos (1990). *A Practical Course in Terminology Processing.* John Benjamins Publishing.

 DOI: https://doi.org/10.6035/clr.7888

Tercedor-Sánchez, Maribel (2011). The cognitive dynamics of terminological variation. *Terminology*. 17(2), 181-197. https://doi.org/10.1075/term.17.2.01ter

TERMCAT. (1990). *Metodologia del treball terminològic*. Generalitat de Catalunya.

Wiese, Ingrid. (2018). Terminology Work in Different Domains: Medical Terminology. In John Humbley, Gerhard Budin, and Christer Laurén (Eds.), *Languages for Special Purposes* (pp. 522-534). De Gruyter Mouton. https://doi.org/10.1515/9783110228014-027

Wright, Sue Ellen (1997). Term selection: The initial phase of Terminology Management. In Sue Ellen Wright, and Gerard Budin. (Eds.), *Handbook of terminology management.* (pp. 13-24). John Benjamins Publishing.

Wright, Sue Ellen and Wright, Leland. (1997). Descriptive Terminology: Terminology Management for Technical Translation. In In Sue Ellen Wright, and Gerard Budin. (Eds.), *Handbook of terminology management* (147-159). John Benjamins Publishing.

Wüster, Eugene (1998). *Introducción a la teoría general de la terminología y a la lexicografía terminológica.* Universidad Pompeu Fabra.