UDC 81'33 **DOI: 10.18413/2313-8912-2022-8-1-0-4**

Irina G. Ovchinnikova[1]
Liana M. Ermakova[2]
Diana M. Nurbakova[3]

Online medical discourse during the COVID-19 pandemic: semantic categories and attitudes

[1] Sechenov First Moscow State Medical University
bld. 2, 8 Trubetskaya St., Moscow, 119992, Russia
E-mail: ovchinnikova.ig@1msmu.ru
ORCID iD: 0000-0003-1726-3360

[2] Université de Bretagne Occidentale
20 rue Duquesne - CS93837 Brest Cedex 3, 29238, France
E-mail: liana.ermakova@univ-brest.fr
ORCID iD: 0000-0002-7598-7474

[3]Institut National des Sciences Appliquées de Lyon
Université de Lyon
20 avenue Albert Einstein, Villeurbanne, 69100, France
E-mail: diana.nurbakova@insa-lyon.fr
ORCID iD: 0000-0002-6620-7771

Received 10 February 2022; accepted 05 March 2022; published 30 March 2022

Abstract. The paper presents the results of the discourse, semantic and sentiment analysis of the medical professional forum publications. Computer-mediated communication (CMC) within the professional community on the portal MirVracha reveals a diversity of peculiarities of medical professional discourse since the portal includes informal posts and chats alongside official and scientific publications. We study materials of MirVracha published in summer 2020 – winter 2021 (dataset includes more than 0.5 million words). We reveal dominant semantic categories of online medical professional discourse and examine the dynamic of topics discussed in the publications against the backdrop of the pandemic situation and social events. The dominant semantic categories of online medical professional discourse during the COVID-19 pandemic are the Diagnosis and Research categories. Based on the results of discourse analysis, we describe the interests and attitudes of the portal users and linguistic means they use to verbalize their attitudes within the professional community. The medical professionals are primarily interested in research materials and outcomes. Based on sentiment analysis, we uncover attitudes revealed in the publications. Medical professional CMC shows mainly neutral attitudes to the topics. In the posts and personal narratives, the portal users discuss political context and express the negative attitude in posts associated with semantic Diagnosis and Bureaucracy categories.

Keywords: Online Medical Discourse; COVID-19; Doctor-Doctor Communication; Discursive Category; Sentiment Analyses

How to cite: Ovchinnikova, I. G., Ermakova, L. M. and Nurbakova, D. M. (2022). Online medical discourse during the COVID-19 pandemic: semantic categories and attitudes, *Research Result. Theoretical and Applied Linguistics*, 8 (1), 57-78. DOI: 10.18413/2313-8912-2022-8-1-0-4

УДК 81'33 **DOI: 10.18413/2313-8912-2022-8-1-0-4**

Овчинникова И. Г.[1]
Ермакова Л. М.[2]
Нурбакова Д. М.[3]

Сетевой медицинский дискурс во время пандемии COVID-19: семантические категории и сентимент-анализ

[1] Сеченовский университет
Трубецкая ул., стр. 8, корп. 2, Москва, 119992, Россия
E-mail: ovchinnikova.ig@1msmu.ru
ORCID iD: 0000-0003-1726-3360

[2] Университет Западной Бретани
ул. Дюкен, 20, CS93837 Брест Седекс 3, 29238, Франция
E-mail: liana.ermakova@univ-brest.fr
ORCID iD: 0000-0002-7598-7474

[3] Национальный институт прикладных наук,
Университет Лиона
авеню А. Эйнштейна, 20, Виллербанн, 69100, Франция
E-mail: diana.nurbakova@insa-lyon.fr
ORCID iD: 0000-0002-6620-7771

Статья поступила 10 февраля 2022 г.; принята 05 марта 2022 г.;
опубликована 30 марта 2022 г.

Аннотация. В статье обсуждаются результаты исследования медицинского профессионального дискурса во время пандемии на материале публикаций на портале МирВрача, предназначенном для медицинских работников. Портал дает возможность изучить разные жанры сетевой медицинской коммуникации (блоги, чаты, научные статьи и дискуссии, официальные сообщения). Датасет включает около 800 сообщений объемом 0.5 млн. слов, опубликованных за полгода пандемии Covid-19 (с 29/06/2020 по 26/02/2021). Для обработки датасета использована модель ‘bag-of-words’. Благодаря применению дискурсивного и семантического анализа, а также сентимент-анализа, выявлены основные семантические категории сетевого медицинского дискурса в профессиональном сообществе, определены наиболее популярные темы в зависимости от стадии развития пандемии и социального контекста. Выявленные семантические категории Диагноз, Исследование, Администрирование, Вмешательство (Лечение) отражают структуру профессиональной коммуникации и соответствуют функциональным моделям общения ‘сбор информации’, ‘предоставление и обсуждение данных’, ‘принятие решения’. В наблюдаемый период доминирующими семантическими категориями оказались Диагноз и Исследование. Семантическая классификация частотных слов позволила выявить, что наибольший интерес у пользователей портала вызывает обсуждение результатов научных исследований, поскольку портал предоставляет обзоры актуальных научных публикаций в ведущих англоязычных медицинских журналах сразу же после выхода номера. Теме пандемии посвящено не менее 15% постов ежемесячно в течение наблюдаемого периода. Сентимент-анализ позволил

выявить некоторое нарастание отрицательного отношения к регулированию работы врачей при пандемии на фоне преобладающей нейтральной тональности публикаций на портале. Отрицательная оценочность присутствует при обсуждении диагностики и лечения пациентов с диагнозом COVID-19.
Ключевые слова: Сетевой Медицинский Дискурс; Коммуникация 'Врач-Врач'; COVID-19; Дискурсивная Категория; Сентимент-Анализ
Информация для цитирования: Овчинникова И. Г., Ермакова Л. М., Нурбакова Д. М. Сетевой медицинский дискурс во время пандемии COVID-19: семантические категории и сентимент-анализ // Научный результат. Вопросы теоретической и прикладной лингвистики. 2022. Т.8. №1. С. 57-78. DOI: 10.18413/2313-8912-2022-8-1-0-4

1. Introduction

Medical (healthcare) professional discourse is studied in various aspects according to social objectives of the healthcare industry. Medical discourse has peculiarities of institutional discourse; therefore, it requires particular ethical norms and strict bureaucratic and legislative regulation (Kuipers, 1989; Waitzkin, 1989). Specific sign systems are involved in the medical communication to define the function and status of the professionals: Dress code, test and observation schedules, Latin terminology, etc. (Good, & Del Vecchio Good, 1981; Khan, 2019). The ideas of medical discourse belong to 'scientific ideology' (Habermas, 1970). Involuntary expressing of sentiments and feelings is restricted by professional ethics (Heritage & Maynard, 2006; Madzhaeva, 2015). Doctors' assessments of politics and social contexts can bring medical misinformation in the public sphere (Caulfield et al., 2019; Ovchinnikova et al., 2021). However, the power of restrictions varies in different cultural and social context of medical professional communication (Good & Del Vecchio Good, 1981). Thus, research of medical communication offline and online in different countries is of importance for understanding the medical professional discourse theory and practice.

Doctor-patient communication attracts attention of researchers due to its social value. Researchers examined the communicative strategies, manifestation of doctor's empathy, types of communicative situations and the classification of patients' behaviour (Ferguson & Candib, 2002; Bogomiagkova & Iskanderova, 2020). Sentiments in doctor-patient communication are examined through semantic and discourse analysis as well as by tools of sentiment analysis (see review in: Zunic et al., 2018). Doctor-doctor communication appears to be less attractive for researchers. Medical discourse within the professional community includes genres and contexts that lack openness and necessity to express empathy (Ovchinnikova et al., 2020).

Discourse among medical professionals offline and online is described in terms of stylistics, terminology, genres and objectives (Akhnina et al., 2018; Shuravina, 2013; Waitzkin, 1989). The genres in medical online communication are partly described in (Akhnina, 2015; Varnavskaya, 2019). Representation of semantic categories and peculiarities of online medical discourse differs from that of offline discourse since network medical communication is less formal (Akhnina et al., 2015). During the COVID-19 pandemic, widespread online medical professional communication provides researchers with an opportunity to study semantic categories, sentiments and cultural peculiarities of discourse in the new technological environment.

In 2020, the healthcare professionals searched effective treatment of COVID-19. They had to lean on inconsistent data and controversial research results published in

medical journals. Inconsistent demands of the WHO administration, uncertainty of the scientific studies and necessity to provide immediate practical solutions put the professionals under pressure. The crisis can transform professional discourse and bring in expressing sentiments in online communication among healthcare professionals. Online medical professional discourse during the current pandemic did not obtain accurate analysis. We would like to cover this gap.

The objective of the paper is to reveal semantic categories of online medical professional discourse and analyze professionals' attitudes (positive, negative and neutral) during the COVID-19 pandemics. We intend to clarify peculiarities of this discourse in the time of the current crisis, carry out sentiment analysis of the publications on the medical professional portal, and examine the transformation of semantics and sentiments from June 2020 to February 2021 when the situation with COVID-19 appeared to be more stable thanks to the beginning of vaccination.

We lean on the following peculiarities of online doctor-doctor discourse:

- The discursive semantic categories are connected to the structure of medical investigation and scientific publication
- Medical professional discourse corresponds to the scientific ideology
- Norms of doctor-doctor professional communication restrict discussing political and social context and expressing emotions.

We aim to answer two research questions:

RQ1: What are the dominant discursive semantic categories in online medical professional discourse?

RQ2: What dominant sentiments are revealed in this discourse during the COVID-19 pandemic?

2. Related Studies

2.1. Peculiarities of Communication among Healthcare Professionals

Bureaucratic and legislative arrangement of medical communication supports the patient centeredness, doctors' professional responsibility and the high value of specialists' opinions (Kuipers, 1989). We expect to obtain a set of means to express scientific knowledge, professional approach and partnership in online posts and publications of healthcare professionals.

Communication within the healthcare professional community presupposes applying peculiar procedures and medical terminology. Sheaff and co-authors (2017) described divergence between content of patience oral narratives and information extracted from these narratives by professionals for medical sense-making. Healthcare professionals discuss this information following the norms and scripts of medical communication, which presupposes priority of scientific knowledge, critical thinking and impersonal professional approach. Kuipers (1989) mentioned that standardized way of medical descriptions allows to smooth stress and to avoid emotional involvement with patients' experience.

Scientific knowledge and medical experience are represented in a set of semantic categories of medical discourse within the professional community. The set contains the semantic categories of ethics, legislative and bureaucratic regulation, contagion, diagnosis, intervention, treatment, and research (Waitzkin, 1989). The set of the discursive semantic categories reflects the structure of the professional sphere, cycles of medical investigation and treatment. Wilce (2009) described text structure in medical discourse: introduction of a problem, review of familiar approaches to solve the problem, explanation of a new idea to contribute to a new solution, clear outcome to apply into medical practice. The structure of medical scientific publications contains templates that match the professional discourse structure. Processing the corpus of annotated abstracts of the medical publications, Nye and co-authors (2018) un-

covered the elements of the templates: data description (patients, diagnoses), selection of treatment and procedures (intervention), outcome (intervention consequences), and deliberation of advantages of a new approach. Thus, transparent and clear structure is a peculiar feature of professional discourse within the healthcare community. Discursive semantic categories are needed to be described in their relation to cycles of medical investigation.

In professional discourse, medical language generally avoids any critical attitude to political conflicts and a social environment (Waitzkin, 1989: 220). However, the evaluation of the social environment indirectly penetrates into medical discourse affecting classification of diseases and drugs. The social and cultural context affects healthcare professional communication. Ferguson and Candib (2002) proved a significant impact of cultural diversity on the efficacy of doctor-patient and doctor-doctor communication: a doctor more smoothly interacts with colleagues who share a similar cultural background. Professional communication, medical care and the distribution of a disease vary by social, cultural, and ecological contexts that generates grounds for diversity in medical discourse (Good & Del Vecchio Good, 1981). Thus, ethnic and cultural variety affects medical professional communication regardless of all restrictions. Studies of medical professional discourse in different countries are of importance for communicative practice and discourse theory.

Medical professional communication is described in the aspect of the functional models. The models emphasize objectives of the act of communication. De Haes and Bensing (2009: 290) offered a set of six function models of medical communication: Fostering the relationship, gathering information, providing information, decision making, enabling disease and treatment related behavior, responding to emotions. The function models are employed in doctor-patient communication and in communication among professionals as well. Decision making and discussing an appropriate treatment concern the interaction among professionals. The priorities and goals of communication with a patient do not match the objectives and intentions in interaction among medical professionals. The most important distinctive feature of the doctor-patient communication is the high level of empathy (Heritage & Maynard, 2006). A healthcare specialist aims to produce an adequate diagnosis and a treatment plan, resolve a decisional conflict, and achieve professional satisfaction while communicating within professional community.

2.2. Studies of Medical Professional Discourse Online

Online healthcare professional discourse includes scientific publications, materials published on websites of healthcare organizations, portals for discussing clinical software and exchanging data, professional forums and networks, etc. Medical professional CMC shows hypertextuality, multiview topical content and diversity of the participants, among whom there are patients, their relatives, students, healthcare professionals who address colleagues as well as non-professionals (Akhnina et al., 2015; Akhnina, 2018). Variety of illustrations belongs to the set of distinctive features of the medical online publications because of the necessity to popularize the advantages of evidence-based medicine (Silletti, 2016).

CMC provides professionals with the possibility to apply algorithms while processing medical communication online and generating data collections (Alnashwan et al., 2019; Goeuriot et al., 2012; Nye et al., 2018). The medical software and CRM became new professional computer tools. Thus, CMC impact is obvious in professionals' communication with patients and lay people as well as in

informal chat within the professional community. Network professional communication is characterized by a combination of edited and spontaneous utterances and instant reactions from community members (Akhnina et al., 2018). Formal CMC in the healthcare professional community follows the standards of professional ethics. However, variety of the semantic categories of online medical discourse requires deep analysis.

The COVID-19 pandemic triggered medicalization of Russian online media discourse; medicalizations brought in medical terminology, neologisms and slang (Severskaya, 2020). Medical terminology is often distorted in media and social networks; different terms substitute each other that brings in topic shifts, confusion and misunderstanding (Ovchinnikova et al., 2021). Nevertheless, it is still unclear whether the medicalization of the online media affects the norms of online professional medical discourse. Thus, the study of discursive semantic categories in CMC among medical professionals during the pandemic allows for understanding relevant topics and sentiments in medical professional discourse in the time of crisis.

2.3. Sentiment Analysis of Medical Discourse

The dominant-specific sentiment analysis of healthcare discourse examines emotions and polarity of materials published on professional websites, forums and online scientific journals. Sentiment analysis allows detecting sentiments and attitudes. The basic sentiment analysis software classifies the input (dataset of texts) and returns the evaluation of the text class: negative or positive polarity, neutral class, skip. Texts from neutral class are unmarked for sentiments. ‘Skip’ class contains unclear cases (noisy texts, borrowed content, etc.). The texts are processed as ‘bag-of-words’ where the software identifies lexemes with emotional connotation; based on the value of the lexemes in the text, it obtains the class (positive or negative polarity). The advanced sentiment analysis is able to recognize basic emotions.

Sentiment analysis in healthcare is carried out with the help of few publicly available domain-specific corpora and lexicons (Denecke et al., 2015). Zunic, Corcoran and Spasic (2020) generated a review of the publications about sentiment analysis in medical discourse mentioned on PubMed and MEDLINE. According to the review, researchers carried out the sentiment analysis on the data from social networks for information exchange among patients and their relatives. As far as we consider, the sentiment analysis is regularly applied to doctor-patient communication. In the clinical narratives, opinionated terms are extremely rare (between 5% and 11%, according to (Denecke et al., 2015)). The attitudes of medical professionals do not attract the researches who consider the attitudes to be predictable. However, healthcare practitioners are interested in the results of sentiment analysis since they intend to enhance professional communication.

Doctor-doctor and medical professional communication attracts the researches of sentiments and attitudes over the last few years. Torshin with co-authors (2020) worked out a system to evaluate sentiments in publication on PubMed and MEDLINE over the last 50 years. The authors intend to apply their results for recognizing expert messages on health forums and evaluating quality of the professional’s publications (Torshin et al., 2020). Putting aside evaluation of the system, we would like to mention that 18.5% of medical texts showed the level of negative polarity inappropriate in scientific discourse. Although the scientific discourse tends to be neutral, recent studies demonstrate that preprint abstracts produced at the beginning of the COVID-19 crisis stand between uncertainty and over-promotion (Bordignon et al., 2021). In the time of the pandemic, the Russian dataset is worked out to study sentiments of social media users. The manually annotated dataset allows for training machine learning and extracting sentiments

from the comments about COVID-19 (Nugamanov et al., 2021).

In the current research, we discuss the new results and evaluate the polarization of medical professional discourse on the different stages of the COVID-19 pandemic.

3. Material and Methods

We study the professional portal Mir Vracha (‘МирВрача’: The doctor’s world) based on RSS data from 29/06/2020 to 26/02/2021 (745 distinct entries in Russian in XML format; 522,744 tokens in total in concatenated texts). The portal was created in 2011 by a medical team. It contains medical professional materials and the blog platform for healthcare workers. The policy of the portal follows the principles of evidence-based medicine. To obtain the full access to the content, a user has to prove that he / she belongs to the medical professional community (to fill in his / her professional ID). Only certified professionals and medical university students are allowed to comment and publish their texts on the MirVracha forum. The web resource counts more than 100,000 users.

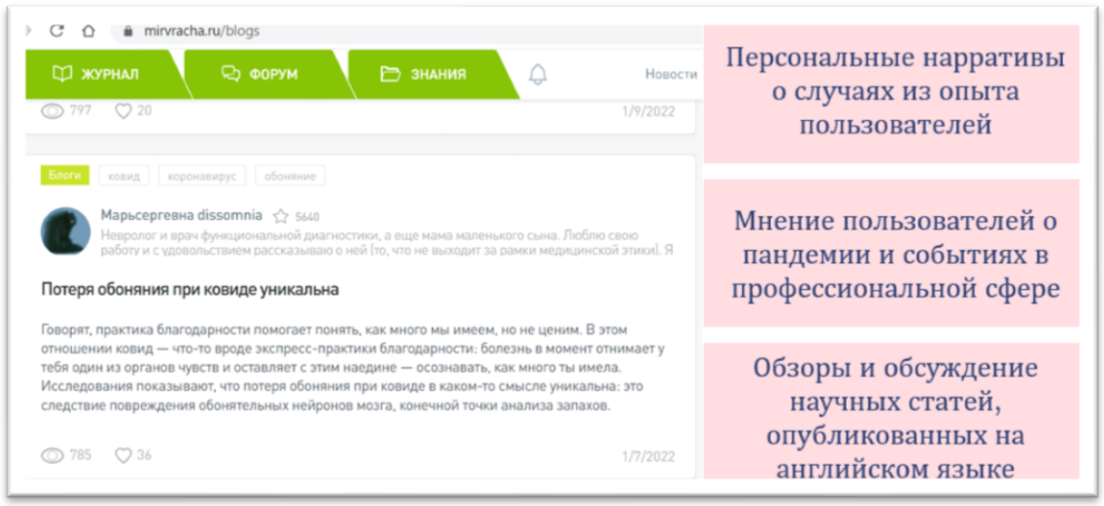
The portal contains scientific and relatively popular materials to share knowledge with specialists from different domains (see an instance in Figure 1). We use the ‘bag-of-words’ approach to the dataset that allows us to analyze the content of the portal ‘en bloc’. We analyze the following publications:

(a) personal narratives describing participants’ professional experience;

(b) participants’ blogs and discussions of articles published in medical scientific journals;

(c) the forum.

Figure 1. MirVracha portal: Blogs and their content



Each entry contains title/topic, link, description, and content, e.g.:

(1) <title>Хроники вакцинной гонки</title>
<link>https://mirvracha.ru/news/hroniki_vacconoy_gonki-05-08-2020</link>
<description>Чиновники рапортуют, что мы впереди планеты всей, специалисты недоумевают, представители всемирной организации здравоохранения разводят руками...</description>

We performed the following procedures applying discourse analysis, semantic analysis and sentiment analysis:

1. Texts from the fields title, description, and content were concatenated, lower-cased and tokenised with nltk.word; thus we arranged dataset into a suitable format and removed confidential data from the material (Ovchinnikova et al., 2020: 102).

2. We automatically generated the word frequency list and manually analyzed semantic groups in the top 10,000 frequent words.

3. The top 10,000 frequent words were compared with a list of terms in the Dictionary of medical terminology[1] that allowed us to distinguish medical terminology and basic lexicon of the Russian language. Medical terminology represents a peculiarity of scientific ideology.

4. We examined the high frequency meaningful words (20+ occurrences) to obtain reliable results. We analyzed the frequent words distributions in our dataset to uncover the topics the words belong to and identified the dominant topics of the publications in the dataset according to their titles, frequent and key words.

5. We applied semantic analysis to distribute the frequent words over the discursive categories.

6. We extracted 4-grams with the references to COVID-19. In the 4-grams, we uncovered typical neighbors and syntactic structures that allowed us to recognize scientific style.

7. We performed sentiment analysis of the posts with the Dostoevsky sentiment analysis Python library for Russian[2]. It processes the given texts and classifies their sentiment orientation as positive, negative or neutral. When the text contains many lexemes with positive connotations, it obtains positive polarity; when lexemes with negative connotations prevail in the text, it obtains negative polarity. Dostoevsky's model is trained on the RuSentiment dataset consisting of more than 30,000 comments from the Russian social network VKontakte[3].

4. Results

4.1. Description of Representativeness of Discursive Semantic Categories

Leaning on the description of the set of the categories in Waitzkin (1989), Kuipers (1989) and Nye et al. (2018), we distinguish four main semantic categories of online medical discourse in our dataset: Diagnosis, Research, Intervention, and Bureaucracy (including regulation, administration and management). To determine a semantic category for a lexeme, we examine its contexts. For instance, lexeme *больной* ‘sick man / woman’ mainly occurs in blogs and forum publication where diagnostic procedures are discussed; thus, it belongs to the Diagnosis category.

The list of the frequent words is given in Table 1. The monthly distribution of the posts is shown in Figure 2. Healthcare professionals participated in discussions and published approximately 70 posts per month. We investigate the monthly distribution of the categories represented by the most frequent words in the posts in order to discover the development of users’ preferences and to understand a connection between social events and topics of the discussions on the portal.

[1] https://medicinskie-terminy.slovaronline.com/
[2] https://github.com/bureaucratic-labs/dostoevsky/issues/146
[3] https://russiannlp.sites.stanford.edu/resources

Table 1. The top high-frequency nouns (100+ occurrence)

№	WORD	COUNT
1	год (лет) ‘year’	1887
2	пациент ‘patient’	1846
3	covid-19	1494
4	риск ‘risk’	847
5	человек (люди) ‘people’	796
6	время ‘time’	706
7	исследования ‘research’	669
8	лечение ‘treatment’	528
9	заболевание ‘disease’	451
10	авторы ‘authors’	449
11	жизнь ‘life’	443
12	течение болезни ‘course of the disease’	407
13	sars-cov-2	398
14	больной ‘sick man / woman’	204
15	данные ‘data’	161

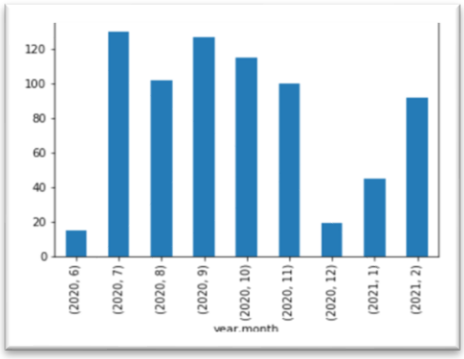


Figure 2. Publications on the MirVracha portal per month

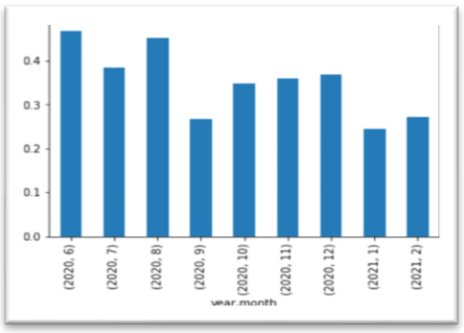


Figure 3. Share of Posts on ‘Diagnosis’ per month

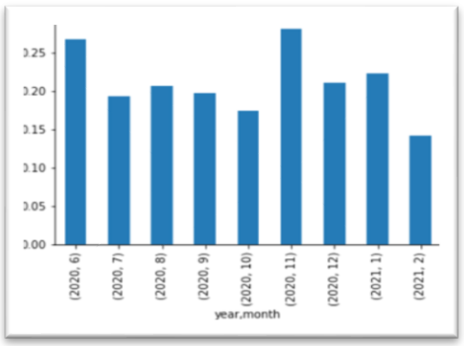


Figure 4. Share of Posts on ‘Research’ per month

4.1.1. General Characteristics of Publications in Healthcare Professional CMC

The frequency of the top 10,000 words varies from 2045 (*это* 'this') to 6 occurrences (566 word forms). We manually examine semantic groups of the words applying componential semantic analysis.

We show the high frequency words distribution over the discursive semantic categories in Table 2 and the share of posts on a given category per month in Figures 3 – 7. A topic of a post was automatically recognized according to its title and keywords. We recognize the most popular topics at different stages of the pandemic thanks to monthly monitoring (see charts in Figures 3 – 7).

The lexemes in Table 1 mostly belong to the basic lexicon. Nevertheless, the term *covid-19* occupies the third line in the list of the high-frequency words; another reference to the disease is a betacoronavirus *sars-cov-2* (see Table 1: 1494 and 398 occurrences, respectively). The lexeme *вакцина* 'vaccine' has 16th rank in the list of high-frequency words.

The discursive categories are represented in the top of the list: Research (*пациент* 'patient', *исследование* 'research'), Diagnosis (*заболевание* 'disease'), Intervention (*лечение* 'treatment'). The patient centeredness of medical professional discourse is revealed in the high frequency of the lexemes referring to those who obtain professional care and assistance.

4.1.2. Discursive Semantic Categories in Online Medical Discourse

We performed semantic analysis on the set of the frequent words (20+ occurrences). The words are manually distributed across the categories of medical discourse according to the results of semantic analysis. We extracted the subcategory Pandemic from the Diagnosis category since our present study is concentrated on online medical professional discourse during the COVID-19 pandemic. The Pandemic subcategory contains names of the disease (*covid-19*, *sars-cov-2*) and peculiar lexemes referring to the prevention of contagion (*маска* 'mask', *карантин* 'quarantine', etc.). The category Bureaucracy contains the legislative and administrative regulation subcategories.

The lexemes are unevenly distributed across the categories (see Table 2). The distribution manifests the priority of the Research and Diagnosis categories over the categories of Intervention and Bureaucracy.

Table 2. Quantitative description of the discourse categories

DISCOURSE CATEGORIES /LEXICAL GROUPS	HIGH-FREQUENCY WORDS BELONGING TO THE CATEGORY / LEXICAL GROUP	TOTAL FREQUENCY
Diagnosis	Инфекция 'infection', боль 'pain', инсульт 'stroke', симптом 'symptom', вирус 'virus'	11974
Research	Автор 'author', данные 'data', пациент 'patient', результат 'result', исследование 'research'	10998
Bureaucracy	Здравоохранение 'healthcare', организация 'organization', Министерство здравоохранения 'Ministry of Health',	9577
Intervention	МРТ 'MRI', терапия 'therapy', лечение 'treatment'	9401
Pandemic	Covid-19, sars-cov-2, коронавирус 'coronavirus'	3492
Basic lexicon	Смерть 'death', здоровье 'health', жизнь 'life', риск 'risk'	9283

According to the data presented in Table 2, the richest category is Diagnosis. The category contains words from different semantic groups and fields. The Diagnosis category includes lexemes denoting diseases, human organs, symptoms and syndromes. It includes many frequent words since references to diseases and diagnostic tools occur in almost every discussion and description of the users' personal professional experience. In Figure 3, the share of the posts corresponding to this category is always higher than 20% of the publications per month; the highest share occurred in summer 2020, when the first wave of COVID-19 provided statistics to determine the diagnostic procedures and variety of the disease course.

The Research category contains lexemes that are typical for scientific research papers and are in use in scientific publications of healthcare professionals. The Research category is of the importance for the portal users since they discuss the current issues of international journals on the portal. A research paper has standard structure and includes certain parts; this structure causes repetitions of the certain words in the discussions regardless of its topic (see the third line in table 2). Thus, the Research category includes the set of the words that refer to the components of research process and research paper structure. The share of posts with the frequent words belonging to the Research category increased 20% of all publications in summer 2020 and November 2020 – January 2021 (Figure 4).

The words that function in the domains 'law', 'administration', 'management' belong to the Bureaucracy category. The share of posts discussing topics of the category does not exceed 6% of the publications (Figure 5). The users had been looking for valuable clarification of doctors' responsibilities in summer 2020 and argued about administrative means to prevent spreading the disease in winter 2020.

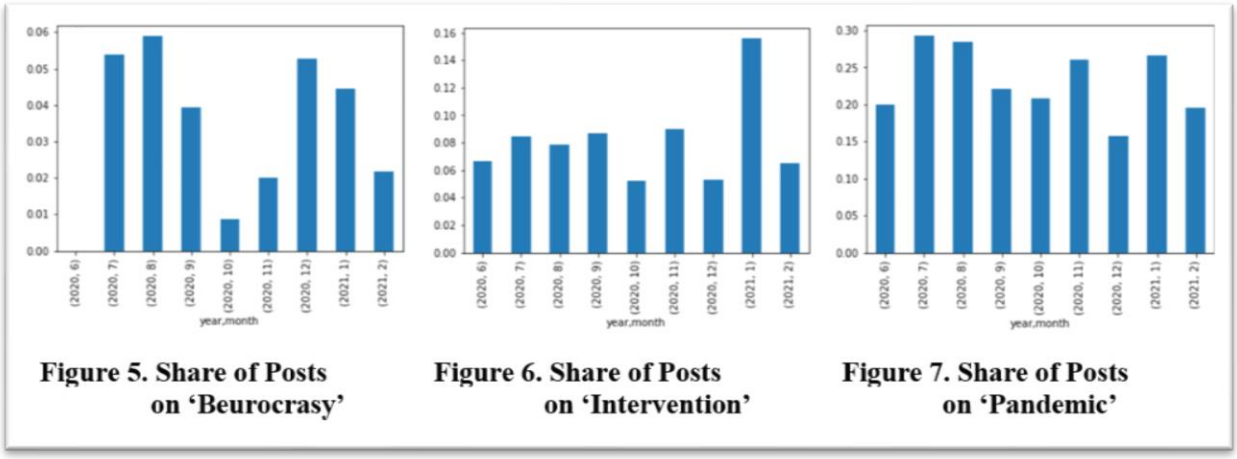


Figure 5. Share of Posts on 'Beurocrasy'

Figure 6. Share of Posts on 'Intervention'

Figure 7. Share of Posts on 'Pandemic'

Intervention was actively discussed in January 2021 (16% of the posts were devoted mostly to the treatment of COVID-19: see Figure 6), when professionals were ready to compare efficacy of different drugs in COVID-19 treatment.

The difference in the quantitative representativeness of the semantic categories through the high-frequency words is clear thanks to the total frequency of the words belonging to the category (see the third column in Table 2). This quantitative difference is not essential for all of the categories but the subcategory Pandemic. Pandemic contains a restricted set of the certain terms and words used to refer to the disease; the most typical of references to the coronavirus are acronyms *COVID-19* and *sars-cov-2*. The share of the posts devoted to COVID-19 varies from 29% in July to 16% in December 2020 (see Figure 7). Distribution of these posts is similar to that of the posts belonging to the Research category.

4.2. Analysis of texts about COVID-19

The posts devoted to the pandemic and the COVID-19 origin cover more than a quarter of the publication in July 2020. In the posts, the bloggers actively discussed the virus, its symptoms and peculiarities (see the text in Figure 1). They also mentioned the drugs (*азитромицин* 'azithromycin', *гидроксихлорохин* 'hydroxychoroquine') and previous pandemics (e.g., '*испанка*' "Spanish" influenza'). The forum participants often used terminology (*REM-сон* 'REM sleep', *энантема* 'enanthema', etc.).

They referred to articles published in the international medical scientific journals (Nature Microbiology, The Lancet, etc.). The style of the discussions is close to those of scientific and popular science publications. The stylistic features become obvious in n-grams with the names of the disease.

Table 3. Frequency of the character and words 4-grams

FREQUENCY INTERVAL	4-GRAMS	FRQUENCY
f ≥20	пациентов, с, covid-19, 'patients with…,'	39
	пациентов, с, covid-19; 'patients with….'.	20
20> f ≥10	covid-19, ,, по, сравнению '… compared with'	16
	пациентов, с, covid-19, и 'patients with… and'	14
	пациентов, с, новой, коронавирусной 'patients with new corona-virus'	12
10> f ≥5	covid-19, знакомит, читателей, с '…introduces readers with'	9
	covid-19, (, клинический, случай '…(clinical case'	9
	пациентов, с, covid-19, в 'patients with … in'	8
	пациентам, с, новой, коронавирусной 'to patients with new coro-navirus	7
	covid-19, в, критическом, состоянии '… in critical condition'	7
	covid-19, ,, а, также '…, and, also'	7
	covid-19, ,, а, также '…, and, also'	5
	covid-19, по, сравнению, с '…in comparison with'	5
5> f ≥2	covid-19, ,, но, и '…, but and'	4
	covid-19, и, при, гриппе '… and with the flu'	4
	covid-19, ., так, '… so / that'	4
	covid-19, ., так, '… so / that'	4
	sars-cov-2, и, других, коронавирусов '… and other coronaviruses'	3
	sars-cov-2, ,, а, также '…, and also'	3
	sars-cov-2, -, результаты, мета-анализа '… - results of meta-analysis'	3
	sars-cov-2, от, матери, плоду '… from mother to her child'	3
	sars-cov-2, ., кроме, того '…. Besides that'	2
	sars-cov-2, с, неинфицированными, . '…with non-infected.'	2
	sars-cov-2, ., информации, об '… . Information about	2
	sars-cov-2, (, 3, '… (3)	2
	sars-cov-2, в, крови, . '… in blood.'	2
	sars-cov-2, ,, а, потому '…, and therefore'	2

Table 3 includes frequent combinations of words with COVID-19 and other names of the disease. The 4-grams include words and punctuation marks that allow us to recognize syntactic constructions. The 4-grams show specific syntactic constructions, which are not

typical for online posts in social networks and forums for non-professionals (e.g., instances 3, 7, 19, 21 in table 3). The authors of the texts and their readers considered clinical cases, compared a course of the disease in different patients, looked for similarity between COVID-19 and flu, juxtaposed different coronaviruses, etc. The term COVID-19 is the most frequent name of the disease. In the 4-grams, the lexeme 'patient' occurs 100 times that shows the patient-centered character of the discussions about COVID-19 on the medical professional portal.

Discussion of the COVID-19 diagnosis and treatment on the forum contains references to users' case studies. The reference to the individual professional experience requires a unique description and attitude. Thus, the 4-grams from the unique descriptions do not appear among frequent combinations of words.

The portal users showed their attitude to COVID-19 and the current pandemic through irony and figurative language. The lexemes referring to feelings and emotions express negative assessment of the situation and its consequences (see lexemes belonging to the Russian basic lexicon in table 4). Using figurative language and tropes, the portal users attempted to avoid the anxiety (*охота за липовыми справками* 'hunt for fake documents'; *благие антикоронавирусные побуждения* 'good anti-COVID impulses'). Colloquialisms and slang manifest irony and sarcasm:

(2) (about the autopsy attendance: *В целом ничего страшного, просто задолбаетесь и получите интересный опыт* 'Generally, that's all right but you'll get bored and will receive an interesting experience';

(3) *Ни один врач-клиницист на вскрытие ковидных больных прийти не удосужился* 'Not a single clinician got around to come to the autopsy of covid patients').

Irony and humor are means to rethink a serious threat by diminishing its scope. Healthcare professionals used irony to avoid stress and pressure of their professional responsibility.

Table 4. Distribution of words with negative connotation over the posts[4]

MEDICAL DISCOURSE CATEGORIES	HIGH FREQUENCY WORDS WITH NEGATIVE CONNOTATIONS BELONGING TO THE CATEGORY	TOTAL WORDS FREQUENCY
Diagnosis	Бесплодный 'infertile', болезнь 'illness', болеть, заболеть 'to get sick', вирус 'virus', заболевание 'disease', инсульт 'stroke', инфекция 'infection', кровотечение 'bleeding', невынашивание 'miscarriage', ожирение 'adiposity', подросток-инвалид 'disabled teenager', симптом 'symptom', тошнить 'to vomit', угри 'acne'	545
Bureaucracy	Виноватый, виновен 'guilty', ложь 'deception', обвиняемый 'accused', обманщик 'deceiver', отставка 'resignation', уголовный 'criminal', тюрьма 'prison'	53
Basic lexicon	Бесполезный 'useless', бить 'to beat', боль 'pain', врать 'to lie', жалоба 'complaint', жуткий 'horrible', запрещенный 'prohibited', злой 'evil', идиотский	992

[4] Based on the data from summer 2020

MEDICAL DISCOURSE CATEGORIES	HIGH FREQUENCY WORDS WITH NEGATIVE CONNOTATIONS BELONGING TO THE CATEGORY	TOTAL WORDS FREQUENCY
	‘idiotic’, кануть в бездну ‘to disappear forever’, кошмар ‘nightmare’, мрачный ‘grim’, мучительный ‘agonizing’, немилость ‘disgrace’, обиженный ‘offended’, опасность ‘danger’, падать ‘to fall’, печаль ‘sadness’, плохой ‘bad’, потеря ‘loss’, проблема ‘trouble’, риск ‘risk’, рок ‘doom’, смерть ‘death’, стенание ‘lamentation’, страшный ‘scary’, угроза ‘threat’, ужас ‘horror’, умирание ‘dying’, умирать, умереть, ‘to die’, умирающий ‘moribund’, хуже ‘worse’	
Pandemic	Коронавирус, коронавирусный ‘coronavirus’ Covid-19, sars-cov-2, коронавирус ‘coronavirus’	248

4.3. Sentiment analysis of the MirVracha publications

The Dostoevsky sentiment analysis library for Python successfully processed our materials; it identifies unclear ‘skip’ sentiment status in 2,6% of the texts. The Negative texts less often occurred in fall 2020 – winter 2021; moreover, the healthcare professionals allowed themselves to show positive sentiments in winter 2021. The sentiment classification manifests a slightly negative attitude because of emotional discussions on the forum. Our results of the sentiment analysis confirm neutral character of medical professional online discourse (see Table 5).

Table 5. Sentiment classification of the texts

CLASSIFICATION	POSITIVE	NEGATIVE	NEUTRAL	SKIP
TEXT DESCRIPTIONS – FALL 2020-WINTER 2021 (774)	13(1.7%)	91(11.7%)	663(85.6%)	7(1.0%)
TEXT DESCRIPTIONS – SUMMER 2020 (185)	1(0.5%)	21(11.4%)	160(86.5%)	3(1.6%)

The high frequency words with negative connotations are unevenly distributed across the dominant semantic categories of medical professional discourse (see Table 4). Negative polarity is mainly associated with basic lexicon functioning in the healthcare domain (*смерть* ‘death’, *риск* ‘risk’, *боль* ‘pain’, etc.). Nevertheless, the Bureaucracy, Diagnosis and Pandemic categories also contain lexemes with negative connotations; thanks to the lexemes, several posts of these categories belong to the class of negative polarity. In Diagnosis, names of diseases have negative connotations that causes the negative polarity of the texts. Thus, deviations from the neutral class in sentiment classification are connected to discussions of diseases, regulation of the professional activity and narratives of the personal experience. The posts on Research and Intervention do not contain lexemes with negative connotations in their lists of high frequency words.

The lexemes with negative connotations refer to the diseases (4), current situation of the pandemic (5), socially inappropriate behaviour and the bureaucratic regulation (6):

(4) <pubDate>Wed, 15 Jul 2020 13:06:02 +0300</pubDate>
<title>Сколько нужно слез?</title>
<link>https://mirvracha.ru/discussion/skolko_nuzhno_slez-14-07-2020</link>
<description>Акушерам-гинекологам кажется, что невынашивания ужасно много. Это не так. Просто это так горько, так сложно, что каждая женщина становится особенной. Хрустальной вазой с ледяной корочкой боли. Исторически в РФ привычным невынашиванием считали пациенток, потерявшим 2 и более беременности в сроке до 24 недели</description>

(5)<pubDate>Thu, 23 Jul 2020 11:55:30 +0300</pubDate>
 Людям страшно жить, потому в тренде спекуляции на тему вакцинации от коронавирусной инфекции.

(6) <pubDate>Mon, 13 Jul 2020 15:38:09 +0300</pubDate>
 Врачам некогда писать извещения, а виноватой становится лживая статистика, не имеющая информации по 33-50% случаям заболеваний.

Healthcare professionals usually followed the ethical norms putting aside their political opinions. However, some texts contain a critical appraisal of the social protests during the COVID-19 pandemic (June 2020 in the USA, January 2021 in Russia: (7)).

(7) <pubDate>Mon, 13 Jul 2020 21:15:26 +0300</pubDate>
<title>Политкорректность съела Кука</title>
Противопоставление между протестами против расизма и контролированием COVID-19 – ложная дихотомия. Приведут ли толпы к большему числу случаев COVID-19? Да. Но к этому приводит и расизм, пишет British Medical Journal. Понимаете, утверждается, что расизм разжигает эпидемию в США! Но ведь нельзя же валить все беды на один фактор.

The openness uncovers that healthcare professionals consider the appraisal of the social context as relevant professional behavior during the current pandemic.

5. Discussion

5.1. General description of the medical professional discourse online

The MirVracha portal presents the variety of publications of different genres written by medical specialists for medical specialists. Medical professional network communication during the COVID-19 pandemic shows the general peculiarities of communication among healthcare professionals:

- Priority of scientific ideology and evidence-based medicine
- concentration on the professional issues
- patient centeredness
- explication of the medical professional discourse semantic categories
- explication of function models
- the professional ethical norms.

Based on the lexical and semantic analysis, we conclude that utterances and texts of medical discourse belong to scientific ideology because they present a search for a solution of the particular problem (Habermas, 1970). Their content corresponds to a ‘technical’ solution and procedures how to handle a patient with COVID-19. The scientific ideology presupposes objectivism,

neglecting personal preferences, analytic cognitive style, impersonalization.

In the portal material, we recognize the following functional models: gathering information, providing and sharing information, decision making, responding to emotions (De Haes and Bensing, 2009: 290). The first model corresponds to the Diagnosis category, the second one is connected to the Research and Intervention categories. Irony and figurative language manifests the latter functional model showing the doctors' responses to their own emotions. Our conclusion is consistent with the results of the studies of doctor-patient communication that show emotional involvement with patients and co-workers within healthcare professional community (Franz and Murphy, 2018). The professionals failed to avoid expressing emotions in doctor-doctor CMC.

The presence of emotions and assessments of the political context in communication among professionals represents a possibility of revealing the potential diversity in evidence-based medicine that Good and Del Vecchio Good (1981) connected to the diversity of social, cultural, ecological preconditions in different countries. As far as we consider, the open appraisal of political context and emotionality during the current pandemic appear in medical professional communication all over the world (Ovchinnikova et al., 2021). However, medical specialists evaluate political news and social events keeping in mind risk of spreading the infection. Penetrating the boundaries of professional ethics, medical professionals openly expressed their negative attitude to the social context since the protests put people in danger of getting infected by the highly contagious coronavirus. As compared to the posts of lay people about the pandemic, the discussions of news on MirVracha lack politicization and reference to conspiracy (Ovchinnikova et al., 2021). We consider the open negative evaluation of social events to be a peculiar characteristic of the medical professional discourse in the time of the pandemic. The assessments of social events are expressed in irony and figurative language. While appealing to irony and using colloquialisms and tropes in their texts, the portal users confronted the crisis in the healthcare area caused by spreading COVID-19. This language behavior allows the professionals to diminish the pressure they have been put under since the outbreak of the pandemic.

Communication among professionals on the MirVracha portal does not match the typical features of CMC. The publications on forum lack illustrations, internet slang, bright emotions, polarization of opinions; syntax of the texts is more complicated than that of the posts published on forums for patients and their family members or social networks (Akhnina et al., 2018; Alnashwan et al., 2019). The distinguished feature of healthcare professional network communication, as compared with the network communication of lay people about COVID-19, is its strict regulation by professional ethics.

5.2. The dominant semantic categories of online discourse of healthcare professionals (RQ1)

Scientific knowledge and concentration on the professional issues in the publications on the MirVracha portal provide its readership with the content associated with the dominant semantic categories of medical discourse. Thanks to the content diversity, the portal allows to uncover the discursive semantic categories that are relevant to the goals and structure of medical communication. The content of reviews and texts published on the portal covers the Diagnosis, Research, Intervention and Bureaucracy categories. The shares of text devoted to the topics that correspond to Diagnosis and Research categories were consistently high (above 15%) during the current pandemic. The lexemes from these categories includes terms that are required in every text of medical discourse since names of diseases, human organs, symptoms and syndromes allow to refer to any case of medical practice. Topics associated with the subcategory Pandemic

were always of great importance for the users. The portal users energetically participated in discussion about COVID-19 using figurative language and irony to experience engagement and satisfaction with online interaction (Yus 2018).

The Intervention and Bureaucracy categories were less popular. The category Intervention was revealed in the publications in January 2021 when the users have begun to discuss various vaccines and vaccination. The topics belonging to the Bureaucracy category generated a backdrop for consideration of social context and evaluation of professional ethical norms. The negative sentiments expressed in the posts on Bureaucracy show contradictions between management and the practitioners who feel overwhelmed by hard work during the pandemic.

The proportions of posts on the topics belonging to the certain semantic category reveal **shifts of the users' attention to particular content**. The high proportion of the publications on Research topics in summer 2020 matches the period of the active discussion of the publication in The Lancet journal about Hydroxycloroquine in the COVID-19 treatment and controversial recommendations of the WHO[5] (Ovchinnikova et al., 2021). In the end of 2020, vaccination and various vaccines were discussed within the healthcare community. The distribution of the posts on Pandemic is very similar to that of the posts devoted to Diagnosis and Research. The users and bloggers appear to be interested in scientific news about COVID-19, complications it causes, and efficacy of drugs. The significant number of posts on Intervention appeared in spring 2021 when medical professionals obtained the sufficient data to compare effect of different vaccines and verified protocols for COVID-19 treatment.

Thus, the shifts in quantitative representativeness of the semantic categories in the forum publications are affected by the enhancement in the research. In July-August 2020, the Russian healthcare professionals had been concentrating on the virus symptoms, the ways to recognize the virus and to prevent complications (Ovchinnikova et al., 2020). In winter 2021, the Russian medical specialists became involved in active discussion of Intervention; they also paid more attention to the management and administrative regulation of their activity. The necessity to reconsider the topics belonging to the Bureaucracy category is caused by new documents appeared at the critical stage of the pandemic.

The topics of medical professional interests **distinguish the professional approach from the lay people's point of view**. The users of social media and healthcare professionals showed the partially-overlapping interests. Abd-Alrazaq and co-authors (2020) clarified the most popular subjects of the COVID-19 discussions in Twitter: origin of the virus; its sources; ways of mitigating the risk of infection; the impact of the pandemic on people, countries and their economy. The professionals discussed origin of the virus, its similarity to other viruses and the risk of infection. As the COVID-19 pandemic evolves, the interests of social media users shift to topics belonging to the Diagnosis, Research and Intervention categories; social media generated and dissipated medical misinformation about Intervention (Arkhipova et al., 2020; Ovchinnikova et al., 2021). Russian speaking users of social media brought in terminology, neologisms and colloquialisms (Severskaya, 2020) that primarily correspond to the Diagnostic and Intervention categories. Fake news and myths about COVID-19 mentioned in the Russian social networks are also associated with the Diagnosis (the ways of getting infected by the virus) and Intervention (medical and alternative ways to treat the disease) categories (Arkhipova et al., 2020). Meanwhile, the politicization of the discussion of COVID-19 in social media is somehow connected with the Bureaucracy

[5] WHO Discontinues Hydroxychloroquine and Lopinavir/Ritonavir Treatment Arms for COVID-19. URL: https://bit.ly/ 3kq6iam (Accessed 15.12.2021).

category. Politicization appeared as a consequence of disappointment in the official policy with its limitations and the quarantine that failed to prevent spread of the coronavirus (Ermakova et al., 2021).

Medical professional discourse online is not affected by the disinfodemic. The results of our study comports with the description of the gap between lay and professional discourse in general (Molek-Kozakowska, 2017) and in healthcare domain in particular (Sheaff et al., 2017). Research appears to be the peculiar discursive semantic category in healthcare professional discourse that distinguishes communication among professionals. Medical professionals show analytic thinking and accurate analysis of accessible data instead of sharing anecdotal evidence and leaning on the opinions of public figures (Ermakova et al., 2021). The healthcare professionals who avoid discussing the peculiar subjects and sharing disinformation in the professional community mark the boundaries of the medical professional discourse.

5.3. Predominance of the publications with neutral polarity (RQ2)

Neutral polarity is expressed in more than 80% of the texts and posts published on the medical professional portal. The results of our sentiment analysis of online medical professional discourse shows high consistency with the offline official and scientific medical discourse of healthcare professionals where manifestation of emotions is strictly limited and regulated (Heritage & Maynard, 2006). Nevertheless, in general, CMC forces participants of online communication to deviate from norms and rules of language behaviour (Akhnina et al., 2018; Bogomiagkova et al., 2020; Kharitonova et al., 2019). This trend is shown through lexical diversity within the set of frequent words with negative connotations from the Russian basic lexicon.

The Diagnosis semantic category evokes emotional discussion since patients and doctors express their attitude to the topics belonging to this category in the online publications. Sentiment analysis of doctor-patient communication shows diversity in sentiments and attitudes. The analysis based on the SentiWordNet lexicon shows that expressed sentiments are associated with the medical condition and medication domain (Yadav, 2018); those domains correspond to the Diagnosis and Intervention semantic categories, respectively. Applying Multiclass Neural Network, Alnashwan et al. (2019) uncover patients' negative attitudes to infection, symptoms and treatment that also matches the Diagnosis and Intervention semantic categories. Thus, patients rather emotionally discuss the treatment (Intervention) while the medical professionals show their emotions discussing the Bureaucracy topics.

Troshin and co-authors (2020) found out that negative polarity and attitudes are contained in the scientific articles from the PubMed/MEDLINE database (Troshin et al., 2020). However, we are unable to compare our results with those in (Troshin et al., 2020) due to significant differences in sentiment analysis tools. In summer 2020 – winter 2021, 21.1% of the materials published on the medical professional portal expressed negative polarity. Since positive polarity was expressed in 1.4% of the materials, the value of the positive polarity in the online medical professional publication was unessential. The current pandemic does not significantly change the balance of neutral and emotional texts in medical professional discourse. However, the course of pandemic leads to decrease in share of positive words in the doctor-doctor communication that is obvious thanks to comparison of our results with the description of the sentiments in the beginning of the pandemic (Bordignon et al., 2021). Negative polarity supersedes the positive polarity in the medical publications.

Sentiment analysis carried out on Russian dataset by Nugumanov et al. (2021) show the distribution of negative and positive attitudes to the different aspects of the pandemic. The dataset shows that the users'

attitudes to Government are mainly negative (Nugamanov et al., 2021: 303). These results are consistent with the sentiments and polarity of the Russian professional portal MirVracha where negative polarity is uncovered in publications about management in healthcare (Ovchinnikova et al., 2020). We confirm the doctors' negative attitude to the regulation of their professional activity during the pandemic. Thus, lay people and healthcare professionals show negative attitudes to means of the official regulation of social and professional activity during the pandemic. Sentiment analysis appears to be the relevant tool to study professional CMC.

6. Conclusions

In our study of online medical professional discourse based on the materials published on the Russian medical portal during the COVID-19 pandemic, we clarified semantic categories and peculiarities of medical discourse in CMC among professionals.

1. Online medical professional discourse has common categories and peculiarities with doctor-doctor medical discourse offline. The distinctive feature of online medical professional discourse during the COVID-19 pandemic is expression of critical appraisal of the social context while discussing the current pandemic.

2. The discursive semantic categories of online medical professional discourse reveal semantic structure of the medical scientific publications and the norms of the medical professional communication including ethical and legislative regulation of the professional activity. Online medical professional discourse expresses the ideas belonging to scientific ideology.

3. Discursive semantic categories Diagnosis, Research, Intervention, and Bureaucracy are almost evenly represented in the list of the high frequency words. The set of the semantic categories covers the area of professional activity and its regulation. However, the publications are mostly devoted to the topics associated with Diagnosis and Research.

4. As the COVID-19 pandemic evolves, the share of online publications devoted to the Intervention category increases when the medical professionals obtain the possibility to discuss efficacy of different drugs and compare different vaccines.

References

Abd-Alrazaq, A., Alhuwail, D., Househ, M., Hamdi, M. and Shah, Z. (2020). Top Concerns of Tweeters During the COVID-19 Pandemic: Infoveillance Study, *Journal of Medical Internet Resources*, 22(4):e19016, April, available at: https://pubmed.ncbi.nlm.nih.gov/32287039/ (Accessed 20 February 2022). https://doi.org/10.2196/19016 *(In English)*

Akhnina, K. V. and Kurilenko, V. B. (2015). Medical net-discourse: essence, qualities, structure, *Polylinguality and Transcultural Practices*, 1, 76-84. *(In Russian)*

Akhnina, K. V., Makarova, M. A. and Yurmanova, S. A. (2018). Network medical discourse: lexical and stylistic features, *Online scientific & educational Bulletin "Health and Education Millennium"*, 20 (6), 48–53. https://dx.doi.org/10.26787/nydha-2226-7417-2018-20-6-48-53 *(In Russian)*

Alnashwan, R., Sorensen, H., O'Riordan, A. and Hoare, C. (2019). Accurate classification of socially generated medical discourse, *International Journal of Data Science and Analytics*, 8(4), 353–365. https://doi.org/10.1007/s41060-018-0128-8 *(In English)*

Arkhipova, A. S., Radchenko, D. A., Kozlova, I. V., Peigin, B. S., Gavrilova, M. V. and Petrov, N. V. (2020). Specifics of Infodemic in Russia: From WhatsApp to the Investigative Committee, *Monitoring of Public Opinion: Economic and Social Changes*, 6, 231–265. https://doi.org/10.14515/monitoring.2020.6.1778 *(In Russian)*

Bogomiagkova, E. S. and Iskanderova, L. V. (2020). Interaction between doctors and patients on medical forums: Analysis of Discourse Practices, *Internet and Modern Society: proceedings of the conference IMS-2020,* Saint-Petersburg, Russia, 22–24, available at: https://ojs.itmo.ru/index.php/IMS/article/view/1134 (Accessed 20 February 2022). *(In Russian)*

Bordignon, F., Ermakova, L. and Noel, M. (2021). Over-Promotion and Caution in Abstracts of Preprints during the COVID-19 Crisis, *Learned Publishing*, 34(4) 622–636. https://doi.org/10.1002/leap.1411 *(In English)*

Caulfield, T., Marcon, A. R., Murdoch, B., Brown, J. M., Perrault, S. T., Jarry, J. ... and Hyde-Lay, R. (2019). Health misinformation and the power of narrative messaging in the public sphere, *Canadian Journal of Bioethics/Revue canadienne de bioéthique*, 2(2), 52–60. https://doi.org/10.7202/1060911ar *(In English)*

De Haes, H. and Bensing, J. (2009). End-points in medical communication research, proposing a framework of functions and outcomes, *Patient education and counselling*, 74(3), 287–294. https://doi.org/10.1016/j.pec.2008.12.006 *(In English)*

Denecke, K. and Deng, Y. (2015). Sentiment analysis in medical settings: New opportunities and challenges, *Artificial Intelligence in Medicine*, 64(1), 17–27. https://doi.org/10.1016/j.artmed.2015.03.006 *(In English)*

Ermakova, L., Nurbakova, D. and Ovchinnikova, I. (2021). Analysis of Users Engaged in Online Discussions about Controversial Covid-19 Treatments, *Adjunct Proceedings of the 29th ACM Conference on User Modelling, Adaptation and Personalization* 2021, *June*, Utrecht, Netherlands, 162–166. https://doi.org/10.1145/3450614.3462239 *(In English)*

Ferguson, W. J. and Candib, L. M. (2002). Culture, language, and the doctor-patient relationship, *Family medicine*, 24, 353–361. *(In English)*

Franz, B. and Murphy, J. W. (2018). Reconsidering the role of language in medicine, *Philosophy, Ethics, and Humanities in Medicine*, 13(1), June. https://doi.org/10.1186/s13010-018-0058-z (Accessed 20 February 2022). *(In English)*

Goeuriot, L., Na, J.-Ch., Kyaing, W. Y. M., Khoo, Ch., Chang, Y.-K., Theng Y.-L. and Kim, J.-J. (2012). Sentiment lexicons for health-related opinion mining, *Proceedings of the 2nd ACM SIGHIT International Health Informatics Symposium, IHI '12, Association for Computing Machinery*, New York, NY, USA, 219–226. https://doi.org/10.1145/2110363.2110390 *(In English)*

Good, B. J. and Del Vecchio Good, M.-J. (1981). The semantics of medical discourse, in Mendelson, E. and Elkana, Y. (eds.), *Sciences and Cultures, Anthropological and Historical Studies of the Science,* Springer, Dordrecht, Netherlands, 177–212. https://doi.org/10.1007/978-94-009-8429-5 *(In English)*

Habermas, J. (1970). Technology and science as 'ideology', in Habermas, J. (ed.), *Towards a Rational Sociaty: Student Protest, Science, and Politics*, trans. Shapiro, J. J, Beacon, Boston, USA, 79–126. https://www.sfu.ca/~andrewf/technology_and_science_as_ideology.pdf (Accessed 2 February 2022) *(In English)*

Heritage, J. and Maynard, D. W. (2006). Problems and prospects in the study of physician-patient interaction: 30 years of research, *Annual Review of Sociology,* 32, 351–374. https://doi.org/10.1146/annurev.soc.32.082905.093959 *(In English)*

Khan, M. (2019). The perspective of medical communication on the biomedical model of practice and patient centeredness: A review of the language of medical case presentation genre, *International Journal of Linguistics, Literature, and Translation*, 2, 71–80. *(In English)*

Kharitonova, N. V., Baryshnikova, S. N. and Monastyretskay O. V. (2019). Peculiarity of lexical component of professionally oriented discourse in the internet (based on medical forum), *Philology. Theory and Practice*, 12, 331–334. https://doi.org/10.30853/filnauki.2019.3.70 *(In Russian)*

Kuipers, J. C. (1989). Medical discourse in anthropological context: Views of language and power, *Medical Anthropology Quarterly*, 3(2), 99–123. https://doi.org/10.1525/maq.1989.3.2.02a00010 *(In English)*

Molek-Kozakowska, K. (2017). Communicating environmental science beyond academia: Stylistic patterns of newsworthiness in popular science journalism, *Discourse & Communication,* 11(1), 69–88. https://doi.org/10.1177/1750481316683294 *(In English)*

Nugumanov, E., Loukachevitch, N. and Dobrov, B. (2021). Extracting Sentiments towards COVID-19 Aspects, *Supplementary 23rd International Conference on Data Analytics and Management in Data Intensive Domains, DAMDID/RCDL 2021*, Moscow, Russia, 299–312. http://ceur-ws.org/Vol-3036/paper24.pdf (Accessed 10 February 2022) *(In English)*

Nye, B., L., Junyi, J., Patel, R., Yang, Y., Marshall, I., Nenkova, A. and Wallace, B. (2018). A corpus with multi-level annotations of patients, interventions and outcomes to support language processing for medical literature, *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers),* Association for Computational Linguistics, Melbourne, Australia, 197–207. https://doi.org/10.18653/v1/P18-1019 *(In English)*

Ovchinnikova, I. G., Ermakova, L. M. and Nurbakova, D. M. (2021). Who Needs the Red Pill: Hydroxychloroquine Wars on Twitter, *Monitoring of Public Opinion: Economic and Social Changes*, 6, 566–583. https://doi.org/10.14515/monitoring.2021.6.1981 *(In Russian)*

Ovchinnikova, I., Ermakova, L. and Nurbakova, D. (2020). Sentiments in Russian Medical Professional Discourse during the Covid-19 Pandemic, *Proceedings of the Third Workshop on Computational Modeling of People's Opinions, Personality, and Emotion's in Social Media.* Barcelona, Spain, 99–108. *(In English)*

Severskaya, O. I. (2020). Covidiots on Coronacation: Coronaviral Lexicon as a Diagnostic Field of Discursive Practices, *Communication studies*, 7 (4), 887–906. https://doi.org/10.24147/2413-6182.2020.7.(4). 887-906 *(In Russian)*

Sheaff, R., Halliday, J., Byng, R., Øvretveit, J., Exworthy, M., Peckham, S. and Asthana, S. (2017). Bridging the discursive gap between lay and medical discourse in care coordination, *Sociology of health & illness*, 39(7), 1019–1034. https://doi.org/10.1111/1467-9566.12553 *(In English)*

Shuravina, L. S. (2013). Medical discourse as a type of institutional discourse, *Herald of Chelyabinsk State University*. *Philology*, *Art History,* 37 (328), 86, 65–67. *(In Russian)*

Silletti, A. M. (2016). The role of illustrations in popularizing medical discourse, *Linguæ &-Rivista di lingue e culture modern,* 14(2), 65–81. https://doi.org/10.7358/ling-2015-002-sill *(In English)*

Torshin, I. Yu., Gromova, O. A., Stakhovskaya, L. V., Vanchakova, N. P., Galustyan, A. N., Kobalava, Zh. D., Grishina, T. R., Gromov, A. N., Ilovaiskaya, I. A., Kodentsova, V. M., Kalacheva, A. G., Limanova, O. A., Maksimov, V. A., Malyavskaya, S. I., Mozgovaya, E. V., Tapilskaya, N. I., Rudakov, K. V. and Semenov V. A. (2020). Analysis of 19.9 million publications of the PubMed/MEDLINE database using artificial intelligence methods: approaches to the generalizations of accumulated data and to the phenomenon of "fake news", *FARMAKOEKONOMIKA, Farmakoekonomika. Modern Pharmacoeconomics and Pharmacoepidemiology*, 13 (2), 146–163. https://doi.org/10.17749/2070-4909/farmakoekonomika.2020.021 *(In Russian)*

Varnavskaya, E. V. (2019). Network Medical Discourse in blogosphere: Sociolect and peculiarity of communication online, *Scientific and methodological e-journal Concept*, available at: https://e-koncept.ru/en/tag/network+medical+discourse (Accessed 9 January 2022). *(In Russian)*

Waitzkin, H. (1989). A critical theory of medical discourse: ideology, social control, and the processing of social context in medical encounters, *Journal of Health and Social Behaviour*, 30(2), 220–239. https://doi.org/2137015 *(In English)*

Wilce, J. M. (2009). Medical discourse, *Annual Review of Anthropology*, 38(1), 199–215. https://doi.org/10.1146/annurev-anthro-091908-164450 *(In English)*

Yadav, Sh., Ekbal, A., Saha, S. and Bhattacharyya, P. (2018). Medical sentiment analysis using social media: Towards building a patient assisted system, *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018),* European Language Resources Association (ELRA), Miyazaki, Japan, 2791–2797. *(In English)*

Yus, F. (2018). Attaching Feelings and Emotions to Propositions. Some Insights on Irony and Internet Communication, *Russian Journal of Linguistics*, 22 (1), 94–107. https://doi.org/10.22363/2312- 9182-2018-22-1-94-107 *(In English)*

Zunic, A., Corcoran, P. and Spasic, I. (2020). Sentiment analysis in health and well-being: Systematic review, *JMIR Medical Informatics*, 8(1):e16023, January, available at: https://pubmed.ncbi.nlm.nih.gov/32012057/ (Accessed 20 February 2022). *(In English)*

Конфликты интересов: у авторов нет конфликта интересов для декларации.

Conflicts of Interest: the authors have no conflict of interest to declare.

Irina Germanovna Ovchinnikova, Doctor of Philology, Professor, *Sechenov First Moscow State Medical University.*
Liana Magdanovna Ermakova, PhD in Computer Science, Associate Professor, Université de Bretagne Occidentale, Brest, France.
Diana Magdanovna Nurbakova, PhD in Computer Science, Associate Professor, Department of Computer Science, Research Laboratory of Informatics in Image and Information Systems, Institut National des Sciences Appliqu´ees de Lyon, Lyon, France.