

Анализ поисковых запросов в «Яндексе», связанных с COVID-19 в Российской Федерации

Д. К. Хорошун, К. Т. Момыналиев, Е. М. Воронин, В. Г. Акимкин

ФБУН «Центральный НИИ эпидемиологии» Роспотребнадзора, Москва

РЕЗЮМЕ

Подходы, основанные на анализе данных поисковых запросов в интернете, могут быть важны для понимания реакции общественности и проведения эпидемиологического надзора за заболеваниями. Одним из таких инструментов может являться сервис «Яндекс.Подбор слов». В дополнение к почти всеобщему доступу населения к поисковым сервисам и возможности сбора данных в режиме реального времени многие пользователи ищут информацию в интернете прежде чем обращаться к врачу, что дает возможность лучше фиксировать начало заболеваний, процессы, связанные с ними и реакцию общества.

Целью нашего ретроспективного описательного исследования COVID-19 в России является использование сервиса «Яндекс.Подбор слов» для описания симптоматики заболевания и осложнений на основе поисковых запросов, а также их связи с общественным интересом в профилактических мерах, тестировании на COVID-19.

Методы. Мы использовали сервис «Яндекс.Подбор слов» – общедоступную онлайн-систему отслеживания поисковых запросов по неделям в поисковой системе «Яндекс». Запросы в «Яндексе» в России анализировали с 10.08.2020 по 28.11.2021. Первоначально мы составили список из 61 поискового запроса в следующих категориях: общие симптомы COVID-19, осложнения, тестирование, использование лекарственных средств, профилактические меры, медицинская помощь, аллергия.

Результаты. Поисковые термины, связанные с симптомами, тестированием и лекарствами, тесно коррелируют с регистрируемыми случаями COVID-19 в России, что указывает на необходимость дальнейших исследований потенциального использования сервиса «Яндекс» в качестве инструмента эпидемиологического надзора за заболеванием.

КЛЮЧЕВЫЕ СЛОВА: COVID-19, цифровая эпидемиология, коронавирус SARS-CoV-2, «Яндекс.Подбор слов», корреляция, запрос.

КОНФЛИКТ ИНТЕРЕСОВ. Авторы заявляют об отсутствии конфликта интересов.

Analysis of Yandex search queries related to COVID-19 in Russian Federation

D. K. Khoroshun, K. T. Momynaliev, E. M. Voronin, V. G. Akimkin

Central Research Institute of Epidemiology, Moscow, Russia

SUMMARY

Approaches based on the analysis of internet search query data can be important for understanding public reaction and conducting disease surveillance. One of these tools may be the Yandex.Wordstat service. In addition to near-universal public access to search services and the ability to collect real-time data, many users search information in the internet before visiting a doctor, which makes it possible to better capture the onset of diseases, the processes associated with them and the reaction of society.

The aim of our retrospective, descriptive study of COVID-19 in Russia is to use Yandex.Wordstat to describe the symptoms of the disease and complications based on search queries, as well as their relationship to the public interest in prevention measures, testing for COVID-19.

Methods. We used the Yandex.Wordstat service, a public online system for tracking search queries by week in the Yandex search engine. Requests to Yandex in Russia were analyzed from 08/10/2020 to 11/28/2021. We initially compiled a list of 61 search terms in the following categories: common symptoms of COVID-19, complications, testing, drug use, preventive measures, medical care, allergies.

Results. Search terms related to symptoms, testing, and drugs closely correlate with reported cases of COVID-19 in Russia, which indicates the need for further research on the potential use of the Yandex service as a disease surveillance tool.

KEY WORDS: COVID-19, digital epidemiology, SARS-CoV-2 coronavirus, Yandex.Wordstat, correlation, query.

CONFLICT OF INTEREST. The authors declare no conflict of interest.

Введение

Пандемия COVID-19 привела к значительной заболеваемости и смертности в России: по состоянию¹ на 05 марта 2022 выявлено 16861 793 случая COVID-19 и 355 537 человек умерло. Реальные, основанные на данных подходы к пониманию ответных мер общественности и проведению эпидемиологического надзора за заболеваемостью имеют первостепенное значение. Одним из таких инструментов является анализ запросов пользователей в сети «Интернет», например, сервис Google Trends успешно используется для описания эпидемий заболеваний, включая Эболу, грипп и ранние ответные меры на COVID-19 [1–5]. В дополнение

к почти всеобщему доступу к поисковым сервисам и сбору данных в режиме реального времени многие пользователи ищут в интернете информацию прежде, чем получить доступ к здравоохранению [6]. Таким образом, возможно лучше фиксировать бремя болезней, процессы болезней и реакцию общества с помощью инструментов анализа поисковых запросов.

Ранее нами был проведен анализ запросов пользователей «Яндекса», связанных с обонянием, в Брянской, Владимирской, Ивановской, Калужской, Московской, Рязанской, Смоленской, Тульской областях и Москве

¹<https://стопкоронавирус.рф>

с 16.03.2020 по 21.02.2021 по неделям. Была показана сильная временная взаимосвязь между самооценкой изменений запаха (запросы, включающие слово «обоняние», например «потеря обоняния») и новыми случаями заражения COVID-19 в областях Центрального региона России [7]. Выявлена корреляция между количеством запросов, связанных с обонянием, и новыми случаями заражения COVID-19: очень высокая для Москвы ($r = 0,96$), высокая для Брянской и Рязанской областей ($r > 0,70$), средняя для Владимирской, Ивановской, Калужской, Московской, Смоленской и Тульской областей ($0,5 < r < 0,7$). Наблюдаемые пики запросов не были связаны с сезонным обострением аллергии в России за предыдущие годы. Для проверки полученных результатов нами также были проанализированы онлайн-запросы российских пользователей «Яндекса» по поводу нарушения обоняния в регионах и крупных городах России с 16.03.2020 по 21.02.2021 [8]. Всего в исследование были включены 81 регион России и несколько городов, таких как Москва, Санкт-Петербург и Нижний Новгород. Была выявлена сильная положительная прямая корреляция ($r > 0,7$) между количеством связанных с запахом запросов в «Яндексе» новых случаев COVID-19 в 59 из 85 регионов России и крупных городов (70%). Во время первого пика заболеваемости COVID-19 в России (апрель-май 2020 года) рост числа запросов, связанных с запахом, опережал рост числа новых случаев на 1–2 недели в 23 из 59 регионов России. Во время второго пика заболеваемости COVID-19 в России (октябрь – декабрь 2020 года) рост числа запросов, связанных с запахом, опережал рост числа новых случаев на 1–2 недели в 36 регионах России, включая Москву.

Цель нашего ретроспективного исследования запросов в интернете, связанных с COVID-19 в России, состоит в том, чтобы оценить корреляцию между тенденциями поиска информации, связанной с симптомами COVID-19, в «Яндексе» и подтвержденным случаями COVID-19 в Российской Федерации за периоды 24.08.2020–27.12.2020 (35–52-я недели 2020 года), 31.05.2021–19.07.2021 (22–29-я недели 2021 года) и 06.09.2021–01.11.2021 (36–44-я недели 2021 года), когда было зафиксировано нарастание случаев заражения, а также общественным интересом к профилактическим мерам, тестированию на COVID-19 и лечению.

Методы исследования

Источник данных

Мы использовали «Яндекс.Подбор слов» – общедоступную онлайн-систему отслеживания поиска «Яндекса» по поисковым запросам, темам, географическому региону и датам. Вместо того чтобы представлять относительное количество поисковых запросов, «Яндекс.Подбор слов» отображает абсолютный объем поиска (АОП) для заранее заданных географических, временных и поисковых терминов. В нашей работе АОП определялся как количество запросов за неделю. Для расчета относительно объема поиска (ООП), количество запросов за определенную неделю нормировались на наибольший объем поиска в неделю за изучаемый период. Таким

Таблица 1
Категории запросов и специфические поисковые запросы, связанные с заболеваемостью COVID-19

Категория	Поисковый запрос
Симптомы	Лихорадка, кашель, головная боль, усталость, боль в груди, потеря обоняния, диарея, рвота, обоняние, сухой кашель, температура, понос, расстройство желудка, утомляемость, боль в горле, одышка, повышение температуры тела, конъюнктивит, чешутся глаза, чихание, резь в глазах, экзема, сыпь на коже, покраснение глаз, насморк, депигментация ногтей, отеки (27 запросов)
Патологический процесс, осложнения	Инфекция, пневмония, насыщение кислородом (3 запроса)
Тестирование	ПЦР, компьютерная томография, D-димер, С-реактивный белок, пульсоксиметрия (5 запросов)
Лекарственные средства	Амоксициллин, цефотаксим, цефтаролин, кларитромицин, моксифлоксацин, левофлоксацин, пиперацillin, тазобактам, цефепим, меропенем, имипенем, ципрофлоксацин, амикацин, линезолид, ванкомицин, азитромицин, ивермектин, дексаметазон (18 запросов)
Медицинское помощь	Кислородные баллоны, больница, врач (3 запроса)
Профилактические меры	Маска, дезинфицирующее средство, карантин (3 запроса)
Аллергия	Аллергия, пыльца (2 запроса)

образом, данные ООП представлялись в виде шкалы от 0 до 100, где 100 соответствует максимальному количеству запросов неделю за исследуемый период, а 0 показывает отсутствие запросов.

Стратегия поиска

В нашей работе проведен поиск запросов в «Яндексе» в России с 10 августа 2020 года (33-я неделя 2020 года) по 28 ноября 2021 года (48-я неделя 2021 года) по неделям с тремя пиками подъема заболеваемости: с 21.12.2020–27.12.2020 (первый пик), 12.07.2021–18.07.2021 (второй пик) и 01.11.2021–07.11.2021 (третий пик). В работе была проведена оценка запросов, связанных с симптомами COVID-19, с осложнениями заболевания, тестированием, использованием лекарственных препаратов, профилактическое поведение, медицинская помощь, аллергия (табл. 1). Всего был проанализирован 61 поисковый запрос.

Статистический анализ

Это ретроспективное исследование в первую очередь было описательным анализом данных «Яндекса.Подбор слов». В данной работе мы определили корреляцию Спирмена между официальными данными о случаях COVID-19 и поисковыми запросами, чтобы определить, были ли поиски сопоставимы с официальными данными.

Результаты

Клинически значимые запросы

Статистика запросов за исследуемый период представлена в таблице 2. Как видно из представленных данных, наибольшее количество запросов (более 135 млн с 10 августа 2020 по 28 ноября 2021 года включает ключевые слова «ма-

Таблица 2
Статистика запросов за исследуемый период (с 10.08.2020 по 28.11.2021)

№	Поисковый запрос	Количество запросов	Количество запросов / количество подтвержденных случаев заражения COVID-19	№	Поисковый запрос	Количество запросов	Количество запросов / количество подтвержденных случаев заражения COVID-19
1	Депигментация ногтей	1012	0,00	32	Дезинфицирующее средство	2571 241	0,3
2	Цефтаролин	16872	0,00	33	Экзема	3168 100	0,4
3	Тазобактам	25358	0,00	34	Конъюнктивит	3293 933	0,4
4	Пиперацillin	32007	0,00	35	Одышка	3496 838	0,4
5	Имипенем	125155	0,01	36	Усталость	3518 314	0,4
6	Пульсоксиметрия	139111	0,02	37	Ципрофлоксацин	3562 375	0,4
7	Резь в глазах	150172	0,02	38	Боль в горле	4419 885	0,5
8	Линезолид	217655	0,03	39	Диарея	60 67036	0,7
9	D-димер	263892	0,03	40	Лихорадка	6070 750	0,7
10	Меропенем	382017	0,04	41	Амоксициллин	6576 202	0,8
11	Сыпь на коже	383791	0,04	42	Сухой кашель	7509 837	0,9
12	Расстройство желудка	427543	0,05	43	Левифлоксацин	8025 899	0,9
13	Ванкомицин	443993	0,10	44	Головная боль	11163 765	1,3
14	Насыщение кислородом	466958	0,10	45	Обоняние	11416 943	1,3
15	Повышение температуры тела	601916	0,10	46	Рвота	11801 878	1,4
16	Цефепим	657377	0,10	47	Азитромицин	13295 572	1,5
17	Утомляемость	659605	0,10	48	Дексаметазон	14130 770	1,6
18	Моксифлоксацин	705547	0,10	49	Насморк	14200 977	1,6
19	Амикацин	723985	0,10	50	Понос	16368 447	1,9
20	Покраснение глаз	791523	0,10	51	Отеки	21777 033	2,5
21	Ивермектин	914474	0,10	52	Аллергия	25691 032	3,0
22	Кислородные баллоны	1582 406	0,20	53	Карантин	29332 462	3,4
23	Чихание	1658 914	0,20	54	Пневмония	29921 691	3,4
24	Потеря обоняния	1669 137	0,20	55	Инфекция	31135 328	3,6
25	Кларитромицин	1670 708	0,20	56	ПЦР	32518 825	3,7
26	Чешутся глаза	1679 495	0,20	57	Кашель	47587 611	5,5
27	Компьютерная томография	1682 702	0,20	58	Маска	135403 303	15,6
28	Цефотаксим	1695 049	0,20	59	Больница	139949 834	16,1
29	Боль в груди	1803 623	0,20	60	Температура	206171 237	23,7
30	Пыльца	1957 410	0,20	61	Врач	248036 845	28,6
31	C-реактивный белок	2414 812	0,30				

ска», «больница», «температура» и «врач», а наименьшее количество (менее 500 тыс.) – связано с лекарственными препаратами Цефтаролин, Тазобактам, Пиперацillin, Имипенем, Линезолид, Меропенем, Ванкомицин, некоторыми симптомами заболевания, например «расстройство желудка», а также с тестированием – «пульсоксиметрия» и «D-димер».

Если оценивать отношение количества запросов к количеству подтвержденных случаев COVID-19 (относительное количество запросов), то следует отметить, что поисковые запросы, связанные с симптомами заболевания («экзема», «конъюнктивит», «одышка», «усталость», «боль в горле», «диарея», «лихорадка», «головная боль», «обоняние», «рвота», «сухой кашель», «насморк») и лекарственными средствами («Азитромицин», «Ципрофлоксацин», «Амоксициллин», «Левифлоксацин», «Дексаметазон»), сравнимы ($\pm 50\%$) с количеством подтвержденных случаев COVID-19.

Поиск информации в «Яндексе», включающей ключевые слова «понос», «отеки», «аллергия», «карантин»,

«пневмония», «инфекция», «ПЦР» и «кашель», превышает количество подтвержденных случаев не менее чем в два раза.

Среди поисков лекарств большинство было связано с ключевыми словами «Дексаметазон», «Азитромицин», «Левифлоксацин», «Амоксициллин», «Ципрофлоксацин», «Цефотаксим», «Кларитромицин» и «Ивермектин».

Корреляция поисковых запросов, связанных с поиском симптомов COVID-19, с подтвержденными случаями заболевания в разные периоды

С 24.08.2020 по 27.12.2020 очень высокая корреляция ($r > 0,9$) с подтвержденными случаями COVID-19 наблюдалась для следующих симптомов заболевания – «покраснение глаз», «одышка», «повышение температуры тела», «температура», «усталость», «лихорадка», «утомляемость», «чихание», «отеки», «рвота» и «экзема», и высокая корреляция ($r > 0,7$) – для запросов «обоняние», «сухой кашель», «чешутся глаза», «кашель», «конъюнктивит» и «боль в груди» (рис. 1).

№	Запрос	Период			
		24.08.2020-27.12.2020	31.05.2021-19.07.2021	06.09.2021-01.11.2021	
1	Аллергия	0,8	0,2	1,0	
2	Азитромицин	0,3	1,0	0,6	
3	Амикацин	0,8	-0,1	0,7	
4	Амоксициллин	0,6	0,6	0,2	
5	Боль в горле	0,4	0,2	-0,8	
6	Боль в груди	0,8	0,3	1,0	
7	Больница	0,5	0,1	0,7	
8	Ванкомицин	0,8	-0,2	1,0	
9	Врач	0,5	0,1	0,2	
10	Головная боль	0,9	-0,8	0,9	
11	Дезинфицирующее средство	0,3	-0,7	0,8	
12	Дексаметазон	0,9	1,0	1,0	
13	Депигментация ногтей	0,2	0,4	-0,6	
14	Диарея	0,5	0,9	-0,6	
15	Ивермектин	0,8	0,9	0,9	
16	Имипенем	0,7	-0,2	0,9	
17	Инфекция	0,8	0,1	0,8	
18	Карантин	-0,6	-0,3	1,0	
19	Кашель	0,8	-0,8	0,5	
20	Кислородные баллоны	0,7	0,8	1,0	
21	Кларитромицин	0,7	-0,2	0,6	
22	Компьютерная томография	0,3	0,1	0,7	
23	Конъюнктивит	0,8	0,8	0,7	
24	Левофлоксацин	0,4	1,0	0,8	
25	Линезолид	0,8	0,6	0,8	
26	Лихорадка	1,0	-0,9	1,0	
27	Маска	0,6	-0,1	0,9	
28	Меропенем	0,6	0,9	0,9	
29	Моксифлоксацин	0,6	1,0	0,9	
30	Насморк	0,1	-0,9	-0,8	
31	Насыщение кислородом	0,8	0,5	0,9	
32	Обоняние	0,7	0,9	0,8	
33	Одышка	0,9	-0,8	1,0	
34	Отеки	1,0	0,2	0,9	
35	Пиперацillin	0,9	-0,7	0,8	
36	Пневмония	0,7	1,0	0,9	
37	Повышение температуры тела	0,9	-0,8	1,0	
38	Покраснение глаз	0,9	0,3	0,7	
39	Понос	-0,2	0,9	-0,7	
40	Потеря обоняния	0,6	0,8	0,8	
41	Пульсоксиметрия	0,7	-0,6	0,8	
42	ПЦР	0,6	0,8	0,9	
43	Пыльца	0,9	-1,0	-0,2	
44	Расстройство желудка	0,1	0,9	-0,7	
45	Рвота	0,9	0,8	0,7	
46	Резь в глазах	0,9	-0,7	0,8	
47	С-реактивный белок	1,0	1,0	1,0	
48	Сухой кашель	0,7	-0,9	0,5	
49	Сыпь на коже	0,9	-0,4	1,0	
50	Тазобактам	0,8	-0,2	0,5	
51	Температура	0,9	0,0	0,6	
52	Усталость	0,9	-1,0	0,6	
53	Утомляемость	1,0	-0,9	1,0	
54	Цефепим	0,7	0,9	0,9	
55	Цефотаксим	0,6	0,9	0,8	
56	Цефтаролин	0,7	-0,6	0,6	
57	Ципрофлоксацин	0,6	0,7	0,5	
58	Чешутся глаза	0,7	0,3	0,3	
59	Чихание	1,0	0,3	0,1	
60	Экзема	1,0	0,1	0,9	
61	D-димер	1,0	1,0	0,1	

Рисунок 1. Корреляция поисковых запросов с подтвержденными случаями COVID-19 в периоды 24.08.2020–27.12.2020, 31.05.2021–19.07.2021 и 06.09.2021–01.11.2021. Красным выделена отрицательная корреляция, синим – положительная корреляция. Жирным шрифтом выделена корреляция, для которой $p < 0,01$.

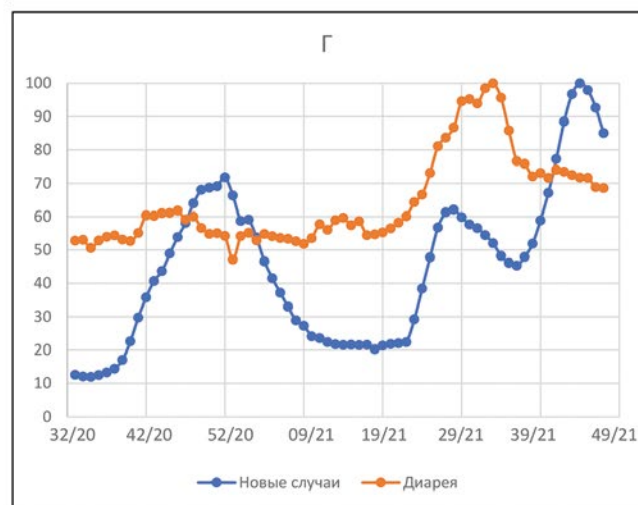
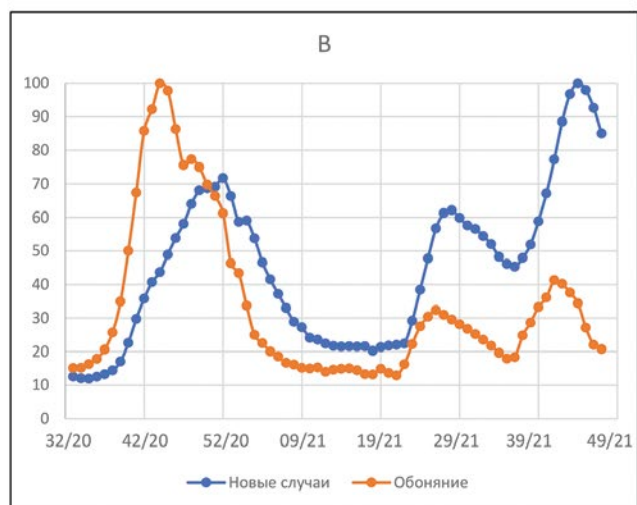
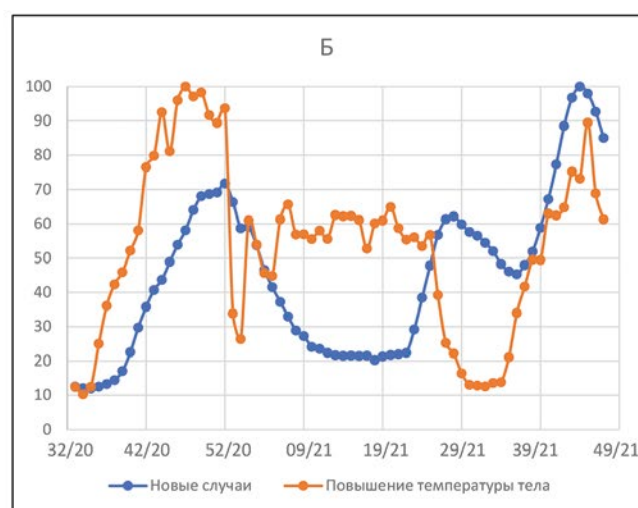
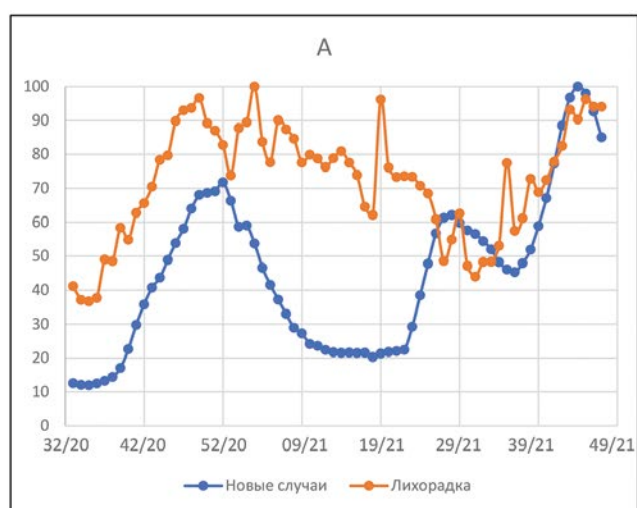


Рисунок 2. Корреляция поисковых запросов, связанных с симптомами: «лихорадка» (А), «повышение температуры тела» (Б), «обоняние» (В) и «диарея» (Г) с подтвержденными случаями COVID-19. Новые случаи – подтвержденное количество случаев COVID-19 в неделю. По оси ординат указано относительное количество подтвержденных случаев COVID-19 и запросов, связанных с поиском соответствующего симптома (данные за каждую неделю были нормализованы на максимальное количество случаев / запросов в неделю соответственно). По оси абсцисс указаны недели, где 32/20 соответствует неделе 03.08.2020–09.08.2020, 42/20 – неделе 12.10.2020–18.10.2020, 52/20 – неделе 21.12.2020–27.12.2020, 09/21 – неделе 01.03.2021–07.03.2021, 19/21 – неделе 10.05.2021–16.05.2021, 29/21 – неделе 19.07.2021–25.07.2021, 39/21 – неделе 27.09.2021–03.10.2021 и 49/21 – неделе 06.12.2021–12.12.2021.

С 31.05.2021 по 19.07.2021 очень высокая корреляция ($r > 0,9$) с подтвержденными случаями COVID-19 наблюдалась только для запросов «понос» и «обоняние», и высокая корреляция ($r > 0,7$) – для запросов «рвота», «потеря обоняния», «конъюнктивит». Для запросов «лихорадка», «сухой кашель», «одышка», «повышение температуры тела», «усталость», «утомляемость», наоборот, наблюдалась высокая отрицательная корреляция. Для запросов, связанных с «температура», «чихание», «экзема», «боль в горле» наблюдалась слабая корреляция с подтвержденными случаями COVID-19. Типичные кривые поисковых запросов, связанных с симптомами, приведены на *рисунке 2*.

С 06.09.2021 по 01.11.2021 запросы «экзема», «отеки», «повышение температуры тела», «утомляемость», «лихорадка», «боль в груди», «одышка» имели очень высокую корреляцию с подтвержденными случаями COVID-19 и высокую корреляцию с запросами «конъюнктивит», «покраснение глаз», «обоняние» и «потеря обоняния». Высокая отрицательная корреляция наблюдалась для запросов «насморк», «боль в горле» и «понос».

В то время как пик запросов для таких симптомов, как «усталость», «боль в груди», «потеря обоняния», «утомля-

емость», «боль в горле», «сыпь на коже», «депигментация ногтей», «одышка», «повышение температуры тела» был за 3–4 недели до первого пика заболеваемости COVID-19 в 21.12.2020–27.12.2020, другие симптомы, такие как «отеки», «экзема», «сыпь на коже» достигали пика за 2 недели до второго пика заболеваемости COVID-19 12.07.2021–18.07.2021.

Запросы, связанные с симптомами «рвота» и «расстройство желудка», достигали пика спустя неделю после второго пика заболеваемости, а «диарея» и «понос» – спустя 4 недели. Другие симптомы («кашель», «конъюнктивит», «чешутся глаза», «насморк» и «чихание») достигали пика запросов за 2–4 недели до третьего пика заболеваемости COVID-19 01.11.2021–07.11.2021.

Корреляция поисковых запросов, связанных с поиском лекарств, с подтвержденными случаями заболевания COVID-19 в разные периоды

С 24.08.2020 по 27.12.2020 высокая корреляция ($r > 0,7$) с подтвержденными случаями COVID-19 наблюдалась для запросов, связанных с лекарствами: «Цефепим», «Кларитромицин», «Ивермектин», «Ванкомицин», «Амикацин» и «Дексаметазон» (*рис. 1*).

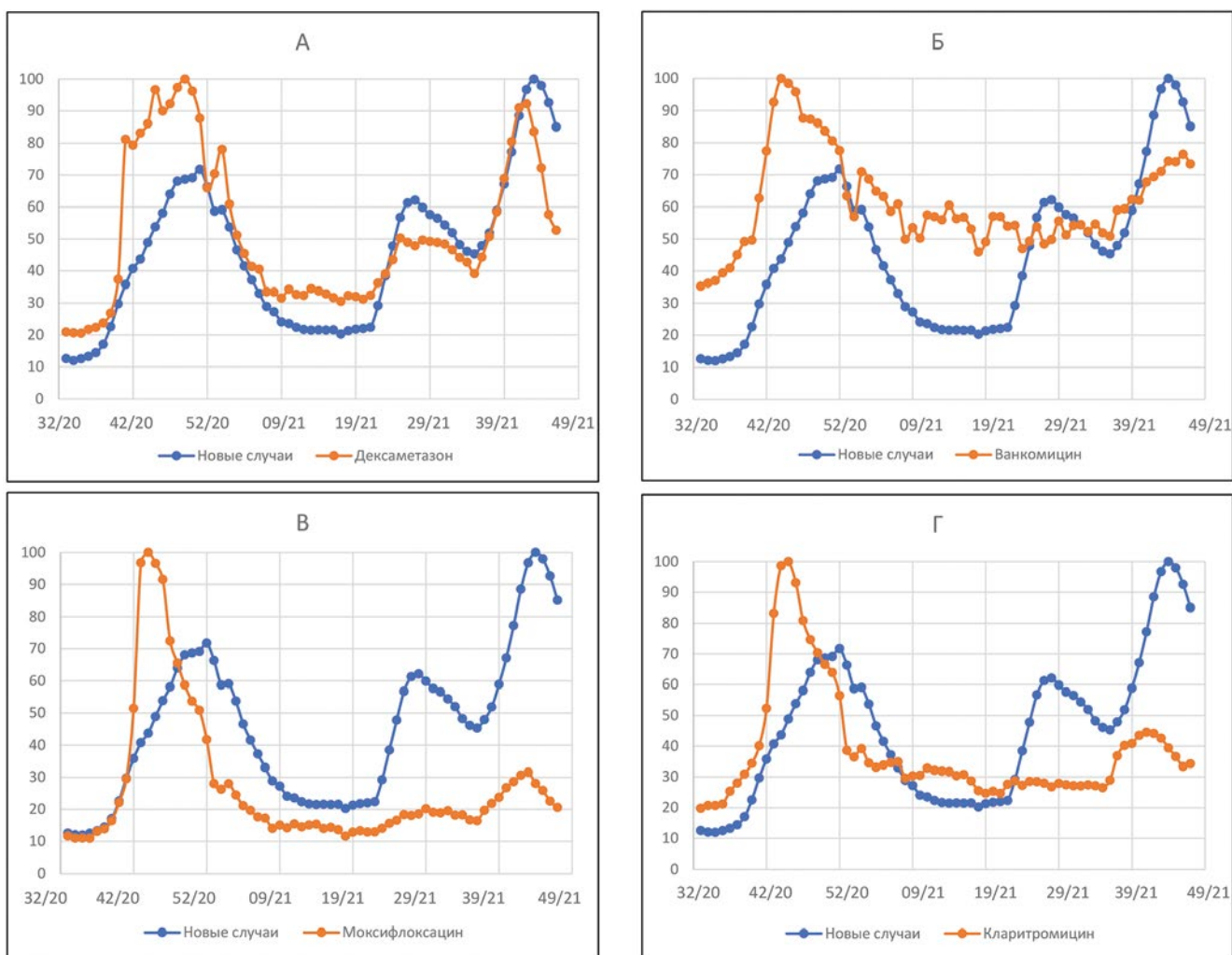


Рисунок 3. Корреляция поисковых запросов, связанных с лекарствами: «Дексаметазон» (А), «Ванкомицин» (Б), «Моксифлоксацин» (В) и «Кларитромицин» (Г) с подтвержденными случаями COVID-19. Новые случаи – подтвержденное количество случаев COVID-19 в неделю. По оси ординат указано относительное количество подтвержденных случаев COVID-19 и запросов, связанных с поиском соответствующего лекарства (данные за каждую неделю были нормализованы на максимальное количество случаев / запросов в неделю соответственно). По оси абсцисс указаны недели, где 32/20 соответствует 03.08.2020–09.08.2020 неделе, 42/20 – неделе 12.10.2020–18.10.2020, 52/20 – неделе 21.12.2020–27.12.2020, 09/21 – неделе 01.03.2021–07.03.2021, 19/21 – неделе 10.05.2021–16.05.2021, 29/21 – неделе 19.07.2021–25.07.2021, 39/21 – неделе 27.09.2021–03.10.2021 и 49/21 – неделе 06.12.2021–12.12.2021.

С 31.05.2021 по 19.07.2021 очень высокая корреляция ($r > 0,9$) с подтвержденными случаями COVID-19 наблюдалась для запросов, связанных с «Цефепим», «Цефотаксим», «Азитромицин», «Оксифлоксацин», «Левофлоксацин», «Дексаметазон», и высокая корреляция ($r > 0,7$) – для запросов «Ципрофлоксацин». Для запросов, связанных с «Пиперацillin» наблюдалась высокая отрицательная корреляция.

С 06.09.2021 по 01.11.2021 запросы, связанные с «Ивермектин», «Моксифлоксацин», «Цефепим», «Ванкомицин» и «Дексаметазон» имели очень высокую корреляцию с подтвержденными случаями COVID-19, «Амикацин» имел высокую корреляцию с подтвержденными случаями COVID-19.

Поисковые запросы, связанные со всеми антибиотиками, достигали пика за 4–8 недель до первого пика заболеваемости COVID-19 21.12.2020–27.12.2020 (рис. 3).

Корреляция поисковых запросов, связанных с поиском осложнений, тестированием, медицинской помощью и профилактическими мерами с подтвержденными случаями заболевания COVID-19 в разные периоды

С 24.08.2020 по 27.12.2020 высокая корреляция ($r > 0,7$) с подтвержденными случаями COVID-19 наблюдалась для

следующих запросов: «насыщение кислородом», «С-реактивный белок», «кислородные баллоны», «пневмония», «инфекция», «аллергия» и «пыльца» (рис. 1).

С 31.05.2021 по 19.07.2021 очень высокая корреляция ($r > 0,9$) с подтвержденными случаями COVID-19 наблюдалась для запросов, связанных с поиском «ПЦР», «С-реактивный белок», «кислородные баллоны» и «пневмония».

С 06.09.2021 по 01.11.2021 запросы, связанные с поиском «пневмония», «маска», «карантин», «аллергия» и «кислородные баллоны», имели очень высокую корреляцию с подтвержденными случаями COVID-19 и высокую корреляцию с запросами «дезинфицирующее средство», «инфекция» и «пневмония», «компьютерная томография», «ПЦР», «насыщение кислородом» и «С-реактивный белок» (рис. 3).

Поисковые запросы, связанные с тестированием и диагностикой («компьютерная томография», «насыщение кислородом», «пульсоксиметрия»), профилактическими мерами («дезинфицирующее средство», «карантин») и осложнениями COVID-19 («пневмония») достигали пика запросов за 5–10 недель до первого пика заболеваемости COVID-19 21.12.2020–27.12.2020.

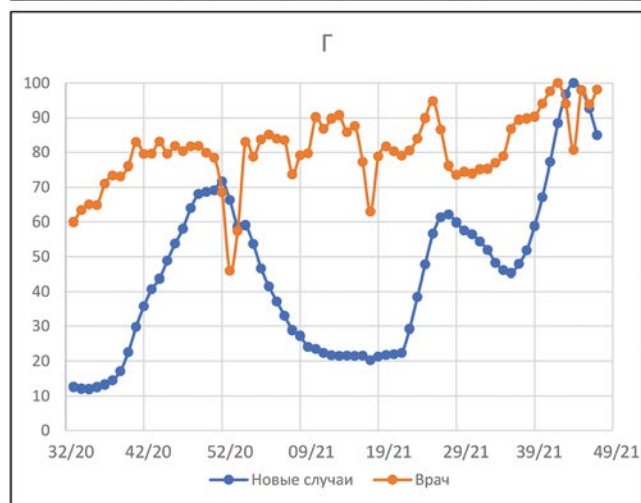
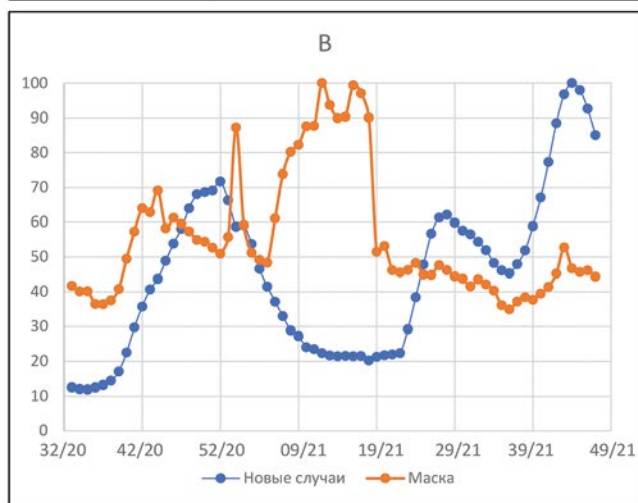
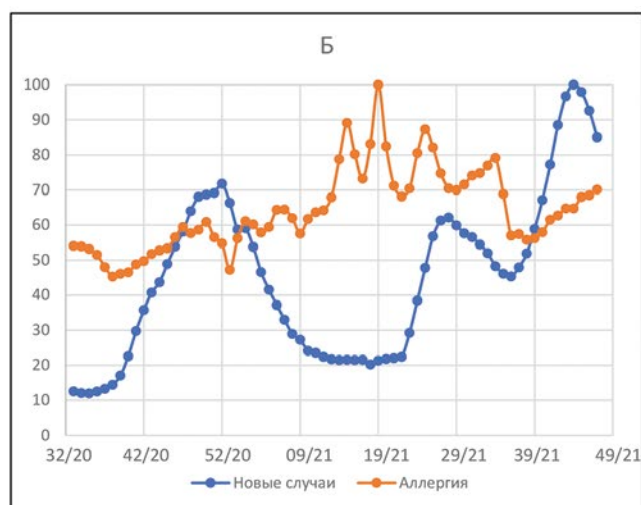
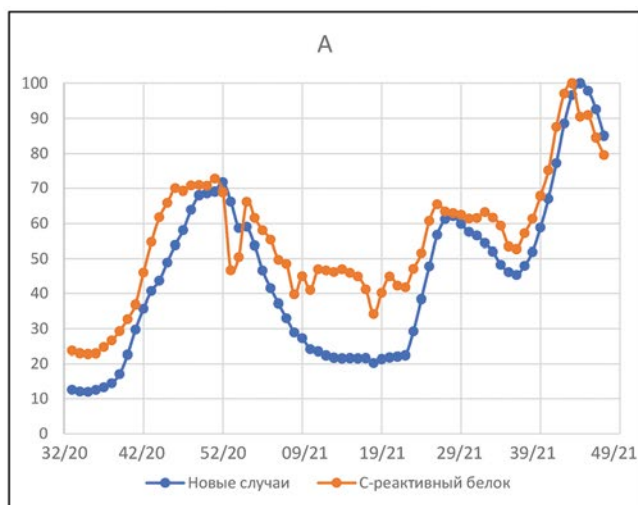


Рисунок 4. Корреляция поисковых запросов, связанных с «С-реактивный белок» (А), «аллергия» (Б), «маска» (В) и «врач» (Г) с подтвержденными случаями COVID-19. Новые случаи – подтвержденное количество случаев COVID-19 в неделю. По оси ординат указано относительное количество подтвержденных случаев COVID-19 и запросов, связанных с поиском соответствующего термина (данные за каждую неделю были нормализованы на максимальное количество случаев / запросов в неделю соответственно). По оси абсцисс указаны недели, где 32/20 соответствует неделе 03.08.2020–09.08.2020, 42/20 – неделе 12.10.2020–18.10.2020, 52/20 – неделе 21.12.2020–27.12.2020, 09/21 – неделе 01.03.2021–07.03.2021, 19/21 – неделе 10.05.2021–16.05.2021, 29/21 – неделе 19.07.2021–25.07.2021, 39/21 – неделе 27.09.2021–03.10.2021 и 49/21 – неделе 06.12.2021–12.12.2021.

Поисковые запросы, связанные с «ПЦР», достигали максимума за неделю до второго пика заболеваемости COVID-19 12.07.2021–18.07.2021.

Поисковые запросы, связанные с «С-реактивный белок», медицинской помощью («больница», «врач»), достигали максимума за неделю до третьего пика заболеваемости COVID-19 01.11.2021–07.11.2021.

Обсуждение

Анализ поисковых запросов, таких как Google Trends, используется для понимания пандемии COVID-19 во всем мире, включая его использование в качестве инструмента цифрового наблюдения [9–13], индикатора психосоциального поведения, симптомов, таких как агевзия и агнозия [14]. На наш взгляд, анализ поисковых запросов имеет множество преимуществ, включая анализ в реальном времени, большой размер выборки и поиск, который может быть упущен в исследованиях, посвященных только пациентам, обратившимся за медицинской помощью. Следует отметить, что, в отличие от Google Trends, сервис «Яндекс.Подбор слов» позволяет получать информацию не в относительном, а в абсолютном количестве – таким

образом, становится возможным сравнивать запросы с реальными случаями в абсолютных количествах, например заражение коронавирусной инфекцией или потребление лекарств и т. п.

В нашем исследовании поисковых запросов, связанных с тремя периодами подъема заболеваемости COVID-19 в России (с 24.08.2020 по 27.12.2020, с 31.05.2021 по 19.07.2021 и с 06.09.2021 по 01.11.2021), показано, что поиск в «Яндексе» выполняется по симптомам (например, «лихорадка», «кашель»), осложнениям (например, «пневмония»), лекарствам (например, «Дексаметазон»), методам тестирования (например, «ПЦР», «компьютерная томография»), медицинской помощью (например, «кислородные баллоны») и профилактическим поведением (например, «дезинфицирующее средство»). 34 запроса из 61 достигли пика еще за несколько недель до первого пика заболеваемости COVID-19 21.12.2020–27.12.202, 8 запросов – до третьего пика заболеваемости COVID-19 01.11.2021–07.11.202.

Только для запросов, которые относятся к симптомам COVID-19 («диарея», «понос», «расстройство желудка», «обоняние», «рвота», «конъюнктивит») и ме-

тодам тестирования («ПЦР», «С-реактивный белок» и «D-димер»), наблюдалась корреляция с подтвержденными случаями COVID-19 во второй период заболеваемости COVID-19 с 31.05.2021 по 19.07.2021, тогда как в первый (24.08.2020–27.12.2020) и третий периоды (06.09.2021–01.11.2021) наибольшую корреляцию с подтвержденными случаями COVID-19 имели запросы «одышка», «повышение температуры тела», «усталость», «лихорадка», «утомляемость», «кашель», «потеря обоняния», что позволяет предположить, что они могут быть наиболее распространенными симптомами COVID-19 [15].

Высокая корреляция запросов, связанных с желудочно-кишечными симптомами («диарея», «понос», «расстройство желудка», «рвота») с подтвержденными случаями COVID-19 во второй период 31.05.2021–19.07.2021 согласуются с данными современной литературы², описывающими симптомы при инфицировании дельта-штаммом SARS-CoV-2. Однако из представленных данных видно, что корреляция указанных симптомов с подтвержденными случаями COVID-19 снизилась в третий период (06.09.2021–01.11.2021) подъема заболеваемости.

Корреляция запросов «головная боль» показала недостоверную корреляцию с подтвержденными случаями COVID-19, что может свидетельствовать о том, что либо головная боль специфична для определенных пациентов при уже исходном высоком уровне запросов (всего 11,163 млн запросов) [16], либо модель поиска симптомов смещена в сторону от головной боли, которую легко не заметить.

Увеличение количества запросов «боль в груди», вероятно, является результатом плевритной боли в груди из-за пневмонии COVID-19, а не боли в груди из-за сердечно-сосудистых заболеваний. Однако Ciofani и соавт. показали среди населения США [17], что количество госпитализаций по поводу инфаркта миокарда сократилось, а количество запросов о боли в груди увеличилось в соответствии со случаями COVID-19. Это может указывать на то, что пациенты (пользователи) могут заниматься самодиагностикой или избегать обращения за помощью по поводу истинной сердечно-сосудистой боли в груди. Интересно, что поиски, включающие «усталость», «боль в груди», «обоняние», «повышение температуры», «утомляемость» достигли пиковых значений почти на 2 недели раньше, чем подтвержденное количество случаев заболевания, что позволяет предположить, что пациенты с COVID-19 могут искать симптомы до того, как пройти тестирование и быть включенными в официальное количество случаев.

Поскольку известно, что тяжелая инфекция COVID-19 вызывает провоспалительное состояние с повышенными воспалительными маркерами и склонностью к образованию тромбов, мы изучили поисковые запросы, связанные с лабораторными исследованиями и осложнением заболевания – пневмонией (данные не приведены). Поиск, включающий «компьютерная томография», «насыщение кислородом», «пульсоксиметрия», «С-реактивный белок», имели сильную корреляцию с поиском «пнев-

мония» ($r > 0,7$). Следует отметить, что поисковые запросы «D-димер», «С-реактивный белок» «насыщение кислородом» и «пневмония» имели сильную корреляцию с подтвержденными случаями COVID-19. Вероятно, потому что указанные тесты проводятся в координации при подозрении на пневмонию, что свидетельствует о том, что люди с тяжелыми осложнениями могут искать информацию в интернете до обращения за помощью.

Помимо клинически релевантных поисков, мы также изучали поисковые запросы, связанные с поиском лекарственных средств. Во всех исследованных периодах подъема поисковые запросы, включающие «Моксифлоксацин» (фторхинолон), «Цефотаксим» (цефалоспорин III поколения), «Цефепим» (цефалоспорин IV поколения), «Ивермектин» (антипаразитарное лекарственное средство), «Линезолид» (оксазолидиноны), «Дексаметазон» (глюкокортикостероид) имели высокую корреляцию с подтвержденными случаями COVID-19, причем поисковые запросы, связанные со всеми антибиотиками, достигали своего пика за 4–8 недель до первого пика заболеваемости COVID-19 (неделя 21.12.2020–27.12.2020), что может свидетельствовать о том, что люди, испытывающие симптомы заболевания, могли заниматься самолечением и в связи с этим искать лекарства. Однако затем активность в поиске антибиотиков спала, что может свидетельствовать уже о принятых мерах государства в части разъяснения рациональности использования антибиотиков и недопустимости самолечения COVID-19. Все лекарственные средства указаны во временных методических рекомендациях Минздрава «Профилактика, диагностика и лечение новой коронавирусной инфекции» при лечении осложнений инфекции, за исключением ивермектина, который массово рекламировался в мире для лечения и профилактики COVID-19. И хотя ВОЗ 31 марта 2021 года опубликовала сообщение³, что не рекомендуется использовать ивермектин для лечения COVID-19 вне клинических испытаний, пользователи в России по-прежнему активно интересовались ивермектином (914 тыс. запросов, пик запросов пришелся на 15.11.2021–21.11.2021) и, таким образом, не исключено, что по-прежнему пытались заниматься самолечением. Соответственно, вероятно, необходимо усилить контроль за распространением недостоверной информации в части лечения заболевания такими лекарствами.

В нашей работе мы также попытались оценить запросы, связанные с профилактическим поведением и медицинскими услугами. Поисковые запросы «карантин», «маска», «больница», «врач» были широко распространены и показывали высокую корреляцию в разные периоды подъема заболевания с официальным подсчетом случаев, но не на протяжении всего периода.

Проведенная работа также показала, что поиск, включающий запрос «аллергия», по крайней мере в два периода (с 24.08.2020 по 27.12.2020 и с 06.09.2021 по 01.11.2021), и «пыльца» в первый период (с 24.08.2020 по 27.12.2020), имел высокую корреляцию с подтвержденными случаями

²<https://www.cdc.gov/coronavirus/2019-ncov/variants/delta-variant.html>

³<https://www.who.int/ru/news-room/feature-stories/detail/who-advises-that-ivermectin-only-be-used-to-treat-covid-19-within-clinical-trials>

COVID-19, что может свидетельствовать о том, что многие пользователи принимали заболевание COVID-19 за аллергическую реакцию.

Следует отметить, что полученные нами результаты следует интерпретировать с осторожностью из-за многих ограничений. Во-первых, дизайн нашего исследования не позволяет нам делать выводы на уровне отдельных пациентов. Увеличение количества поисковых запросов может, по крайней мере частично, быть связано с увеличением присутствия связанных тем в СМИ, а не с отдельными ситуациями. Например, массовая реклама ивермектина способствовала подъему интереса к данному лекарственному средству больше, чем к другим лекарствам. Кроме того, поиск симптомов может быть связан не только с COVID-19. Однако, учитывая высокую распространенность COVID-19 в России и значительную корреляцию с подтвержденными случаями COVID-19, разумно предположить, что рост числа связанных поисковых запросов связан с COVID-19. Хотя мы пытались выбрать термины, которые охватили бы наибольший процент связанных терминов, это могло снизить специфичность поиска. Например, тема поиска «врач», а не «врач инфекционной больницы», или «врач скорой помощи», или «участковый врач» может охватывать все предполагаемые поисковые запросы с высокой чувствительностью, хотя при этом снижается специфичность поиска, поскольку также в статистику будут попадать врачи других специальностей, например «врач-гинеколог» или «врач-уролог». Несмотря на эти ограничения, большой размер выборки «Яндекс.Подбор слов», данные в реальном времени, непредвзятость в отношении групп населения, которые могут быть исключены из традиционных исследований, являются полезным инструментом для понимания COVID-19 [17]. Дальнейшие исследования могут также быть сосредоточены на понимании связи между поисковыми запросами и официальным подсчетом смертей в России, включая прогнозирование данных на основе моделей «Яндекс.Подбор слов».

Список литературы / References

1. Mollema L., Harmsen I. A., Broekhuizen E., Clijnk R., De Melker H., Paulussen T., Kok G., Ruiter R. & Das E. Disease detection or public opinion reflection? Content analysis of tweets, other social media, and online newspapers during the measles outbreak in The Netherlands in 2013. *Journal of medical Internet research*. 2015. 17 (5), e128. DOI: <https://doi.org/10.2196/jmir.3863>

2. Chen Y., Zhang Y., Xu Z., Wang X., Lu J., & Hu W. Avian Influenza A (H7N9) and related Internet search query data in China. *Scientific reports*. 2019. 9 (1), 10434. DOI: <https://doi.org/10.1038/s41598-019-46898-y>
3. Nurul Azmawati M., Hariz M. S., Mohd Dzulhairi M. R., Shalinawati R., Ilna I. Knowledge, attitude and practice on bats-borne diseases among village residents: A pilot study. *Med & Health*. 2018. 13 (2): 48–57. DOI: <https://doi.org/10.17576/MH.2018.1302.05>
4. Zeraatkar K., Ahmadi M. Trends of infodemiology studies: a scoping review. *Health Info Libr. J.* 2018. 35 (2): 91–120. DOI: <https://doi.org/10.1111/hir.12216>
5. Tang L., Bie B., Park S. E., Zhi D. Social media and outbreaks of emerging infectious diseases: A systematic review of literature. *Am. J. Infect. Control*. 2018. 46 (9): 962–72. DOI: <https://doi.org/10.1016/j.ajic.2018.02.010>
6. Shaughnessy AF, Slawson DC, Duggan AP. "Alexa, Can You Be My Family Medicine Doctor?" The Future of Family Medicine in the Coming Techno-World. *The Journal of the American Board of Family Medicine*. 2021. Mar 1; 34 (2): 430–4. DOI: <https://doi.org/10.3122/jabfm.2021.02.200194>
7. Момыналиев К. Т., Хорошун Д. К., Акимкин В. Г. Интернет-запросы, связанные с обонянием, как маркер оценки эпидемической ситуации в связи с коронавирусом SARS-CoV-2. Эпидемиология и инфекционные болезни. Актуальные вопросы. 2021; № 2: 42–48. DOI: <https://dx.doi.org/10.18565/epidem.2021.11.2.42-8>
8. Momynaliev K. T., Khoroshun D. K., Akimkin V. G. Study of olfaction-related Internet queries to evaluate the effectiveness of anti-epidemic measures against COVID-19. *Epidemiology and infectious diseases*. Current items. 2021; No. 2: 42–48. DOI: <https://dx.doi.org/10.18565/epidem.2021.11.2.42-8>
9. Momynaliev K. T., Khoroshun D. K., Akimkin V. G. Online queries as a criterion for evaluation of the epidemiological status and effectiveness of COVID-19 epidemic control measures. *medRxiv* 2021.07.11. 21260148; DOI: <https://doi.org/10.1101/2021.07.11.21260148>
10. Husnayin A., Fuad A., Su EC. Applications of Google Search Trends for risk communication in infectious disease management: A case study of the COVID-19 outbreak in Taiwan. *International Journal of Infectious Diseases*. 2020 Jun 1; 95: 221–3. DOI: <https://doi.org/10.1016/j.ijid.2020.03.021>
11. Ayyoubzadeh SM, Ayyoubzadeh SM, Zahedi H, Ahmadi M, Kalhori SR. Predicting COVID-19 incidence through analysis of google trends data in Iran: data mining and deep learning pilot study. *JMIR public health and surveillance*. 2020. 6 (2): e18828. DOI: 10.2196/18828.
12. Venkatesh U., Gandhi PA. Prediction of COVID-19 outbreaks using google trends in India: A retrospective analysis. *Healthcare informatics research*. 2020. 26 (3): 175–84. DOI: 10.4258/hir.2020.26.3.175.
13. Ciaffi J., Meliconi R., Landini MP, Ursini F. Google trends and COVID-19 in Italy: could we brace for impact? *Internal and Emergency Medicine*. 2020 May 25: 1–5. DOI: <https://doi.org/10.1007/s11739-020-02371-7>
14. Brodeur A., Clark AE, Fleche S, Powdthavee N. COVID-19, lockdowns and well-being: Evidence from Google Trends. *Journal of public economics*. 2021. 193: 104346. DOI: 10.1016/j.jpubeco.2020.104346.
15. Walker A., Hopkins C., Surda P. Use of Google Trends to investigate loss-of-smell-related searches during the COVID-19 outbreak. *International forum of allergy & rhinology* 2020. Vol. 10, No. 7, pp. 839–847. DOI: 10.1002/alr.22580.
16. Grant MC, Geoghegan L, Arbyn M, Mohammed Z, McGuinness L, Clarke EL, Wade RG. The prevalence of symptoms in 24,410 adults infected by the novel coronavirus (SARS-CoV-2; COVID-19): A systematic review and meta-analysis of 148 studies from 9 countries. *PLoS one*. 2020. 15 (6): e0234765. DOI: 10.1371/journal.pone.0234765.
17. Halpin S., O'Connor R., Sivan M. Long COVID and chronic COVID syndromes. *Journal of medical virology*. 2021. 93 (3): 1242–3. DOI: 10.1002/jmv.26587.
18. Visaria A., Palamarasetti P., Reddy S. и др. Characterization of the Second Wave of the COVID-19 Pandemic in India: A Google Trends Analysis. *medRxiv* 2021.05.19. 21257473; DOI: <https://doi.org/10.1101/2021.05.19.21257473>

Статья поступила / Received 14.03.22

Получена после рецензирования / Revised 02.04.22

Принята к публикации / Accepted 05.04.22

Сведения об авторах

Хорошун Димаш, специалист организационно-методического отдела. E-mail: dimash.mom@gmail.com. ORCID: 0000-0002-2445-5689

Момыналиев Куват Темиргалиевич, д.б.н., доцент, в.н.с. организационно-методического отдела. E-mail: dhoroshun@gmail.com. ORCID: 0000-0003-4656-1025

Воронин Евгений Михайлович, к.м.н., рук. научной группы математических методов и эпидемиологического прогнозирования. E-mail: emvoronin@yandex.ru. ORCID: 0000-0001-5925-7757

Акимкин Василий Геннадьевич, д.м.н., проф., академик РАН, директор. E-mail: vgakimkin@yandex.ru. ORCID: 0000-0003-4228-9044

ФБУН «Центральный НИИ эпидемиологии» Роспотребнадзора, Москва

Автор для переписки: Хорошун Димаш. E-mail: dimash.mom@gmail.com

About authors

Khoroshun Dimash, specialist of Organizational and Methodical Dept. E-mail: dimash.mom@gmail.com. ORCID: 0000-0002-2445-5689

Momynaliev Kuvat T., DBio Sci (habil.), associate professor, leading researcher at Organizational and Methodical Dept. E-mail: dhoroshun@gmail.com. ORCID: 0000-0003-4656-1025

Voronin Evgeny M., PhD Med, head of Scientific Group of Mathematical Methods and Epidemiological Forecasting. E-mail: emvoronin@yandex.ru. ORCID: 0000-0001-5925-7757

Akimkin Vasily G., DM Sci (habil.), professor, academician of RAS, director. E-mail: vgakimkin@yandex.ru. ORCID: 0000-0003-4228-9044

Central Research Institute of Epidemiology, Moscow, Russia

Corresponding author: Khoroshun Dimash. E-mail: dimash.mom@gmail.com

Для цитирования: Хорошун Д. К., Момыналиев К. Т., Воронин Е. М., Акимкин В. Г. Анализ поисковых запросов в «Яндексе», связанных с COVID-19 в Российской Федерации. Медицинский алфавит. 2022; (18): 14–22. <https://doi.org/10.33667/2078-5631-2022-18-14-22>.

For citation: Khoroshun D. K., Momynaliev K. T., Voronin E. M., Akimkin V. G. Analysis of Yandex search queries related to COVID-19 in Russian Federation. *Medical alphabet*. 2022; (18): 14–22. <https://doi.org/10.33667/2078-5631-2022-18-14-22>.

