
**СИСТЕМНЫЙ АНАЛИЗ ПРИБОРОВ
И ИЗМЕРИТЕЛЬНЫХ МЕТОДИК**

УДК 001.1+002.513.5+614.1

© А. Л. Буляница, А. А. Евстапов, 2021

**ОСНОВНЫЕ ТЕНДЕНЦИИ ПУБЛИКАЦИОННОЙ АКТИВНОСТИ
ПО ТЕМАТИКЕ, СВЯЗАННОЙ С ИНФЕКЦИОННЫМИ
ЗАБОЛЕВАНИЯМИ: COVID, SARS, EBOLA И ПТИЧИЙ ГРИПП**

Статья содержит наукометрический анализ публикаций, индексируемых в международной базе цитирования Scopus. Рассматривались группы публикаций, связанные по ключевым словам с COVID / COVID-19, вызываемым его вирусом SARS, а также публикации, посвященные пандемиям, вызванным вирусами Эбола (Ebola) и "птичий грипп" (различные варианты, объединенные поисковым запросом "Bird* flu"). Предполагалось выявить группы связей между развитием заболеваний и динамикой публикационной активности по соответствующей тематике, выполнить анализ выборок публикаций, формируемых путем последовательных (двух- и трехэтапных) поисковых запросов. Помимо динамики числа публикаций анализировались распределения по а) областям знаний (отраслям науки), б) по ассоциирующимся с указанной областью набором ключевых слов, в) по государственной принадлежности авторов и, применительно к авторам из России, значимости вклада организаций в выборку публикаций.

Из-за текущих количественных изменений числа публикаций за период подбора материала (середина марта – конец июня 2021 г.) некоторые количественные оценки могут незначительно измениться. При этом качественные выводы сохраняются.

В большинстве своем полученные выводы представляются ожидаемыми.

Кл. сл.: COVID (или COVID-19); SARS; Ebola; птичий грипп; пандемия; число публикаций; цитирования; отрасль знаний; Scopus

ВСТУПЛЕНИЕ

Цель статьи заключается в анализе основных тенденций публикационной активности по тематике, связанной с пандемиями, вызванными вирусами SARS, Ebola и птичий грипп. Отбор публикаций выполнен по библиометрической базе SCOPUS в период март–июнь 2021 г.

Основными направлениями являлись следующие:

1. Оценка распределения публикаций по областям знаний.
2. Определение основных методов и техник исследований, обнаружения и анализа указанных заболеваний.
3. Выявление ключевых (наиболее цитируемых) статей по разным направлениям исследований или/и областям знаний.
4. Анализ основных тенденций публикационной активности, в частности динамики числа публикаций и их связи со вспышками эпидемий.

Отбор осуществлялся по схемам поиска с различным выбором ключевых слов или/и их сочетаний. При этом не оценивалась степень соответствия содержания публикации и выбранного ключевого слова или/и комбинации слов.

Основными задачами были сравнительный анализ следующих распределений публикаций базы данных SCOPUS по тематике, ассоциированной со словами COVID, SARS, Ebola и "bird* flu" (символ * допускает разные варианты последующих букв или их отсутствие) по данным на период март–июнь 2021 г., сгруппированных:

- 1) по областям знаний;
- 2) по языкам публикаций;
- 3) по государственной принадлежности авторов;
- 4) при дополнительных поисковых ограничениях (метод анализа, сочетание двух и более заболеваний, дополнительные ассоциирующиеся слова и выражения, т.е. двух- или трехэтапный последовательный отбор / ограничение публикаций).

А также оценка динамики числа публикаций и "присутствия" авторов из РФ в публикациях рассматриваемой тематики.

**1. АНАЛИЗ ПУБЛИКАЦИЙ, СВЯЗАННЫХ
С КЛЮЧЕВЫМ СЛОВОМ COVID**

Поиск 14.03.2021 дал по ключевым словам covid — 114 202 публикаций, covid-19 — 113 041

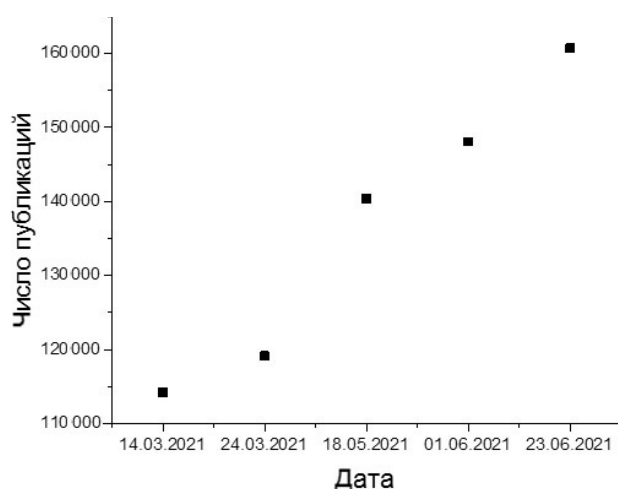


Рис. 1. Изменение числа публикаций базы данных Scopus при задании ключевого слова COVID

публикаций. С точки зрения выбора ключевых слов — практически одно и то же. На 24.03.2021 число публикаций выросло до 119 145, а на 18.05.2021 до 140 379, т.е. примерно на 25%. На 01.06.2021 отмечено 148 064 публикации, на 23.06.2021 число публикаций достигло примерно 160 700. На рис. 1 проиллюстрирован рост числа публикаций за краткосрочный период работы над статьей.

Рост числа публикаций близок к линейному в пределах рассматриваемого 3-месячного интервала. При этом увеличение числа публикаций за 5 дней составляет примерно 2%.

Лишь небольшая часть увеличения числа публикаций объясняется дополнительно найденными

публикациями 2020 г.: 84 874 на середину марта и 85 131 на середину мая 2021 г. В то же время число публикаций 2021 г. выросло от 29 273 до 55 185, или примерно на 88%. Публикаций 2018–2019 гг. около 50, что составляет менее 0.04% от общего числа публикаций по данной тематике. Основным видом публикаций были статьи — 68 095 (59.6%), письма — 14 919 (13.1%), обзоры (11.4%).

В табл. 1 показано распределение публикаций по наиболее значимым по вкладу областям знаний. Так как публикация может быть отнесена к нескольким областям одновременно, то сумма может превышать 100%. Далее в большинстве случаев ограничиваемся рассмотрением 5 областей знаний с наибольшим вкладом.

Принципиальных различий в распределениях публикаций 2020 и 2021 гг. не наблюдается. Следует отметить незначительное возрастание доли публикаций по иммунологии и микробиологии, а также в области техники и психологии за счет сокращения публикаций, отнесенных к медицине. Эта тенденция может отражать не столько изменение приоритетов исследования, сколько различие в среднем времени подготовки материала к публикации (например, техническое направление может требовать более продолжительного периода, чем медицинская публикация).

Отметим, что самая цитируемая на 24.03.2021 статья [1] (с датой публикации 15–21.02.2020) имеет 12 997 ссылок практически за год после публикации (по данным на 01.06.2021 число цитирований превысило 15 000), поскольку журнал The Lancet относится к высокорейтинговым изданиям, признаваемым международной научной общественностью.

Табл. 1. Распределение публикаций, отобранных по ключевому слову COVID, по областям знаний (в таблице — области с наибольшим относительным вкладом)

Область знаний	Доля, 2020–2021 гг., %	Доля, 2020 г., %	Доля, 2021 г., %
Медицина	66.5	67.8	62.5
Социальные науки	12.0	12.0	12.0
Биохимия, генетика, молекулярная биология	10.4	10.1	11.3
Иммунология и микробиология	6.4	6.2	7.0
Компьютерные науки	4.8	4.8	4.9
Фармакология, токсикология	4.0	4.1	3.8
Техника	4.0	3.8	4.5
Психология	3.5	3.2	4.3

Распределение по языкам публикаций: английский — 107 576 (94.2%), испанский — 2700 (2.4%), китайский — 1282 (1.1%), французский — 1211 (1.1%), русский 580 (0.5%).

В табл. 2 представлены данные по общему числу публикаций на 10.04.2021 в зависимости от дополнительных ключевых слов. В табл. 3 иллюстрируется долевое распределение публикаций по областям знаний для отдельных групп рассматриваемой выборки, а именно публикаций по COVID и другому поисковому признаку (ключевому выражению).

Наблюдаются существенные различия в распределениях по областям знаний для трех рассматриваемых подгрупп публикаций по COVID, ограниченных различными аналитическими методами.

Основные тенденции

1. Доля публикаций, отнесенных к медицине, для PCR-RT (ПЦР в реальном времени) домини-

рующая (более 80%) и вдвое превышающая аналогичные показатели для двух других направлений.

2. По-видимому, наблюдается "переход" части публикаций из области формальной медицины в направления биофизика, генетика и молекулярная биология (около 17% для PCR-RT и около 45–48% для других подгрупп).

3. Доля компьютерных наук в целом невелика, но для направления Isothermal. Ampl. она много больше (6% против 3 и 1.5%).

4. Область фармакологии составляет большую часть публикаций, связанных с масс-спектрометрией. По-видимому, это — задачи по выявлению лекарственных препаратов.

5. Собственно техническое направление значимо представлено только применительно к изотермической амплификации (18%), а в других направлениях порядка 2–3%.

Табл. 2. Показатели публикационной активности в зависимости от сочетания ключевого слова COVID и дополнительной комбинации ключевых слов

Комбинация ключевых слов	Публикации, кол-во	Цитирования / индекс Хирша	Ссылки на самую цитируемую статью
Covid & Nature journal	*	8987 / 42	1684 [2]
Covid & PCR-RT	143	3523 / 23	1073[3]
Covid & mass-spectrometry	285	2605 / 22	731 [4]
Covid & isothermal & amplification	234	3196 / 29	342 [5]
Covid & DNA	1221	21 903 / 55	3622 [6]
Covid & nanopore & sequence	24	376 / 5	281 [7]

*Число публикаций сильно менялось на протяжении рассматриваемого периода

Табл. 3. Распределение по формальным областям знаний для разных групп публикаций

Область знаний	Covid & PCR-RT, %	Covid & Mass-spectr., %	Covid & Isothermal Ampl., %
Медицина	80.9	36.6	43.0
Биохимия, генетика, молекулярная биология	17.6	48.2	45.2
Иммунология и микробиология	24.4	14.4	21.1
Компьютерные науки	1.5	3.1	6.1
Фармакология, токсикология...	2.3	18.3	5.3
Техника	3.1	1.9	18.0

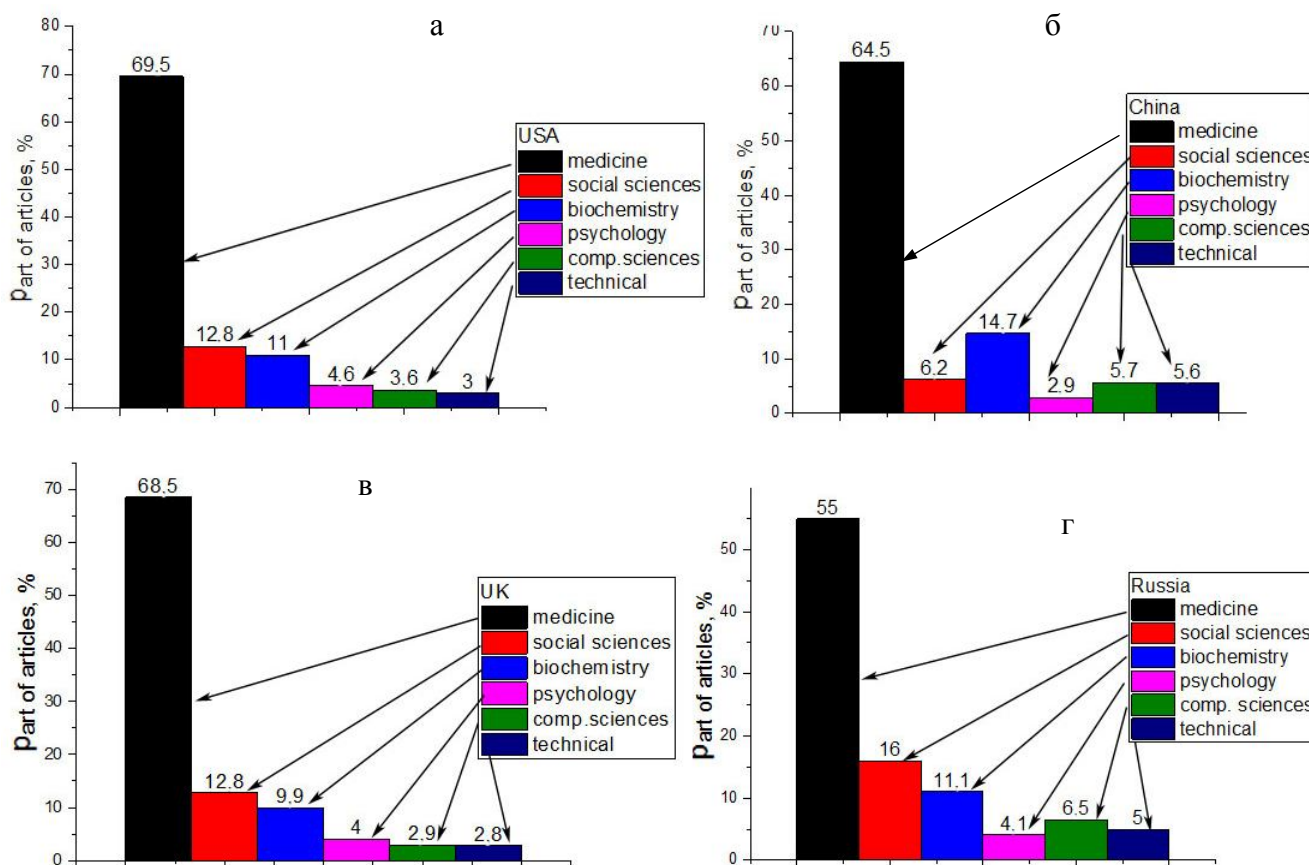


Рис. 2. Распределение публикаций по областям знаний для авторов различной государственной принадлежности

Различие распределений, возможно, связано не только с развитием техники исследований, но и с отличием в тематике журналов, в которых опубликованы результаты.

На рис. 2 (а–г) представлены распределения публикаций по областям знаний для авторов различной государственной принадлежности. Используются англоязычные названия областей знаний.

Общие тенденции

1. Доля публикаций в области медицины составляет в основном 65–70%, лишь для авторов из РФ она существенно ниже (55%).

2. Доля публикаций по социальным наукам в большинстве случаев составляет более 12%, лишь для авторов из Китая около 6%.

3. Вклад области биохимии составляет 10–15% во всех случаях, для области психологии аналогичная доля 3–5%.

4. Доля публикаций по компьютерным и техническим наукам для США и Великобритании около 3%, а для РФ и Китая примерно вдвое большая (5.5–6.5%)

Эти тенденции могут объясняться как приоритетами в научно-технической политике указанных государств, так и в менталитете или/и особенностях социальной политики.

2. СРАВНЕНИЕ С ПУБЛИКАЦИЯМИ ПО ПОИСКОВОМУ СЛОВУ SARS (БОЛЬШОЙ ПЕРИОД ВРЕМЕНИ)

Общее число публикаций — 150 764 на дату 14.03.2021. Динамика за последний период: 2018 — 6272, 2019 — 6462, 2020 — 37181, 2021 — 13 966. Из этих публикаций: статьи — 88 376 (58.6%), материалы конференций — 36 533 (24.2%), обзоры — 12 085 (8.0%). К 24.03.2021 число публикаций выросло до 154 311. Доля статей близка к публикациям по COVID, а доля материалов конференций значительно большая. Это может объясняться тем, что публикации по COVID относятся к периоду 2020–2021 гг., когда из-за соответствующих ограничений число конференций уменьшилось. В табл. 4 приведены данные публикационной активности по тематике, относящейся к SARS, на 10.04.2021.

Табл. 4. Распределение публикаций по комбинациям ключевых слов, содержащих SARS

Комбинация ключевых слов	Публикации	Цитирования / индекс Хирша	Самая цитируемая статья
SARS & PCR-RT	170	4278/28	1073 [3]
SARS & Mass-spectrometry	1269	28449/76	731 [4]
SARS & isothermal & amplification	270	4791/39	342 [5]
SARS & DNA	2470	Очень велико	3622 [6]
SARS & nanopore & sequence	31	382/5	281 [7]

Табл. 5. Распределение публикаций по областям знаний для группы, выбранной по ключевому слову SARS

Область знаний	Число публикаций	Доля, %*
Медицина	43 485	28.8
Земля и планетарные науки	29 233	19.4
Биохимия, генетика, молекулярная биология	24 700	16.4
Физика и астрономия	14 804	9.8
Компьютерные науки	32 724	21.7
Фармакология, токсикология...	16 408	10.9
Техника	34 851	23.1

*Суммарная доля может превышать 100%, поскольку некоторые публикации могут быть отнесены одновременно к нескольким областям

Самые цитируемые статьи по подгруппам, например PCR-RT в сочетании либо с SARS, либо с COVID, полностью совпадают. То есть наиболее цитируемые статьи связаны с исследованием совместно и SARS, и COVID. В табл. 5 приведено распределение публикаций, ассоциирующихся с SARS, по областям знаний. Указаны самые большие по вкладу области.

Очевидные тенденции

Для статей по SARS, публиковавшихся в течение более длительного периода, доля относящихся к медицине многократно меньше, чем для COVID (29% в сравнении с 65%), но доли относящихся к технике и компьютерным наукам — многократно больше (23 и 22% против соответственно 4 и 5%).

При этом в публикациях по COVID большую долю составляют социальные науки и психология (суммарно около 16–17%). Публикации по отраслям, формально относимым к генетике, биохимии,

молекулярной биологии, имеют устойчивое представительство в обеих группах публикаций (соответственно 11 и 16%). В табл. 6 представлено распределение публикаций разных лет по государственной принадлежности авторов.

Расхождение распределений публикаций по SARS и COVID для выбранной восьмерки стран можно оценить, используя информационное расстояние Кульбака – Лейтнера. Это расхождение примерно равно 0.155 нит. Такой величине соответствует, к примеру, расхождение двух нормально распределенных случайных величин с единичной дисперсией при сдвиге математических ожиданий примерно на 0.39. Такое расхождение нельзя считать значительным.

Отдельно рассмотрим динамику публикационной активности за период с 1950 г. по тематике SARS для авторов разной государственной принадлежности (аффилированные организации) на 24.03.2021. Эта информация приведена в табл. 7.

Табл. 6. Распределение авторов по государственной принадлежности (на 14.03.2021)

Страна	SARS		COVID	
	число	место	число	место
USA	34 094	1	28 764	1
China	30 379	2	11 958	3
Italy	11 923	3	9975	4
Germany	10 910	4	4558	8
UK	9599	5	12 079	2
India	7826	6	8243	5
France	6823	7	4320	10
Russia	2046	18	1372	22*

*Выше РФ в рейтинге стоят Канада, Австралия, Иран, Турция, Япония, Саудовская Аравия и ряд других стран, помимо упомянутых в таблице

Табл. 7. Динамика публикационной активности авторов по тематике SARS из США, Китая, Великобритании и России

Годы	Количество публикаций	Доля (%) / ранг по публикационной активности			
		США	Китай	Великобритания	РФ
1950–70	158	18.4 / 1	—	14.6 / 2	0.6 / 21
1971–80	881	31.6 / 1	—	6.0 / 2	0.6 / 18
1981–90	3612	26.9 / 1	0.6 / 14	5.5 / 2	0.2 / 27
1991–2000	11 099	34.2 / 1	3.0 / 8	7.4 / 4	1.9 / 11
2001–2005	14 258	23.8 / 1	15.8 / 2	5.6 / 6	1.1 / 18
2006–2010	20 319	22.2 / 2	25.3 / 1	5.4 / 5	1.1 / 21
2011–2015	25 221	19.5 / 2	28.3 / 1	4.6 / 6	1.1 / 24
2016–2017	11 294	17.2 / 2	31.1 / 1	4.7 / 6	1.7 / 14
2018	6272	16.5 / 2	32.1 / 1	5.2 / 6	1.9 / 14
2019	6459	14.8 / 2	36.4 / 1	4.5 / 6	2.1 / 11
2020	37 371	22.9 / 1	16.0 / 2	8.0 / 4	1.6 / 19
2021	17 319	26.8 / 1	11.4 / 2	9.8 / 3	0.9 / 31

Символ "—" означает отсутствие данных.

За рассматриваемый период публикаций, относящихся к SARS, ассоциированных с Российской Федерацией, насчитывается 2088 (по базе SCOPUS на 24.03.2021). Индекс Хирша этих публикаций — 55. Доля публикаций, по сравнению с США, Китаем и Великобританией, невелика — не превышает 2.1%.

Основные области знаний: медицина — 540, или 25.9%; Земля и планетарные науки — 494 (23.7%); компьютерные науки — 429 (20.5%); техника — 383 (18.3%); биохимия, генетика и молекулярная биология — 324 (15.5%).

Это распределение достаточно слабо отличается от общего распределения публикаций по SARS

безотносительно государственной принадлежности авторов.

Отметим рост публикационной активности по данной тематике, наблюдающийся с 1950 г. Основные исследования проводятся в США и Великобритании. На 2001–2005 гг. приходится всплеск публикационной активности. К исследованиям подключаются ученые из Китая. И с 2006 по 2019 г. основная доля публикаций имеет аффилированность с Китаем. В период пандемии соотношение по числу публикаций меняется в пользу США.

На временном интервале до 2019 г. включительно, когда отсутствовало влияние пандемии COVID, динамика среднегодового числа публика-

ций Y достаточно точно описывается степенным законом:

$$Y = 1.067 \cdot 10^{-3} \cdot (t - 1950)^{3.7},$$

где t — рассматриваемый год публикации. Коэффициент детерминации зависимости, приведенной к линейной в $\log - \log$ координатах, превышает 0.983, что свидетельствует об адекватности найденной зависимости.

Соответственно, компенсировав этот тренд, можно оценить воздействие ковидного фактора на публикационную активность. В этом случае некомпенсированными остаются примерно 30 000 публикаций 2020 г. и 14 800 публикаций 2021 г. (примерно 30% от годового объема). Соотношение числа публикаций по COVID за 2020 и 2021 гг. также близко к 2 : 1, а именно 85 тыс. к 55 тыс. (но 55 тыс. публикаций отмечено на 18.05.2021).

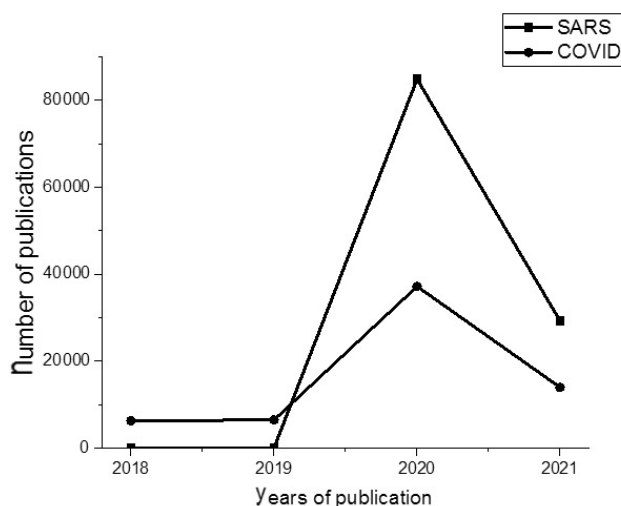
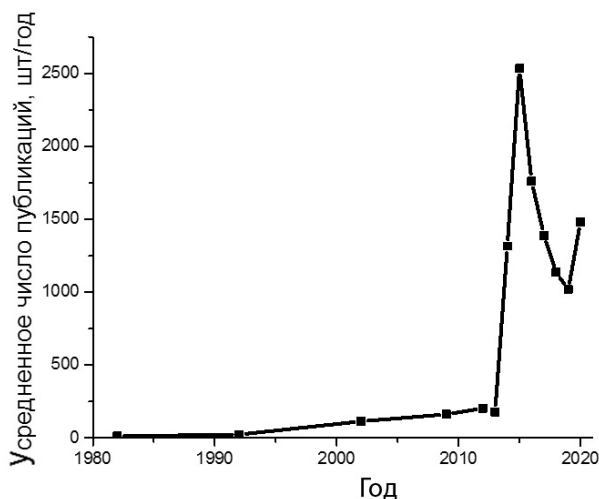


Рис. 3. Сопоставление числа публикаций в Scopus, связанных с ключевыми словами COVID и SARS



Следовательно, пандемия оказала сопоставимое влияние на публикационную активность как связанную с ключевым словом COVID, так и связанную с ключевым словом SARS.

За период 2014–2019 гг. доля публикаций, отнесенных к компьютерным наукам, составляла 30% и более. После начала пандемии COVID эта доля многократно уменьшилась за счет перераспределения публикаций в другие области: медицина, фармакология, биохимия, психология и ряд др.

Рис. 3 отражает сравнительную динамику публикационной активности по COVID и SARS. Наблюдается связь между лавинообразным ростом числа публикаций по COVID и соответствующим увеличением публикаций по SARS, как это было отмечено ранее.

3. ТЕНДЕНЦИИ ПУБЛИКАЦИОННОЙ АКТИВНОСТИ, АССОЦИИРОВАННОЙ С КЛЮЧЕВЫМ СЛОВОМ EBOLA

Общее число публикаций — 13 941 статей. Доля публикаций до начала 1980-х гг. составляет около 1%. Динамика публикационной активности проиллюстрирована рис. 4.

Первые 3 значения усредняются по 10 годам, значение периода 2007–2011 гг. по 5 годам, а далее представлено ежегодное число публикаций по тематике, ассоциирующей с ключевым словом Ebola. Наблюдается кратковременный всплеск публикационной активности в период 2016–2018 гг. До этого число годовых публикаций в период 1982–2013 гг. росло практически линейно, увеличиваясь на 6–7 статей в год. Распределение публикаций по тематике, связанной с ключевым словом Ebola, отражено в табл. 8.

Рис. 4. Динамика среднегодового числа публикаций, ассоциированных с Ebola

Табл. 8. Распределение по областям знаний (с самым большим относительным вкладом) публикаций, связанных с ключевым словом Ebola

Область знаний	Доля, %
Медицина	61.0
Социальные науки	7.4
Биохимия, генетика, молекулярная биология	15.9
Иммунология и микробиология	19.9
Фармакология, токсикология...	7.8

Распределение числа публикаций по государственной принадлежности авторов:

США — 6026 (43.2%),
Великобритания — 1688 (12.1%),
Франция — 866 (6.2%),
Германия — 846 (6.1%),
Канада — 840 (6.0%),
Китай — 708 (5.1%).

Россия находится на 23-м месте со 133 публикациями (чуть менее 1%). Для сравнения, представители Сьерра-Леоне находятся на 10-м месте с 427 публикациями, что достаточно легко объясняется. Это один из очагов выявления вируса Эбола. Наиболее цитируемая публикация [8] имеет 3439 цитирований.

Рассмотрим комбинацию ключевых слов, а именно поиск по двум ключевым словам Ebola & PCR. Выявлено 308 статей, наиболее цитируемая из них — [9] указана в 462 ссылках. Общее число цитирований 9650, т.е. в среднем 31.3 цитирования на статью. Индекс Хирша — 49. Распределение публикаций по областям знаний и по государственной принадлежности авторов мало отличается от общих данных по тематике Эбола: первые четыре места занимают США, Великобритания, Германия и Китай. Франция и Канада переместились на 6-е и 7-е место с долями публикаций 11 и 9% соответственно, а на 5-е место переместились авторы из Сьерра-Леоне (13%). Доля публикаций авторов из РФ также около 1%, но место в рейтинге стран ниже (34-е).

При выборе ключевых слов Ebola & mass-spectrum* выявлено 100 статей с 1914 цитированиями. Индекс Хирша — 25. Наиболее цитируемая статья — [10] с 229 цитированиями. Здесь примерно равные доли публикаций — от 35 до 39% приходятся на 3 области знаний: иммунология и микробиология; медицина; биохимия, генетика и др.

При поиске публикаций с использованием комбинации слов Ebola & DNA найдено 893 публика-

ции. Статья с наибольшим цитированием — [11] имеет 748 ссылок. Всего цитирований указанных публикаций по данным на 03.06.2021 отмечено 36 270. Индекс Хирша — 96.

Распределение по областям знаний аналогично предыдущему: те же 3 области с вкладом только от 35 до 41%. Из предыдущей ведущей семерки стран "выпала" Сьерра-Леоне, и Германия с Великобританией поменялись местами. Доля публикаций авторов из РФ достигла 2.1% (19 публикаций), что соответствует 11-му месту. Главным образом авторы из РФ представляют НПО "Вектор", Новосибирский университет и РАН.

В соответствии с поисковым запросом Ebola & nanopore & sequence выявлено всего 17 статей. Это нерепрезентативная выборка.

Выбор сочетания ebola & isothermal & ampl* выделяет всего 45 публикаций с общим числом цитирований 1048 (индекс Хирша — 18). Самая цитируемая статья со 157 ссылками — [12].

Сохранена ведущая роль авторов США (37.8%), но второе место заняли авторы из Китая (24.5%). За авторами из РФ числится 1 статья (места в рейтинге с 12-го по 23-е).

4. ТЕНДЕНЦИИ ПУБЛИКАЦИОННОЙ АКТИВНОСТИ, АССОЦИИРОВАННОЙ С КЛЮЧЕВЫМ ВЫРАЖЕНИЕМ "ПТИЧИЙ ГРИПП"

Поскольку англоязычные версии написания ключевого выражения допускают определенное разнообразие, было выбрано обобщающее выражение "bird* flu". Символ * допускает разные варианты последующих букв или их отсутствие.

Всего выявлено 1553 статьи по данной тематике. Статья [13] на 03.06.2021 имеет максимум цитирований 908 (на 08.06.2021 число цитирований — 912).

Табл. 9. Распределение публикаций по тематике "птичий грипп" по областям знаний

Область знаний	Доля, %
Медицина	49.8
Междисциплинарные науки	11.6
Биохимия, генетика, молекулярная биология	11.3
Иммунология и микробиология	9.7
Сельское хозяйство и биология	10.9

Распределение публикаций по областям знаний позволяет выявить самые большие по вкладу области. Эта информация приведена в табл. 9.

Распределение по государственной принадлежности авторов:

США — 287 (18.5%),
Китай — 122 (7.9%),
Великобритания — 88 (5.7%),
Индия — 69 (4.4%),
Франция — 49 (3.2%),
Таиланд — 48 (3.1%).

Далее — Япония, Канада и Германия. Авторы из РФ имеют 16 публикаций (традиционно около 1%) и 23-е место в рейтинге. Видно, что столь явного доминирования авторов США, которое наблюдалось практически во всех остальных выборах, здесь не выявлено.

Распределение публикационной активности по времени имеет колоколообразный характер с максимумом вблизи 2006 г. Из-за малых пульсаций ежегодного числа публикаций достоверность аппроксимирующих зависимостей не очень велика.

Наиболее адекватные аппроксимирующие зависимости — это гауссова кривая с ненулевой фоновой добавкой (примерно 55 ежегодных публикаций) и максимумом 273 или распределение Лоренца с фоном 47 публикаций и максимумом 517. Собственно вершина, соответствующая 2006 г., описывается плохо, но при аппроксимации участка ± 6 лет от максимума наблюдается приемлемое количественное совпадение. Коэффициенты детерминации для обеих зависимостей равны 0.80 и 0.84, что считается допустимым при признании адекватной аппроксимирующей зависимости.

По набору ключевых слов: bird's flu & DNA отобрано всего 65 статей с общим цитированием 1953, индекс Хирша — 19. Статья [14] имеет наибольшие в указанной группе 454 ссылки.

Здесь примерно 23-процентное участие авторов США и Китая, далее вдвое меньшее число публикаций из Великобритании. Авторы РФ и Кореи имеют по 3 публикации (рейтинг 6–7 с долей около 4.5%).

По SARS & bird's flu найдено 108 статей с общим числом цитирований 1600 и индексом Хирша 20. Максимум 218 цитирований имеет статья [15]. Число публикаций за последнее десятилетие совершает колебания в пределах от 1 до 8 статей в год. О резком интересе к тематике свидетельствует рост цитирований в 2020–2021 гг. Если за 2017–2019 гг. число цитирований — порядка 100, то в 2020 г. это число возросло до 280, а цитирования 2021 г. уже превысили 240.

По областям знаний первое место по-прежнему занимают публикации, отнесенные к медицине 46.3%, но на 2-м и 3-м местах в рейтинге областей находятся: социальные науки (20.4%) и бизнес, менеджмент и т.п. (примерно 9%). Распределение по государственной принадлежности авторов отличается от предыдущих: если первые два места рейтинга США и Китай достаточно традиционны, то далее идут представители стран Азии, затронутых ранее птичьим гриппом, а именно Гонконг и Сингапур. Авторы, аффилированных с РФ, нет, но примерно в 13% случаев авторы указаны без государственной принадлежности.

5. ПУБЛИКАЦИИ С КЛЮЧЕВЫМ ВЫРАЖЕНИЕМ SARS & EBOLA

Сочетание слов SARS & Ebola позволяет найти 1012 статей с общим числом цитирований 25 484 (на 03.06.2021). Индекс Хирша — 79. Максимум цитирований имеет статья [16] — 2139.

За период 2013–2019 гг. число статей колебалось в пределах от 20 до 50; в 2020 г. произошло практически десятикратное увеличение числа публикаций (до 388), а за первую половину 2021 г. уже 195 публикаций. Распределения публикаций по областям знаний и по государственной принадлежности авторов похожи на большинство предыдущих:

а) медицина — 5.8%, иммунология и микробиология — 21.7%, биохимия, генетика и т.п. — 20.8%,

фармакология и токсикология — 13.9%, социальные науки — 7.8%, остальное меньше;

б) авторы из США — 40.4%, Великобритания — 12.0%, Китая — 7.7%, Индии — 6.8%, Германии — 6.1% и т.д. Авторы из РФ имеют 4 публикации (рейтинг — 49).

Число цитирований публикаций этой тематики также дает выраженное примерно 4-кратное увеличение от примерно 1500–1700 цитирований в 2017–2019 гг. до 7200 за 2020 г. и уже 5500 за первую половину 2021 г.

6. АНАЛИЗ СОДЕРЖАНИЯ НАИБОЛЕЕ ЦИТИРУЕМЫХ ПУБЛИКАЦИЙ ПО ВЫДЕЛЕННЫМ РАЗДЕЛАМ (КОМБИНАЦИИ КЛЮЧЕВЫХ СЛОВ)

Представлен краткий разбор содержания наиболее цитируемых по соответствующей тематике статей.

[1] (Разд. 1): Отмечены случаи пневмонии в Ухане, Китай, которые были вызваны новым бета-коронавирусом, новым коронавирусом 2019 г. (2019-nCoV). Проанализированы эпидемиологические, клинические, лабораторные и рентгенологические характеристики пациентов. Методы сбора данных о пациентах: ПЦР в реальном времени и секвенирование следующего поколения. Проводилось сравнение данных пациентов из отделения интенсивной терапии и пациентов, не принадлежащих к этой группе. Наиболее интересная часть результатов приведенного статистического анализа: большинство инфицированных пациентов были мужчинами (73%); менее половины имели основные заболевания, включая диабет (20%), гипертонию (15%) и сердечно-сосудистые заболевания (15%). Средний возраст составил 49 лет (IQR 41.0–58.0). 2/3 посещали рынок морепродуктов Хуанань. Общими симптомами в начале заболевания были лихорадка, кашель и миалгия или усталость; менее распространенными симптомами были выделение мокроты, относительно редко встречались головная боль, кровохарканье и диарея. Одышка развилась у более половины пациентов в среднем на 8-й день. У всех пациентов была пневмония с аномальными результатами КТ грудной клетки. Осложнения включали острый респираторный дистресс-синдром, РНК-анемию, острую сердечную травму и вторичную инфекцию. В первой группе пациентов смертность составила около 15%. Выявлено сходство с коронавирусом тяжелого острого респираторного синдрома. Главные проблемы связаны с происхождением заболевания и в эпидемиологии, продолжительности передачи инфекции от человека и клиническом спектре заболеваний.

[2] (табл. 2): COVID-19 — это острая инфекция дыхательных путей, возникшая в конце 2019 г. Первоначально около 14% случаев с тяжелыми течениями и 6% случаев с критическими течениями.

Основной вопрос — отсутствие информации о репликации вируса, иммунитете и инфекционности дыхательных путей. Выделение вируса из глотки показало его высокое содержание в течение первой недели симптомов с пиком по числу копий РНК на мазок из горла на 4-й день. Вирус легко выделяется из образцов пробы, полученных из горла или легких, но не из образцов кала. Анализ крови и мочи также не выявили вируса. Возможно обнаружение различных по последовательности популяций вирусов в образцах пробы из горла и легких одного пациента, что доказывает независимую репликацию. Оценен срок выработки антител: за 7 дней у 50% пациентов и к 14-му дню у всех пациентов. Кроме острой формы COVID-19 может проявляться как легкое заболевание верхних дыхательных путей.

[3] (Разд. 2, табл. 4): COVID-19 вызывает серьезные вспышки в общинах и внутрибольничных условиях. Полные данные о респираторной вирусной нагрузке и реакции с образованием антител у пациентов, инфицированных SARS-CoV-2, отсутствуют. Описаны результаты исследования пациентов в двух больницах Гонконга (анализ слюны и сыворотки крови). Серийную вирусную нагрузку определяли методом количественной ПЦР с обратной транскрипцией. Секвенирование всего генома было проведено для выявления возможных мутаций, возникающих во время инфекции. Период исследования составлял 22.01–12.02 2020 г. Выводы: а) необходим строгий инфекционный контроль, б) слюна — хороший неинвазивный образец для анализа как для пациентов, так и для медицинских работников.

В отличие от тяжелого острого респираторного синдрома, пациенты с COVID-19 имели самую высокую вирусную нагрузку вблизи очага, что могло объяснить быстро распространяющийся характер этой эпидемии. Этот вывод подчеркивает важность строгого инфекционного контроля и раннего использования мощных противовирусных средств отдельно или в комбинации, для лиц с высоким риском. Серологический анализ может дополнить RT-qPCR для диагностики.

Для разного вида иммуноглобулинов указаны вероятности формирования антител (во всех случаях более 88% на 14-й день).

[4] (табл. 2, 4): Коронавирус 2 (SARS-CoV-2), который является возбудителем коронавирусной болезни COVID-19, заразил более 2.3 млн чел., привел к смерти более 160 тыс. чел. (Epub 2020, Apr 30). В указанный момент еще не существовало противовирусных препаратов с доказанной клини-

ческой эффективностью для лечения COVID-19, а также вакцин, предотвращающих заражение SARS-CoV-2. Показаны результаты клонирования, маркировки и экспрессии 26 из 29 белков SARS-CoV-2 в клетках человека, а также идентификации человеческих белков, связанных с каждым из белков SARS-CoV-2. Для этого использована масс-спектрометрия аффинной очистки с выявлением 332 высоконадежных белково-белковых взаимодействий между SARS-CoV-2 и белками человека, среди них 66 лекарственных белков человека. Определено два набора фармакологических агентов, которые проявляли противовирусную активность: ингибиторы трансляции мРНК и предсказанные регуляторы рецепторов сигма-1 и сигма-2.

[5] (табл. 2, 4): В статье содержится информация о разработке быстрого (до 40 мин), простого в реализации и точного метода анализа на основе CRISPR-Cas12, обеспечивающего обнаружение SARS-CoV-2 из экстрактов РНК дыхательных мазков. Проверка выполнена по 36 пациентам с COVID-19 и 42 пациентам с другими вирусными респираторными инфекциями. Предложенный метод анализа на основе CRISPR обеспечивает визуальную и более быструю альтернативу анализу ПЦР в реальном времени. Диагностические возможности метода по контролю заболеваний SARS-CoV-2 — 95-процентное положительное и 100-процентное отрицательное прогностическое совпадение.

[6] (табл. 2, 4): По состоянию на 26.01.2020 подтверждено более 2000 случаев заражения (2019-nCoV), большинство из которых касались людей, живущих в Ухане или посещающих его. Была подтверждена передача инфекции от человека к человеку. Проведено секвенирование образцов от девяти стационарных пациентов, восемь из которых посетили рынок морепродуктов Хуанань в Ухане. Моделирование гомологии было проведено для изучения вероятных рецептор-связывающих свойств вируса. В результате: десять последовательностей генома 2019-nCoV, полученных от девяти пациентов, были чрезвычайно похожи, демонстрируя идентичность последовательности более 99.98%.

Этот вирус тесно связан (с идентичностью 88%) с двумя коронавирусами, вызывающими тяжелый острый респираторный синдром (SARS), полученными от летучих мышей. 2019-nCoV имеет структуру домена, связывающего рецепторы, аналогичную структуре SARS-CoV, несмотря на изменение аминокислот в некоторых ключевых остатках. Интерпретация: 2019-nCoV достаточно отличается от SARS-CoV, чтобы считаться новым бетакоронавирусом, инфицирующим человека. Выдвинуто предположение, что летучие мыши могут быть первоначальным хозяином этого виру-

са; животное, продаваемое на рынке морепродуктов в Ухане, может представлять собой промежуточного хозяина, способствующего появлению вируса у людей. Также заявлена потребность срочного расследования эволюции, адаптации и распространения этого вируса.

[7] (табл. 2, 4): Констатация — SARS-CoV-2 это бетакоронавирус, ответственный за пандемию COVID-19 с неизвестной транскриптомной архитектурой. Два взаимодополняющих метода секвенирования позволили получить карту транскриптома и эпитранскриптома SARS-CoV-2 с высоким разрешением. В дополнение к каноническим геномным и 9 субгеномным РНК SARS-CoV-2 производит транскрипты, кодирующие неизвестные фрагменты со слиянием, делецией и/или сдвигом фрагментов. Технология нанопористого прямого секвенирования РНК позволила найти по меньшей мере 41 сайт модификации РНК на вирусных транскриптах с наиболее частым мотивом AAGAA и некоторые другие проявления модификации.

[8] (Разд. 3): Произведен анализ возникающих инфекционных заболеваний и делается вывод об их обусловленности социально-экономическими и экологическими факторами. Проанализирована база данных из 335 эпидемий в период с 1940 по 2004 г. и продемонстрированы неслучайные глобальные закономерности некоторых из них: а) большинство эпидемий 71.8% происходят из дикой природы (например, тяжелый острый респираторный вирус, вирус Эбола) и значительно усиливаются с течением времени; б) более половины случаев вызваны бактериями или риккетсиями, что отражается в устойчивости микробов к лекарствам. Отмечен глобальный географический эффект: существенный риск эпидемий, происходящих в более низких широтах.

[9] (Разд. 3): Вирусная геморрагическая лихорадка — это острые инфекции с высокой летальностью. Одним из возможных источников являются вирусы Эбола и Марбург, вирус Ласса и ряд др. Их клинически трудно диагностировать и различать; в подозрительных случаях требуется быстрая и надежная лабораторная диагностика. Установлены шесть одноэтапных ПЦР-анализов обратной транскрипции в реальном времени для этих патогенов на основе смеси ферментов обратной транскриптазы и Taq-полимеразы. Разработаны новые праймеры и/или зонды. Продукты ПЦР были обнаружены в режиме реального времени на приборе LightCycler с использованием технологии 5'-нуклеазы или интеркаляции красителя SybrGreen. Для этого красителя ослаблено ингибирующее действие и установлены универсальные условия циклирования для его обнаружения. Пределы обнаружения, с вероятностью не менее 95%, определены методом регрессионного анализа

и меняются от 1545 до 2835 эквивалентов вирусного генома/мл сыворотки (от 8.6 до 16 копий РНК на анализ).

[10] (Разд. 3): Описано средство ремдесивир, аналог аденозиннуклеотидного пролекарства, который проявляет противовирусную активность широкого спектра действия в отношении нескольких РНК-вирусов. Это соединение находится в стадии клинической разработки для лечения болезни, вызванной вирусом Эбола. Проведен анализ кинетики ферментов и дана оценка клинической селективности и эффективности в сравнении с АТФ.

[11] (Разд. 3): Отмечена крупнейшая вспышка болезни, вызванная вирусом Эбола, распространившаяся по Гвинее, Либерии, Сьерра-Леоне и Нигерии. При секвенировании 99 геномов вируса Эбола от 78 пациентов было выявлено быстрое накопление генетических вариаций. Рассмотренный западноафриканский вариант отличался от центральноафриканских линий (2004 г.), перешел из Гвинеи в Сьерра-Леоне в мае 2014 г. и впоследствии демонстрировал устойчивую передачу от человека к человеку.

[12] (Разд. 3): В обзоре также имеется указание на происхождение нового коронавируса — город Ухань, Китай. Первоначальные данные геномного секвенирования этого вируса не совпадают с ранее секвенированными CoV, что указывает на новый штамм CoV (2019-nCoV), который в настоящее время называется тяжелым острым респираторным синдромом CoV-2 (SARS-CoV-2). Предполагается, что COVID-19 происходит от животного-хозяина с последующей передачей от человека к человеку, но не следует исключать возможность других путей, таких как передача через пищу. По сравнению с заболеваниями, вызванными ранее известными человеческими CoV, а также Эбола и птичий грипп, COVID-19 демонстрирует менее тяжелый патогенез, но более высокую способность к передаче. Предполагается, что в качестве основного источника вируса выступают летучие мыши. Диагностика с помощью ПЦР в реальном времени и секвенирования следующего поколения облегчила идентификацию патогена на ранней стадии. В обзоре рассматриваются эпидемиологические, диагностические, клинические и терапевтические аспекты, включая перспективы применения вакцин и других профилактических мер, которые были рекомендованы во всем мире.

[13] (Разд. 4): Авторы использовали библиотеку фаговых дисплеев антител человека и эктодомен геммагглютинаина H5 для отбора десяти нейтрализующих антител, которые были эффективны против всех протестированных вирусов гриппа 1-й

группы, включая H5N1 "птичий грипп" и H1N1 "испанский грипп". Кристаллическая структура одного такого антитела, связанного с H5, показывает, что он блокирует инфекцию. Девять из рассмотренных антител используют ген зародышевой линии VH1-69, и все они, по-видимому, используют один и тот же механизм нейтрализации. Данные свидетельствуют о том, что иммунотерапия на основе антитела является перспективной стратегией защиты от сезонных и пандемических вирусов гриппа.

[14] (Разд. 4): Оценено изменение патогенности вирусов птичьего гриппа H5N1 здоровых домашних уток в материковом Китае с 1999 по 2002 г. и отмечено ее повышение. Представлена гипотеза, объясняющая механизм этого эволюционного направления. Исследован двадцать один вирус, подтвержденный как вирус гриппа А подтипа H5N1. Все вирусы высокопатогенны для птенцов. Частично их удалось реплицировать.

[15] (Разд. 4): Ссылка относится в большей степени к обзорным аналитическим бизнес-публикациям и в меньшей степени содержит научно обоснованные решения и выводы. Исследование направлено на: 1) выявление рисков во время турпоездов за границу; 2) изучение того, влияют ли такие предполагаемые риски на решения туристов во время кризисов; 3) исследование того, смягчаются ли такие риски каким-либо образом предыдущим опытом путешественника в зарубежной стране, и 4) определение влияния терроризма и воздействия заболевания, например птичьего гриппа, на индустрию гостеприимства Таиланда. Опасения влияют не на отказ от туризма в целом, а в большей степени на смену маршрута.

[16] (Разд. 5): Статья представляет собой обзор данных о развитии терапии, связанной с РНК-интерференцией, за десятилетний период, прошедший с момента ее открытия. В аннотации есть указание на присуждение за это достижение Нобелевской премии в области медицины. Однако имена лауреатов — Эндрю Файр (Andrew Z. Fire) и Крэйг Меллоу (Craig C. Mello) — не упомянуты.

7. СПЕЦИФИКА ПУБЛИКАЦИЙ ПО РАССМАТРИВАЕМОЙ ТЕМАТИКЕ В ЖУРНАЛЕ NATURE

Журнал Nature выбран, исходя из его популярности, высокого рейтинга, а также наличия в нем публикаций, относящихся к различным областям знаний, авторов, имеющих разную государственную принадлежность. Импакт-фактор журнала достаточно высок (он примерно равен 28). Дополнительные ключевые слова: COVID, AIDS, Flu

и аналогичные им, связанные с распространенными заболеваниями.

Набор комбинации ключевых слов: Journal Nature & Covid на 15.06.2021. Выявлено 437 публикаций (237 относится к 2020 г., 200 к 2021 г.). Все они формально отнесены к междисциплинарной области знаний (multidisciplinary). Среднее цитирование статей по тематике COVID на 15.03.2021 было более 110, что в 4 раза превышает среднее цитирование статей в Nature.

Распределение по государственной принадлежности авторов опять показывает уверенное преобладание США (72, или 16.5%), далее Великобритания (28), Германия (24), Китай (19), Швейцария (13). С авторами из РФ аффилированы 3 публикации (место — 22).

Подвыборки публикаций статей Nature с дополнительными поисковыми словами, помимо COVID, а именно — flu, aids и oncology, — дает соответственно 1, 2 и 0 публикаций.

В то же время публикации в Nature, ассоциированные с некоторыми другими заболеваниями, присутствуют.

Так, с ключевым словом "oncology" выявлено в Nature 120 публикаций. Первая статья датируется 1961 г., далее в течение 20 лет имеются одиночные публикации с интервалами 4–5 лет.

За 10-летие 2001–11 гг. число ежегодных публикаций менялось в пределах от 1 до 14 (некоторое повышение над общим фоном в 2007–8 гг. соответствовало 11 и 14 статьям).

За период 2012–21 гг. число ежегодных публикаций колеблется от 1 до 6. Общее число цитирований — 65 062. Индекс Хирша — 70, статья с максимальным цитированием 7457 относится

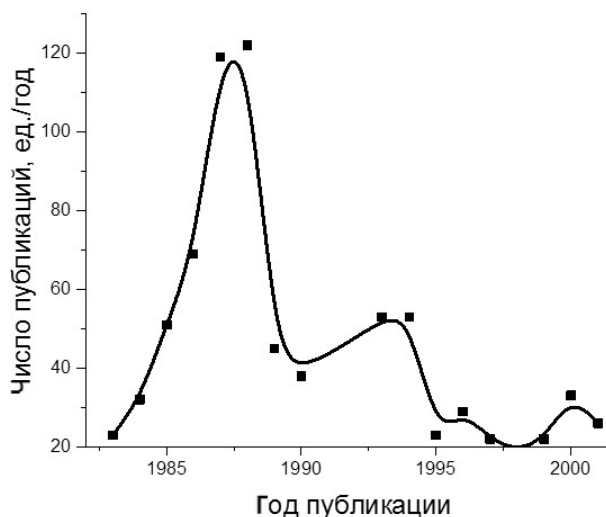


Рис. 5. Динамика числа публикаций по СПИД (AIDS) в журнале Nature

к июню 2002 г. Число авторов, аффилированных с США равно 73 (60.8%), далее многократно меньше авторов из Великобритании, Канады, Италии. Авторы из РФ отмечены только в одной статье.

При формировании поискового запроса Aids & Nature (journal) выявлено 1122 публикации. Хронологическое распределение до 2001 г. приведено на рис. 5. Далее идет спад публикационной активности: В 2002–2011 гг. — колебания от 10 до 18 публикаций, в 2012–2021 гг. — колебания от 2 до 13 статей. Начальный рассматриваемый период 1980–1982 гг. публикаций не более одной ежегодно. На графике (рис. 5) отражен наиболее информативный период 1983–2001 гг. Собственно, согласно Википедии, 1983 г. и считается открытием ВИЧ.

Максимум публикаций (порядка 120) достигается через 3–4 года после открытия ВИЧ, а именно в 1986–1987 гг.

Все эти публикации формально отнесены к междисциплинарной области знаний. Кроме того, указаны следующие области: биохимия, генетика, молекулярная биология (504), медицина (180), иммунология и микробиология (130) и др.

Более всего авторов из США (697), далее Великобритания (206), Франция (77), Германия (48), Канада (43). Относительно малый вклад Китая — 22 (12-е место). РФ по степени участия ниже — 39-е место. Низкая доля участия Китая, по видимому, связана с тем, что значительное развитие публикационной активности китайских ученых связано с XXI в, а большинство публикаций относятся ко второй половине 1980-х гг.

Поиск по запросу Flu & Nature (journal) позволяет сделать следующие наблюдения: наибольшее число публикаций приходится на 2005–2006, 2009 и 2012 гг. (соответственно 36, 23, 21 и 22). После 2012 г. публикационная активность заметно снизилась (не более 10 статей в год).

По-прежнему лидеры — авторы из США (34), далее многократно меньше из Великобритании (7), Нидерландов (6), Франции и Германии — по 5.

8. СТАТИСТИКА ПУБЛИКАЦИЙ, ХАРАКТЕРИЗУЮЩИХ МЕТОДИКИ АНАЛИЗА ЗАБОЛЕВАНИЙ

Целесообразно использовать ключевые слова либо protein, либо immunoassay.

Комбинация Covid & protein позволяет выявить 16 197 публикаций на 15.06.2021. Практически все, за исключением 4 статей, относятся к 2020 г. (9213) или 2021 г. (6980).

По областям знаний: медицина — 10 740

(66.3%), биохимия, генетика, молекулярная биология — 4709 (29.1%), иммунология и микробиология — 3213 (19.8%), фармакология, токсикология — 1696 (10.5%), химия — 930 (5.7%).

По государственной принадлежности авторов: США — 4104, Китай — 2886, Италия — 1707, Индия — 1453, Великобритания — 1295, Германия — 820, Франция — 724, ..., Россия — 224 (20-е место).

Из российских организаций более всего авторов представляют Московский государственный медицинский университет им. И.М. Сеченова (67), РАН (46), МГУ (33), Минздрав РФ (25).

При формировании запроса Covid & immunoassay найдено 1227 публикаций (631 в 2020 г., 596 в 2021 г.). По областям знаний: медицина — 915; биохимия, генетика, молекулярная биология — 309; иммунология и микробиология — 291; химия — 86; техника и фармакология, токсикология — по 57.

По государственной принадлежности авторов: США — 290, Китай — 193, Италия — 123, Великобритания — 118, Германия — 91, Испания — 55, Франция — 50, ..., Россия — 11 (25-е место).

По запросу Ebola & protein отобрано 2798 статей. Примерная динамика числа публикаций: до 1997 г. — не более 10 статей в год, 1998–2002 гг. — от 20 до 40, 2003–2009 гг. — от 50 до 70, далее от 2010 г. — соответственно 89, 128, 109, 87, 140, 290, 252, 231, 238, 208, 275 и 105 статей. Более 200 статей в год — уровень публикаций, достигнутый к 2015 г.

По областям знаний: иммунология и микробиология — 1240, медицина — 1050, биохимия, генетика, молекулярная биология — 985, сельскохозяйственные и биологические науки — 369, фармакология, токсикология — 343.

По государственной принадлежности авторов: США — 1627, Германия — 323, Великобритания — 228, Канада — 212, Китай — 210, Франция — 195, ..., Россия — 56 (11-е место).

Из Российских организаций представлены: НПО "Вектор" (22), РАН (11), СО РАН (9), МГУ и Минздрав РФ (по 7).

Для сравнения: при запросе Ebola & immunoassay выделено 122 статьи с общим числом цитирований 2253. Индекс Хирша — 25, наибольшее число цитирований — 166.

Примерная динамика числа публикаций за последние несколько лет начиная с 2013 г. — 1, 2, 8, 13, 11, 14, 17, 13 и 8 соответственно.

По областям знаний: медицина — 73, иммунология и микробиология — 46, биохимия, генетика, молекулярная биология — 33, химия — 10, техника — 8.

По государственной принадлежности авторов:

США — 65, Великобритания — 13, Китай — 10, Швейцария — 9, Канада, Сьерра-Леоне и Россия — по 8.

9. ОСОБЕННОСТИ ОТБОРА ПУБЛИКАЦИЙ ПРИ ЗАДАНИИ СОЧЕТАНИЯ КЛЮЧЕВЫХ СЛОВ: COVID И РАВНОЗНАЧНЫЙ ЕМУ ДИАГНОЗ

В нашем случае речь идет о сопутствующих COVID диагнозах (заболеваниях).

Запрос COVID & flu позволил выявить 1101 публикаций (2020 г. — 686, 2021 г. — 415).

По областям знаний распределение достаточно типичное: медицина — 783, биохимия, генетика, молекулярная биология — 144, иммунология и микробиология — 134, фармакология, токсикология — 92, социальные науки — 72.

По государственной принадлежности авторов: США — 248, Италия — 170, Индия — 135, Великобритания — 99, Китай — 69, Бразилия — 67, ..., Россия — 14 (23-е место).

По запросу COVID & AIDS отобрано 502 публикации (257 и 241 в 2020 и 2021 гг. соответственно).

По областям знаний: медицина — 309; социальные науки — 77; биохимия, генетика, молекулярная биология — 47; иммунология и микробиология — 39; компьютерные науки — 32.

Видно, что данное распределение по областям знаний заметно отличается от предыдущего.

По государственной принадлежности авторов: США — 168, Великобритания — 57, Индия — 39, Китай — 35, Австралия — 24, Италия — 21, ..., Россия — 9 (18-е место).

При запросе COVID & oncology выявлено 1539 публикаций (1000 и 539 в 2020 и 2021 гг. соответственно).

По областям знаний: медицина — 1449; биохимия, генетика, молекулярная биология — 448; фармакология, токсикология — 51.

По государственной принадлежности авторов: США — 502, Италия — 275, Великобритания — 183, Франция — 114, Индия — 111, ..., Китай — 62 (11-е место), ..., Россия — 10 (36-е место).

10. ОСОБЕННОСТИ ОТБОРА ПУБЛИКАЦИЙ С ДВУХ- И ТРЕХЭТАПНЫМ ПОСЛЕДОВАТЕЛЬНЫМ ЗАДАНИЕМ КЛЮЧЕВЫХ СЛОВ ИЛИ ИХ КОМБИНАЦИЙ

1-й этап: определение публикаций, содержащих два медицинских диагноза, включая COVID, а также медико-диагностическую информацию.

COVID-ассоциирующиеся ключевые слова и выражения и интерпретации соответствующего отбора публикаций с первым ограничением по ключевому слову COVID.

2-й этап: далее анализируются наиболее значимые подгруппы публикаций по дополнительным признакам (ключевым словам).

А) SARS-CoV-2 — 42 197 упоминаний в 157 871 публикации (на 16.06.2021), или 26.7%.

Б) Вирусная пневмония — 35 375, или 22.4%.

В) Показательно "торжество" гендерного равенства — 30 079 упоминаний женщин и 29 144 упоминаний мужчин.

Г) "Фактор риска" — 9656 и "факторы риска" — 5054 публикаций; суммарно 14 170 публикаций, или 9.3%. Таким образом, выбор очень чувствителен к формальностям запроса.

Д) Психология — 7378 упоминаний, или 4.7%.

Далее примерно 4500 упоминаний о С-реактивном белке и гипертензии.

При выполнении на 1-м этапе анализа ассоциаций с ключевым словом COVID при формальном ограничении публикаций, только отнесенных к области знаний "медицина", имеем следующее.

Число отобранных публикаций снизилось до 101 345, или 64.2% (это достаточно стабильный показатель рассматриваемой доли публикаций).

Суммарное число упоминаний факторов риска в различных вариантах написаний составило 12 840 в сравнении с 14 170 публикациями, указанными выше в п. Г). То есть примерно 12.5% упоминаний о факторах риска COVID не относятся к медицинской сфере непосредственно.

Число упоминаний С-реактивного белка и гипертензии снизилось примерно до 4100 и 3900 соответственно. Очевидно, что данные ключевые слова (выражения) относятся к медицинским аспектам. Тем не менее некоторые публикации, в которых упомянуты данные выражения в качестве ключевых, формально не отнесены к медицине либо по характеру публикации, либо по выбору источника.

Заметим, что, например, психология и медицина рассматриваются в качестве различных областей знаний.

Также интересно по крайней мере трехвариантное упоминание полимеразной цепной реакции, а именно: собственно ПЦР — 2849 раз, ПЦР с обратной транскрипцией — 3409 или ПЦР в реальном времени — 2543.

3-й этап: анализируется роль дополнительных ограничений, включающих конкретные факторы риска.

Можно оценить рейтинг факторов риска по их упоминанию. Возможно, это косвенно отразит

и степень значимости этих факторов. Хотя последнее представляется дискуссионным.

Некоторые упомянутые факторы риска COVID: гипертензия — 13.2%, диабет — 12.6%, депрессия — 6.5%, возрастной фактор — 6.1%, хроническая обструктивная болезнь легких (ХОБЛ) — 4.3%, курение — 3.0%.

Эти факторы упоминаются в различных публикациях, не только научных, в качестве осложняющих при COVID.

11. ОБСУЖДЕНИЕ И РЕКОМЕНДАЦИИ

Ассоциации со словом COVID локализованы во времени периодом 2019-2021 гг. Кроме того, при наукометрическом анализе существенным фактором формирования выборки публикаций является порядок задания ключевых слов.

Отбор, начиная с COVID, представляется менее информативным, если требуется искать сопутствующие диагнозы. Лучше начинать со второго диагноза.

В Разд. 1 статьи (табл. 1) рассмотрено условное разделение публикаций, ассоциирующихся со словом COVID, на 8 основных областей знаний. В данном разделе рассмотрим набор ключевых слов, чаще всего встречающихся в публикациях по 5 из указанных областей. При этом не анализируется содержание публикации. Рассмотрению подлежали: а) слова, указанные в соответствующей выборке публикаций в качестве ключевых; б) оценка соответствия нашего представления об отрасли знаний ключевому слову (выражению); в) влияние вариаций написания одного и того же по смыслу ключевого слова или выражения; г) основные тенденции, связанные с выбором ключевых слов (выражений).

Первая по значимости область знаний — медицина (103 044 публикации на 23.06.2021).

Слова и выражения, которые наиболее часто ассоциируются с публикациями по COVID и областью знаний медицина: human(s), pandemic(s), SARS-CoV-2, Virus Pneumonia, coronavirus (disease или infection). Эти англоязычные выражения легко переводимы на русский язык.

Несколько реже встречаются слова, априорно ассоциирующиеся с медициной: эпидемиология (более 11 700 раз), госпитализация (около 8000), психология (примерно 5680 упоминаний). Интересным обстоятельством является уже отмеченное ранее практически равное присутствие ключевых слов женщина (26 406) и мужчина (25 662). Является ли это случайным совпадением или своеобразной данью толерантности (гендерному равенству), определить достаточно трудно.

Следующая область знаний — социальные науки (20 712 публикаций).

Наиболее распространенные ключевые слова и выражения практически полностью совпадают с предыдущими. Несколько чаще встречаются ключевые слова: эпидемиология (более 900) и психология (более 650). Также практически равное упоминания женщин (802) и мужчин (759). В этом плане можно говорить об общности медицины с социальными науками.

Однако, в отличие от предыдущей подгруппы, гораздо чаще встречаются ключевые слова и выражения, связанные с влиянием пандемии на общество. В частности, слова образование и обучение в различных вариантах написания: education, distance education, learning, e-learning, online learning, distance learning. Также достаточно распространены выражения, не только по существу, но и формально связанные с социальной сферой, а именно social media, organization and management.

Computer sciences — ожидаемо специфическая область знаний (9136 публикаций). В этой выборке публикаций наиболее существенные отличия от рассмотренных ранее. В частности, довольно серьезное изменение набора ключевых слов, входящих в первую десятку по распространенности по сравнению с двумя предыдущими выборками. Так, на 5-м и 10-м местах присутствует deep-learning и machine-learning, соответственно 816 и 626 упоминаний. Ключевые слова "студенты" (552 раза) и "социальные сети" (347) более свойственны социальным наукам. В то же время достаточно часто встречается оборот "Big Data", который специфичен именно для компьютерных наук.

Еще пара выборок по двум близким областям знаний имеет достаточно сходные особенности распределений ключевых слов и выражений.

Область знаний, объединяющая биохимию, генетику и молекулярную биологию (15 765 публикаций). Ключевое слово "генетика" встречается не слишком часто (всего в 1834 публикациях, что является 30-м показателем по частоте). Здесь, как и в некоторых предыдущих случаях, число ссылок на мужчин и женщин практически равное. Вторая тенденция — присутствие слов и выражений, устойчиво ассоциирующихся с данными областями знаний. Например, достаточно часто упоминаются антивирусный агент, патогенетика, лекарственный эффект, химия и ряд других. Обращает внимание частое упоминание различных слов и выражений, связанных с возрастом: возраст, средний возраст, взрослые, юные и т.п.

Заключительная из 5 рассматриваемых областей знаний объединяет иммунологию и микробиологию (10 240 публикаций). Практически полное совпадение первых 10–15 ключевых выражений с предыдущей группой. Среди распростра-

ненных ключевых выражений присутствуют как априорно связанные с иммунологией, так и в большей степени ассоциирующиеся с медициной выражения: пневмония, иммунология, вирусология, бета-коронавирус, генетика, метаболизм, иммунный ответ.

В целом можно утверждать, что не всегда наиболее распространенные в публикациях ключевые слова априорно ассоциируются у нас с соответствующими областями знаний. Возможно, что отнесение публикации при ее отборе к определенной отрасли науки определяется не столько содержанием публикации, сколько формальными факторами: названием журнала, тематикой конференции, материалами которой рассматриваются в публикации и т.п.

ЗАКЛЮЧЕНИЕ

При анализе публикаций по рассматриваемой тематике (в основном COVID, SARS), индексируемых в базе данных Scopus, выявлены некоторые тенденции.

Публикации, ассоциированные с ключевым словом COVID, преимущественно ожидаемо отнесены к области медицины, их доля составила 65–70%, и лишь для авторов из РФ она существенно ниже (55%). Доля публикаций, отнесенных к социальным наукам, для авторов из большинства стран составляет около 10–12%, лишь для авторов из Китая не более 6%; вклад области знаний биохимия составляет 10–15%; для области психология доля составляет 3–5%; доля публикаций по компьютерным наукам и технике для США и Великобритании — около 3%, а для РФ и Китая — примерно вдвое большая (5.5–6.5%). Эти тенденции могут объясняться как приоритетами научно-технической стратегии развития указанных государств, так и менталитетом или/и особенностями социальной политики.

Публикации, связанные с SARS, хронологически более растянуты. Доля статей, формально отнесенных к медицине, — многократно меньше, чем для COVID (29% в сравнении с 65%); также отмечены многократно большие доли публикаций по технике и компьютерным наукам, а именно около 23 и 22% (для сравнения, в публикациях по COVID эти доли 4–5%).

Публикации по отраслям, формально относящимся к генетике, биохимии, молекулярной биологии, имеют устойчивое представительство в обеих группах публикаций (соответственно 11 и 16%).

Рост публикационной активности постоянен с 1950 г. Основные публикации приходятся на авторов из США и Великобритании, а с начала 2000-х гг. резко возрастает участие авторов из Ки-

тая, особенно в период 2006–2019 гг.

Пандемия 2020–2021 гг. оказала сопоставимое влияние на публикационную активность как связанную с ключевым словом COVID, так и с ключевым словом SARS, что отражено в многократном увеличении числа публикаций по COVID и соответствующим увеличением публикаций по SARS. На этот период сохранена ведущая роль авторов США (более 37%), второе место заняли авторы из Китая (чуть менее 25%). У авторов из РФ место в рейтинге с 12-го по 23-е. Главным образом авторы из РФ представляют НПО "Вектор", Новосибирский университет и РАН.

Приведены результаты статистического анализа выборок по соответствующим группам публикаций, связанных с развитием наиболее известных пандемий конца XX и начала XXI вв.

По публикациям, ассоциированным со словом Эбола (Ebola), наблюдается кратковременный всплеск публикационной активности в период 2016–2018 гг. До этого длительный период (20 лет, с 1982 г.) число годовых публикаций росло практически линейно, увеличиваясь на 6–7 статей в год. Столь явного доминирования авторов США, которое наблюдалось практически во всех остальных выборках, здесь не выявлено. Распределение публикационной активности по времени имеет колоколообразный характер с максимумом вблизи 2006 г.

Еще одна очевидная тенденция: география распространения пандемии косвенно отражается в государственной принадлежности авторов. В частности, доля публикаций, связанных с вирусом Эбола, для авторов из стран Центральной Африки явно выше средних значений. Аналогичная тенденция проявляется при оценке доли авторов из Восточной Азии в публикациях, относящихся к птичьему гриппу.

В выборке публикаций по тематике, связанной с птичьим гриппом (Bird's flu) и ДНК, авторы из РФ имеют наивысший рейтинг из рассматриваемых групп публикаций: рейтинг 6–7 с долей около 4.5%. Если первые два места рейтинга для США и Китая достаточно традиционны, то далее идут представители стран Азии, затронутых ранее птичьим гриппом, а именно Гонконг и Сингапур.

Отдельно проанализировано содержание наиболее цитируемых по каждой рассматриваемой группе статей. В ряде случаев число цитирований превышает 10 000 менее чем за 2 года, прошедших после публикации.

Наиболее цитируемые статьи главным образом связаны с описанием симптомов COVID, историей происхождения, оценками возможностей вакцин и терапии в целом.

Однако часть наиболее цитируемых публикаций связана не с чисто научными или научно-технологическими аспектами, но с анализом (описанием) социальных аспектов, в частности, последствий различных пандемий для туристического бизнеса.

При анализе публикаций, связанных с COVID или аналогичными пандемиями, в журнале Nature выявлены как закономерности, совпадающие с общими тенденциями, так и специфические особенности. Одной из общих закономерностей является резкое возрастание публикаций по СПИДу, относящееся к 1986–1987 гг. Практически наблюдаем 3-летнюю задержку по отношению к дате открытия СПИДа.

Отбор публикаций при задании сложного (многоэтапного) запроса также выявил ряд тенденций. Основные из них:

1. Существенную роль играет последовательность задания ключевых слов.

2. Выбор чувствителен к формальному написанию ключевого слова или выражения (например, отдельно отбираются публикации при задании слов: human или humans)

3. В большинстве случаев наблюдается априорно определенное соответствие между рассматриваемой областью знаний и ключевыми словами и выражениями. Например, для медицины набор ключевых слов может включать диагностические признаки (одышка, гипертензия и т.п.), для области знаний компьютерные науки вполне ожидаемы термины (алгоритм, машинное обучение, Big Data и т.п.).

4. Отдельные тенденции не имеют научного обоснования, но являются отражением социальных аспектов. Например, практически равное упоминание мужчин и женщин даже применительно к выборкам пациентов с очевидным доминированием мужчин.

Анализ уровня публикационной активности российских ученых в рассматриваемых областях знаний свидетельствует о том, что входжение в пятерку мировых государств по числу международных публикаций представляется трудно достижимой задачей. Даже если учитывать один из ключевых принципов статистики — неравномерность распространения закономерностей, выявленных в специальной группе, на более общую группу данных. В данном случае рассматриваемые научные публикации относятся в основном к естественнонаучным областям знаний (физика, биология, химия, медицина, техника, компьютерные науки, аналитические методы и т.д.), соответствующие оценки вклада ученых из РФ в публикации базы Scopus все-таки дают представление об их вкладе в научные публикации в целом.

Благодарности

Авторы выражают благодарность преподавателям курсов повышения квалификации на базе С.-Петербургского политехнического университета Петра Великого по направлению "Научная публикация в международных изданиях: подготовка, подача, продвижение", в первую очередь Ватину Николаю Ивановичу и Якубсон Вере Михайловне за данные ими советы и рекомендации, которые позволили повысить эффективность работы с библиометрическими базами данных.

Работа выполнена в рамках Государственного задания Министерства науки и высшего образования Российской Федерации № 075-00280-21-00.

СПИСОК ЛИТЕРАТУРЫ

- Huang C., Wang Y., Li X., Ren L., Zhao J. et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China // *The Lancet*. 15–21 February, 2020. Vol. 395, iss. 10223. P. 497–506. DOI: 10.1016/S0140-6736(20)30183-5
- Wölfel R., Corman V.M., Guggemos W., Seilmaier M. et al. Virological assessment of hospitalized patients with COVID-2019 // *Nature*. 28 May, 2020. Vol. 581, iss. 7809. P. 465–469. DOI: 10.1038/s41586-020-2196-x
- To K.K.-W., Tsang O.T.-Y., Leung W.-Sh. et al. Temporal profiles of viral load in posterior oropharyngeal saliva samples and serum antibody responses during infection by SARS-CoV-2: an observational cohort study // *The Lancet Infectious Diseases*. 2020. Vol. 20, iss. 5. P. 565–574. DOI: 10.1016/s1473-3099(20)30196-1
- Gordon D.E., Jang G.M., Bouhaddou M. et al. A SARS-CoV-2 protein interaction map reveals targets for drug repurposing // *Nature*, 2020. Vol. 583. P. 459–468. DOI: 10.1038/s41586-020-2286-9
- Broughton J.P., Deng X., Yu G. et al. CRISPR–Cas12-based detection of SARS-CoV-2 // *Nat Biotechnol*. 2020. Vol. 38. P. 870–874. DOI: 10.1038/s41587-020-0513-4
- Lu R., Zhao X., Li J., Niu P. et al. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding // *The Lancet*. 22–28 February, 2020. Vol. 395, iss. 10224. P. 565–574. DOI: 10.1016/S0140-6736(20)30251-8
- Kim D., Lee J.-Y., Yang J.-S., Kim J.W., Kim V.N., Chang H. The Architecture of SARS-CoV-2 Transcriptome // *Cell*. May 14, 2020. Vol. 181, iss. 4. P. 914–921. DOI: 10.1016/j.cell.2020.04.011.
- Jones K.E., Patel N.G., Levy M.A., Storeygard A., Balk D., Gittleman J.L., Daszak P. Global trends in emerging infectious diseases // *Nature*. February, 2008. Vol. 451, iss. 7181. P. 990–993. DOI: 10.1038/nature06536
- Drosten C., Götting S., Schilling S., Asper M., Panning M., Schmitz H., Günther S. Rapid detection and quantification of RNA of Ebola and Marburg viruses, Lassa virus, Crimean-Congo hemorrhagic fever virus, Rift Valley fever virus, dengue virus, and yellow fever virus by real-time reverse transcription-PCR // *Journal of Clinical Microbiology*. 2002. Vol. 40, iss. 7. P. 2323–2330.
- Tchesnokov E.P., Feng J.Y., Porter D.P., Götte M. Mechanism of inhibition of ebola virus RNA-dependent RNA polymerase by remdesivir // *Viruses*. 2019. Vol. 11, iss. 4. 326. DOI: 10.3390/v11040326
- Gire S.K., Goba A., Andersen K.G., Sealfon R.S.G., Park D.J., Kanneh L., Jalloh S., Momoh M., Fullah M., Dudas G., Wohl S., Moses L.M. Genomic surveillance elucidates Ebola virus origin and transmission during the 2014 outbreak // *Science*. 12 September, 2014. Vol. 345, iss. 6202. P. 1369–1372. DOI: 10.1126/science.1259657.
- Dhama K., Khan S., Tiwari R., Sircar S., Bhat S., Malik Y.S., Singh K.P., Chaicumpa W., Bonilla-Aldana D.K., Rodriguez-Morales A.J. Coronavirus disease 2019–COVID-19 // *Clinical Microbiology Reviews*. October, 2020. Vol. 33, iss. 4. P. 1–48. DOI: 10.20944/preprints202003.0001.v2
- Sui J., Hwang W.C., Perez S., Wei G. et al. Structural and functional bases for broad-spectrum neutralization of avian and human influenza A viruses // *Nature Structural and Molecular Biology*. March, 2009. Vol. 16, iss. 3. P. 265–273. DOI: 10.1038/nsmb.1566
- Chen H., Deng G., Li Z., Tian G., Li Y., Jiao P., Zhang L., Liu Z., Webster R.G., Yu K. The evolution of H5N1 influenza viruses in ducks in southern China // *Proceedings of the National Academy of Sciences of the United States of America*. 13 July, 2004. Vol. 101, iss. 28. P. 10452–10457. DOI: 10.1073/pnas.0403212101
- Rittichainuwat B.N., Chakraborty G. Perceived travel risks regarding terrorism and disease: The case of Thailand // *Tourism Management*. June, 2009. Vol. 30, iss. 3. P. 410–418. DOI: 10.1016/j.tourman.2008.08.001
- Whitehead K.A., Langer R., Anderson D.G. Knocking down barriers: Advances in siRNA delivery // *Nature Reviews Drug Discovery*. 2009. Vol. 8. P. 129–138. DOI: 10.1038/nrd2742

Институт аналитического приборостроения РАН, Санкт-Петербург (Буляница А.Л., Евстапов А.А.)

Санкт-Петербургский политехнический университет Петра Великого (Буляница А.Л.)

Контакты: Буляница Антон Леонидович,
antbulyan@yandex.ru

Материал поступил в редакцию 20.07.2021

THE MAIN TRENDS OF PUBLICATION ACTIVITY ON TOPICS RELATED TO INFECTIOUS DISEASES: COVID, SARS, EBOLA AND BIRD FLU

A. L. Bulyanitsa^{1,2}, A. A. Evstrapov¹

¹*Institute for Analytical Instrumentation of RAS, Saint Petersburg, Russia*

²*Peter the Great St. Petersburg Polytechnic University, Russia*

The article contains a scientometric analysis of publications indexed in the international citation database Scopus. Groups of publications containing the keywords COVID / COVID-19, the SARS virus that causes it, as well as publications on pandemics caused by Ebola and bird flu / avian influenza (various options combined by the search query "bird* flu") were considered. It was supposed to identify groups of links between the development of diseases and the dynamics of publication activity on the relevant topic, to analyze the samples of publications formed by sequential (two- and three-stage) search queries. In addition to the dynamics of the number of publications, distributions were analyzed by a) fields of knowledge (branches of science), b) by a set of keywords associated with a specified area, c) by nationality of authors and, in relation to authors from Russia, the significance of the contribution of organizations to the sample of publications.

Due to the current quantitative changes in the number of publications for the period of material selection (mid-March – late June 2021), some quantitative estimates may change slightly. At the same time, the qualitative conclusions are preserved.

For the most part, the findings appear to be expected.

Keywords: COVID (or COVID-19), SARS, Ebola, bird flu, pandemic, number of publications, cites, fields of knowledge, Scopus

INTRODUCTION

The purpose of the article is to analyze the main trends in publication activity on topics related to pandemics caused by SARS, Ebola and bird flu / avian influenza viruses. The selection of publications was carried out using the SCOPUS bibliometric database in March – June 2021.

The main tasks were:

1. Assessment of the distribution of publications by areas of knowledge.
2. Determination of the main methods and techniques of research, detection and analysis of the indicated diseases.
3. Identification of key (most cited) articles in different areas of research and / or areas of knowledge.
4. Analysis of the main trends in publication activity, in particular, the dynamics of the number of publications and their relationship with outbreaks of epidemics.

The selection was carried out according to search schemes with a different choice of keywords and / or their combinations. At the same time, the degree of relevance between the content of the publication and the selected keyword and / or combination of words was not evaluated.

The key direction was a comparative analysis of the following distributions of publications in the

SCOPUS database on topics associated with the words: COVID, SARS, Ebola and bird* flu (the symbol "*" allows different variants of subsequent letters or their absence) according to data for March – June 2021, grouped by:

- 1) areas of expertise;
- 2) languages of publications;
- 3) the nationality of the authors;
- 4) those with additional search restrictions (analysis method, combination of two or more diseases, additional associated words and expressions). That is, a two- or three-stage sequential selection (limitation) of publications.

Also, the subject of consideration was the dynamics of the number of publications and the determination of the share of authors from the Russian Federation in publications on the topic under consideration.

1. ANALYSIS OF PUBLICATIONS RELATED TO THE KEYWORD COVID

Keyword search on 14.03.2021: COVID — 114 202 publications, COVID-19 — 113 041 publications, that is, in terms of the choice of keywords, it is almost the same. On 24.03.2021, the number of publications increased to 119 145, and on 18.05.2021 to 140 379, that is, by about 25%. As of 01.06.2021

148 064 publications were noted, on 23.06.2021 the number of publications reached approximately 160 700. Fig. 1 illustrates the growth in the number of publications in the short-term period of work on the article.

Fig. 1. Changing the number of publications in the Scopus database when specifying the COVID keyword

The growth in the number of publications is close to linear within the considered 3-month interval. At the same time, the increase in the number of publications within 5 days is approximately 2%.

Only a small part of the increase in the number of publications is explained by additionally found publications of 2020 year: 84 874 by mid-March and 85 131 by mid-May 2021. At the same time, the number of publications of 2021 increased from 29 273 to 55 185, or by about 88%. Publications of 2018–19 are about 50, which is less than 0.04% of the total number of publications on this topic. The main type of publications were articles — 68 095 (59.6%), letters — 14 919 (13.1%), reviews — (11.4%).

Tab. 1 shows the distribution of publications in the five most significant fields of knowledge. Since the publication can be attributed to several areas at the same time, then the amount can exceed 100%.

Tab. 1. Distribution of publications by areas of knowledge (only areas with the largest relative contribution)

Fundamental differences in the distributions of publications of 2020 and 2021 do not exist. It should be noted a slight increase in the share of publications in immunology and microbiology, as well as in the field of technology and psychology due to the reduction of publications related to medicine. This trend may reflect not so much a change in research priorities as a difference in the average time to prepare a material for publication (for example, a technical direction may require a longer period than a medical publication.)

Note that the most cited article on 24.03.2021 [1] has 12 997 citations within nearly a year after publication (as of 01.06.2021, the number of citations exceeded 15000), since The Lancet magazine is one of the top-rated publications recognized by the international scientific community.

Distribution by language of publications:

English — 107 576 (94.2%), Spanish — 2700 (2.4%), Chinese — 1282 (1.1%), French — 1211 (1.1%), Russian — 580 (0.5 %).

Tab. 2 presents data on the total number of publications on 10.04.2021, depending on additional keywords. Tab. 3 illustrates the share distribution of publications by areas of knowledge for individual groups of the sample under consideration, namely, publications on COVID and another search criterion (key

expression). There are significant differences in the distributions by area of expertise for the three subgroups of COVID publications under consideration, limited by different analytical methods.

Tab. 2. Indicators of publication activity depending on the combination of the COVID keyword and the additional combination of keywords

Tab. 3. Distribution by formal areas of knowledge for different groups of publications

Main trends:

1. The share of publications related to medicine for PCR-RT is dominant (over 80%) and is twice as high as for the other two areas.

2. Apparently, there is a "transition" of some publications from the field of formal medicine to biophysics, genetics and molecular biology (about 17% for PCR-RT and about 45-48% for other subgroups).

3. The share of computer science in general is small, but for the direction of isothermal amplification it is much higher (6% versus 3 and 1.5%).

4. The field of pharmacology makes up most of the publications related to mass spectrometry. Apparently, these are tasks associated with the identification of drugs.

5. The actual technical direction is critically presented only in relation to isothermal amplification (18%), and in other directions about 2–3%.

The difference in distributions may be due not only to the development of research techniques, but also to the difference in the subject of the journals in which the results are published.

Fig. 2 (a–r) shows the distribution of publications by area of knowledge for authors of different nationalities.

Fig. 2. Distribution of publications by areas of knowledge for authors of different nationalities

General trends:

1. The share of publications in the field of medicine is mainly 65–70%, only for authors from the Russian Federation it is significantly lower (55%)

2. The share of publications on social sciences in most cases is more than 12%, only for authors from China is about 6%.

3. The contribution of the field of biochemistry is 10–15% in all cases, for the field of psychology a share is 3–5%.

4. The share of publications on computer and technical sciences for the USA and Great Britain is about 3%, and for the Russian Federation and China it is approximately twice as large (5.5–6.5%).

These trends can be explained both by the priorities in the scientific and technical policy of these

states, and by the mentality and / or peculiarities of social policy.

2. COMPARISON WITH PUBLICATIONS BY SEARCHWORD SARS (LONG PERIOD OF TIME)

Total number of publications is 150 764 on 14.03.2021. Dynamics for the last period: 2018 — 6272, 2019 — 6462, 2020 — 37 181, 2021 — 13 966, of which: articles — 88 376 (58.6%), conference proceedings — 36 533 (24.2%), reviews — 12 085 (8.0%). By 24.03.2021 the number of publications increased to 154 311. The share of articles is close to publications on COVID, and the share of conference materials is much larger. This may be due to the fact that COVID publications are from the 2020–2021 period, when the number of conferences decreased due to the corresponding restrictions.

Tab. 4 shows the publication data on topics related to SARS on 10.04.2021.

Tab. 4. Distribution of publications by keyword combinations containing SARS

The most cited articles by subgroup, such as PCR-RT combined with either SARS or COVID, are exactly the same. That is, the most cited articles are related to the research jointly with SARS and COVID. Tab. 5 shows the distribution of publications associated with SARS by area of expertise. The areas that are the largest in terms of contribution are noted.

Tab. 5. Distribution of publications by area of expertise for the group selected by the SARS keyword

Obvious trends.

For articles on SARS published over a longer period, the portion related to medicine is many times less than on COVID (29% compared to 65%), but portions related to engineering and computer science are many times more — 23% and 22% against, respectively 4% and 5%.

At the same time, social sciences and psychology account for a large share in publications on COVID (in total, about 16–17%). Publications in the branches formally related to genetics, biochemistry, and molecular biology have a stable representation in both groups of publications (11 and 16%, respectively). Tab. 6 shows the distribution of publications of different years according to the nationality of the authors.

Tab. 6. Distribution of authors by nationality (on 14.03.2021)

The discrepancy between the distributions of publications on SARS and COVID for the selected eight

countries can be estimated using the Kullback–Leibler divergence. This discrepancy is approximately 0.155 nits. Such a value corresponds, for example, to the divergence of two normally distributed random variables with unit variance when the mathematical expectations are shifted by about 0.39. This discrepancy cannot be considered significant.

We will separately consider the dynamics of publication activity for the period since 1950 on the SARS for authors of various nationalities (through affiliated organizations) on 24.03.2021. This information is given in Tab. 7.

Tab. 7. Dynamics of SARS publication activity of authors from the USA, China, Great Britain and Russia

During the period under review, there are 2088 publications related to SARS associated with the Russian Federation (according to the SCOPUS database on 24.03.2021). The Hirsch index of these publications is 55. The share of publications, in comparison with the USA, China and Great Britain authors, is small and does not exceed 2.1%.

Main areas of knowledge: medicine — 540 or 25.9%, Earth and planetary sciences — 494 (23.7%), computer science — 429 (20.5%), engineering — 383 (18.3%), biochemistry, genetics and molecular biology — 324 (15.5%).

This distribution is slightly different from the general distribution of publications on SARS, regardless of the nationality of the authors.

Note the growth of publication activity on this topic, which has been observed since 1950. The main research has been carried out in the United States and Great Britain. There is a surge in publication activity in 2001–2005. And from 2006 to 2019 the main share of publications is affiliated with China. During a pandemic, the publication ratio changes in favor of the United States.

In the time interval up to 2019 inclusive, when there was no influence of the COVID pandemic, the dynamics of the average annual number of publications Y is quite accurately described by the power law:

$$Y = 1.067 \cdot 10^{-3} \cdot (t - 1950)^{3.7},$$

where t — is the considered publication year. The coefficient of determination of the dependence, reduced to linear in log-log coordinates, exceeds 0.983, which indicates the adequacy of the found dependence.

Accordingly, by compensating for this trend, it is possible to assess the impact of the covid factor on publication activity. In this case, approximately 30 000 publications in 2020 and 14 800 publications in 2021 (approximately 30% of the annual volume) remain uncompensated. The ratio of the number of pub-

lications on COVID for 2020 and 2021 is also close to 2: 1, namely, 85 thsd to 55 thsd (but 55 thsd publications were noted on 18.05.2021).

Consequently, the pandemic had a comparable impact on publishing activity, both related to the COVID keyword and related to the SARS keyword.

For the period 2014–2019 the share of publications classified as computer science was 30% or more. After the outbreak of the COVID pandemic, this share decreased many times due to the redistribution of publications into other areas: medicine, pharmacology, biochemistry, psychology and a number of others.

Fig. 3 shows the comparative dynamics of publication activity for COVID and SARS. There is a relation between the explosion in COVID publications and a corresponding increase in SARS publications, as noted earlier.

Fig. 3. Comparison of the number of publications in Scopus, related to keywords COVID and SARS

3. PUBLICATION TRENDS ASSOCIATED WITH THE EBOLA KEYWORD

The total number of publications is 13 941 articles. The share of publications before the early 1980s is about 1%. The dynamics of publication activity is illustrated in Fig. 4.

Fig. 4. Dynamics of the average annual number of publications associated with Ebola

The first 3 values are averaged over 10 years, the value of the period 2007–2011 — over 5 years, and then the annual number of publications on topics associated with the Ebola keyword is presented. There is a short-term surge in publication activity in the period 2016–2018. Prior to this, the number of publications per year in 1982–2013 grew almost linearly, increasing by 6–7 articles per year. The distribution of publications on topics related to the Ebola keyword is shown in Tab. 8.

Tab. 8. Distribution of publications related to the Ebola keyword by area of expertise (with the largest contribution)

Distribution of the number of publications by nationality of authors:

USA — 6026 (43.2%),
Great Britain — 1688 (12.1%),
France — 866 (6.2%),
Germany — 846 (6.1%),
Canada — 840 (6.0%),
China — 708 (5.1%).

Russia is in 23rd place with 133 publications (just under 1%). For comparison, the representatives of Sierra Leone are in 10th place with 427 publications, which is quite easily explained. This is one of the infection focuses for the Ebola virus. The most cited publication [8] has 3439 citations.

Let's consider a combination of keywords, namely, a search for two keywords Ebola & PCR. 308 articles were identified, the most cited of them [9] has 462 citations. The total number of citations is 9650, that is, an average of 31.3 citations per article. Hirsch index is 49. The distribution of publications by area of expertise and nationality of authors differs little from general data on the Ebola topic: the first four places are occupied by the United States, Great Britain, Germany and China. France and Canada moved up to 6th and 7th places with publication shares of 11 and 9%, respectively, while authors from Sierra Leone moved up to 5th place (13%). The share of publications by authors from the Russian Federation is also about 1%, but the place in the ranking of countries is lower (34th).

Choosing the key words Ebola & mass-spectrum*, 100 articles were identified with 1914 citations. Hirsch index is 25. Most cited article [10] has 229 citations. Here, approximately equal shares of publications from 35 to 39% per cent fall on 3 areas of knowledge: Immunology and microbiology; Medicine; Biochemistry, genetics and etc.

Searching for publications using the combination of the words Ebola & DNA, 893 publications were found. The most cited article [11] has 748 citations. The total number of citations for these publications as of 03.06.2021 was 36 270. Hirsch index is 96.

The distribution by areas of knowledge is similar to the previous one: the same 3 areas with a contribution of only 35 to 41%. From the previous leading seven countries, Sierra Leone dropped out and Germany and Great Britain changed places. The share of publications by authors from the Russian Federation reached 2.1% (19 publications), which corresponds to the 11th place. Mainly, authors from the Russian Federation represent NPO Vector, Novosibirsk University and the Russian Academy of Sciences.

According to the search query Ebola & nanopore & sequence, a total of 17 articles is found. This is not a representative sample.

Combination Ebola & isothermal & ampl* identifies only 45 publications with a total number of citations of 1048 (Hirsch index is 18). The most cited article has 157 references [12]. The leading role of authors from the USA (37.8%) was retained, but authors from China (24.5%) took the second place. Authors from the Russian Federation have 1 article (places in the rating from 12 to 23).

4. TRENDS IN PUBLICATION ACTIVITY ASSOCIATED WITH THE KEY PHRASE AVIAN INFLUENZA

Since the English versions of the spelling of the key expression allow for a certain variety, the generic expression bird* flu was chosen. The * symbol allows different variants of subsequent letters or their absence.

A total of 1553 articles were identified on this topic. Article [13] has 908 maximum citations on 03.06.2021 (as of 08.06.2021 the number of citations is 912).

The distribution of publications by areas of knowledge allows to identify the largest areas in terms of contribution. This information is summarized in Tab. 9.

Tab. 9. Distribution of publications on avian influenza by area of expertise

Distribution by nationality of authors:

USA — 287 (18.5%),
China — 122 (7.9%),
UK — 88 (5.7%),
India — 69 (4.4%),
France — 49 (3.2%),
Thailand — 48 (3.1%).

Further — Japan, Canada and Germany. Authors from the Russian Federation have 16 publications (traditionally, about 1%) and 23rd place in the rating. It is evident that such a clear dominance of US authors, which was observed in almost all other ratings, was not revealed here.

The distribution of publication activity over time is bell-shaped with a maximum near 2006. Due to the small fluctuations in the annual number of publications, the reliability of the approximating dependences is not very high.

The most adequate approximating dependences are the Gaussian curve with a nonzero background addition (approximately 55 publications per year) and a maximum of 273 or the Lorentz distribution with a background of 47 publications and a maximum of 517. The actual peak corresponding to 2006 is poorly described, but when approximating the area ± 6 years from the maximum, an acceptable quantitative coincidence is observed. The coefficients of determination for both dependences are equal to 0.80 and 0.84, which is considered acceptable if an approximating dependence is recognized as adequate.

By the set of keywords bird's flu & DNA, only 65 articles were selected with a total citation of 1953, Hirsch index is 19. Article [14] has the largest 454 references in the indicated group.

Here, about 23% contribution of authors from the USA and China, and then half the number of publications from the UK. Authors from the Russian Federation and Korea have 3 publications each (rating 6–7 with a share of about 4.5%).

On SARS & bird's flu, 108 articles were found with a total number of citations of 1600 and a Hirsch index of 20. The article [15] has a maximum of 218 citations. The number of publications over the past decade fluctuates between 1 and 8 articles per year. The sharp interest in the topic is evidenced by the growth of citations in 2020–2021. If in 2017–2019 the number of citations was about 100, then in 2020 this number increased to 280, and citations in 2021 have already exceeded 240.

In terms of areas of knowledge, the first place is still occupied by publications related to medicine, 46.3%, but the 2nd and 3rd places in the ranking of areas are occupied by: social sciences 20.4% and business, management, etc. (about 9%). The distribution by nationality of the authors differs from the previous ones: if the first two places of the United States and China in the rating are quite traditional, then representatives of Asian countries previously affected by bird flu, namely Hong Kong and Singapore, go further. There are no authors affiliated with the Russian Federation, but in about 13% of cases the authors are listed without state affiliation.

5. PUBLICATIONS WITH THE KEY EXPRESSION SARS & EBOLA

The combination of words SARS & Ebola allows to find 1012 articles with a total number of citations of 25 484 (on 03.06.2021). The Hirsch index is 79. The article [16] has a maximum of 2139 citations.

During the period 2013–2019, the number of articles ranged from 20 to 50; in 2020, there was an almost tenfold increase in the number of publications (up to 388), and in the first half of 2021 there were already 195 publications. The distribution of publications by field of knowledge and by nationality of authors is similar to most of the previous ones:

a) medicine — 51.8%, immunology and microbiology — 21.7%, biochemistry, genetics, etc. — 20.8%, pharmacology and toxicology — 13.9%, social sciences — 7.8%, the rest is less;

b) authors from the USA — 40.4%, Great Britain — 12.0%, China — 7.7%, India — 6.8%, Germany — 6.1%, etc. Authors from the Russian Federation have 4 publications (rating 49).

The number of citations of publications on this topic also gives a pronounced about 4-fold increase from about 1500–1700 citations in 2017–19 up to 7200 in 2020 and already 5500 in the first half of 2021.

6. ANALYSIS OF THE CONTENT OF THE MOST CITED PUBLICATIONS BY SELECTED SECTIONS (COMBINATIONS OF KEYWORDS)

A brief analysis of the content of the most cited article on the relevant topic is presented.

[1] (Sec. 1): There have been cases of pneumonia in Wuhan, China caused by the novel beta-coronavirus, the novel coronavirus of 2019 (2019-nCoV). Epidemiological, clinical, laboratory and radiological characteristics of patients were analyzed. Methods for collecting patient data: real-time PCR and next generation sequencing. A comparison was made between ICU patients and non-ICU patients. The most interesting part of the results of this statistical analysis: the majority of infected patients were men, 73%; less than half had underlying medical conditions, including diabetes 20%, hypertension 15%, and cardiovascular disease 15%. The average age was 49 years (IQR 41 0-58 0). 2/3 of them visited Huanan Seafood Market. The common symptoms at the onset of the disease were fever, cough, and myalgia or fatigue; sputum production was less common, and headache, hemoptysis, and diarrhea were relatively rare. Dyspnea developed in more than half of the patients on average on the 8th day. All patients had pneumonia with abnormal chest CT findings. Complications included acute respiratory distress syndrome, RNA anemia, acute heart injury, and secondary infection. In the first group of patients, the mortality rate was about 15%. There was a similarity revealed between the coronavirus and the severe acute respiratory syndrome. The main problems relate to the origin of the disease and epidemiology, duration of human transmission and the clinical spectrum of diseases.

[2] (Tab. 2): COVID-19 is an acute respiratory tract infection that originated at the end of 2019. Initially, there were about 14% of cases with severe courses and 6% of cases with critical courses.

The main issue is the lack of information about virus replication, immunity and infectivity of respiratory tract. Pharyngeal virus isolation showed high levels during the first week of symptoms, with a peak in RNA copy number per throat swab on the fourth day. The virus is readily isolated from throat or lung specimens, but not stool specimens. Blood and urine tests also showed no virus. It is possible to detect sequentially different virus populations in throat and lung samples from the same patient, which proves independent replication. The period of antibody production was estimated: within 7 days in 50% of patients and by the 14th day in all patients. In addition to the acute form, COVID-19 can manifest itself as a mild illness of the upper respiratory tract.

[3] (Sec. 2, Tab. 4): COVID-19 is causing serious outbreaks in communities and hospital settings. Complete data on respiratory viral load and antibody response in SARS-CoV-2 infected patients are not available. The results of a study of patients in two hospitals in Hong Kong (analysis of saliva and blood serum) are described. Serial viral load was determined by quantitative reverse transcription PCR. Whole genome sequencing was performed to identify possible mutations that occur during infection. The study period was 22.01–12.02.2020. Conclusions are following: a) strict infection control is necessary, b) saliva is a good non-invasive sample for analysis, both for patients and for medical staff.

In contrast to severe acute respiratory syndrome, COVID-19 patients had the highest viral load near the sites, which could explain the rapidly spreading nature of this epidemic. This finding highlights the importance of strict infection control and early use of potent antiviral agents, alone or in combination, for those at high risk. Serologic testing can complement RT-qPCR for diagnosis.

For different types of immunoglobulins, the probabilities of antibody formation are indicated (in all cases, more than 88% on the 14th day).

[4] (Tab. 2, 4): Coronavirus 2 (SARS-CoV-2), the causative agent of the coronavirus disease COVID-19, has infected more than 2.3 mln people and has resulted in more than 160 thsd deaths (Epub 2020 Apr 30). At that time, there were no antiviral drugs with proven clinical efficacy for the treatment of COVID-19, as well as vaccines that prevent infection with SARS-CoV-2. The results of cloning, labeling and expression of 26 of 29 SARS-CoV-2 proteins in human cells are shown, as well as the identification of human proteins associated with each of the SARS-CoV-2 proteins. For this, affinity purification mass spectrometry was used to reveal 332 highly reliable protein-protein interactions between SARS-CoV-2 and human proteins, including 66 medicinal human proteins. Two sets of pharmacological agents were identified that exhibited antiviral activity: inhibitors of mRNA translation and predicted regulators of sigma-1 and sigma-2 receptors.

[5] (Tab. 2, 4): The article contains information on the development of a fast (up to 40 min), easy to implement and accurate method of analysis based on CRISPR-Cas12 providing the detection of SARS-CoV-2 from RNA extracts of respiratory smears. The test was performed on 36 patients with COVID-19 and 42 patients with other viral respiratory infections. The proposed CRISPR-based assay method provides a visual and faster alternative to real-time PCR assay. The diagnostic capabilities of the SARS-CoV-2 disease control method are 95% positive and 100% negative predictive coincidence.

[6] (Tab. 2, 4): On 26.01.2020, more than 2000 cases of infection (2019-nCoV) have been confirmed, most of which involved people living in or visiting Wuhan. Human-to-human transmission has been confirmed. Samples were sequenced from nine inpatients, eight of whom visited the Huanan Seafood Market in Wuhan. Homology modeling was performed to investigate the likely receptor-binding properties of the virus. As a result: ten 2019-nCoV genome sequences obtained from nine patients were extremely similar, showing over 99.98% sequence identity.

This virus is closely related (with 88% identity) to two coronaviruses that cause severe acute respiratory syndrome (SARS), derived from bats. 2019-nCoV has a receptor-binding domain structure similar to that of SARS-CoV, despite changes in amino acids in some key residues. Interpretation: 2019-nCoV is different enough from SARS-CoV to be considered a novel beta-coronavirus infecting humans. It has been suggested that bats may be the original host for this virus, and that an animal sold at the Wuhan seafood market may be an intermediate host that promotes the virus in humans. The need for urgent investigation of the evolution, adaptation and spread of this virus is also stated.

[7] (Tab. 2, 4): It's stated that SARS-CoV-2 is the beta-coronavirus responsible for the COVID-19 pandemic with unknown transcriptomic architecture. Two complementary sequencing techniques have produced a high-resolution map of the SARS-CoV-2 transcriptome and epitranscriptome. In addition to the canonical genomic and 9 subgenomic RNAs, SARS-CoV-2 produces transcripts encoding unknown ORFs with fusion, deletion and / or shift of fragments. The technology of nanoporous direct sequencing of RNA has made it possible to find at least 41 RNA modification sites on viral transcripts with the most frequent AA-GAA motif and some other manifestations of modification.

[8] (Sec. 3): Analysis of emerging infectious diseases is carried out and the conclusion about their dependence on socio-economic and environmental factors is presented. A database of 335 epidemics from 1940 to 2004 was analyzed, and non-random global patterns were demonstrated, some of which are: a) the majority of epidemics 71.8% originate in the wild (for example, severe acute respiratory virus, Ebola virus) and are significantly increasing over time, b) more than half of the cases are caused by bacteria or rickettsiae, which is reflected in the drug resistance of microbes. A global geographic effect is noted: a significant risk of epidemics occurring in low latitudes.

[9] (Sec. 3): Viral hemorrhagic fever is an acute infection with a high mortality rate. Among the possible sources are Ebola and Marburg viruses, Lassa virus and several others. They are clinically difficult to di-

agnose and distinguish; in suspicious cases, fast and reliable laboratory diagnostics are required. Six one-step real-time reverse transcription PCR assays for these pathogens based on a mixture of reverse transcriptase and Taq polymerase enzymes have been identified. New primers and / or probes have been developed. PCR products were detected in real time on LightCycler device using 5'-nuclease or SybrGreen dye intercalation technology. For this dye, the inhibitory effect is weakened and universal cycling conditions for its detection have been determined. The detection limits with a probability of at least 95% are defined by the method of regression analysis and vary from 1545 to 2835 equivalents of the viral genome / ml of serum (from 8.6 to 16 RNA copies per analysis).

[10] (Sec. 3): Described is remdesivir, an analogue of adenosine nucleotide prodrug, which exhibits broad-spectrum antiviral activity against several RNA viruses. This compound is in clinical development for the treatment of Ebola virus disease. An analysis of the kinetics of enzymes was carried out and an assessment of clinical selectivity and effectiveness in comparison with adenosine triphosphate (ATP) was given.

[11] (Sec. 3): The largest outbreak of Ebola virus disease has been reported in Guinea, Liberia, Sierra Leone and Nigeria. Sequencing of 99 Ebola genomes from 78 patients revealed a rapid accumulation of genetic variations. The considered West African variant differed from the Central African lines (2004), crossed from Guinea to Sierra Leone in May 2014 and subsequently showed sustained person-to-person transmission.

[12] (Sec. 3): The review also provides an indication of the origin of the new coronavirus in Wuhan, China. Initial genomic sequencing data for this virus does not match previously sequenced CoVs, indicating a new strain of CoV (2019-nCoV), which is currently referred to as severe acute respiratory syndrome CoV-2 (SARS-CoV-2). COVID-19 is thought to originate from a host animal with subsequent human-to-human transmission, but other routes, such as foodborne transmission, should not be ruled out. COVID-19 exhibits less severe pathogenesis, but higher transmission potential compared to diseases caused by previously known human CoV, as well as Ebola and avian influenza. Bats are thought to be the main source of the virus. Real-time PCR and next-generation sequencing have facilitated early identification of the pathogen. The review examines epidemiological, diagnostic, clinical and therapeutic aspects, including the prospects for the use of vaccines and other preventive measures that have been recommended around the world.

[13] (Sec. 4): The authors used the human antibody phage display library and the H5 hemagglutinin ectodomain to select ten neutralizing antibodies that were effective against all Group 1 influenza viruses tested, including H5N1 Avian Flu and H1N1 Spanish Flu. The crystal structure of one such antibody, coupled to H5, indicates that it blocks infection. Nine of the antibodies examined use the VH1-69 germline gene, and they all appear to use the same neutralization mechanism. Evidence suggests that antibody-based immunotherapy is a promising strategy for protection against seasonal and pandemic influenza viruses.

[14] (Sec. 4): Changes in the pathogenicity of H5N1 avian influenza viruses in healthy domestic ducks in mainland China from 1999 to 2002 have been assessed and an increase in the pathogenicity is noted. A hypothesis is presented that explains the mechanism of this evolutionary direction. Twenty-one viruses were tested, confirmed as influenza A virus subtype H5N1. All viruses are highly pathogenic for chicks. They were partially replicated.

[15] (Sec. 4): The link refers more to review business analytical publications and to a lesser extent contains evidence-based decisions and conclusions. The study aims to 1) identify risks when traveling abroad; 2) examine whether such perceived risks affect travelers' decisions during crises; 3) investigate whether such risks are mitigated in any way by the traveler's previous experience in foreign country, and 4) determine the impact of terrorism and the impact of disease, such as bird flu, on the Thai hospitality industry. Fears do not affect the abandonment of tourism in general, but to a greater extent, the change in route.

[16] (Sec. 5): The article is a review of data on the development of therapy associated with RNA interference over the 10-year period since its discovery. The annotation contains an indication of the award of the Nobel Prize in Medicine for this achievement. However, the names of the winners — Andrew Z. Fire and Craig C. Mello — were not mentioned.

7. SPECIFICITY OF PUBLICATIONS ON THE TOPIC UNDER CONSIDERATION IN JOURNAL NATURE

The journal Nature was selected based on its popularity, high rating, as well as covering various areas of knowledge and interaction with authors of different nationalities. The journal's impact factor is quite high (it is approximately equal to 28). Additional keywords: COVID, AIDS, Flu, and similar related to common diseases.

A set of a combination of keywords Journal Nature & COVID on 15.06.2021. 437 publications were identified (237 refers to 2020, 200 to 2021). All of them

are formally referred to as a multidisciplinary field of knowledge. The average citation of articles on COVID on 15.03.2021 was over 110, which is 4 times the average citation of articles in Nature.

The distribution by nationality of authors again shows a confident dominance of the United States (72 or 16.5%), followed by Great Britain (28), Germany (24), China (19), Switzerland (13). 3 publications are affiliated with authors from the Russian Federation (position 22nd).

Subsamples of Nature article publications with additional search terms besides COVID, namely flu, AIDS and oncology give 1, 2 and no publications.

At the same time, there are publications in Nature associated with several other diseases.

So 120 publications with the keyword oncology revealed in Nature. The first article dates back to 1961, then for 20 years there are single publications at intervals of 4–5 years.

Over the 10 years of 2001–2011, the number of annual publications varied from 1 to 14 (a slight increase over the general background in 2007–2008 corresponded to 11 and 14 articles).

For the period 2012–2021, the number of annual publications ranges from 1 to 6. The total number of citations is 65 062. The Hirsch index is 70, the article with the maximum citation of 7457 refers to June 2002. The number of authors affiliated with the United States is 73 (60.8%), then many times less authors from Great Britain, Canada, Italy. Authors from the Russian Federation are noted in only one article.

During the formation of the search query AIDS & Nature (journal) 1122 publications were identified. The chronological distribution up to 2001 is shown in Fig. 5. Then there is a decline in publication activity: in 2002–2011 fluctuations from 10 to 18 publications, 2012–2021 fluctuations from 2 to 13 articles. The initial period under review is 1980–1982, there are no more than one publications annually. The graph (Fig. 5) shows the most informative period of 1983–2001. Actually, according to Wikipedia, 1983 is considered the year of HIV discovery.

Fig. 5. Dynamics of the number of publications on AIDS in the journal Nature

The maximum number of publications (about 120) is reached in 3–4 years after the discovery of HIV, namely, in 1986–1987.

All of these publications are formally referred to as a multidisciplinary field of knowledge. In addition, the areas are indicated: biochemistry, genetics, molecular biology — 504, medicine — 180, immunology and microbiology — 130, etc.

Most of the authors are from the USA — 697, then Great Britain — 206, France — 77, Germany — 48, Canada — 43. A relatively small contribution of Chi-

na — 22 (12th place). Russia in terms of participation is below — 39th place. The low participation of China seems to be due to the fact that a significant development of the publication activity of Chinese scientists is associated with the 21st century, and most of the publications date back to the second half of the 1980s.

Searching for Flu & Nature (journal) allows to make the following observations: the largest number of publications falls on 2005–2006, 2009 and 2012 (36, 23, 21 and 22, respectively). After 2012, publication activity has noticeably decreased (no more than 10 articles per year).

As before, the leaders are authors from the USA (34), then many times less from Great Britain (7), the Netherlands (6), France and Germany — 5 each.

8. STATISTICS OF PUBLICATIONS, CHARACTERIZING METHODS OF DISEASE ANALYSIS

It is reasonable to use the keywords either protein or immunoassay.

The combination of COVID & protein identifies 16 197 publications as of 15.06.2021. Almost all, with the exception of 4 articles, refer to 2020 (9213) or 2021 (6980).

By fields of knowledge: medicine — 10 740 (66.3%), biochemistry, genetics, molecular biology — 4709 (29.1%), immunology and microbiology — 3213 (19.8%), pharmacology, toxicology — 1696 (10.5%), chemistry — 930 (5.7%).

By nationality of authors: USA — 4104, China — 2886, Italy — 1707, India — 1453, Great Britain — 1295, Germany — 820, France — 724, ..., Russia — 224 (20th place).

Of the Russian organizations, the majority of authors represents Sechenov Moscow Medical University — 67, RAS — 46, Moscow State University — 33, Ministry of Healthcare of the Russian Federation — 25.

When forming a query for COVID & immunoassay, 1227 publications were found (631 in 2020, 596 in 2021) In the fields of knowledge: 915 medicine, biochemistry, genetics, molecular biology 309, immunology and microbiology 291, chemistry 86, engineering and pharmacology, toxicology — 57 each.

By nationality of authors: USA — 290, China — 193, Italy — 123, Great Britain — 118, Germany — 91, Spain — 55, France — 50, ..., Russia — 11 (25th place).

2798 articles were selected for the request of Ebola & protein. Approximate dynamics of the number of publications: no more than 10 articles per year up to 1997, from 20 to 40 — in 1998–2002, from — 50 to 70 in 2003–2009, then from 2010, respectively 89,

128, 109, 87, 140, 290, 252, 231, 238, 208, 275 and 105 articles. More than 200 articles per year is the level of publication activity achieved by 2015.

By fields of knowledge: immunology and microbiology — 1240, medicine — 1050, biochemistry, genetics, molecular biology — 985, agricultural and biological sciences — 369, pharmacology, toxicology — 343.

By nationality of authors: USA — 1627, Germany — 323, Great Britain — 228, Canada — 212, China — 210, France — 195, ..., Russia — 56 (11 place).

The Russian authors are affiliated with NPO Vector — 22, RAS — 11, Siberian Branch of RAS — 9, Moscow State University and the Ministry of Healthcare of the Russian Federation — 7 each.

For comparison, the query Ebola & immunoassay highlighted 122 articles with a total number of citations of 2253. Hirsch index is 25, the highest number of citations is 166.

The approximate dynamics of the number of publications over the past few years, starting from 2013 — 1, 2, 8, 13, 11, 14, 17, 13 and 8, respectively.

By areas of knowledge: medicine — 73, immunology and microbiology — 46, biochemistry, genetics, molecular biology — 33, chemistry — 10, technology — 8.

By nationality of authors: USA — 65, UK — 13, China — 10, Switzerland — 9, Canada, Sierra Leone and Russia — 8 each.

9. FORMATION OF A QUERY INCLUDING A PAIR OF KEYWORDS OF THE SAME TYPE

In our case, we are considering concomitant COVID diagnoses (diseases).

COVID & flu query identified 1101 publications (2020 — 686, 2021 — 415).

By areas of knowledge, the distribution is quite typical: medicine — 783, biochemistry, genetics, molecular biology — 144, immunology and microbiology — 134, pharmacology, toxicology — 92, social sciences — 72.

By nationality of authors: USA — 248, Italy — 170, India — 135, Great Britain — 99, China — 69, Brazil — 67, ..., Russia — 14 (23rd place).

502 publications were selected for COVID & AIDS (257 and 241 in 2020 and 2021, respectively)

By areas of expertise: medicine — 309, social sciences — 77, biochemistry, genetics, molecular biology — 47, immunology and microbiology — 39, computer science — 32.

It can be seen that this distribution by areas of knowledge is noticeably different from the previous one.

By nationality of authors: USA — 168, UK — 57, India — 39, China — 35, Australia — 24, Italy — 21, ..., Russia — 9 (18th place).

COVID & oncology query identified 1 539 publications (1000 and 539 in 2020 and 2021, respectively).

By fields of knowledge: medicine — 1449, biochemistry, genetics, molecular biology — 448, pharmacology, toxicology — 51.

By nationality of authors: USA — 502, Italy — 275, Great Britain — 183, France — 114, India — 111, ..., China — 62 (11th place), ..., Russia — 10 (36th place).

10. FEATURES OF THE SELECTION OF PUBLICATIONS WITH TWO- AND THREE-STAGE SEQUENTIAL ASSIGNMENT OF KEYWORDS OR THEIR COMBINATIONS

Stage 1: identification of publications containing two medical diagnoses, including COVID, as well as medical diagnostic information.

COVID-associated keywords, expressions and interpretations of the corresponding selection of publications with the first restriction on the COVID keyword.

Stage 2: Further, the most significant subgroups of publications are analyzed by additional criteria (keywords):

A) SARS-CoV-2 — 42 197 mentions with 157 871 publication (on 16.06.2021) or 26.7%.

B) Viral pneumonia — 35 375 or 22.4%.

C) Indicative is the gender equality: 30 079 mentions of women and 29 144 mentions of men.

D) Risk factor — 9656 and risk factors — 5054 publications; total of 14 170 publications or 9.3%. Thus, the choice is very sensitive to the formalities of the request.

E) Psychology — 7378 mentions or 4.7%.

Further, approximately 4500 mentions of C-reactive protein and hypertension.

If, at the Stage 1, an analysis of associations with the COVID keyword is performed, considering only publications that are formally classified as medicine, then we have the following results.

The number of mentions dropped to 101 345 or 64.2% (this is a fairly stable indicator of the considered share of publications).

The total number of mentions of risk factors in different spellings was 12 840, compared with 14 170 publications in the previous subsection D. That is, approximately 12.5% of mentions of COVID risk factors do not relate directly to the medical field.

The number of mentions of C-reactive protein and hypertension dropped to about 4100 and 3900, respectively. Obviously, these keywords (expressions) refer to medical aspects. However, some publications in which these expressions are mentioned as key ones are not formally classified as medical either by the nature of the publication or by the choice of the source.

Note that, for example, psychology and medicine are considered as different areas of knowledge.

Also interesting is the at least three-variant mention of the polymerase chain reaction, namely, PCR itself — 2849 times, PCR with reverse transcription — 3409 or real-time PCR — 2543.

Stage 3: The role of additional restrictions, including specific risk factors, is analyzed.

One can assess the rating of risk factors by their mention. Perhaps this will indirectly reflect the degree of significance of these factors. Although the latter seems to be a subject of discussion.

Some of the mentioned risk factors for COVID are: hypertension — 13.2%, diabetes — 12.6%, depression — 6.5%, age factor — 6.1%, chronic obstructive pulmonary disease (COPD) — 4.3%, smoking — 3.0%.

These factors are cited in various publications, not only scientific ones, as complicating COVID.

11. DISCUSSION AND RECOMMENDATIONS

Associations with the word COVID are localized in time for the period 2019–2021. In addition, in scientometric analysis, an essential factor in the formation of a sample of publications is the order of assigning keywords.

Selection, starting with COVID, appears to be less informative when looking for concomitant diagnoses. Better to start with a second diagnosis.

In Sec.1 (Tab. 1), the conditional division of publications associated with the word COVID into 8 main areas of knowledge is considered. In this section, we consider a set of keywords that are most often found in publications in each of these areas. At the same time, the content of the publication is not analyzed. The following were subject to consideration: a) the words indicated in the corresponding sample of publications as key words, b) assessment of the correspondence of our understanding of the branch of knowledge to the keyword (expression), c) the influence of variations in spelling of keyword or expression, the same in the sense, d) basic trends related to the choice of keywords (expressions).

The first most important area of knowledge is medicine (103 044 publications on 23.06.2021).

Words and phrases most commonly associated with COVID publications and the medical field: human(-s), pandemic(-s), SARS-CoV-2, Virus Pneumonia, coronavirus (disease or infection). These English language expressions are easily translated into Russian.

Words that are a priori associated with medicine are somewhat less common: epidemiology (more than 11 700 times), hospitalization (about 8000), psychology (about 5680 references). An interesting circumstance is the previously noted almost equal presence of the keywords woman (26 406) and man (25 662). Whether this is a coincidence or a kind of tribute to tolerance (gender equality) is difficult to determine.

The next area of expertise is social sciences (20 712 publications).

The most common keywords and phrases are almost identical to the previous ones. The keywords are somewhat more common: epidemiology (over 900) and psychology (over 650). Also, almost equal mentions of women (802) and men (759). In this regard, we can talk about the commonality of medicine with the social sciences.

However, unlike the previous subgroup, keywords and phrases related to the impact of the pandemic on society are much more common. In particular, such words as education and training in various writing options: education, distance education, learning, e-learning, online learning, distance learning. Also, expressions are quite common, not only in substance, but also formally related to the social sphere, namely, socialmedia, organization and management.

Computer sciences is an expectedly specific area of expertise (9136 publications). In this sample of publications, there are the most significant differences from those considered earlier. In particular, a rather significant change in the set of keywords included in the top ten in terms of prevalence compared to the two previous samples. So on the 5th and 10th places there are deep-learning and machine-learning, respectively, 816 and 626 mentions. The keywords students (552 times) and social networks (347) are more common in social sciences. At the same time, the big data keywords, which are specific to computer science, is quite common.

Another couple of samples in two closely related areas of knowledge have fairly similar features of the distributions of keywords and expressions.

An area of expertise combining biochemistry, genetics and molecular biology (15 765 publications). The keyword genetics is not encountered very often (only in 1834 publications, which is the 30th most frequent). Here, as in some previous cases, the number of references to men and women is almost equal. The second trend is the presence of words and expressions that are consistently associated with these areas

of knowledge. For example, antiviral agent, pathogenetics, drug effect, chemistry and a number of others are often mentioned. Attention is drawn to the frequent mention of various words and expressions related to age: age, middle age, adults, youth, etc.

The final of the 5 considered areas of knowledge combines immunology and microbiology (10 240 publications). Almost complete coincidence of the first 10–15 key expressions with the previous group. Among the common key expressions, there are both a priori associated with immunology, and more associated with medicine expressions: pneumonia, immunology, virology, beta-coronavirus, genetics, metabolism, immune response.

In general, it can be argued that not always the most common keywords in publications are a priori associated with the corresponding areas of knowledge. It is possible that the assignment of a publication to a certain branch of science in its selection is determined not so much by the content of the publication as by formal factors: the name of the journal, the topic of the conference, the materials of which are considered in the publication, etc.

CONCLUSION

When analyzing publications on the topic under consideration (mainly COVID, SARS), indexed in the Scopus database, some trends were revealed:

Publications associated with the keyword COVID, as expected, are mainly attributed to the field of medicine, the share of which is 65–70%, and only for authors from the Russian Federation it is significantly lower (55%). Percentage of publications classified to social sciences for authors from most countries is about 10–12%, only for authors from China is no more than 6%; the contribution of biochemistry is 10–15%; for psychology the share is 3–5%; the share of publications on computer science and technology for the USA and Great Britain is about 3%, and for the Russian Federation and China it is about twice as large (5.5–6.5%). These trends can be explained both by the priorities of the scientific and technical development strategy of these states, and by the mentality and / or peculiarities of social policy.

Articles related to SARS have been published for more lengthy periods. The share of articles formally classified as medicine is many times less than for COVID (29% compared to 65%). The share of articles related to technology and computer science is many times higher (about 23 and 22%, respectively) than in publications on COVID (shares of 4–5%).

Publications in branches formally related to genetics, biochemistry, and molecular biology have a stable representation in both groups of publications (11 and 16%, respectively).

The growth of publication activity is constant since 1950. The main publications are by authors from the United States and the United Kingdom, and since the early 2000s, the participation of authors from China has sharply increased, especially in the period 2006–2019.

Pandemic 2020–2021 had a comparable impact on publication activity both related to the COVID keyword and related to the SARS keyword, which is reflected in a manifold increase in the number of publications on COVID and a corresponding increase in publications on SARS. During this period, the leading role of authors from the United States (more than 37%) was preserved, the second place was taken by authors from China (just under 25%). Authors from the Russian Federation have a place in the rating from 12 to 23. Mainly authors from the Russian Federation represent NPO Vector, Novosibirsk University and RAS.

The results of a statistical analysis of samples for the corresponding groups of publications related to the development of the most famous pandemics of the late 20th and early 21st centuries are presented.

For publications associated with the word Ebola, there is a short-term surge in publication activity in the period 2016–2018. Before that, a long period (20 years since 1982), the number of annual publications grew almost linearly, increasing by 6–7 articles per year. Such a clear dominance of the US authors, which was observed in almost all other samples, was not revealed here. The distribution of publication activity over time is bell-shaped with a maximum near 2006.

Another obvious trend: the geography of the spread of the pandemic is indirectly reflected in the nationality of the authors. In particular, the share of publications related to the Ebola virus for authors from Central African countries is clearly above average. A similar trend appears when assessing the share of authors from East Asia in publications related to bird flu.

In a sample of publications on topics related to bird flu and DNA, the authors from the Russian Federation have the highest rating among the groups of publications under consideration: rating 6–7 with a share of about 4.5%. If the first two places in the rating of the United States and China are quite traditional, then representatives of Asian countries previously affected by bird flu, namely Hong Kong and Singapore, come next.

The content of the most cited articles for each considered group of articles is analyzed separately. In some cases, the number of citations exceeds 10 000 in less than 2 years after publication.

The most cited articles are mainly related to the description of COVID symptoms, the history of origin, assessments of the possibilities of vaccines and therapy in general.

However, some of the most cited publications are not related to purely scientific or scientific and technological aspects, but to the analysis (description) of social aspects, in particular, the consequences of various pandemics for the tourism business.

When analyzing publications related to COVID or similar pandemics in the journal Nature, both patterns coinciding with general trends and specific features were identified. One of the common patterns is the sharp increase in AIDS publications dating back to 1986–87. We are practically observing a 3-year delay in relation to the date of AIDS discovery.

The selection of publications when setting a complex (multi-stage) search query also revealed a number of trends.

The main ones:

1. An essential role is played by the sequence of assigning keywords.

2. The choice is sensitive to the formal spelling of the keyword or expression (for example, publications are selected separately when specifying the words: human or humans).

3. In most cases, there is an a priori definite correspondence between the considered area of knowledge and keywords and expressions. For example, for medicine, a set of keywords can include diagnostic signs (shortness of breath, hypertension, etc.), for the field of computer science terms are quite expected (algorithm, machine learning, big data, etc.).

4. Some trends have no scientific basis, but are a reflection of social aspects. For example, almost equal mention of men and women, even when applied to samples of patients with an obvious dominance of men.

Analysis of the level of publication activity of the Russian scientists in the areas of knowledge under consideration indicates that the entry in the top five world states in terms of the number of international publications seems to be a difficult task. Even if we take into account one of the key principles of statistics — the illegality of the distribution of the patterns identified in a special group to a more general group of data. In this case, the scientific publications under consideration relate mainly to natural science fields of knowledge (physics, biology, chemistry, medicine, engineering, computer science, analytical methods, etc.), the corresponding assessments of the contribution of scientists from the Russian Federation to the publications of the Scopus database are all still give an idea of their contribution to scientific publications in general.

Acknowledgments

The authors express their gratitude to the teachers of advanced training courses at Peter the Great St. Petersburg Polytechnic University in the direction of scientific publication in international journal (preparation, submission, promotion), first of all, to Vatin Nikolai Ivanovich and

Yakubson Vera Mikhailovna for the advice and recommendations given by them, which made it possible to increase the efficiency of work with bibliometric databases.

The work was carried out within the framework of the State Order of the Ministry of Science and Higher Education of the Russian Federation No. 075-00280-21-00.

REFERENCES

1. Huang C., Wang Y., Li X., Ren L., Zhao J. et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *The Lancet*, 15–21 February, 2020, vol. 395, iss. 10223, pp. 497–506. DOI: 10.1016/S0140-6736(20)30183-5
2. Wölfel R., Corman V.M., Guggemos W., Seilmaier M. et al. Virological assessment of hospitalized patients with COVID-2019. *Nature*, 28 May, 2020, vol. 581, iss. 7809, pp. 465–469. DOI: 10.1038/s41586-020-2196-x
3. To K.K.-W., Tsang O.T.-Y., Leung W.-Sh. et al. Temporal profiles of viral load in posterior oropharyngeal saliva samples and serum antibody responses during infection by SARS-CoV-2: an observational cohort study. *The Lancet Infectious Diseases*, 2020, vol. 20, iss. 5, pp. 565–574. DOI: 10.1016/s1473-3099(20)30196-1
4. Gordon D.E., Jang G.M., Bouhaddou M. et al. A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature*, 2020, vol. 583, pp. 459–468. DOI: 10.1038/s41586-020-2286-9
5. Broughton J.P., Deng X., Yu G. et al. CRISPR–Cas12-based detection of SARS-CoV-2. *Nat Biotechnol.*, 2020, vol. 38, pp. 870–874. DOI: 10.1038/s41587-020-0513-4
6. Lu R., Zhao X., Li J., Niu P. et al. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *The Lancet*, 22–28 February, 2020, vol. 395, iss. 10224, pp. 565–574. DOI: 10.1016/S0140-6736(20)30251-8
7. Kim D., Lee J.-Y., Yang J.-S., Kim J.W., Kim V.N., Chang H. The Architecture of SARS-CoV-2 Transcriptome. *Cell*, May 14, 2020, vol. 181, iss. 4, pp. 914–921. DOI: 10.1016/j.cell.2020.04.011
8. Jones K.E., Patel N.G., Levy M.A., Storeygard A., Balk D., Gittleman J.L., Daszak P. Global trends in emerging infectious diseases. *Nature*, February, 2008, vol. 451, iss. 7181, pp. 990–993. DOI: 10.1038/nature06536
9. Drosten C., Götting S., Schilling S., Asper M., Panning M., Schmitz H., Günther S. Rapid detection and quantification of RNA of Ebola and Marburg viruses, Lassa virus, Crimean-Congo hemorrhagic fever virus, Rift Valley fever virus, dengue virus, and yellow fever virus by real-time reverse transcription-PCR. *J. Clin. Microbiol.*, 2002, vol. 40, iss. 7, pp. 2323–2330. DOI: 10.1128/JCM.40.7.2323-2330.2002
10. Tchesnokov E.P., Feng J.Y., Porter D.P., Götting M. Mechanism of inhibition of ebola virus RNA-dependent RNA polymerase by remdesivir. *Viruses*, 2019, vol. 11, iss. 4, 326. DOI: 10.3390/v11040326
11. Gire S.K., Goba A., Andersen K.G., Sealfon R.S.G., Park D.J., Kanneh L., Jalloh S., Momoh M., Fullah M., Dudas G., Wohl S., Moses L.M. Genomic surveillance elucidates Ebola virus origin and transmission during the 2014 outbreak. *Science*, 12 September, 2014, vol. 345, iss. 6202, pp. 1369–1372. DOI: 10.1126/science.1259657
12. Dhama K., Khan S., Tiwari R., Sircar S., Bhat S., Malik Y.S., Singh K.P., Chaicumpa W., Bonilla-Aldana D.K., Rodriguez-Morales A.J. Coronavirus disease 2019–COVID-19. *Clin. Microbiol. Rev.*, October, 2020, vol. 33, iss. 4, pp. 1–48. DOI: 10.20944/preprints202003.0001.v2
13. Sui J., Hwang W.C., Perez S., Wei G. et al. Structural and functional bases for broad-spectrum neutralization of avian and human influenza A viruses. *Nature Structural and Molecular Biology*, March, 2009, vol. 16, iss. 3, pp. 265–273. DOI: 10.1038/nsmb.1566
14. Chen H., Deng G., Li Z., Tian G., Li Y., Jiao P., Zhang L., Liu Z., Webster R.G., Yu K. The evolution of H5N1 influenza viruses in ducks in southern China. *Proceedings of the National Academy of Sciences of the United States of America*, 13 July, 2004, vol. 101, iss. 28, pp. 10452–10457. DOI: 10.1073/pnas.0403212101
15. Rittichainuwat B.N., Chakraborty G. Perceived travel risks regarding terrorism and disease: The case of Thailand. *Tourism Management*, June, 2009, vol. 30, iss. 3, pp. 410–418. DOI: 10.1016/j.tourman.2008.08.001
16. Whitehead K.A., Langer R., Anderson D.G. Knocking down barriers: Advances in siRNA delivery. *Nature Reviews Drug Discovery*, 2009, vol. 8, pp. 129–138. DOI: 10.1038/nrd2742

Contacts: *Bulyanitsa Anton Leonidovich*,
antbulyan@yandex.ru

Article received by the editorial office on 20.07.2021