
Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

World Patent Information 68 (2022) 102094

Contents lists available at ScienceDirect

World Patent Information

journal homepage: www.elsevier.com/locate/worpatin

# Searching and Analyzing Patent-relevant COVID-19 Information

Luca Falciola [a,*], Massimo Barbieri [b]

[a] Scibilis SRL, Ixelles, Belgium
[b] Technology Transfer Office, Politecnico di Milano, Milan, Italy

ARTICLE INFO

*Keywords:*
SARS-CoV-2
Databases
Patent classifications
Open access
Scientific publishing
Patent searching

ABSTRACT

The COVID-19 pandemic has prompted several institutions to offer free, dedicated websites and tools to foster research and access to urgently needed innovative solutions by facilitating the search and analysis of information within the large amount of scientific and patent literature which was published since January 2020. This situation is clearly exceptional and challenging for patent information users searching for relevant disclosures at a given date in a reliable manner. This article provides an overview of search criteria and strategies, main databases and websites, number of publications, biological sequence information and experimental data sets covering COVID-19 findings within scientific and patent literature have been disclosed between January and August 2020. The analysis of non-patent literature has been focused on the identification, date assignment, disambiguation, and access to experimental data. The analysis of patent literature has been focused on the trends found within the earliest filed and published patent documents in representative jurisdictions worldwide. Some practical advice and strategies for technical, medical, or patentability assessment of COVID-19-related innovations across different information formats and resources are proposed.

## 1. Introduction

Beyond the human tragedy, the COVID-19 emergency has required most communities and economic sectors worldwide to pursue their activities in an unprecedented and unpredictable environment. Compared to infectious diseases during the last few decades, the COVID-19 pandemic combines many of their features (such as geographic distribution and transmissibility) but has an impact amplified by imbalances and weaknesses of global ecosystems [1]. The public and private research institutions are called to very quickly adapt their governance and strategies to support innovation, strengthening collaboration efforts to overcome this "Research and Development gap" [2] in response to the pandemic. The COVID-19 crisis is also impacting R&D budgets and priorities for financing innovation, pushing for major and structural changes in both high technology and traditional sectors, affecting work life and everyday activities.

Patenting activities contribute to the dissemination of novel technical solutions and products, and of innovation in general, and COVID-19 emergency affected activities of patent offices and the overall "patent industry" as well. This extraordinary situation has triggered some changes to the legislative and regulatory frameworks, as summarized by World Intellectual Property Organization (WIPO) in a dedicated webpage [3] or by the World Trade Organization (WTO) in a working paper about the national patent-related policies in WTO Member states during COVID-19 pandemic [4]. Patent offices, courts, governments and health authorities may decide to enact specific country-specific policies applicable to patent rights, in particular by taking into account both financial and human rights issues related to the pandemic and its consequences. Thus, there are many propopsals to accept far more exceptions to legal and commercial policies than those generally applied, with at least partially conflicting attitudes towards open access to methods, drugs, and equipment [5,6].

Since the official declaration of the COVID-19 global outbreak, many publishers and institutions have made several tools available for extracting and aggregating potentially relevant information from "classical" sources (namely peer-reviewed articles and patent literature) or those less formally defined and structured (unreviewed publications, experimental and clinical data, blogs, information from social

*Abbreviations:* IPC, International Patent Classification; CPC, Cooperative Patent Classification; WIPO, World Intellectual Property Organization; EPO, European Patent Office; SARS-CoV-2, Severe Acute Respiratory Syndrome Coronavirus 2; COVID-19, Coronavirus Disease 2019; NCBI, National Center for Biotechnology Information; EBI, European Bioinformatics Institute; EMBL, European Molecular Biology Laboratory; MeSH, Medical Subject Headings.

* Corresponding author.
*E-mail addresses:* LFalciola@scibilis.be (L. Falciola), massimo.barbieri@polimi.it (M. Barbieri).

https://doi.org/10.1016/j.wpi.2022.102094
Received 4 December 2020; Received in revised form 16 January 2022; Accepted 16 January 2022
Available online 20 January 2022


**Fig. 1.** COVID-19 conspiracy theory on the street. The image was taken by the corresponding author in a street near the Université Libre de Bruxelles, Brussels (Belgium) in early April 2020.

networks). Such a situation is made possible by newly available technologies and is hardly comparable to past viral outbreaks (for instance, those related to Ebola virus), in terms of dimension and speed. Direct access and continuous stream of publications of uneven quality and consistency have far-reaching consequences on the attitude of public opinion. For instance, some French websites indicated a European patent filed in 2004 (!), referring to a coronavirus strain that was isolated in Hanoi (Vietnam), as the "Coronavirus Patent" (Fig. 1).

This article intends to provide patent information professionals with an overview of criteria, data, and sources useful to COVID-19 searches as made available until November 2020. Moreover, an analysis of major sources of patent and scientific information (including biological sequences and datasets for drug and diagnostic development) was performed for the period between January and August 2020, summarizing main qualitative and quantitative features of COVID-19 information that has been disclosed in this period. Main features of the earliest published patent publications that refer explicitly to COVID-19 were reviewed, identifying some key issues for establishing and performing search strategies in this domain.

## 2. Methodology

The databases and webpages cited from the websites of the aforementioned scientific publishers, patent offices, or other institutions were initially accessed, searched and reviewed for this article during July and August 2020. The final data reported in tables and figures covering publications in the period between January and August 2020 were generated using the indicated search strategies and databases between September and October 2020. All the website addresses, references, datasets, databases, and contents herein are cited as available on Jan. 11th, 2022.

The details of webpage names and address for searching non-patent literature, clinical information, and experimental data cited in Section 3 are listed in Supplementary File 1. The scientific literature that was published during the period Jan.–Aug. 2020 and explicitly citing COVID-19, or relevant synonyms, in the title, abstract, and/or indexing was extracted from PubMed and The Lens Scholarly Works using the criteria listed in Table 1, according to their specific search functions and language. The results of searches were downloaded in csv or xlsx format, manually reviewed, compared to each other and to the published COVID-19 datasets (as indicated in Section 3), normalized for format, and consolidated using spreadsheets that list the main bibliographic and indexing details. These files have been used to generate Table 2 and the figures that were elaborated for Section 3.

The details of providers, webpage names and Internet address for searching patent literature cited in Section 4, are listed in Table 3. The patent documents that explicitly cite COVID-19, or relevant synonyms, in the title, abstract, and/or claims and that were filed and published during the period Jan.–Aug. 2020 were searched using the Group (1) and Group (2) criteria listed in Table 1 in the following databases and websites, according to their specific search functions and language: Derwent Innovation, Patentscope, DepatisNet, Lens Patents, Espacenet

| Provider | Disclosing Covid-19 search criteria | Covering only data published in 2020 |
|---|---|---|
| Ovid | YES | NO |
| EuropePMC | YES | NO |
| Semantic Scholar | NO | NO |
| Mendeley | NO | NO |
| WHO | NO | YES |
| Outbreak.info | NO | NO |
| ScienceOPEN | YES | NO |
| NCBI (PMC) | YES | NO |
| NCBI (PubMed) | YES | YES |
| NCBI (LitCovid) | NO | YES |
| L·OVE | YES | NO |
| Elsevier (Scopus) | NO | NO |
| Publons | NO | YES |
| Cochrane Library | NO | YES |

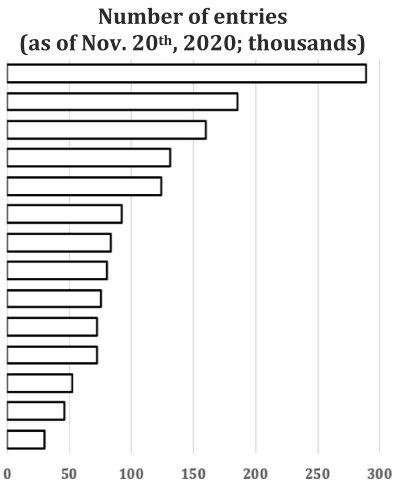



**Fig. 2.** Websites providing dedicated databases for searching for COVID-19 scientific literature and data. Further details for these and other cited databases are available in Supplementary File 1.

**Table 1**
Category of search criteria for COVID-19 topics.

| Group | Combination of relevant keywords |
|---|---|
| *Group (1)* | "covid-19″ OR "covid 19″ OR "covid19″<br>"SARS-CoV-2″ OR "SARS-CoV2" OR "sarscov2″<br>"2019 ncov" OR ″2019-nCoV" OR ″2019nCoV"<br>"covid-2019″ OR "covid 2019″ OR "COVID2019″<br>"severe acute respiratory syndrome coronavirus 2″<br>"2019 novel coronavirus" OR "coronavirus disease 2019″<br>"novel corona virus" OR "novel coronavirus" OR "new corona virus" OR "new coronavirus"<br>"Wuhan coronavirus" |
| *Group (2)* | *"Alternative phrasing"*<br>"SARS-Coronavirus-2″ OR "severe acute respiratory syndrome 2″ OR "severe acute respiratory coronavirus 2″ OR "coronavirus disease-19" OR "coronavirus 19" OR "coronavirus 2019" OR "coronavirus disease-2019" OR "coronavirus-2″ OR "coronavirus 2″ OR coronavirus2<br>*"Location phrasing"*<br>(wuhan OR hubei OR huanan) AND ("severe acute respiratory" OR outbreak OR betacoronavir* OR coronavir* OR virus OR "Middle east respiratory" OR SARS OR MERS)<br>*"Alternative acronyms"*<br>"novel cov" OR "ncov-2019" OR "ncov-19" OR "ncov" OR "cov 2″ OR CoV2 OR "MERS-COV" OR SARS2 OR "sars-cov" OR Covid<br>*"General Coronavirus"*<br>Betacoronavir* OR coronavir* OR "Corona virus" OR "corona viruses" |

and EP Full text Search (both available in the EPO website). The search results were downloaded in csv or xlsx format, manually reviewed, and compared with the results of similar searches within the registers and databases made available by national patent offices (namely in South Korea, USA, India, Italy, United Kingdom, Australia, Israel, Singapore, Russia, and Spain). The format of the patent data was normalized and consolidated using a spreadsheet into a single xlsx file (the "Early COVID-19 Patent Dataset"; Supplementary File 2) that contains the references to 1130 patent documents, including their publication numbers, type of document, title (in English and/or in the original language), details on the patent classification as IPC Groups (if available on Nov. 2020), relevant dates by month (earliest priority, publication, and grant), range in the number of days between earliest priority date and publication and/or grant date (in 15-day intervals), and technological field (as identified on the basis of the definition of indicated IPC Subclasses, IPC Groups and/or title). This xlsx file was used to generate Tables 4–6 and the figures that were elaborated for Section 4).

The details of biological sequences and databases cited in Section 5 are listed in Supplementary File 5.

Additional details are provided in the text commenting on the tables and figures.

## 3. Profiling COVID-19 scientific literature and data as published in Jan.–Aug. 2020

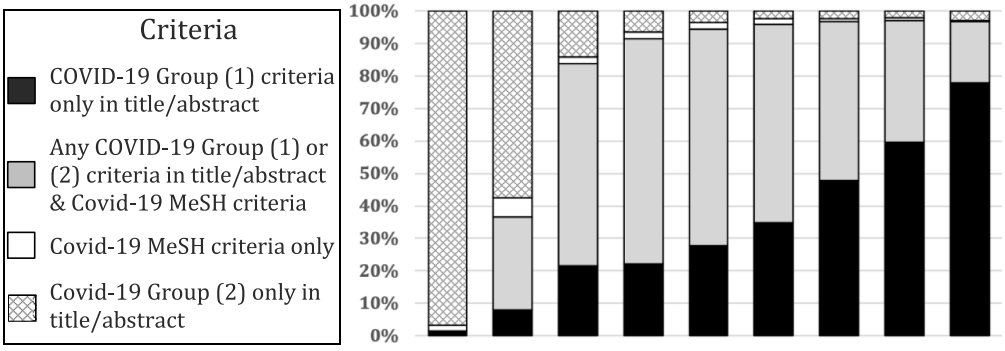
### *3.1. COVID-19 scientific information: providers, dedicated websites and nomenclature*

A review of the main webpages or portals dedicated to the COVID-19 scientific literature has been performed, consolidating the URLs in Supplementary File 1. The contents of such databases have been compared (Fig. 2) to highlight those providing access to the largest repository of documentation that is selected as being relevant for COVID-19. These documents are mostly regular scientific articles and books but, in many cases, other types of disclosures are also included: preprints, news, reports on clinical activities, experimental datasets, and other electronic-only disclosures. These absolute figures regarding the size of these databases should not be used to rank such resources accordingly, as many other criteria should also be taken into consideration including: the coverage (i.e., any coronavirus research or data published even before 2020, or focused only on COVID-19 findings in 2020); the list of keywords and other selection criteria that have been used to establish the database; the available features for searching for and extracting either specific information or full text documents; the frequency of update within the database; the reliability in checking for duplicates, format, or completeness of contents.

A major challenge for the patent examination and evaluation may come from the many, possibly conflicting sources about the content and date of disclosure of COVID-19 information. As a preliminary step, patent information specialists should verify which scientific keywords are the most appropriate to search for COVID-19 disclosures in scientific or patent literature. Similarly to previous situations where a novel entity of major scientific importance is identified and, therefore, competition for establishing some kind of primacy or leadership starts among investigators, the process of naming COVID-19 biological and medical concepts to be consistently and univocally used in literature and databases (and thus also to be used for future searches) has not been linear or fully standardized. Until end of 2019, six coronaviruses were known to infect humans: HCoV-229E, HCoV-OC43, HCoV-NL63, HCoV-HKU1, severe acute respiratory syndrome (SARS-CoV-1), and Middle East Respiratory syndrome (MERS-CoV). As summarized in the World Health Organization (WHO) website [7], the infectious agent responsible of the novel pneumonia detected in Wuhan, China, during December 2019 was described as a new coronavirus (CoV or nCoV; order Nidovirales, Family Coronaviridae, Subfamily Coronavirinae, and Genus Betacoronavirus) and thus initially named as "2019 novel Coronavirus", "2019-nCoV" or "Wuhan coronavirus". Only in early February 2020 this novel Coronavirus was defined as SARS-CoV-2, meaning "severe acute respiratory syndrome coronavirus 2″, to make this virus distinguishable from "SARS-CoV-1 which was isolated as the agent responsible for outbreak with less dramatic consequences in 2002–2004 and presents some sequence similarity. It was also agreed that the illness caused by SARS-CoV-2 and overall symptoms of the SARS-CoV-2 infection should be called COVID-19 (meaning "Coronavirus Disease 2019").

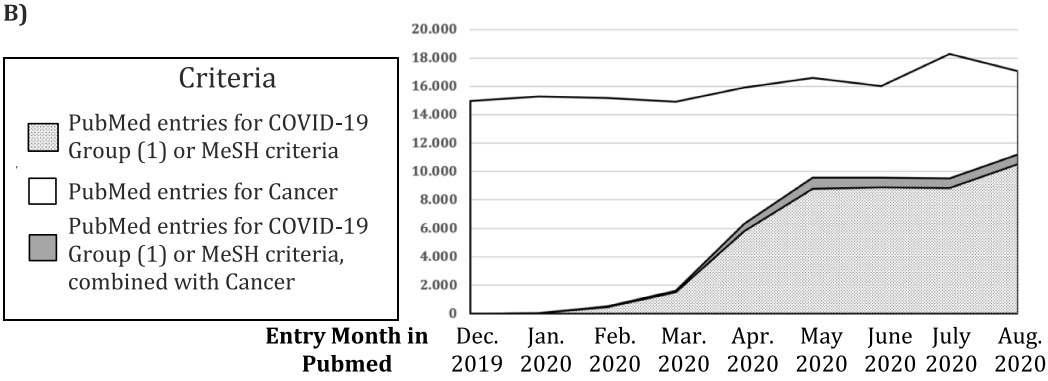
Even a quick search in the scientific literature shows that not only "typos" of the official names or acronyms of the virus and the disease (for instance COVID-2019 or SARS-CoV2) but also more generic names found in newspapers, websites, or social networks such as SARS, Coronavirus, Corona, or COVID are quite common, introducing potential ambiguity with respect to previously identified Coronaviruses (such as

A)




| Entry Month in Pubmed | Dec. 2019 | Jan. 2020 | Feb. 2020 | Mar. 2020 | Apr. 2020 | May 2020 | June 2020 | July 2020 | Aug. 2020 |
|---|---|---|---|---|---|---|---|---|---|
| **Covid-19 Group (1) or MeSH-related Pubmed entries** | 2 | 43 | 504 | 1.593 | 6.330 | 9.562 | 9.589 | 9.499 | 11.206 |
| **Malaria-related Pubmed entries** | 357 | 310 | 329 | 312 | 363 | 360 | 351 | 431 | 385 |
| **Total Pubmed entries (k=thousands)** | 101k | 104k | 104k | 104k | 116k | 121k | 117k | 134k | 131k |

B)




C)

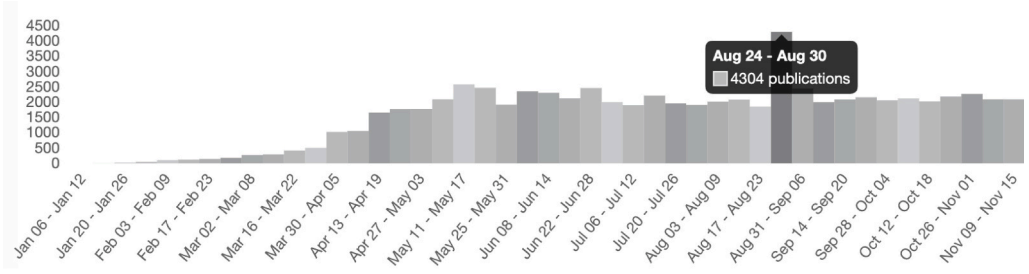


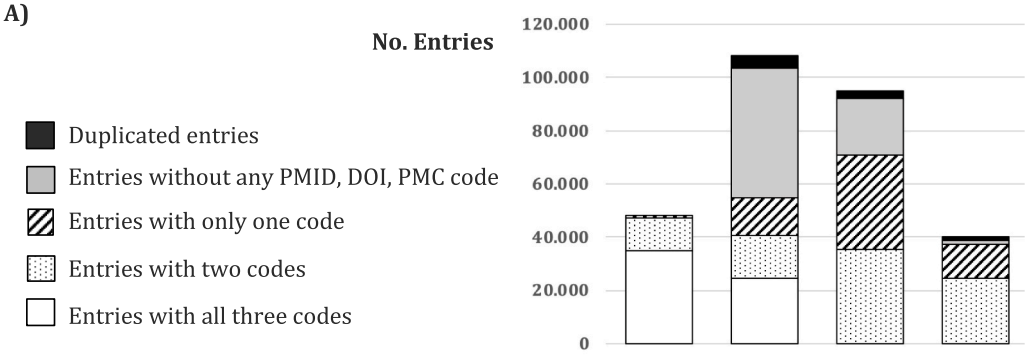

**Fig. 3.** Trend in COVID-19 scientific publications present in PubMed in Jan.–Aug. 2020. Percentage and absolute values of COVID-19 Pubmed entries compared with either total Pubmed entries or PubMed entries for malaria (A). Comparison between cancer-related and COVID-19 Pubmed entries in the same period (B). Weekly entries in PubMed based database LitCovid, as available on Nov. 2020 (C).

SARS, MERS, or SARS-CoV-1). Further sources of confusion may affect the recall and precision features of a search in generic databases or when using full-text search engines are the names of companies similar to COVID or SARS, and some acronyms used in scientific literature. Examples of such potentially confusing acronyms are NCO-based names referring to Cyanate (NCO) chemical group or CoV-based acronyms referring to Coefficient of Variation (NCoV may be used to indicate "Naïve Coefficient of Variation").

This problem was recognized by databases providers, starting from the alternative names that the National Center for Biotechnology Information (NCBI) has listed for official scientific indexation (Severe Acute Respiratory Syndrome Coronavirus 2, SARS-CoV-2 [8]and COVID-19 [9]). Some of the websites listed in Supplementary File 1 provide users with details of the criteria that have been used for creating literature listings. Main examples are the listing of websites and databases updated regularly by the EPPI (Evidence for Policy and Practice Information)-Center [10], or the list of keywords that were used by the Center for Disease Control and Prevention (CDC) to establish a database for COVID-19 literature covering several sources in October 2020 [11] and later integrated in the WHO global COVID-19 literature database [12].

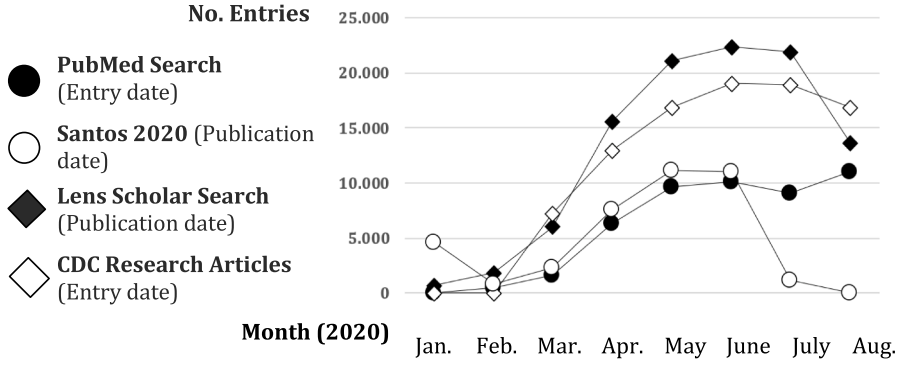
The analysis of search criteria in the pages cited above together with

A)




| Dataset (period Jan-Aug. 2020) | PubMed Search (COVID-19 Group (1) or MeSH criteria) | Lens Scholar Search | CDC Research Articles | Santos 2020 |
|---|---|---|---|---|
| % Unique entries having PMID code | 100% | 41% | 37% | 64% |
| % Unique entries having DOI code | 98% | 51% | 76% | 95% |
| % Unique entries having PMC code | 73% | 25% | 4% | 0% |

B)




C)

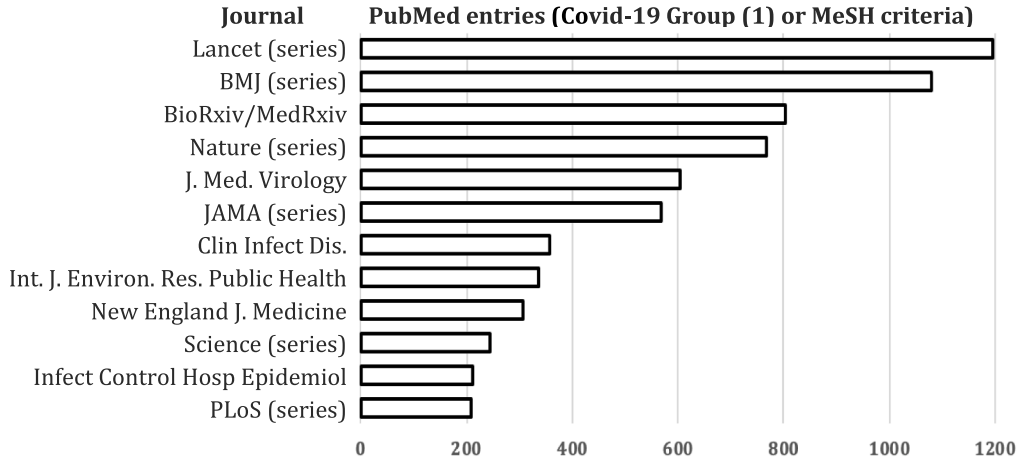



**Fig. 4.** Trend in COVID-19 scientific publications in PubMed compared to other scientific literature databases and repositories. The content of four datasets of scientific publications, either determined using search criteria (in PubMed or Lens Scholar) or as made available to the public (by the CDC website or in an article identified as Santos 2020) were compared for the period Jan.–Aug. 2020. The content of the datasets, as an absolute number of entries or their indexing (using PMID, DOI, or PMC identifying codes) was compared (A). No publications having the three codes have been found in both CDC Research Articles (since the DOI is indicated in combination only with one other code) or in Santos 2020 (since PMC is not indicated). The content is also compared with respect to the indicated months and time criteria (B). No content appeared both in Santos 2020 in Aug. 2020 (since this month was not covered in the article) and for CDC Research Articles in Jan. 2020 and Feb. 2020 (since such file was started in March 2020 and the publication date is not provided for any entry). The journals (or series of journals) more often present in the COVID-19 PubMed dataset generated with the COVID-19 Group (1) or MeSH criteria for the period Jan.–Aug. 2020 are indicated with the corresponding number of PubMed entries (C).

the browsing of search results looking for sources of potential name variants (alternative spellings, acronyms, or typos), has been used to establish a list of COVID-19 keywords that were categorized into two main Groups (Table 1). The Group (1) criteria includes the most relevant criteria that would present the lowest possibilities of confusion, including most frequent, unambiguous words and acronyms. The Group (2) criteria includes other four main types of name variants:

- Alternative phrases referring to "novel coronavirus" and other elements from Group (1);
- Phrases or keyword combinations referring to coronavirus identified in Wuhan area;
- Short acronyms related to SARS or CoV;
- Coronavirus in general.

However, it is important to point out that the different criteria categorized in Group (2) criteria, as well as new ones that may appear in the next future (defining new variants, for example), would still be relevant for elaborating on search strategies for performing the most complete COVID-19 prior art search, even if documents relating to Coronaviruses other than SARS-CoV-2, or even completely different topics, may be found with them. The search results obtained using Group (2) criteria may require a further review for evaluating their actual relevance or may be filtered by combining these results with other criteria such as a publication date later than Dec. 2019 and/or specific names of authors, institutions, compounds, or medical topics related to COVID-19.

### 3.2. Analysis of COVID-19 scientific publications in PubMed

A detailed bibliometric analysis of the COVID-19 scientific literature is not an objective of this article. Some articles have been published throughout 2020 in which detailed analyses of scientific production have been pursued using various criteria and visualization approaches for selecting databases and sources, selecting and combining search criteria, categorizing and linking publications, focusing on the production of specific countries, institutions or authors [13–15]. A good starting point to evaluate which publication or indexing details are relevant when examining a publication before or after a given date is the PubMed database in the NCBI website, the most popular free database for biomedical non-patent literature that indexes publications with Medical Sub-Heading (MeSH) system and that has been updated in 2020 with a new layout and new "PubMed 2.0" features [16]. The COVID-19 pandemic has prompted not only the introduction of two distinct Supplementary Concepts in PubMed MeSH system ("severe acute respiratory syndrome coronavirus 2″ and "COVID-19" [17]) but also the changing of their own policies for including preprints [18] for the rapid identification and assessment of scientific publications.

A previous paper [19] compared the number of publications until mid-June 2020that are obtained by combining different types and formats of keywords. Our analysis has taken a different approach by considering both the global number of entries in PubMed and the indexing speed in PubMed during a nine-month period from Dec. 2019 (when COVID-19 literature was absent) until Aug. 2020 (Fig. 3). The results were generated by comparing the Group (1) criteria with other criteria described above but referring specifically to the entry month, i.e. the month when the article is indexed in PubMed, independently from the official publication date. This approach, focusing on the actual availability and searchability of the scientific literature in PubMed, shows that less than 80 entries are found in Dec. 2019 (and less than 900 entries for the whole year 2019) by searching PubMed for "Coronavirus", while in 2020 several thousands of publications relating to COVID-19 Group (1) criteria have been entered each month, at least since April 2020. These data are compared not only with those obtained by including COVID-19 Group (2) criteria but also with overall PubMed entries, indicating that this rapid increase has made COVID-19 the topic of almost 9% of all entries in PubMed in Aug. 2020. This number of publications largely outweighs the newest publications about other important medical topics, such as malaria which is still a major health problem in some parts of the world. Even if only a small number of publications is identified by using the COVID-19 Group (2) criteria, an important percentage of the publications presenting the COVID-19 Group (1) criteria in the title or abstract are not indexed with either specific MeSH Supplementary Concepts, at least as of Nov. 2020. When elaborating and performing searches in PubMed. It is consequently important to take into account this delayed and possibly incomplete indexing process within PubMed, which may be pursued retrospectively over several months after an article is formally added into the database.

The importance of integrating at least main name variants when drafting COVID-19 search strategies in PubMed (and possibly in other scientific or patent databases) is suggested by other qualitative and quantitative observations upon COVID-19 searches in PubMed summarized in Fig. 3. For instance, since Apr. 2020 the number of publications found in PubMed with COVID-19 Group (1) criteria reaches around 70% of those indexed for a broad health subject such as cancer. An important fraction of publications is indexed for both cancer and COVID-19 Group (1), indicating the interest in evaluating the impact of COVID-19 on the medical management of other pathologies, such as cancer [20], and which may become of interest also for patenting activities. A further topic of interest is also how the increase in COVID-19 during Aug. 2020 may actually be a consequence of a specific review process, with previously published documents that have been most recently added (or newly indexed) in PubMed during a short period of time. LitCovid, the NCBI webpage proposing a subset of PubMed entries intended to be specific for COVID-19 [21], shows weekly data confirming that approximately 2000 COVID-19 publications are regularly entered in PubMed every week since May 2020 until Nov. 2020. Other discrepancies were observed with respect to one of the first articles about COVID-19 pandemic, highlighting the potential for international spread via commercial air travel of a *"pneumonia of unknown etiology*", indicated in PubMed as having its earliest date on Jan. 17th, 2020 [22] when the journal website indicates the publication on Jan. 14th, 2020 [23]. These observations confirm how it is important to carefully verify the official publication date of search results in PubMed, given the variable discrepancy or delay (from a few days up to several months) between such date and the other date information for a given article in this database, potentially impacting both actual searchability of articles within PubMed and the conclusions about applicable prior art for

**Table 2**
Other sources including publications not or poorly covered in PubMed.

| Categories | CDC Research Articles | Lens Scholar | Santos 2020 |
|---|---|---|---|
| ArXiv, SSRN, and other Open Access or preprint providers (other than BiorXiv and MedrXiv) | ✓ | ✓ | ✓ |
| References with obsolete PMID, incomplete details on the origin of publication or identification codes | ✓ | ✓ | ✓ |
| Homeland Security Digital Library | ✓ | ✓ | |
| WHO publications | ✓ | ✓ | |
| National & intl. journals in law, politics, economics | ✓ | ✓ | |
| National journals in biomedical sciences (in non-English language, mainly) | ✓ | ✓ | |
| References to Clinical Trials | ✓ | | |

**Table 3**
Webpages dedicated to COVID-19 patent information.

| Provider | Webpage name | Main features<br>Web Address |
|---|---|---|
| **Patentscope (WIPO)** | COVID-19 INDEX | Webpage listing almost 100 search strategies for 10 technological areas, including IPC codes, with or without Coronavirus-related keywords (launched on Apr. 21st, 2020)<br>https://patentscope.wipo.int/search/en/covid19.jsf |
| **EPO** | Fighting coronavirus | Webpage listing almost 200 search strategies in Espacenet format grouped in four broad themes and multiple sub-categories, with xlsx files and links, including CPC codes, IPC codes, and/or Coronavirus-related keywords (launched on June 22nd, 2020, last updated in Jun. 2021)<br>https://www.epo.org/news-events/in-focus/fighting-coronavirus.html |
| **Lens** | COVID-19 Datasets: Patents | Search strategies and datasets covering 16 technologies, including CPC codes and/or Coronavirus-related keywords (launched on Jan. 28th, 2020, last updated on May 19th, 2020)<br>https://about.lens.org/covid-19/ |
| **China Patent Information Center (CNPAT)** | Information platform for patents on COVID-19 | Two separate datasets of Chinese and non-Chinese, "world patent documents", divided into 9 categories and several categories, linked to reports analyzing the data (in Chinese only, established on April 29th, 2020; without explicit description of search strategy to define them) https://ncp.patentstar.cn/en/Home/SpecialDB |
| **IPAustralia** | Information on COVID-19 technology | Patent information on the latest COVID-19 technology is organized in six sections with graphical analysis and related search strategies (for Derwent Innovation and EPOQUE syntax) that are provided in a separate xls file https://www.ipaustralia.gov.au/tools-resources/publications-reports/information-covid-19-technology |
| **South Korean Intellectual Property Office (KIPO)** | COVID-19 Patent Information Navigation | Five separate datasets of patent documents (no explicit description of search strategies) that are divided into several categories (established on Mar 24th, 2020)<br>https://www.kipo.go.kr/kpo/BoardApp/ApplInfoAppE |
| **Medicine Patents Pool** | Databases for patents covering vaccines and medicines, searchable for the patented medicines being tested for COVID-19, and organized by country | |
| | VaxPal | https://medicinespatentpool.org/what-we-do/vaxpal |
| | MedsPal, | https://www.medspal.org/?keywords=covid&page=1 |
| **Spanish Patent & Trademark Office (SPTO)** | Since March 2020 website offers two distinct tools dedicated to new patent publications related to COVID-19 in the form of a quarterly bulletin (PDF file, in Spanish only) and weekly updated technological alerts (as html, xls or PDF file, also in English), without explicit description of search strategies | |
| | Coronavirus: Diagnosis and therapy (alerts) | http://www.oepm.es/en/informacion_tecnologica/informacion_gratuita/Alertas_Tecnologicas/detalle.html?id=68400&n=CORONAVIRUS:& |
| | Coronavirus: Diagnosis and therapy (bulletin) | https://www.oepm.es/en/informacion_tecnologica/informacion_gratuita/boletines_de_vigilancia_tecnologica/boletines_oepm/coronavirus/index.html |
| **Italian Patent & Trademark Office (UIBM)** | Fighting coronavirus | List of patent applications filed in Italy between 2009 and 2019 disclosing products and technologies applicable to COVID-19, using categories and selection method adapted from WIPO Patentscope COVID-19 Index (in Italian only) https://www.uibm.gov.it/biotech/covid-19.html |

evaluating patentability.

### 3.3. Analysis of COVID-19 scientific publications in other literature databases

The relevance of observations made above for COVID-19 publications in PubMed, can be compared with the features of datasets available from other providers of COVID-19 scientific literature. Three other COVID-19 scientific literature datasets were used for comparison: the scientific publications that can be extracted using the COVID-19 Group (1) criteria from the databases of scientific publications in The Lens Scholarly Works website [24], the COVID-19 database [25] that was established by CDC in March 2020 until October 2020 (when it was integrated within the corresponding resource in the WHO website), and the dataset published together with a publication about COVID-19 scientific publications [26]. All three providers list PubMed as a source of the literature but it is interesting to evaluate how much they actually overlap with PubMed (and with each other) and how many and which type of references may be additionally found in these datasets which integrate other sources. However, this analysis for aggregating or disambiguating references in the literature is extremely difficult given the different format of authors' names, titles and/or bibliographic details among databases. The references common to two or more sources can be precisely defined on the basis of identification codes such as the PubMed ID (PMID), the PubMed Central ID (PMC), or the Digital Object Identifier (DOI) only. If the identification codes for references extracted from PubMed were fairly complete and reliable, the situation is unfortunately quite different for documents that are selected in these other sources as being relevant for COVID-19, mainly because a large fraction of references in these databases do not list, or list in incorrect format, such identification codes, leading also to potentially duplicated entries.

Still taking into account such limitations, it is possible to make some quantitative and qualitative evaluation of COVID-19 publications made available through these four sources using COVID-19 Group (1) criteria for extracting them (Fig. 4). At the total number of references, The Lens Scholarly Works and the CDC datasets provide at least double the references present in PubMed (even if there are some duplication and uncertain identification issues). The trends over the first eight months of 2020 show how the datasets in the article and in The Lens Scholarly Works, which were meant to cover the period until Aug. 2020, do not include all the latest COVID-19 publications already included in PubMed. A further analysis can be carried out for the origin of publications indexed in different databases. As expected, the majority of publication sources extracted from PubMed, and almost entirely included also in the other three datasets, are leading publishers of biomedical and scientific literature, preprint services (bioRxiv/medRxiv) [27], or open access publishers such as PLoS [28].

Together with some specific publications in scientific domains such as epidemiology, virology, or environment, this selection of journals comprises approximately 15% of all COVID-19 publications that were indexed in PubMed during Jan.–Aug. 2020. A summary of additional sources found in the scientific literature datasets other than PubMed was made (Table 2). Aside from undetermined references, the datasets obtained from CDC and Lens.org include publications from a wide range of sources of potential interest depending on the specific objective of the search: additional preprints or open-access publication platforms, journals having more limited distribution (and often published in a non-English language), reports and communications from institutions, journals in domains other than purely biomedical ones (generally covered by PubMed in a limited manner), or references to clinical trials.

This analysis of scientific publications by means of the proposed COVID-19 search criteria and the databases selection is far from being complete but still provides some insights and issues relating to the search and analysis of COVID-19 scientific literature in general and for patent evaluation. The extension of the search in two or more databases of scientific literature is even more fundamental than usual in order to appropriately cover literature potentially relevant for evaluating the patentability requirements of inventions claimed as useful for fighting COVID-19, but it requires being careful to check a number of details when searching such databases. On one hand, the choice and the evolution of names or keywords used by investigators since the beginning of the pandemic should be attentively evaluated. On the other hand, generic scientific database and specific datasets containing COVID-19 literature may differ quite significantly for both their coverage and overall searchability in terms of indexing, presence of identification codes, search tools, and delay from official publication date. The continuous flow of several thousands of COVID-19 articles published each month since April 2020 is making the search in databases with quality, consistency, and format issues already before the pandemic even more difficult. It should also be highlighted that preprints, or even publications that were made hastily available through websites and databases without a proper peer- and publishing review, may later on be republished with modifications, or even retracted, as already indicated in a specific webpage [29].

A further uncertainty remains about whether and how these datasets will be maintained in the future to allow patent information specialists to evaluate prior evidence correctly with respect to current legal standards. This situation may evolve in the future towards a more structured and integrated way of making available scientific publications. These events further justify specific attention when searching for, re-extracting, or maybe even saving copies of all relevant COVID-19 information as available at a given date, if to be considered at a later time during the patent proceedings or legal challenges. For instance, the notice of concurrent patent filings or license may be mentioned within the dedicated section of the text for a number of COVID-19 articles published in journals that commit authors to notify potential conflict-of-interest or financial matters related to the content of article. This information of potential patent or business relevance is irregularly, or not at all, indexed in PubMed using a specific search field (Conflict of Interest Statements) and thus it is hardly searchable and quantifiable without a detailed article-by-article analysis [30].

### 3.4. Expanding COVID-19 scientific literature: business, clinical & experimental information

Additional resources for COVID-19 information beyond scientific literature databases may be important to identify details relevant to expand the search criteria and provide patent information users with more complete reports. This consideration applies to searches that are related not only to getting a more global coverage of a specific product or technology but also to the origin or confirmation of statements about COVID-19 found in social or general media that may have received more or less deserved attention.

A first example of such resources are dedicated websites of journals with a structured and commented-on selection of COVID-19 medical and scientific news reports, where further details relevant to patent-related topics may be found for products and technologies that investigators and entities may disclose in parallel or even before the actual scientific publication. Major scientific publishers such as *Science* [31], *Nature* [32], *Cell* [33], *JAMA* [34], or *BMJ* [35] have consolidated news, links, and other COVID-19 contents within a specific webpage in their portals, covering their own and other publications. Other information webpages have established fully or partially free portals to consolidate their articles and news coverage where business or social consequences of the pandemic are also reported with potentially useful details and insights. Among them, the dedicated webpages of *The Economist* [36], *The Scientist* [37], and *Bioworld* [38] are worth a mention for their higher standards in selecting and reporting specific COVID-19 news.

A second example are the websites reporting clinical trials, and in general the activities that clinical institutions have established to validate findings of diagnostic and therapeutic relevance in human subjects but such trials are often cited in scientific publications or news in a not

very detailed or updated manner. The search and analysis of such clinical information that is disclosed between the initial discovery and the much later formal publication of final clinical results at conferences or in journals has always been difficult due to the wide variety of completeness, standards, and search features in dedicated databases. The COVID-19 crisis has provoked an intense interest from a large audience in any clinical progress, even at most preliminary stage, and this has pushed traditional providers of clinical information to consolidate and update data on COVID-19 clinical trials in a more complete and transparent manner. The main website covering clinical trials worldwide, Clinicaltrials.gov, has established specific webpages with a summary [39] and an indexing system [40] for COVID-19 clinical research. Updated listings and links to national clinical registries are also accessible through specific links in the International Clinical Trials Registry Platform at WHO [41]. Additional information resources associating clinical research and official status for drug and vaccine authorization can be found in the dedicated webpages established of regulatory authorities such as the Food & Drug Administration (FDA) [42] or by other organizations such as the Milken Institute [43].

A third example is experimental data that authors cite in a publication but are not fully reported in the corresponding official PDF file as searchable full text, tables, or graphics of the document. This definition covers multimedia contents (for example, videos or audio descriptions of protocols) and functional and structural information over chemical or biological compounds, often including information related to experimental values as measured over a number of samples, (such as inhibitory, stimulatory, or antiviral activity) or in a human population (such as clinical or epidemiological data). Biological sequences or chemical coordinates are also available in well-established, dedicated databases that are freely accessible worldwide and allocate unique accession numbers to the molecules and data that is deposited for each molecule. Investigators can use this indexing in their publications and for performing searches for specific entries by applying text, sequence, and/or chemical criteria (further details are provided in Section 5 below). Such data may be provided by the authors by means of either links in the corresponding webpage of the publication or to a stand-alone file within a dedicated public repository. These types of disclosures are generally briefly cited in the official text of the article as supporting information, supplementary materials, or otherwise, and made available to readers for further analysis as links to web-only material in various formats (PDF or Word files, csv/xlsx Excel-compatible files, txt or xml files, or other data-specific formats that may be not easily used).

The experimental data or other files associated with a publication within an article can be extracted and compared only by reviewing the webpage of the article. At most, only specific databases, such as PubMed or PMC have an “associated data” filter that would allow searchers to identify relevant articles presenting such files [44], but this search feature does not guarantee that all articles with additional files will be actually identified. This publication trend started in the early 2000s, in parallel with pressure upon authors and publishers for more transparency about the access and quality of original data used for a publication. Technical and cost accessibility of new technologies allow for the generation and elaboration of an exponentially increasing amount of data for a wide range of applications. Similar files, in alternative formats and/or versions, can be uploaded by authors even in locations other than the publisher’s website, such as webpages or repositories that are maintained by many universities and organizations. Obviously, a number of COVID-19 articles have also been published alongside similar additional files, for example an article that provides readers with a supporting spreadsheet file summarizing predicted or experimental binding affinities of more than 8000 compounds as candidate inhibitors of a SARS-CoV-2 protease [45]. Some websites have tried to consolidate such data in electronic repositories where these files, and other datasets made available by investigators independently from any publication or submission as preprint, are aggregated in a more or less organized manner into databases, allowing data mining across documentation in various formats. Among these searchable databases listed in Supplementary File 1, the most interesting are COVID-19 Open Research Dataset (CORD-19 [46]), Mendeley COVID-19 Dataset, Zenodo COVID-19 Data, Figshare COVID-19 Open Research Data, and Google COVID-19 Open Data.

## 4. Extracting and comparing COVID-19 patent literature as published in Jan.–Aug. 2020

### 4.1. COVID-19 patent information: providers and criteria

Since the official announcement of the COVID-19 outbreak in Feb. 2020, patent offices have started a series of initiatives and improved their internet services to support applicants and other users of patent system to perform actions from remote locations with respect to the formalities, the payments, the filing and the examination of patent applications, as well as other proceedings usually performed at patent offices such as the extension (automatic or requested by the applicant) of some deadlines. The World Intellectual Property Organization (WIPO) has started providing patent information users at large (investigators, companies, policymakers) with information collected from patent offices worldwide, reports, and specific tools. Together with the COVID-19 IP Policy Tracker [47], the WIPO website also provides users with a glossary containing key concepts related to the technologies applicable to COVID-19 in English and nine other languages [48] and a dedicated COVID-19 Search facility within their own patent search tool Patentscope [49]. Similar initiative taken by WIPO as well as other patent offices have been reported in *World Patent Information* [50] and reviewed in this Section together with other patent information tools identified by authors in the websites of patent offices in the webpages listed in Table 3.

These webpages and the quick description that is provided are clearly not exhaustive, but some general observations can be made, also in view of the benchmarking made in a presentation at EPOPIC 2020 [51]. It should be noted that the two major patent offices - the United States Patent and Trademark Office (USPTO) and the Japanese Patent Office (JPO) - have preferred (at least until end November 2020) not to establish any specific search facilities on their websites, and instead to concentrate on a series of procedures and programs providing support to inventors and patent applicants as indicated in a joint statement issued in June 2020 [52]. These patent offices have proposed compiling a database of patent licensing data, as the one accessible through the USPTO’s COVID-19 Response Resource Center [53]. A further observation is that the COVID-19 patent search facilities can be divided into those providing users with full details about the search criteria used to generate the datasets on their websites (thus, giving the possibility to adapt them to users’ needs) and those not explicitly describing the criteria used to establish and update their datasets. The first category includes only the dedicated patent search tools in WIPO (Patentscope), EPO (Espacenet), IP Australia, and The Lens Patents. Aside from patent offices, the Lens.org and Medicine Patent Pools websites provide users with patent datasets that are either covering biological sequences (in the former) or are more drug-focused (in the latter). In general, all such listings broadly cover products and technologies applicable to Coronaviruses in general, other virologic pathogens, and epidemics, without imposing any time limit.

The analysis of the patent literature related to COVID-19 has been pursued by some authors, as of Nov. 2020, mostly covering pre-2020

**Table 4**
Geographic distribution of patent documents in Early COVID-19 Patent Dataset.

| Patent Office | No. of patent document | Commentaries |
|---|---|---|
| China | 822 | Including 34 utility models and 26 granted patents |
| India | 119 | All patent applications whose bibliographic information are published in the Official Journal of the Indian Patent Office |
| United Kingdom | 55 | Including 54 provisional patent applications (only one regular patent application) |
| Australia | 43 | Including 15 Innovation Patents and 27 provisional patent applications (only one regular patent application) |
| South Korea | 15 | Including 7 granted patents |
| USA | 14 | Including 10 Continuation-in-Part applications, with earlier priority dates between 2013 and 2019 (one already granted) and 4 patent applications filed in 2020 (one already granted), mostly relating to either diagnostic or therapeutic technologies |
| Singapore | 14 | Including 13 provisional patent applications (only one regular patent application) |
| Germany | 12 | All utility applications, for either diagnostic or protective means |
| Russia | 12 | Including 9 granted patents, all related to either diagnostic or therapeutic technologies |
| Israel | 6 | |
| Brazil | 5 | |
| Spain | 4 | |
| Philippines | 2 | |
| WIPO | 1 | |
| EPO | 1 | |
| Colombia | 1 | |
| Chile | 1 | |
| Italy | 1 | |
| Taiwan | 1 | |
| Norway | 1 | |
| **TOTAL** | **1130** | Prevalence of patent documents from China and India (>80%) |

patent filings, sometimes combining references from scientific literature and patent-based findings about coronavirus on the basis of previous outbreaks caused by SARS-CoV-1 and MERS-CoV that require confirmation as being applicable during this pandemic. These publications are not always describing the search criteria and methodology used to establish their selection of patent literature starting from technical information and needs related to COVID-19. For instance, some publications have focused on potentially relevant patent filings related to vaccines and other drugs of various chemical nature, highlighting specific molecular sequences or scaffolds and patent classifications [54,55], antiviral strategies involving small molecules and biologics [56–59], natural compounds or traditional medicine products with anti-viral properties [60,61], or nanotechnology-based solutions [62].

### 4.2. The Early COVID-19 patent dataset: construction

It is not easy, when comparing the features and outcomes of the search strategies proposed in these publications (where available) or in the free patent information facilities listed above, to evaluate which are most interesting for future searches. However, it seems interesting to consolidate the search strategies that were explicitly disclosed and meant to be applied by users in the same websites proposing them, such as those present in the WIPO, EPO, and The Lens websites between Feb. and Nov. 2020. The main features of these and other free databases for searching patent information have been previously review [63] but some major improvements have been introduced since 2019 and have been exploited to design tools supporting COVID-19 search.

An overview of the search criteria defined in Patentscope, Espacenet and The Lens Patents, for a total of almost 300 search strategies, have been consolidated in xls file and analyzed. This review has been performed by separating keywords, the International Patent Classification (IPC) codes, and the Cooperative Patent Classification (CPC) codes that have been assigned to a large variety of technologies, and not only those of a strictly biomedical nature, applicable to COVID-19. The listing shows some specificities of choices made by either each provider compared to the others or by the same provider for different topics. For instance, Patentscope search strategies all present at least a technology-specific IPC code (but no CPC codes) and, where present, always the same combination of Coronavirus-related keywords (including viral classifications and synonyms but, interestingly, not the official virus name SARS-CoV-2). The Espacenet search strategies presented in the "Fighting Coronavirus" website [64], associate Coronavirus-specific keywords (but in various alternative strategies comprising different names, acronyms, and even some non-English terms, and again never SARS-CoV-2) to technology-specific keywords, IPC and/or CPC codes, sometimes choosing very specific classification codes. The search strategies proposed by The Lens Patents are those more broadly referring to Coronavirus-related matters (by keywords only) and applicable technologies (by IPC only). At the level of specific topics, it is interesting to compare some broader choices of patent classification codes made in Patentscope compared to those made for searching Espacenet, for instance with respect to informatics, medical equipment, or drug classifications (with Espacenet not listing any IPC code under A61P, quite strangely).

The above-mentioned observations regarding COVID-19 patent literature analysis and search strategies have motivated the authors to evaluate patent publication by restricting the search, on one hand, to the main names and acronyms found in the Covid-19 scientific literature analyzed in Section 3 and, on the other hand, to the patent documents actually filed and published during the same first eight months period (between January and August 2020). The search was performed in a selection of patent search facilities and confirmed by later searching in the patent registry and databases of major patent offices, limiting the presence of the main COVID-19 keywords (those identified as Group (1) criteria in Section 3) to title, abstract, and claims, where the potential usefulness in any aspect of COVID-19 medical management is more explicitly indicated. Obviously, such a search cannot be considered as a complete search of COVID-19 findings in patent literature but is intended only to point out how and where users of the patent system have chosen to file patent applications and improve the visibility of their own patent activities under such unusual circumstances, even by requesting an accelerated examination and publication in many cases. The search was initially performed in databases covering several patent authorities and then confirmed or refined using the databases and registers of national patent offices. Further details on the search process and of the construction of this patent literature dataset, named "Early COVID-19 Patent Dataset" and comprising 1130 entries, are provided in Section 2 above and in Supplementary File 2.

### 4.3. The Early COVID-19 patent dataset: general filing trends

A first level of analysis of the Early COVID-19 Patent Dataset is made with respect to jurisdictions and types of documents filed in such jurisdictions (Table 4) to evaluate how the different types of patent protection available in each country, as summarized for PCT Contracting States in the dedicated WIPO webpage [65], have been preferred by applicants for early patent examination and disclosure. This overview clearly shows major differences, with patent filings in China and India representing more than 80% of total patent documents, with only occasional filings submitted at major international offices (WIPO and EPO), or in most other countries. Another interesting trend is the one observed in countries where it is possible to search the early publication of details for provisional patent applications or Innovation Patents, such as in the United Kingdom, Australia and Singapore. In this subset of patent document (totaling almost 10% of Early COVID-19 Patent Dataset), the actual level of disclosure is very uneven, with titles being mostly quite generic but sometimes with details about chemical or biological

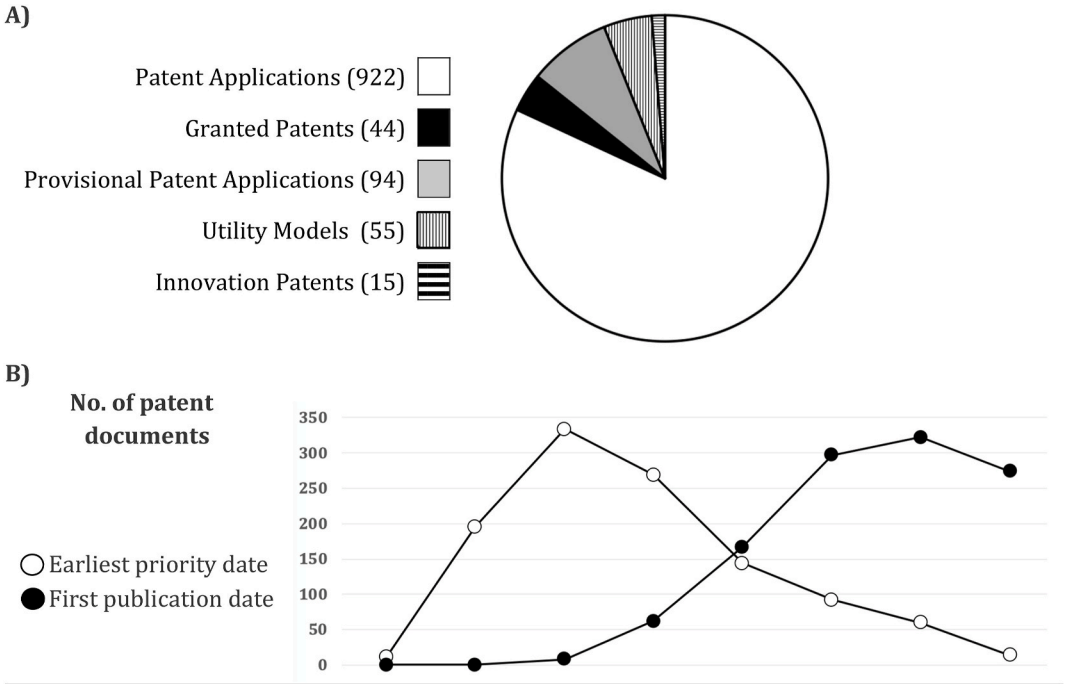
products. This information points, nonetheless, to inventors and applicants that may file regular applications on these topics using provisional patent filings as priority documents in the following 12 months, similarly to those authors of scientific publications that have declared the filing of a patent application comprising data present in the article (a specific article about this topic is in preparation).

The Early COVID-19 Patent Dataset also shows that applicants have completed the full process from the examination up to patent grant in less than 6 months in countries such as USA, China, South Korea, and Russia. The number of US patent filings is mostly consequential to the use of US-specific provisions to continue and re-submit pending applications with additional information. Applicants from other countries like Germany, Israel, Brazil, or Spain show some limited interest in achieving earlier patent protection even at a national level, sometimes by appplying for utility models. This data seems to indicate a general preference for making use of national rather than international authorities to accelerate the patent proceedings but with major differences across countries. In China, according to a study of listed Chinese companies [66], the applicants appear engaged in a kind of "patent race", but other considerations may then affect the preference of domestic applicants with respect to the patent examination and renewal process [67]. A recent study also suggests that non-Chinese companies could expect to be disadvantaged during patent examination by "technology protection" [68]. In any case, the Chinese Patent Office provided an unusual visibility to applicants that obtain an early patent grant, as shown in the press release about a patent covering a COVID-19 vaccine [69].

A second level of analysis in the Early COVID-19 Patent Dataset is made with respect to the type of patent documents, in general and by earlier filing or publication date (Fig. 5). The data show that, aside to standard patent proceedings, but the preference for utility models or provisional patent applications is clear in specific countries. When these data are analyzed by the month indicated as the earliest priority date (or initial filing) and the publication date between Jan. and Aug. 2020, the earliest patent filing date found was in mid-January, with the earliest publication dates in March 2020. The global trend of the two dates is fairly predictable, with an increase of publications over the selected period of time and the peak filings in March 2020 which follow the same pattern as the number of months between initial filing and publication (on average slightly higher than 3 months). However, a burst of filings appeared, especially those of provisional patent applications and utility models filed between February and April 2020. Since May 2020, the number of already published patent filings reached a kind of plateau, if not decreased but it can be expected when taking into account that, for example, provisional patent applications have a statutory, public notification of less than two months from filing).

### 4.4. *The Early COVID-19 patent dataset: patent classification symbols and trends in main technological domains*

A third type of analysis in the Early COVID-19 Patent Dataset was carried out with respect to the technological domains where early patent




| Month (2020) | | Jan. | Feb. | Mar. | Apr. | May | June | July | Aug. |
|---|---|---|---|---|---|---|---|---|---|
| By Priority Month | Standard or Innovation Patent Applications | 1% | 18% | 31% | 23% | 11% | 10% | 4% | 1% |
| | Provisional Patent Applications | 0% | 0% | 21% | 29% | 15% | 8% | 19% | 8% |
| | Utility Models | 2% | 37% | 17% | 24% | 12% | 8% | 0% | 0% |

**Fig. 5.** Early COVID-19 Patent Dataset: analysis by document and technology type. The Early COVID-19 Patent Dataset is analyzed with respect to the type of documents, with the total numbers indicated in parentheses (A) and the earliest month indicated in the official databases for claiming priority rights (or as direct filing) and for publication (B). The percentage of documents, by priority month, for each type of document is shown in the table below the graph.

filings and publications were pursued (Fig. 6). The authors have defined four technological areas on the basis of the categories defined in the literature cited above], the categorization of criteria for searching COVID-19 patent literature in Patentscope and Espacenet, the IPC classification (at the level of group and where available), and by looking at the English titles of the documents in the Early COVID-19 Patent Dataset. These four technological areas are Diagnostic (means to identify or predict SARS-CoV-2 infected subjects), Therapeutic (means to treat COVID-19 patients and SARS-CoV-2 infected subjects), Protection (means to avoid SARS-CoV-2 infection by blocking virus contact or propagation mechanically), and Cleaning (means to avoid SARS-CoV-2 infection by removing or destroying the virus). This categorization was used to simplify the analysis of topics that are present in the early filed and published patent documents and should not be obviously considered as a definitive status, with patent documents that may disclose findings relevant or exploitable in more than one domain. As main patent classification criteria, the IPC Group was chosen since it appears providing a granularity sufficient to identify main technological features.

When the consistency and the number of IPC Groups is compared, first within each technological domain and then between technological domains, some distinctively featured documents can be identified.

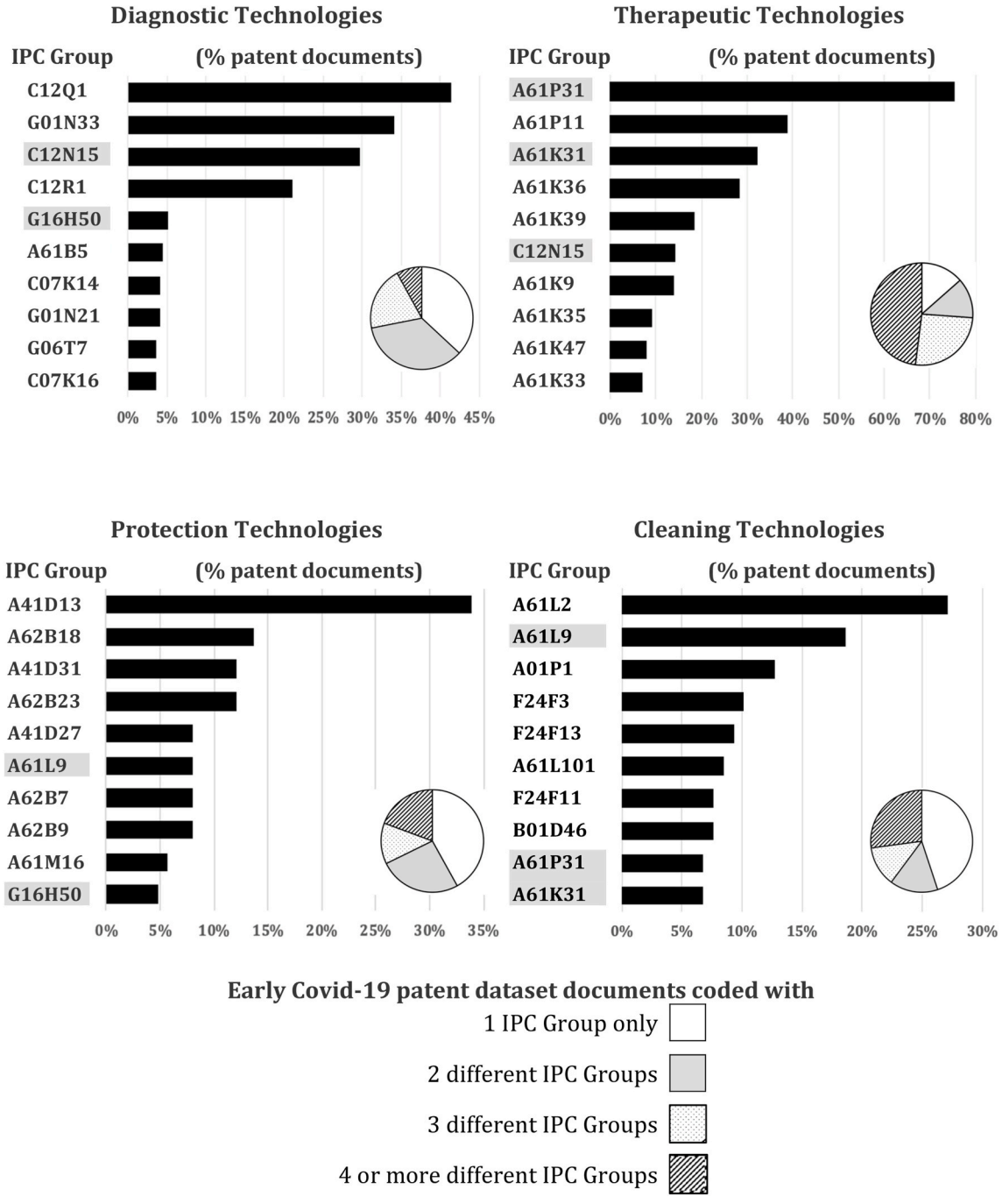



**Fig. 6.** Early COVID-19 Patent Dataset: category of technologies and related IPC Groups. The 10 most frequent IPC Groups in the documents published with at least one IPC code (1033 out of 1130) are listed under the four technological domains defined by the authors The percentage of documents within each category of technology that present either a given IPC Group (in the histogram) or a given number of distinct IPC Groups (pie chart with the key at the bottom of the figure) is indicated. The greyed IPC Groups are those found common to two technological domains.

**Table 5**
Patent documents in Early COVID-19 Patent Dataset that claim biological or chemical entities.

| IPC Group | No. of patent documents in Early COVID-19 Patent Dataset | Commentaries |
|---|---|---|
| **C07K14** | 38 | Protein sequences containing more than 20 amino acids from SARS-CoV-2 (or human ones recognizing SARS-CoV-2 sequences), to be used for developing diagnostic and/or therapeutic applications |
| **C07K16** | 31 | Antigens from SARS-CoV-2 (or human ones recognizing SARS-CoV-2 sequences) or related antibodies, to be used for developing diagnostic and/or therapeutic applications |
| **C12N5** | 22 | Human, animal or plant cell products and technologies, to be used for developing diagnostic and/or therapeutic products |
| **C07K19** | 19 | Peptides-based products to be used for diagnostic and/or therapeutic products |
| **C12N7** | 16 | Bacterial cell products and technologies, to be used for developing diagnostic and/or therapeutic products |
| **C12N1** | 8 | Viral products and technologies, to be used for developing diagnostic and/or therapeutic products |
| **C07F9** | 6 | Compounds containing elements of Groups 5 or 15 of the Periodic System (essentially Phosphorus-containing compounds) |

Within Diagnostic Technologies, the IPC Groups C12Q1, G01N33 and C12R1 (identifying processes and assays for investigating, measuring or testing or analyzing materials, using microorganisms, nucleic acids, enzymes, etc.) are largely prevalent, where patent applications may combine two or more Subgroups within these IPC Groups. A frequently present IPC Group, C12N15 (referring to genetic engineering and related products such as plasmids), may be used to define both Diagnostic and Therapeutic Technologies. However, patent documents that refer to Therapeutic Technologies are strongly associated with specific A61P Groups associated with anti-infectives or disorders of the respiratory system and to a generally wider range of IPC groups, given the frequent association of specific IPC Groups under A61K and A61P subclasses. This finding would suggest that the number IPC Groups may be used to facilitate a preliminary categorization of patent documents relating to drugs and vaccines. Many IPC Groups within the C07 Subclass that define molecular entities having different chemical nature also apply when referring to the pharmaceutical use and preparations of chemical, biological, or otherwise defined compounds. A large variety of IPC Groups under A61K Subclass are present in patent documents categorized under Therapeutic Technologies, with the expected prevalence of A61K31 (defining medicinal preparations containing organic active ingredients, often associated with novel reformulation and/or administration of known drugs for further medical uses) and A61K39 (broadly defining medicinal preparations containing antigens or antibodies, including vaccines). The frequency of patent filings under IPC Group A61K36 (defining medicinal preparations that contain ingredients with undetermined constitutions from algae, fungi, or plants) is relatively high in China and India where such preparations are extensively used in traditional medicine.

This analysis further shows that Protection Technologies are mostly associated with IPC Groups under IPC Subclass A41D (which includes a series of outerwear, accessories, and protective garments, in particular for medical personnel and uses) and IPC Subclass A62B (which includes a series of life-saving devices and products such as those related to respiratory apparatus, helmets and filters for breathing-protection purposes). Patent documents categorized under Protection Technologies may share the IPC Group G16H50 (referring to Information and Communication Technologies for medical diagnosis or other medical uses) or IPC Subclass A61L (referring to disinfection, sterilization or purification methods for human safety and health) with Diagnostic or Cleaning Technologies, given the possibility that such devices may combine features or be useful in other domains. Patent documents categorized under Cleaning Technologies also share a comparable distribution in a number of different IPC Groups (somehow between the distribution observed for Diagnostic and Therapeutic Technologies) but present specific IPC Groups related to preparations, chemical compounds, methods or apparatus for disinfecting, sterilizing materials (under A61L2, A01P1, and A61L101) or air-conditioning and ventilation (under groups within IPC Subclass F24F).

The analysis of IPC Groups in the Early COVID-19 Patent Dataset may suggest other trends. For instance, the IPC Groups related to novel chemicals or biologicals are not strongly represented, possibly because

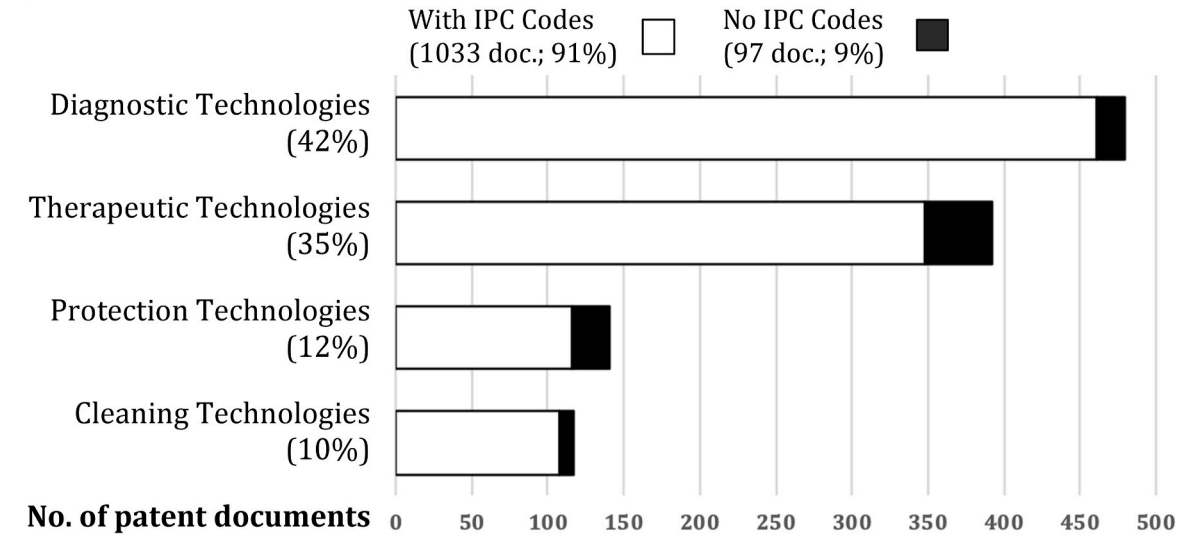
**Table 6**
Potentially relevant IPC (Sub)classes or Groups poorly represented absent in the Early COVID-19 Patent Dataset.

| IPC Symbol | Short definition of Technology or Product |
|---|---|
| **A41D19** | Gloves (with anti-viral properties) |
| **A47B-D; E04B-C** | Tables, chairs, partitions, panels, and other specially adapted furniture |
| **A47K** | Sanitary equipment (for disinfection) |
| **A61B18,42** | Surgical instruments, gloves, containers for surgical/diagnostic tools |
| **A61F7** | Thermal appliances for medical or therapeutic treatment |
| **A61G3** | Medical transportations |
| **A61H31** | Artificial respiration or heart stimulation |
| **A61M1** | Suction or pumping devices for medical purposes |
| **A61P33** | Antiparasitic agents |
| **A62D5,7** | Materials for coverings, clothing protecting from harmful chemicals |
| **A62D9** | Composition of chemical substances for use in breathing apparatus |
| **B08; B15** | Chambers or hoods for preventing escape of dirt or fumes |
| **B25J9** | Programme-controlled manipulators, robot |
| **B29C; B33Y** | Additive/3D manufacturing or plastics (for protections, face masks) |
| **B32B** | Layered products (with imprinted surfaces having antiviral properties) |
| **B60N; B60R** | Arrangements, seats specially adapted for vehicle |
| **B65D83** | Sterilizing air, objects involving spraying |
| **C07D** | Heterocyclic compounds (with anti-viral properties) |
| **C07H** | Sugars and nucleotide derivatives (with anti-viral properties) |
| **C09D** | Coating composition with inactivating or repelling effect against viruses |
| **C12M3** | Cell or virus culture apparatus |
| **C12P** | Fermentation or enzyme-based processes for preparing compounds |
| **F24F9, -12, -24** | Screen, ventilation to prevent entry or contact with virus |
| **G02B21,23,37** | Microscopy, instruments for viewing inside bodies |
| **G06F3,19** | Digital computing or data processing equipment or methods |
| **G06N5,7** | Artificial Intelligence for medical uses |
| **G16C** | Chemoinformatics |

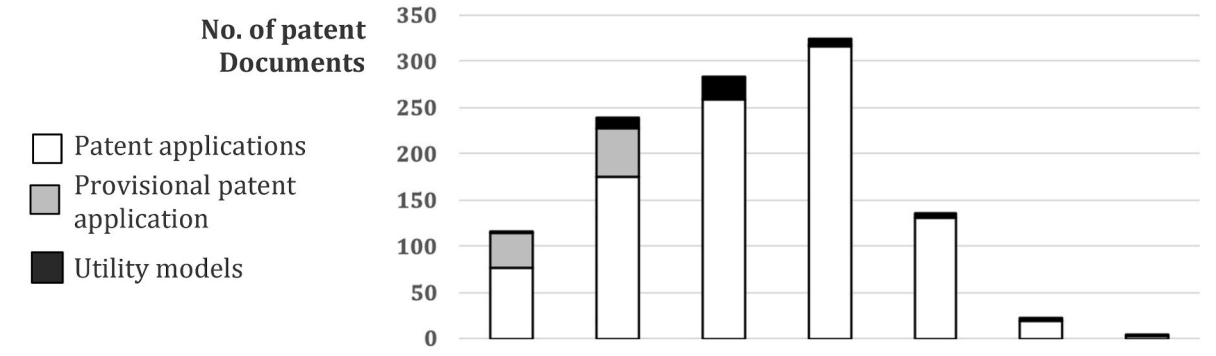
generating and characterizing novel compounds requires more time and/or because applicants having the technical and financial resources to generate them prefer to pursue patent proceedings in a more traditional manner, taking advantage of the 12 month-priority with publication at 18 months. Earlier filed and published patent documents referring to COVID-19 applications seem to mostly refer to novel uses or adaptations of existing compounds, preparations, or methods. However, even if present in a limited number of patent documents, novel compounds can be identified according to IPC Groups, the most frequent being listed in Table 5. Among them, there is a clear prevalence of proteins (in the form of antigens or antibodies, for instance), but Groups under IPC Subclass C12 N defining cell-based products and technologies are also present, and in fact are more present than classical IPC Groups for small molecules such as those under IPC Subclasses C07D or C07C. The presence of IPC Group C07F9 is explained by the fact that some phosphorus-containing compounds are known as nucleotide analogues having anti-viral properties, including the candidate COVID-19 drug Remdesivir [70].

The representation and trends described for the Early COVID-19 Patent Dataset using the IPC Groups may be compared to the results

A)




B)




| Priority to publication (days) | 0-30 | 31-60 | 61-90 | 91-120 | 121-150 | 151-180 | >180 |
|---|---|---|---|---|---|---|---|
| Diagnostic Technologies | 7% | 21% | 22% | 32% | 15% | 2% | 1% |
| Therapeutic Technologies | 8% | 21% | 26% | 31% | 11% | 2% | 1% |
| Protection Technologies | 18% | 24% | 32% | 17% | 7% | 1% | 1% |
| Cleaning Technologies | 17% | 27% | 27% | 20% | 7% | 1% | 0% |

**Fig. 7.** Early COVID-19 Patent Dataset: analysis by technological domain. The number of patent documents in the Early COVID-19 Patent Dataset are assigned to each of the four technological domains defined by the authors, indicating the percentage in the Early COVID-19 Patent Dataset in parentheses (A). The Early COVID-19 Patent Dataset is also analyzed with respect to the days between the date of priority filing (or direct filing) and of publication, and the four technological domains (B).

obtained using the almost 300 COVID-19 search strategies proposed by Patentscope, Espacenet, and The Lens Patents. It is interesting to note which IPC Groups are shared with the approaches that are proposed by these patent search platforms. As expected, a large number of IPC Groups are present in both the predefined searches and in the Early COVID-19 Patent Dataset. The large majority of the IPC Groups listed in Fig. 6 are present in one or more of the published search strategies with a few exceptions, mainly for Cleaning Technologies (A61L101, A01N59, A01P1, and F24F13). However, quite a number of IPC Groups proposed in Patentscope, Espacenet, and The Lens Patents search strategies are rarely or not at all represented in the Early COVID-19 Patent Dataset, as summarized in Table 6. These IPC Groups cover quite a variety of topics, in particular with respect to transportation, appliances, materials, information technology, disinfection, or medical uses that are either more often categorized in the Early Covid-19 Patent Dataset using other IPC Groups or, apparently, absent from the Early Covid-19 Patent Dataset.

Future searches in the COVID-19 patent literature will allow a conclusion of whether the observations made above on the basis of Early COVID-19 Patent Dataset are fully or partially reflecting the applicants' attitude that are active in specific countries and/or in technological areas (lack of interest, patent filings still to be completed, or simply filed and prosecuted in a more standard, not accelerated manner). In any case, extending the search to include both CPC codes and keywords related to COVID-19 in the full-text and in the granted claims will return a more complete analysis of innovation trends consequent to or triggered by this crisis during the year 2020. Further trends in the Early COVID-19 Patent Dataset may be defined on the basis of the above mentioned four technological domains and filing trends (Fig. 7). As expected, a majority of patent documents in the Early COVID-19 Patent Dataset are associated with either diagnostic or therapeutic uses, with an even higher preference for Therapeutic Technologies among provisional patent applications. However, the distribution according to the days between the filing and publication month suggest that, independently from the filing month, those assigned to Cleaning or Protection Technology were often published earlier (at least 70% were published less than three months from filing) than those for Diagnostic or Therapeutic Technologies (at least 70% were published between two and four months). Regarding the distribution and prevalence of the IPC Groups described above, often only later, more in-depth searches in the patent literature explain and confirm these initial observations about prevalence or preference of choices made by applicants about filing and disclosing patent applications in connection to COVID-19.

## 5. COVID-19 chemical information and biological sequence information

### 5.1. Access to COVID-19 chemical & biological structural information

The access and searchability of biological sequences and chemical structures within scientific and patent literature is facilitated by the dedicated databases established by many organizations and their specific policies for managing such information. In particular, two institutions, NCBI (National Center for Biotechnology Institutions, USA) and EBI (European Biotechnology Institute, UK), should be mentioned since they have made a major effort to aggregate and offer a structured access to biological and chemical data about SARS-CoV-2 biology. Main COVID-19 resources in NCBI and EBI websites are summarized in Table 7. A repository of raw, primary functional data available through NCBI is associated with NCATS (National Center for Advancing Translational Science) which organizes its own screening and biological data into four main sections, as described in a related publication [71]. A further NIH COVID-19 dedicated webpage summarizes its own and external resources for data under separate, searchable sections. EBI has announced a series of initiatives about providing timely information on SARS-CoV-2 and COVID-19 research, the main COVID-19 resources being grouped in a separate portal [72] that is intended also to facilitate data submission, sharing, and analysis from external users.

Among these data, those related to biological properties of small organic molecules that may be validated as COVID-19 candidate therapeutic compounds are also of particular interest for patenting analysis. In this domain, the NCBI and EBI databases that cover chemical compounds of biological interest (PubChem [73] and ChEMBL [74], respectively) present chemical and experimental data about potential SARS-CoV-2 activities of specifically indexed compounds listed therein. It is not easy to list, let alone to compare, similar websites or dedicated sections in terms of origin, amount, and overall quality of their content. It is equally difficult to evaluate how many publications, and possibly patent applications already filed, may have exploited data found in these repositories.

### 5.2. Identifying COVID-19 protein sequences & related publications

A separate analysis should be dedicated to the search and analysis of biological sequences (i.e., nucleic acids and proteins). The NCBI and EBI portals usually covering this type of information have established some specific webpages and other resources for the analysis of newly identified SARS-CoV-2 sequence data, such as in COVID-19 UniProtKB [75] or Ensembl COVID-19 [76] for protein and nucleic acid sequences, respectively. These websites and resources in other websites may also cover and compare nucleic acids or proteins of previously identified Coronaviruses (being potentially relevant for present SARS-CoV-2

**Table 7**
Webpages dedicated to bioinformatic and chemoinformatic data related to COVID-19 research.

| | | |
|---|---|---|
| NIH-NCBI data resources | **OpenData COVID-19 (NCATS)** | **NCBI SARS-CoV-2 data hub** |
| | *OpenData Browser*, activities of compounds in the Approved Drug Collection and other collections in relevant biological assays) | Listings of deposited SARS-CoV-2 protein and genome sequences as deposited by researchers worldwide |
| | *Omics Efforts*, listing publications that contain publicly accessible original (or re-analyzed) data from studies in proteomics, genomics, epigenomics, etc. as generated using human samples, cell/animal models, SARS-Cov-2 variants, and/or clinical data | Links to bioinformatic tools for analyzing, geo-localizing, and comparing such sequences |
| **EMBL-EBI data resources** | **COVID-19 Data Portal** | **Other EBI Resources** |
| | *Sequences,* being *viral* (SARS-CoV-2 ones) or *host* (relevant human ones) | *UniProt COVID-19 Portal*, with latest available SARS-CoV-2 protein sequences |
| | *Expression* data at gene/protein level of human genes involved in SARS-CoV-2 activity or infection | *Ensembl SARS-CoV-2 genome browser*, with links to latest SARS-CoV-2 nucleotide sequences |
| | *Proteins,* with sequence and functional data on SARS-CoV-2 and human proteins | *PDBe-KB COVID-19 Data Portal*, listing protein sequences, structures and interactions data |
| | *Biochemistry,* with information about relevant human pathways, interactions, targets, compounds | *IntAct coronavirus dataset*, listing molecular interaction data from literature |
| | Links to other Data hubs, bioinformatic resources, and European research projects | *Europe PMC COVID-19 pre-prints,* indexing separately pre-prints with open data |

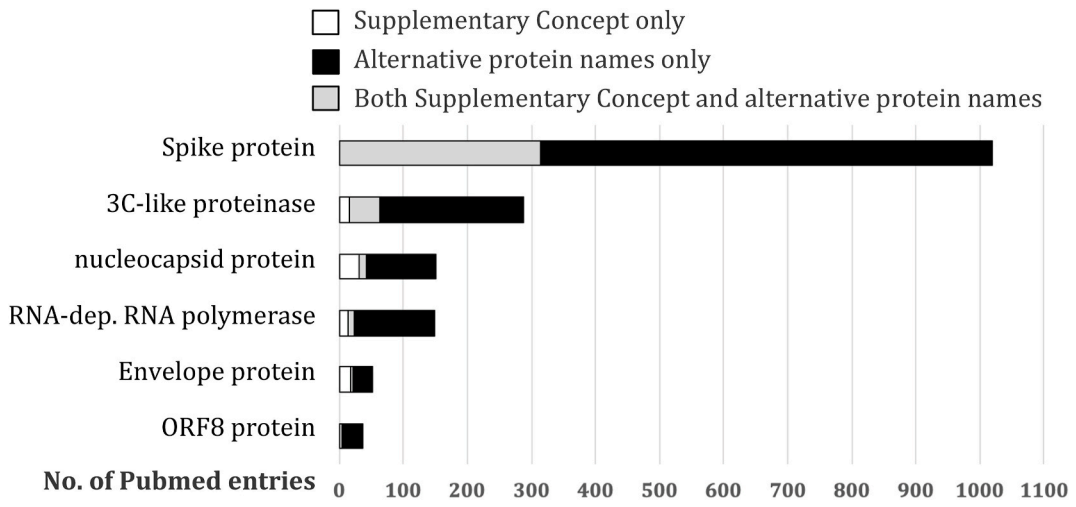



**Fig. 8.** PubMed Entries related to SARS-Cov-2 proteins. The NCBI MeSH Supplementary Concept and related alternative names for each SARS-Cov-2 protein in EBI, NCBI, and MeSH related records, in combination with Covid-19 keywords (see COVID-19 Group (1) criteria and MeSH Supplementary Concepts), were used to search within PubMed on Sept. 16th 2020 for records entered between Jan.–Aug. 2020. The six most commonly cited SARS-Cov-2 proteins are shown. Additional details about main names and codes for SARS-CoV-2 proteins are provided in Supplementary File 3.

patent and research activities), often in association with the human sequences that appear interacting with SARS-CoV-2 molecular components or are somehow involved in SARS-CoV-2 biology or pathology. Two further tools of interest are those available in the "Human Coronaviruses Data Initiative" portal in Lens.org [77], using their own PatSeq Finder search tool to extract prior Coronavirus-related sequences that are described in the patent literature (as of May 2020), as well as a database established at Oxford University and named CoV-AbDab dedicated to antibodies recognizing SARS-CoV-2 proteins [78].

A deeper analysis of these resources and information is beyond the scope of this article but patent information professionals should be aware that, as in many other biomedical domains, for SARS-CoV-2 there is also a wide number of codes, names, and representations with respect to the same biological sequence that may be found in different databases or publications. A main example is represented by proteins that are encoded by SARS-CoV-2 genome whose RNA sequence was first released on Jan. 10th, 2020 [79] but then integrated on the basis of several findings accumulated since then by analyzing clinical samples worldwide. The details of SARS-CoV-2 biology, genomic variability, and candidate targets for therapeutic intervention are extensively reviewed [80,81]. The majority of the SARS-Cov-2 RNA genome is translated within human infected cells as two large polyproteins (named ORF1a polyprotein, or Replicase polyprotein 1a, and ORF1ab polyprotein, or Replicase polyprotein 1 ab). These proteins undergo a proteolytic process generating further 16 proteins (named as NSP1-NSP16, which have mainly replicative roles or enzymatic activities). The remaining part of the genome codes for 4 proteins mainly performing structural functions (generally indicated as S, E, M, and N proteins) and for a series of accessory proteins whose actual existence in vivo, number, and function is still an object of intense study. They appear coded by Open Reading Frames (ORFs) in the SARS-CoV-2 genome that are partially overlapping with each other or with the ORF sequences of structural proteins. Apart from the 6 main accessory proteins (named as ORF3a, ORF6, ORF7a, ORF7b and ORF10), at least other 4 sequences have been identified as new potential ORFs by different investigators.

Some important details need to be taken into account when searching for names and sequences of the SARS-Cov-2 proteins are indexed in NCBI, EBI, and Lens.org websites, as of Nov. 2020. The names and codes identifying SARS-Cov-2 proteins for a total of 29 sequences (including some sequences that appear either unreviewed or as errors) are assigned in NCBI webpages to separate records as protein sequences and/or as separate Supplementary Concepts in MeSH thesaurus. On the EBI website, SARS-CoV-2 proteins are identified in the UniprotKB protein database under a specific taxonomy code which is associated with only 13 reviewed entries (10 directly coded proteins, the 2 polyproteins, plus another ORF not specifically identified as such in NCBI) but UniProtKB further lists thousands of different, unreviewed, partial protein entries. Aside from obviously different accession numbers between EBI and NCBI databases, it should be also noted that some preferred names for SARS-Cov-2 proteins in the two databases are different, with a variable number of alternative names for each protein in the database record. Thus, if the search of patent-relevant information involves the specific structure, activity, and/or interactions of a specific SARS-CoV-2 protein, it is important to identify and use correctly at least the names that NCBI, MeSH, and UniprotKB list. An even more detailed, and complex, prior analysis of such databases would be then required to evaluate naming, coding, and actual sequences for SARS-CoV-2 nucleic acids, in particular for diagnostic uses and early detection of mutations within clinical samples that are associated toSARS-COV-2 variants-of-concern [82].

SARS-Cov-2 gene and protein names can be used to search PubMed, using both MeSH Supplementary Concept and the main scientific names to identify those protein that may be more extensively cited in patent applications as well (Fig. 8). The details on SARS-Cov-2 proteins (names, size, and codes) that were identified in NCBI and EBI resources are consolidated in Supplementary File 3. This analysis shows that only few among the annotated SARS-Cov-2 proteins are frequently cited in scientific articles listed by PubMed for the Jan.–Aug. 2020 period, with protein-specific Supplementary Concept being associated with a variable, but still quite low, fraction of total PubMed entries presenting the protein name in the title or abstract. In particular, four of them (S/Spike protein, NSP5/3C-like proteinase, NSP12/RNA-dependent RNA polymerase, and N/nucleocapsid protein) appear as those most extensively studied in COVID-19 scientific literature and thus, possibly, will be the most present in future patent literature as well. However, scientific literature has highlighted how both the biological activities and coding capacity of SARS-CoV-2 are still partially unexplored, with other poorly characterized, minor viral ORFs having potential regulatory or functional relevance [83–85]. Future studies about the sequence and properties of SARS-CoV-2 proteins and their interactions with human molecules may suggest additional targets and strategies for developing novel therapeutic or diagnostic applications, and thus opening new opportunities for patent filings.

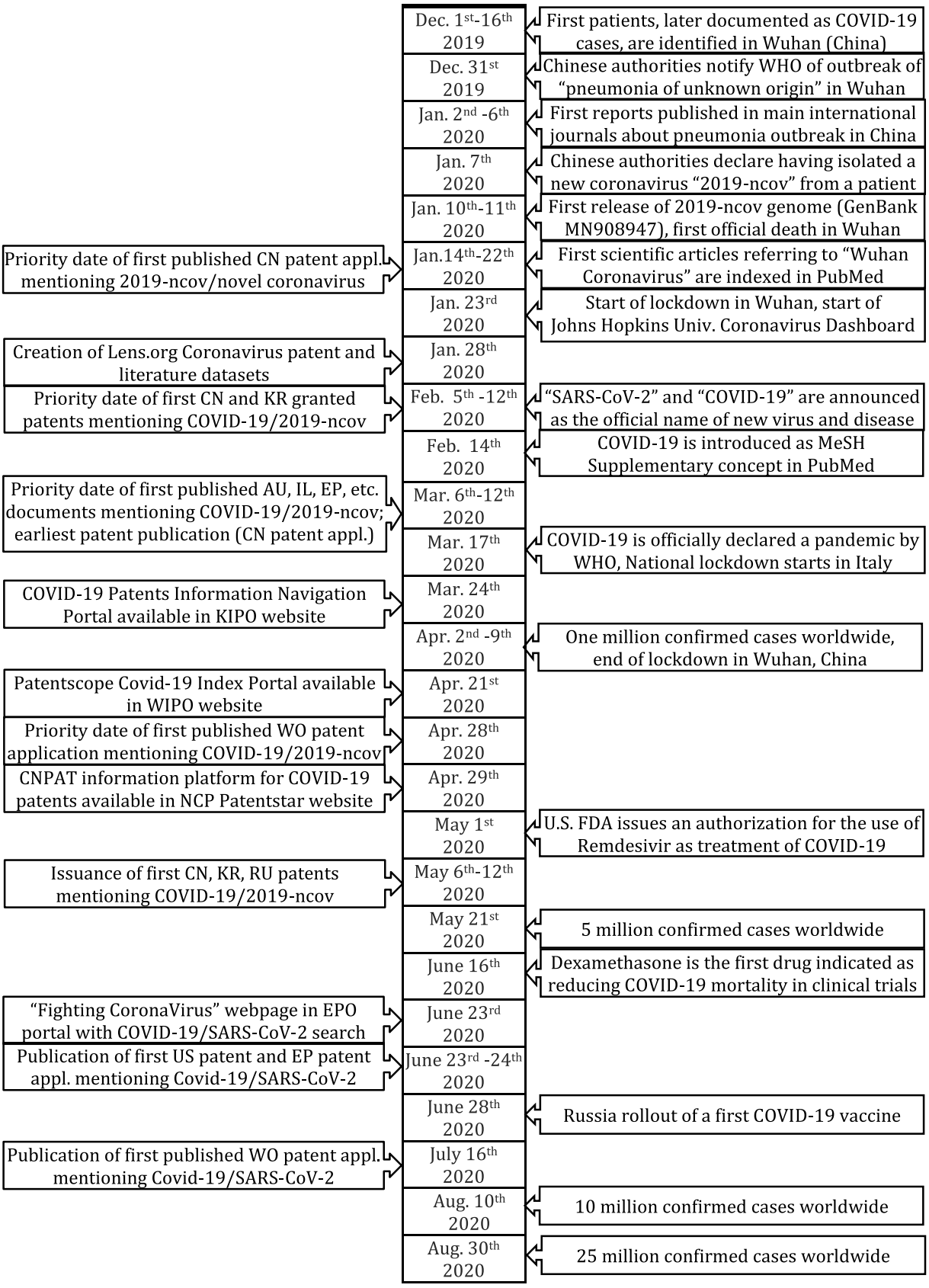



**Fig. 9.** COVID-19 patent and pandemic timeline during Dec. 2019–Aug.2020. The listed events and the related dates have been selected from the patent information and related resources (as available on November 2020) that has been presented in this article (left) and the cited literature (right).

## 6. Conclusions

This paper was drafted during the fourth quarter of 2020 and does not (and cannot) list and analyze, all the relevant websites and data sources in this ever-changing landscape but our findings suggest some considerations about future searches and analysis of patent information. The searches and the websites listed in the text or in Tables, together with the cited references, can be used to select patent-related publications and actions to be compared with the sequence of major COVID-19 medical and scientific milestone events, showing how the situation quickly evolved just before the declaration of the COVID-19 pandemic, and for the first 6 months following (Fig. 9).

The quantitative and qualitative analysis of information made by this report, as of November 2020, is not able to predict any future findings in patent literature released after the statutory delay for patent publication (18 months from an earlier, priority date, in general), nor conclude whether or how valuable "COVID-19 specific" datasets and documentary repositories will be still available in the future for patent-related analysis. Already during the preparation and the revision of this article, the contents of quite a number of websites dedicated to COVID-19 research have been aggregated, ceased to be regularly updated, or have simply disappeared already. Given the increasing amount and often rapid obsolescence of information, there is no guarantee that all providers will continue to allocate the time and resources required to keep the databases updated and accessible, thus leading to potential confusion and inconsistency for future technical or patentability assessments. In fact, it has been observed that possibly hundreds of online-only journals may have already "vanished" [86]. Such issues related to stability and availability of scientific publications may affect the prior art evaluation process, and in particular for biomedical technologies, where non-patent literature is often predominant over patent literature as relevant prior art [87,88].

It should also be noticed that, lacking any specific treatment for COVID-19 during the first months after the pandemic outbreak, the only immediately available option was to face the emergency by exploiting possibly incomplete, COVID-19 findings that are generated by using equipment and drugs already commercially available (or otherwise known from virologic or respiratory clinical research). This approach is generally been defined as "Drug Repurposing", but in the present case it may lead to a potential "tsunami" of such projects based on serendipitous observations or more systematic, large scale analysis of candidate compounds [89]. Patentability of findings about known materials, compounds, or other products used to fight against COVID-19 may then be challenged by means of open access preprint repositories or by other approaches summarized in a report published by WIPO Magazine [90]. Indeed, the impact of the flood of non-patent literature and data freely available during 2020 still has to be assessed for the patentability of novel biological or chemical entities that will be claimed in patent applications filed and published from mid-2020 onwards, after pursuing more "normal" and lengthy drug discovery programs.

Another important aspect is how the "frontrunners" having not only filed patent applications earlier but also having had their publications, examinations, and even grants accelerated, will be provided with fully enforceable patent rights. Aside from the pressure of public opinion and governments requiring immediate, "IP rights-free" access to potentially life-saving technologies, some uncertainties about patent granting process are consequent to the complexity of both defining the disclosure date for relevant publications in 2020 and anticipating how pre-2020 disclosures about Coronavirus-related matters in general would be considered relevant for the patentability of claims related to SARS-CoV-2 biology and COVID-19 medical management. Indeed, the complex access to SARS-Cov-2 biological samples and the lack of appropriate validation protocols in the early months of the pandemic would suggest that only a limited number of early patent filings contain reliable and specific data supporting the broadest scope of claims.

On Jan. 5th, 2020, WHO has published what may be considered as the first official report about a pneumonia of unknown etiology detected in Wuhan City, Hubei Province of China [91] and, as of November 2020, factors associated with the SARS-Cov-2 infectivity and origin are still partially known. Since then, COVID-19 has possibly irrevocably changed not only many aspects of daily life and health systems but also economy and innovation, with major efforts are still required to overcome consequences of this pandemic on other clinical conditions or in specific populations. The patent system will not avoid new challenges not only to legal framework but also to attitudes and policies well established among users and stakeholders over the last few decades.

## Funding statement

The authors declare that they have no known competing financial interest or personal relationships which have, or could be perceived to have, influenced the work reported in this article.

## Declaration of competing interest

The authors declare no conflict of interest.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.wpi.2022.102094.

## References

[1] D.M. Morens, A.S. Fauci, Emerging pandemic diseases: how we got to COVID-19, Cell 182 (2020) 1077–1092, https://doi.org/10.1016/j.cell.2020.08.021.
[2] C.M. Kinsella, P.D. Santos, I. Postigo-Hidalgo, A. Folgueiras-González, T. C. Passchier, K.P. Szillat, J.O. Akello, B. Álvarez-Rodríguez, J. Martí-Carreras, Preparedness needs research: how fundamental science and international collaboration accelerated the response to COVID-19, PLoS Pathog. 16 (2020), https://doi.org/10.1371/journal.ppat.1008902 e1008902.
[3] WIPO, WIPO's COVID-19 Response. https://www.wipo.int/covid-19/en/, 2020.
[4] Wu, B, P. Khazin, Patent-related actions taken in WTO members in response to Covid-19 pandemic, World Tour. Barom. (Engl. version) (2020). https://www.wto.org/english/res_e/reser_e/ersd202012_e.pdf.
[5] L.M. Khachigian, 2020. Pharmaceutical patents: reconciling the human right to health with the incentive to invent, Drug Discov. Today 25 (2020) 1135–1141, https://doi.org/10.1016/j.drudis.2020.04.009.
[6] F. Tietze, P. Vimalnath, L. Aristodemou, J. Molloy, Crisis-critical intellectual property: findings from the COVID-19 pandemic. Centre for Technology Management Working Paper Series, Cambridge Univ., 2020 April. https://doi.org/10.17863/CAM.51142.
[7] WHO, Naming the coronavirus disease (COVID-19) and the virus that causes it. https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/naming-the-coronavirus-disease-(covid-2019)-and-the-virus-that-causes-it, 2020.
[8] National Center for Biotechnology Information, MeSH [Supplementary Concept] severe acute respiratory syndrome coronavirus 2. https://www.ncbi.nlm.nih.gov/mesh/2052180, 2020.
[9] National Center for Biotechnology Information, MeSH [supplementary concept] COVID-19. https://www.ncbi.nlm.nih.gov/mesh/2052179, 2020.
[10] EPPI-Centre, Resources relating to COVID-19. https://eppi.ioe.ac.uk/cms/Projects/DepartmentofHealthandSocialCare/Publishedreviews/COVID-19Livingsystematicmapoftheevidence/COVID-19Resources/tabid/3767/Default.aspx, 2021.
[11] CDC, Search request: systematic literature review on COVID-19. https://www.cdc.gov/library/docs/covid19/SearchStrategy_9Oct2020.pdf, 2020.
[12] WHO, COVID-19 Global literature on coronavirus disease. https://search.bvsalud.org/global-literature-on-novel-coronavirus-2019-ncov/, 2021.
[13] S. Gianola, T.S. Jesus, S. Bargeri, G. Castellini, Characteristics of academic publications, preprints, and registered clinical trials on the COVID-19 pandemic, PLoS One 15 (2020), https://doi.org/10.1371/journal.pone.0240123 e0240123.
[14] P. Radanliev, D. De Roure, R. Walton, Data mining and analysis of scientific research data records on Covid-19 mortality, immunity, and vaccine development -

in the first wave of the Covid-19 pandemic, Diabet Metab. Syndr. 14 (2020) 1121–1132, https://doi.org/10.1016/j.dsx.2020.06.063.
[15] L. Zhang, W. Zhao, B. Sun, Y. Huang, W. Glänzel, How scientific research reacts to international public health emergencies: a global analysis of response patterns, Scientometrics 124 (2020) 747–773, https://doi.org/10.1007/s11192-020-03531-4.
[16] J. White, PubMed 2.0, Med. Ref. Serv. Q. 39 (2020) 382–387, https://doi.org/10.1080/02763869.2020.1826228.
[17] NLM, New MeSH supplementary concept record for the 2019 novel coronavirus, wuhan, China. NLM Tech Bull. 432 (2020) b3, https://www.nlm.nih.gov/pubs/techbull/jf20/brief/jf20_mesh_novel_coronavirus.html (see update about "COVID-19 and SARS-CoV-2 MeSH Terms-2021" published in NLM Technical Bulletin on Dec. 4th 2020;, https://www.nlm.nih.gov/pubs/techbull/nd20/nd20_mesh_covid_terms.html.
[18] NLM. NLM announces NIH preprint pilot to provide early access to COVID-19 research, NLM Tech. Bull. 434 (2020) e3, https://www.nlm.nih.gov/pubs/techbull/mj20/mj20_preprint_pilot.html.
[19] J.V. Lazarus, et al., Searching PubMed during the COVID-19 pandemic, J. Med. Internet Res. 22 (2020), https://doi.org/10.2196/23449 e23449.
[20] O. Alhalabi, V. Subbiah, Managing cancer care during the COVID-19 pandemic and beyond, Trends Cancer 6 (2020) 533–535, https://doi.org/10.1016/j.trecan.2020.04.005.
[21] Q. Chen, A. Allot, Z. Lu, LitCovid: an open database of COVID-19 literature, Nucleic Acids Res. (2020), https://doi.org/10.1093/nar/gkaa952.
[22] I.I. Bogoch, et al., Pneumonia of unknown aetiology in Wuhan, China: potential for international spread via commercial air travel, PubMed ID 31943059, https://pubmed.ncbi.nlm.nih.gov/31943059/?format=pubmed, 2020.
[23] I. Bogoch, et al., Pneumonia of unknown aetiology in Wuhan, China: potential for international spread via commercial air travel, J. Trav. Med. 27 (2) (2020) taaa008, https://academic.oup.com/jtm/article/27/2/taaa008/5704418.
[24] The Lens, COVID-19 data set. https://about.lens.org/covid-19, 2020.
[25] B. CDC-Stephen, Thacker Library Branches, Covid-19 research articles database. https://www.cdc.gov/library/researchguides/2019novelcoronavirus/researcharticles.html, 2020.
[26] B.S. Santos, et al., COVID-19: a scholarly production dataset report for research analysis, Data Brief 32 (2020) 106178, https://doi.org/10.1016/j.dib.2020.106178.
[27] bioRxiv, A list of COVID-19 SARS-CoV-2 preprints from medRxiv and bioRxiv. https://connect.biorxiv.org/relate/content/181, 2020.
[28] PLOS, COVID-19 pandemic (2019-21). https://collections.plos.org/collection/covid-19, 2020.
[29] Retraction Watch, Retracted coronavirus (COVID-19) papers. https://retractionwatch.com/retracted-coronavirus-covid-19-papers, 2020.
[30] L. Falciola, Pharma & biotech patent searching in public scientific databases in: CEPIUG 10th year anniversary conference, in: CEPIUG 10th Year Anniversary Conference, Milan (Italy), CEPIUG, 2018.
[31] Science AAAS, Coronavirus: research, commentary, and news. https://www.sciencemag.org/tags/coronavirus, 2020.
[32] Nature, SARS-CoV-2-Latest research and news. https://www.nature.com/subjects/sars-cov-2, 2020.
[33] Cell Press, COVID-19 content archive. https://www.cell.com/covid-19-archive.
[34] JAMA network, Coronavirus disease 2019 (COVID-19). https://jamanetwork.com/journals/jama/pages/coronavirus-alert, 2020.
[35] The BMJ, BMJ's Coronavirus (covid-19) Hub. https://www.bmj.com/coronavirus, 2020.
[36] The Economist, Our coverage of the coronavirus. https://www.economist.com/coronavirus-pandemic, 2020.
[37] The Scientist, Covid 19 news, articles. https://www.the-scientist.com/tag/covid-19, 2020.
[38] BioWorld, Coronavirus - free access to BioWorld coronavirus articles. https://www.bioworld.com/articles/topic/517, 2020.
[39] Clinicaltrialsgov, Views of COVID-19 studies listed on ClinicalTrials.gov (beta). https://clinicaltrials.gov/ct2/covid_view, 2020.
[40] Clinicaltrialsgov, COVID-19 search Strategy. https://clinicaltrials.gov/ct2/results?cond=COVID-19, 2020.
[41] WHO, International clinical trials registry platform (ICTRP). https://www.who.int/clinical-trials-registry-platform, 2020.
[42] FDA, Coronavirus (COVID-19) drugs. https://www.fda.gov/drugs/emergency-preparedness-drugs/coronavirus-covid-19-drugs, 2020.
[43] Milken Institute covid-19 treatment and vaccine tracker. https://covid-19tracker.milkeninstitute.org/#treatment_antibodies.
[44] NLM, Data filters in PMC and PubMed, NLM Tech. Bull. 421 (2018) b8, https://www.nlm.nih.gov/pubs/techbull/ma18/brief/ma18_pmc_data_filters.html.
[45] K. Gao, D.D. Nguyen, J. Chen, R. Wang, G. Wei, Repositioning of 8565 existing drugs for COVID-19, J. Phys. Chem. Lett. (2020) 5373–5382, https://doi.org/10.1021/acs.jpclett.0c01579, 11 13.
[46] L. Lu Wang, et al., CORD-19: the Covid-19 Open Research Dataset, ArXiv, 2020. https://arxiv.org/abs/2004.10706v2.
[47] WIPO, COVID-19 IP policy tracker. https://www.wipo.int/covid19-policy-tracker/#/covid19-policy-tracker/ipo-operations, 2020.
[48] WIPO, WIPO pearl, COVID-19 glossary. https://wipopearl.wipo.int/en/covid19, 2020.
[49] WIPO, Patentscope COVID-19 INDEX. https://patentscope.wipo.int/search/en/covid19.jsf, 2020.
[50] J. List, Intellectual Property information in the time of corona virus, World Patent Inf. 61 (2020) 101971, https://doi.org/10.1016/j.wpi.2020.101971.
[51] N.S. Clarke, 2020. Fighting coronavirus – background to the EPO's response, in: EPOPIC, EPO, 2020. http://documents.epo.org/projects/babylon/eponot.nsf/0/9150186096F3FE21C125861F002907E7/$File/EPOPIC_2020_Nigel_Clarke.pdf.
[52] Joint Message from the USPTO and the JPO, For the future of innovation. https://www.uspto.gov/about-us/news-updates/joint-message-uspto-and-jpo-future-innovation, 2020.
[53] USPTO, USPTO Covid-19 response resource center. https://www.uspto.gov/coronavirus/uspto-covid-19-response-resource-center, 2020.
[54] J.A.C. Nascimento Junior, et al., SARS, MERS and SARS-CoV-2 (COVID-19) treatment: a patent review, Expert Opin. Ther. Pat. 30 (2020) 567–579, https://doi.org/10.1080/13543776.2020.1772231.
[55] F. Machuca-Martinez, R.C. Amado, O. Gutierrez, Coronaviruses: a patent dataset report for research and development (R&D) analysis, Data Brief 30 (2020) 105551, https://doi.org/10.1016/j.dib.2020.105551.
[56] C. Liu, et al., Research and development on therapeutic agents and vaccines for COVID-19 and related human coronavirus diseases, ACS Cent. Sci. 6 (2020) 315–331, https://doi.org/10.1021/acscentsci.0c00272.
[57] Q.A. Zhou, et al., Potential therapeutic agents and associated bioassay data for COVID-19 and related human coronavirus infections, ACS Pharmacol Transl Sci 3 (2020) 813–834, https://doi.org/10.1021/acsptsci.0c00074.
[58] Anonymous, Recent patents related to vaccines and methods of treatment of coronaviruses Nature Biotechnology, 38, 2020, p. 695, https://doi.org/10.1038/s41587-020-0559-3.
[59] P. Musyuni, G. Aggarwal, M. Nagpal, R.K. Goyal, A case study: analysis of patents of coronaviruses and covid-19 for technology assessment and future research, Curr. Pharmaceut. Des. (2020), https://doi.org/10.2174/1381612826666200720233947.
[60] S.A.M. Khalifa, et al., Screening for natural and derived bio-active compounds in preclinical and clinical studies: one of the frontlines of fighting the coronaviruses pandemic, Phytomedicine (2020) 153311, https://doi.org/10.1016/j.phymed.2020.153311.
[61] K.L. Yang, et al., [Analysis of traditional Chinese medicine from patent information sharing platform of coronavirus disease 2019 (COVID-19)], Zhongguo Zhongyao Zazhi 45 (2020) 3001–3006, https://doi.org/10.19540/j.cnki.cjcmm.20200330.501.
[62] E. Ruiz-Hitzky, et al., Nanotechnology responses to COVID-19, Adv Healthc Mater (2020), https://doi.org/10.1002/adhm.202000979 e2000979.
[63] B. Jürgens, N. Clarke, Study and comparison of the unique selling propositions (USPs) of free-to-use multinational patent search systems, World Patent Inf. 52 (2018) 9–16, https://doi.org/10.1016/j.wpi.2018.01.001.
[64] EPO, Fighting coronavirus. https://www.epo.org/news-events/in-focus/fighting-coronavirus.html, 2020.
[65] WIPO, Types of protection available via the PCT in PCT contracting states. https://www.wipo.int/pct/en/texts/typesprotection.html, 2020.
[66] J. Gu, Spatiotemporal dynamics of the patent race: empirical evidence from listed companies in China, Asian J. Technol. Innovat. (2020), https://doi.org/10.1080/19761597.2020.1830813.
[67] G. Zhang, L. Xiong, H. Duan, D. Huang, Obtaining certainty vs. creating uncertainty: does firms' patent filing strategy work as expected? Technological Forecasting and Social Change, 160, 2020, p. 120234, https://doi.org/10.1016/j.techfore.2020.120234.
[68] G. de Rassenfosse, E. Raiteri, Technology protectionism and the patent system: strategic technologies in China, SSRN (2020), https://doi.org/10.2139/ssrn.2803379.
[69] CNIPA (China National Intellectual Property Administration), Vaccine against coronavirus is granted. http://english.cnipa.gov.cn/art/2020/8/19/art_1347_151000.html, 2020.
[70] H. Yu, H. Yang, E. Shi, W. Tang, Development and clinical application of phosphorus-containing drugs med drug discov, 8, 2020, p. 100063, https://doi.org/10.1016/j.medidd.2020.100063.
[71] K. Brimacombe, et al., An OpenData Portal to Share COVID-19 Drug Repurposing Data in Real Time, BioRxiv, 2020, p. 135046. https://www.biorxiv.org/content/10.1101/2020.06.04.135046v1.
[72] EBI, The COVID-19 data portal. https://www.covid19dataportal.org, 2020.
[73] S. Kim, et al., 2020. PubChem in 2021: new data content and improved web interfaces, Nucleic Acids Res. (2020), https://doi.org/10.1093/nar/gkaa971.
[74] A. Gaulton, ChEMBL_27 SARS-CoV-2 release. http://chembl.blogspot.com/2020/05/chembl27-sars-cov-2-release.html, 2020.
[75] EBI - Uniprot, COVID-19 UniProtKB. https://covid-19.uniprot.org/uniprotkb?query=*, 2020.
[76] EBI, Ensembl COVID-19. https://covid-19.ensembl.org/index.html, 2020.
[77] Lens, Human coronaviruses data initiative. https://about.lens.org/covid-19/, 2020.
[78] M.I.J. Raybould, A.A. Kovaltsuk, C. Marks, C.M. Deane, CoV-AbDab: the coronavirus antibody database, Bioinformatics (2020), https://doi.org/10.1093/bioinformatics/btaa739 btaa73.
[79] Virologicalorg, Novel 2019 coronavirus genome. https://virological.org/t/novel-2019-coronavirus-genome/319, 2020.
[80] Y.A. Helmy, M. Fawzy, A. Elaswad, A.S.P. Sobieh, A.A. Kenney, Shehata, The COVID-19 pandemic: a comprehensive review of taxonomy, genetics, epidemiology, diagnosis, treatment, and Control, J. Clin. Med. 9 (2020) 1225, https://doi.org/10.3390/jcm9041225.
[81] L. Chen, L. Zhong, Genomics functional analysis and drug screening of SARS-CoV-2, Genes Dis (2020), https://doi.org/10.1016/j.gendis.2020.04.002.
[82] WHO, SARS-CoV-2 variants, working definitions and actions taken. https://www.who.int/en/activities/tracking-SARS-CoV-2-variants/, 2020.

[83] Y. Finkel, et al., The coding capacity of SARS-CoV-2, Nature (2020), https://doi.org/10.1038/s41586-020-2739-1.
[84] A.K. Banerjee, et al., SARS-CoV-2 disrupts splicing, translation, and protein trafficking to suppress host defenses, Cell 183 (2020) 1325–1339, https://doi.org/10.1016/j.cell.2020.10.004.
[85] H. Yao, et al., Molecular architecture of the SARS-CoV-2 virus, Cell 183 (2020) 730–738, https://doi.org/10.1016/j.cell.2020.09.018, e13.
[86] J. Brainard, Dozens of scientific journals have vanished from the internet, and no one preserved them. https://www.sciencemag.org/news/2020/09/dozens-scientific-journals-have-vanished-internet-and-no-one-preserved-them, 2020.
[87] Q. Ke, An analysis of the evolution of science-technology linkage in biomedicine, J Inform 14 (2020) 101074, https://doi.org/10.1016/j.joi.2020.101074.
[88] O.A. Jefferson, A. Jaffe, D. Ashton, B. Warren, D. Koellhofer, U. Dulleck, A. Ballagh, J. Moe, M. DiCuccio, K. Ward, G. Bilder, K. Dolby, R.A. Jefferson, Mapping the global influence of published research on industry and innovation, Nat. Biotechnol. 36 (2018) 31–39, https://doi.org/10.1038/nbt.4049.
[89] H.A.M. Mucke, COVID-19 and the drug repurposing tsunami, Assay Drug Dev. Technol. 18 (2020) 211–214, https://doi.org/10.1089/adt.2020.996, 2020.
[90] J. Nurton, Drug repurposing and the COVID-19 pandemic, WIPO Magazine 2 (2020). https://www.wipo.int/wipo_magazine/en/2020/02/article_0004.html.
[91] WHO, Pneumonia of unknown cause – China. https://www.who.int/csr/don/05-january-2020-pneumonia-of-unkown-cause-china/en/, 2020.