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.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Integrating heterogeneous data to facilitate COVID-19 drug repurposing

Lucía Prieto Santamaría\textsuperscript{a,b,c,1}, Marina Díaz Uzquiano\textsuperscript{b,1}, Esther Ugarte Carro\textsuperscript{b}, Nieves Ortiz-Roldán\textsuperscript{d}, Yuliana Pérez Gallardo\textsuperscript{c}, Alejandro Rodríguez-González\textsuperscript{a,b,*}

\textsuperscript{a}ETS Ingenieros Informáticos, Universidad Politécnica de Madrid, 28660 Boadilla del Monte, Madrid, Spain
\textsuperscript{b}Centro de Tecnología Biomédica, Universidad Politécnica de Madrid, 28660 Boadilla del Monte, Madrid, Spain
\textsuperscript{c}Ezeris Networks Global Services S.L., 28028 Madrid, Spain
\textsuperscript{d}Facultativo Especialista Área (FEA), Anestesiología y Reanimación, Hospital Sierrallana, Servicio Cántabro de Salud, 39300 Torrelavega, Cantabria, Spain

In the COVID-19 pandemic, drug repositioning has presented itself as an alternative to the time-consuming process of generating new drugs. This review describes a drug repurposing process that is based on a new data-driven approach: we put forward five information paths that associate COVID-19-related genes and COVID-19 symptoms with drugs that directly target these gene products, that target the symptoms or that treat diseases that are symptomatically or genetically similar to COVID-19. The intersection of the five information paths results in a list of 13 drugs that we suggest as potential candidates against COVID-19. In addition, we have found information in published studies and in clinical trials that support the therapeutic potential of the drugs in our final list.

Keywords: COVID-19; SARS-COV-2; Drug repurposing; Data-driven approaches

Introduction

The novel coronavirus SARS-CoV-2 emerged in December 2019, and since then, there has been an increasing interest in finding new ways of treating COVID-19. This is not a new scenario: three highly pathogenic coronaviruses, SARS-CoV, MERS-CoV, and SARS-CoV-2, have emerged within the past 20 years. Nonetheless, there are currently no approved drugs tailored against SARS-CoV-2, nor for human coronaviruses in general.\textsuperscript{1,2}

These pathogen emergences are an urgent situation in which drug development seems to be an inefficient option. There is, however, the possibility of reusing already-approved drugs that can help to treat the disease. This paper describes and validates drug-repurposing (DR) methods using an existing medical data platform, DISNET,\textsuperscript{3} which holds information about the relations between diseases and symptoms, symptoms and drugs, diseases and genes, and drugs and targets. The DR methods that we used consider the gene associations and symptomatology of a disease of interest in order to find reusable drugs. In this case study, the disease of interest is COVID-19.

Our objective was to assess the different repositioning procedures that can be derived from DISNET, execute them for COVID-19, and explore the current literature and ongoing clinical trials relating to COVID-19, in order to validate our results and methodology.

Thus, we present a simple and straightforward pipeline that uses DISNET to unveil potential drugs that can be repurposed to treat a particular disease, in our case, COVID-19. It is a straightforward method because DISNET presents the data necessary to describe disease–drug, symptom–drug, and target–drug associations, all summarized in one platform. It is a simple method because it can be easily re-done to find repurposable drugs for diseases other than COVID-19, for instance, potential future coronavirus emergences.
Data-driven COVID-19 DR approaches
The proposed analysis intends to provide a list of already-existing and approved drugs that could show promise in treating COVID-19. With this aim, we studied five different paths to explore the relationships described by DISNET information as a means to find repurposable drugs:

- Path 1: COVID – symptoms – drugs. We identified COVID-19 symptoms, and then obtained the drugs that are directly indicated for those symptoms.
- Path 2: COVID – symptoms – diseases – drugs. We identified COVID-19 symptoms, then extracted the diseases that are associated with those symptoms, and finally the drugs that are used to treat such diseases.
- Path 3: COVID – symptoms – diseases – genes – targets – drugs. We identified COVID-19 symptoms, then obtained the diseases that had symptoms in common with COVID-19. We then extracted the genes related to such diseases. Afterwards, the targets associated with those genes and, finally, drugs related to those targets were identified.
- Path 4: COVID – genes – diseases – drugs. Having identified genes that are associated with COVID-19, we obtained diseases that have gene associations in common with COVID-19. Drugs that are indicated for those diseases were then identified.
- Path 5: COVID – genes – targets – drugs. Having identified COVID-19-associated genes, we extracted the proteins that are associated with these genes. These proteins were then considered as drug targets, and the drugs that are associated with these targets were obtained.

A graphical representation of the different pathways used in the analysis is presented in Fig. 1.

All information used was provided by the DISNET platform, queried in January 2021. Drugs in the list resulting from the intersection of the five paths were further analyzed: the drugs were validated by searching both the literature and clinical trials data for evidence and hypothesized mechanisms to support their potential to be repurposed for COVID-19.

Data acquisition and integration
All data employed in this work were obtained from the DISNET project database. DISNET aims to extract biomedical knowledge related to diseases, including (but not limited to) symptoms, genes, and drugs, by mining and querying public sources. The disease information is stored in a heterogeneous database that is structured in three levels: the phenotypical layer (with mainly disease–symptoms associations), the biological layer (containing the associations of diseases with genes and proteins, among other associations), and the drugs layer (which stores drug-related data, including the associations of drugs with diseases). The integration of this variety of information makes it possible to discover patterns and uncover knowledge that would otherwise remain unknown.

Intermediate data files providing further results can be accessed in the public repository specified in the Data availability section. Summaries of the data used are included in two tables in the Appendix A. Supplementary data section: SM1—Supplementary Table 1 provides details of data used to describe entities (diseases, symptoms, genes, proteins and...
targets, drugs, and COVID-19 symptoms and genes), whereas SM1—Supplementary Table 2 provides insights into the data used to describe relationships (disease—symptom, disease—gene, gene—protein, drug—disease, drug—target, and drug—symptom associations). Both tables cover data descriptions, counts, identifiers, provenance, source, and date of access for each layer when possible. Given the heterogeneity, diversity, and large number of data sources in the life sciences domain, COVID-19 has been differently represented and codified in multiple terminologies. We have included several entities that have been used to represent COVID-19 in different codifications (SM1—Supplementary Table 3). All of these instances have been treated as a single entity when considering COVID-19 in DISNET.

Diseases are represented in the DISNET database by different codifications, depending on the layer and source. When the analysis was performed, there was a total of 9,225 different diseases in the phenotypical layer (codified by DISNET’s own identifiers, 24,314 in the biological layer (identified by Unified Medical Language System (UMLS) Concept Unique Identifiers (CUIs)), and 9,116 in the drugs layer (represented by UMLS CUIs).

Drugs and drug information (such as the molecular type, the chemical structure, or the International Chemical Identifier (InChI) key) were extracted from ChEMBL4–5 (https://www.ebi.ac.uk/chembl/) in May 2020. There were 3,944 different drugs codified by ChEMBL identifiers. Of those, 1,821 drugs were related to the 9,116 diseases identified in the drugs layer. Such associations were obtained from CTD6 (The Comparative Toxicogenomics Database, http://ctdbase.org/). Symptoms were extracted from the mining of textual sources. MetaMap version 2016v2 was the natural language processing (NLP) tool utilized for this task. For more details about its configuration and the source specifications, further explanations can be checked at http://disnet.ctb.upm.es/3,5 A total of 2,248 symptoms and 211,362 disease—symptom associations were available when the analysis was performed.

Three textual sources were mined to identify COVID-19 symptoms: Wikipedia (https://www.wikipedia.org), the European Centre for Disease Prevention and Control, (ECDC; https://www.ecdc.europa.eu/en) and the Mayo Clinic (https://www.mayoclinic.org/). COVID-19 documents from the Mayo Clinic were monitored through time and represented in the different DISNET snapshots captured from March 1, 2020 to February 1, 2021. Three documents from Wikipedia and the ECDC were specifically mined in December 2020. Their web addresses are: https://en.wikipedia.org/wiki/Coronavirus_disease_2019, https://en.wikipedia.org/wiki/Symptoms_of_COVID-19, and https://www.ecdc.europa.eu/en/covid-19/latest-evidence/clinical.

A total of 76 different COVID-19 symptoms were identified, with Wikipedia texts providing 37 symptoms, ECDC documentation 23, and the Mayo Clinic sources 52. Of the 76 symptoms, 12 were found in all three sources. For the symptoms associated with other diseases, DISNET contains information from three sources—Wikipedia, Mayo Clinic, and PubMed—mined at different times as explained by Lagunes-García et al.4 The current analysis was carried out with the information available in all DISNET snapshots up to February 1, 2021. In addition, symptoms were also included in DISNET’s drug layer as indications for each drug. The relationships between drugs and their indicated symptoms were extracted from ChEMBL in May 2020. A total of 10,475 drug—symptoms associations (involving 1,857 different drugs and 851 different symptoms) were recorded in DISNET.

Genes and gene information (National Center for Biotechnology Information (NCBI) identifiers, names, symbols, and so on), as well as disease—gene associations (DGAs), were extracted from DisGeNET8 (https://www.disgenet.org/) in May 2020. The extracted data included 20,610 genes and 358,209 DGAs.

COVID-19 genes were downloaded from DisGeNET COVID-19 Data (https://www.disgenet.org/downloads) in November 2020. A total of 75 genes were stored in DISNET’s knowledge base. Proteins encoded by these genes were extracted from DisGeNET in May 2020. They were identified by UniProt9 (https://www.uniprot.org/) accession numbers. In the DISNET biological layer, there was a total of 18,521 proteins and 15,770 gene—protein associations. In addition, proteins that featured as targets in DISNET’s drug layer were extracted from ChEMBL in May 2020. Associations between drugs and targets were obtained from both ChEMBL and DrugBank10 (https://www.drugbank.com/). There were 1,594 targets, 7,727 different drug—target associations and 2,874 drugs associated with available targets.

DISNET DR information paths

The inputs for the current analysis were the symptoms or genes related to COVID-19 mentioned above. Information paths 1, 2, and 3 started from COVID-19 symptoms, whereas paths 4 and 5 used COVID-19 genes as the entry point (Fig. 1). The drug lists derived from each path were obtained as output. The following information paths are described below.

- Path 1: COVID – symptoms – drugs. Of the 76 COVID-19 symptoms identified with UMLS CUIs, two mapped to phenotypic effects that were considered to be indications for ChEMBL drugs, particularly ‘Insulin-dependent diabetes mellitus’ and ‘Infection’.
- Path 2: COVID – symptoms – disease – drugs. The 76 COVID-19 symptoms were associated with other 2,630 diseases, which were in turn associated with 1,806 different drugs. Examples of diseases that are phenotypically related to COVID-19 (that is, that have symptoms in common with COVID-19) would be ‘Pulmonary edema’, ‘Pneumonia’, ‘Pericarditis’, ‘Sarcoidosis’, ‘Lyme disease’, ‘Cardiac tamponade’, ‘Leptospirosis’, ‘Churg-Strauss syndrome’, ‘Kawasaki disease’ or ‘Hypothyroidism’. From now on, these diseases will be referred to as DPCs (Diseases Phenotypically similar to COVID-19). The drug that was associated with the largest number of DPCs was ‘Valproic acid’ followed by ‘Acetaminophen’.
- Path 3: COVID – symptoms – diseases – genes – targets – drugs. In the third information path (P3 in Fig. 1), which starts from the 76 COVID-19 symptoms, a total of 2,395 DPCs were associated with 1,209 different genes, which were in turn targets for 2,177 drugs. ‘Fostamatinib’ was the drug associated with the largest number of DPC-related targets (260 targets).
Validation with clinical trial data
In order to validate these DR paths, we compared the drugs obtained from the five pathways with those currently under testing in clinical trials for COVID-19. This comparison would validate our DR approach if it covered drugs that are thought to treat or evidenced to treat COVID-19 effectively, based not only on computational methods but also on empirical knowledge. That is, our DR approach would be validated if it predicted medications that appeared in the clinical trials and that therefore had some empirical evidence to support their efficacy in treating COVID-19. The drug trials were obtained from the DrugBank COVID-19 clinical trials section (https://go.drugbank.com/indications/DBCOND0126697#drug-trials). By the end of February 2021, there were 1,654 clinical trials for COVID-19, studying 712 different drugs.

DrugBank considers two types of drugs: small molecule drugs and biotech drugs. However, the drugs from DISNET’s database that were considered in this study were those defined by a DrugBank code and a CHEMBL identifier simultaneously, excluding some compounds that DrugBank but not DISNET may consider as to be drugs. Of the 712 small molecules and biotech drugs that were being tested for COVID-19, 344 met DISNET’s criteria.

Potential drugs for COVID-19
Diseases phenotypically related to COVID-19 (DPCs) and diseases biologically related to COVID-19 (DBC) were studied for their similarities with COVID-19, which derived new avenues in DR possibilities. The mean numbers of DPC-related-symptoms were obtained for each of the disease categories that has symptoms similar to those of COVID-19. Likewise, the mean numbers of DBC-related-genes helped to reveal the disease categories that were genetically similar to COVID-19. DPCs and DBCs were classified according to ICD-10-CM (International Classification of Diseases 10 Clinical Modification, https://www.cdc.gov/nchs/icd/icd10cm.htm) (Fig. 2).

From a phenotypical standpoint, the classes that seemed to be most tightly related to COVID-19 (that is, that presented higher mean numbers of COVID-19-related symptoms) were ‘E00-E90: Endocrine, nutritional and metabolic diseases’, ‘A00-B99: Certain infectious and parasitic diseases’, ‘I00-I99: Diseases of the circulatory system’ and ‘K00-K93: Diseases of the digestive system’. These classes contained the most symptoms that were shared between DPCs and COVID-19. From a biological standpoint, ICD classes ‘C00-D48: Neoplasms’, ‘I00-I99: Diseases of the circulatory system’, ‘K00-K93: Diseases of the digestive system’ and ‘A00-B99: Certain infectious and parasitic diseases’ contained the largest number of genes that were shared between DBCs and COVID-19.

These results showed that COVID-19 is more likely be treated by drugs that are indicated for diseases in ICD classes A00-B99 (Infectious), I00-I99 (Circulatory), or K00-K93 (Digestive), as these classes share the greatest number of symptoms and genes with COVID-19. Furthermore, the fact that COVID-19 mostly shared genes with diseases classified as C00-D48 (Neoplasms), suggests that COVID-19 and cancer could have some similarities from a molecular and immunological standpoint.

The drug lists that resulted from each of the five repurposing paths had 13 drugs in common (Table 1), meaning that these drugs are used to: treat COVID-19 symptoms, treat diseases similar to COVID-19 (similar in terms of common symptoms and in terms of common genes), and target proteins associated with the COVID-19 disease, as well as target proteins associated with diseases similar to COVID-19 in terms of symptomatology. For more details on the different intersections between the DR paths, the Appendix A. Supplementary data section (SM 2—Path intersections) provides further analysis.

The 13 drugs identified by all of the repurposing pathways are classified into five categories according to the MeSH-pharmacological action (PA) therapeutic uses Table 1: ‘Anti-infective’, ‘Cardiovascular’, ‘Antineoplastic’, ‘Respiratory System’ or ‘Antirheumatic’ agents’. The most repeated classes for these drugs were ‘Anti-infective agents’ and ‘Cardiovascular agents’, confirming the potential of these drug classes for reuse according to the virus and the disease’s known nature and biology. In addition, the presence of ‘Antineoplastic agents’ supports the previously mentioned hypothesis that COVID-19 shares some commonalities with cancer in that both can extend from a primary point to other anatomic locations2 and, in some cases, can be mediated similarly by the immune system. Treating COVID-19 with drugs that have antineoplastic properties has been an important study subject in recent months.

The distribution of the numbers of symptoms and genes related to DPCs and DBCs, respectively, are shown in Fig. 3. The average number of in DPC-associated symptoms was 7.88, whereas the mean number of DBC-associated genes was 4.26. In the case of DBC-associated gene numbers, the distribution was centered around the mode (which was 2), meaning that most DBCs shared a small number of genes with COVID-19. In the case of DPC-associated symptoms, the distribution is more dispersed (DPCs tended to share higher numbers of symptoms with COVID-19). The 13 drugs that intersected in the five paths were associated with at least one DPC and DBC, situated in percentile 0.955 of the explained symptoms distribution and percentile 0.925 of the explained genes distribution, respectively. Thus, the 13 drugs are indicated for at least one DPC and one DBC.
DBC with the greatest numbers of symptoms and genes shared with COVID-19.

Moreover, and corroborating the insights of the present work, the 13 drugs have already been tested or hypothesized as potential treatments for COVID-19. In the rest of this section, we discuss how these drugs have already been related to COVID-19 in the literature. Further explanations of each drug’s mechanism of action are included in the Appendix A. Supplementary data section (SM 3–Intersecting drugs mechanisms of action).

**Aldesleukin**

Aldesleukin is a cytokine (interleukin 2 (IL2)) used to boost the immune system for the treatment of some types of kidney and skin cancer. DISNET associated the IL2 receptor alpha gene (IL2RA) (along with other IL2 receptors) with COVID-19 and with DPCs.

Some studies suggest the use of immunotherapy to treat the heightened inflammatory response and the pulmonary complications that SARS-CoV-2 produces in some patients.\(^{14,15}\) IL2 can be a suitable immunomodulator for improving the state of COVID-
FIGURE 3
Distributions of DPC-associated symptoms and DBC-associated genes. On the left is the boxplot for the 2,630 DPC-associated symptoms, having a mean of 7.88 and a maximum of 41. On the right is the boxplot for the 1,943 DBC-associated genes, having a mean of 4.26 and a maximum of 42.

19 patients as it induces the proliferation and maturation of cells (regulatory T cells) that play a role in the control of excessive inflammation.15 There is currently a clinical trial that aims to demonstrate the efficacy of low-dose IL2 in COVID-19 patients who develop acute respiratory distress syndrome.16

Cefazolin
Not much research has been done on the use of the antibiotic cefazolin for COVID-19. It has appeared only as a candidate SARS-CoV-2 inhibitor in a study that used chemoinformatics and machine learning to detect FDA-approved drugs that could target essential SARS-CoV-2 structures.17 DISNET data showed a relationship between cefazolin and the SARS-CoV-2-associated gene, IL2.

The main use of cefazolin is to kill bacteria by disrupting their cell wall. However, recent studies have shown in-silico and in-vitro evidence that cefazolin can bind to some cytokine receptors (including the IL2 receptor), and therefore inhibits the proliferation and differentiation of immune cells, and the production of cytokines.18,19 This could help to mitigate the effects of the excessive immune response that SARS-CoV-2 induces in some patients. In addition, cefazolin could treat secondary bacterial pneumonia induced by COVID-19.

Epinephrine
Epinephrine is usually employed in the treatment of severe allergic reactions. It is often used for the treatment of anaphylaxis, and as a second-line agent in septic shock, acute asthma, cardiac arrest, and cardiopulmonary resuscitation. There is currently not much evidence as to whether the use of epinephrine would be beneficial for mitigating COVID-19 symptoms.20,21

Everolimus and sirolimus
Everolimus and sirolimus (or rapamycin) are mTOR inhibitors, a group of immunosuppressants that are used in transplant patients. DISNET associated mTOR with COVID-19 disease and with these two drugs. In the literature, it has been proposed that mTOR inhibitors can reduce the severity of COVID-19 infection.22,23 A small number of studies have produced clinical evidence to suggest that everolimus may be beneficial to hospitalized transplant patients who are suffering from COVID-19.24,25 Other studies have proposed that rapamycin should be used for its ability to inhibit protein synthesis and the synthesis of pro-inflammatory molecules, blocking viral replication and the cytokine storm that occurs in acute COVID-19 infections.26,27 In addition, a network-based DR study found that sirolimus is a potential anti-SARS-CoV-2 drug because it targets the human–coronavirus interactome.28

Hydroxychloroquine
Hydroxychloroquine has shown conflicting results when used in the treatment of COVID-19.29 This drug was first evaluated for use against SARS-CoV-2 because it is a less toxic derivative of chloroquine, a drug that had demonstrated antiviral effects against SARS-CoV and MERS-CoV.30 However, accumulating data from controlled trials suggest that neither chloroquine nor hydroxychloroquine provides a clinical benefit for patients with COVID-19. In a randomized, blinded, placebo-controlled trial of 479 hospitalized patients with COVID-19, hydroxychloroquine did not improve 14-day clinical status or 28-day mortality.31 In the DISNET data, hydroxychloroquine is associated with the coronavirus-related gene, ACE2, which encodes the ACE2 protein used by the virus to enter the host cell.

Angiotensin II receptor Blockers (ARBs) and Angiotensin Converting Enzyme inhibitors (ACEi)
Losartan and candesartan cilexetil are classified as Angiotensin II Receptor Blockers (ARBs), drugs that are used to treat high blood pressure and heart failure. The same biological effect is produced by another class of medications, Angiotensin Converting Enzyme inhibitors (ACEi), which include enalapril and ramipril. ACEis and ARBs are notably used in diabetic and in hypertensive patients.

According to the DISNET data, COVID-19 and many DPCs are associated with AGTR1 (Angiotensin II Receptor Type 1) and ACE (Angiotensin I Converting Enzyme), which are targets for ARB and ACEi proteins, respectively.

There has been some controversy over the use of these types of medications to treat COVID-19.32 On the one hand, some studies in animal models have shown that ACEis and ARBs upregulate the production of ACE2.33,34 In addition, the upregulation of ACE2 has been shown to be protective against acute respiratory distress syndrome and acute lung injury.35–37 On the other hand, because SARS-CoV-2 uses ACE2 to enter the host cell,38 it has been hypothesized that increased ACE2 expression may produce a higher risk of COVID-19 infection.39 However, this risk has not proven to be an issue in a number of clinical studies reviewed by Liu et al.,40 and there is no elevated risk of infection in patients undergoing ACEi/ARB therapy. Furthermore, the same review showed that patients with hypertension who were receiving ACEi/ARB treatment were linked to a lower mortality rate than their non-receiving counterparts with hypertension.40
More specifically, a study on ARBs proved that these molecules are safe and effective against SARS-CoV-2: molecular-docking and molecular-dynamics evidence for candesartan and losartan, as well as in-vitro evidence for losartan, showed that these are suitable new drugs against SARS-CoV-2. Moreover, clandesartan cilexetil has been studied further for its potential to downregulate the cytokine storm that SARS-CoV-2 induces in some patients.41

Minocycline
Minocycline, a tetracycline antibiotic, is another drug that has been studied repeatedly as a COVID-19 DR.42–45 Tetracyclines have proven to be efficient against other positive-sense RNA viruses,46–48 and they have anti-inflammatory properties that could inhibit the inflammatory response in acute SARS-CoV-2 infections.44,45,47,48 Furthermore, a combinatorial computational study has found minocycline to be a potent inhibitor of the SARS-CoV-2 main protease.50 In the DISNET data, minocycline was indicated for the positive RNA-virus-borne disease West Nile Fever, which shares 23 out of the 76 symptoms associated with COVID-19.

Sitagliptin and vildagliptin
Sitagliptin and vildagliptin are inhibitors of the DPP-4 protein (DPP-4i), which are used as anti-diabetic drugs for type-2 diabetes. DPP-4 is a ubiquitous enzyme that is expressed on the surface of most cell types. It deactivates a variety of other bioactive peptides, including glucose-dependent insulinotrophic polypeptide and glucagon-like peptide 1; therefore, its inhibition could potentially affect glucose regulation through multiple effects.

The data from DISNET showed DPP4–COVID-19 and DPP4–DPCs associations. DISNET data also showed that the diseases ‘Type 2 diabetes’ and ‘Diabetes mellitus type 2’ shared around 14% of symptoms with COVID-19 (11 and 10 symptoms, respectively, out of a total of 76). More generally, the diseases ‘Diabetes mellitus’ and ‘Diabetes’ shared around a 27% of symptoms with COVID-19 (22 and 20 symptoms, respectively). Some studies have concluded that the benefits of DPP-4i drugs for the treatment of COVID-19 are unclear.51,52 However, a small study showed that diabetic patients who used DPP-4i had a lower mortality rate and were less likely to need non-invasive mechanical ventilation than non-DPP-4i diabetic users.53

More specifically, the DPP-4i sitagliptin has been proposed as a potential treatment for COVID-19 in a couple of computational DR studies.17,54 In addition, it has been reported that diabetic patients treated with the DPP-4i sitagliptin at the time of hospitalization had a lower mortality rate, better clinical outcomes, and more hospital discharges than their diabetic counterparts not taking sitagliptin.55

The fact that these 13 drugs are being studied or discussed in the literature as potential anti-SARS-CoV-2 treatments demonstrates that DISNET and the suggested paths are reliable DR approaches. Not only that, but comparisons with ongoing clinical trials further validated the reliability of these approaches: of the 344 DISNET drugs undergoing clinical trials, 329 (or 95.64% of the on-trial DISNET-compliant drugs) were identified by our DR methodology. Of these drugs, 100 were identified by path 1, 272 by path 2, 263 by path 3, 272 by path 4, and 56 by path 5.

Out of the 13 final drugs identified by all five pathways, eight were being tested in clinical trials, namely: aldesleukin, ramipril, enalapril, losartan, sirolimus, minocycline, sitagliptin, and hydroxychloroquine. Finding these drugs both in the intersection of the five paths and in the ongoing or concluded trials corroborates the idea that they can potentially be repurposed for the target disease. Furthermore, the five remaining drugs that are present in the intersecting list but not in the clinical trials have a similar molecular underlying process or the same classifications as the eight that are in clinical trials: candesartan cilexetil is a cardiovascular agent classified as an ARB, just like losartan; everolimus is an immunosuppresser mTOR inhibitor, like sirolimus; and vildagliptin is an anti-diabetic drug, as is sitagliptin. Cefazolin is an antibiotic that may target viral structures, whereas epinephrine could be used for some COVID-19 symptoms (although the association with this disease was established in the different paths very nonspecifically). The plausibility of 11 of the 13 intersecting drugs as agents against the virus has been validated by evidence or hypotheses in scientific publications, or by their presence in official clinical trials in the context of COVID-19.

Although not further studied in the current manuscript, intersections between other paths include promising repurposeable drugs for COVID-19. The most important paths are thought to be paths 2, 4, and 5, as they provide drugs that are indicated for the treatment of diseases that are symptomatologically and genetically similar to COVID-19 and that drugs target COVID-19 related genes directly. These three paths converge upon 81 different drugs, 38 of which (46.91%) are present in clinical trials. An example would be baricitinib, which has been studied as a treatment for COVID-19, sometimes in combination with remdesivir, because of its capacity to reduce systemic inflammation and lung injury.56–58

Concluding remarks
This present study proposes a new pipeline consisting of five information paths that identify already-existing drugs that are potentially reusable for COVID-19. These five paths led to five different drug lists, which shared 13 drugs in common, namely: aldesleukin, candesartan cilexetil, cefazolin, enalapril, epinephrine, everolimus, hydroxychloroquine, losartan, minocycline, ramipril, sirolimus, sitagliptin, and vildagliptin. Some of these drugs (for example, hydroxychloroquine) have already been tested as treatments for COVID-19, whereas others (such as ARB and ACEI drugs or DDP-4i drugs) might seem like more remote possibilities. Nevertheless, all 13 drugs have been mentioned throughout the literature as potentially useful treatments against COVID-19.

What is more, of the 13 drugs identified by all five pathways, eight have been or are being directly tested against COVID-19 in clinical trials. The remaining five drugs have very similar mechanisms of action and are classified in the same categories as some of these eight drugs. In addition, 95.64% of all existing drugs in COVID-19 clinical trials have been included in at least one of the five lists resulting from our DR pathways, once again demonstrating the exploitable facets of the present study.
Some additional information that can be derived from this study is that the 13 final drugs were associated with DPCs and DBCs that suggest that these drugs are more likely to belong to the MeSH-PA therapeutic use categories ‘Anti-infective’, ‘Cardiovascular’ and ‘Antineoplastic’.

Notwithstanding these findings, the current work might have some limitations concerning the difficulties of integrating and standardizing data. DISNET has helped to overcome this challenge by providing several integrated sources. Nevertheless, the heterogeneous nature of biomedical data, and the resulting difficulty of combining many sources into one, could have concealed some underlying knowledge that was missed in the current study. Another limitation is that we did not consider combinatorial treatments, which may be effective against COVID-19.\(^ {57,59,60}\)

Future lines of research will include studying the differences that could be observed when considering COVID-19 in adults and infants as two distinct entities, as these two patient groups are thought to have different symptomatology. Moreover, having numerical data to represent symptom prevalence (for example, a score to weight the COVID-19–symptoms relationships) would improve the analysis. Other means to achieve DR might come from machine learning or chemical structures similarity approaches.

As a final note, we can conclude that DISNET has proven to be a successful DR tool, with our results confirming the data obtained from the literature and from official clinical trials.

**Competing interests**

The authors declare no competing interests.

**Acknowledgements**

The work described here is a result of the project “DISNET (Creation and analysis of disease networks for drug repurposing from heterogeneous data sources applied to rare diseases)”, developed under grant RTI2018-094576-A-100 from the Spanish Ministerio de Ciencia, Innovación y Universidades. LPS’s work is supported by the “Programa de fomento de la investigación y la innovación (Doctorados Industriales)” from the Comunidad de Madrid (grant IND2019/TIC-17159).

**Data availability**

All results and files used in the study are accessible in a public repository: https://medal.ctb.upm.es/internal/gitlab/disnet/disnet_covid_paper/.

**Appendix A. Supplementary data**

Supplementary data to this article can be found online at https://doi.org/10.1016/j.drudis.2021.10.002.

**References**

1. X. Zhao, Y. Ding, J. Du, Y. Fan, 2020 update on human coronaviruses: one health, one world, Med Nov Technol Devices 8 (2020), https://doi.org/10.16/j. medn.2020.100043 100043.

2. A. Zumla, J.F.W. Chan, E.I. Azhar, D.S.C. Hui, K.-Y. Yuen, Coronaviruses—drug discovery and therapeutic options, Nat Rev Drug Discov 15 (2016) 327–347, https://doi.org/10.1038/nrd.2015.37.

3. G. Lagunes-García, A. Rodríguez-González, L. Prieto-Santamaría, E.P. García del Valle, M. Zanin, E. Menasalvas-Ruiz, DISNET: a framework for extracting phenotypic disease information from public sources, PeerJ 8 (2020), https:// doi.org/10.7717/peerj.8580 e8580.

4. M. Davies, M. Nowotka, G. Papadatos, N. Dedman, A. Gaulton, F. Atkinson, et al., ChEMBL web services: streamlining access to drug discovery data and utilities, Nucleic Acids Res 43 (2015) W162–W260, https://doi.org/10.1093/nar/gkv352.

5. D. Mendez, A. Gaulton, A.P. Bento, J. Chambers, M. De Veij, E. Félix, et al., ChEMBL: towards direct deposition of bioassay data, Nucleic Acids Res 47 (2019) D930–D940, https://doi.org/10.1093/nar/gky1075.

6. A.P. Davis, C.J. Grondin, R.J. Johnson, D. Scialy, J. Wiegens, T.C. Wiegens, C.J. Mattingly, Comparative Toxicogenomics Database (CTD): update 2021, Nucleic Acids Res 49 (2021) D1138–D1143, https://doi.org/10.1093/nar/gkaa891.

7. G. Lagunes-García, A. Rodríguez-González, L. Prieto-Santamaría, E.P. García del Valle, M. Zanin, E. Menasalvas-Ruiz, How Wikipedia disease information evolve over time? An analysis of disease-based articles changes, Inf Process Manag 57 (2020), https://doi.org/10.1016/j.ipm.2020.102225 102225.

8. J. Pfeifer, J.M. Ramirez-Anguita, J. Saich-Pitarach, F. Ronzano, E. Centeno, F. Sanz, L.I. Furlong, The DisGeNET knowledge platform for disease genomics: 2019 update, Nucleic Acids Res 48 (2020) D845–D855, https://doi.org/10.1093/nar/gkz1021.

9. UniProt Consortium, UniProt: the universal protein knowledgebase in 2021, Nucleic Acids Res 49 (2021) D480–D489, https://doi.org/10.1093/nar/gkaa1100.

10. D.S. Wishart, Y.D. Feunang, A.C. Guo, E.J. Lo, A. Marcu, J.R. Grant, et al., DrugBank 5.0: a major update to the DrugBank database for 2018, Nucleic Acids Res 46 (2018) D1074–D1082, https://doi.org/10.1093/nar/gkx1037.

11. National Library of Medicine. Pharmacological Actions. Documentation. https://www.nlm.nih.gov/mesh/ps_abt.html. Reviewed October 5, 2020. Accessed March 15, 2021.
32. F. Soria Arcos, A. Romero Puche, V.T. Vicente, Controversy regarding ACE
inhibitors and p53 activation, MicroRNAs: the possible role in mediating cell death in renal injury by regulating the balance between angiotensin-converting enzyme and angiotensin-converting enzyme 2 protects from severe acute lung failure, Nature 436 (2005) 112
29. A.V. Hernandez, Y.M. Roman, V. Pasupuleti, J.J. Barboza, C.M. White,
30. M. Wang, R. Cao, L. Zhang, X. Yang, J. Liu, M. Xu, et al., Remdesivir and interferon-β may be a viable treatment option against SARS-CoV-2, Med Hypotheses 144 (2020), https://doi.org/10.1016/j.mehy.2020.110054.110054.
31. M. Dalal, Is DPP4 inhibition a comrade or adversary in COVID-19 infection, https://doi.org/10.23812/20-Editorial-Kritas.
35. F. Chaudhry, S. Lavandero, X. Xie, B. Sabharwal, Y.-Y. Zheng, A. Correa, et al., Sitagliptin treatment at the time of hospitalization was associated with reduced mortality in patients with type 2 diabetes and COVID-19: a case series from an academic hospital in Lombardy, Italy. Diabetes Care 43 (2020) 3006–3016, https://doi.org/10.2337/dc20-1521.
45. H. Singh, A.K. Kakkar, P. Chauhan, Repurposing minocycline for COVID-19 management: mechanisms, opportunities, and challenges, Expert Rev Anti Infect Ther 18 (2020) 997–1003, https://doi.org/10.1080/17478720.2017.182190.
50. S. Bharadwaj, K.E. Lee, V.D. Dwivedi, S.G. Kang, Computational insights into tetracyclines as inhibitors against SARS-CoV-2 Mpro via combinatorial molecular simulation calculations, Life Sci 257 (2020), https://doi.org/10.1016/j.lfs.2020.110653.110653.
58. F. Cantini, L. Niccoli, C. Nannini, D. Matarrese, M.E. Di Natale, P. Lotti, et al., Impact of comorbidities and glycemia at admission and Dipeptidyl Peptidase 4 inhibitors in patients with type 2 diabetes with COVID-19: a case series from an academic hospital in Lombardy, Italy. Diabetes Res Clin Pract 164 (2020), https://doi.org/10.1016/j.diabres.2020.108215.108215.
59. P.S. Sen Gupta, M.K. Rana, Ivermectin, famotidine, and doxycycline: a suggested strategy, J Biol Regul Homeost Agents 34 (2020) 9–14, https://doi.org/10.3821/20-Editorial-Kritas.
63. A. Kalil, T. Patterson, A. Mehta, K. Tomashek, C. Wolfe, V. Ghazaryan, et al., Baricitinib in COVID-19 moderate pneumonia: multicentre, open-label, randomized, non-inferiority trial, Lancet Respir Med 8 (2020) 487–488, https://doi.org/10.1016/j.lrrp.2020.06.052.