

RESEARCH

Open Access



Real-world evidence with a retrospective cohort of 15,968 COVID-19 hospitalized patients suggests 21 new effective treatments

Carlos Loucera^{1,2} , Rosario Carmona^{1,2,3} , Marina Esteban-Medina^{1,2} , Gerrit Bostelmann¹ , Dolores Muñoz-Muñoz⁴ , Román Villegas⁴ , María Peña-Chilet^{1,2,3} and Joaquín Dopazo^{1,2,3,5*}

Abstract

Purpose Despite the extensive vaccination campaigns in many countries, COVID-19 is still a major worldwide health problem because of its associated morbidity and mortality. Therefore, finding efficient treatments as fast as possible is a pressing need. Drug repurposing constitutes a convenient alternative when the need for new drugs in an unexpected medical scenario is urgent, as is the case with COVID-19.

Methods Using data from a central registry of electronic health records (the Andalusian Population Health Database), the effect of prior consumption of drugs for other indications previous to the hospitalization with respect to patient outcomes, including survival and lymphocyte progression, was studied on a retrospective cohort of 15,968 individuals, comprising all COVID-19 patients hospitalized in Andalusia between January and November 2020.

Results Covariate-adjusted hazard ratios and analysis of lymphocyte progression curves support a significant association between consumption of 21 different drugs and better patient survival. Contrarily, one drug, furosemide, displayed a significant increase in patient mortality.

Conclusions In this study we have taken advantage of the availability of a regional clinical database to study the effect of drugs, which patients were taking for other indications, on their survival. The large size of the database allowed us to control covariates effectively.

Keywords Drug repurposing, COVID-19, Real world data, Real world evidence, Retrospective cohort

Introduction

During the COVID-19 pandemic, population-wide person-level electronic health record (EHR) data has increasingly gained importance for exploring, modeling, and reporting disease trends to inform healthcare and public health policy [1]. The increasing availability of COVID-19 digital health data has fostered the interest in the use of real-world data (RWD) [2], defined as patient data collected from their EHRs, which can be analyzed to generate real-world evidence (RWE) [3]. Actually, RWE can provide a better image of the actual clinical environments in which medical interventions are

*Correspondence:

Joaquín Dopazo

joaquin.dopazo@juntadeandalucia.es

¹Computational Medicine Platform, Andalusian Public Foundation Progress and Health-FPS, Sevilla, Spain

²Institute of Biomedicine of Seville, IBIS, University Hospital Virgen del Rocío/CSIC/University of Seville, Sevilla, Spain

³Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER), FPS, Hospital Virgen del Rocío, Sevilla, Spain

⁴Subdirección Técnica Asesora de Gestión de la Información. Servicio Andaluz de Salud, Sevilla, Spain

⁵FPS/ELIXIR-ES, Fundación Progreso y Salud (FPS), CDCA, Hospital Virgen del Rocío, Sevilla, Spain



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

carried out when compared to conventional randomized clinical trials (RCTs), given that RWD includes detailed data on patient demographics, comorbidities, adherence, and simultaneous prescriptions [4, 5]. Moreover, RWE studies are not only cheaper than RCTs but can also be accomplished much faster, an advantage in scenarios in which an urgent decision must be taken, as in the case of a pandemic. In particular, discovering new drugs that could be used as efficient COVID-19 therapies is still an urgent need. Interestingly, much information on drugs, prescribed in COVID-19 patients for other indications, that could affect the progression of the disease is currently available in EHRs. For example, RWE has recently demonstrated that vitamin D has a significant protective effect on COVID-19 hospitalized patients [6]. Therefore, RWD opens the door to carry out massive drug repurposing studies as well as research on potential adverse effects or interactions of drugs with COVID-19 progression.

Since 2001, the Andalusian Public Health System has systematically stored all the electronic health record (EHR) data of Andalusian patients in the Health Population Base (BPS) [7], which is currently one of the largest repositories of clinical data in the world (with over 13 million of comprehensive patient registries) [7]. Because of its size and the detail of the data stored, BPS constitutes a unique and privileged environment to carry out large-scale RWE studies.

Results

Data analysis

Clinical data for a total of 15,968 COVID-19 patients hospitalized in Andalusia between January and November 2020 were requested from the BPS. The data was transferred from BPS to the Infrastructure for secure

real-world data analysis (iRWD) [8] at the Foundation Progress and Health of the Andalusian Public Health System.

The endpoint considered was COVID-19 death during the first 30 days of hospital stay (see Methods). To elucidate if any given treatment could potentially reduce the mortality in COVID-19 inpatients a covariate balance analysis, which considers confounders (covariates that present an *a priori* possibility of confounding the association between a treatment and the survival outcome: sex, obesity, hypertension, cancer, pulmonary diseases, hypertension, asthma, age, and mental diseases; see Methods and Table 1), was carried out to determinate the viability of further covariate-adjusted analysis. For these drugs eligible for covariate-adjusted analysis, survival was estimated using a weighted Cox Proportional Hazard model (See Methods), conditioned to the confounders of interest (Table 1). A total of 864 treatments were identified in the BPS drug archive among the patients analyzed.

Since clinical data on laboratory analyses is also available in the BPS, lymphocyte progression, high levels of which account for a favorable progression, was assessed along with the drug treatment by a Linear Mixed Effects analysis, weighting the model with the same schema as in the survival analysis (see Methods for details).

Drugs with significant effect on patient survival

Survival estimations showed that a total of 21 drugs have a significant effect on patient survival and, simultaneously, showed a significant increase in lymphocyte counts, after correction for the possible confounding covariables and for multiple testing (see Fig. 1; Table 2). Figure 2 shows the pattern of lymphocyte counts along the infection in the period studied for Enoxaparin (Fig. 2A), which displays a clear trend of high levels of lymphocyte progression, for calcifediol (Fig. 2B), with protective effect already reported [6], supported also by high levels of lymphocyte progression, and, as a counter-example, furosemide, here linked to an increase in death risk, with lymphocyte levels below the average population (Fig. 2C). Table S2 contains an exhaustive list of the results obtained for the drugs tested.

Validation of previous machine learning predictions

Interestingly, a number of the drugs found to affect COVID-19 patient survival were predicted as potentially active against COVID-19 [9] using machine learning and mathematical modeling [10] of the recently proposed COVID-19 the disease map [11] (see the last two columns from Table S2). It is interesting to note that, among the drugs eligible for the covariate-adjusted analysis (those in Table S2) there is a significant enrichment of drugs predicted as repurposable by the machine learning model among those with a significant protective

Table 1 Association between each covariate and the end point using chi-squared tests, along with the test p-value, counts and proportions with respect to the end point

covariate	survival	death	p-value
Total N	13,116	2678	
Sex (female)	6024 (45.9)	1129 (42.2)	< 0.001
Flu vaccine	5465 (41.7)	1746 (65.2)	< 0.001
Pneumococcal vaccine	3441 (26.2)	1111 (41.5)	< 0.001
Diabetes	3856 (29.4)	1167 (43.6)	< 0.001
Circulatory diseases	8111 (61.8)	2261 (84.4)	< 0.001
Cancer	1550 (11.8)	545 (20.4)	< 0.001
Respiratory diseases	2896 (22.1)	828 (30.9)	< 0.001
Dementia	964 (7.3)	536 (20.0)	< 0.001
Other mental diseases	1764 (13.4)	407 (15.2)	0.018
Anxiety and mood disorders	3382 (25.8)	784 (29.3)	< 0.001
Age			< 0.001
18–40	1399 (10.7)	20 (0.7)	
41–67	5971 (45.5)	380 (14.2)	
68–99	5746 (43.8)	2278 (85.1)	

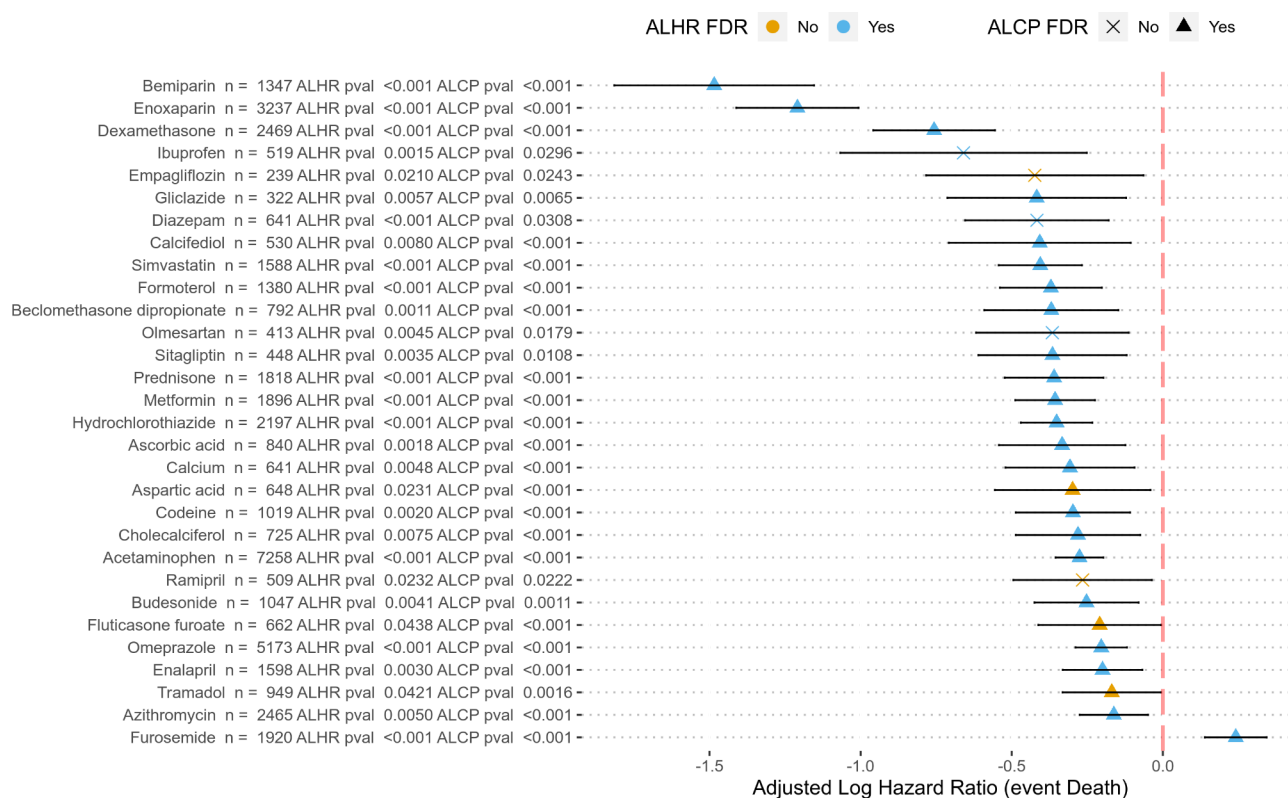


Fig. 1 Impact of drugs on patient survival. Adjusted log-hazard ratios 95% confidence intervals for all the eligible treatments that were significant in both analyses (survival and lymphocyte count progression) before and after FDR adjustment

effect with respect to the covariate-adjusted survival test ($X^2=4.003$, $pvalue=0.0454$), which supports the validity of the predictions previously made [9].

Discussion

The drugs associated to the highest survival, bemiparin (DB09258), logarithm of Hazard ratio (LHR)= -1.62, with a 95%, confidence interval (CI) of [-1.95,-1.31], and a False Discovery Rate (FDR) adjusted $pvalue=<10^{-11}$ and Enoxaparin (LHR= -1.17, 95% CI [-1.36,-0.98], FDR $pvalue=<10^{-11}$), are antithrombotic used, as other heparins, to prevent thrombotic and thromboembolic complications in hospitalized patients. While for bemiparin only weak evidence of its protective effect has been found in the literature [12], a lower rate mortality in COVID-19 patients was described for enoxaparin when compared to other heparins [13], in agreement with the results found here. However, this protective effect is not shared by other anticoagulants, such as tinziparin (LHR= -0.34, 95% CI[-1.38, 0.69], FDR $pvalue=1$), despite its use in pulmonary embolism, or Fondaparinux (LHR=-0.33, 95% CI[-1.64, 0.97], FDR $pvalue=1$). Calcifediol and Cholecalciferol, already described by us in a previous work [6], are significantly associated with better patient survival, probably due to the protective role of vitamin D due to its pro-immune and antiinflammatory properties. Other

studies suggest also a protective effect of ascorbic acid (vitamin C) [14]. Table S2 contains an exhaustive list of the results obtained from the drugs tested.

One of drugs with a significant protective effect is simvastatin, a widely used statin, a group of drugs that reduce the blood level of low-density lipoprotein (LDL) cholesterol. Statins are also known for their pleiotropic effect, exerting an anti-inflammatory and antithrombotic action by inhibiting the NF- κ B pathway which directly reduces inflammatory cytokines (IL1, IL6, TNF- α), CRP, and neutrophils [15]. Furthermore, a retrospective study performed in COVID-19 hospitalized patients showed that statins inhibit RAS activation and reduce angiotensin II proinflammatory effects, therefore improving endothelial function and remodeling after vascular injury [16]. A recent *in-vitro* study demonstrates that simvastatin pretreatment in human Calu-3 epithelial lung cells inhibited SARS-CoV-2 binding and entry to the cell by inducing a redistribution of ACE2 receptors, lowering its concentration on the plasma membrane [17]. Recent retrospective studies also point to the relationship between statin consumption and a reduced risk of mortality in COVID-19 patients [16, 18]. Another predicted drug is hydrochlorothiazide, a diuretic drug, often combined with ACE-inhibitors such as enalapril as antihypertensive therapy [19]. It has been reported that patients with

Table 2 Log Hazard ratios obtained for the drugs tested, along with standard deviations (SDs), FDR-adjusted p-values, and Lymphocyte proliferation values (see Methods) along with FDR-adjusted p-values. The two last columns indicate the drugs used in the machine learning drug repurposing prediction study and the significance of the prediction

Drugbank ID	Drug	N	LHR estimate	LHR SD	LHR CI 5%	LHR CI 95%	LHR p-value	LHR FDR p-value	Lym-phocyte count	LC CI 5%	LC CI 95%	LC p-value	LC FDR p-value
DB09258	Bemiparin	1347	-1.48449	0.16861	-1.81495	-1.15402	< 0.00001	< 0.00001	6.19153	0.00990	0.01908	< 0.00001	< 0.00001
DB01225	Enoxaparin	3237	-1.20921	0.10316	-1.41141	-1.00702	< 0.00001	< 0.00001	11.31073	0.01233	0.01750	< 0.00001	< 0.00001
DB01234	Dexamethasone	2469	-0.75679	0.10273	-0.95813	-0.55545	< 0.00001	< 0.00001	9.42789	0.01233	0.01881	< 0.00001	< 0.00001
DB01050	Ibuprofen	519	-0.65959	0.20808	-1.06743	-0.25175	0.00153	0.01095	2.18280	0.00080	0.01503	0.02961	0.07372
DB09038	Empagliflozin	239	-0.42376	0.18355	-0.78352	-0.06401	0.02096	0.06730	2.26829	0.00126	0.01714	0.02430	0.06177
DB01120	Gliclazide	322	-0.41723	0.15098	-0.71315	-0.12131	0.00572	0.02492	2.74039	0.00257	0.01549	0.00649	0.02262
DB00829	Diazepam	641	-0.41669	0.12112	-0.65408	-0.17930	0.00058	0.00473	2.16427	0.00057	0.01141	0.03084	0.07525
DB00146	Calcifediol	530	-0.40717	0.15363	-0.70827	-0.10607	0.00804	0.03164	3.40110	0.00502	0.01869	0.00073	0.00328
DB00641	Simvastatin	1588	-0.40558	0.07026	-0.54328	-0.26789	< 0.00001	< 0.00001	5.16632	0.00523	0.01163	< 0.00001	0.00001
DB00983	Formoterol	1380	-0.37040	0.08596	-0.53888	-0.20192	0.00002	0.00018	5.13579	0.00650	0.01453	< 0.00001	0.00001
DB00394	Beclomethasone dipropionate	792	-0.36901	0.11319	-0.59085	-0.14717	0.00111	0.00849	4.15390	0.00594	0.01656	0.00004	0.00045
DB00275	Olmesartan	413	-0.36562	0.12879	-0.61803	-0.11320	0.00453	0.02209	2.37903	0.00137	0.01416	0.01786	0.05002
DB01261	Sitagliptin	448	-0.36512	0.12515	-0.61040	-0.11984	0.00353	0.01956	2.56016	0.00179	0.01355	0.01077	0.03552
DB00635	Prednisone	1818	-0.35982	0.08332	-0.52311	-0.19652	0.00002	0.00018	5.43363	0.00644	0.01371	< 0.00001	< 0.00001
DB00331	Metformin	1896	-0.35648	0.06733	-0.48845	-0.22451	< 0.00001	< 0.00001	3.88744	0.00273	0.00827	0.00010	0.00073
DB00999	Hydrochlorothiazide	2197	-0.35148	0.06036	-0.46979	-0.23318	< 0.00001	< 0.00001	3.98046	0.00273	0.00802	0.00007	0.00067
DB00126	Ascorbic acid	840	-0.33321	0.10674	-0.54241	-0.12400	0.00180	0.01219	3.66682	0.00406	0.01344	0.00026	0.00127
DB01373	Calcium	641	-0.30728	0.10895	-0.52082	-0.09375	0.00480	0.02250	3.75161	0.00479	0.01529	0.00019	0.00103
DB00128	Aspartic acid	648	-0.29845	0.13140	-0.55598	-0.04091	0.02313	0.06911	3.66511	0.00543	0.01791	0.00027	0.00127
DB00318	Codeine	1019	-0.29747	0.09647	-0.48655	-0.10840	0.00204	0.01247	3.92417	0.00441	0.01324	0.00009	0.00073
DB00169	Cholecalciferol	725	-0.28067	0.10502	-0.48651	-0.07483	0.00753	0.03062	3.66953	0.00437	0.01440	0.00026	0.00127
DB00316	Acetaminophen	7258	-0.27573	0.04035	-0.35481	-0.19665	< 0.00001	< 0.00001	6.53570	0.00361	0.00671	< 0.00001	< 0.00001
DB00178	Ramipril	509	-0.26557	0.11700	-0.49489	-0.03624	0.02323	0.06911	2.29420	0.00102	0.01294	0.02222	0.06025
DB01222	Budesonide	1047	-0.25260	0.08800	-0.42508	-0.08012	0.00410	0.02174	3.28003	0.00289	0.01151	0.00108	0.00460
DB08906	Fluticasone furoate	662	-0.20888	0.10361	-0.41196	-0.00580	0.04381	0.11921	4.00962	0.00547	0.01595	0.00007	0.00067
DB00338	Omeprazole	5173	-0.20373	0.04349	-0.28897	-0.11848	< 0.00001	0.00004	3.92296	0.00189	0.00566	0.00009	0.00073
DB00584	Enalapril	1598	-0.19957	0.06725	-0.33138	-0.06775	0.00300	0.01745	3.76049	0.00297	0.00945	0.00018	0.00103
DB00193	Tramadol	949	-0.16897	0.08313	-0.33191	-0.00604	0.04209	0.11921	3.16350	0.00252	0.01072	0.00162	0.00636
DB00207	Azithromycin	2465	-0.16216	0.05779	-0.27542	-0.04889	0.00502	0.02267	4.36750	0.00338	0.00888	0.00001	0.00018
DB00695	Furosemide	1920	0.24131	0.05212	0.13916	0.34346	< 0.00001	0.00005	-3.74328	-0.00805	-0.00251	0.00019	0.00103

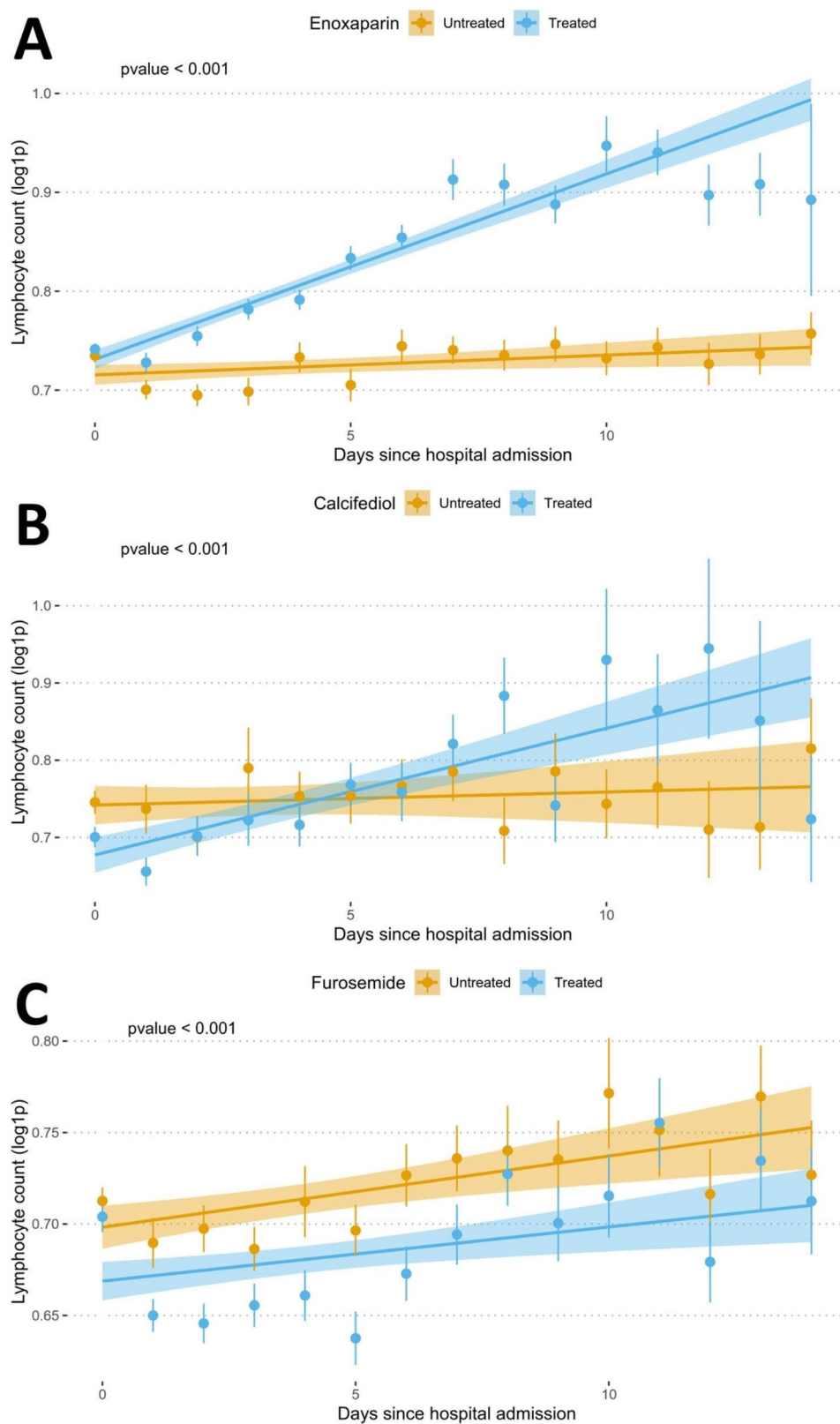


Fig. 2 Lymphocyte counts. Plots showing the evolution of lymphocyte counts along the time studied (15 days since hospital admission) for (A) enoxaparin, (B) Calcifediol and (C) Furosemide

hypertension present a higher susceptibility to a severe COVID-19 prognosis [20], underlying hypertension as a risk factor for increased mortality in infected patients. Although the effect of antihypertensive drugs on COVID-19 patients with hypertension is controversial, the upregulation of ACE2 by ACE-inhibitors was linked to a dampened hyperinflammation and increased intrinsic antiviral responses of the cell in hypertensive COVID-19 patients [21]. The results presented here, together with these previous reports, suggest that ACE-inhibitors may have a protective effect, in addition to helping to improve the prognosis of hypertensive patients. Dexamethasone has been studied in the context of COVID-19 disease due to its property as an anti-inflammatory drug [22]. Although ibuprofen and other analgesic like acetaminophen was initially discommended for COVID-19 treatment [23], further studies based on observational data could not confirm the theoretical risks of ibuprofen and other Nonsteroidal Anti-Inflammatory Drugs (NSAIDs) in SARS-CoV-2 infection [24]. Moreover, other studies suggested that some NSAID could have antiviral activity in coronaviruses, including SARS-CoV-2 [25], an activity demonstrated here for ibuprofen. Similarly, tramadol is an opioid analgesic used to treat moderate to severe pain, that was initially deemed as bad prognosis [26], but further studies suggested a potential therapeutic effect [27].

The empagliflozin is an inhibitor of the sodium-glucose cotransporter 2 (*SGLT2*) used in the treatment of type 2 diabetes, whose potential utility in patients with COVID-19 has been suggested [28] but not demonstrated yet. Also for diabetes patients, the available evidence suggests that Sitagliptin may be beneficial in treating COVID-19, particularly in patients with type 2 diabetes who appear to be at high risk of mortality and of cardiorenal or cerebrovascular complications [29]. Another diabetes treatment, metformin, has also been suggested as an effective in the treatment of COVID-19 [30].

It has been suggested that steroids used for asthma treatment could have a protective effect in COVID-19 [31], although specifically beclometasone dipropionate was not assessed. It has also been reported that corticosteroids, including prednisone, are effective in reducing mortality in COVID-19 patients within their therapeutic window [27], or reduce hospitalization times, like budesinone [32]. Some studies suggest that formoterol could be used to improve lung function and assist symptom control in COVID-19 patients [33] however, the available evidence does not suggest any significant interaction between formoterol and COVID-19 [34]. A recent study suggested that olmesartan could alleviate renal fibrosis induced by SARS-CoV-2 envelope protein by regulating HMGB1 release and autophagic degradation of TGF- β 1 [35]. In the case of omeprazole, a proton pump inhibitor used to treat gastroesophageal reflux disease

(GERD), peptic ulcer disease, and other acid-related disorders, several studies have indicated an anti-viral effect [36], as well as a therapeutic role in combination with other antiviral [36]. Finally, zithromycin is an antibiotic with potential antiviral and anti-inflammatory properties [37] although the consensus is that there is no evidence to support the use of azithromycin for the treatment of COVID-19 [38].

On the other hand, a study suggested that furosemide, a diuretic medication used to treat fluid build-up due to heart failure, liver scarring, or kidney disease and high blood pressure, may have potential therapeutic benefits for COVID-19 patients with acute respiratory distress syndrome [36], contrarily to what we observed here, supported by the lymphocyte count data (see Table 2 and Supplementary Table S2). It is important to note that other drugs, which are marginally non-significant because of small sample sizes, have also a potential negative effect on COVID-19 patient survival. These drugs have different mechanisms of action and are used to treat different conditions: latanoprost, used to treat glaucoma and ocular hypertension, ciprofloxacin, an antibiotic, tamsulosin, an alpha-blocker, trazodone and lormetazepam, used to treat insomnia and anxiety, and lormetazepam, a benzodiazepine.

To our knowledge, previous studies either did not identify evidence suggesting any effect on COVID-19 prognosis, or they not have detected the significant protective effects we observed in this study for certain drugs such as diazepam, gliclazide, hydrochlorothiazide, calcium, aspartic acid, codeine, ramipiril, fluticasone furoate, fluticasone furoate, zithromycin and enalapril. The large sample size of this study and the appropriate management of confounding variables allowed us to validate some proposed therapeutic interventions and to expand the number of potential COVID-19 treatments.

Conclusions

The Andalusian Population Health Database was used to explore drug repurposing using data from 15,968 COVID-19 patients hospitalized in Andalusia between January and November 2020. The study identified 21 drugs associated with improved patient survival and lymphocyte progression. This finding offers potential treatment options for COVID-19. However, one drug, furosemide, was linked to increased patient mortality, requiring further investigation. This study demonstrates the value of drug repurposing strategies in addressing emergent health challenges. Additionally, it underscores the importance of comprehensive clinical databases in advancing medical knowledge and patient care.

Materials and methods

Design and patient selection

This study uses a retrospective cohort which includes Andalusian patients with COVID-19 diagnosis, hospitalized during the period January 2020 to November 2021.

The Ethics Committee for the Coordination of Biomedical Research in Andalusia approved the study (29th September, 2020, Acta 09/20) and waived informed consent for the secondary use of clinical data for research purposes.

Data management

Clinical data corresponding to COVID-19 patients hospitalized in Andalusia between January and November 2020 was requested to the Health Population Base (BPS), and from there transferred to the Infrastructure for secure real-world data analysis (iRWD) at the Foundation Progress and Health (FPS) of the Andalusian Public Health System for further analysis. In particular, the data listed in Table S1 was extracted in BPS from the electronic health records for each patient and transferred to FPS for a total of 15,968 COVID-19 patients that fulfilled the inclusion criteria.

Data preprocessing

Medication data in the office and hospital pharmacy records were found for 864 treatments. Individuals are considered as treated with a specific drug if prescriptions and the corresponding pharmacy dispensations (therein after a valid pharmacy order) were found within a period from 15 days before the hospital admission until the discharge up to 14 days (or death). Otherwise they were considered untreated.

The endpoint studied was COVID-19 death (certified death events during hospitalization). As in previous studies, the first 30 days of hospital stay were considered for survival calculations [39]. The time variable in the models corresponds to the length (in days) of hospital stay. The stays that imply one or more changes of hospital units are combined in a single stay where the admission and discharge dates are set to either the start of the first or the end of the last combined stay. Only the first stay for each patient was considered to reduce potential biases due to reinfection.

Covariate definition

Following previous studies [40] the ICD codes were grouped into conditions as diabetes mellitus (ICD code E11), diseases of the circulatory system (ICD10 codes I00-I99), diseases of the respiratory system (ICD10 codes J00-J99), neoplasms (ICD10 codes C00-D49), dementia (ICD10 codes F00-F03), anxiety or mood disorders (ICD10 codes F30-F48), and other mental diseases (ICD10 codes F04-F29 and F50-F99). Obesity and other

associated conditions (ICD10 codes D5-D8) with a possible confounding effect with the COVID-19 outcome were checked but no evidence was found in our database (non-significant χ^2 association test). The age was categorized in the following ranks: [18, 40], [41, 67] and [68, 99]. Gender was also considered as a known covariate. Table 1 displays the association between each covariate and the end point considered here, death, using chi-squared tests, along with the test p-value. Counts and proportions with respect to the end point are also provided.

Statistical analysis

To elucidate if any given treatment could potentially reduce the mortality in COVID-19 inpatients three statistical tests were conducted, considering covariates that present an a priori possibility of confounding the association between a treatment and the survival outcome [41] (see previous section).

Firstly, the survival outcome was estimated using a Cox Proportional Hazard model weighted using the Inverse Probability of Treatment Weighting (IPTW) technique, with the weights computed using a logistic regression model and adjusted for estimating the Average Treatment effect on the Treated population (ATT) conditioned to the confounders of interest using the whole cohort. ATT is the most used weighting approximation to estimate treatment effects [42]. To obtain an accurate measure of the variability of the marginal hazard ratios the closed-form estimator previously proposed [43] was used.

Then, the lymphocyte progression, a marker of COVID-19 severity [44], is established as the different measurements of lymphocyte counts with respect to the initial day of hospitalization up to 14 days [45]. Dates outside the hospitalization date range were omitted. The association between the positive linear trends in daily lymphocyte counts and reduced mortality in COVID-19 is obtained by comparing the trends in a treated population versus a control untreated population. A Linear Mixed Effects (LME) analysis was conducted to estimate if there was an increasing linear trend in the log-transformed lymphocyte progression due to being under a given treatment and the statistical significance was checked using an ANOVA analysis of the model [46]. The model was weighted following the same weighting schema as in the survival analysis. In addition, a covariate balance analysis to determine the viability of the weighting schema [47] was carried out.

For each treatment, the Inverse Probability Weighting (IPW) was used, based on propensity scores (IPW) generated using the WeightIt R package (v 0.12) [48]. Here, the exposed condition is either having valid pharmacy order for the treatment during the 15 days prior to the beginning of the hospitalization event or during the first 14 days of the hospitalization. To assess the viability of

the IPW analysis the proportion of covariates that could be effectively balanced was checked using the standardized mean differences test as implemented in the Cobalt R package (v 4.3.1) [49], using the 0.05 threshold [47]. A treatment is eligible if all the covariates could be properly balanced, resulting in 122 eligible treatments out of the 864 initially found.

In both cases, p-values are corrected for multiple testing with False Discovery Rate (FDR) [50]. Significance is achieved at level 0.05 and 95% confidence intervals are provided.

Software used

Weights for IPW are computed with the WeightIt R package (v 0.12) [48]. IPW covariate suitability was computed using the Cobalt R package (v 4.3.1) [49]. The survival estimation was conducted with R package HrIPW (v 0.1.2) [51]. The LME analysis was conducted with R package lme4 (v 1.1–27) [52]. The ANOVA analysis of the LME model was conducted with R package lmerTest (v 3.1-3) [46]. R version 3.6.3 (2020-02-29).

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12985-023-02195-9>.

Additional file 1: Supplementary Table S1. Data imported from BPS for each patient: code and definition of the variable. Supplementary Table S2. Log Hazard ratios obtained for the drugs tested, along with standard deviations (SDs), upper and lower coefficient intervals (CI), nominal and FDR-adjusted p-values. Also, Lymphocyte proliferation values (see Methods) along with standard deviations (SDs), upper and lower coefficient intervals (CI), nominal and FDR-adjusted p-values. The two last columns indicate the drugs used in the machine learning drug repurposing prediction study and the proteins targeted by the drug. Targets marked with an * were those analyzed in the machine learning drug repurposing

Authors' contributions

C.L. carried out the statistical analysis and write the Methods section, R.C. participated in the interpretation of the results, M.E.-M. participated in the interpretation of the results, G.B. carried out data curation, D.M.-M. and R.V. did the data extraction, M.P.C. coordinated the interpretation of results and write the discussion, J.D. conceived the work and wrote the manuscript. All authors read and approved the final manuscript.

Funding

This research was funded by Spanish Ministry of Science and Innovation (grant PID2020-117979RB-I00), the Instituto de Salud Carlos III (ISCIII), co-funded with European Regional Development Funds (ERDF) (grant IMP/00019), and has also been funded by Consejería de Salud y Consumo, Junta de Andalucía (grants COVID-0012-2020, PS-2020-342 and IE19_259 FPS) and the postdoctoral contract of Carlos Loucera (PAIDI2020- DOC_00350) co-funded by the European Social Fund (FSE) 2014–2020.

Data availability

The data supporting the findings of this study is stored in the Andalusian Population Health Database (<https://www.sspa.juntadeandalucia.es/servicioandaluzdesalud/profesionales/sistemas-de-informacion/base-poblacional-de-salud>) but restrictions apply to the availability of these data, which were used under the ethics committee permission for the current study, and so are not publicly available.

Declarations

Competing interests

The authors declare no competing interests.

Ethics approval

This project was reviewed and approved by the Ethics Committee for the Coordination of Biomedical Research in Andalusia (29th September, 2020, Acta 09/20).

Conflict of interest

The authors have no competing interests to declare that are relevant to the content of this article.

Received: 6 December 2022 / Accepted: 27 September 2023

Published online: 06 October 2023

References

1. Wood A, Denholm R, Hollings S, Cooper J, Ip S, Walker V et al. Linked electronic health records for research on a nationwide cohort of more than 54 million people in England: data resource. *BMJ*. 2021;373.
2. Corrigan-Curay J, Sacks L, Woodcock J. JAMA: the journal of the American Medical Association. 2018;320(9):867–8. <https://doi.org/10.1001/jama.2018.10136>. Real-World Evidence and Real-World Data for Evaluating Drug Safety and Effectiveness.
3. Sherman RE, Anderson SA, Dal Pan GJ, Gray GW, Gross T, Hunter NL, et al. Real-world evidence—what is it and what can it tell us. *N Engl J Med*. 2016;375(23):2293–7.
4. Hannan EL. Randomized clinical trials and observational studies: guidelines for assessing respective strengths and limitations. *JACC: Cardiovasc Interventions*. 2008;1(3):211–7.
5. Cowie MR, Blomster JL, Curtis LH, Duclaux S, Ford I, Fritz F, et al. Electronic health records to facilitate clinical research. *Clin Res Cardiol*. 2017;106(1):1–9.
6. Loucera C, Peña-Chilet M, Esteban-Medina M, Muñoz-Muñoz D, Villegas R, Lopez-Miranda J, et al. Real world evidence of calcifediol or vitamin D prescription and mortality rate of COVID-19 in a retrospective cohort of hospitalized andalusian patients. *Sci Rep*. 2021;11(1):1–12.
7. Muñoz-Muñoz D, Goicoechea-Salazar J, García-León F, Laguna-Tellez A, Larrocha-Mata D, Cardero-Rivas M. Health record linkage: andalusian health population database. *Gac Sanit*. 2019;34(2):105–13.
8. Infrastructure for secure generation of. Real World Evidence from Real World Data from the Andalusian Health Population Database. 2020.
9. Loucera C, Esteban-Medina M, Rian K, Falco MM, Dopazo J, Peña-Chilet M. Drug repurposing for COVID-19 using machine learning and mechanistic models of signal transduction circuits related to SARS-CoV-2 infection. *Signal Transduct Target Therapy*. 2020;5(1):290. <https://doi.org/10.1038/s41392-020-00417-y>.
10. Rian K, Esteban-Medina M, Hidalgo MR, Çubuk C, Falco MM, Loucera C, et al. Mechanistic modeling of the SARS-CoV-2 disease map. *BioData Min*. 2021;14(1):1–8.
11. Ostaszewski M, Niarakis A, Mazein A, Kuperstein I, Phair R, Orta-Resendiz A, et al. COVID19 Disease Map, a computational knowledge repository of virus–host interaction mechanisms. *Mol Syst Biol*. 2021;17(10):e10387. <https://doi.org/10.15252/msb.2021.10387>.
12. Varona JF, Núñez E, Fernández Félix BM, Castellano Vázquez JM, Cubillo A. Efficacy and safety of therapeutic vs. prophylactic bempiparin in noncritically ill patients with COVID-19 pneumonia. *Eur J Intern Med*. 2022;99:106–8. <https://doi.org/10.1016/j.ejim.2022.01.031>.
13. Pawlowski C, Venkatakrishnan A, Kirkup C, Berner G, Puranik A, O'Horo JC, et al. Enoxaparin is associated with lower rates of mortality than unfractionated heparin in hospitalized COVID-19 patients. *EClinicalMedicine*. 2021;33:100774.
14. Olczak-Pruc M, Swieczkowski D, Ladny JR, Pruc M, Juarez-Vela R, Rafique Z, et al. Vitamin C supplementation for the treatment of COVID-19: a systematic review and Meta-analysis. *Nutrients*. 2022;14(19). <https://doi.org/10.3390/nu14194217>.

15. Zhou Q, Liao JK. Pleiotropic effects of statins. - basic research and clinical perspectives. *Circulation Journal: Official Journal of the Japanese Circulation Society*. 2010;74(5):818–26. <https://doi.org/10.1253/circj.10-0110>.
16. Kouhpeikar H, Tabasi HK, Khazir Z, Naghipour A, Moghadam HM, Forouzanfar H et al. Statin use in COVID-19 hospitalized patients and outcomes: a retrospective study. *Front Cardiovasc Med*. 2022;9.
17. Teixeira L, Temerozo JR, Pereira-Dutra FS, Ferreira AC, Mattos M, Gonçalves BS et al. Simvastatin downregulates the SARS-CoV-2-Induced inflammatory response and impairs viral infection through disruption of lipid rafts. *Front Immunol*. 2022;13.
18. Umakanthan S, Senthil S, John S, Madhavan MK, Das J, Patil S, et al. The protective role of statins in COVID-19 patients: a retrospective observational study. *Transl Med Commun*. 2021;6(1):22. <https://doi.org/10.1186/s41231-021-00102-4>.
19. Ruoff G. ACE inhibitors and diuretics. The benefits of combined therapy for hypertension. *Postgrad Med*. 1989;85(3):127–32. <https://doi.org/10.1080/00325481.1989.11700602>. 37 – 9.
20. Huang S, Wang J, Liu F, Liu J, Cao G, Yang C, et al. COVID-19 patients with hypertension have more severe disease: a multicenter retrospective observational study. *Hypertens Res*. 2020;43(8):824–31.
21. Peng M, He J, Xue Y, Yang X, Liu S, Gong Z. Role of hypertension on the severity of COVID-19: a review. *J Cardiovasc Pharmacol*. 2021;78(5):e648–e55. <https://doi.org/10.1097/fjc.0000000000001116>.
22. Chappell L, Horby P, Lim WS, Emberson JR, Mafham M, Bell JL et al. Dexamethasone in hospitalized patients with Covid-19-preliminary report. *The New England Journal of medicine*. 2020.
23. Day M. Covid-19: ibuprofen should not be used for managing symptoms, say doctors and scientists. *British Medical Journal Publishing Group*; 2020.
24. Moore N, Bosco-Levy P, Thurin N, Blin P, Droz-Perroteau C. NSAIDs and COVID-19: a systematic review and meta-analysis. *Drug Saf*. 2021;44(9):929–38.
25. Amici C, Caro AD, Ciucci A, Chiappa L, Castilietti C, Martella V, et al. Indomethacin has a potent antiviral activity against SARS coronavirus. *Antivir Ther*. 2006;11(8):1021–30.
26. Kifle ZD, Enyew EF, Mekuria AB. A recent achievement in the discovery and development of vaccines and therapeutic agents in the race for COVID-19 protection and treatment. *J Evidence-Based Integr Med*. 2021;26:2515690X211003727.
27. Li Y, Zhou X, Li T, Chan S, Yu Y, Ai JW, et al. Corticosteroid prevents COVID-19 progression within its therapeutic window: a multicentre, proof-of-concept, observational study. *Emerg Microbes Infect*. 2020;9(1):1869–77. <https://doi.org/10.1080/22221751.2020.1807885>.
28. MacIsaac RJ, Jerums G, Ekinci EI. Cardio-renal protection with empagliflozin. *Annals of Translational Medicine*. 2016;4(20).
29. Mikhael EM, Ong SC, Sheikh Ghadzi SM. Efficacy and safety of Sitagliptin in the treatment of COVID-19. *J Pharm Pract*. 2023;36(4):980–7. <https://doi.org/10.1177/08971900221102119>.
30. Varghese E, Samuel SM, Liskova A, Kubatka P, Büsselberg D. Diabetes and coronavirus (SARS-CoV-2): molecular mechanism of metformin intervention and the scientific basis of drug repurposing. *PLoS Pathog*. 2021;17(6):e1009634. <https://doi.org/10.1371/journal.ppat.1009634>.
31. Ramakrishnan S, Nicolau DV, Langford B, Mahdi M, Jeffers H, Mwasuku C, et al. Inhaled budesonide in the treatment of early COVID-19 (STOIC): a phase 2, open-label, randomised controlled trial. *The Lancet Respiratory Medicine*. 2021;9(7):763–72.
32. Yu L-M, Bafadhel M, Dorward J, Hayward G, Saville BR, Gbinigie O, et al. Inhaled budesonide for COVID-19 in people at high risk of complications in the community in the UK (PRINCIPLE): a randomised, controlled, open-label, adaptive platform trial. *The Lancet*. 2021;398(10303):843–55. [https://doi.org/10.1016/S0140-6736\(21\)01744-X](https://doi.org/10.1016/S0140-6736(21)01744-X).
33. Yu Y, Fang B, Yang X-D, Zheng Y. One stone two birds: anti-inflammatory bronchodilators as a potential pharmacological strategy for COVID-19. *Front Pharmacol*. 2023;14:1185076.
34. Wijesinghe M, Weatherall M, Perrin K, Harwood M, Beasley R. Risk of mortality associated with formoterol: a systematic review and meta-analysis. *Eur Respir J*. 2009;34(4):803–11. <https://doi.org/10.1183/09031936.00159708>.
35. Zhou S, Yu Z, Chen Z, Ning F, Hu X, Wu T, et al. Olmesartan alleviates SARS-CoV-2 envelope protein induced renal fibrosis by regulating HMGB1 release and autophagic degradation of TGF- β 1. *Front Pharmacol*. 2023;14:1187818. <https://doi.org/10.3389/fphar.2023.1187818>.
36. Ray A, Sharma S, Sadasivam B. The potential therapeutic role of Proton Pump inhibitors in COVID-19: hypotheses based on existing evidences. *Drug Res (Stuttg)*. 2020;70(10):484–8. <https://doi.org/10.1055/a-1236-3041>.
37. Kournoutou GG, Dinos G. Azithromycin through the Lens of the COVID-19 treatment. *Antibiot (Basel)*. 2022;11(8). <https://doi.org/10.3390/antibiotics11081063>.
38. Gbinigie K, Frie K. Should azithromycin be used to treat COVID-19? A rapid review. *BJGP open*. 2020;4(2).
39. Sterne JA, Murthy S, Diaz JV, Slutsky AS, Villar J, Angus DC, et al. Association between administration of systemic corticosteroids and mortality among critically ill patients with COVID-19: a meta-analysis. *JAMA: The Journal of the American Medical Association*. 2020;324(13):1330–41.
40. Hoertel N, Sánchez-Rico M, Vernet R, Beeker N, Jannot A-S, Neuraz A, et al. Association between antidepressant use and reduced risk of intubation or death in hospitalized patients with COVID-19: results from an observational study. *Mol Psychiatry*. 2021;26(9):5199–212.
41. Gutiérrez-Gutiérrez B, del Toro MD, Borobia AM, Carcas A, Jarrín I, Yllescas M, et al. Identification and validation of clinical phenotypes with prognostic implications in patients admitted to hospital with COVID-19: a multicentre cohort study. *Lancet Infect Dis*. 2021. [https://doi.org/10.1016/S1473-3099\(21\)00019-0](https://doi.org/10.1016/S1473-3099(21)00019-0).
42. Imbens GW. Nonparametric estimation of average treatment effects under exogeneity: a review. *Rev Econ Stat*. 2004;86(1):4–29.
43. Hajage D, Chauvet G, Belin L, Lafourcade A, Tubach F, De Rycke Y. Closed-form variance estimator for weighted propensity score estimators with survival outcome. *Biom J*. 2018;60(6):1151–63.
44. Lu C, Liu Y, Chen B, Yang H, Hu H, Zhao Y. Prognostic value of lymphocyte count in severe COVID-19 patients with corticosteroid treatment. *Signal Transduct Target Therapy*. 2021;6(1):1–3.
45. Manenti L, Maggiore U, Fiaccadori E, Meschi T, Antoni AD, Nouvenne A, et al. Reduced mortality in COVID-19 patients treated with colchicine: results from a retrospective, observational study. *PLoS ONE*. 2021;16(3):e0248276. <https://doi.org/10.1371/journal.pone.0248276>.
46. Kuznetsova A, Brockhoff PB, Christensen RH. lmerTest package: tests in linear mixed effects models. *J Stat Softw*. 2017;82(1):1–26.
47. Stuart EA, Lee BK, Leacy FP. Prognostic score–based balance measures can be a useful diagnostic for propensity score methods in comparative effectiveness research. *J Clin Epidemiol*. 2013;66(8):84–S90. e1.
48. Greifer N, Weighltt. Weighting for Covariate Balance in Observational Studies. 2021.
49. Greifer N. cobalt: Covariate Balance Tables and Plots. 2021.
50. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J Royal Stat Soc Ser B*. 1995;57(1):289–300.
51. Hajage D, HrlPW. Hazard Ratio Estimation using Cox Model Weighted by the Estimated Propensity Score. 2020.
52. Bates D, Maechler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. *J Stat Softw*. 2015;67:1–48.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.