



A machine learning model exploring the relationship between chronic medication and COVID-19 clinical outcomes

Berta Miró¹ · Natalia Díaz González¹ · Juan-Francisco Martínez-Cerdá² · Clara Viñas-Bardolet² · Alex Sánchez-Pla^{3,4} · Adrián Sánchez-Montalvá^{5,6,7} · Marta Miarons^{8,9}

Received: 20 February 2025 / Accepted: 25 May 2025 / Published online: 28 July 2025
© The Author(s) 2025

Abstract

Background The impact of chronic medication on COVID-19 outcomes has been a topic of ongoing debate since the onset of the pandemic. Investigating how specific long-term treatments influence infection severity and prognosis is essential for optimising patient management and care.

Aim This study aimed to investigate the association between chronic medication and COVID-19 outcomes, using machine learning to identify key medication-related factors.

Method We analysed 137,835 COVID-19 patients in Catalonia (February–September 2020) using eXtreme Gradient Boosting to predict hospitalisation, ICU admission, and mortality. This was complemented by univariate logistic regression analyses and a sensitivity analysis focusing on diabetes, hypertension, and lipid disorders.

Results Participants had a mean age of 53 (SD 20) years, with 57% female. The best model predicted mortality risk in 18 to 65-year-olds (AUCROC 0.89, CI 0.85–0.92). Key features identified included the number of prescribed drugs, systemic corticoids, 3-hydroxy-3-methylglutaryl coenzyme A (HMG-CoA) reductase, and hypertension drugs. A sensitivity analysis identified that hypertensive participants over 65 taking angiotensin-converting enzyme (ACE) inhibitors or angiotensin II receptor blockers (ARBs) had lower mortality risk (OR 0.78 CI 0.68–0.92) compared to those on other antihypertensive medication (OR 0.8 CI 0.68–0.95). Treatment with inhibitors of dipeptidyl peptidase 4 was associated to higher mortality in participants aged 18–65, while metformin showed a protective effect in those over 65 (OR 0.79, 95% CI 0.68–0.92).

Conclusion Machine learning models effectively distinguished COVID-19 outcomes. Patients under ACEi or ARBs or biguanides should continue their prescribed medications, which may offer protection over alternative treatments.

Keywords ACE inhibitors · ARBs · COVID-19 · HMG-CoA reductase · Machine learning · Metformin · Mortality · Polypharmacy · Prediction models

✉ Adrián Sánchez-Montalvá
Adrian.Sanchez.Montalva@uab.cat

¹ Statistics and Bioinformatics Unit, Vall d’Hebron Institut de Recerca (VHIR), Barcelona, Spain

² Agency for Health Quality and Assessment of Catalonia (AQuaS), Barcelona, Spain

³ Genetics, Microbiology and Statistics Department, Universitat de Barcelona, Barcelona, Spain

⁴ Centro de Investigación Biomédica en Red Fragilidad y Envejecimiento Saludable, Instituto de Salud Carlos III, 28029 Madrid, Spain

⁵ Department of Medicine, Universitat Autònoma de Barcelona, Barcelona, Spain

⁶ International Health Unit Vall d’Hebron (PROSICS), Infectious Diseases Department, Vall d’Hebron University Hospital, Vall d’Hebron Institute of Research, Barcelona, Spain

⁷ Centre for Biomedical Research in Infectious Diseases Network (CIBERINFEC), Institute of Health Carlos III, Madrid, Spain

⁸ Pharmacy Department, Vall d’Hebron Hospital Universitari, Vall d’Hebron Barcelona Hospital Campus, Barcelona, Spain

⁹ Pharmacy Department, Consorci Hospitalari de Vic, Barcelona, Spain

Impact statements

- Machine learning-driven models effectively predict hospitalisation, ICU admission, and death in COVID-19 patients.
- ACEi, ARBs, and biguanides (metformin) demonstrated potential protective effects.
- Findings support continued use of these medications without switching to alternatives.
- The study methodology is scalable and can be applied at a low additional cost using AI-driven analysis of administrative health databases.

Introduction

COVID-19 emerged as an unprecedented public health crisis, placing immense strain on national health systems. Globally, the outbreak disrupted economies, hindered trade, and restricted mobility [1], created challenges in medication supply, and pharmacy services [2].

Various studies have been conducted to determine risk factors associated with poor COVID-19 outcomes [3–8], although a reduced number used health record data [9–13]. While unmodifiable factors like age, sex, and comorbidities have been linked to severe COVID-19 outcomes [14, 15], potentially modifiable factors such as chronic medication use have also drawn considerable attention. Since the first wave, studies have examined how long-term treatments [16–21], particularly cardiovascular drugs like angiotensin-converting enzyme (ACE) inhibitors and angiotensin II receptor blockers (ARBs), may influence disease severity. Evidence suggests there is no increase in mortality risk [22–24]. Corticosteroids and other anti-inflammatories have also been used in managing COVID-related respiratory distress [25].

The management of chronic conditions and associated medications has been a key focus [26], emphasizing the importance of effective disease control with drugs that either avoid harmful interactions or offer protection against COVID-19, thereby reducing potential complications [27].

Aim

This study aimed to investigate the association between chronic medication and COVID-19 outcomes, using machine learning to identify key medication-related factors.

Ethics approval

This study obtained ethical clearance from the Ethics Committee from Vall d'Hebrón in 2020 (reference number: PR (AG) 240/2020) and received approval through a public call for grants to utilize the PADRIS databases in research projects.

Method

Study population, registry features, and data acquisition

The Data Analytics Program for Research and Innovation in Health (PADRIS) provided the data used in this study, drawing administrative data primarily from the Electronic Health System of Catalonia, centralized by the Agency for Health Quality and Assessment of Catalonia (AQuaS). This study was not in any clinical trial registry.

The final PADRIS pseudonymised database was built by integrating six administrative databases: patient demographics and socioeconomic data, COVID-19 diagnoses, comorbidities, drugs dispensed by community pharmacies, drugs dispensed by hospital pharmacies, and ICU admissions. It included 138,218 individuals diagnosed with COVID-19 in Catalonia between February 27 and September 14, 2020.

Medical diagnosis and comorbidities adhered to the International Statistical Classification of Diseases and Related Health Problems 10th Edition (ICD-10) [28]. The pharmacy invoicing data followed the Anatomical Therapeutic Chemical classification established by the World Health Organization Collaborating Centre (Supplementary Material 1) [29]. Chronic medication use was defined as a drug dispensed from the community or hospital pharmacy for at least 30 days prior to COVID-19 diagnosis, with continued use at diagnosis. For patients that initiated a medication within 30 days of diagnosis, that exposure was classified as non-use and excluded, consistent with previous literature [30, 31].

SARS-CoV-2 infection was identified via positive nucleic acid amplification tests, approved rapid antigen tests or serology tests indicating recent infection, or clinical symptoms with close epidemiological links to confirmed cases [32].

The Adjusted Morbidity Groups (GMA) categorised the population into five groups attributed to the risk of hospital admission and overall survival, and were used for comorbidity risk assessment [33]. The socioeconomic status information classified individuals into four groups based on annual gross income.

Data processing

Data pre-processing was performed using Pentaho Data Integration 9.0.0.0–423. During extraction, missing or inconsistent data, non-study participants, and duplicates were identified and corrected, or removed. Participants missing GMA values ($n = 7089$) were excluded, with no significant impact on variable distribution. Transformation steps included homogenisation of formats, unifying ICD or ATC nomenclatures (Supplementary Material 1), and generating binary variables for drugs and diseases. Comorbidities were cross-validated against administered medications, revealing under-registration of HIV, which was corrected by reclassifying participants on HIV-specific treatments. No data imputation was performed. In the load phase, data were de-identified and integrated into a unified dataset. Privacy measures included numeric value replacement and removal of variables with fewer than 10 participants. The resulting database was reviewed by the data protection committee before use.

Statistical analyses

Three main analyses were conducted. First, a univariate descriptive analysis was performed for all variables, including comparative analyses of potential predictive factors, stratified by age and sex. Next, prediction models for each COVID-19 outcome considered (hospitalisation, ICU admission, mortality, and in-hospital mortality) were developed using eXtreme Gradient Boosting (XGBoost). Medication and comorbidities were modelled separately, with both approaches incorporating demographic and socioeconomic features.

The dataset was split into a validation set (10%), and then outcomes were balanced by undersampling before being further divided into training (70%) and testing (30%) sets. The data partitioning process was uniformly applied across all datasets to maintain methodological consistency and model fairness. However, if there were underrepresented groups in the original data, they remained equally underrepresented in each partition. Categorical variables were one-hot encoded, and grid search hyperparameter tuning was applied to optimize the area under the receiver operating characteristic curve (AUROC) using log-loss. Models were trained using tenfold cross-validation, repeated 10 times. To enhance interpretability, SHapley Additive exPlanations (SHAP) were used to determine feature attributions [34]. SHAP summary plots were generated to assess feature importance and their effects on model predictions.

Finally, a sensitivity analysis was conducted on participants with arterial hypertension, diabetes mellitus, and lipid disorders to evaluate the impact of specific prescribed medications on the outcomes. Odds ratios (OR) with 95%

confidence intervals (CI) were calculated. In all analyses, participants were stratified into two age groups: 18 to 65 and over 65.

Analyses were conducted using R software (version 4.3.0), with specific packages and code accessible in GitHub (https://github.com/bertamiro/XGBoost_code_COVID_outcomes).

Results

Characteristics of the cohort

The final dataset had 121,650 participants (Fig. 1), with a mean age of 53 years (SD 20), 68,528 (57%) of which were female (Supplementary Material 2). In the logistic regression analysis for participants aged 18 to 65, age at exposure was the strongest predictor of hospitalisation (60–64 years OR 12.75, CI 11.03–14.82). Other factors were rituximab

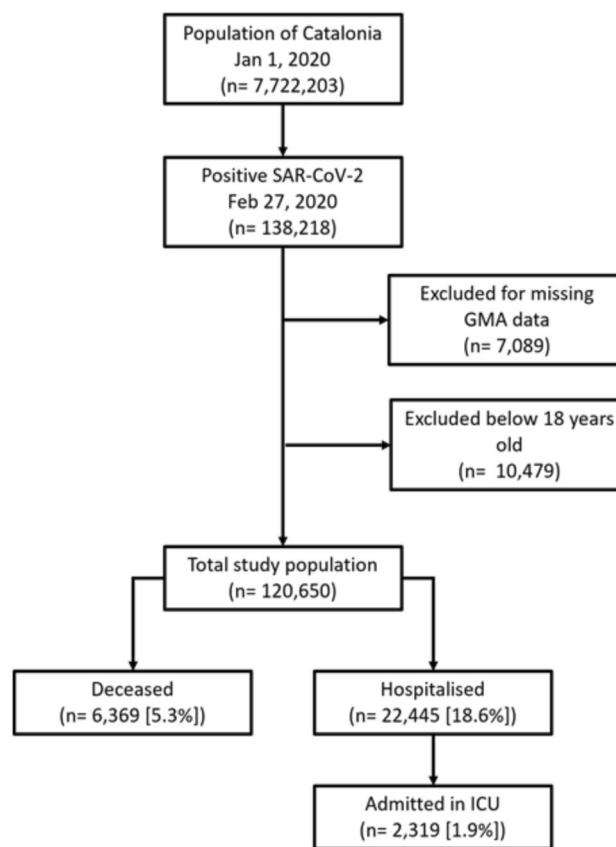


Fig. 1 Description of the cohort used in the study. Data was collected from several data administrative databases centralized by AQUAS within the PADRIS framework. Those participants with missing GMA data were excluded from the study, as those younger than 18 years old. The final dataset used for analysis had 121,650 participants and included 255 variables spanning sociodemographic factors, comorbidities, chronic medication, and clinical outcomes

use (OR 15.19, CI 7.06–35.26), other antivirals (OR 14.37, CI 1.38–309.14), polypharmacy (OR 8.15, CI 4.06–16.52), and heparins (OR 8.07, CI 6.74–9.66). Among participants over 65, rituximab (OR 3.08 CI 1.56–6.31), calcineurin inhibitors (OR 3.01 CI 2.15–4.26), and Janus Kinase inhibitors (OR 2.96 CI 1.3–7.12) were most strongly associated with hospitalisation risk (Supplementary Material 3). Age was less important in the older groups (75–79 age OR 1.22, CI 1.13–1.31).

Risk of mortality among participants aged 18 to 65 was also determined by age (OR 74.21, CI 28.33–299.93). Other factors included rituximab use (OR 18.91, CI 4.49–54.1), polypharmacy (OR 16.3, CI 3.9–46.04), heart failure (OR 16.05, CI 11.35–22.09), and systemic corticosteroids (OR 15.39, CI 11.69–19.96) (Supplementary Material 4). In participants over 65, age remained a key determinant of mortality risk (85–89 years OR 4.15, CI 3.7–4.66), GMA score (very high risk GMA OR 2.87, CI 2.54–3.25), calcineurin inhibitors (OR 2.15, CI 1.48–3.07), and selective immunosuppressants (OR 1.97, CI 1.45–2.64; Supplementary Material 5).

The risk of ICU admission was also strongly associated with age (60–64 years OR 18.42, CI 11.82–30.71;), rituximab use (OR 14.87, CI 4.99–36.13), calcineurin inhibitors (OR 8.66, CI 5.51–13.02), direct-acting anticoagulants (OR 7.99, CI 4.57–13.03), and polypharmacy (OR 7.06, CI 1.69–19.9, Supplementary Material 6). Among participants over 65, the most relevant factors for ICU admission were calcineurin inhibitors (OR 3.53, CI 1.94–5.94), a history of transplants (OR 3.13, CI 1.72–5.24), male sex (OR 3.01, CI 2.65–3.43), and the use of selective immunosuppressants (OR 2.53, CI 1.46–4.1; Supplementary Material 7).

Finally, features associated with in-hospital death in participants aged 18 to 65 were age (60–64 years OR 8.53, CI 2.7–51.77), use of calcineurin inhibitors (OR 6.38, CI 3.72–10.39), selective immunosuppressants (OR 5.8, CI 3.49–9.2) or polypharmacy (OR 5.69, CI 1.31–17.51; Supplementary Material 8). Comparable characteristics to mortality risk, whether occurring in or outside the hospital were found relevant in participants over 65 (Supplementary Material 9).

Machine learning model performance and interpretability

Eight distinct models were evaluated, one for each outcome and age group combinations. This approach allowed isolating the impact of pharmacological variables while minimizing collinearity with the disease-related variables for which these medications were prescribed. XGBoost was selected due to its robust tree-based ensemble, enabling the modelling of non-linear relationships, mitigation of predictor collinearity, and superior performance compared to K-Nearest

Neighbors, AdaBoost, and Support Vector Machine methods (Supplementary Material 10). Pharmacological models demonstrated similar performance overall (Figs. 2 and 3). However, models for participants older than 65 performed worse. For hospitalisation risk, the model for participants aged 18 to 65 (Fig. 2a) achieved an AUC of 0.70 (CI 0.69–0.71), compared to 0.58 (CI 0.57–0.60) for those older than 65 (Fig. 3a). Similarly, for mortality risk in participants aged 18 to 65 (Fig. 2b) showed strong discrimination (AUC 0.89, CI 0.85–0.92), versus 0.65 (CI 0.63–0.67) model in older participants (Fig. 3b).

For ICU admission, the model for participants aged 18 to 65 (Fig. 2c) exhibited a similar performance to that of the model for older participants (Fig. 3c), with AUC values of 0.74 (CI 0.70–0.78) and 0.73 (CI 0.67–0.79), respectively. In contrast, the models predicting the risk of death during hospitalisation demonstrated greater variability: it achieved an AUC of 0.53 (CI 0.51–0.55) for participants aged 18 to 65, while it reached an AUC of 0.65 (CI 0.64–0.66) for those over 65 (Figs. 2d and 3d).

Models based on disease-related features performed similarly to those using medication data. The best performing model predicted risk of death in participants aged 18 to 65 (AUC 0.92, CI 0.89–0.95), while the hospitalisation model for those older than 65 showed the lowest performance (AUC 0.59, CI 0.58–0.61; Supplementary Material 11).

Variable importance

Most models consistently identified age and sex as the most influential features, with older age and male sex linked to worst COVID-19 outcomes. In participants aged 18 to 65, hospitalisation risk was associated with the number of prescribed drugs and heparin use (Fig. 4a). In participants older than 65, the use of systemic corticosteroids was also linked to a higher risk (Fig. 4b).

In mortality models, key features for participants aged 18 to 65 included the use ACEi, ARBs, and HMG-CoA reductase inhibitors, all linked to increased risk (Fig. 5a); in those older than 65, ACEi and systemic corticosteroids were the main determinants (Fig. 5b).

Predictors of ICU admission in both age groups included the use of HMG-CoA reductase inhibitors and ACEi or ARBs (Fig. 6a, b). While ACEi and ARBs were linked to a slightly lower ICU admission risk, participants using HMG-CoA reductase inhibitors were at higher risk.

In models evaluating COVID-19 outcomes and chronic diseases, main predictors of hospitalisation in participants aged 18 to 65 included age, sex, obesity, social status, smoking, and metabolic disorders (Supplementary Material 12), while in those older than 65, key features were age, dementia, sex, and social status (Supplementary Material 12). Mortality risk models identified age, sex, neoplasms, social

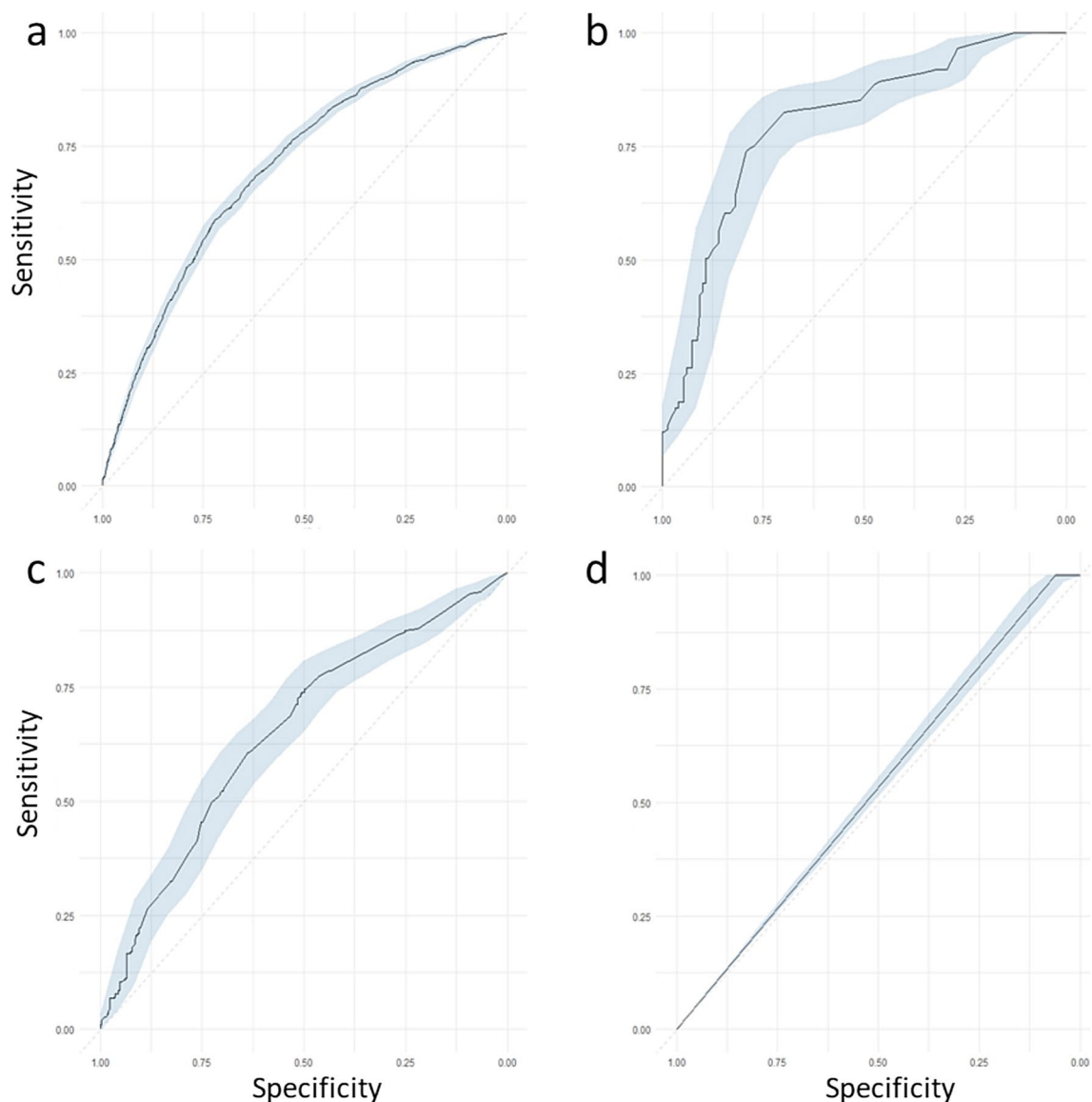


Fig. 2 ROC values with a 95% CI for the pharmacological models evaluated in participants 18 to 65. XGBoost models for prescription medication-related features were built for the four outcomes: **a** risk of

hospitalisation from COVID-19, **b** risk of death, **c** risk of ICU admission, and **d** risk of death while hospitalized separately in participants aged 18 to 65 years old

status, COPD, and diabetes as significant predictors in the 18 to 65 age group (Supplementary Material 13); whereas age, sex, renal insufficiency, and metabolic disorders were the most relevant factors in those older than 65 (Supplementary Material 13).

Finally, the top predictors of ICU admission in participants aged 18 to 65 were sex, hypertension, obesity, and diabetes (Supplementary Material 14), whereas in those older than 65, age, sex, dementia, metabolic disorders, and GMA were most relevant (Supplementary Material 14). Local model predictions for hospitalisation risk in three representative participants aged 18 to 65 are shown in Supplementary Material 15, with accurate predictions at $p < 0.05$.

Sensitivity analysis

A multivariate analysis was conducted to evaluate the relationship between selected comorbidities and their associated medications, specifically hypertension (ACEi and ARBs), diabetes (biguanides and DPP4 inhibitors), and metabolic disorders (statins).

Participants receiving ACEi or ARBs for hypertension did not show an increased risk for any of the outcomes compared to those with compared to patients not using ACEi or ARBs to treat hypertension.. However, in individuals over 65, treatment with ACEi and ARBs was associated with a reduced risk of death (OR 0.78, 95% CI 0.68–0.92 and OR

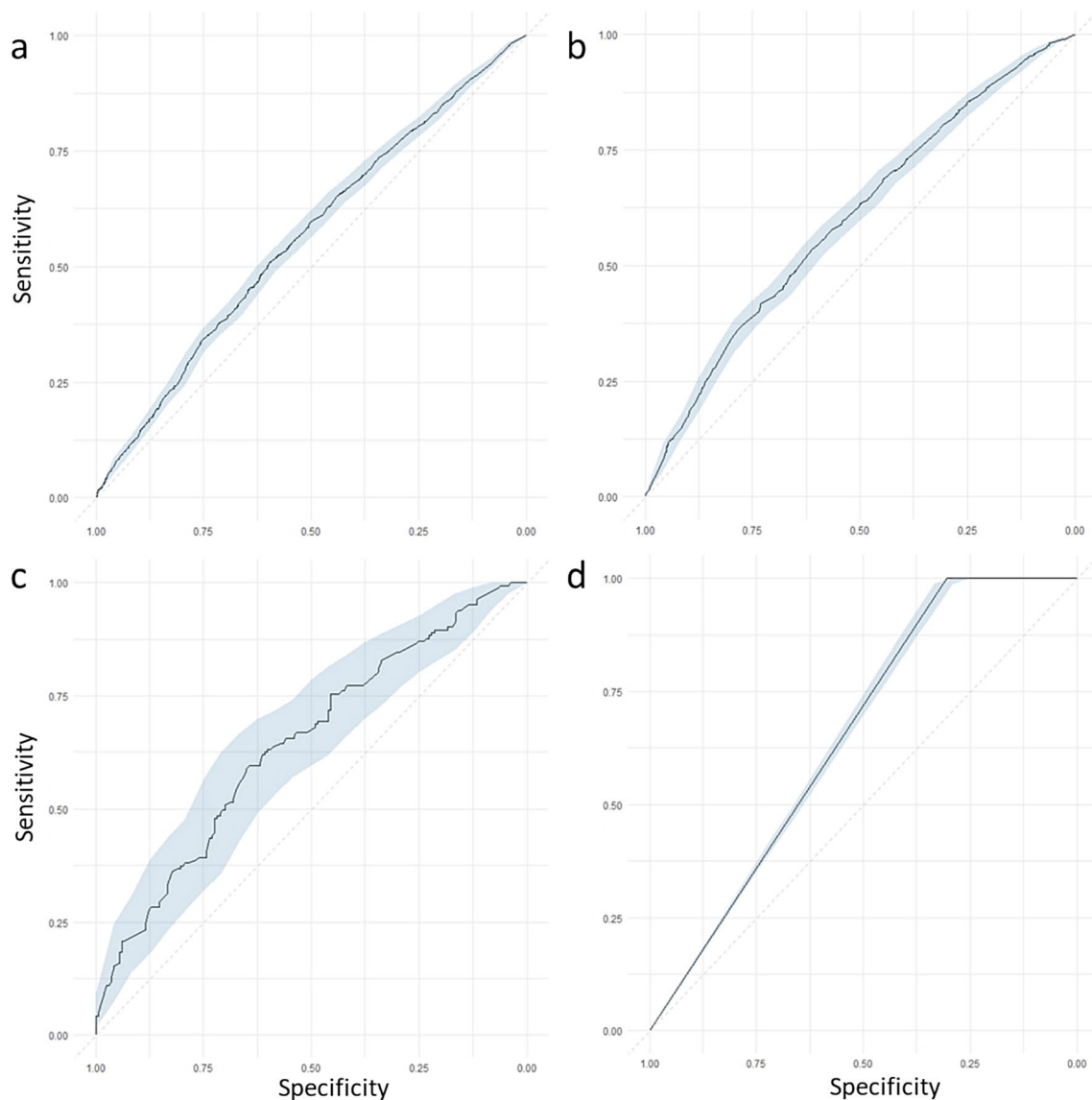


Fig. 3 ROC values with a 95% CI for the pharmacological models evaluated in participants above 65. XGBoost models for prescription medication-related features were built for the four outcomes: **a** risk of

hospitalisation from COVID-19, **b** risk of death, **c** risk of ICU admission, and **d** risk of death while hospitalized separately in participants aged older than 65 years

0.8, 95% CI: 0.68–0.95, respectively, Supplementary Material 17). Among participants aged 65 or below, none of the analysed medications were linked to an increased risk of hospitalisation or ICU admission (Supplementary Material 16, Supplementary Material 18).

For diabetic participants, treatment with DPP4 inhibitors in the 18 to 65 age group was associated with a higher risk of death (OR 3.03, 95% CI 1.17–7.00, Supplementary Material 17). In contrast, among participants over 65, biguanides demonstrated a protective effect on overall mortality (OR 0.79, 95% CI 0.68–0.92), an effect not observed with DPP4 inhibitors (OR 1.2, 95% CI 0.96–1.49). Furthermore, when comparing biguanides to DPP4 inhibitors in participants

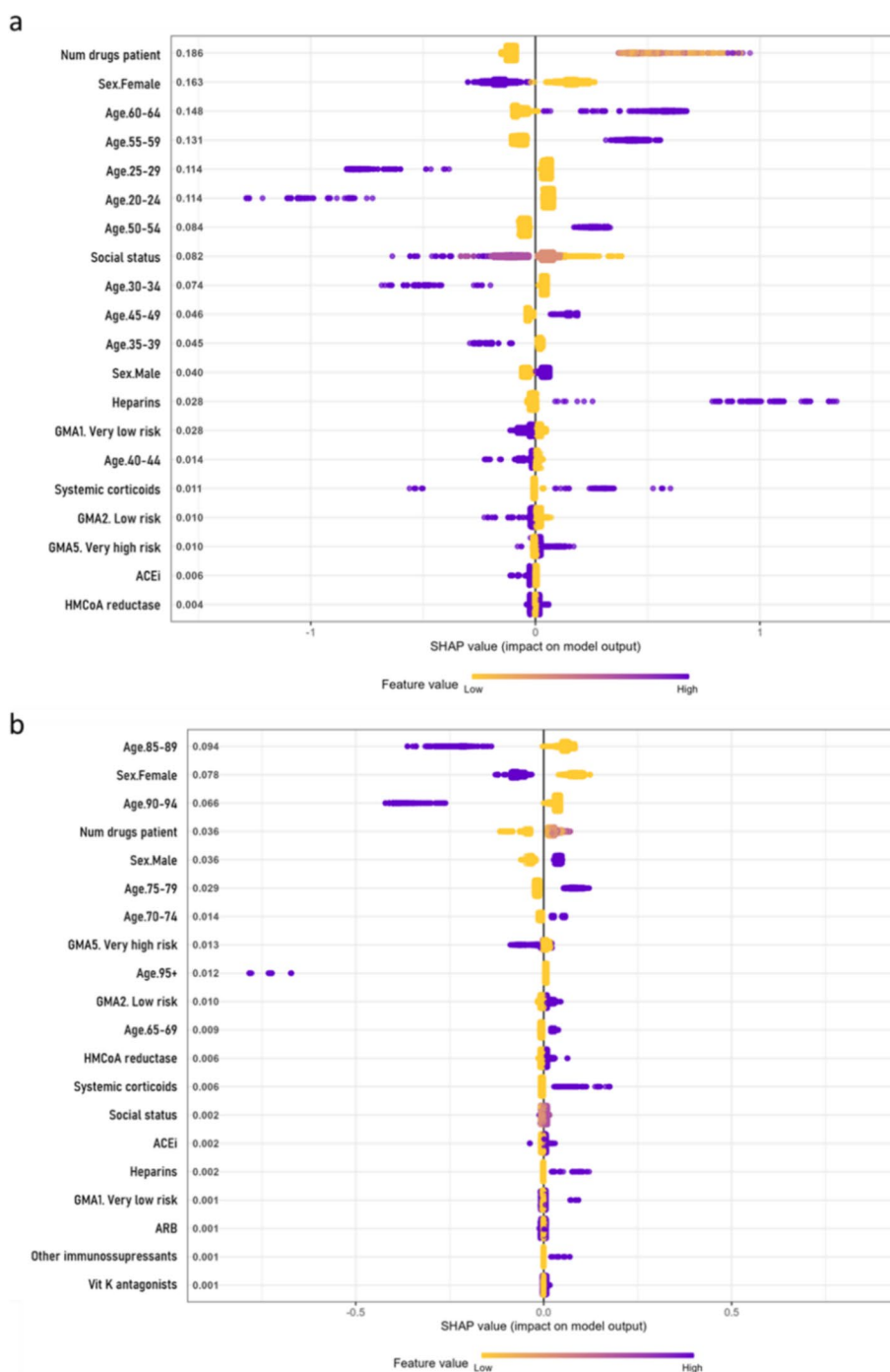
older than 65, biguanides consistently showed better outcomes across all measures. Statins did not provide a protective effect for any of the assessed outcomes in either age group.

Discussion

Statement of key findings

This study showed that models for chronic medication exhibited AUCROC values ranging from 0.70 to 0.89 for COVID-19 participants aged 18 to 65; however, these predictions

Fig. 4 SHAP summary plot for hospitalisation risk in the chronic medication and demographics model. Variable importance (SHAP values) for the XGBoost model to determine the risk of hospitalisation from COVID-19 in participants by age groups: **a** from 18 to 65 years old, and **b** above 65 years old, for the pharmacological features model. SHAP values are ranked in descending order based on the absolute value of their influence on the XGBoost model; high values mean a higher probability of hospitalisation. Purple represents high variable values and yellow represents lower variable values (in the categorical variables one-hot-encoded, it is 0 yellow and 1 purple). Each point represents an instance (participant) for that variable in the dataset



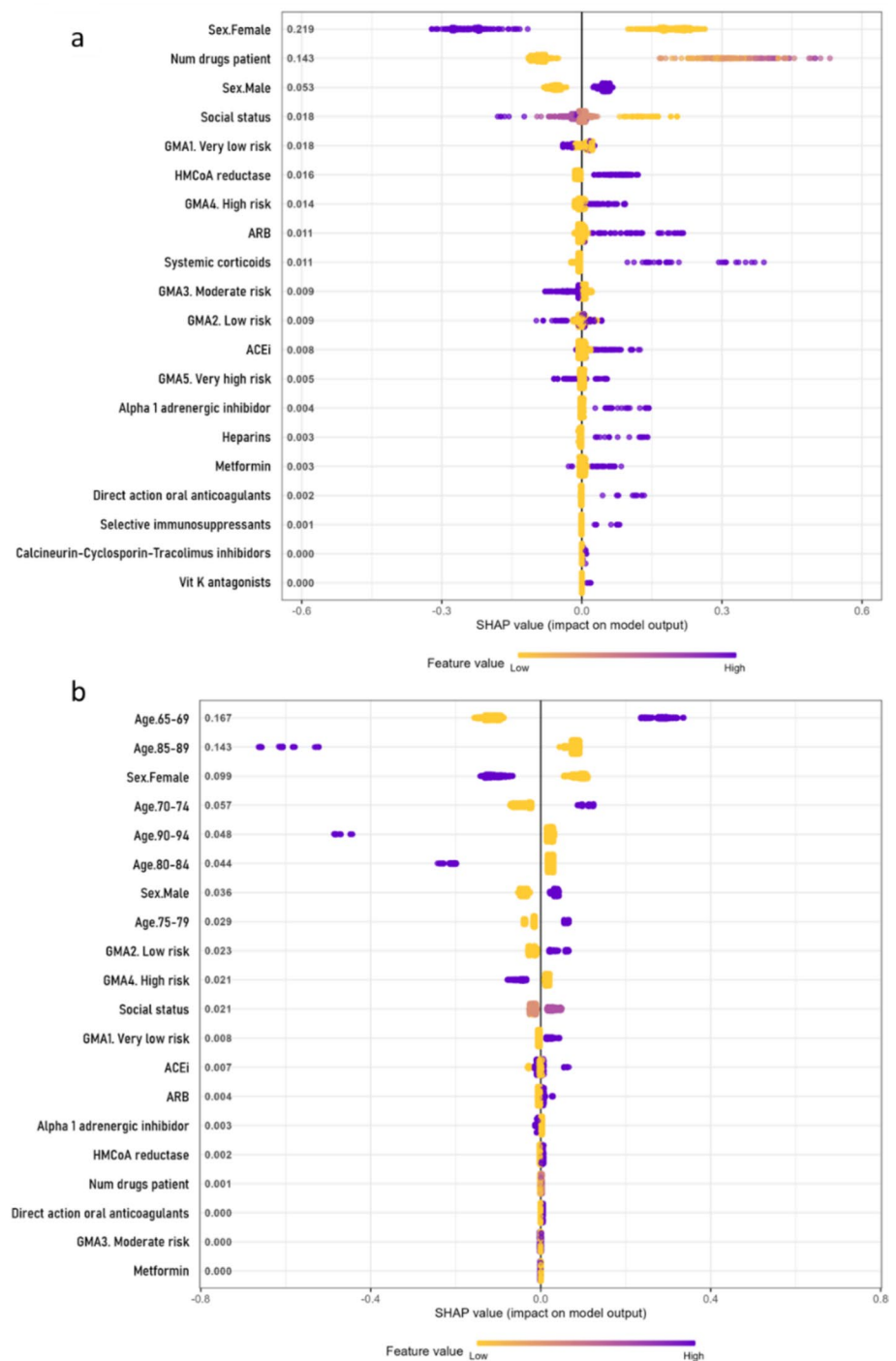
showed a poorer fit in participants older than 65 for all tested outcomes (AUROC between 0.58 and 0.73). This could be attributed to the more complex context observed in elderly participants. Age, sex, number of drugs, and previous pathologies were consistently important predictors across models. Hypertension drugs such as ACEi and ARBs were associated especially in the risk of death and ICU admission in participants aged 18 to 65. By contrast, the sensitivity study found that in participants with arterial hypertension

older than 65, ACEi and ARBs were protective factors for death when compared with participants with arterial hypertension without other medications (OR 0.78, and OR 0.8, respectively).

Interpretation

We conducted a retrospective study of chronic medication use employing XGBoost to identify key factors and predict

Fig. 5 SHAP summary plot for mortality risk in the chronic medication and demographics model. Variable importance (SHAP values) for the XGBoost model to determine the risk of death from COVID-19 in participants by age groups: **a** from 18 to 65 years old, and **b** above 65 years old, for the pharmacological features model. SHAP values are ranked in descending order based on the absolute value of their influence on the XGBoost model; high values mean a higher probability of hospitalisation. Purple represents high variable values and yellow represents lower variable values (in the categorical variables one-hot-encoded, it is 0 yellow and 1 purple). Each point represents an instance (participant) for that variable in the dataset

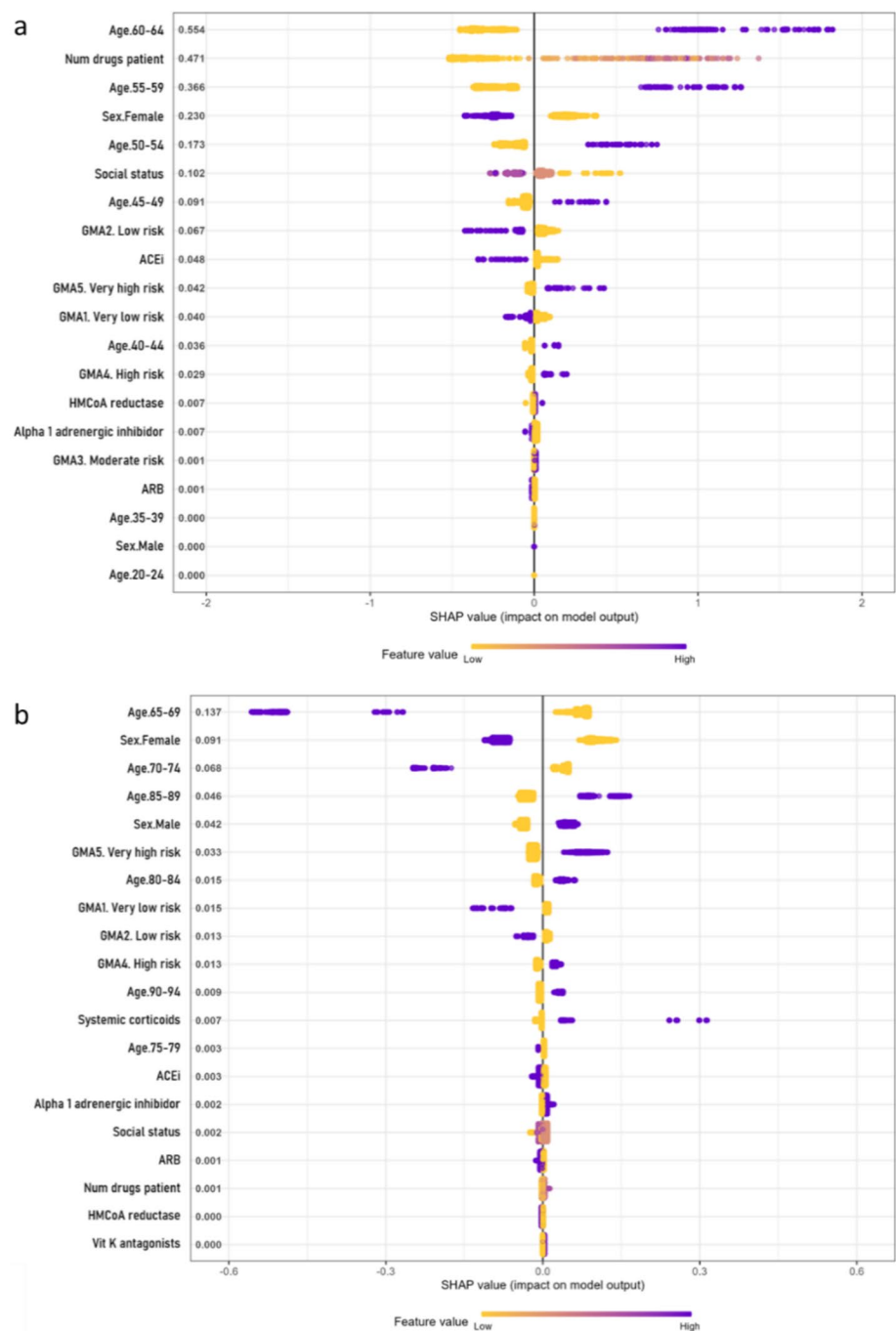


relevant outcomes. The models achieved robust predictive power for COVID-19 participants aged 18 to 65, but showed diminished performance in participants over 65, likely reflecting population heterogeneity and unaccounted confounding factors [35]. While prior research examined chronic medication, many studies had smaller sample sizes or relied on population-based statistical methods [36]. This study highlighted the role of age, sex, chronic medication, socioeconomic factors, and pre-existing conditions in

predicting COVID-19 outcomes. Our findings align with previous evidence identifying these factors and pre-existing comorbidities as key determinants of COVID-19 severity [37–41].

As pathologies increase with age, so does medication use and its associated side effects, all of which influence COVID-19 severity [42]. We found that the number of prescribed drugs correlated with a higher risk of negative outcomes. Other studies have linked polypharmacy to increased

Fig. 6 SHAP summary plot for ICU admission risk in the chronic medication and demographics model. Variable importance (SHAP values) for the XGBoost model to determine the risk of ICU admission from COVID-19 in participants by age groups: **a** from 18 to 65 years old, and **b** above 65 years old, for the pharmacological features model. SHAP values are ranked in descending order based on the absolute value of their influence on the XGBoost model; high values mean a higher probability of hospitalisation. Purple represents high variable values and yellow represents lower variable values (in the categorical variables one-hot-encoded, it is 0 yellow and 1 purple). Each point represents an instance (participant) for that variable in the dataset



vulnerability [27, 43] and a higher risk of hospitalisation and mortality from COVID-19.

Conflicting findings regarding ACEi and ARBs emphasise the need for subpopulation-specific analyses [44–46]. Although these drugs were associated with a higher risk of

death when compared with participants who did not take antihypertensive drugs, they were inversely associated with ICU admission risk in the same age group. While XGBoost suggested an association between these drugs and increased mortality in participants aged 18 to 65, our sensitivity

analysis did not confirm this. Instead, we found a protective effect in hypertensive participants taking these drugs compared with hypertensive participants taking other anti-hypertensive drugs [47]. The discrepancy likely arises from differences in cohort composition; XGBoost included both hypertensive and non-hypertensive participants, but the sensitivity analysis focused solely on hypertensive individuals.

We also observed a potential protective effect of metformin against severe COVID-19 outcomes, likely due to its immunomodulatory properties [48]. This aligns with hypotheses suggesting its role in mitigating hyper-inflammatory responses. In the sensitivity study, treatment with DPP4 inhibitors conferred higher probability of death in younger diabetic participants [49], whereas biguanides had a protective effect on overall fatality rates in diabetic participants over 65 [50]. Metformin could reduce viral affinity for angiotensin-converting enzyme 2 and inhibit inflammation via adenosine monophosphate-activated protein activation, potentially protecting against severe COVID-19 [51, 52].

ML-driven patient prioritization tools are needed in emergency settings [53]; our analysis of medications associated with severe COVID-19 outcomes represents a foundational step toward their development.

Our results highlight strong associations between socioeconomic vulnerability and COVID-19 severity and mortality [54, 55].

Strengths and weaknesses

This study uniquely analyses the relationship between prescribed chronic medication and COVID-19 outcomes using a large cohort. Our findings are concordant with other populations studied with different methodologies, reinforcing their robustness. Sensitivity analyses further provided insights into specific medications within hypertensive and diabetic subpopulations.

However, the study has limitations. Results depend on age and sex, factors that may limit the impact of other predictors. Additionally, medication data, sourced from community and hospital pharmacy invoicing systems, cannot guarantee adherence. To mitigate this, only medications dispensed within three months prior to inclusion were considered. Moreover, medications were grouped by active ingredient, without accounting for dosage variations. Furthermore, the evolving COVID-19 management criteria during the pandemic introduced potential inconsistencies. Lastly, models performed poorly for participants over 65, limiting results interpretation for this relevant age group.

Future research

Future studies should investigate the mechanisms underlying the protective effects of ACEi, ARBs, and metformin in specific subpopulations. Additionally, research to develop predictive models tailored to elderly participants is still needed, since they are more prone to severe COVID-19 outcomes. It would be interesting to conduct prospective studies to validate these findings and explore causal relationships between chronic medication and COVID-19 outcomes. Finally, the impact of socioeconomic factors on COVID-19 outcomes requires further exploration to design equitable healthcare interventions.

Conclusion

Machine learning predictive models have high performance predicting hospitalisation, ICU admission, and death in COVID-19 participants. Key factors influencing outcomes include age, sex, comorbidities, socioeconomic status and chronic medication. Our findings support decreased risks of poor outcomes in participants with arterial hypertension under ACEi or ARBs treatment. Additionally, biguanides showed potential protective effects for participants with diabetes. Participants taking metformin, ACEi, or ARBs should not be switched to alternative medications in relation to COVID-19.

Our findings leveraged administrative databases and advanced AI querying tools. Given their regular updates and compliance with EU data protection regulations, the methodology is readily applicable with limited additional system cost.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s11096-025-01955-7>.

Funding Open Access Funding provided by Universitat Autònoma de Barcelona. This study was supported by the Spanish Foundation for Hospital Pharmacy (FEFH) and the Spanish Society of Hospital Pharmacy (SEFH) through the call for research grants 2022–2023. The Agency for Health Quality and Assessment of Catalonia (AQuaS) provided two data scientists for extracting, cleaning, merging and anonymizing the data from institutional databases for the development of this project. ASM is supported by a Juan Rodés (JR18/00022) postdoctoral fellowship from ISCIII. ASP is supported by the Spanish Ministerio de Ciencia e Innovación, grant PID2019-104830RB-I00, and by the Departament d'Economia i Coneixement de la Generalitat de Catalunya, grant 2021SGR01421 (GRBIO). The funding sources had no involvement in any aspect of this manuscript.

Conflicts of interest The authors declare they have no conflict of interest.

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/>.

References

1. Zhou P, Yang XL, Wang XG, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*. 2020;579(7798):270–3. <https://doi.org/10.1038/s41586-020-2012-7>.
2. Liu S, Luo P, Tang M, et al. Providing pharmacy services during the coronavirus pandemic. *Int J Clin Pharm*. 2020;42:299–304. <https://doi.org/10.1007/s11096-020-01017-0>.
3. Clift AK, Coupland CAC, Keogh RH, et al. Living risk prediction algorithm (QCOVID) for risk of hospital admission and mortality from coronavirus 19 in adults: national derivation and validation cohort study. *BMJ*. 2020;371: m3731. <https://doi.org/10.1136/bmj.m3731>.
4. Das AK, Mishra S, Saraswathy GS. Predicting CoVID-19 community mortality risk using machine learning and development of an online prognostic tool. *PeerJ*. 2020;8: e10083. <https://doi.org/10.7717/peerj.10083>.
5. van Halem K, Bruyndonckx R, van der Hilst J, et al. Correction to: Risk factors for mortality in hospitalized patients with COVID-19 at the start of the pandemic in Belgium: a retrospective cohort study. *BMC Infect Dis*. 2020;20(1):956. <https://doi.org/10.1186/s12879-020-05690-4>.
6. Ikemura K, Bellin E, Yagi Y, et al. Using automated machine learning to predict the mortality of patients with COVID-19: prediction model development study. *J Med Internet Res*. 2021;23(2): e23458. <https://doi.org/10.2196/23458>.
7. Ramos-Rincon JM, Buonaiuto V, Ricci M, et al. Clinical characteristics and risk factors for mortality in very old patients hospitalized with COVID-19 in Spain. *J Gerontol A Biol Sci Med Sci*. 2021;76(3):e28–37. <https://doi.org/10.1093/gerona/glaa243>.
8. Rinderknecht MD, Klopfenstein Y. Predicting critical state after COVID-19 diagnosis: model development using a large US electronic health record dataset. *NPJ Digit Med*. 2021;4(1):113. <https://doi.org/10.1038/s41746-021-00482-9>.
9. Cheng J, Sollee J, Hsieh C, et al. COVID-19 mortality prediction in the intensive care unit with deep learning based on longitudinal chest X-rays and clinical data. *Eur Radiol*. 2022;32(7):4446–456. <https://doi.org/10.1007/s00330-022-08588-8>. Epub 2022 Feb 19. Erratum in: *Eur Radiol*. 2022;32(7):5034. <https://doi.org/10.1007/s00330-022-08680-z>.
10. Jeffrey K, Woolford L, Maini R, et al. Prevalence and risk factors for long COVID among adults in Scotland using electronic health records: a national, retrospective, observational cohort study. *eClinicalMedicine*. 2024;71:102590. <https://doi.org/10.1016/j.eclinm.2024.102590>.
11. Dillon P, Siadimas A, Roumpanis S, et al. Electronic health record data for assessing risk of hospitalisation for COVID-19: methodological considerations applied to multiple sclerosis. *Mult Scler Relat Disord*. 2023;71: 104512. <https://doi.org/10.1016/j.msard.2023.104512>.
12. Zang C, Zhang Y, Xu J, et al. Data-driven analysis to understand long COVID using electronic health records from the RECOVER initiative. *Nat Commun*. 2023;14(1):1948. <https://doi.org/10.1038/s41467-023-37653-z>.
13. Barda N, Riesel D, Akriv A, et al. Developing a COVID-19 mortality risk prediction model when individual-level data are not available. *Nat Commun*. 2020;11(1):4439. <https://doi.org/10.1038/s41467-020-18297-9>.
14. Caruso C, Marcon G, Accardi G, et al. Role of sex and age in fatal outcomes of COVID-19: women and older centenarians are more resilient. *Int J Mol Sci*. 2023;24(3):2638. <https://doi.org/10.3390/ijms24032638>.
15. Pojero F, Candore G, Caruso C, et al. The role of immunogenetics in COVID-19. *Int J Mol Sci*. 2021;22:2636. <https://doi.org/10.3390/ijms22052636>.
16. Kuno T, Sahashi Y, Kawahito S, et al. Prediction of in-hospital mortality with machine learning for COVID-19 patients treated with steroid and remdesivir. *J Med Virol*. 2022;94(3):958–64. <https://doi.org/10.1002/jmv.27393>.
17. Tiwari S, Dogan O, Jabbar MA, et al. Applications of machine learning approaches to combat COVID-19: A survey. In: Kaklauskas A, Abraham A, Okoye K, Guggari S, editors. *Lessons from COVID-19*. Academic Press; 2022. p. 263–87. <https://doi.org/10.1016/B978-0-323-99878-9.00014-5>.
18. Malekpour MR, Abbasi-Kangevari M, Shojaei A, et al. Effect of chronic medication use on outcome measures of hospitalized COVID-19 patients: evidence from big data. *Front Public Health*. 2023;11:1061307. <https://doi.org/10.3389/fpubh.2023.1061307>.
19. Larrosa-García M, García-García S, Louro J, et al. Use of chronic medications and risk of severe death due to COVID-19 in hospitalised patients. *Eur J Hosp Pharm*. 2024;31:247–52. <https://doi.org/10.1136/ejpharm-2021-003186>.
20. Israel A, Schäffer AA, Cicurel A, et al. Identification of drugs associated with reduced severity of COVID-19: a case-control study in a large population. *Elife*. 2021;10:e68165. <https://doi.org/10.1101/2020.10.13.20211953>.
21. Zou W, Yao X, Chen Y, et al. An elastic net regression model for predicting the risk of ICU admission and death for hospitalized patients with COVID-19. *Sci Rep*. 2024;14(1):14404. <https://doi.org/10.1038/s41598-024-64776-0>.
22. Flacco ME, Acuti Martellucci C, Bravi F, et al. Treatment with ACE inhibitors or ARBs and risk of severe/lethal COVID-19: a meta-analysis. *Heart*. 2020;106(19):1519–24. <https://doi.org/10.1136/heartjnl-2020-317336>.
23. González-Cebrián A, Borrás-Ferrís J, Ordovás-Baines JP, et al. Machine-learning-derived predictive score for early estimation of COVID-19 mortality risk in hospitalized patients. *PLoS ONE*. 2022;17(9): e0274171. <https://doi.org/10.1371/journal.pone.0274171>.
24. Casillas N, Ramón A, Torres AM, et al. Predictive model for mortality in severe COVID-19 patients across the six pandemic waves. *Viruses*. 2023;15(11):2184. <https://doi.org/10.3390/v15112184>.
25. Bahsoun A, Fakhri Y, Zareef R, et al. Corticosteroids in COVID-19: pros and cons. *Front Med (Lausanne)*. 2023;10:1202504. <https://doi.org/10.3389/fmed.2023.1202504>.
26. Kretschy IA, Asiedu-Danso M, Kretschy JP. Medication management and adherence during the COVID-19 pandemic: perspectives and experiences from low-and middle-income countries. *Res Social Adm Pharm*. 2021;17(1):2023–6. <https://doi.org/10.1016/j.sapharm.2020.04.007>.
27. Abou J, Elders PJM, Huijts D, et al. Pharmacist-led deprescribing of cardiovascular and diabetes medication within a clinical medication review: the LeMON study (less medicines in older patients in the Netherlands), a cluster randomized controlled trial. *Int J Clin Pharm*. 2025. <https://doi.org/10.1007/s11096-025-01863-w>.

28. World Health Organization. ICD-10: international statistical classification of diseases and related health problems: tenth revision. WHO. 2004. <https://icd.who.int/browse10/2019/en>. Accessed 15 May 2025.
29. Anatomical Therapeutic Chemical (ATC) Classification. <https://www.who.int/tools/atc-ddd-toolkit/atc-classification>. Accessed 15 May 2025.
30. Taitel M, Fensterheim L, Kirkham H, et al. Medication days' supply, adherence, wastage, and cost among chronic patients in Medicaid. *Med Care*. 2012;50(6):552–6. <https://doi.org/10.5600/mmrr.002.03.a04>.
31. Miani C, Martin A, Exley J, et al. Clinical effectiveness and cost-effectiveness of issuing longer versus shorter duration (3-month vs. 28-day) prescriptions in patients with chronic conditions: systematic review and economic modelling. *Health Technol Assess*. 2017;21(78):1–128. <https://doi.org/10.3310/hta21780>.
32. COVID19_Estrategia_vigilancia_y_control_e_indicadores.pdf. https://www.sanidad.gob.es/profesionales/saludPublica/ccayes/alertasActual/nCov/documentos/COVID19_Estrategia_vigilancia_y_control_e_indicadores.pdf. Accessed 15 May 2025.
33. Monterde D, Vela E, Clèries M, et al. Adjusted morbidity groups: a new multiple morbidity measurement of use in primary care. *Aten Primaria*. 2016;48(10):674–82. <https://doi.org/10.1016/j.aprim.2016.06.003>.
34. Anjum M, Khan K, Ahmad W, et al. New SHapley Additive ExPlanations (SHAP) approach to evaluate the raw materials interactions of steel-fiber-reinforced concrete. *Materials* (Basel). 2022;15(18):6261. <https://doi.org/10.3390/ma15186261>.
35. Trevisan C, Del Signore S, Fumagalli S, et al. Assessing the impact of COVID-19 on the health of geriatric patients: The European GeroCovid Observational Study. *Eur J Intern Med*. 2021;87:29–35. <https://doi.org/10.1016/j.ejim.2021.01.017>.
36. Malekpour MR, Abbasi-Kangevari M, Shojae A, et al. Effect of the chronic medication use on outcome measures of hospitalized COVID-19 patients: evidence from big data. *Front Public Health*. 2023;11:1061307. <https://doi.org/10.3389/fpubh.2023.1061307>.
37. Boateng A, Maposa D, Mokobane R, et al. Analysis of COVID-19 cases and comorbidities using machine learning algorithms: a case study of the Limpopo Province, South Africa. *Sci Afr*. 2023;21:e01840. <https://doi.org/10.1016/j.sciaf.2023.e01840>.
38. Reina Reina A, Barrera JM, Valdivieso B, et al. Machine learning model from a Spanish cohort for prediction of SARS-CoV-2 mortality risk and critical patients. *Sci Rep*. 2022;12:5723. <https://doi.org/10.1038/s41598-022-09613-y>.
39. Passarelli-Araujo H, Passarelli-Araujo H, Urbano MR, et al. Machine learning and comorbidity network analysis for hospitalized patients with COVID-19 in a city in Southern Brazil. *Smart Health*. 2022;26: 100323. <https://doi.org/10.1016/j.smhl.2022.100323>.
40. Aktar S, Talukder A, Ahamad MM, et al. Machine learning approaches to identify patient comorbidities and symptoms that increased risk of mortality in COVID-19. *Diagnostics* (Basel). 2021;11(8):1383. <https://doi.org/10.3390/diagnostics11081383>.
41. Azizi Z, Shiba Y, Alipour P, et al. Importance of sex and gender factors for COVID-19 infection and hospitalisation: a sex-stratified analysis using machine learning in UK Biobank data. *BMJ Open*. 2022;12:e050450. <https://doi.org/10.1136/bmjopen-2021-050450>.
42. Ghasemi H, Darvishi N, Salari N, et al. Global prevalence of polypharmacy among the COVID-19 patients: a comprehensive systematic review and meta-analysis of observational studies. *Trop Med Health*. 2022;50:60. <https://doi.org/10.1186/s41182-022-00456-x>.
43. Sirois C, Boiteau V, Chiu Y, et al. Exploring the associations between polypharmacy and COVID-19-related hospitalisations and deaths: a population-based cohort study among older adults in Quebec, Canada. *BMJ Open*. 2022;12: e060295. <https://doi.org/10.1136/bmjopen-2021-060295>.
44. Fang L, Karakiulakis G, Roth M. Are patients with hypertension and diabetes mellitus at increased risk for COVID-19 infection? *Lancet Respir Med*. 2020;8(4): e21. [https://doi.org/10.1016/S2213-2600\(20\)30116-8](https://doi.org/10.1016/S2213-2600(20)30116-8).
45. Beyersstedt S, Casaro EB, Rangel ÉB. COVID-19: angiotensin-converting enzyme 2 (ACE2) expression and tissue susceptibility to SARS-CoV-2 infection. *Eur J Clin Microbiol Infect Dis*. 2021;40(5):905–19. <https://doi.org/10.1007/s10096-020-04138-6>.
46. Provenzani A, Polidori P. Covid-19 and drug therapy, what we learned. *Int J Clin Pharm*. 2020;42:833–6. <https://doi.org/10.1007/s11096-020-01049-6>.
47. Meng J, Xiao G, Zhang J, et al. Renin-angiotensin system inhibitors improve the clinical outcomes of COVID-19 patients with hypertension. *Emerg Microb Infect*. 2020;9(1):757–60. <https://doi.org/10.1080/22221751.2020.1746200>.
48. Mathews Samuel S, Varghese E, Büßelberg D. Therapeutic potential of metformin in COVID-19: reasoning for its protective role. *Trends Microbiol*. 2021;29(10):894–907. <https://doi.org/10.1016/j.tim.2021.03.004>.
49. Pedrosa AR, Martins DC, Rizzo M, et al. Metformin in SARS-CoV-2 infection: a hidden path—from altered inflammation to reduced mortality. A review from the literature. *J Diabetes Complicat*. 2023;37(2):108391. <https://doi.org/10.1016/j.jdiacomp.2022.108391>.
50. Chen X, Shi S, Sun H, et al. Metformin alleviates inflammatory response and severity rate of COVID-19 infection in elderly individuals. *Sci Rep*. 2025;15:11340. <https://doi.org/10.1038/s41598-025-96294-y>.
51. Almarsdóttir AB, Haq R, Nørgaard JDSV. Prioritizing patients for medication review by emergency department pharmacists: a multi-method study. *Int J Clin Pharm*. 2023;45:387–96. <https://doi.org/10.1007/s11096-022-01515-3>.
52. Ojeda-Fernández L, Foresta A, Macaluso G, et al. Metformin use is associated with a decrease in the risk of hospitalization and mortality in COVID-19 patients with diabetes: a population-based study in Lombardy. *Diabetes Obes Metab*. 2022;24(5):891–8. <https://doi.org/10.1111/dom.14648>.
53. Koufakis T, Pavlidis AN, Metallidis S, et al. Sodium-glucose cotransporter 2 inhibitors in COVID-19: meeting at the crossroads between heart, diabetes and infectious diseases. *Int J Clin Pharm*. 2021;43:764–7. <https://doi.org/10.1007/s11096-021-01256-9>.
54. Wachtler B, Michalski N, Nowossadeck E, et al. Socioeconomic inequalities and COVID-19—a review of the current international literature. *J Health Monit*. 2020;5(Suppl 7):3–17. <https://doi.org/10.25646/7059>.
55. Roman YM. COVID-19 pandemic: the role of community-based pharmacy practice in health equity. *Int J Clin Pharm*. 2022;44:1211–5. <https://doi.org/10.1007/s11096-022-01440-5>.

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