Title: Brief Report: Identifying common pharmacotherapies associated with reduced COVID-19 morbidity using electronic health records

Running Title: COVID-19 repositioning

Authors: Victor M. Castro MS (1), Rachel A. Ross MD, PhD (2,3), Sean M. McBride MD, PhD (4), Roy H. Perlis MD, MSc (1)*

Affiliations:
1. Center for Quantitative Health, Division of Clinical Research, Massachusetts General Hospital, 185 Cambridge Street, Boston, MA, USA
2. Division of Anxiety and Depressive Disorders, McLean Hospital, Belmont, MA, USA
3. Department of Psychiatry, Harvard Medical School, Boston, MA, USA.
4. Departments of Genetics and Psychiatry, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

Correspondence:
Roy H. Perlis, MD MSc
rperlis@mgh.harvard.edu
Massachusetts General Hospital
185 Cambridge Street, 6th Floor
Boston, MA 02114

Word count: 1744

Key Points

Question: Can electronic health records identify medications that may be associated with diminished risk of COVID-19 morbidity?

Findings: This cohort study across 5 hospitals identified medications enriched among individuals who did not require hospitalization for COVID-19 despite a positive test.

Meaning: While preliminary and subject to confounding, our results suggest that electronic health records may complement efforts to identify novel therapeutics for COVID-19 by identifying FDA-approved compounds with potential benefit in reducing COVID-19-associated morbidity.

Abstract

Importance: Absent a vaccine or any established treatments for the novel and highly infectious coronavirus-19 (COVID-19), rapid efforts to identify potential therapeutics are required.

Objective: To identify commonly-prescribed medications that may be associated with lesser risk of morbidity with COVID-19 across 5 Eastern Massachusetts hospitals.

Design: In silico cohort using electronic health records between 7/1/2019 and 4/07/2020.

Setting: Outpatient, emergency department and inpatient settings from 2 academic medical centers and 3 community hospitals.

Participants: All individuals presenting to a clinical site and undergoing COVID-19 testing.

Main Outcome or Measure: Inpatient hospitalization; documented requirement for mechanical ventilation.

Results: Among 12,818 individuals with COVID-19 testing results available, 2271 (17.7%) were test-positive, and 707/2271 (31.1%) were hospitalized in one of 5 hospitals. Based on a comparison of ranked electronic prescribing frequencies, medications enriched among test-positive individuals not requiring hospitalization included ibuprofen, valacyclovir, and naproxen. Among individuals who were hospitalized, mechanical ventilation was documented in 213 (30.1%); ibuprofen and naproxen were also more commonly prescribed among individuals not requiring ventilation.

Conclusions and Relevance: These preliminary findings suggest that electronic health records may be applied to identify medications associated with lower risk of morbidity with COVID-19, but larger cohorts will be required to address confounding by indication. Larger scale efforts at repositioning may help to identify FDA-approved medications meriting study for prevention of COVID-19 morbidity and mortality.

Funding: none.

Introduction

The rapid spread and mortality associated with the 2019 novel coronavirus (2019-nCoV; severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), or COVID19) necessitates efforts to find therapeutic strategies with unprecedented speed. Absent any treatments known to be effective against the novel virus, randomized trials have been launched for antiviral agents already in development for other indications[1], and in vitro studies[2] or in silico modeling efforts[3] have identified compounds with potential antiviral benefit.

If medications already known to be safe and approved for human use exhibited benefit, they could rapidly be deployed in clinical settings once efficacy was established in randomized trials. Such repositioning has been embraced enthusiastically in other contexts[4]. For example, we previously demonstrated that one commonly-prescribed medication could shorten, rather than prolong, QT interval[5]. However, with COVID-19, preliminary efforts to reposition the antimalarial medications hydroxychloroquine or chloroquine have been dogged by the recognition that enthusiasm far outstripped data[6].

In the present study, we aimed to identify medications that might be associated with decreased risk of COVID-19-associated morbidity, including hospitalization and mechanical ventilation. In particular, we examined medications more commonly prescribed in the period prior to hospitalization to individuals with documented infection who did not experience such morbidity. While adequately-powered and designed randomized trials are required to demonstrate efficacy, we hypothesized that a simple pharmacovigilance strategy examining enrichment of individual medications in the morbidity-free group would facilitate identification of candidate drugs for efficacy studies and prioritization of medications implicated by cellular screens.

Methods

Subjects

We included all individuals undergoing COVID-19 testing through one of 5 hospitals, including 2 academic medical centers and 3 community affiliate hospitals, between March 1, 2020 and April 7, 2020. Data for all of these individuals were drawn from the Partners Research Patient Data Registry (RPDR)[7] and used to generate an i2b2 datamart.[8] Specific data included age, sex, race/ethnicity, as well as all available narrative clinical notes from inpatient, outpatient, and emergency settings. Age-adjusted Charlson comorbidity index, a measure of overall burden of illness, was calculated using coded ICD9 and 10 diagnostic codes drawn from the EHR as previously described[9]. Presence of a coronavirus test, and test result, was determined from the enterprise laboratory feed (LOINC:94309-2).

The Partners HealthCare Human Research Committee approved the study protocol. As no participant contact was required in this study based on secondary use of data arising from routine clinical care, the committee waived the requirement for informed consent as detailed by 45 CFR 46.116.

Medication exposure

Medication exposures were defined based upon presence of at least one prescription event (categorized as RxNorm ingredient) in the electronic health record in the 8 months between July 1, 2019 and February 29, 2020, the last date for which medication information was available at time of

final analysis. (A sensitivity analysis requiring 2 prescription events did not meaningfully change results and is not described further.)

Hospital Course Characterization from Narrative Clinical Notes

Hospital services provided would typically be captured by coded clinical data (i.e., ICD-10 diagnosis plus CPT codes). However, such data is not available for some time after discharge, and not updated in real time in the RPDR; as such, it is poorly suited for rapidly detecting outcomes in the context of an emerging pandemic. Therefore, we developed and validated simple string-based classifiers to identify mechanical ventilation based on narrative clinical notes. A preliminary curated list of tokens was used to query a corpus of narrative notes; the strings resulting from these queries were manually reviewed in context to determine specificity, and iteratively revised. We also identified date of clinical service (emergency department evaluation, inpatient hospitalization) and date of onset of mechanical ventilation on the basis of index note of a given type.

Study Design and Analysis

Primary analysis sought to examine whether medications were over-represented among a subset of COVID-19 tested individuals. Specifically, to detect medications that might diminish morbidity, we considered individuals hospitalized versus not hospitalized among those testing positive, and secondarily individuals requiring mechanical ventilation during the study period versus not among those hospitalized.

Numerous methods exist for detecting enrichment between two groups. We elected to apply a simple rank-based method analogous to gene set enrichment analysis[10,11], in which we rank-ordered

individual medications (defined by RxNorm ingredients) [and secondarily, medication classes (defined by National Drug File - Reference Terminology groups, drawn from the UMLS ontology[12]] on the basis of frequency within each group - i.e., the group with versus without the outcome of interest. We then examined the medications with greatest *shift* in rank - that is, postulating that those which move 'up' the list are more likely to be beneficial. This approach postulates that medications with lower rank (ie, greater frequency) in the better-outcome group would be more likely to be associated with that better outcome. We then used multiple logistic regression with outcome as dependent variable and medication or group as predictor, with adjustment for age, sex, race, Charlson score, and site of COVID-19 test. As a hypothesis-generating study, we did not Bonferroni-correct for multiple contrasts, and focused on estimates of odds ratio. In keeping with RPDR practice to minimize risk of re-identification, in tables, all groups including fewer than 6 individuals are obfuscated. All analyses utilized R 3.6.0[13].

Results

A total of 12,818 individuals received COVID-19 testing with results available; 2271 (17.7%) were test positive. Among those with a positive test, 707/2271 (31.1%) were hospitalized in one of the 5 hospitals. Table 1 reports sociodemographic and clinical characteristics of individuals testing positive, distinguishing documented inpatient hospitalization or lack of hospitalization.

Table 2 indicates those medications with greatest enrichment (i.e., increase in rank-order) between hospitalized and not-hospitalized individuals and frequency among test-positive individuals of at least 1%, as well as odds ratios for hospitalization associated with each medication after adjustment for age, sex, race, ethnicity, site, and Charlson score. Electronically prescribed medications with greatest increase in rank among test-positive individuals not requiring hospitalization, and associated with odds ratios less than 1, included ibuprofen, valacyclovir, and naproxen; confidence intervals for

ibuprofen and naproxen exclude 1. (For all medications with frequency greater than 1% among test-positive individuals, see Supplemental Table 1; for most enriched medications by category, see Supplemental Table 2 and 3).

Among individuals who were hospitalized, mechanical ventilation was documented in 213 (30.1%); features of this cohort are reported in Table 3. Adjusted odds ratios for ventilation associated with each medication are listed in Table 2 (right columns); ibuprofen and naproxen were again associated with odds ratios for ventilation less than 1, albeit with confidence intervals including 1 (small counts preclude estimates for valacyclovir).

Discussion

This preliminary electronic health records study examined outcomes among 2,271 individuals who tested positive for COVID-19 and were hospitalized in one of 5 Boston-area hospitals between 3/1/2020 and 4/7/2020. As anticipated, among the most enriched medication classes in rank-based analysis (Supplemental Table 2) are those associated with lower-risk clinical groups. For example, oral contraceptives indicate premenopausal women; acne medications indicate younger individuals in general. While confirming the risk of confounding, they also suggest assay sensitivity: the rank-based analysis does recover lower-risk groups.

We also identify individual candidate medications that, if further supported in other cohorts, might be screened for repositioning as COVID-19 treatment or for prevention of morbidity. Most notably, two nonsteroidal antiinflammatory drugs appear to be enriched among patients not requiring hospitalization, and show qualitatively similar enrichment among patients not requiring mechanical ventilation during hospitalization. We emphasize that we cannot exclude confounding in this context

as well. However, it also seems plausible that suppression of an inflammatory response may diminish the immune consequences of COVID-19 infection that may contribute to pathogenesis[14]. Initial case reports of greater morbidity among NSAID-treated patients, postulated to arise from increases in the cellular COVID-19 receptor ACE-2, have subsequently been questioned[15]. Rank-based analysis also identifies valacyclovir as enriched among individuals not requiring hospitalization; while this too may reflect age differences in prescribing, it may also provide support for antiviral repositioning efforts[6].

Multiple FDA-approved medications have been suggested in in vitro studies to diminish coronavirus infection in general, or COVID19 in particular. For example, prior work identified multiple neurotransmitter inhibitors, including chlorpromazine, as inhibiting coronavirus infection in kidney epithelial cells[2]; a similar effort identified 4 compounds (chloroquine, chlorpromazine, loperamide, and lopinavir) as inhibiting coronavirus replication[16]. A recent in silico study examining affinity for the key COVID-19 protease[3] identified ziprasidone as a potential covalently-binding inhibitor. Our analysis does not provide further support for chloroquine, and is not informative regarding the others because of low prescribing frequency.

We emphasize the preliminary nature of these findings, which will require replication and extension in other data sets prior to clinical investigation. Two key limitations must be emphasized. First, as in any nonrandomized study, the risk for bias - and particularly for confounding by indication - is high. That is, other differences between the groups being compared, such as differences in comorbidity, may be proxied by the medication exposure; causal relationships cannot be inferred. The observed associations with oral contraceptives, and with anti-acne medication, likely illustrate this risk even as they suggest assay sensitivity.

Second, the power afforded by this cohort, despite incorporating a network of hospitals, is modest. In particular, absence of effect does not preclude potential efficacy, particularly as our approach will exclude rarer medications, such as those typically prescribed for time-limited periods such as short-term antibiotics.

Further limitations include the reliance on medications electronically prescribed, rather than filled, in the 9 months prior to hospitalization. The lag in availability of current medication data precludes characterization of current medications on admission. This lag, as well as the possibility that medications may be discontinued or not filled, would tend to bias results toward the null hypothesis, by introducing additional heterogeneity. While sensitivity analysis requiring multiple prescriptions for definition of exposure did not meaningfully change results, further studies with more precise exposure characterization will be critical once larger cohorts are available.

Despite these limitations, medication repositioning represents an appealing strategy for responding to COVID-19 because of the rapidity with which promising interventions can be transitioned to clinical trials, and potentially to clinical application[4]. As the safety profile of these medications is well-understood, trials can focus on detection of short-term efficacy, without requiring the longer-term safety data typically required for FDA registration[4,6]. A particularly powerful strategy may integrate clinical informatics and cellular/translational data, as such data become available. At minimum, we hope these results will spur others to pursue similar investigations, yielding the larger, richer data sets required for detection of more modest effects and control of confounding.

Acknowledgements

No funding was received for this study. Dr. Perlis has received consulting fees from Burrage Capital, Genomind, RID Ventures, and Takeda. He holds equity in Outermost Therapeutics and Psy Therapeutics. Mr. Castro, Dr. Ross, and Dr. McBride report no conflict of interest.

Role of the Funding Source

No funding source contributed to any aspect of study design, data collection, data analysis, or data interpretation. The corresponding author (RHP) had full access to all the data in the study. All authors shared the final responsibility for the decision to submit for publication.

Figures and Tables

Table 1. Demographic comparison of hospitalized and not hospitalized among patients who tested positive for SARS-NCOV-2.

	COVID+ Not Hospitalized (N=1564)	COVID+ Hospitalized (N=707)	Total (N=2271)	p value
Male gender	621 (39.7%)	414 (58.6%)	1035 (45.6%)	< 0.001
Age (Mean [SD])	47.058 (18.608)	62.359 (17.957)	51.821 (19.721)	< 0.001
Race				0.009
Asian	56 (3.6%)	28 (4.0%)	84 (3.7%)	
Black	275 (17.6%)	101 (14.3%)	376 (16.6%)	
Other	219 (14.0%)	138 (19.5%)	357 (15.7%)	
Unknown	323 (20.7%)	136 (19.2%)	459 (20.2%)	
White	691 (44.2%)	304 (43.0%)	995 (43.8%)	
Hispanic ethnicity	148 (9.5%)	68 (9.6%)	216 (9.5%)	0.907
Age-adjusted Charlson Comorbidity Index (Mean [SD])	2.201 (3.110)	5.164 (4.838)	3.123 (3.978)	< 0.001
Hospital Type				0.162
Academic medical centers	1103 (70.5%)	478 (67.6%)	1581 (69.6%)	
Community hospitals	461 (29.5%)	229 (32.4%)	690 (30.4%)	

Table 2. Medications most enriched among COVID+ patients (n=2,271) who were not, admitted to the hospital

medication name	hospital - yes	hospital - no	vent - yes	vent - no	rank change	adjusted OR (hospital)	[95%	CI]	adjusted OR (vent)	[95%	CI]
Ibuprofen	34	125	#	29	29	0.649	0.415	0.992	0.421	0.139	1.046
Valacyclovir	#	25	#	#	20	0.477	0.134	1.318	NA	NA	NA
Naproxen	7	26	#	6	18	0.388	0.144	0.930	0.362	0.019	2.269
Oseltamivir	6	24	#	#	17	0.403	0.141	0.990	1.595	0.21	8.954
Sertraline	10	27	#	9	16	0.834	0.351	1.847	0.259	0.014	1.47
Ketorolac tromethamine	35	55	7	28	16	1.697	1.025	2.775	0.753	0.291	1.722
Atropine	16	32	#	13	15	0.635	0.312	1.251	0.658	0.146	2.155
Flumazenil	10	26	#	8	15	0.734	0.311	1.609	0.713	0.104	3.048
Doxycycline	17	31	#	15	13	0.741	0.364	1.456	0.348	0.053	1.313
Amoxicillin/clavulanate	20	34	6	14	13	1.226	0.637	2.302	1.408	0.477	3.733
Naloxone	19	33	#	16	13	1.294	0.671	2.430	0.541	0.123	1.699
Ketoconazole	11	24	#	8	12	0.71	0.302	1.583	0.831	0.176	3.005
Hydrochlorothiazide	29	43	7	22	12	0.716	0.405	1.241	0.776	0.291	1.851
Bupropion	10	23	#	#	12	0.854	0.34	1.993	3.621	0.963	13.631
Atenolol	13	25	#	11	11	0.434	0.195	0.919	0.485	0.072	1.920

OR, odds ratio; #, obfuscated value (5 or less)

Table 3. Demographic comparison of COVID+ hospitalized patients with and without mechanical ventilation

	COVID+ hospitalized with no mechanical ventilation (N=494)	**COVID+ hospitalized with mechanical ventilation** (N=213)	**Total** (N=707)	**p value**
Male gender	274 (55.5%)	140 (65.7%)	414 (58.6%)	0.011
Age (Mean [SD])	62.219 (19.060)	62.685 (15.131)	62.359 (17.957)	0.751
Race				0.005
Asian	22 (4.5%)	6 (2.8%)	28 (4.0%)	
Black	68 (13.8%)	33 (15.5%)	101 (14.3%)	
Other	99 (20.0%)	39 (18.3%)	138 (19.5%)	
Unknown	78 (15.8%)	58 (27.2%)	136 (19.2%)	
White	227 (46.0%)	77 (36.2%)	304 (43.0%)	
Hispanic ethnicity	42 (8.5%)	26 (12.2%)	68 (9.6%)	0.125
Age-adjusted Charlson Comorbidity Index (Mean [SD])	5.405 (4.955)	4.606 (4.517)	5.164 (4.838)	0.044
Hospital type				0.009
Academic medical centers	319 (64.6%)	159 (74.6%)	478 (67.6%)	
Community hospitals	175 (35.4%)	54 (25.4%)	229 (32.4%)	

References

1. Li G, De Clercq E. Therapeutic options for the 2019 novel coronavirus (2019-nCoV). *Nat Rev Drug Discov*. 2020;19(3):149-150. doi:10.1038/d41573-020-00016-0

2. Dyall J, Coleman CM, Hart BJ, et al. Repurposing of clinically developed drugs for treatment of Middle East respiratory syndrome coronavirus infection. *Antimicrob Agents Chemother*. 2014;58(8):4885-4893. doi:10.1128/AAC.03036-14

3. Liu S, Zheng Q, Wang Z. Potential covalent drugs targeting the main protease of the SARS-CoV-2 coronavirus. *Bioinforma Oxf Engl*. April 2020. doi:10.1093/bioinformatics/btaa224

4. Power A, Berger AC, Ginsburg GS. Genomics-Enabled Drug Repositioning and Repurposing: Insights From an IOM Roundtable Activity. *JAMA*. 2014;311(20):2063-2064. doi:10.1001/jama.2014.3002

5. Castro VM, Clements CC, Murphy SN, et al. QT interval and antidepressant use: a cross sectional study of electronic health records. *BMJ*. 2013;346:f288. doi:10.1136/bmj.f288

6. Kupferschmidt K, Cohen J. Race to find COVID-19 treatments accelerates. *Science*. 2020;367(6485):1412-1413. doi:10.1126/science.367.6485.1412

7. Nalichowski R, Keogh D, Chueh HC, Murphy SN. Nalichowski, R., Keogh, D., Chueh, H.C. and Murphy, S.N., 2006. *AMIA Annu Symp Proc*. 2006;2006:1044.

8. Murphy SN, Weber G, Mendis M, et al. Serving the enterprise and beyond with informatics for integrating biology and the bedside (i2b2). *J Am Med Inform Assoc JAMIA*. 2010;17(2):124-130. doi:10.1136/jamia.2009.000893

9. Charlson M, Szatrowski TP, Peterson J, Gold J. Validation of a combined comorbidity index. *J Clin Epidemiol*. 1994;47(11):1245-1251.

10. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A*. 2005;102(43):15545-15550. doi:10.1073/pnas.0506580102

11. Mootha VK, Lindgren CM, Eriksson K-F, et al. PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat Genet*. 2003;34(3):267-273. doi:10.1038/ng1180

12. National Drug File - Reference Terminology - Summary | NCBO BioPortal. https://bioportal.bioontology.org/ontologies/NDFRT/?p=summary. Accessed April 9, 2020.

13. R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing; 2019. www.R-project.org.

14. Wu C, Chen X, Cai Y, et al. Risk Factors Associated With Acute Respiratory Distress Syndrome and Death in Patients With Coronavirus Disease 2019 Pneumonia in Wuhan, China. *JAMA Intern Med*. March 2020. doi:10.1001/jamainternmed.2020.0994

15. Sodhi M, Etminan M. Safety of Ibuprofen in Patients with COVID-19; Causal or Confounded? *Chest*. March 2020:S0012369220305729. doi:10.1016/j.chest.2020.03.040

16. de Wilde AH, Jochmans D, Posthuma CC, et al. Screening of an FDA-Approved Compound Library Identifies Four Small-Molecule Inhibitors of Middle East Respiratory Syndrome Coronavirus Replication in Cell Culture. *Antimicrob Agents Chemother*. 2014;58(8):4875-4884. doi:10.1128/AAC.03011-14