



MOX-Report No. 04/2021

**Development and validation of a clinical risk score to
predict the risk of SARS-CoV-2 infection from
administrative data: A population-based cohort study
from Italy**

Orlando, V.; Rea, F.; Savaré, L.; Guarino, I; Mucherino, S.;
Perrella, A.; Trama, U.; Coscioni, E.; Menditto, E.; Corrao, G.

MOX, Dipartimento di Matematica
Politecnico di Milano, Via Bonardi 9 - 20133 Milano (Italy)

mox-dmat@polimi.it

<http://mox.polimi.it>

RESEARCH ARTICLE

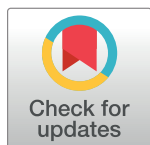
Development and validation of a clinical risk score to predict the risk of SARS-CoV-2 infection from administrative data: A population-based cohort study from Italy

Valentina Orlando^{1,2‡}, Federico Rea^{3,4‡}, Laura Savaré^{5,6*}, Ilaria Guarino¹, Sara Mucherino^{1,2}, Alessandro Perrella⁵, Ugo Trama⁶, Enrico Coscioni⁷, Enrica Menditto^{1,2‡}, Giovanni Corrao^{3,4‡}

1 CIRFF, Center of Drug Utilization and Pharmacoeconomics, University of Naples Federico II, Naples, Italy, **2** Department of Pharmacy, University of Naples Federico II, Naples, Italy, **3** National Centre for Healthcare Research & Pharmacoepidemiology, at the University of Milano-Bicocca, Milan, Italy, **4** Department of Statistics and Quantitative Methods, Laboratory of Healthcare Research & Pharmacoepidemiology, Unit of Biostatistics, Epidemiology and Public Health, University of Milano-Bicocca, Milan, Italy, **5** Infectious Disease of Healthcare Direction, AORN Antonio Cardarelli, Naples, Italy, **6** Regional Pharmaceutical Unit, Campania Region, Naples, Italy, **7** Division of Cardiac Surgery, AOU San Giovanni di Dio e Ruggi d'Aragona, Salerno, Italy

‡ VO and FR contributed equally to this work and served as co-first authors. EM and GC also contributed equally to this work and served as co-lead authors.

* laura.savare@unimib.it



OPEN ACCESS

Citation: Orlando V, Rea F, Savaré L, Guarino I, Mucherino S, Perrella A, et al. (2021) Development and validation of a clinical risk score to predict the risk of SARS-CoV-2 infection from administrative data: A population-based cohort study from Italy. PLoS ONE 16(1): e0237202. <https://doi.org/10.1371/journal.pone.0237202>

Editor: Giordano Madeddu, University of Sassari, ITALY

Received: July 18, 2020

Accepted: December 12, 2020

Published: January 20, 2021

Copyright: © 2021 Orlando et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: Permission use anonymized data to this study was granted to the researchers of the Centro di Ricerca in Farmacoconomia e Farmacoutilizzazione (CIRFF) by the governance board of Unità del Farmaco della Regione Campania. The CIRFF has a regional decree that allow for conducting research by making secondary use of administrative data (DGR n 276 23/05/2017). Furthermore, General Director of Health Care Services authorized the CIRFF to access the COVID-19 patients data in

Abstract

Background

The novel coronavirus (SARS-CoV-2) pandemic spread rapidly worldwide increasing exponentially in Italy. To date, there is lack of studies describing clinical characteristics of the people at high risk of infection. Hence, we aimed (i) to identify clinical predictors of SARS-CoV-2 infection risk, (ii) to develop and validate a score predicting SARS-CoV-2 infection risk, and (iii) to compare it with unspecific scores.

Methods

Retrospective case-control study using administrative health-related database was carried out in Southern Italy (Campania region) among beneficiaries of Regional Health Service aged over than 30 years. For each person with SARS-CoV-2 confirmed infection (case), up to five controls were randomly matched for gender, age and municipality of residence. Odds ratios and 90% confidence intervals for associations between candidate predictors and risk of infection were estimated by means of conditional logistic regression. SARS-CoV-2 Infection Score (SIS) was developed by generating a total aggregate score obtained from assignment of a weight at each selected covariate using coefficients estimated from the model. Finally, the score was categorized by assigning increasing values from 1 to 4. Discriminant power was used to compare SIS performance with that of other comorbidity scores.

Campania Region (Protocol n. 0173642 of 26/03/2020). Requests for information on data access can be directed to Dr. Enrica Menditto, Director of CIRFF (contact via enrica.menditto@unina.it).

Funding: The authors received no specific funding for this work.

Competing interests: Giovanni Corrao received research support from the European Community, the Italian Agency of Drug (AIFA), and the Italian Ministry for University and Research (MIUR). He took part to a variety of projects that were funded by pharmaceutical companies (i.e. Novartis, GSK, Roche, AMGEN and BMS). He also received honoraria as member of Advisory Board from Roche. Other authors declare that they have no conflict of interest to disclose. There are no patents, products in development or marketed products to declare. This does not alter our adherence to PLOS ONE policies on sharing data and materials.

Results

Subjects suffering from diabetes, anaemias, Parkinson's disease, mental disorders, cardiovascular and inflammatory bowel and kidney diseases showed increased risk of SARS-CoV-2 infection. Similar estimates were recorded for men and women and younger and older than 65 years. Fifteen conditions significantly contributed to the SIS. As SIS value increases, risk progressively increases, being odds of SARS-CoV-2 infection among people with the highest SIS value ($SIS = 4$) 1.74 times higher than those unaffected by any SIS contributing conditions ($SIS = 1$).

Conclusion

Conditions and diseases making people more vulnerable to SARS-CoV-2 infection were identified by the current study. Our results support decision-makers in identifying high-risk people and adopting of preventive measures to minimize the spread of further epidemic waves.

Introduction

Since December 2019, the novel coronavirus (SARS-CoV-2) pandemic spread rapidly from the Hubei province in China to 185 countries causing over 3,000,000 cases [1]. The epidemic spread to and increased exponentially in Italy, earlier than in any other western Country, having generated at the current time (June 15) over 236,000 confirmed SARS-CoV-2 infections [2]. SARS-COV-2 causes a Coronavirus disease 2019 (Covid-19), for which minor symptoms are anosmia, ageusia, gastrointestinal symptoms, headache, and cutaneous manifestations and major symptoms are fever, cough, dyspnoea [3, 4]. Due to these major symptoms it may be considered necessary to hospitalize patients for respiratory complications [5].

Several hospital-based studies [6–10], including a systematic review of literature and meta-analysis [11], focused on the attempt for predicting the progression of the disease towards developing critical manifestations or death. These studies are important for the clinical practice point of view for identifying patients at whom early treatment must be guaranteed. However, as most infections are not life-threatening [7], for the public health point of view it becomes increasingly important stratifying population for identifying people at higher risk of infection. Despite this, at our best knowledge, no studies on this topic have been still published.

We therefore performed a large investigation based upon healthcare utilization database from the Italian Region of Campania aimed (1) to identify clinical predictors of the risk of SARS-CoV-2 infection, (2) to develop and validate a score overall predicting the risk of SARS-CoV-2 infection, and (3) to compare discriminant power of such a score with that from unspecific scores of clinical profile.

Methods

Target population and data source

Residents in Campania who were beneficiaries of the Regional Health Service (RHS) aged 30 years or older formed the target population (almost 3.9 million people, around 9% of the Italian population of that age group). Italian citizens have equal access to essential healthcare services provided by the National Health Service [12]. An automated system of healthcare

utilization (HCU) databases allows managing NHS within each Italian region, including Campania. HCU data report a variety of information drawn from services provided fully or in part free of charge from NHS to beneficiaries of NHS (e.g. the ICD-CM-9 codes of inpatient diagnoses and services supplied from public or private hospitals, the ATC codes of outpatient drugs dispensed from pharmacies). This allowed to Campania Region of designing, building and routinely managing the so-called Campania Region Database (CaReDB) which formed the data source for the current study. Completeness and data validity of CaReDB being elsewhere reported [13–21]. Selected characteristics of CaReDB are described in [S1 Table](#).

From the beginning of the Covid-19 epidemic, a surveillance system was implemented to detect all cases identified by reverse transcription-polymerase chain reaction (RT-PCR) testing for SARS-CoV-2. Diagnostic algorithm was based on the protocol released by the World Health Organization (WHO) [22], i.e., on nasopharyngeal swab specimens tested with at least two real-time RT PCT assays targeting different genes (E, RdRp and M) of SARS-CoV-2.

These various types of data (i.e., CaReDB and Covid-19 registry) can be interconnected, since a single individual identification code is used by all databases for each citizen enrolled. To preserve privacy, each identification code was automatically deidentified, the inverse process being allowed only to the Regional Health Authority on request from judicial authorities. Permission for using anonymized data for this study was granted to the researchers of the Centro di Ricerca in Farmacoeconomia e Farmacoutilizzazione (CIRFF) by the governance board of Unità del Farmaco della Regione Campania. According with Italian Data protection Authority, neither Ethical Committee approval, nor informed consent were required for carrying out observational studies based on HCU data as the our [23]. Our research protocol adhered to the tenets of the Declaration of Helsinki 1975 and its later amendments.

Cases and controls

The date of SARS-CoV-2 infection diagnosis was considered as the index date and patients were extracted from the Covid-19 registry until June 10, 2020. A total of 4,629 subjects positive to SARS-CoV-2 were identified. Among these, we excluded i) patients with missing demographic information ($N = 469$) and ii) patients younger than 30 years at the index date ($N = 663$). Finally, 3,497 patients were included into the study as cases. Among them, 453 patients died during the observational period.

For each case, up to five controls were randomly selected from the target population to be matched for gender, age at index date and municipality of residence. The density incidence approach was used for selecting controls since patients who had a confirmed diagnosis of SARS-CoV-2 infection were eligible as potential controls until they became cases, and all matches had to be at risk of SARS-CoV-2 infection.

Identifying clinical predictors of SARS-CoV-2 infection

A list of 47 diseases and conditions potentially predicting the risk of SARS-CoV-2 infection was developed starting from the lists included in several comorbidities scores, i.e., the Charlson [24], Elixhauser [25], Chronic Disease Scores [26] and RxRiskV Index [27, 28], and in some systematic reviews on Covid-19 risk factors [11, 29–31]. The algorithms for detecting patients who suffer from each of them from the above-mentioned databases were carefully chosen according to previously published papers on case-identification algorithms based on Italian healthcare data [32–36]. Individuals were classified as having one of the conditions listed if they received at least two consecutive dispensations of a drug for treatment of a specific class of disease and/or one hospital discharge with the diagnoses coded with the specific ICD-9-CM ([S2 Table](#)).

Conditional logistic regression was used to estimate odds ratios (ORs), with 90% confidence intervals (CIs), for the association between candidate predictors and the odds of SARS-CoV-2 infection. Predictors entered as dichotomous covariates into the model, i.e., with value 0 or 1 according to whether the specific condition was not or was recorded at least once within two-years prior baseline (2018–2019). Unadjusted and mutually adjusted models were fitted by including one by one covariate, and all covariates together, respectively. Power considerations suggested of excluding covariates with prevalence $\leq 0.12\%$ among controls, i.e., predictors for which our sample size was not enough for detecting OR of at least 3, with a 0.80 power, and by accepting a 0.10 two-sided first type error. In addition, some conditions were grouped together when strong uncertainty of algorithm did not allow for distinguishing them.

With the aim of testing the hypothesis that predictors may affect severity of clinical manifestations of SARS-CoV-2 infection, rather than infection *per se*, analyses were restricted to strata having fatal infection. Stratifications for sex and age categories (<65 years, ≥ 65 years) were performed as secondary analyses.

Developing and validating a score to predict SARS-CoV-2 infection

Seven out of ten of the 3,497 1:5 case-control sets were randomly selected to form the so-called training (derivation) set. The conditional logistic regression model was fitted to compute the ORs as above described. The least absolute shrinkage and selection operator (LASSO) method was applied for selecting the diseases / conditions able to independently predict the SARS-CoV-2 infection [37]. The coefficients estimated from the model were used for assigning a weight at each selected covariate. A weight was assigned to each coefficient by multiplying it by 10 and rounding it to the nearest whole number [38]. The weights thus obtained were then summed to generate a total aggregate score. To simplify the system, i.e., with the aim of accounting for excessive heterogeneity of the total aggregate score, the latter was categorized by assigning increasing values of 1, 2, 3 and 4 to the categories of the aggregate score of 0, 1–2, 3–4, ≥ 5 , respectively. The so obtained index was denoted SARS-CoV-2 Infection Score (SIS).

Performance of SIS was explored by applying the corresponding weights to the so-called validation set consisting of the 1,048 1:5 case-control sets who did not enter the training set. To evaluate the clinical utility of SIS for predicting infection, we considered the receiver operating characteristic (ROC) curve analysis and used area under the ROC curve (AUC) as a global summary of the discriminatory capacity of the scores [39].

Comparing specific and unspecific predictors of SARS-CoV-2 infection

Some unspecific scores surrogating general clinical profile of each case and control included into the study were considered. In particular, the number of drugs with different 3rd level ATC dispensed to, and comorbidities with different ICD-9-CM experienced by each case and control within two-years prior baseline (2018–2019) were recorded. Categorization was made by assigning increasing values of 1, 2, 3 and 4 to 0, 1–4, 5–9 and ≥ 10 drugs (comedication score) and 1, 2, 3 and 4 to 0, 1–2 and ≥ 3 comorbidities (comorbidity score). In addition, cases and controls were categorized according to the Multisource Comorbidity Score (MCS), a new index of patients' clinical status derived from inpatients diagnostic information and outpatient drug prescriptions provided by the regional Italian data and validated for outcome prediction [40, 41]. To simplify comparisons, the original five categories of worsening clinical profile (0, 1, 2, 3 and 4) as defined by MCS, were reduced to milder (MCS = 0), middle ($1 \leq \text{MCS} \leq 3$) and severe (MCS ≥ 4) categories.

With the aim of comparing discriminatory ability of specific (SIS) and unspecific (comedications, comorbidities and MCS) predictors of SARS-CoV-2 infection, ROC curves and corresponding AUCs were again used.

All analyses were performed using SAS 9.4 (Cary, NC). A 2-sided p-value of 0.10 or less was considered significant.

Results

Clinical predictors of SARS-CoV-2 infection

Owing to their low prevalence, fourteen conditions were excluded from this analysis (tuberculosis, weight loss, disorders involving the immune mechanisms, disorders of fluid, electrolyte and acid-base balance, coagulation defects, bipolar disorders, alcohol abuse, drug addiction, multiple sclerosis, cystic fibrosis, chronic and acute pancreatitis, ankylosing spondylitis, systemic sclerosis, systemic sclerosis). Among the 33 remaining conditions, two were grouped, i.e., chronic pulmonary obstructive disease with asthma (chronic respiratory disease), and chronic renal disease with or without dialysis.

The characteristics of the cohort members are shown in [Table 1](#). Among the 31 remaining conditions, 23 (74%) showed significant association with the risk of SARS-CoV-2 infection from univariate regression.

[Table 2](#) reports multivariate association between the considered diseases/conditions and the risk of SARS-CoV-2 infection which results significant for 12 conditions (39%).

In particular, patients suffering from diabetes, anaemias, mental disorders (dementia / Alzheimer's disease, psychosis and anxiety), Parkinson's disease, glaucoma, diseases of the circulatory system (heart failure and hypertension), chronic respiratory, inflammatory bowel, and rheumatologic conditions showed statistical evidence of increased risk of infection with respect to patients who did not suffer from them. Likely because of low power, only 7 conditions resulted significantly associated with the risk of fatal Covid-19 disease, but there was no relevant difference in the estimates with respect to the risk of SARS-CoV-2 infection ([Table 2](#)).

Anaemias, dementia/Alzheimer, psychosis, anxiety, epilepsy, heart failure, kidney diseases and particularly cystic fibrosis increased the risk of SARS-CoV-2 infection among women, whereas higher risk of infection was observed among men suffering from diabetes, psychosis, anxiety, Parkinson, arrhythmia, chronic pulmonary disease, inflammatory bowel diseases and particularly dementia/Alzheimer and rheumatologic conditions ([S3 Table](#)).

Estimates were similar for Covid-19 patients younger and older than 65 years. Among the former group, a significant higher risk of infection was observed for diabetes, anxiety, Parkinson's disease, arrhythmia, inflammatory bowel and chronic pulmonary diseases, particularly dementia/Alzheimer, whereas patients older than 65 years suffering from thyroid disorders, anaemias, dementia/Alzheimer, psychosis, anxiety, epilepsy and heart failure showed a significant higher risk infection ([S4 Table](#)).

SARS-CoV-2 Infection Score (SIS)

Fifteen conditions significantly contributed to the SIS, the corresponding weights being reported in [Table 3](#). Factors which most contributed to the total aggregate score were dementia / Alzheimer's disease, kidney disease, psychosis, inflammatory bowel disease and rheumatologic conditions, while diabetes, anaemias, anxiety, Parkinson's disease, glaucoma, heart failure, hypertension, arrhythmia, thyroid disorders and chronic respiratory disease provided small, although significant, contributions. [Fig 1](#) shows that, as the SIS value increases, the OR progressively increases, being the odds of SARS-CoV-2 infection among people with the highest SIS value (SIS = IV), 1.74 times higher than those unaffected by any SIS contributing conditions (SIS = I). The prevalence of controls stratified according to the SIS score gradually decreases from 50% (SIS = I) to 12% (SIS = IV).

Table 1. Baseline characteristics of cohort members (Covid-19 cases and related controls), individual (one by one, univariate) Odds Ratio (OR), and 90% Confidence Intervals (CI), for the relationship between selected diseases/conditions and the risk of SARS-CoV-2 infection.

	Cases (N = 3,497)	Controls (N = 17,358)	Individual OR (90% CI)
Male gender	1,945 (55.6%)	9,640 (55.5%)	MV
Age (years)			
30–64	2,375 (67.9%)	11,829 (68.1%)	MV
≥65	1,122 (32.1%)	5,538 (31.9%)	
Infectious and parasitic diseases			
HIV infection	68 (1.9%)	301 (1.7%)	1.12 (0.90 to 1.41)
Neoplasms			
Malignant neoplasms	155 (4.4%)	661 (3.8%)	1.18 (1.01 to 1.37)
Endocrine, nutritional and metabolic diseases, and immunity disorders			
Thyroid disorders	225 (6.4%)	920 (5.3%)	1.25 (1.10 to 1.42)
Diabetes	411 (11.8%)	1732 (10%)	1.22 (1.10 to 1.35)
Hyperlipidaemia	729 (20.8%)	3708 (21.4%)	0.97 (0.89 to 1.05)
Obesity	48 (1.4%)	153 (0.9%)	1.58 (1.20 to 2.08)
Hyperuricemia/Gout	180 (5.1%)	711 (4.1%)	1.28 (1.11 to 1.48)
Diseases of the blood and blood-forming organs			
Anaemias	265 (7.6%)	927 (5.3%)	1.48 (1.31 to 1.67)
Mental disorders			
Dementia / Alzheimer	48 (1.4%)	89 (0.5%)	2.79 (2.06 to 3.79)
Psychosis	124 (3.5%)	303 (1.7%)	2.10 (1.75 to 2.52)
Depression	233 (6.7%)	1,003 (5.8%)	1.17 (1.03 to 1.33)
Anxiety	1,369 (39.1%)	5,615 (32.3%)	1.37 (1.29 to 1.47)
Diseases of the nervous system and sense organs			
Parkinson's disease	67 (1.9%)	188 (1.1%)	1.78 (1.40 to 2.26)
Epilepsy	176 (5%)	660 (3.8%)	1.35 (1.17 to 1.55)
Glaucoma	119 (3.4%)	482 (2.8%)	1.25 (1.05 to 1.48)
Diseases of the circulatory system			
Ischaemic Heart Disease/Angina	213 (6.1%)	841 (4.8%)	1.29 (1.12 to 1.47)
Heart failure	281 (8%)	1,005 (5.8%)	1.49 (1.31 to 1.69)
Arrhythmia	196 (5.6%)	738 (4.3%)	1.36 (1.18 to 1.57)
Valvular diseases	43 (1.2%)	180 (1%)	1.18 (0.88 to 1.57)
Vascular diseases	52 (1.5%)	186 (1.1%)	1.41 (1.08 to 1.84)
Cerebrovascular diseases	127 (3.6%)	445 (2.6%)	1.46 (1.22 to 1.74)
Hypertension	826 (23.6%)	3,731 (21.5%)	1.15 (1.07 to 1.25)
Diseases of the respiratory system			
Chronic respiratory diseases (COPD and asthma together)	244 (7%)	908 (5.2%)	1.37 (1.21 to 1.56)
Diseases of the digestive system			
Liver cirrhosis and other liver chronic diseases	54 (1.5%)	216 (1.2%)	1.23 (0.96 to 1.59)
Inflammatory bowel diseases	54 (1.5%)	169 (1%)	1.60 (1.23 to 2.07)
Diseases of the genitourinary system			
Kidney disease with or without dialysis	67 (1.9%)	210 (1.2%)	1.60 (1.26 to 2.03)
Diseases of the skin and subcutaneous tissues			
Psoriasis	23 (0.7%)	113 (0.7%)	1.02 (0.70 to 1.48)
Diseases of the musculoskeletal system and connective tissue			
Rheumatologic conditions	28 (0.8%)	79 (0.5%)	1.77 (1.23 to 2.56)
Other conditions			
Transplantation	13 (0.4%)	59 (0.3%)	1.10 (0.66 to 1.82)

(Continued)

Table 1. (Continued)

	Cases (N = 3,497)	Controls (N = 17,358)	Individual OR (90% CI)
Chronic pain	89 (2.5%)	378 (2.2%)	1.19 (0.97 to 1.45)
Inflammation, not elsewhere specified	410 (11.7%)	2,244 (12.9%)	0.89 (0.81 to 0.98)

Abbreviation: MV, matching variable.

<https://doi.org/10.1371/journal.pone.0237202.t001>

Comparing with unspecific predictors of SARS-CoV-2 infection

Generic/unspecific scores surrogating clinical profile showed to be associated with the risk of SARS-CoV-2 infection, showing patients with ≥ 10 drug treatments, those with ≥ 3 comorbidities, and those with MCS value ≥ 4 , increased risk of 65%, 36% and 45% with respect to patients cotreatments, comorbidities and MCS value = 1, respectively (Table 4).

AUC (90% CI) of SIS, cotreatment and comorbidity scores and MCS respectively had values of 0.54 (0.52 to 0.56), 0.52 (0.50 to 0.54), 0.53 (0.51 to 0.55), and 0.53 (0.51 to 0.55) (Fig 2). There was no evidence that specific and unspecific scores had different discriminatory ability.

Discussion

Our study shows that several diseases and conditions are significantly and independently associated with the risk of SARS-CoV-2 infection. Beyond conditions making particularly vulnerable the respiratory system (e.g., chronic obstructive pulmonary disease and asthma), comorbidities positively associated with the infection practically included all diagnostic categories. Predictors belonging to nutritional and metabolic (diabetes), cardiovascular (heart failure and hypertension) and renal diseases were widely expected, since it has accepted that SARS-CoV-2 has major implications for the cardiovascular system. Indeed, patients with heart failure [42], diabetes [43–45], hypertension [12] and kidney disease [46–48] have been consistently identified as particularly vulnerable populations, and these findings were consistently found in our study. In addition, we confirmed that people with weakened immune systems from a medical condition or treatment are at a higher risk. Among these, those living with haemoglobin disorders [49], inflammatory bowel disease [50] and immune-rheumatological diseases [51] must be considered vulnerable groups for SARS-CoV-2 infection. Mental health and cognitive function might have independent utility in understanding the burden of respiratory disease, since they may influence the risk of contracting the infection, at least in part by impairing innate or adaptive immunity [52] and diminishing the precautions taken to minimize risk. Another explanation of our findings is that people with history of depression [53], psychosis [54] and stress disorders [55] could experience elevated rates of an array of respiratory infections because these conditions often require treatment in a psychiatric care facility, and the risk of infection can be particularly high in these structures. Finally, our study adds evidence regarding the impact of diseases and conditions on the risk of SARS-CoV-2 infection between men and women. As pointed out by a recent study [56], sex and age disaggregated data are essential for understanding the distributions of risk infection in the population and the extent to which they affect clinical outcomes.

Despite our results confirm that a wide range of diseases and conditions likely increase vulnerability to SARS-CoV-2 infection, and probably its more severe clinical manifestations, we have not been able to develop a score that accurately may predict the risk of infection. In addition, we found that predictive ability of the score obtained by weighting risk factors of SARS-CoV-2 infection did not overcome that of some generic scores of comorbidities and comedications. This expands upon previous findings of individual comorbidities as

Table 2. Independent (all together, multivariate) Odds Ratio (OR), and 90% Confidence Intervals (CI), for the relationship between selected diseases/conditions and the risk of SARS-CoV-2 infection as a whole (3,497 cases and corresponding 17,358 controls), as well as the risk of fatal SARS-CoV-2 infection (435 cases and corresponding 2,154 controls).

	All Covid-19 cases		Fatal Covid-19 cases	
	Cases / Controls	OR (90% CI)	Cases / Controls	OR (90% CI)
Infectious and parasitic diseases				
HIV infection	68 / 301	1.07 (0.85 to 1.34)	11 / 47	1.04 (0.58 to 1.86)
Neoplasms				
Malignant neoplasms	155 / 661	0.99 (0.85 to 1.16)	35 / 147	0.99 (0.70 to 1.42)
Endocrine, nutritional and metabolic diseases, and immunity disorders				
Thyroid disorders	225 / 920	1.13 (0.99 to 1.29)	31 / 133	0.93 (0.64 to 1.37)
Diabetes	411 / 1,732	1.15 (1.03 to 1.28)	88 / 327	1.30 (1.01 to 1.67)
Hyperlipidaemia	729 / 3,708	0.86 (0.79 to 0.94)	131 / 703	0.69 (0.55 to 0.86)
Obesity	48 / 153	1.18 (0.89 to 1.57)	6 / 18	1.08 (0.46 to 2.56)
Hyperuricemia/Gout	180 / 711	1.08 (0.93 to 1.27)	56 / 175	1.29 (0.95 to 1.76)
Diseases of the blood and blood-forming organs				
Anaemias	265 / 927	1.24 (1.09 to 1.41)	63 / 184	1.45 (1.07 to 1.95)
Mental disorders				
Dementia / Alzheimer	48 / 89	2.14 (1.55 to 2.96)	14 / 27	1.92 (1.02 to 3.63)
Psychosis	124 / 303	1.71 (1.40 to 2.08)	35 / 57	2.68 (1.76 to 4.08)
Depression	233 / 1,003	0.98 (0.86 to 1.12)	49 / 149	1.21 (0.88 to 1.67)
Anxiety	1,369 / 5,615	1.26 (1.17 to 1.36)	217 / 7,824	1.33 (1.07 to 1.65)
Diseases of the nervous system and sense organs				
Parkinson's disease	67 / 188	1.32 (1.02 to 1.70)	18 / 47	1.32 (0.80 to 2.18)
Epilepsy	176 / 660	1.10 (0.94 to 1.28)	44 / 99	1.57 (1.11 to 2.22)
Glaucoma	119 / 482	1.22 (1.03 to 1.46)	27 / 95	1.32 (0.89 to 1.97)
Diseases of the circulatory system				
Ischaemic Heart Disease/Angina	213 / 841	0.99 (0.84 to 1.15)	58 / 186	1.24 (0.88 to 1.76)
Heart failure	281 / 1,005	1.24 (1.07 to 1.44)	86 / 268	1.41 (1.04 to 1.90)
Arrhythmia	196 / 738	1.14 (0.98 to 1.33)	51 / 198	0.95 (0.68 to 1.31)
Valvular diseases	43 / 180	0.80 (0.59 to 1.09)	11 / 42	0.84 (0.44 to 1.60)
Vascular diseases	52 / 186	1.00 (0.75 to 1.32)	9 / 40	0.69 (0.36 to 1.33)
Cerebrovascular diseases	127 / 445	1.00 (0.83 to 1.21)	37 / 108	1.06 (0.72 to 1.56)
Hypertension	826 / 3,731	1.12 (1.01 to 1.24)	167 / 712	1.16 (0.94 to 1.43)
Diseases of the respiratory system				
Chronic respiratory diseases (COPD and asthma together)	244 / 908	1.18 (1.03 to 1.35)	50 / 178	1.14 (0.83 to 1.55)
Diseases of the digestive system				
Liver cirrhosis and other liver chronic diseases	54 / 216	0.93 (0.71 to 1.21)	14 / 35	1.42 (0.79 to 2.56)
Inflammatory bowel diseases	54 / 169	1.47 (1.13 to 1.91)	6 / 32	0.72 (0.33 to 1.56)
Diseases of the genitourinary system				
Kidney disease with or without dialysis	67 / 210	1.10 (0.84 to 1.42)	23 / 60	1.07 (0.60 to 1.90)
Diseases of the skin and subcutaneous tissues				
Psoriasis	23 / 113	0.93 (0.63 to 1.36)	2 / 19	0.39 (0.11 to 1.44)
Diseases of the musculoskeletal system and connective tissue				
Rheumatologic conditions	28 / 79	1.54 (1.06 to 2.23)	5 / 18	1.21 (0.50 to 2.91)
Other conditions				
Transplantation	13 / 59	0.87 (0.52 to 1.46)	3 / 8	1.45 (0.42 to 4.97)
Chronic pain	89 / 378	1.06 (0.86 to 1.31)	21 / 78	1.08 (0.69 to 1.70)
Inflammation, not elsewhere specified	410 / 2,244	0.85 (0.77 to 0.93)	74 / 316	1.12 (0.87 to 1.45)

<https://doi.org/10.1371/journal.pone.0237202.t002>

Table 3. Weights, assigned to diseases that were significantly associated with the risk of Covid-19 disease, used to construct the SARS-CoV-2 Infection Score (SIS).

Disease / Condition	Log (OR)	Weights
Thyroid disorders	0.08	1
Diabetes	0.08	1
Anaemias	0.23	2
Dementia / Alzheimer	0.98	10
Psychosis	0.46	5
Anxiety	0.23	2
Parkinson's disease	0.27	3
Glaucoma	0.15	2
Heart failure	0.27	3
Arrhythmia	0.12	1
Hypertension	0.12	1
Chronic Pulmonary disease	0.15	2
Inflammatory bowel diseases	0.40	4
Kidney dialysis	0.75	8
Rheumatologic conditions	0.55	6

<https://doi.org/10.1371/journal.pone.0237202.t003>

independent risk factors for SARS-CoV-2 infection [57, 58], and confirms our substantial inability to predict the risk of SARS-CoV-2 infection. This can be explained by several limitations of our approach, which generate estimates biased towards the null. First, exposure misclassification regards our inability to carefully capturing conditions and diseases through algorithms based on healthcare utilization databases [59]. Second, it is well known that outcome misclassification can bias epidemiologic results. For Covid-19, suboptimal test

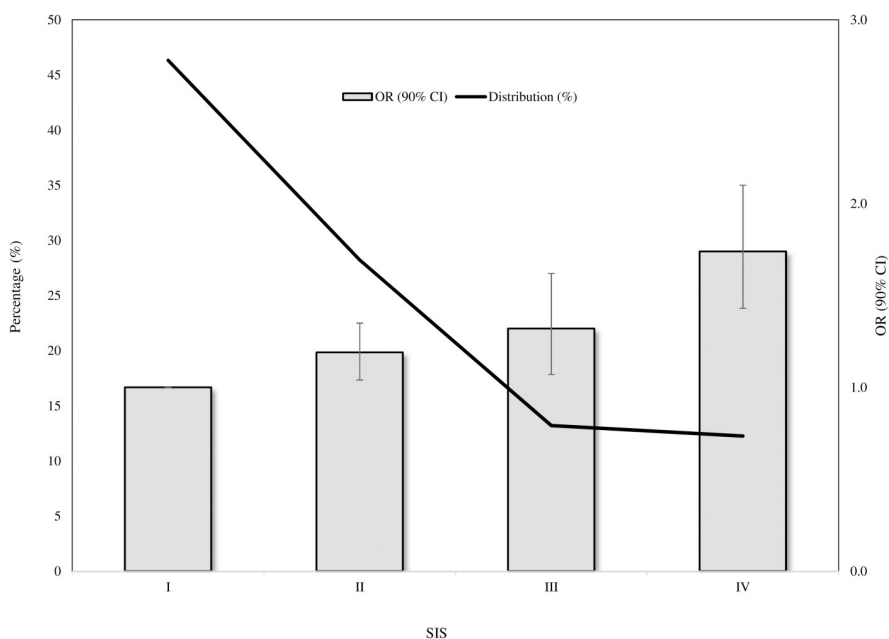


Fig 1. SARS-CoV-2 Infection Score (SIS) distribution among controls, and corresponding trend in odds ratios (and 90% confidence intervals) along categories of SIS. SARS-CoV-2 Infection Score: I, II, III and IV to 0, 1–2, 3–4 and ≥ 5 .

<https://doi.org/10.1371/journal.pone.0237202.g001>

Table 4. Relationship between selected score and the risk of SARS-CoV-2 infection.

Scores	OR (90% CI)
SARS-CoV-2 Infection Score (SIS)	
I (0)	1.00 (Ref.)
II (1–2)	1.19 (1.03 to 1.36)
III (3–4)	1.32 (1.10 to 1.58)
IV (≥ 5)	1.74 (1.44 to 2.10)
Number of comedications	
I (0)	1.00 (Ref.)
II (1–4)	1.05 (0.91 to 1.21)
III (5–9)	1.17 (0.97 to 1.41)
IV (≥ 10)	1.65 (1.25 to 2.19)
Number of comorbidities	
I (0)	1.00 (Ref.)
II (1–2)	1.21 (1.05 to 1.38)
III (≥ 3)	1.36 (1.15 to 1.60)
Multisource Comorbidity Score (MCS)	
I (0)	1.00 (Ref.)
II (1–3)	1.21 (1.03 to 1.41)
III (≥ 4)	1.45 (1.23 to 1.70)

<https://doi.org/10.1371/journal.pone.0237202.t004>

sensitivity, despite excellent specificity, results in an overestimation of cases in the early stages of an outbreak, and substantial underestimation of cases as prevalence increases [60]. It should be noticed, however, that both, exposure and outcome misclassification likely drew estimates

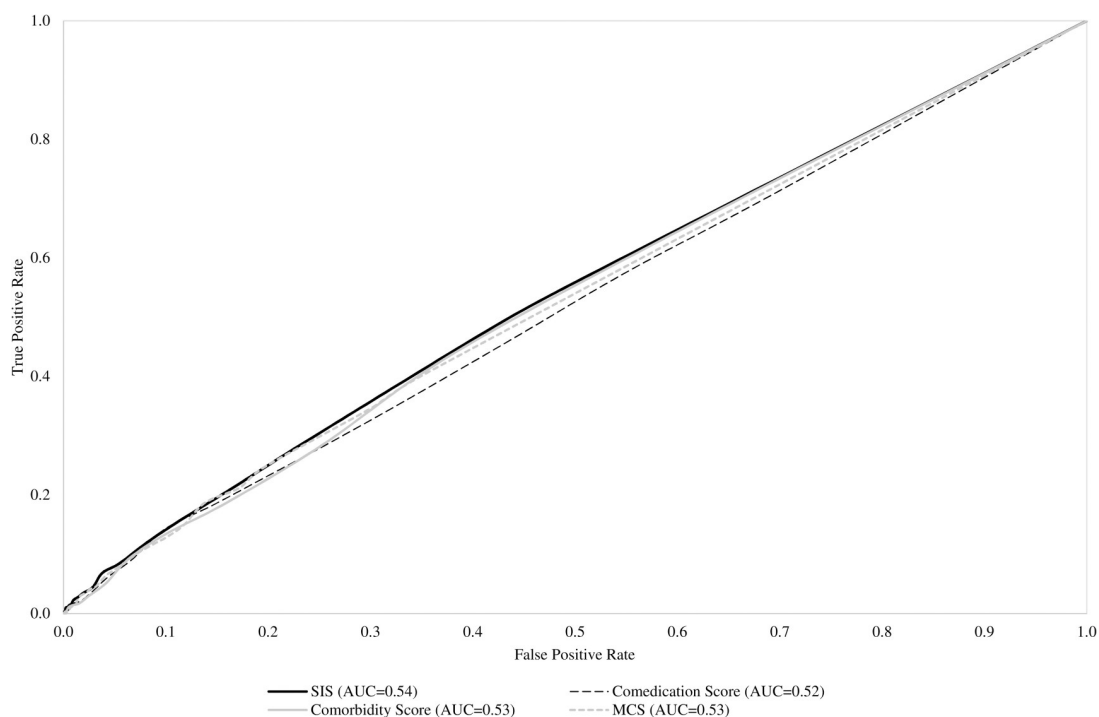


Fig 2. Receiver Operating Characteristics (ROC) curves comparing discriminant power of SARS-CoV-2 Infection Score (SIS), and selected unspecific score surrogating clinical profile (cotreatments, comorbidities and Multisource Comorbidity Score).

<https://doi.org/10.1371/journal.pone.0237202.g002>

towards the null (i.e., underestimate the strength of the association between their presence and the outcome risk) so generating uncertainty for the weighting approach of score developing. Third, the lack of information on biologic markers potentially able to predict infection, and severity of its clinical manifestations, is another limitation of our study. For example, according to the current literature, some laboratory hallmarks have been shown to predict infection, particularly in more severe cases [61]. Finally, our choice of accepting a 0.10 first type error, and of consequently reporting 90% confidence intervals, is justified by the exploratory nature of our study, but at the same time likely generate false positive signals, so limiting discriminant power of the score.

Three other elements of weakness should be acknowledged. First, the lack of data regarding the clinical outcome experienced by SARS-CoV-2 positive patients in terms of home isolation, hospitalization and admission in intensive care. Second, because few people aged less than 30 years were diagnosed to be affected by SARS-CoV-2 infection during the investigated period, and few of them suffered from chronic conditions such as those considered in our study, patients with less than 30 years were excluded from the analysis. Although this reduced the uncertainty of the results, the generalisability of our findings requires extreme caution. Finally, because data on stays in long-term facilities are not recorded in our database, we cannot exclude that the higher risks associated with mental disorders observed in our study could be explained by confounding, i.e., patients who suffered from these conditions are often hospitalized in these structures where the risk of infection can be particularly high.

In conclusion, taking the limitations we discussed into account, we identified conditions and diseases that make people more vulnerable to SARS-CoV-2 infection. These findings contribute to inform public health, and clinical decisions regarding risk stratifying. However, further research is need for developing a score reliably predicting the risk, possibly by integrating healthcare utilization with clinical and biological data.

Our results can be an important tool supporting all clinical and political stakeholders allowing the identification of the population most at risk of contracting SARS-CoV-2 infection and facilitating the provision of appropriate preventive/therapeutic measures, especially with the hypothetical prediction of a new autumn outbreak. Adopting preventive measures can help to minimize the damage generated by a potential new relapse that the health systems will face.

Supporting information

S1 Table. Campania Region Database (CaReDB) characteristics. ATC = Anatomical Therapeutic Chemical; ICD-9-CM = *International Classification of Diseases*, 9th Revision, Clinical Modification. ^aTime span covered was 2009–2018 for hospital-discharge records and 2014–2019 for outpatient pharmacy records.

(DOCX)

S2 Table. List of diseases and conditions candidate for predicting SARS-CoV-2 infection, and corresponding ICD-CM and ATC codes used for detecting they.

(DOCX)

S3 Table. Odds Ratio (OR), and 90% Confidence Intervals (CI), for the relationship between selected diseases/conditions and the risk of SARS-CoV-2 infection, stratified according to gender.

(DOCX)

S4 Table. Odds Ratio (OR), and 90% Confidence Intervals (CI), for the relationship between selected diseases/conditions and the risk of SARS-CoV-2 infection, stratified

according to age categories (i.e., younger and older 65 years).
(DOCX)

Author Contributions

Conceptualization: Enrica Menditto, Giovanni Corrao.

Data curation: Ilaria Guarino, Sara Mucherino.

Formal analysis: Federico Rea, Laura Savaré, Ilaria Guarino, Sara Mucherino.

Methodology: Federico Rea, Laura Savaré, Giovanni Corrao.

Supervision: Valentina Orlando, Enrica Menditto, Giovanni Corrao.

Visualization: Valentina Orlando, Alessandro Perrella, Ugo Trama, Enrico Coscioni, Enrica Menditto, Giovanni Corrao.

Writing – original draft: Giovanni Corrao.

Writing – review & editing: Valentina Orlando, Federico Rea, Laura Savaré, Ilaria Guarino, Sara Mucherino, Alessandro Perrella, Ugo Trama, Enrico Coscioni, Enrica Menditto.

References

1. Dong Ensheng, Du Hongru, Gardner Lauren. An interactive web-based dashboard to track COVID-19 in real time. *The Lancet Infectious Diseases* 2020; 20:533–4. [https://doi.org/10.1016/S1473-3099\(20\)30120-1](https://doi.org/10.1016/S1473-3099(20)30120-1) PMID: 32087114
2. Onder G, Rezza G, Brusaferro S. Case-Fatality Rate and Characteristics of Patients Dying in Relation to COVID-19 in Italy [published online ahead of print, 2020 Mar 23]. *JAMA*. 2020; <https://doi.org/10.1001/jama.2020.4683> PMID: 32203977
3. Vaira LA, Hopkins C, Salzano G, Petrocelli M, Melis A, Cucurullo M, et al. Olfactory and gustatory function impairment in COVID-19 patients: Italian objective multicenter-study. *Head Neck*. 2020 Jul; 42(7):1560–1569. Epub 2020 May 21. <https://doi.org/10.1002/hed.26269> PMID: 32437022
4. Vaira LA, Deiana G, Fois AG, Pirina P, Madeddu G, De Vito A, et al. Objective evaluation of anosmia and ageusia in COVID-19 patients: Single-center experience on 72 cases. *Head Neck*. 2020 Jun; 42(6):1252–1258. Epub 2020 Apr 29. <https://doi.org/10.1002/hed.26204> PMID: 32342566
5. De Vito A, Geremia N, Fiore V, Princic E, Babudieri S, Madeddu G. Clinical features, laboratory findings and predictors of death in hospitalized patients with COVID-19 in Sardinia, Italy. *Eur Rev Med Pharmacol Sci* 2020; 24 (14): 7861–7868. https://doi.org/10.26355/eurrev_202007_22291 PMID: 32744714
6. Liang W, Liang H, Ou L, Chen B, Chen A, Li C, et al. Development and Validation of a Clinical Risk Score to Predict the Occurrence of Critical Illness in Hospitalized Patients With COVID-19. *JAMA Intern Med*. 2020;e202033. <https://doi.org/10.1001/jamainternmed.2020.2033> PMID: 32396163
7. Xie J, Hungerford D, Hui Chen H, Abrams ST, Li S, Wang G, et al. Development and external validation of a prognostic multivariable model on admission for hospitalized patients with COVID-19. *medRxiv preprint* <https://doi.org/10.1101/2020.03.28.20045997>
8. Li Q, Zhang J, Ling Y, Li W, Zhang X, Luet H, et al. A simple algorithm helps early identification of SARS-CoV-2 infection patients with severe progression. *Infection*. 2020; 1–8. <https://doi.org/10.1007/s15010-019-01387-2> PMID: 31919762
9. Sun Y, Koh V, Marimuthu K, Tek Ng O, Young B, Vasoo S, et al. Epidemiological and Clinical Predictors of COVID-19. *Clin Infect Dis*. 2020;ciaa322. <https://doi.org/10.1093/cid/ciaa322> PMID: 32211755
10. Haimovich A, Ravindra NG, Stoytchev S, Young HP, Wilson FP, van Dijk D, et al. Development and validation of the COVID-19 severity index (CSI): a prognostic tool for early respiratory decompensation. *medRxiv preprint*. <https://doi.org/10.1101/2020.05.07.20094573>
11. Zheng Z, Peng F, Xu B, Zhao J, Liu H, Peng J, et al. Risk factors of critical & mortal COVID-19 cases: A systematic literature review and meta-analysis. *J Infect*. 2020; S0163-4453(20)30234-6. <https://doi.org/10.1016/j.jinf.2020.04.021>
12. Mancina G, Rea F, Ludergrani M, Apolone G, Corrao G. Renin-Angiotensin-Aldosterone System Blockers and the Risk of Covid-19. *N Engl J Med*. 2020; 382(25):2431–2440. <https://doi.org/10.1056/NEJMoa2006923> PMID: 32356627

13. Moreno Juste A, Menditto E, Orlando V, Monetti VM, Gimeno Miguel A, González Rubio F, et al. Treatment Patterns of Diabetes in Italy: A Population-Based Study. *Front Pharmacol*. 2019; 10:870. Published 2019 Aug 6. <https://doi.org/10.3389/fphar.2019.00870> PMID: 31447672
14. Guerriero F, Orlando V, Monetti VM, Russo V, Menditto E. Biological therapy utilization, switching, and cost among patients with psoriasis: retrospective analysis of administrative databases in Southern Italy. *Clinicoecon Outcomes Res*. 2017; 9:741–748. Published 2017 Dec 1. <https://doi.org/10.2147/CEOR.S147558> PMID: 29238210
15. Russo V, Monetti VM, Guerriero F, Trama U, Guida A, Menditto E, et al. Prevalence of antibiotic prescription in southern Italian outpatients: real-world data analysis of socioeconomic and sociodemographic variables at a municipality level. *Clinicoecon Outcomes Res*. 2018; 10:251–258. Published 2018 May 3. <https://doi.org/10.2147/CEOR.S161299> PMID: 29765241
16. Iolascon G, Gimigliano F, Orlando V, Capaldo A, Di Somma C, Menditto E. Osteoporosis drugs in real-world clinical practice: an analysis of persistence. *Aging Clin Exp Res*. 2013; 25 Suppl 1:S137–S141. <https://doi.org/10.1007/s40520-013-0127-5> PMID: 24046038
17. Orlando V, Guerriero F, Putignano D, Monetti VM, Tari DU, Farina G, et al. Prescription Patterns of Anti-diabetic Treatment in the Elderly. Results from Southern Italy. *Curr Diabetes Rev*. 2015; 12(2):100–106. <https://doi.org/10.2174/1573399811666150701120408> PMID: 26126718
18. Menditto E, Cahir C, Aza-Pascual-Salcedo M, Bruzzese D, Poblador-Plou B, Malo S, et al. Adherence to chronic medication in older populations: application of a common protocol among three European cohorts. *Patient Prefer Adherence*. 2018; 12:1975–1987. Published 2018 Oct 5. <https://doi.org/10.2147/PPA.S164819> PMID: 30323567
19. Casula M, Catapano AL, Piccinelli R, Menditto E, Manzoli L, De Fendi L, et al. Assessment and potential determinants of compliance and persistence to antiosteoporosis therapy in Italy. *Am J Manag Care*. 2014; 20(5):e138–e145. PMID: 25326928
20. Orlando V, Monetti VM, Moreno Juste A, Russo V, Mucherino S, Trama U, et al. Drug Utilization Pattern of Antibiotics: The Role of Age, Sex and Municipalities in Determining Variation. *Risk Manag Healthc Policy*. 2020; 13:63–71. Published 2020 Jan 29. <https://doi.org/10.2147/RMHP.S223042> PMID: 32099490
21. Orlando V, Coscioni E, Guarino I, Mucherino S, Perrella A, Trama U, et al. Drug-utilisation Profiles and COVID-19: Retrospective Cohort Study in Italy, 29 May 2020, PREPRINT (Version 1) available at Research Square. <https://doi.org/10.21203/rs.3.rs-31829/v>
22. Corman VM, Landt O, Kaiser M, Molenkamp R, Meijer A, Chu DKW, et al. Detection of 2019 novel coronavirus (2019-nCoV) by 9 real-time RT-PCR. *Euro Surveill*. 2020 Jan; 25(3):2000045. <https://doi.org/10.2807/1560-107917.ES.2020.25.3.2000045>
23. Italian Data Protection Authority. General authorisation to process personal data for scientific research purposes— 1 March 2012 [1884019]. 10.1094/PDIS-11-11-0999-PDN.
24. Charlson ME, Pompei P, Ales KL, et al. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis*. 1987; 40(5):373–83. [https://doi.org/10.1016/0021-9681\(87\)90171-8](https://doi.org/10.1016/0021-9681(87)90171-8) PMID: 3558716
25. Elixhauser A, Steiner C, Harris DR, Coffey RM. Comorbidity measures for use with administrative data. *Med Care*. 1998 Jan; 36(1):8–27. <https://doi.org/10.1097/00005650-199801000-00004> PMID: 9431328
26. Von Korff M, Wagner EH, Saunders K. A chronic disease score from automated pharmacy data. *J Clin Epidemiol*. 1992 Feb; 45(2):197–203. [https://doi.org/10.1016/0895-4356\(92\)90016-g](https://doi.org/10.1016/0895-4356(92)90016-g) PMID: 1573438
27. O'Shea M, Teeling M, Bennett K. The prevalence and ingredient cost of chronic comorbidity in the Irish elderly population with medication treated type 2 diabetes: A retrospective cross-sectional study using a national pharmacy claims database. *BMC Health Serv Res* 2013; 13(1):1. <https://doi.org/10.1186/1472-6963-13-23> PMID: 23324517
28. Pratt NL, Kerr M, Barratt JD, Kemp-Casey A, Kalisch Ellett LM, Ramsay E, et al. The validity of the Rx-Risk Comorbidity Index using medicines mapped to the Anatomical Therapeutic Chemical (ATC) Classification System. *BMJ Open* 2018; 8(4):1–8. <https://doi.org/10.1136/bmjopen-2017-021122> PMID: 29654048
29. Wynants L, Van Calster B, Collins GS, Riley RD, Heinze G, Schuit E, et al. Prediction models for diagnosis and prognosis of covid-19 infection: systematic review and critical appraisal. *BMJ*. 2020 Apr 7; 369:m1328. Erratum in: *BMJ*. 2020 Jun 3; 369:m2204. <https://doi.org/10.1136/bmj.m1328> PMID: 32265220
30. Aghagholi G, Gallo Marin B, Soliman LB, Sellke FW. Cardiac involvement in COVID-19 patients: Risk factors, predictors, and complications: A review. *J Card Surg*. 2020 Jun; 35(6):1302–1305. <https://doi.org/10.1111/jocs.14538> PMID: 32306491

31. Wang B, Li R, Lu Z, Huang Y. Does comorbidity increase the risk of patients with COVID-19: evidence from meta-analysis. *Aging (Albany NY)*. 2020 Apr 8; 12(7):6049–6057. <https://doi.org/10.18632/aging.103000>
32. Maio V, Yuen E, Rabinowitz C, Louis D, Jimbo M, Donatini A, et al. Using pharmacy data to identify those with chronic conditions in Emilia Romagna, Italy. *Journal of health services research & policy*, 2005, 10.4: 232–238. <https://doi.org/10.1258/135581905774414259>
33. Canova C, Danieli S, Barbiellini CA, Simonato L, Di RD, Cappai G, et al. A Systematic Review of Case-Identification Algorithms Based on Italian Healthcare Administrative Databases for Three Relevant Diseases of the Nervous System: Parkinson's Disease, Multiple Sclerosis, and Epilepsy. *Epidemiol Prev*. 2019; 43(4 Suppl 2):62–74. <https://doi.org/10.19191/EP19.4.S2.P062.093> PMID: 31650807
34. Bezzini D, Policardo L, Meucci G, Olivelli M, Bartalini S, Profili F, et al. Prevalence of Multiple Sclerosis in Tuscany (Central Italy): A Study Based on Validated Administrative Data. *Neuroepidemiology*. 2016; 46(1):37–42. <https://doi.org/10.1159/000441567> PMID: 26618996
35. Bargagli AM, Colais P, Agabiti N, Mayer F, Buttari F, Centonze D, et al. Prevalence of multiple sclerosis in the Lazio region, Italy: use of an algorithm based on health information systems. *J Neurol*. 2016; 263(4):751–759. <https://doi.org/10.1007/s00415-016-8049-8> PMID: 26886201
36. Chini F, Pezzotti P, Orzella L, Borgia P, Guasticchi G. Can we use the pharmacy data to estimate the prevalence of chronic conditions? a comparison of multiple data sources. *BMC Public Health*. 2011; 11:688. Published 2011 Sep 5. <https://doi.org/10.1186/1471-2458-11-688> PMID: 21892946
37. Tibshirani R. The lasso method for variable selection in the Cox model. *Stat Med*. 1997; 16(4):385–395. [https://doi.org/10.1002/\(sici\)1097-0258\(19970228\)16:4<385::aid-sim380>3.0.co;2-3](https://doi.org/10.1002/(sici)1097-0258(19970228)16:4<385::aid-sim380>3.0.co;2-3) PMID: 9044528
38. Gagne JJ, Glynn RJ, Avorn J, Levin R, Schneeweiss S. A combined comorbidity score predicted mortality in elderly patients better than existing scores. *J Clin Epidemiol*. 2011; 64(7):749–759. <https://doi.org/10.1016/j.jclinepi.2010.10.004> PMID: 21208778
39. Xu H, Qian J, Paynter NP, Zhang X, Whitcomb BW, Tworoger SS, et al. Estimating the receiver operating characteristic curve in matched case control studies. *Stat Med*. 2019; 38(3):437–451. <https://doi.org/10.1002/sim.7986> PMID: 30467878
40. Corrao G, Rea F, Di Martino M, De Palma R, Scondotto S, Fusco D, et al. Developing and validating a novel multisource comorbidity score from administrative data: a large population-based cohort study from Italy. *BMJ Open*. 2017; 7(12):e019503. Published 2017 Dec 26. <https://doi.org/10.1136/bmjopen-2017-019503> PMID: 29282274
41. Corrao G, Rea F, Carle F, Di Martino M, De Palma R, Francesconi P, et al. Measuring multimorbidity inequality across Italy through the multisource comorbidity score: a nationwide study [published online ahead of print, 2020 May 20]. *Eur J Public Health*. 2020;ckaa063. <https://doi.org/10.1093/eurpub/ckaa063> PMID: 32433750
42. Wang BX. Susceptibility and prognosis of COVID-19 patients with cardiovascular disease. *Open Heart*. 2020; 7(1):e001310. <https://doi.org/10.1136/openhrt-2020-001310> PMID: 32587104
43. Carey IM, Critchley JA, DeWilde S, Harris T, Hosking FJ, Cook DG. Risk of Infection in Type 1 and Type 2 Diabetes Compared With the General Population: A Matched Cohort Study. *Diabetes Care*. 2018; 41(3):513–521. <https://doi.org/10.2337/dc17-2131> PMID: 29330152
44. Geerlings SE, Hoepelman AI. Immune dysfunction in patients with diabetes mellitus (DM). *FEMS Immunol Med Microbiol*. 1999; 26(3–4):259–265. <https://doi.org/10.1111/j.1574-695X.1999.tb01397.x> PMID: 10575137
45. Peleg AY, Weerathna T, McCarthy JS, Davis TM. Common infections in diabetes: pathogenesis, management and relationship to glycaemic control. *Diabetes Metab Res Rev*. 2007; 23(1):3–13. <https://doi.org/10.1002/dmrr.682> PMID: 16960917
46. Perico L, Benigni A, Remuzzi G. Should COVID-19 Concern Nephrologists? Why and to What Extent? The Emerging Impasse of Angiotensin Blockade. *Nephron*. 2020; 144(5):213–221. <https://doi.org/10.1159/000507305> PMID: 32203970
47. Cheng Y, Luo R, Wang K, Zhang M, Wang Z, Dong L et al. Kidney disease is associated with in-hospital death of patients with COVID-19. *Kidney Int*. 2020; 97(5):829–838. <https://doi.org/10.1016/j.kint.2020.03.005> PMID: 32247631
48. Diao B, Feng Z, Wang C, Feng Z, Tan Y, Wang H, et al. Human kidney is a target for novel severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection. *medRxiv*. In press. <https://doi.org/10.1101/2020.03.04.20031120>
49. Chowdhury SF, Anwar S. Management of Hemoglobin Disorders During the COVID-19 Pandemic. *Front Med (Lausanne)*. 2020; 7:306. Published 2020 Jun 9. <https://doi.org/10.3389/fmed.2020.00306> PMID: 32582745

50. de León-Rendón JL, Hurtado-Salazar C, Yamamoto-Furusho JK. Aspects of inflammatory bowel disease during the COVID-19 pandemic and general considerations. Aspectos y consideraciones generales en la enfermedad inflamatoria intestinal durante la pandemia por COVID-19. *Rev Gastroenterol Mex.* 2020;S0375-0906(20)30054-9. <https://doi.org/10.1016/j.rgmx.2020.05.001>
51. Favalli EG, Ingegnoli F, De Lucia O, Cincinelli G, Cimaz R, Caporali R. COVID-19 infection and rheumatoid arthritis: Faraway, so close!. *Autoimmun Rev.* 2020; 19(5):102523. <https://doi.org/10.1016/j.autrev.2020.102523> PMID: 32205186
52. Cohen S, Tyrrell DA, Smith AP. Psychological stress and susceptibility to the common cold. *N Engl J Med.* 1991; 325(9):606–612. <https://doi.org/10.1056/NEJM199108293250903> PMID: 1713648
53. Andersson NW, Goodwin RD, Okkels N, Gustafsson LN, Taha F, Cole SW et al. Depression and the risk of severe infections: prospective analyses on a nationwide representative sample. *Int J Epidemiol.* 2016; 45(1):131–139. <https://doi.org/10.1093/ije/dyv333> PMID: 26708840
54. Seminog OO, Goldacre MJ. Risk of pneumonia and pneumococcal disease in people with severe mental illness: English record linkage studies. *Thorax.* 2013; 68(2):171–176. <https://doi.org/10.1136/thoraxjnl-2012-202480> PMID: 23242947
55. Jiang T, Farkas DK, Ahern TP, Lash TL, Sørensen HT, Gradus JL. Posttraumatic Stress Disorder and Incident Infections: A Nationwide Cohort Study. *Epidemiology.* 2019; 30(6):911–917. <https://doi.org/10.1097/EDE.0000000000001071> PMID: 31584893
56. Sharma G, Volgman AS, Michos ED. Sex Differences in Mortality from COVID-19 Pandemic: Are Men Vulnerable and Women Protected? *JACC Case Rep.* 2020;10.1016/j.jaccas.2020.04.027. <https://doi.org/10.1016/j.jaccas.2020.04.027> PMID: 32373791
57. Guan WJ, Liang WH, Zhao Y, Liang H, Chen Z, Li Y et al. Comorbidity and its impact on 1590 patients with COVID-19 in China: a nationwide analysis. *Eur Respir J.* 2020; 55(5):2000547. Published 2020 May 14. <https://doi.org/10.1183/13993003.00547-2020> PMID: 32217650
58. Christensen DM, Strange JE, Gislason G, Torp-Pedersen C, Gerds T, Fosbøl E et al. Charlson Comorbidity Index Score and Risk of Severe Outcome and Death in Danish COVID-19 Patients. *J Gen Intern Med.* 2020;1–3. <https://doi.org/10.1007/s11606-020-05991-z> PMID: 32583345
59. Burstyn I, Yang Y, Schnatter AR. Effects of non-differential exposure misclassification on false conclusions in hypothesis-generating studies. *Int J Environ Res Public Health.* 2014; 11(10):10951–10966. Published 2014 Oct 21. <https://doi.org/10.3390/ijerph111010951> PMID: 25337942
60. Burstyn I, Goldstein ND, Gustafson P. Towards reduction in bias in epidemic curves due to outcome misclassification through Bayesian analysis of time-series of laboratory test results: Case study of COVID-19 in Alberta, Canada and Philadelphia, USA. Preprint. medRxiv. 2020;2020.04.08.20057661. Published 2020 Apr 11. <https://doi.org/10.1101/2020.04.08.20057661>
61. Vultaggio A, Vivarelli E, Virgili G, Lucenteforte E, Bartoloni A, Nozzoli C et al. Prompt predicting of early clinical deterioration of moderate-to-severe COVID-19 patients: usefulness of a combined score using IL-6 in a preliminary study. *J Allergy Clin Immunol Pract* 2020:S2213-2198(20)30611-5. <https://doi.org/10.1016/j.jaip.2020.06.013> PMID: 32565226

MOX Technical Reports, last issues

Dipartimento di Matematica
Politecnico di Milano, Via Bonardi 9 - 20133 Milano (Italy)

- 02/2021** Calissano, A.; Feragen, A; Vantini, S.
Graph-Valued Regression: Prediction of unlabelled networks in a Non-Euclidean Graph-Space
- 03/2021** Torti, A.; Marika, A.; Azzone, G.; Secchi, P.; Vantini S.
Bridge closure in the road network of Lombardy: a spatio-temporal analysis of the socio-economic impacts
- 01/2021** Pegoraro, M.; Beraha, M.
Projected Statistical Methods for Distributional Data on the Real Line with the Wasserstein Metric
- Franco, N.r.; Massi, M.c.; Ieva, F.; Manzoni, A.; Paganoni, A.m.; Zunino, P.; Et al.
Development of a method for generating SNP interaction-aware polygenic risk scores for radiotherapy toxicity
- 83/2020** Hron, K.; Machalova, J.; Menafoglio, A.
Bivariate densities in Bayes spaces: orthogonal decomposition and spline representation
- 84/2020** Vergara, C.; Stella, S.; Maines, M.; Catanzariti, D.; Demattè, C.; Centonze, M.; Nobile, F.; Quaranta
Computational electrophysiology to support the mapping of coronary sinus branches for cardiac resynchronization therapy
- 85/2020** Cavinato, L.; Sollini, M.; Kirienko, M.; Biroli, M.; Ricci, F.; Calderoni, L.; Tabacchi, E.; Nanni
PET radiomics-based lesions representation in Hodgkin lymphoma patients
- 82/2020** Vismara, F; Benacchio, T.; Bonaventura, L.
A seamless, extended DG approach for hyperbolic-parabolic problems on unbounded domains
- 81/2020** Antonietti, P. F.; Mascotto, L.; Verani, M.; Zonca, S.
Stability analysis of polytopic Discontinuous Galerkin approximations of the Stokes problem with applications to fluid-structure interaction problems
- 80/2020** Zingaro, A.; Dede', L.; Menghini, F.; Quarteroni, A.
Hemodynamics of the heart's left atrium based on a Variational Multiscale-LES numerical model