PDSS: A Pharmacological Decision Support System for Diabetics Patients with COVID-19

Isabel AMAYA-RODRIGUEZ,¹, Nekane LARBURU, Maria Rollán MARTINEZ-HERRERA, Kristin REBESCHER, Iván MACIA, Miguel A. ARMENGOL DE LA HOZ, Cristina RUBIO-ESCUDERO and Alba GARIN-MUGA

¹Department of Computer Science, University of Seville, Spain
²Biodonostia Health Research Institute, (Bioengineering Area), eHealth Group, San Sebastián 20014, Spain
³Applied Chest Imaging Laboratory, Brigham and Women’s Hospital, Harvard Medical School, Boston, MA, USA
⁴Big Data Department - Regional Ministry of Health of Southern Spain, Spain

ORCiD ID: Isabel Amaya-Rodríguez https://orcid.org/0000-0002-1073-7116, Nekane Larburu https://orcid.org/0000-0003-0248-7783, Maria Rollán Martinez-Herrera https://orcid.org/0000-0002-4667-521X, Kristin Rebescher https://orcid.org/0000-0002-8169-9655, Ivan Macia https://orcid.org/0000-0003-0448-7840, Miguel A. Armengol de la Hoz https://orcid.org/0000-0002-7012-2973, Cristina Rubio-Escudero https://orcid.org/0000-0001-5732-9139, Alba Garín-Muga https://orcid.org/0000-0002-7160-1191

Abstract. With the advent of SARS-CoV-2, several studies have shown that there is a higher mortality rate in patients with diabetes and, in some cases, it is one of the side effects of overcoming the disease. However, there is no clinical decision support tool or specific treatment protocols for these patients. To tackle this issue, in this paper we present a Pharmacological Decision Support System (PDSS) providing intelligent decision support for COVID-19 diabetic patient treatment selection, based on an analysis of risk factors with data from electronic medical records using Cox regression. The goal of the system is to create real world evidence including the ability to continuously learn to improve clinical practice and outcomes of diabetic patients with COVID-19.

Keywords. Diabetes; Clinical decision support systems; COVID-19; Risk Factor

1. Introduction

On January 30th, 2020, the World Health Organization (WHO) declared a public health emergency of international concern due to coronavirus disease (COVID-19). COVID-19 is caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and is characterised by its strong transmission capacity, and although most patients experience mild symptoms or moderate illness, according to the WHO, 10-15% of cases progress to
severe disease. As of December 15th, 2022, there were 646,740,524 confirmed COVID-19 cases [1] and approximately 26% of detected COVID-19 patients have Diabetes Mellitus (DM) [2] depending on the study. Due to the high incidence of diabetics with COVID-19, several studies have analysed and reported on the negative impact that DM may have on patients admitted to the hospital with COVID-19 based on survival analysis (SA) [2]. However, to the best of our knowledge, no article in the literature proposes the development of a pharmacological decision support system for DM patients with COVID-19 based on SA. In this article, we propose a recommendation system based on risk factors attending to their impact on survival of various drugs administered DM patients with COVID-19 treated in HM Hospitales (HM), a private hospital group based in Spain. The results obtained may be useful in clarifying the optimal pharmacological protocol for COVID-19 diabetes patients. Proposed analysis methodologies may also apply to other PDSS in other disease areas.

2. Methods

Given the unprecedented nature of the pandemic and the wealth of information held within electronic medical records, HM, recognized early on the importance of making anonymised COVID-19 data publicly available. Through its “Covid Data Saves lives” project, HM published anonymised clinical, demographic, and laboratory data for the first time on April 15th 2020 to identify trends and narrow in on establishing a standard of care for COVID-19 patients more rapidly. The study was approved by the Ethics committee of the HM Hospitals (approval number 22.04.2005-GHM). After access granting, the raw data (N=2561) was cleaned to improve its completeness and usability removing patients with missing values or outliers and those cases without confirmed COVID-19 diagnosis or incomplete data, which were approximately 21% of the total. The final cohort included 2,101 patients diagnosed with SARS-Cov-2 infection. The variables in the dataset included demographic information, laboratory analyses, comorbidities, and administered pharmaceutical treatment as part of their care during their hospital stay, among others.

The purpose of this study is to evaluate the effectiveness and factors affecting outcomes for drugs administered to COVID-19 patients with DM. Therefore, patients that had the following ICD-10 diagnosis codes (E11.9, E11.5, E11.22, E11.319, and O24.410) were considered patients with DM, which were 18.18% of the total. Additionally, we selected the patients who were administered the most common pharmaceutical treatments against COVID-19 such as Hydroxychloroquine, Azithromycin, Azithromycin/Hydroxychloroquine combination, Lopinavir/Ritonavir combination, Antiviral HIV, Tocilizumab, Oseltamivir, Interleukin 1 inhibitors, Interferons, Glucocorticoids, Dexamethasone, Hydrocortisone, Heparine and Organic Nitrate.

A preliminary extraction of features was performed to simplify the variables that were collected from the EHR of the patients involved in the analysis. The variables selected in this pre-processing were chosen following clinical criteria and those that had statistical significance (p<0.005) between diabetic and non-diabetic patients in the data set. The characteristics analysed were age, sex, hospitalization days (1-7 days, 8-14 days, 15-30 days, 30-90 days, 90+ days), discharge status (home, voluntary, other hospital, social-sanitary centre, death), comorbidities and ICU admission (yes/no).
Once these variables were identified, the analysis then focused on obtaining which drugs improve or worsen the survival of diabetics with COVID-19 and what are the influencing factors for survival. Thus, given these influencing factors, potential hypothesis to generate real world evidence (RWE) in the form of clinical rules were obtained following the procedure of evidence-based clinical guideline development [3]. Next, we present the components of our PDSS for DM patients with COVID-19.

2.1. Pharmacological Decision Support System Components

The PDSS consists of two main components: i) the Risk Factors Discovery Engine that analyses the key risk factors related to the different treatments for COVID-19 in diabetic patients and ii) the Pharmacological Rule System, that making usage of the key factors formalizes new knowledge and, based on input clinical information from the patient's EHR, returns as a result a set of drug recommendations or contraindications for a diabetic patient [4].

2.2. Risk Factors Discovery Engine

As mentioned in the methods, the factors included in the risk analysis were obtained from the preliminary extraction to guide the identification of the relevant factors that affect the drug administration. They were selected by applying the different statistical tests mentioned below using a value of p < 0.05 as the threshold applied to determine statistical significance. Categorical data were analysed using the Chi-square test and in the case of numerical data, Fisher's exact test or the Mann-Whitney U test was applied.

Once the selected variables were included, we need to identify which are the influencing factors to generate the hypothesis for our study. The following steps were carried out:

1. For each drug:
   a. Diabetic patients were divided into two groups, those who had received the drug and those who had not.
   b. Survival analysis with Kaplan-Meier curves was performed with both groups and the Log Rank test was applied to demonstrate differences between cohorts when p<0.05 (Table 1).

2. If differences were demonstrated (1.b.), Cox regression analysis was performed to look for factors influencing patient survival (Table 1)

When the hypothesis is validated with enough evidence (RWE), the corresponding rules are implemented in the Pharmacological Rule System (PRB), which incorporate the relevant risk factors obtained (2).

2.3. Pharmacological Rule System

Once the risk factors were identified, RWE in the form of digital rules needed to be implemented, following similar approaches to guideline-based CDSS systems [4]. First the subjects that are involved in the study need to be identified. In this case DM patients with COVID-19. Next, the PRS is able to receive as input the relevant patient clinical EHR data, collect the data from the variables obtained in the preliminary extraction and
return whether the drug may be recommended for the patient or not. Additionally, the system may evolve and automatically create new RWE based rules improving the knowledge base [4]. The following image shows the workflow of the PDSS system:

![Figure 1. PDSS workflow](image)

3. Results

Table 1 includes the significant results of the survival analysis comparison in patients with DM who received or not a treatment after applying Kaplan-Meier and Log-Rank. As result, the drugs with relevant impact in the survival rate of diabetics were Hydroxychloroquine, Azithromycin, Azithromycin/Hydroxychloroquine combination, Lopinavir/Ritonavir combination, Antiviral HIV and Heparin.

The resulting preliminary extraction variables were the age, the hospitalization days between 1-7 days and several comorbidities such as Hypertension, Obesity, Allergies, Hyperlipidemia/Dyslipidemia, Apnoea, Chronic Kidney Disease, Chronic Heart Disease and Acute Renal Failure. Despite of sex and ICU admission not being statistically significative in the preliminary extraction, they were added in the risk factors discovery engine due to previous studies showed that they are important facts in the survival of COVID-19 patients and due to this it was added [5,6]. The age was divided in two groups <65 years and >= 65 years as was applied by [2].

![Table 1. Survival and risk factors results by treatment](table)

<table>
<thead>
<tr>
<th>Drug Name</th>
<th>N(%) Diabetics with the treatment</th>
<th>Log Rank (p value)</th>
<th>Factors do not Improve Survival</th>
<th>Factors improving Survival</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hydroxychloroquine</td>
<td>308 (83.6)</td>
<td>&lt;0.005</td>
<td>-</td>
<td>ICU and Hypertension and Allergies</td>
</tr>
<tr>
<td>Azithromycin</td>
<td>215 (58.4)</td>
<td>.04</td>
<td>-</td>
<td>ICU and Hypertension and Allergies</td>
</tr>
<tr>
<td>Azithromycin / Hydroxychloroquine combination</td>
<td>202 (54.8)</td>
<td>&lt;.005</td>
<td>-</td>
<td>ICU and Hypertension and Allergies</td>
</tr>
<tr>
<td>Lopinavir/ritonavir combination</td>
<td>154 (41.8)</td>
<td>.03</td>
<td>Age &lt; 65 years</td>
<td>ICU and Hyperlipidemia /Dyslipidemia and Acute Renal Failure</td>
</tr>
<tr>
<td>Antiviral HIV</td>
<td>157 (42.6)</td>
<td>.02</td>
<td>Age &lt; 65 years</td>
<td>ICU and Hyperlipidemia /Dyslipidemia and Acute Renal Failure</td>
</tr>
<tr>
<td>Heparine</td>
<td>10 (84.2)</td>
<td>&lt;0.005</td>
<td>Sex == Female</td>
<td>ICU and Allergies</td>
</tr>
</tbody>
</table>

1 Log Rank (p value) for DM patients who had received the drug and those who had not; 2 Factors do not Improve Survival obtained in COX regression analysis; 3 Factors improving Survival obtained in COX regression analysis
The results presented in Table 1 could be transformed to multiple hypothetic rules form which the conditions would be the identified risk factors. As an example, below one potential hypothetic RWE-based rule for a patient under 65 years with COVID-19 and DM is presented:

\[
\text{IF} \ (\text{Age} < 65) \ && \ (\text{DIABETES} == \text{TRUE}) \ && \ (\text{COVID-19} == \text{TRUE}) \\
\text{THEN SHOULD NOT} \ (\text{“Lopinavir/ritonavir combination”}) \\
\text{AND SHOULD NOT} \ (\text{“HIV antivirals”})
\]

4. Discussion and Conclusion

We have proposed a methodology to obtain hypothesis for COVID-19 patients with DM based on a preliminary analysis of risk factors affecting outcomes for given treatments. This methodology allows to generate RWE on cases based on statistical analysis using survival analysis (Kaplan-Meier and Cox regression) and translate this evidence into rule-based decision support systems. The development of such a system is relevant in scenarios such a pandemic or health crisis situations, where evidence needs to be generated quickly in absence of further evidence from clinical trials. Similar approaches may be followed for other diseases in real-world data applications such as survival analysis in value-based healthcare scenarios. As future work, the recommendations obtained will be validated by clinical professionals. Besides, further variables should be analysed to improve the risk factors recommendations. For example, allergies or specific variables related to comorbidities, such as Creatinine in the case of Acute Renal Failure. Finally, data-driven approaches, such as machine learning models, could be incorporated for complex data scenarios, to complement statistical approaches in outcome prediction.

5. References