SMART COVID Navigator: A Clinical Decision Support for COVID-19

Varun Suraj, Catherine Del Vecchio Fitz, Laura B. Kleiman, Suresh Bhavnani, Chinmay Jani, Surbhi Shah, Rana Mckay, Jeremy Warner, Gil Alterovitz

Submitted to: Journal of Medical Internet Research
on: April 02, 2021

Table of Contents

SMART COVID Navigator: A Clinical Decision Support for COVID-19

Varun Suraj[1]; Catherine Del Vecchio Fitz[2]; Laura B. Kleiman[2]; Suresh Bhavnani[3*] PhD; Chinmay Jani[4*] MD; Surbhi Shah[5*] MBBS; Rana Mckay[6]; Jeremy Warner[7]; Gil Alterovitz[8*]

[1]Brigham and Women's Hospital Boston US
[2]Reboot Rx Boston US
[3]Preventive Medicine and Population Health Institute for Translational Sciences University of Texas Medical Branch University of Texas Health Science Center in Houston Houston US
[4]Department of Internal Medicine, Mount Auburn Hospital- Harvard Medical School Cambridge US
[5]Hematology, oncology and bone marrow transplantation Mayo Clinic Arizona Phoenix US
[6]University of California San Diego San Diego US
[7]VUMC Nashville US
[8]Brigham and Women's Hospital/Harvard Medical School BOSTON US
[*]these authors contributed equally

Corresponding Author:
Gil Alterovitz
Brigham and Women's Hospital/Harvard Medical School
75 Francis Street,
BOSTON
US

Abstract

Background: The coronavirus disease 2019 (COVID-19) caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has infected 219 million individuals at the time of writing. A large volume of research findings from observational studies about disease interactions with COVID-19 are being produced almost daily, making it difficult for physicians to keep track of the latest information on COVID-19's effect on patients with certain pre-existing conditions.

Objective: In this paper, we describe the creation of a clinical decision support tool, the SMART COVID Navigator, a web application to assist clinicians in treating COVID-19 patients. Our app allows clinicians to access a patient's electronic health records and identify disease interactions from a large set of observational research studies that affect severity and fatality due to COVID-19.

Methods: The SMART COVID Navigator takes a two-pronged approach to clinical decision support. The first part is a connection to electronic health record servers, allowing the app to access a patient's medical conditions. The second is accessing datasets with information on various observational studies to determine the latest research findings about COVID-19 outcomes for patients with certain medical conditions. By connecting these two data sources, users can see how a patient's medical history will affect their COVID-19 outcomes.

Results: The SMART COVID Navigator aggregates patient health information from multiple Fast Healthcare Interoperability Resources-enabled electronic health record systems. This allows physicians to see a comprehensive view of patient health records. The application accesses two datasets of over 1,100 research studies to provide information on fatality and severity of COVID-19 for several pre-existing conditions. We also analyze the results of the collected studies to determine which medical conditions result in an increased chance of severity and/or fatality of COVID-19 progression. We find that certain conditions result in a higher likelihood of severity and fatality probabilities. We also analyze various cancer tissues and find that the probabilities for fatality vary greatly depending on the tissue being examined.

Conclusions: The SMART COVID Navigator allows physicians to predict fatality and severity of COVID-19 progression given a particular patient's medical conditions. This can allow physicians to determine how aggressively to treat patients infected with COVID-19 and to prioritize different patients for treatment taking into account their prior medical conditions.

(JMIR Preprints 02/04/2021:29279)
DOI: https://doi.org/10.2196/preprints.29279

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

✓ **Please make my preprint PDF available to anyone at any time (recommended).**
Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
Only make the preprint title and abstract visible.
No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**
Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v
Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http

Original Manuscript

SMART COVID Navigator: A Clinical Decision Support for COVID-19

Varun Suraj, Catherine Del Vecchio Fitz, Laura B. Kleiman, Suresh Bhavnani, Chinmay Jani, Surbhi Shah, Rana R. McKay, Jeremy Warner, Gil Alterovitz

Background:

The coronavirus disease 2019 (COVID-19) caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has infected 219 million individuals at the time of writing. A large volume of research findings from observational studies about disease interactions with COVID-19 are being produced almost daily, making it difficult for physicians to keep track of the latest information on COVID-19's effect on patients with certain pre-existing conditions.

Objective:

In this paper, we describe the creation of a clinical decision support tool, the SMART COVID Navigator, a web application to assist clinicians in treating COVID-19 patients. Our app allows clinicians to access a patient's electronic health records and identify disease interactions from a large set of observational research studies that affect severity and fatality due to COVID-19.

Methods:

The SMART COVID Navigator takes a two-pronged approach to clinical decision support. The first part is a connection to electronic health record servers, allowing the app to access a patient's medical conditions. The second is accessing datasets with information on various observational studies to determine the latest research findings about COVID-19 outcomes for patients with certain medical conditions. By connecting these two data sources, users can see how a patient's medical history will affect their COVID-19 outcomes.

Results:

The SMART COVID Navigator aggregates patient health information from multiple Fast Healthcare Interoperability Resources-enabled electronic health record systems. This allows physicians to see a comprehensive view of patient health records. The application accesses two datasets of over 1,100 research studies to provide information on fatality and severity of COVID-19 for several pre-existing conditions. We also analyze the results of the collected studies to determine which medical conditions result in an increased chance of severity and/or fatality of COVID-19 progression. We find that certain conditions result in a higher likelihood of severity and fatality probabilities. We also analyze various cancer tissues and find that the probabilities for fatality vary greatly depending on the tissue being examined.

Conclusions:

The SMART COVID Navigator allows physicians to predict fatality and severity of COVID-19 progression given a particular patient's medical conditions. This can allow physicians to determine how aggressively to treat patients infected with COVID-19 and to prioritize different patients for treatment taking into account their prior medical conditions.

Keywords: COVID-19; clinical decision support; precision medicine; web application; FHIR

1. Introduction

Precision medicine, as defined by the National Institutes of Health's Precision Medicine Initiative, is "an emerging approach to disease treatment and prevention that takes into account individual variability in genes, environment, and lifestyle for each person" [1]. Recent increases in the availability of electronic patient data have facilitated the development of precision medicine. For instance, EHRs, or electronic health records, store all information collected in hospitals, such as blood tests, X-rays, any diagnostic tests, and any biographical information about a patient, such as their age, weight or height [2]. This approach to medical treatment could prove to be useful in dealing with the coronavirus disease 2019 (COVID-19) caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), which has been ravaging the US and the world [3, 4]. Currently, a vast amount of research is being conducted to understand how patient underlying health conditions interact with the virus' progression. For example, certain conditions such as diabetes and heart disease have been found to raise the severity and fatality rates for patients who become infected with the virus [5]. Applying the tools of precision medicine can help doctors customize treatment for patients affected by COVID-19 based on the patient's underlying conditions. Attempts at clinical decision support systems for COVID-19 have been made using information on risk factors and biomarker measurements [6, 7].

With the rapid spread of COVID-19 and scarcity of physician resources and time, there is an immediate need for a clinical decision support system that provides patient and disease interaction information to clinicians to allow them to practice precision medicine. In this paper, we describe the construction of the SMART COVID Navigator, a web-based app designed to assist clinicians to relate patient risk factors to the growing amount of research that identifies how various underlying patient conditions affect the progression of COVID-19. The app identifies patient risk factors based on integrating comprehensive data available through multiple electronic health records (EHRs). The app then allows clinicians to quickly access a large set of research studies based on their patient's medical history in an easy-to-access format, helping promote updated information on COVID-19 and its risks for a diverse community. The tool will simplify a clinician's search for relevant research and findings, and support clinical treatment.

2. Background

The SMART COVID Navigator connects patient information from multiple EHR servers to two databases of COVID-19 research studies. This will allow clinicians to access data-driven research based on a particular patient's risk factors. The SMART COVID Navigator builds on the framework developed by the SMART Cancer Navigator, which offered clinical decision support by connecting patient EHR information to cancerous gene-variants [8, 9]. The Navigator is a further step in creating FHIR-protocol based tools to support personalized medicine [10].

The web-based app was built by researchers at the Biomedical Cybernetics Laboratory in Harvard Medical School. It was created using an Angular and bootstrap front-end framework. The code is available at https://github.com/smart-covid-navigator/Covid-Application.

3. Methods

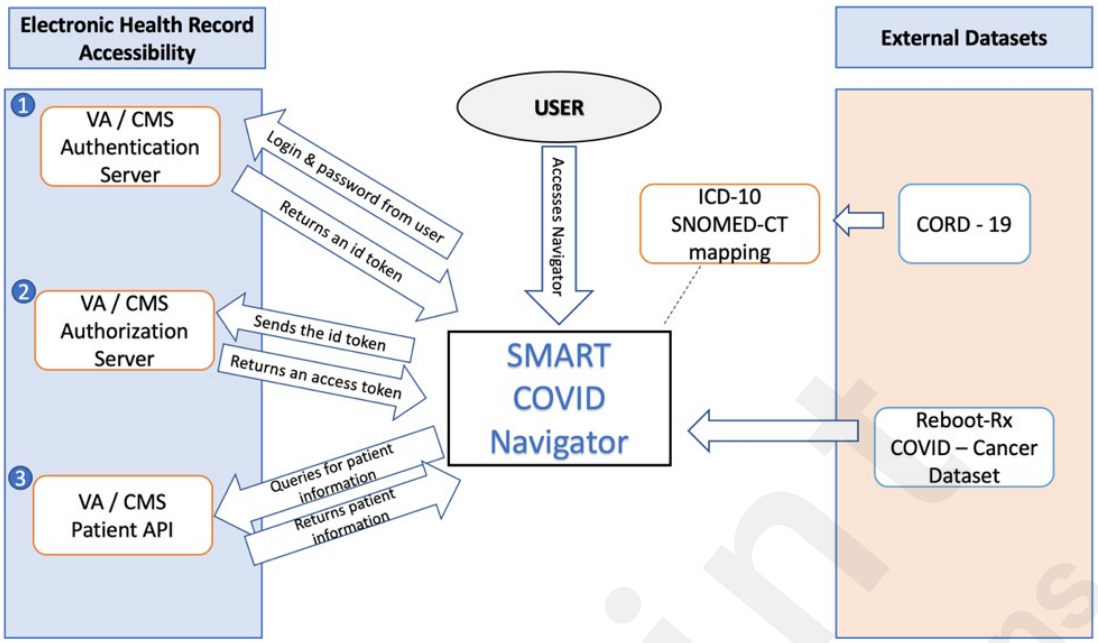

Figure 1: Architecture diagram of the SMART COVID Navigator.

3.1. *Accessing Multiple FHIR Servers*

The SMART COVID Navigator allows for the user to log into two EHRs: the Veterans Affairs (VA) and the Center for Medicare and Medicaid Services (CMS). There are significant advantages to the app being able to log into multiple EHR servers [11]. Doctors viewing their patient's medical information through the app have the ability to discern any discrepancies between the data sources. Similarly, doctors and patients can be sure that they are getting the most up-to-date information regarding their health records, as any information not captured in one of the EHRs would most likely be shown in the other.

The VA and CMS servers follow the Fast Healthcare Interoperability Resources (FHIR) standard [12]. The FHIR platform allows for the interoperability of the navigator with other EHRs which follow the FHIR protocol. The app is registered with the relevant EHRs. The system can be expanded to access additional health records provided their APIs follow the FHIR standard. The EHRs provide the app with a client ID and a client secret, which is used for authentication purposes. This section describes the login process.

To log into the two API-enabled EHR systems, the app implements the OAuth2 [13] and OpenID Connect [14] standards to achieve secure authentication and authorization. The authentication process involves a user entering their login credentials, while the authorization process requires a request to the EHR's server to obtain an access token. Upon receiving the access token, the app is authorized to retrieve relevant patient information.

The left side of Figure 1 depicts the system architecture for the EHR access process. The login process begins when the user clicks one of the login buttons. The user is redirected to the login portal for the EHR of their choice to enter their login credentials. The OAuth2 process redirects the user back to the app with an ID token in the URL (step 1 in Figure 1). With this ID token, the app then requests an access token from the relevant API to get access to the patient's medical information. The app accomplishes this second step by sending the ID token - along with additional information such as the app's client ID and client secret - back to the EHR server through an HTML POST request. The server then returns an access token, which the app can use to gain access to any part of the patient's profile (step 2 in Figure 1). The access token is saved to the local storage of the app; thus, if the user refreshes the app or attempts to log into the other EHR, they will still retain

access to the first EHR. This two-step process ensures better security of sensitive medical data.

From the EHRs, the following information is retrieved: 1) the patient's name, 2) their location zip code, 3) their date of birth (which is used to calculate their age), and 4) a list of medical conditions associated with the patient, along with a numerical code presenting the condition (step 3 in Figure 1).

The EHR server results are visually displayed at the top of the screen. This menu appears when a user logs either one of the servers. It displays demographic information such as the patient's name, age, and current zip code. It also shows a list of medical conditions retrieved from the patient's profile, with the list of conditions appearing in a dropdown menu (Fig. 2). If the user is logged into both the VA and CMS servers, then the app retrieves demographic information from the VA server, while the condition list is a combined list from both EHRs. Please note that the patient information displayed in Figure 2 and in other figures below are based on information taken from VA and CMS simulated patient data; this data is not from a real patient, so there are no HIPAA concerns.

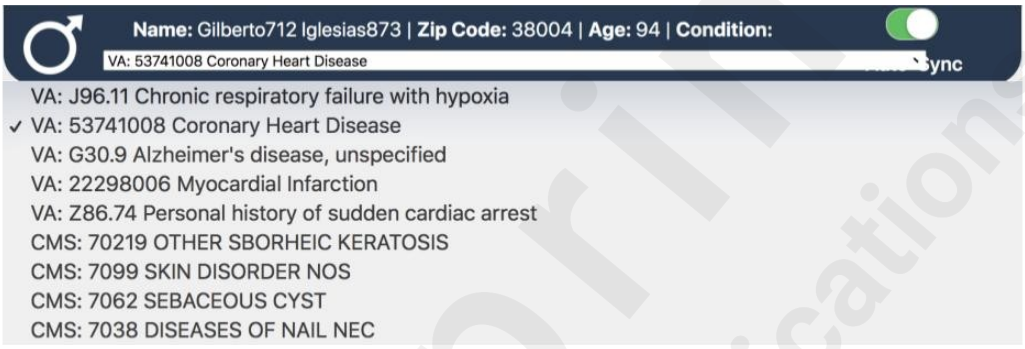

Fig. 2. The disease condition list for a VA test patient. This dropdown appears from the patient header.

With these patient conditions now available in the Navigator, the next step is to connect these conditions to risk factors shown to affect COVID-19 progression.

3.2. *Accessing AI-Powered Knowledge Base from CORD-19*

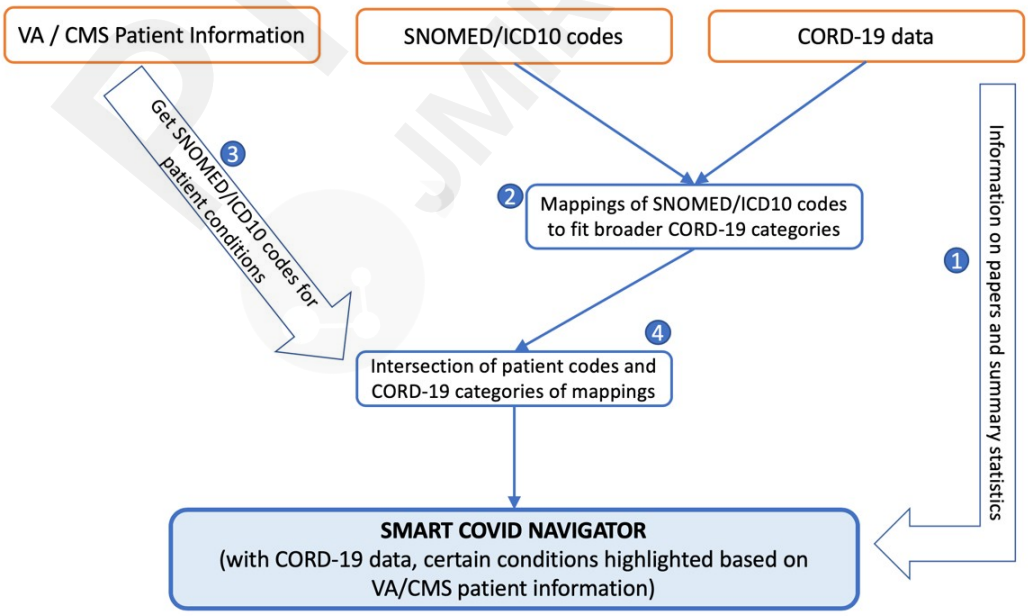

Fig. 3. Architecture diagram of the interactions between the VA and CMS EHR systems, SNOMED/ICD-10 codes, and the CORD-19 dataset.

The SMART COVID Navigator collects risk factor information from two external datasets, as shown

on the right side of Figure 1, and relates it to patient medical conditions. The first dataset is the COVID-19 Open Research Dataset (CORD-19), created by the White House and a set of leading research groups such as the Georgetown University's Center for Security and Emerging Technology and the National Library of Medicine – National Institutes of Health [15]. The resource currently contains over 200,000 scholarly articles related to COVID-19. Researchers apply AI techniques to create this knowledge base. CORD-19 sources papers from the World Health Organization, PubMed Central, bioRxiv, and medRxiv. These can be in the form of physical print-outs, PDF files, or XML files. The papers and preprints are then collected by Semantic Scholar [16], and the resulting metadata are harmonized and deduplicated. The full text of the papers is then extracted [17]. This knowledge base is stored on Kaggle, an online community of data scientists and machine learning practitioners [18]. The Navigator utilizes the "risk factors" section of CORD-19 to support clinical assessment of how a patient's medical conditions affect their chances of having a severe or fatal COVID-19 infection (step 1 in Figure 3).

The CORD-19 data is publicly available and is stored in comma separated value (CSV) files. Each file tracks studies relating to a specific condition (for example, hypertension and heart disease) that is being tracked. Each row in the CSV file represents a study. The study name, link, date, and significance of the severity and fatality of the patients are all provided. Table 1 provides a list of risk factors provided at CORD-19.

As shown in table 1, the CORD-19 dataset consists of 28 risk factors - 22 of which are disease conditions, while the other six are patient biographical characteristics. The comprehensive clinical terminology coding systems, SNOMED-CT used by the VA [19] and ICD-10 used by CMS [20], consists of thousands of more specific disease identifiers than the broader disease categories tracked by CORD-19. We mapped the identifiers in the SNOMED-CT and ICD-10 code sets to the CORD-19 risk factors to be able to match patient conditions to disease categories tracked by CORD-19 (step 2 in Figure 3). This mapping was accomplished through word matching. For each of the 28 risk factors present in CORD-19, the name of the risk factor – or a part of the name – was matched with every instance of that name occurring in the SNOMED-CT and ICD-10 code sets using a Python script created by the authors. For instance, the risk factor "diabetes" links to any disease classification in the code sets that has the phrase "diabetes" or variations such as "diabetic". Because the disease identifiers in the code sets are more specific than the broad risk factors tracked by CORD-19, many risk factors align with hundreds of more specific disease names from the code sets. Since the mapping involved matching words or parts of words using Python code (and was not done manually), we believe there is unlikely to be incorrect mapping. For example, we searched for the string "diabet" which would capture both "diabetes" and "diabetic". Nevertheless, it is possible that some conditions in the two code sets might not have been captured despite our attempts at careful mapping.

Table 1. Risk Factors Tracked by CORD-19

Risk Factors	
Age	Endocrine diseases
Asthma	Ethnicity: Hispanic vs. non-Hispanic
Autoimmune disorders	Heart Disease
COPD	Heart Failure
Cancer	Hypertension
Cardio- and cerebrovascular disease	Immune system disorders
Cerebrovascular disease	Male gender
Chronic digestive disorders	Neurological disorders
Chronic kidney disease	Overweight or obese

Chronic liver disease	Race: Asian vs. White
Chronic respiratory diseases	Race: Black vs. White
Dementia	Race: Other vs. White
Diabetes	Respiratory system diseases
Drinking	Smoking Status

Visually in the Navigator, the information from CORD-19 regarding studies related to COVID-19 risk factors appears as a menu of conditions (Fig. 4). When the user clicks a particular condition, a pop-up screen is displayed, providing all information collected about studies relevant to that condition (Fig. 5). This will allow a user to view studies and use the findings for clinical decision support. The study name (with a hyperlink to the full study), date, and significance of the severity and fatality statistics are shown for each study. At the top of the pop-up, an overview of the studies associated with the chosen condition is displayed. The information included are the total number of studies for that condition, the percentage of studies that found the selected condition to cause a significant change in severity of COVID-19 progression (out of the studies that measure for severity), and the percentage of studies that found a significant change in fatality due to COVID-19. We alert the user if the proportion of papers finding a significant result is greater than 50 percent. If a particular risk factor is found to be associated with severity or fatality from COVID-19 infection by a high proportion of observational studies, then physicians should pay more attention to patients with that risk factor. However, individual physicians should make their own decisions based on the information in the SMART COVID Navigator and the severity of their patients' health conditions.

The patient biographical and health information retrieved from the electronic health records (see section 3.2) can be used to better guide clinicians on the health conditions they should be concerned about in relation to COVID-19. The Navigator uses the previously described mappings from SNOMED-CT and ICD-10 identifiers to the CORD-19 risk factors; wherever there is a match between the code in a patient's EHR condition list (step 3 of Figure 3) and those associated with each CORD-19 risk factor, the condition highlights orange (step 4 of Figure 3), thereby alerting the user to presence of the condition of concern (Fig. 4). If the patient is above 60 years of age, then the "Age" condition is highlighted as well. The clinician should click on any highlighted conditions to access further information from the CORD-19 database, as this information is likely to be relevant for the patient in question.

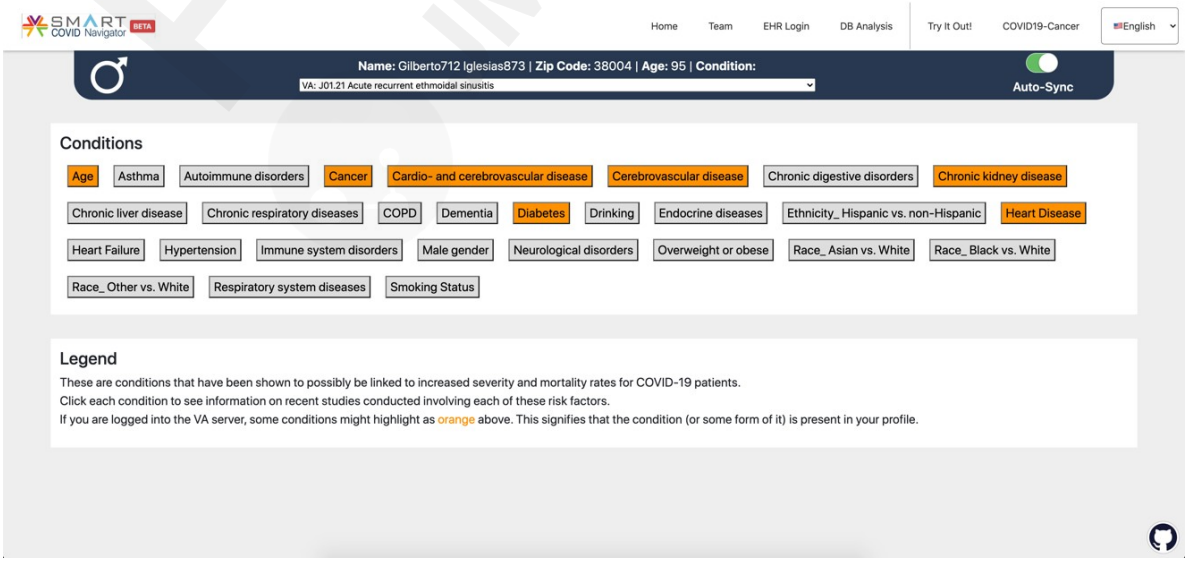

Fig. 4. A view of the app logged into a test VA profile. A button appears for each condition that is being tracked. Based on the conditions in the patient's profile, certain risk factors on the screen are highlighted orange.

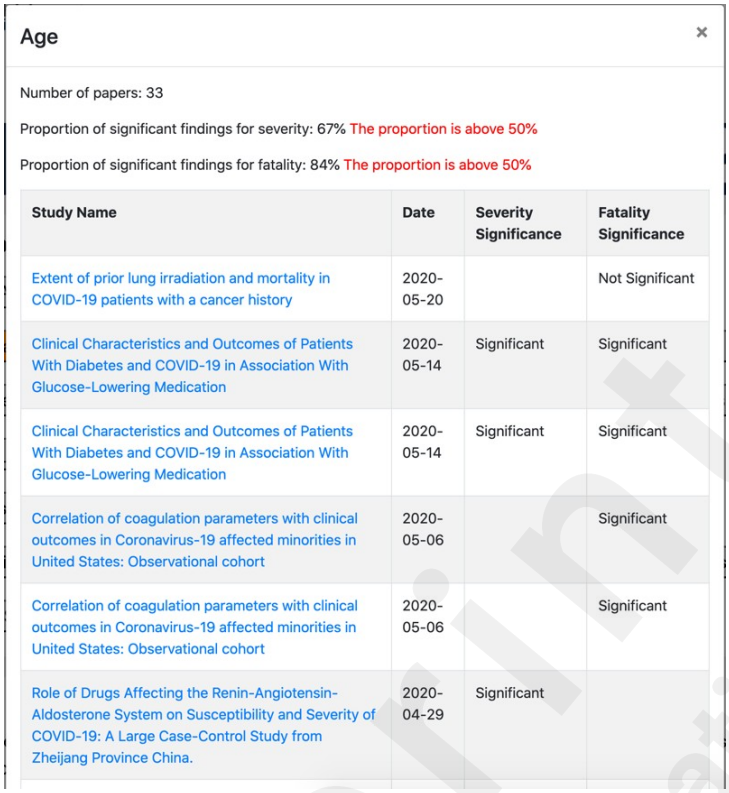

Fig. 5. This figure shows how the app presents information from CORD-19 relating to a particular risk factor (in this case, the example is for the “Age” risk factor).

3.3. *COVID-19 in Cancer Patients*

Cancer has been shown to be a risk factor for COVID-19, and different types of cancers affect COVID-19 progression differently [21], therefore it is important for physicians to understand disease progression if cancer patients get infected with COVID-19. To help with this, the SMART COVID Navigator accesses RebootRx’s Reboot: COVID-Cancer Project datasets in order to display information regarding the impact of COVID-19 and its treatment on cancer patients [22]. The Reboot: COVID-Cancer Project identifies relevant published clinical studies and extracts and aggregates the data from those studies into two datasets. The first dataset examines COVID-19 disease progression for cancer patients. The second dataset examines what effect drugs currently being tested for COVID-19 treatment could have on cancer (independent of their effect on COVID-19). The datasets are publicly available on the Reboot: COVID-Cancer Project website via interaction dashboards and can be requested in the form of Excel files. This feature gives an organized tool for clinicians to better understand COVID-19 outcomes for different types of cancer.

Table 2 shows each of the tissue types that are examined in the datasets (the numbers in parentheses indicate which datasets the tissue type is in). As shown in table 2, the datasets consist of 30 tissue types. Tissue types were chosen to distinguish the cancers instead of cancer types, as the number of tissue types was more manageable and therefore would be easier to access for the user of the application.

Table 2: Tissue Types Tracked by the Reboot: COVID-Cancer Project Datasets

Tissue Types		
Bladder/Unitary Tract (1, 2)	Head and neck (1, 2)	Pancreas (2)
Bone and soft tissue (1)	Hematologic not specified (1)	Pleura (1)
Bowel (1, 2)	Kidney (1, 2)	Prostate (1, 2)
Bowel, Esophagus/Stomach (2)	Liver (1, 2)	Sarcoma (2)

Brain/CNS (1, 2)	Lung (1, 2)	Skin (1, 2)
Breast (1, 2)	Lymphoid (1, 2)	Soft tissue (2)
Cervix (1)	Lymphoid, Myeloid (2)	Testis (2)
Esophagus/Stomach (2)	Myeloid (1, 2)	Thoracic (1)
Genitourinary (1)	Not specified (1, 2)	Thymus (1, 2)
Gynecological (1)	Ovary/Fallopian Tube (1, 2)	Thyroid (2)

When a user navigates to the “COVID19-Cancer” tab of the SMART COVID Navigator, they will see a menu of tissue types, similar to how the CORD-19 risk factors were displayed, as well as another button that displays summary information for all of the tissues in one pop-up, colored blue to distinguish it from the other buttons representing tissue types (Fig. 6). When one of the buttons is clicked, a pop-up appears with two tabs. The first tab displays information retrieved from dataset 1 regarding patient outcomes, and the second tab displays information from dataset 2 about COVID-19 drugs that could be useful for cancer treatment (Fig. 7). When the summary tab is clicked, information retrieved from dataset 1 about all tissue types in that dataset are displayed, so that clinicians can easily see on one central page all of the patient outcome statistics (Fig. 8).

Please note that the use of dataset 1 in the SMART COVID Navigator is to display patient outcomes for the treatment of COVID-19 in cancer patients. This use is similar to how we used the CORD-19 dataset. The display of dataset 2 is for a different purpose – to inform physicians treating cancer how drugs that are being used for COVID-19 treatment have an application in cancer treatment.

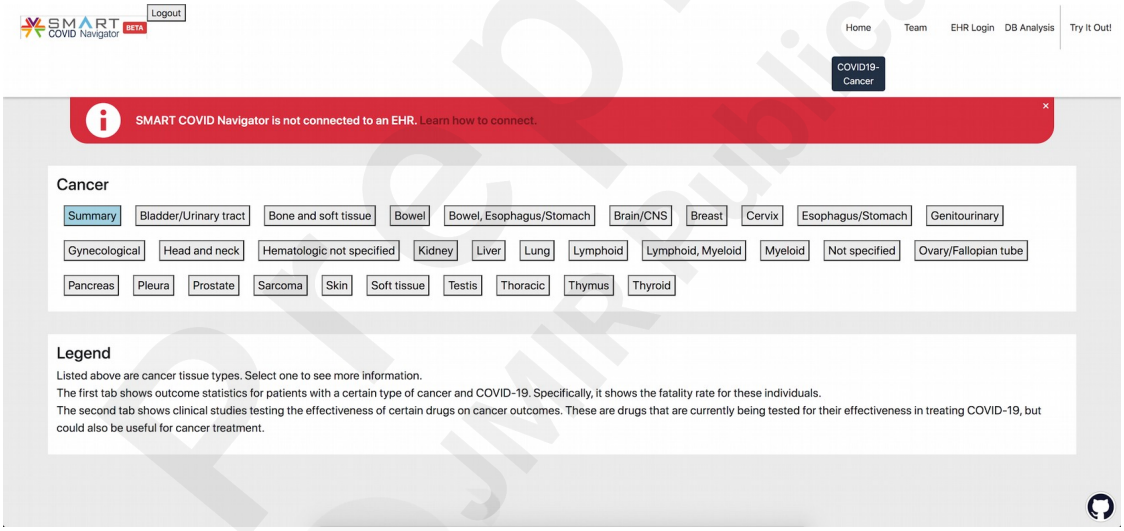

Fig. 6. The standard view of the COVID19-Cancer tab of the COVID Navigator. The “Summary” button is highlighted blue to distinguish it from the other buttons that represent tissue types.

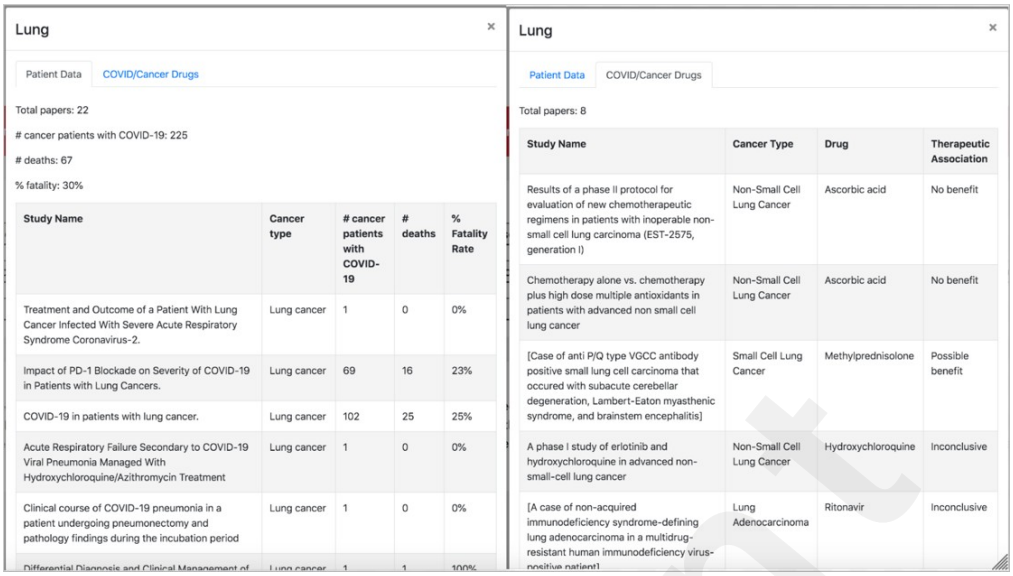

Fig. 7. When a tissue is clicked, a pop-up appears with two tabs (this figure shows the pop-up for lung tissue). The first tab (on the left) tracks patient outcomes for lung cancer patients who contract COVID-19. The second tab (on the right) tracks drugs being tested for COVID-19 that could have an impact on cancer of the selected tissue.

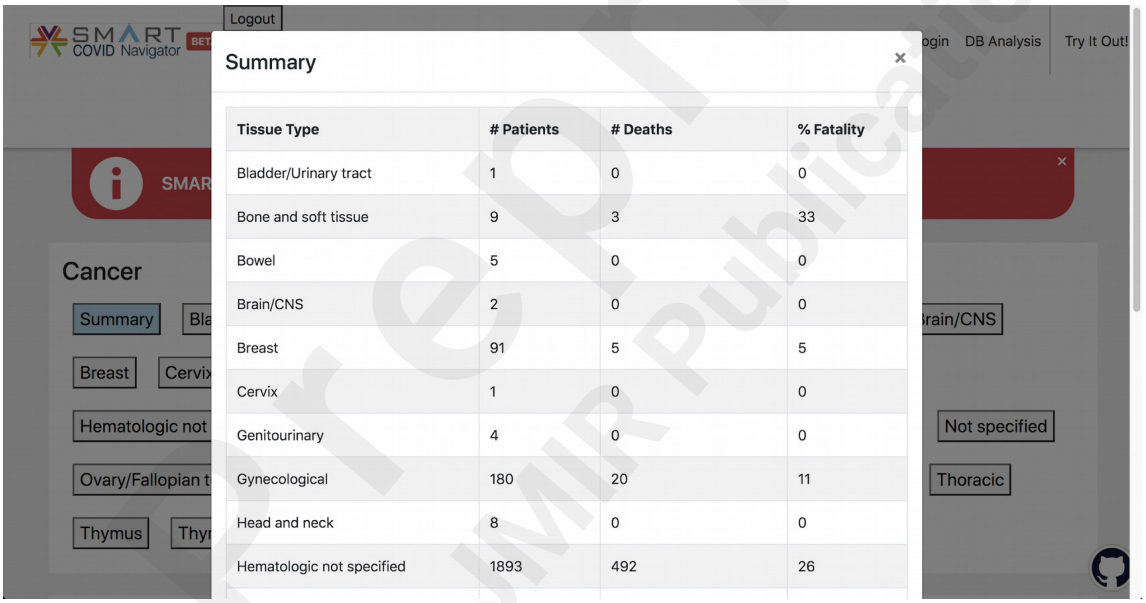

Fig. 8. When the blue “Summary” button is clicked, this pop-up appears, which displays the patient outcome summary statistics for each tissue type in a concise manner.

3.4. *Data Governance*

There are two sources of data used in the application. The first is patient data from the VA and CMS, accessed when a patient logs into their accounts in these electronic health records. The SMART COVID Navigator does not store any patient data. The data is accessed by the patient or physician under authorization by the patient. This data is used to identify patient risk factors when the physician/patient are using the application. When they are done using the application and log out, the patient data is deleted from the local machine which is being used to access the app. At no point is patient data stored in the main server running the application.

The second source of data are from the publicly available CORD-19 and RebootRx COVID-Cancer Datasets for use in the application. Given these are publicly available repositories of research papers, we do not implement any specific data governance with respect to this information.

4. Knowledge Base Analysis Results

In addition to the development of the SMART COVID Navigator, we performed analyses on the research studies in the CORD-19 and the RebootRx COVID-Cancer Project knowledge bases. Given the large number of studies both in the CORD-19 knowledge base and the COVID-Cancer datasets, it would be difficult for a clinician to quickly assess if a particular patient's disease condition and/or cancer type interacts significantly with COVID-19. For the CORD-19 knowledge base, we explored whether on average the studies found significant results for each of the 28 conditions examined, and for the RebootRx COVID-Cancer datasets, we explored the likelihood of mortality for a particular cancer type based on the aggregate data collected from the datasets.

4.1. *CORD-19 Knowledge Base Analysis Results*

First, we discuss the results from the CORD-19 knowledge base relating to risk factors. We found that out of the 816 studies examined in the risk factors section, 355 studies (43.5%) studied only how the particular disease condition affected the severity of COVID-19 progression, 328 studies (40.2%) examined only how the disease condition affected fatality, and 89 studies (10.9%) examined both severity and fatality. 44 studies (5.4%) had no entry for severity nor fatality.

We examined each of the 28 individual risk conditions to identify which of them have been found more frequently to be a significant factor in COVID-19 severity and fatality rates. This data is displayed in tables 3 and 4. Table 3 presents the data for each condition regarding severity, and table 4 presents the data regarding fatality.

Table 3 documents severity statistics; specifically, the number and proportion of studies that find if a given condition results in a statistically significant change in the likelihood of a patient having a severe COVID-19 infection. For example, out of the 98 studies looking at heart disease, 54 of them study severity and 42 (78%) found a significant result for severity. On the other hand, chronic liver disease seems to not have a significant impact on COVID-19 severity, as out of the 12 papers that study the condition, 6 study severity and only 1 (18%) found a significant correlation. The app alerts users if more than 50 percent of papers report a significant finding.

Table 3: Severity Statistics for All Risk Factors

Risk Factor	Total Papers	# Studying Severity	# With Significant Severity	Proportion of Severe Findings
Age	33	15	10	67%
Asthma	6	3	0	0%
Autoimmune disorders	3	1	0	0%
Cancer	26	16	8	50%
Cardio- and cerebrovascular disease	10	7	4	57%
Cerebrovascular disease	20	8	6	75%
Chronic digestive disorders	5	2	0	0%
Chronic kidney disease	38	19	10	53%
Chronic liver disease	12	6	1	17%
Chronic respiratory diseases	29	12	6	50%
COPD	41	23	18	78%
Dementia	5	2	1	50%

Diabetes	100	57	33	58%
Drinking	1	1	0	0%
Endocrine diseases	4	4	2	50%
Ethnicity: Hispanic vs. non-Hispanic	15	6	0	0%
Heart Disease	98	54	42	78%
Heart Failure	17	3	1	33%
Hypertension	100	60	35	58%
Immune system disorders	6	4	3	75%
Male gender	100	59	29	49%
Neurological disorders	6	3	1	33%
Overweight or obese	43	28	18	64%
Race: Asian vs. White	5	2	0	0%
Race: Black vs. White	21	7	2	29%
Race: Other vs. White	9	3	0	0%
Respiratory system diseases	7	6	5	83%
Smoking Status	56	33	16	48%

Table 4 is similar to table 3, except that instead of documenting findings about COVID-19 severity, it documents statistics regarding fatality. We can again look at some examples of medical conditions. As in the case of severity of COVID-19 progression, heart disease is a useful predictor of COVID-19 fatality; out of the 51 papers that measure fatality statistics for the condition, 43 of them (84%) found a significant correlation. Being a smoker, however, does not seem to have the same high degree of correlation; only 5 (19%) out of the 26 papers that measure fatality statistics for smoking status found a significant result. Similarly to severity, the app alerts users if more than 50 percent of papers report a significant finding for fatality.

Table 4: Fatality Statistics for All Risk Factors

Risk Factor	Total Papers	# Studying Fatality	# With Significant Fatality Value	Proportion of Fatal Findings
Age	33	19	16	84%
Asthma	6	4	2	50%
Autoimmune disorders	3	3	2	67%
Cancer	26	13	6	46%
Cardio- and cerebrovascular disease	10	3	2	67%
Cerebrovascular disease	20	14	10	71%
Chronic digestive disorders	5	2	0	0%
Chronic kidney disease	38	23	14	61%
Chronic liver disease	12	5	3	60%
Chronic respiratory diseases	29	19	13	68%
COPD	41	19	8	42%
Dementia	5	4	4	100%
Diabetes	100	46	26	57%
Drinking	1	0	0	N/A
Endocrine diseases	4	0	0	N/A
Ethnicity: Hispanic vs. non-Hispanic	15	11	6	55%

Heart Disease	98	51	43	84%
Heart Failure	17	14	9	64%
Hypertension	100	43	21	49%
Immune system disorders	6	4	4	100%
Male gender	100	39	20	51%
Neurological disorders	6	3	3	100%
Overweight or obese	43	20	12	60%
Race: Asian vs. White	5	4	3	75%
Race: Black vs. White	21	17	7	41%
Race: Other vs. White	9	9	3	33%
Respiratory system diseases	7	2	2	100%
Smoking Status	56	26	5	19%

There are conditions for which the papers tracked by CORD-19 do not offer a definitive answer. Cancer is one of these conditions; 50% of studies found a significant finding for severity, while 46% found a significant finding for fatality. A possible explanation for this inconclusive result is the existence of various types of cancers in different tissues of the body, resulting in a mixed result when grouping various cancer types together (as analyzed in section 4.2). For such cases, clinicians will need to access additional information from other papers.

As discussed before, we display the severity and fatality percentage summary statistics in the app for each condition, so clinicians can get a quick overview of the significance of that condition without having to review each study in detail. However, we caution that clinicians might want to review the studies in greater detail, depending on the medical condition of the concerned patient.

4.2. *Reboot: COVID-Cancer Project Knowledge Base Analysis Results*

In addition to the CORD-19 analysis, we conducted an analysis based on the papers in the Reboot: COVID-Cancer Project dataset relating to patient outcomes. This dataset stores the number of patients with a specific type of cancer and with COVID-19, the number of those patients that died, and the percent fatality found. The Navigator aggregates this data to offer the user summary statistics for each of the tissue types.

Table 5 documents the aggregate patient outcomes for each tissue, compiled by adding the number of patients with COVID-19 and number of deaths for all papers relating to the given tissue type. Some cancer tissue types in the table can be seen to be associated with a relatively high fatality rate. For instance, out of the 200 patients with thoracic cancer, 66 died (33%). Similarly, 152 out of the 504 COVID-19 patients with lymphoid cancer died (30%), and 67 of the 225 patients with lung cancer died (30%), indicating that cancers of these tissue types generally result in a higher fatality rate for COVID-19. On the other hand, 5 out of the 91 COVID-19 patients with breast cancer died (5%), indicating that breast cancer is associated to a lesser degree to COVID-19 fatality than some other cancer tissue types.

Table 5: Outcome Statistics for All Tissue Types

Tissue Type	**# Patients**	**# Deaths**	**% Fatality**
Bladder/Urinary Tract	1	0	0%
Bone and soft tissue	9	3	33%
Bowel	5	0	0%
Brain/CNS	2	0	0%
Breast	91	5	5%

Cervix	1	0	0%
Genitourinary	4	0	0%
Gynecological	180	20	11%
Head and neck	8	0	0%
Hematologic not specified	1893	492	26%
Kidney	18	2	11%
Liver	5	1	20%
Lung	225	67	30%
Lymphoid	504	152	30%
Myeloid	29	5	17%
Not specified	13265	2632	20%
Ovary/Fallopian Tube	2	0	0%
Pleura	1	0	0%
Prostate	124	28	23%
Skin	3	0	0%
Thoracic	200	66	33%
Thymus	1	0	0%

For some tissue types, the data available through the Reboot: COVID-Cancer Project does not contain an adequate number of patients to support clinical decisions. For instance, based on the fatality rate of 33%, it would seem that bone and soft tissue cancer is associated with a high risk of COVID-19 fatality. However, the data only contains 9 patients with bone and soft tissue cancer, so more information is needed to make clinical decisions for COVID-19 patients with bone and soft tissue cancer. Clinicians can access additional information from the full text of papers from the Navigator in order to make better decisions about patient treatment.

5. Discussion

5.1. *Views from Clinician Users*

We surveyed clinicians who tested the SMART COVID Navigator to elicit their assessment. Their feedback highlighted the following uses of the application in patient care and as an educational resource.

The clinicians noted that the app allows for a real-time assessment of co-morbidities that any given patient may have that could impact severity and fatality risk from COVID-19, and that it also allows for the assessment of patients in which multiple factors may be at play. The clinicians also noted that the app could be used at the point of care, and filled an unmet clinical need. They stated that the app made it easier to narrow down the literature and find out the relevant patient management-related answers quickly. They also pointed out that the fatality and severity rates displayed by the CORD-19 data could help them in triaging and stratifying patients in limited resource conditions, allowing them to decide whether aggressive treatment was warranted immediately or not. This would also help them involve palliative care early enough if needed. Finally, they commented that the app would help them in shared decision making with patients.

Apart from patient care, the clinicians felt that the SMART COVID Navigator could serve as an educational resource in teaching medical students, residents, journal clubs, and in perhaps even in continuing medical education for physicians. Some also felt that the app can help in making institutional guidelines.

The clinicians made suggestions for future work related to the SMART COVID Navigator. Real-time use of the application would be enhanced by creating a smartphone interface in addition to the current desktop-based web interface, considering that physicians often perform literature reviews on smartphones. While the app currently links to the Veterans Affairs and Center for Medicare and Medicaid Services, some clinicians felt that the application would benefit from being linked to other

EHRs such as EPIC. In fact, the same type of user interface and backend application programming interface (SMART-on-FHIR) can be leveraged and is supported by these as well. Another suggestion was to link the application to a patient's COVID-19 vaccination records to further improve the physician's ability to treat the patient.

5.2. *Additional Comments and Extensions*

SMART COVID Navigator was created in response to a growing need for precision medicine tools to assist doctors dealing with COVID-19 infected patients with risk factors shown to affect COVID-19 progression. The Navigator achieves this by connecting patient medical information – in the form of electronic health records – with datasets giving comprehensive information about patient outcomes.

We show in this paper how to create a clinical decision support system for physician to understand how a new disease (COVID-19) interacts with a broad set of patient risk factors, through the use of the CORD-19 dataset. In addition, we expand one existing patient condition – namely, cancer – and provide physicians information on how various types of cancers interact with COVID-19. This can be extended to other conditions, such as heart disease, since certain chronic conditions appear with a large number of variations in patients.

This app benefits from the creation of research study datasets like CORD-19 and the Reboot: COVID-Cancer Project, thereby showing the value of such collaborative efforts to collect current research findings in one place. Currently, these datasets do not provide API access; therefore, we have to periodically download these datasets for use in the app. In the future, we recommend that such datasets provide API access to enable real-time updating.

We note that in the CORD-19 dataset, some disease categories are examined by a limited number of papers; thus, the results may not be meaningful for our understanding of COVID-19's impact on patients with that risk factor. Similarly, certain cancer tissue types are not well represented in the Reboot: COVID-Cancer Project dataset. We expect that as research into COVID-19 progresses further, there will be better representation across disease categories and tissue types.

The ability for users to log into two FHIR-supported electronic health record systems is a useful feature of the application. In addition to providing physicians the latest information about patient's health records, it can also alert users to discrepancies between different databases, giving them an opportunity to correct their health records. The app is not limited to just the two electronic health record databases currently used (the Veterans Affairs and the Center for Medicare and Medicaid Services); more EHRs can be added to the app as long as they follow the FHIR format, allowing for even more expansion of patient record access.

A possible next step for the SMART COVID Navigator could be to better consolidate the research available in CORD-19 and the Reboot: COVID-Cancer Project using meta-analysis of the various studies to provide more information to the clinician [23]. For instance, more sophisticated weighting of the results in different papers could be implemented. Another extension would be to provide treatment information through the Navigator platform, depending upon the patient's health profile. In addition, advancements in artificial intelligence and machine learning technologies could play an important role in the improvement of the Navigator platform, as well as precision medicine as a whole, by offering predictions about disease progression based on past patient health information, thereby tailoring treatment for the individual and also allowing for increased efficiency in treatment [24]. This could take the form of creating patient risk scores, generated based on data from electronic health records.

6. Conclusion

Precision medicine has long been hailed as the next step in advancing patient care, with

institutions such as the White House creating initiatives to further precision medicine technology advancement [25]. The SMART COVID Navigator is a clinical decision support tool designed to allow physicians to provide precision medicine in the context of COVID-19. Using the app, clinicians can identify patient risk factors from multiple electronic health records and connect that to databases of COVID-19 research. Without a clinical decision support system supporting COVID-19 precision medicine, a clinician would need to examine multiple studies in order to fully understand disease progression and fatality outcomes given a particular disease that their patient has, which is costly in terms of time and effort. The number of studies examining particular disease conditions and their relationship to COVID-19 is growing at a rapid pace. The app simplifies this by providing summary statistics on the risk factors' effect on COVID-19 severity and fatality. While this application currently focuses on COVID-19, it can be a readily available platform for quickly expanding into any potential new diseases that emerge.

References

1. What is precision medicine? MedlinePlus. URL: https://medlineplus.gov/genetics/understanding/precisionmedicine/definition/ [accessed 2021-07-06]
2. What is an electronic health record (EHR)? HealthIT.gov. URL: https://www.healthit.gov/faq/what-electronic-health-record-ehr [accessed 2021-07-06]
3. Li Q., Guan X., Wu P., Wang X., Zhou L., Tong Y. *et al*. 2020 Early Transmission dynamics in Wuhan, China, of Novel Coronavirus–Infected pneumonia. *N. Engl. J. Med.***382**, 1199–1207.
4. M. Wadman, J. Couzin-Frankel, J. Kaiser, and C. Matacic, "How does coronavirus kill? clinicians trace a ferocious rampage through the body, from brain to toes," Science, doi:10.1126/science.abc3208.
5. Zaki, N., Alashwal, H., & Ibrahim, S. (2020). Association of hypertension, diabetes, stroke, cancer, kidney disease, and high-cholesterol with COVID-19 disease severity and fatality: A systematic review. *Diabetes & metabolic syndrome*, *14*(5), 1133–1142. Advance online publication. https://doi.org/10.1016/j.dsx.2020.07.005
6. McRae MP, Simmons GW, Christodoulides NJ, et al. Clinical decision support tool and rapid point-of-care platform for determining disease severity in patients with COVID-19. Lab Chip. 2020;20:2075-2085.
7. Wu, G., Yang, P., Xie, Y., Woodruff, H. C., Rao, X., Guiot, J., Frix, A. N., Louis, R., Moutschen, M., Li, J., Li, J., Yan, C., Du, D., Zhao, S., Ding, Y., Liu, B., Sun, W., Albarello, F., D'Abramo, A., Schininà, V., … Lambin, P. (2020). Development of a Clinical Decision Support System for Severity Risk Prediction and Triage of COVID-19 Patients at Hospital Admission: an International Multicenter Study. *The European respiratory journal*, 2001104. Advance online publication. https://doi.org/10.1183/13993003.01104-2020
8. Warner, J. L., Prasad, I., Bennett, M., Arniella, M., Beeghly-Fadiel, A., Mandl, K. D., & Alterovitz, G. (2018). SMART Cancer Navigator: A Framework for Implementing ASCO Workshop Recommendations to Enable Precision Cancer Medicine. JCO precision oncology, 2018, 10.1200/PO.17.00292. https://doi.org/10.1200/PO.17.00292
9. Li X and Warner JL (2020) A Review of Precision Oncology Knowledgebases for Determining the Clinical Actionability of Genetic Variants. Front. Cell Dev. Biol. 8:48. doi: 10.3389/fcell.2020.00048
10. Overby, C. L. & Tarczy-Hornoch, P. Personalized medicine: challenges and opportunities for translational bioinformatics. *Personalized Med.* **10**, 453–462 (2013).
11. Li, Z., Roberts, K., Jiang, X. & Long, Q. Distributed learning from multiple EHR databases: contextual embedding models for medical events. *J. Biomed. Inform.* **92**, 103138 (2019).
12. Alterovitz, G., Heale, B., Jones, J. et al. FHIR Genomics: enabling standardization for precision

medicine use cases. npj Genom. Med. 5, 13 (2020). https://doi.org/10.1038/s41525-020-0115-6
13. OAuth2. URL: https://oauth.net/2/ [accessed 2021-07-06]
14. OpenID Connect. URL: https://openid.net/connect/ [accessed 2021-07-06]
15. Kricka, L. J., Polevikov, S., Park, J. Y., Fortina, P., Bernardini, S., Satchkov, D., Kolesov, V., & Grishkov, M. (2020). Artificial Intelligence-Powered Search Tools and Resources in the Fight Against COVID-19. EJIFCC, 31(2), 106–116.
16. Semantic Scholar. URL: https://www.semanticscholar.org/ [accessed 2021-07-06]
17. Lu Wang, L., Lo, K., Chandrasekhar, Y., Reas, R., Yang, J., Eide, D., Funk, K., Kinney, R., Liu, Z., Merrill, W., Mooney, P., Murdick, D., Rishi, D., Sheehan, J., Shen, Z., Stilson, B., Wade, A. D., Wang, K., Wilhelm, C., Xie, B., … Kohlmeier, S. (2020). CORD-19: The Covid-19 Open Research Dataset. ArXiv, arXiv:2004.10706v2.
18. COVID-19 Open Research Dataset Challenge (CORD-19). URL: https://www.kaggle.com/allen-institute-for-ai/CORD-19-research-challenge [accessed 2021-07-06]
19. SNOMED International. URL: http://www.snomed.org/ [accessed 2021-07-06]
20. HL7 FHIR: ICD-10. URL: http://hl7.org/fhir/sid/icd-10 [accessed 2021-07-06]
21. Kuderer, N. M., et al. Clinical impact of COVID-19 on patients with cancer (CCC19): a cohort study. *Lancet* https://doi.org/10.1016/S0140-6736(20)31187-9 (2020).
22. *Reboot: COVID-Cancer Project.* Reboot Rx, Inc. URL: https://rebootrx.org/covid-cancer. [accessed 2021-07-06]
23. Zaki, N., Mohamed, E., Ibrahim, S., & Khan, G. (2020, January 01). The influence of comorbidity on the severity of COVID-19 disease: Systematic review and analysis. Retrieved August 16, 2020, from https://www.medrxiv.org/content/10.1101/2020.06.18.20134478v2
24. Schork N. J. (2019). Artificial Intelligence and Personalized Medicine. *Cancer treatment and research*, *178*, 265–283. https://doi.org/10.1007/978-3-030-16391-4_11
25. Collins, F. S., & Varmus, H. (2015). A New Initiative on Precision Medicine. New England Journal of Medicine, 372(9), 793–795. https://doi.org/10.1056/NEJMp1500523

Supplementary Files

Figures