Implementation of regional COVID-19 registry in Hormozgan (RCovidRH), Iran: Rationale and study protocol

Farid Khorrami1, Mehraban Shahi1, Nasrin DavariDolatabadi1, Nader Alishan Karami1, Mehdi HasaniAzad2, Fereshteh Jafariyan2, Abbas Sheikhtaheri*3

Received: 26 Apr 2020                   Published: 11 Aug 2020

Abstract

Background: The rapid outbreak of COVID-19 has resulted in a global pandemic in 2020. Information sources such as disease registries through accessing quality, valid, accurate, and timely data empower researchers and health authorities to study and develop appropriate actions. Our study describes the protocol for implementation of regional COVID-19 registry in Hormozgan province (RCovidRH).

Methods: We followed approved phases for the development of RCovidRH to cover the population in Hormozgan. Missioned to develop and implement the protocol, the registry’s steering committee was made up of 10 members from subject fields of the registry at the core and 5 subgroups. The main purpose of the registry is to provide a comprehensive information profile of demographic, clinical, laboratory, imaging, and treatment data of confirmed and probable COVID-19 patients in Hormozgan. The data is retrospectively and prospectively collected. Case report form (CRF) was mainly based on International Severe Acute Respiratory and Emerging Infection Consortium (ISARIC) CRF. A web-based, 2-language software was also developed to facilitate data collection and storage. Data analysis is to be conducted with collaboration of clinical physicians, data-mining specialists, and epidemiologists after reaching appropriate sample size.

Results: We included data related to demographic and identification, onset and admission, signs and symptoms at hospital admission, admission signs and symptoms, comorbidities, pathogen testing, assessment, laboratory, imaging, complications, treatment and medication, and outcomes. We found this registry was limited by incomplete clinical data for small fraction of outpatients, incomplete or inaccurate address by referred people due to fear of social rejection, delay in data entry at the facilities due to workload.

Conclusion: This registry via organizing clinical and epidemiological COVID-19 data increases the potentiality of joint studies. Recognition and coordination of a registry is highly important to solve its limitations to collect data. Other universities and provinces can apply our model to develop COVID-19 registries or data sets for this disease.

Keywords: COVID-19, Hormozgan, Iran, Registries, SARS-CoV-2, Study Protocol, COVID-19 registry, Data elements

Introduction

The coronavirus family contains many viruses found in nature, birds, and mammals (1-3). The virus was first recognized in the 1960s by a series of respiratory infections in children and adults (1-3). The new variant of this virus, severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), causes Coronavirus disease 2019 (COVID-19).

↑What is “already known” in this topic:
High quality data are needed to fight against COVID-19 pandemic.

→What this article adds:
This registry by proving accurate data about confirmed and probable COVID-19 cases enables us to better understand the epidemiology, prognosis, and clinical characteristics of the disease.

1 Health Information Technology, Faculty of Paramedicine, Hormozgan University of Medical Sciences, Bandar Abbas, Iran
2 Infectious and Tropical Diseases Research Center, Hormozgan Health Institute, Hormozgan University of Medical Sciences, Bandar Abbas, Iran
3 Health Management and Economics Research Center, Department of Health Information Management, School of Health Management and Information Sciences, Iran University of Medical Sciences, Tehran, Iran
The outbreak of COVID-19 originated in Wuhan, China, in early December 2019 and rapidly spread to different parts of the world within a month (1, 4). Since January 2020, the virus has been reported in most parts of China as well as 206 other countries (5). The COVID-19 pandemic is a global phenomenon and is estimated to continue. In other words, the geographical expansion of the virus is unrestricted, and a significant number of the world's population will be affected, and it has been announced as a public health emergency of international concern by the World Health Organization (WHO) (4-6). According to WHO report (May 18, 2020), more than 4.6 million people worldwide have already been infected (5) and the USA is ranked first with 1,432,265 infected cases (7). The disease quickly gained worldwide attention because of the growing number of new cases and deaths (4).

Unfortunately, Iran is one of the countries with the highest number of COVID-19 cases and is among high infected countries in the world as of May 18, 2020 (7). In Iran, 120,198 people have already been infected by COVID-19, which have led to 6988 deaths (7). Since February 20, 2020 to the date of submission of this paper, there has been an increasing trend of confirmed cases and deaths in Iran (7). Statistics regarding COVID-19 in Hormozgan province, south of Iran, show 1938 confirmed cases and 31 deaths (8).

In facing such pandemics, the Ministry of Health and Medical Education (MOHME) of the Islamic Republic of Iran recommends to execute control programs developed based on scientific evidence to prevent the loss of financial and human resources and to access quality, valid, accurate, and timely data. One of such programs is to establish disease registries (9). In fact, establishing COVID-19 registries is a solution to collect quality data related to those infected by COVID-19 in different geographical areas. A registry is an organized system that systematically collects clinical and nonclinical data using standard observational methods. In fact, registries combine data from all cases of a particular disease or all individuals with specified characteristics and are implemented in a specific population. Systematic data collection allows for the evaluation of specific outcomes, health-related characteristics, and outcomes according to epidemiological criteria (10-12). Given the rapid spread of COVID-19, there is an urgent need to identify specific patients’ characteristics and disease outcomes in different populations. Hence, it seems registries can play a critical role in providing quality data for conducting a wide range of research. Registries also follow other objectives, including identifying cases for clinical trials, and collecting timely and accurate data which in turn can assist researchers in various studies (10, 13). In other words, registries improve our knowledge regarding a disease or a health outcome (12). Additionally, joint studies using quality data at national and international levels are necessary. Therefore, standardization organizations at national and international levels should take the necessary measures to establish standards for data exchange among registries developed for this purpose. Thus, organizations such as WHO and Ministry of Health will be able to conduct complete and comprehensive programs for prevention, control, and treatment of the disease. Therefore, developing a registry based on a standard method and data set can also be used as a model for other registries. In this study, we aimed to describe our protocol for the development and implementation of regional COVID-19 registry in Hormozgan province (RCovidRH), Iran.

Methods

According to the standards and literature related to disease registries (14-17), several phases should be followed to develop a disease registry program: identifying stakeholders and developing a steering committee, purpose formulation, and definition of the target population; determining the method of case finding; identifying all reportable cases; data set formulation and developing the case report form (CRF); data gathering method; data storage method; data quality assessment protocol; and methods for data analysis and report production.

Steering committee

This regional and clinical disease registry was developed under the supervision of Hormozgan University of Medical Sciences (HUMS), the vice-chancellor for Research (Bandar Abbas, Iran) in 2020. In the beginning, we developed a steering committee, including 10 experts in related fields such as health information management, information technology, biostatistics and epidemiology, health care administration, virology, and infectious diseases. The committee was responsible to develop the registry protocol, formulate the purposes, secure the funds and operational technical staff, and supervise and evaluate the registry. We developed some operational subgroups, including IT subgroup with the responsibility of developing web-based software; data collection subgroup to find cases and gather and enter the data; data quality subgroup to evaluate the quality of data, provide feedback and improve quality; information management subgroup to review the literature to develop the dataset and CRF; and finally, data analysis subgroup to develop data analysis plan and to conduct data analysis (Fig. 1). Recently, we have made the decision to add biobank and patients’ follow-up subgroups.

http://mjiri.iums.ac.ir

![Fig. 1. The organizational structure of the RCovidRH registry](image-url)


**Purposes of the RCovidRH registry**

After several online meetings, the steering committee decided to develop this registry as a clinical-research registry with the main purpose of developing a comprehensive information profile of demographic, clinical, laboratory, imaging, and treatment data of confirmed and probable COVID-19 patients in Hormozgan province, Iran. The following specific purposes were also formulated:

1. To determine the frequency and demographic patterns of the confirmed and probable COVID-19 patients referred to health care facilities in the region;
2. To determine the proportion of death to incidence among the confirmed and probable COVID-19 patients referred to health care facilities in the region;
3. To determine the frequency of the confirmed and probable COVID-19 patients by gender, age, geographical area, and other demographic factors;
4. To determine the frequency of sign, symptoms, and complications of patients;
5. To determine the frequency of outcomes among confirmed and probable COVID-19 patients by gender, age, and geographical area;
6. To determine the frequency of outcomes based on treatment methods;
7. To determine the association between comorbidities among confirmed and probable COVID-19 patients;

**Study design and study population**

The RCovidRH registry was set up to record the data obtained from all patients with/suspected to have COVID-19 who referred to 20 specialized health care facilities affiliated to HUMS. The data are retrospectively and prospectively collected because the registry was launched just a month after the first reported case of the disease in Hormozgan. The population of the study is Hormozgan province with 1 750 627 inhabitants in 14 cities in the south of Iran with warm climate (Fig. 2). Moreover, data related to patients from other provinces referred to above facilities are also recorded in the registry.

**Case definition, inclusion and exclusion criteria**

In this study, the registry collects demographic and clinical data of all patients with/suspected to have COVID-19 who referred to health care facilities in different parts of the province before and after the launch of the project. According to the latest statistic by HUMS (May 18, 2020), 2948 patients have been admitted as acute respiratory patients in HUMS affiliated facilities so far. COVID-19 tests on 12 233 cases revealed 1938 positive cases (8). All of the patients’ data are retrospectively registered. In addition, new cases will be registered prospectively from the launch date of the registry. The data related to nonresidents of Hormozgan province will be separately analyzed.

**Data elements and case report form**

To develop the dataset, we initially searched the existing datasets and CRFs. We then reviewed and compared the CRFs suggested by the US Center for Disease Prevention and Control (CDC) (19), WHO (20), and International Severe Acute Respiratory and Emerging Infection Consortium (ISARIC) (21). Then, in the steering committee meetings, we finally selected the ISARIC CRF as the basis for the development of our CRF, mainly due to its comprehensiveness. We also added some questions extracted from other CRFs as well as some questions based on the experiences of the clinicians in our team encountering patients who referred to the HUMS facilities. Also, we developed a follow-up form based on the selected data items. Table 1 shows our complete dataset. According to the decision recently made by the steering committee, we are going to include patients’ serologic and biobank data to RCovidRH soon.

**Software development**

After modeling the program and being approved by the committee, the database (physical design) was developed in SQL Server Microsoft. The network-based system and user interface were designed in Microsoft Visual Studio.
COVID-19 registry in Hormozgan

Table 1. Data items included in the RCovidRH registry

<table>
<thead>
<tr>
<th>Data class</th>
<th>Data items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demographic and identification data</td>
<td>Name, Family name, Father name, National ID, Gender, Birth Date, Age, Phone number, Occupation, Education, Nationality, Ethnicity, Address, Case type (confirmed, probable), Employed as a healthcare worker, Employed in a microbiology laboratory, Having relatives with this disease and the relations, Being a passenger and from where, Pregnancy status (and Gestational week), Post-partum status (and pregnancy outcome, delivery date, Test for baby, Result of baby test, Method of baby test), Infant &lt; 1 year (Birth weight, Gestational outcome, Breastfeeding status)</td>
</tr>
<tr>
<td>Onset and admission</td>
<td>Onset date of first/earliest symptom, Admission date at this facility, Time of admission, Transfer from other facility (Name and Date of admission), Travel in the 14 days prior to first symptom onset (Date and Place), Contact with animals, raw meat or insect bites in the 14 days prior to symptom onset (Type of animal)</td>
</tr>
<tr>
<td>Signs and symptoms at hospital admission</td>
<td>Temperature, HR, RR, Systolic BP, Diastolic BP, Severe dehydration, Sternal capillary refill time &gt; 2 seconds, Oxygen saturation, On(Room air or Oxygen therapy)</td>
</tr>
<tr>
<td>Admission signs and symptoms</td>
<td>History of fever, Cough, sputum, bloody sputum/hemoptysis, Sore throat, Runny nose (Rhinorrhea), Ear pain, Wheezing, Chest pain, Muscle aches (Myalgia), Joint pain (Arthralgia), Fatigue / Malaise, Shortness of breath (Dyspnea), Lower chest wall in drawing, Headache, Altered consciousness/confusion, Seizures, Abdominal pain, Vomiting / Nausea, Diarrhea, Conjunctivitis, Sweating, Vertigo, Chills, Palpitation, Skin rash, Skin ulcers, Lymphadenopathy, Bleeding (Hemorrhage), Site of bleeding, Loss sense of smell and taste</td>
</tr>
<tr>
<td>Co-morbidities</td>
<td>Chronic cardiac disease (including congenital heart disease not hypertension), Hypertension, Chronic pulmonary disease (not asthma), Asthma (physician diagnosed), Chronic kidney disease, Liver disease (Moderate or severe or Mild), Chronic neurological disorder, Malignant neoplasm, Chronic hematologic disease, AIDS/HIV, Obesity, Diabetes with complications, Diabetes without complications, Rheumatologic disorder, Dementia, Malnutrition, Other Smoking</td>
</tr>
<tr>
<td>Pathogen testing</td>
<td>Influenza (Type), Coronavirus (Type), RSV, Adenovirus, Bacteria, Other infectious respiratory diagnosis (Type), Clinical pneumonia, Type of Biospecimen collected (e.g. nasal swap, sputum), Date, Laboratory test method (PCR, Culture, Other), Results, Pathogen tested/detected</td>
</tr>
<tr>
<td>Assessment (Daily)</td>
<td>Date, Admission to ICU, FiO2, SaO2, PaO2 at time of FiO2, PaO2 sample type, PO2 (from same blood gas record as PaO2), pH, HCO3, Base excess, AVPU Alert, AVPU Verbal, AVPU Pain, AVPU Unresponsive, Glasgow Coma Score, Richmond Agitation-Sedation Scale (RASS), Riker Sedation-Agitation Scale (SAS), Systolic Blood Pressure, Diastolic Blood Pressure, Mean Arterial Blood Pressure, Urine flow rate, Non-invasive ventilation (e.g. BIPAP, CPAP), Invasive ventilation, Extra corporeal life support (ECLS), High-flow nasal cannula oxygen therapy, Dialysis/Hemofiltration, Any vasopressor/inotropic support (If yes, Type), Neuromuscular blocking agents, Inhaled Nitric Oxide, Tracheostomy inserted, Prone positioning, Other intervention or procedure</td>
</tr>
<tr>
<td>Laboratory</td>
<td>Date, Laboratory name, Hemoglobin, WBC Count, Lymphocyte count, Neutrophil Count, Hematocrit, Platelets, APTT/APTR, PT, INR, ALT/SGPT, LDH, Total Bilirubin, AST/SGOT, Glucose, Blood Urea Nitrogen (BUN), Lactate, Creatinine, Sodium, Potassium, Mg, Troponin, Potassium, Procalcitonin, CRP, Blood Group, FBS, Troponin, IgG, IgM, ESR</td>
</tr>
<tr>
<td>Imaging</td>
<td>Chest X-Ray, Presence of infiltrates, CT Scan, Results of CT</td>
</tr>
<tr>
<td>Complications</td>
<td>Viral pneumonitis, Bacterial pneumonia, Acute Respiratory Distress Syndrome (and Severity), Pneumothorax, Pleural effusion, Cryptogenic organizing pneumonia (COP), Bronchiolitis, Meningitis / Encephalitis, Seizure, Stroke / Cerebrovascular accident, Congestive heart failure, Endocarditis / Myocarditis / Pericarditis, Cardiac arrhythmia, Cardiac ischemia, Cardiac arrest, Bacteremia, Coagulation disorder / Disseminated Intravascular Coagulation, Anemia, Rhabdomyolysis / Myositis, Acute renal injury / Acute renal failure, Gastrointestinal Hemorrhage, Pancreatitis, Liver dysfunction, Hyperglycemia, Hypoglycemia, Other</td>
</tr>
</tbody>
</table>

2019 using C#.Net with Net Core 2.2. Moreover, to enhance user experience in mobile devices with different operating systems, we used PWA feature. Software content was developed based on the confirmed dataset, and contains various forms based on the above-mentioned data classes. Different user access levels were also designed. Users are authorized to access different parts of the system based on their roles in the registry. The system is currently available at http://www.corona.hdpsco.ir (Fig. 3).
Data collection and quality assurance

Data related to cases before and after launching the project are being collected. To identify the retrospective patients, the data resources are (1) the medical records of patients discharged with diagnosis of COVID-19 or those suspected to have the disease (ICD-10 Code: U07.1 and U07.2), and (2) the laboratory PCR test report of outpatients issued by HUMS Virology Laboratory. The liaisons at health care facilities collect and scan paper-based medical records and dispatched them to the registry staff for abstracting and completing the CRFs. Then, data entry is performed by the staff. Data and coding quality are also important (22). Therefore, 2 nursing staff supervised by an infectious disease specialist is responsible to control the quality of the entered data. The feedback regarding data quality issues are referred to the registry staff for modifications and corrections. Also, validation rules have also been considered in the software to control full completion of essential data elements and to ensure the accuracy of the data items. Figure 4 depicts the process for retrospective data collection. To collect prospective data more conveniently, we are also collecting related data right after admission and during treatment via a CRF attached to patients' medical records to be completed by infectious diseases specialists.

Implementation

To implement the registry system, an executive manager was appointed as the supervisor of data collection, data entry, and data quality subgroups. The manager was introduced to all health facilities affiliated to HUMS, which admit patients with respiratory conditions. The facilities introduced their liaisons to scan and send the medical records of potential cases. The liaisons were mostly from health information management departments at the facilities. The executive manager and the staff at the steering committee follow up by sending the medical records by liaisons. After receiving the records, 2 nursing staff extract data from paper medical records, summarize them in CRF, and enter data in the registry system. In the next step, the nursing staffs supervised by an infectious diseases specialist check the data for completeness, comprehensiveness, and accuracy. If the records are complete, the process is over. Otherwise, they follow up the case via the related liaisons to modify the incomplete or inaccurate data items.

Table 1. Ctd

<table>
<thead>
<tr>
<th>Data class</th>
<th>Data items</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Treatment and medication</strong></td>
<td>ICU admission, total ICU duration, date of ICU admission and discharge, Oxygen therapy, Non-invasive ventilation, Invasive ventilation, Invasive ventilation total duration, Prone ventilation, Inhaled Nitric Oxide, Inserted Tracheostomy, Extracorporeal support, Renal replacement therapy (RRT) or dialysis, Inotropes/vasopressors (start and end date), Other intervention or procedure, Antiviral agent (Type of medication), Antibiotic (Type of medication), Corticosteroid (Route, type and dose), Antifungal agent, Others</td>
</tr>
<tr>
<td><strong>Outcome</strong></td>
<td>Date, Outcome type (Discharged alive, Hospitalization, Transfer to other facility, Death, Palliative discharge), Ability to self-care at discharge, Post-discharge treatment, Oxygen therapy, Dialysis/renal treatment, Other intervention or procedure, Transferred facility name</td>
</tr>
</tbody>
</table>
Data analysis
The registry committee will perform data analysis with collaboration of clinical physicians, data-mining specialists, and epidemiologists after reaching appropriate sample size. We will conduct appropriate descriptive and inferential data analysis to achieve the registry purposes.

Ethical statements
The proposal for this study was approved by the ethics committee of Hormozgan University of Medical Sciences with ethical code HUMS.REC.1398.482.

Discussion
In this article, we described the rationale and the fundamental characteristics of RCovidRH. Given the prevalence of this disease worldwide and its growing trend, development and implementation of COVID-19 registries in different regions is essential. The purpose of this registry was to
collect demographic and clinical data to support epidemiological studies and to explore the unknown relationships between different variables of this disease in Hormozgan province. Since the commencement of the project (March 25, 2020), we managed to collect and record approximately 1400 positive inpatient and outpatient cases and more than 560 suspected cases from 14 different health care facilities affiliated to HUMS.

CRF data elements initially developed based on ISARIC (Version 1.3, 24 Feb) were updated after an expert panel meeting of the steering committee members. According to the panel, the modifications made to the CRF were as follow:

1. Adding patient’s “place of residence” to the demographic and identification section because Hormozgan province is a tourist attraction region during the Persian new year holidays (Nowruz, the Persian New Year).
3. Adding “hypertension” to the comorbidities section due to the request of cardiologists.
4. Adding “blood group” for all patients; adding “troponin” and “magnesium” for patients with hypertension; and “FBS” for diabetic patients to the laboratory section.
5. Adding “performing CT-SCAN”, “PACS No.” and “CT-SCAN results” to the imaging section.
6. Adding ICD-10 codes for probable and confirmed cases based on WHO and MOHME guidelines.

Designing the registry application based on international standards allows for data exchange with other data collecting applications of the disease designed based on the standards. Since we have developed our registry based on a standard method suggested by literature and that we have adopted our data set based on international standards, our solution could easily be customized for all other medical universities and provinces nationwide. The features of our registry allows for centrally host other COVID-19 clinical registries and other data related to COVID-19 in Iran.

The limitations which have already been experienced since the beginning of the implementation phase are as follow:

1. In the first month of the registry, we were unable to collect data related to outpatients suspected to have COVID-19. However, following organizational coordination and recognition of our registry in the university, we are now collecting data related to outpatient cases. However, because of the time limitation for these patients, collecting complete clinical data is difficult. Following organizational coordination and obtaining permission to contact patients, we are going to organize a follow-up (both in person and telephone) team to contact patients for completing the missing data. Furthermore, we are planning to add biobanks to our registry. Therefore, all patients will be called to return for collecting specimens. This will be another opportunity to complete our data regarding these patients.
2. Due to fear of social rejection some patients failed to provide accurate address. For example, some passengers did not provide exact address. Therefore, their data will be excluded from geographical analysis.
3. It was not possible to collect data at bedside due to the workload of the health care professionals at the facilities. Given the rapid spread of the disease, data entry was slightly delayed. We have recently coordinated with some hospital managers to add our CRF forms to hospital routines. We hope this intervention will increase the timeliness of registration and also data quality.

Conclusion
In this registry, we included a standard data set developed by international organizations. The data classes include demographic and identification data, onset and admission, signs and symptoms at hospital admission, admission signs and symptoms, comorbidities, pathogen testing, daily assessment, laboratory, imaging, complications, treatment and medication, and outcome. Being launched according to the data elements of international standards, this registry provides the possibility of data exchange. Moreover, organizing clinical and epidemiological data, the registry will allow for responding clinical and epidemiological questions. Our registry was limited by incomplete clinical data of a small fraction of outpatients, incomplete demographic data (e.g. address) by referred people due to fear of social rejection, and delay in data entry at the facilities due to workload. We found that organizational coordination is highly important to run a disease registry and improve its implication. After 2 months, our registry is now very well recognized in the university, even at the level of top managers and other clinical departments. We did some predictions based on these data for decision-makers. Some departments, e.g. genetic department, are going to join us to develop a biobank and also add genetics and biobank data to our registry. Following this coordination and recognition, we could solve some of our limitations. For example, we were able to obtain the permission to contact patients, add our CRFs to hospital routines, and collect outpatients’ data.

Acknowledgements
This registry was financially supported by the grant provided by Hormozgan University of Medical Sciences. (Main project number at HUMS-Vice Chancellor for Research and Technology: 980464). We would like to express our gratitude to Dr Teamur Aghamolaei (the HUMS Vice-chancellor for Research and Technology) and all the staff of Health Information Management Departments at health care facilities affiliated to HUMS who supported us to collect data related to COVID-19 patients.

Conflict of Interests
The authors declare that they have no competing interests.

References
COVID-19 registry in Hormozgan