



Design and development of a web-based registry for Coronavirus (COVID-19) disease

Hadi Kazemi-Arpanahi¹, Khadijeh Moulaei², Mostafa Shanbehzadeh*³

Received: 15 Apr 2020

Published: 25 Jun 2020

Abstract

Background: The 2019 coronavirus (COVID-19) is a highly contagious disease associated with a high morbidity and mortality worldwide. The accumulation of data through a prospective clinical registry enables public health authorities to make informed decisions based on real evidence obtained from surveillance of COVID-19. This registry is also fundamental to providing robust infrastructure for future research surveys. The purpose of this study was to design a registry and its minimum data set (MDS), as a valid and reliable data source for reporting and benchmarking COVID-19.

Methods: This cross sectional and descriptive study provides a template for the required MDS to be included in COVID-19 registry. This was done by an extensive literature review and 2 round Delphi survey to validate the content, which resulted in a web-based registry created by Visual Studio 2019 and a database designed by Structured Query Language (SQL).

Results: The MDS of COVID-19 registry was categorized into the administrative part with 3 sections, including 30 data elements, and the clinical part with 4 sections, including 26 data elements. Furthermore, a web-based registry with modular and layered architecture was designed based on final data classes and elements.

Conclusion: To the best of our knowledge, COVID-19 registry is the first designed instrument from information management perspectives in Iran and can become a homogenous and reliable infrastructure for collecting data on COVID-19. We hope this approach will facilitate epidemiological surveys and support policymakers to better plan for monitoring patients with COVID-19.

Keywords: Minimum data set, MDS, Registry system, COVID-19, Coronavirus

Conflicts of Interest: None declared

Funding: Ilam University of Medical Sciences

*This work has been published under CC BY-NC-SA 1.0 license.

Copyright© Iran University of Medical Sciences

Cite this article as: Kazemi-Arpanahi H, Moulaei Kh, Shanbehzadeh M. Design and development of a web-based registry for Coronavirus (COVID-19) disease. *Med J Islam Repub Iran.* 2020 (25 Jun);34:68. <https://doi.org/10.34171/mjiri.34.68>

Introduction

In December 2019, a series of cases of pneumonia with unknown etiology occurred in Wuhan, Hubei Province, China. On January 7, 2020, the novel coronavirus (COVID-19), previously known as severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2 or 2019-nCoV), was identified as the causative organism (1, 2). It is classified as a type of RNA virus that belongs to the family of coronaviruses, which primarily leads to a respiratory sys-

tem infection (3). COVID-19 is highly contagious that rapidly spread to other countries. The World Health Organization (WHO) has recently declared the COVID-19 a public health emergency (4).

Given the significant burdens associated with COVID-19, decision was made to adopt information technology and data infrastructures to bolster efficient research, surveillance, and treatment of this emerging outbreak. Clini-

Corresponding author: Dr Mostafa Shanbehzadeh, shanbehzadeh-m@medilam.ac.ir

¹ Department of Health Information Technology, Abadan Faculty of Medical Sciences, Abadan, Iran

² Department of Health Information Management, School of Health Management and Information Sciences, Kerman University of Medical Sciences, Kerman, Iran

³ Department of Health Information Technology, School of Paramedical, Ilam University of Medical Sciences, Ilam, Iran

↑What is “already known” in this topic:

Disease registries are collections of secondary data related to patients with a specific diagnosis, condition, or procedure, and they play an important role in clinical and managerial decisions.

→What this article adds:

COVID-19 registry facilitates studying the real clinical practice, capturing quality metrics, monitoring the disease and healthcare delivery patterns, and tracing clinical outcomes.

cal registry is one of such information platforms that standardize the collection of highly generalizable data and reinforce research infrastructure (5, 6). Clinical registries have a great potential for epidemiological surveillance, evaluating health-care delivery patterns, tracking clinical outcomes, describing disease natural progression, evidence-based therapy, and comparing the effectiveness of different interventions and post marketing drug surveillance. Moreover, clinical registries allow the study of the designed parameters and recruitment of participants for clinical trials (7-10). The COVID-19 registry will serve as a data source to standardize collection of comprehensive data related to many unclear aspects of COVID-19, such as transmission patterns, severity, clinical phenotype, prognostic factors, therapeutics' plans effectiveness and complications, survival estimation, incidence and prevalence of disease across country, and thereby allowing collaboration on research and surveillance of COVID-19.

However, despite the advantages afforded by clinical registries, some considerations need to be addressed from a data management perspective: the design of an effective data capture system and determination of the required data elements and validity of their corresponding values (11-13). Therefore, in this study, the required data elements for COVID-19 were defined and a clinical registry platform that met these requirements was designed.

Methods

This was a cross-sectional and descriptive study in 2020 that was conducted in two phases: in the first one, the aim was to identify required data elements and validated data capture template to be included in the COVID-19 registry, and the second one was designing a registry system for COVID-19 on the web platform.

COVID-19 registry data element determination

Literature review

First, an extensive literature review to identify the COVID-19 MDS data elements was performed. In the first step to retrieve related resources, the Web of Science, ScienceDirect, Embase, Scopus, Elsevier, Cochran, PubMed and Google Scholar were reviewed, and the follow-

ing search terms were used: (designed using English MeSH keywords and Emtree terms): "COVID-19" or "Novel coronavirus 2019" or "2019 nCoV" and "clinical characteristics" or "para clinical characteristics" or "epidemiological characteristics". After selecting the advance search interface in the mentioned databases using title, title/abstract, title/abstract/keyword and topic fields, and setting up Boolean operators (AND, OR) and implementing the input and output criteria (full text English articles from Dec 2019- Mar 2020), 18 articles were included in the study (3, 14-30). Data were extracted from the related retrieved resources and entered into the checklist with 2 administrative and clinical sections.

Questionnaire development

A questionnaire was developed using the data elements of the checklist and included 5 columns: "very important", "important", "neutral", "low important", and "very important" for each data item (eg, patient name, visit number, vital sign, exposure and etc.). To add necessary data elements by experts, a blank row was provided at the end of the questionnaire. The content validity of the questionnaire was assessed by an expert panel, including 2 infectious specialists and 3 health information management (HIM) experts. Also, test-retest was used to evaluate the reliability of the questionnaire.

Delphi phase

The initial MDS content was validated by Delphi technique using 2 rounds by a group of multidisciplinary experts working in hospitals affiliated to Ilam University of Medical Sciences (west of Iran). Table 1 shows the demographic characteristics of these experts. The experts were asked to review the initial data list to score each item according to their importance perceived by them based on a 5-point Likert scale, ranging from 1 to 5, where 1 indicated "not important" for inclusion and 5 indicated "highly important" for inclusion.

Agreement was reached for data elements based on experts' agreement level. After initial ranking, items with less than 50% agreement were deleted, those with more than 75% agreement excluded from the second round, and those with 50% to 75% agreement were surveyed in the second round. The checklists were individually presented

Table 1. Demographic characteristics of Delphi participants

| Variables | Frequency | Percentage |
|-------------------------|-----------|------------|
| Specialty | | |
| Infectious disease | 9 | 39.12 |
| Internal medicine | 8 | 34.79 |
| Radiologist | 6 | 26.09 |
| Gender | | |
| Female | 8 | 34.79 |
| Male | 15 | 65.21 |
| Age (years) | | |
| 30-40 | 9 | 39.12 |
| 40-50 | 6 | 26.09 |
| 50-60 | 6 | 26.09 |
| >60 | 2 | 8.7 |
| Work experience (years) | | |
| <10 | 6 | 26.09 |
| 10-20 | 11 | 47.82 |
| 20-30 | 5 | 21.74 |
| >30 | 1 | 4.35 |
| Total | 23 | 100 |

to the experts who were blind to the scores of other experts, and if there was 75% consensus over a subject, it was included into the final MDS.

COVID-19 registry software development tools

We used Visual Studio 2019 to design COVID-19 web-based registry because of its numerous benefits (eg, cost-effectiveness, scalability and accessibility, user friendliness, fast and convenience, custom search, improved intelligence, clipboard and refactoring attributes) (31, 32). The proposed system was implemented with cascading style sheets (CSS) technology as a web-based program. CSS, along with the Hypertext Markup Language (HTML), was used to describe the presentation of documents and set the document syntax, layout, display format, and visual effects (eg, font type, color, spacing, and sizes). The code was written in Java script language for designing the website. Finally, Structured Query Language (SQL) was used to create the relational database (RDB). SQL provides efficient and systematic storage of data with high performance, availability, scalability, flexibility, management,

and security (33).

Results

The results of this study are divided into 3 phases:

Determining the proposed MDS for COVID-19

The proposed COVID-19 MDS was divided into the nonclinical section with 4 data classes, including 43 data elements, and the clinical data category with 4 data classes, including 44 data items. The nonclinical section includes sociodemographic, identification number, and patient disposition classes, and the clinical category includes diagnostic, exposure, physical examination, and medical / diagnostic procedure.

Determining final minimum data Set for COVID-19

The potential participants who determined the final data elements of the MDS of the COVID-19 registry were 25 medical specialists. However, 2 specialists did not participate in the study. Table 1 shows the demographic characteristics of the experts. The results of the 2 Delphi rounds

Table 2. Examples of nonclinical and clinical data classes for COVID-19 MDS

| Data classes | Total number of elements | First round of Delphi | | | Second round of Delphi | | | Final |
|----------------------------------|--------------------------|-----------------------|--------|-------|------------------------|--------|-------|-------|
| | | < 50% | 50-75% | 75% < | < 50% | 50-75% | 75% < | |
| Administrative data category | | | | | | | | |
| Sociodemographic characteristics | 18 | 3 | 10 | 5 | 1 | 0 | 4 | 14 |
| Identification | 10 | 3 | 4 | 3 | 3 | 0 | 1 | 4 |
| Patient disposition | 15 | 2 | 10 | 3 | 1 | 0 | 2 | 12 |
| Clinical data category | | | | | | | | |
| Diagnostic | 15 | 3 | 9 | 3 | 2 | 0 | 1 | 10 |
| Exposure | 8 | 3 | 2 | 3 | 2 | 0 | 1 | 3 |
| Physical examination | 11 | 3 | 4 | 4 | 2 | 0 | 2 | 6 |
| Medical procedures | 10 | 1 | 6 | 3 | 2 | 0 | 1 | 7 |
| Total | 87 | 18 | 45 | 24 | 12 | 0 | 12 | 56 |

Table 3. Weighting of data items after the second round of Delphi

| Data classes | No | Data elements | Mean | Percentage | Final decision |
|--------------------|----|-------------------------|------|------------|----------------|
| Sociodemographic | 1 | Patient's name | 4.1 | 82 | Kept |
| | 2 | Father's name | 3.85 | 77 | Kept |
| | 3 | Spouse / partner's name | 2.1 | 42 | Removed |
| | 4 | Age (in years) | 4.15 | 83 | Kept |
| | 5 | Sex | 4.5 | 90 | Kept |
| | 6 | Date of birth | 3.95 | 79 | Kept |
| | 7 | Place of birth | 4.25 | 85 | Kept |
| | 8 | Marital status | 4.09 | 81.8 | Kept |
| | 9 | Income | 2.3 | 46 | Removed |
| | 10 | Religion | 2.5 | 50 | Removed |
| | 11 | Employment status | 3.95 | 79 | Kept |
| | 12 | Occupation | 4.01 | 80.2 | Kept |
| | 13 | Educational level | 4.12 | 82.4 | Kept |
| | 14 | Race/ nationality | 3.55 | 77 | Kept |
| | 15 | Home address | 4.05 | 81 | Kept |
| | 16 | Postal / zip code | 3.93 | 78.6 | Kept |
| | 17 | Phone number | 3.88 | 77.6 | Kept |
| | 18 | Fax no | 1.8 | 36 | Removed |
| Identified numbers | 19 | National ID | 4.2 | 84 | Kept |
| | 20 | Visit number | 3.98 | 79.6 | Kept |
| | 21 | Medical record number | 2.5 | 50 | Removed |
| | 22 | Social security number | 3.9 | 78 | Kept |
| | 23 | Physician ID | 3.96 | 79.2 | Kept |
| | 24 | Specimen ID | 2.8 | 56 | Removed |
| | 25 | Hospital ID | 2.2 | 44 | Removed |
| | 26 | Report ID | 1.8 | 36 | Removed |
| | 27 | Insurance ID | 1.88 | 37.6 | Removed |
| | 28 | Family ID | 1.6 | 32 | Removed |

are presented in Tables 2 and 3.

The experts participated in 2 rounds by completing the questionnaire.

At the end of the first Delphi round, 18 data elements were deleted (< 50%), 45 moved to the next round (50%-75%), and 24 marked as definitive (75% <). In addition,

Table 3. Ctd

| Data classes | No | Data elements | Mean | Percentage | Final decision |
|---------------------|-------------------|----------------------------|------|------------|----------------|
| Patient disposition | 29 | Admission date | 4.35 | 87 | Kept |
| | 30 | Reason for admission | 4.20 | 84 | Kept |
| | 31 | Type of admission | 4.09 | 81.8 | Kept |
| | 32 | Readmission | 4.22 | 84.4 | Kept |
| | 33 | Length of stay | 4.1 | 82 | Kept |
| | 34 | Discharge date | 4.3 | 86 | Kept |
| | 35 | Discharge status | 4.3 | 86 | Kept |
| | 36 | Underlying cause of death | 4.1 | 82 | Kept |
| | 37 | Date of death | 4.05 | 81 | Kept |
| | 38 | Discharge location | 3.8 | 76 | Kept |
| | 39 | Discharge recommendations | 2.4 | 48 | Removed |
| | 40 | Discharge/ referral date | 2.15 | 43 | Removed |
| | 41 | Discharge /referral type | 1.95 | 39 | Removed |
| | 42 | Discharge Prescribed drugs | 4.1 | 82 | Kept |
| 43 | Date of follow-up | 4.23 | 84.6 | Kept | |

Table 3. Ctd

| Data Classes | No | Data elements | Mean | Percentage | Final Decision |
|----------------------|----|--------------------------------------|------|------------|----------------|
| Diagnostic | 1 | Disease history | 4.8 | 96 | Kept |
| | 2 | Comorbidity | 4.6 | 92 | Kept |
| | 3 | Family history | 2.8 | 56 | Removed |
| | 4 | Disease status | 4.2 | 84 | Kept |
| | 5 | Disease severity status | 4.25 | 85 | Kept |
| | 6 | Mental condition | 2.2 | 44 | Removed |
| | 7 | Case classification | 4 | 80 | Kept |
| | 8 | Vital sign | 2.4 | 48 | Kept |
| | 9 | Sing and symptoms | 4.65 | 93 | Kept |
| | 10 | symptoms types (if symptomatic) | 4.6 | 92 | Kept |
| | 11 | Symptom onset date | 3.01 | 60.2 | Removed |
| | 12 | Chief complaint | 3.03 | 60.6 | Removed |
| | 13 | Days from exposure to symptom onset | 3.85 | 77 | Kept |
| | 14 | Time between diagnosis and treatment | 3.93 | 78.6 | Kept |
| | 15 | Date of diagnosis | 4.45 | 89 | Kept |
| Exposure | 16 | Exposed to high risk agent | 4.1 | 82 | Kept |
| | 17 | Exposure type | 2.85 | 57 | Removed |
| | 18 | Cause of exposure | 2.3 | 46 | Removed |
| | 19 | Exposure history | 4.78 | 95.6 | Kept |
| | 20 | Activity on exposure | 1.07 | 21.4 | Removed |
| | 21 | Location of exposure | 2.63 | 52.6 | Removed |
| | 22 | Number of exposures | 2.1 | 42 | Removed |
| | 23 | Date of exposure | 4.5 | 90 | Kept |
| Physical Examination | 24 | Respiratory rate: per minute | 4.4 | 85 | Kept |
| | 25 | Pulse | 2.6 | 52 | Removed |
| | 26 | Waist circumference | 2.7 | 54 | Removed |
| | 27 | Temperature: °C | 4.5 | 90 | Kept |
| | 28 | Brachial Index | 2.4 | 48 | Removed |
| | 29 | Blood group | 3.90 | 78 | Kept |
| | 30 | Body Mass Index | 3.05 | 61 | Removed |
| | 31 | Blood pressure: mmHg | 4 | 80 | Kept |
| | 32 | Lung examination | 4.15 | 83 | Kept |
| | 33 | Heart rate: bit per minute | 3.9 | 78 | Kept |
| | 35 | Weight /height | 1.8 | | Removed |
| Procedures | 36 | Quarantine / isolation | 2.1 | 42 | Removed |
| | 37 | Oxygen support | 4.45 | 89 | Kept |
| | 38 | Immunization/ vaccination | 2.3 | 46 | Removed |
| | 39 | Radiology | 4.8 | 96 | Kept |
| | 40 | CT features | 4.95 | 99 | Kept |
| | 41 | Lung segment involvement | 4.5 | 90 | Kept |
| | 42 | Prescription / medication | 3.2 | 64 | Removed |
| | 43 | LAB test name | 4.9 | 98 | Kept |
| | 44 | Test result | 4.9 | 98 | Kept |
| | 44 | Test time | 3.89 | 77.8 | Kept |

no new data elements were suggested by the experts. After the second round, in general, 13 data elements for the nonclinical and 18 elements for the clinical category were excluded from the report template. Therefore, the experts agreed on 30 data elements from 43 data elements of the nonclinical category. The second category was the clinical data involving 4 data classes with 26 data elements. The ultimate data elements for the nonclinical and clinical categories were 30 and 26, respectively (Table 2). The results of weighing of data elements after the second round of Delphi are displayed in Table 3.

In Table 4, data classes, elements, and their formats and standard contents (recording template) were defined for 2 nonclinical and clinical data categories.

The COVID-19 registry framework

In the development of the software, our focus was on accessibility and user-friendliness of the system to expe-

dite reporting time. Our designed system uses an advanced search capability to enable custom search, contact to site administrator, provide useful news about disease (prevention, self-care, treatment information, etc.), rendering daily statistics and multimedia instructions. Access to the registry is provided to registered members on the system home page (user name & password boxes)- Each user has a unique identification password and username to log into the system. Figures 1 and 2 display the designed web-based registry screen of COVID-19.

Discussion

The lesson learned from previous global pandemics and the widespread prevalence of zoonotic viral diseases (eg, SARS and MERS), highlights the importance of patient registries in the field of new emerging outbreaks such as COVID-19 (34, 35). In this regard, for proper implementation of a public health surveillance system (PHSs), clini-

Table 4. Nonclinical and clinical MDS description for COVID-19 registry system

| Administrative data category | | | Clinical data category | | |
|------------------------------|---|-------------|-------------------------------------|---|-------------|
| Data elements | Content definition | Data Format | Data elements | Content definition | Data Format |
| Sociodemographic | | | Diagnostic | | |
| Patient's name | First / middle / last name | String | Disease history | Free text | String |
| Father's name | First / middle / last name | String | Comorbidity | Free text | String |
| Age (in years) | *Infant: x<1y, *child: 1y<x<5y, *teenage: 5y<x<17y, *young: 17y<x<34y, *middle age: 34y<x<65y, *aged: x>65y | Categorical | Disease status | *Active, *inactive, *recovered | Categorical |
| | | | Disease severity status | *General, *severe, *critical | Categorical |
| Sex | *M *F | Categorical | Case classification | *Final, *suspicious *probable | Categorical |
| Date of birth | yyyy /mm/ dd | Date | Sing and symptoms | *Symptomatic, *a symptomatic | Binary |
| Place of birth | Geographical location: province, city, village | Categorical | symptoms types (if symptomatic) | Free text | String |
| Marital status | *Single *married *widowed, *other | Categorical | Days from exposure to symptom onset | Number of days | Integer |
| Employment status | *Unemployed, *employed, *retired, *student, *other | Categorical | Time between diagnosis to treatment | number of days | Integer |
| Occupation | Free text | String | Date of diagnosis | yyyy /mm/ dd | Date |
| | | | Exposure | | |
| Race/ nationality | Iranian: *Persian, *Kurdish, *Turkish, *other | Categorical | Exposed to high risk agent | *Yes, *no, *unknown | Categorical |
| Educational level | *Illiterate, *less than high school diploma, *bachelor, *master of science or above, *unspecified | Categorical | Exposure history | *Person-to-person *Animal-to-person *Contact with contaminated surfaces *Food/water born *Other, *unknown | Categorical |
| | | | Date of exposure | yyyy /mm/ dd | Date |
| | | | Physical Examination | | |
| Home address | Province-city-street-alley-house no | String | Respiratory rate: per minute | * ≤ 24 breaths per min * >24 breaths per min | Categorical |
| Postal / zip code | Ten digits with dash | Integer | Temperature: °C | * <37.3, *37.3 – 38, *38.1 – 39, * >39.0 | Categorical |
| Phone number | Ten digits with +98 Identifier | Integer | Heart rate: bit per minute | * <60, *between 60-100, * >100, *unknown | Categorical |
| National ID | Validated numerical range | Integer | Blood group | RH positive: A, B, AB, O RH negative: A, B, AB, O | Categorical |
| Visit number | Validated numerical range | Integer | Blood pressure: mmHg | * <120, *between 120-129, *between 130-139, * >140, *unknown | Categorical |
| Social security number | Validated numerical range | Integer | Lung examination | *Clear/normal, *rales, *decreased breath sounds, *rhonchi, *wheezing | Categorical |
| Physician ID | Validated numerical range | Integer | | | |

Table 4. Ctd

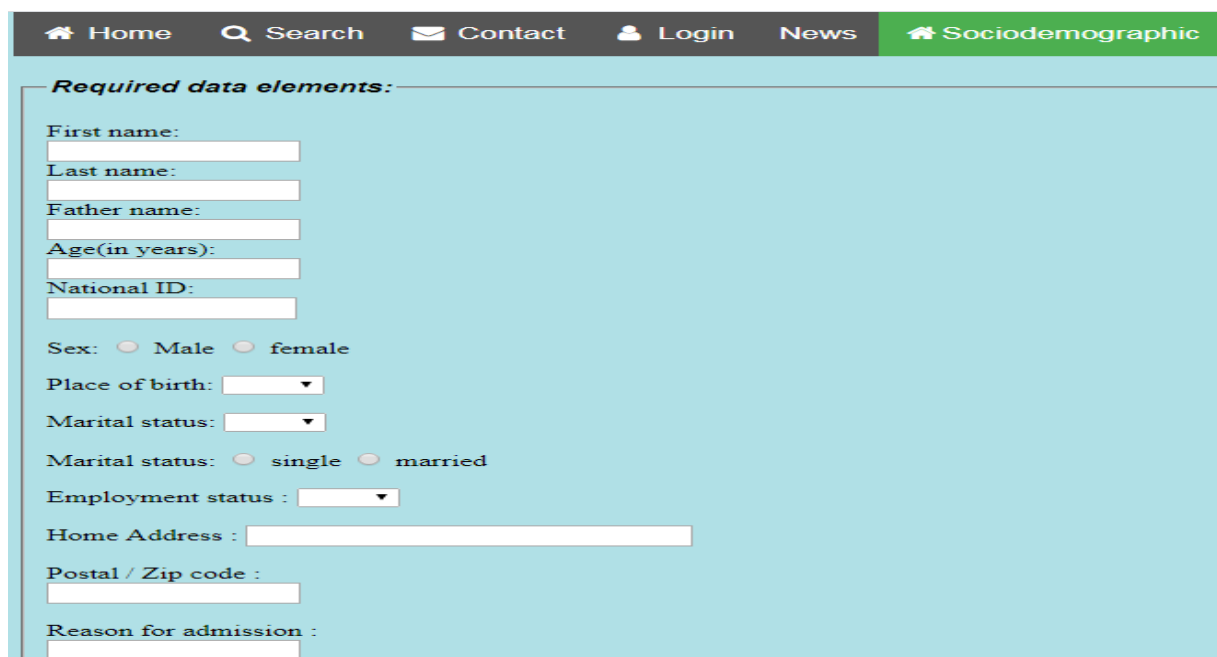
| Administrative data category | | | Clinical data category | | |
|------------------------------|--|--------------|--------------------------|---|-------------|
| Data elements | Content definition | Data Format | Data elements | Content definition | Data Format |
| | Patient Disposition | | | Medical / Diagnostic Procedures | |
| Admission date | yyyy /mm/ dd | Date | Oxygen support | *Non-invasive mechanical ventilator *invasive mechanical ventilator *Extracorporeal Membrane Oxygenation (ECMO) *other | Categorical |
| Reason for admission | Unstructured free text | String | | | |
| Type of admission | *Urgent, *programmed, *unknown | Categorical | Radiology | *Bilateral chest CT-scan *Unilateral chest CT-scan | Categorical |
| Readmission | *yes *no | Binary | | | |
| Length of stay | *≤7day, *>7 day | Force Choice | CT Features | *Ground-Glass Opacity (GGO) | Categorical |
| Discharge date | yyyy /mm/ dd | Date | | | |
| Discharge status | *Deceased, * full recovery, * partial recovery, * other | Categorical | | *Consolidation interlobular thickening *Crazy paving pattern *Bronchial wall thickening *Lymph adenopathy *pericardial effusion *pneumonia *pneumothorax *plural effusion *interstitial abnormalities *other | |
| Underlying cause of death | *Related to current disease *unrelated to current disease *not applicable * unknown | Categorical | | | |
| Date of death | *yyyy /mm/ dd, | Date | Lung segment involvement | *unilateral *bilateral | Binary |
| Discharge location | *Deceased * home *other care facilities * unknown | Categorical | LAB test name | *Real time - PCR: RT- PCR *Complete Blood Count: CBC * alanine /aspartate transaminase :AST/ ALT *Blood Urea Nitrogen: BUN *C-reactive protein level: CRP *Platelet: PLT, * Lymphocyte: LYM *Neutrophil: NEU, * Creatinine: Cr *Lactate dehydrogenase: LDH, *other | Categorical |
| Discharge Prescribed drugs | Drug name | String | Test result | * Positive CoV * Negative CoV | Binary |
| Date of follow-up | yyyy /mm/ dd | Date | Test time | yyyy /mm/ dd | Date |

cal registries offer enhanced progresses in systematic collection, analysis, comparison, and integration of population-based data. Clinical registries allow ongoing monitoring and benchmarking of clinical management and treatment outcomes, and thus are considered among the most effective strategies for quality healthcare improvement. They are powerful and comprehensive tools for conducting research and detecting eligible subjects to contribute in a particular study or clinical trial (5, 36-38). COVID-19 registry has been developed on the web platform enabling scientific teamwork in the field of COVID-19 and purposes to achieve a collaborative multi setting research study with a flexible registry structure. Moreover, no extra on-site software installation, configuration, or hardware is needed because of the web-based platform.

One of the basic steps in building a registry was determining a minimum and yet inclusive required data set that would be standardized across organizations and could pave the way for collaboration between researchers (39). To identify the necessary data on COVID-19 across clinical and public health information systems, initially a list of potential data elements gathered through conducting

extensive literature review. Also, the number of data elements was reduced after the expert panel's discussion and vote, and ultimately the MDS were finalized for inclusion in the COVID-19 registry. The COVID-MDS aims to harmonize the collection process, and increase the comparability of clinical care data across COVID-19 registries and databases, and facilitate pooled analyses to address clinical research questions.

The quality of research results depends on generalizable and high quality data (40). As Ieva et al (2014) stated, "when reliable data capture system recording data regarding a disease natural history progression, researchers can design studies more creditably and identify suitable subjects." (41) Also, the International Conference on Harmonization (ICH) guideline E6 on good clinical practice (GCP) necessitates that clinical trials be conducted based precise, comprehensive and verifiable data to guarantee patient safety and data quality. Thus, the credibility of registry is emphasized by researchers (42). Studies derived from well-designed and well-implemented registries provide a more realistic view of clinical procedures, patient outcomes, safety and efficacy, and measurable effective



Home Search Contact Login News Sociodemographic

Required data elements:

First name:

Last name:

Father name:

Age(in years):

National ID:

Sex: Male female

Place of birth:

Marital status:

Marital status: single married

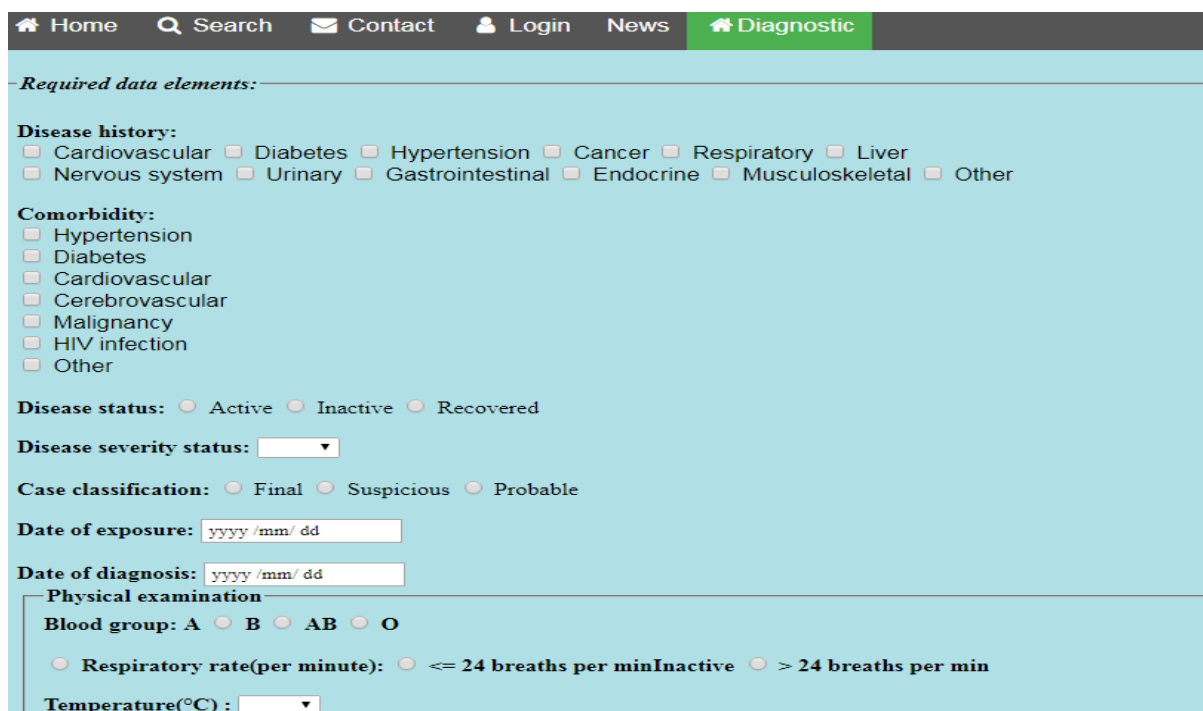
Employment status :

Home Address :

Postal / Zip code :

Reason for admission :

Fig. 1. Administrative data entry screen



Home Search Contact Login News Diagnostic

Required data elements:

Disease history:

Cardiovascular Diabetes Hypertension Cancer Respiratory Liver
 Nervous system Urinary Gastrointestinal Endocrine Musculoskeletal Other

Comorbidity:

Hypertension
 Diabetes
 Cardiovascular
 Cerebrovascular
 Malignancy
 HIV infection
 Other

Disease status: Active Inactive Recovered

Disease severity status:

Case classification: Final Suspicious Probable

Date of exposure:

Date of diagnosis:

Physical examination

Blood group: A B AB O

Respiratory rate(per minute): ≤ 24 breaths per min Inactive > 24 breaths per min

Temperature($^{\circ}$ C) :

Fig. 2. Clinical data entry screen

tiveness, and support the decision-making and evidence-based design process (43).

The quality of clinical registers can be restricted due to poor uptake or unreliable data entry process. The manual data entry is time-consuming for clinical staff and is vulnerable to documentation errors, such as inaccuracies and omissions (9, 44). In COVID-19 registry, an electronic web-based data entry is provided to automatically reject incorrect values or those that are outside the range; furthermore, the manual entry of data is avoided as much as

possible.

Furthermore, to comply with other data quality criteria, such as data consistency and comparability in COVID-19 registry, first, most required data elements and their values were determined for reporting COVID-19 in a consistent manner across Iran's health system. COVID-19 registry is comprehensive and can provide an in-depth description of specific patient cohorts rather than delivering epidemiological data. Another key feature for any registry is interoperability with other health information systems that

can be helpful to avoid duplication of data entry and reduce the workload on care givers. Bergin et al (2016) also recognized some possible challenges in developing a registry that hinder comprehensive and accurate data capture, including the increased workload of health care providers and proper integration of data capture into daily clinical workflow (45). Therefore, it is valuable to harmonize data elements, data descriptions, and process for uniform capturing of each item (46, 47). Thus, in this study, both COVID19 MDS and detailed categories (levels) and data formats for data capturing were defined. For future studies, working on technical aspects of data exchanging to automated pool data in the registry is the next challenge.

Given some unfamiliar aspects of the COVID-19, further development and adjustments are required; thus, conducting a pilot study, including a further Delphi step to refine the MDS, is recommended. Moreover, this MDS may need to be evaluated from the perspectives of larger group of medical and public health experts to be applicable at the national level. Further, we used the Delphi consensus approach to reach an agreement on COVID-19 MDS. This technique has been demonstrated to be suitable for assessment information systems requirements (48). However, one of its restrictions is that most views are marginalized. Despite the aforementioned limitations, this registry provides a standardized and agreed dataset on COVID-19 to accumulate patients, so gradually larger cohorts will be available in the future. In addition, this registry can collect large volumes of data from multiple settings and lay the foundation to conduct in-depth analyses by the artificial intelligence (AI) technique on many unfamiliar aspects of COVID-19. In addition, it is expected to push quickly towards better scientific collaboration for COVID-19. Registry implementation also allowed us to evaluate the quality of care and to help inform best practice in controlling COVID-19.

Conclusion

This study represents a fundamental effort towards building a national registry that uses information management approaches to improve accuracy, completeness, comparability, and interoperability of data about COVID-19 across the health care sector. This registry helps to conduct surveys to study various aspects of COVID 19 using a set of variables that were included in the registry according to the experts' opinions.

Acknowledgments

This article is extracted from a research project supported by Ilam University of Medical Sciences (IR.MEDILAM.REC.1399.040). We also thank the Research Deputy of Ilam University of Medical Sciences for financially supporting this project.

Conflict of Interests

The authors declare that they have no competing interests.

References

1. Lin X, Gong Z, Xiao Z, Xiong J, Fan B, Liu J. Novel coronavirus

2. pneumonia outbreak in 2019: computed tomographic findings in two cases. *Korean J Radiol.* 2020;21(3):365-8.
3. Shi H, Han X, Jiang N, Cao Y, Alwalid O, Gu J, et al. Radiological findings from 81 patients with COVID-19 pneumonia in Wuhan, China: a descriptive study. *Lancet Infect Dis.* 2020.
4. Yang W, Cao Q, Qin L, Wang X, Cheng Z, Pan A, et al. Clinical characteristics and imaging manifestations of the 2019 novel coronavirus disease (COVID-19): A multi-center study in Wenzhou city, Zhejiang, China. *J Infect.* 2020.
5. Guan WJ, Ni ZY, Hu Y, Liang WH, Ou CQ, He JX, et al. Clinical characteristics of coronavirus disease 2019 in China. *N Engl J Med.* 2020.
6. Alonso-Pérez J, Segovia S, Domínguez-González C, Olivé M, Grimón MDM, Fernández-Torrón R, et al. Spanish Pompe registry: baseline characteristics of first 49 patients with adult onset of Pompe disease. *Med Clin (Barc).* 2020.
7. Hornik CP, Gelfand AA, Szperka CL, Pezzuto T, Utevsky A, Kessel S, et al. Development of a Prospective Real-World Data Clinical Registry of Children and Adolescents With Migraine. *Headache.* 2020;60(2):405-15.
8. Bellgard MI, Render L, Radochowski M, Hunter A. Second generation registry framework. *Source Code Biol Med.* 2014;9(1):14.
9. Boulanger V, Schlemmer M, Rossvo S, Seebald A, Gavin P. Establishing Patient Registries for Rare Diseases: Rationale and Challenges. *Pharmaceut Med.* 2020:1-6.
10. Yang M, Loeb DF, Sprowell AJ, Trinkley KE. Design and implementation of a depression registry for primary care. *Am J Med Qual.* 2019;34(1):59-66.
11. Hurvitz EA, Gross PH, Gannotti ME, Bailes AF, Horn SD. Registry-based Research in Cerebral Palsy: The Cerebral Palsy Research Network. *Phys Med Rehabil Clin N Am.* 2020;31(1):185-94.
12. Irie H, Okamoto H, Uchino S, Endo H, Uchida M, Kawasaki T, et al. The Japanese Intensive care PATient Database (JIPAD): A national intensive care unit registry in Japan. *J Crit Care.* 2020;55:86-94.
13. Kazemi-Arpanahi H, Vashghani-Farahani A, Baradaran A, Mohammadzadeh N, Ghazisaeedi M. Developing a Minimum Data Set (MDS) for Cardiac Electronic Implantable Devices Implantation. *Acta Inform Med.* 2018;26(3):164-8.
14. Kazemi-Arpanahi H, Vashghani-Farahani A, Baradaran A, Ghazisaeedi M, Mohammadzadeh N, Bostan H. Development of a minimum data set for cardiac electrophysiology study ablation. *J Educ Health Promot.* 2019;8:101-.
15. Chang TH, Wu JL, Chang LY. Clinical characteristics and diagnostic challenges of pediatric COVID-19: A systematic review and meta-analysis. *J Formos Med Assoc.* 2020.
16. Fang Z, Yi F, Wu K, Lai K, Sun X, Zhong N, et al. Clinical Characteristics of 2019 Coronavirus Pneumonia (COVID-19): An Updated Systematic Review. *medRxiv.* 2020.
17. Fu H, Xu H, Zhang N, Xu H, Li Z, Chen H, et al. Association between Clinical, Laboratory and CT Characteristics and RT-PCR Results in the Follow-up of COVID-19 patients. *medRxiv.* 2020.
18. Fu L, Wang B, Yuan T, Chen X, Ao Y, Fitzpatrick T, et al. Clinical characteristics of coronavirus disease 2019 (COVID-19) in China: a systematic review and meta-analysis. *J Infect.* 2020.
19. Huang Y, Zhou H, Yang R, Xu Y, Feng X, Gong P. Clinical characteristics of 36 non-survivors with COVID-19 in Wuhan, China. *medRxiv.* 2020.
20. Jia J, Hu X, Yang F, Song X, Dong L, Zhang J, et al. Epidemiological characteristics on the clustering nature of COVID-19 in Qingdao City, 2020: a descriptive analysis. *Disaster Med Public Health Prep.* 2020:1-17.
21. Sahin AR, Erdogan A, Agaoglu PM, Dineri Y, Cakirci AY, Senel ME, et al. 2019 novel coronavirus (COVID-19) outbreak: a review of the current literature. *EJMO.* 2020;4(1):1-7.
22. Lyu P, Liu X, Zhang R, Shi L, Gao J. The performance of chest CT in evaluating the clinical severity of COVID-19 pneumonia: identifying critical cases based on CT characteristics. *Invest Radiol.* 2020.
23. Mao Y, Lin W, Wen J, Chen G. Clinical and pathological characteristics of 2019 novel coronavirus disease (COVID-19): a systematic reviews. *MedRxiv.* 2020.
24. Pormohammad A, Ghorbani S, Baradaran B, Khatam A, Turner R, Mansournia MA, et al. Clinical Characteristics, Laboratory Findings, Radiographic Signs and Outcomes of 52,251 Patients with Confirmed

- COVID-19 Infection: A Systematic Review and Meta-Analysis. Preprints. 2020.
24. Surveillances V. The epidemiological characteristics of an outbreak of 2019 novel coronavirus diseases (COVID-19)—China, 2020. *China CDC Wkly.* 2020;2(8):113-22.
 25. Tabata S, Imai K, Kawano S, Ikeda M, Kodama T, Miyoshi K, et al. The clinical characteristics of COVID-19: a retrospective analysis of 104 patients from the outbreak on board the Diamond Princess cruise ship in Japan. *medRxiv.* 2020.
 26. Wang L, Duan Y, Zhang W, Liang J, Xu J, Zhang Y, et al. Epidemiologic and Clinical Characteristics of 26 Cases of COVID-19 Arising from Patient-to-Patient Transmission in Liaocheng, China. *Clin Epidemiol.* 2020;12:387.
 27. Wang X, Fang J, Zhu Y, Chen L, Ding F, Zhou R, et al. Clinical characteristics of non-critically ill patients with novel coronavirus infection (COVID-19) in a Fangcang Hospital. *Clin Microbiol Infect.* 2020.
 28. Xia XY, Wu J, Liu HL, Xia H, Jia B, Huang WX. Epidemiological and initial clinical characteristics of patients with family aggregation of COVID-19. *J Clin Virol.* 2020:104360.
 29. Zhao X, Zhang B, Li P, Ma C, Gu J, Hou P, et al. Incidence, clinical characteristics and prognostic factor of patients with COVID-19: a systematic review and meta-analysis. *medRxiv.* 2020.
 30. Zheng Y, Xu H, Yang M, Zeng Y, Chen H, Liu R, et al. Epidemiological Characteristics and Clinical Features of 32 Critical and 67 Noncritical Cases of COVID-19 in Chengdu. *J Clin Virol.* 2020:104366.
 31. Šimović V, Varga M. Applying SQL database query to access SQL server 2019—visual studio 2019. *J Agron.* 2019;6.
 32. Strauss D. Being More Productive in Visual Studio. Exploring Advanced Features in C#: Springer; 2019. p. 221-81.
 33. Petrovečki M, Kovačević Ž. Featured characteristics of SQL Server 2019. *Polytechnic Design.* 2019;7(1):1-7.
 34. Allam Z, Jones DS, editors. On the coronavirus (COVID-19) outbreak and the smart city network: universal data sharing standards coupled with artificial intelligence (AI) to benefit urban health monitoring and management. *Healthcare*; 2020: Multidisciplinary Digital Publishing Institute.
 35. Fisher D, Heymann D. Q&A: The novel coronavirus outbreak causing COVID-19. *BMC Med.* 2020;18(1):1-3.
 36. Choinière M, Ware M, Pagé M, Lacasse A, Lanctôt H, Beaudet N, et al. Development and implementation of a registry of patients attending multidisciplinary pain treatment clinics: The Quebec Pain Registry. *Pain Res Manag.* 2017;2017.
 37. Akand M, Sarıca K, Kiremit MC, Soytas M, Güven S. Development of a prospective data registry system for retrograde intrarenal surgery in renal stones: Turkish Academy of Urology Prospective Study Group (ACUP study). *Turk J Urol.* 2020;46(1):57.
 38. Hoque DME, Kumari V, Hoque M, Ruseckaite R, Romero L, Evans SM. Impact of clinical registries on quality of patient care and clinical outcomes: a systematic review. *PLoS One.* 2017;12(9):e0183667.
 39. Golan R, Bernstein A, Sedrakyan A, Daskivich TJ, Du DT, Ehdaie B, et al. Development of a nationally representative coordinated registry network for prostate ablation technologies. *J Urol.* 2018;199(6):1488-93.
 40. Shanbehzadeh M, Ahmadi M. Identification of the necessary data elements to report AIDS: A systematic review. *E-physician.* 2017;9(12):5920.
 41. Ieva F, Gale CP, Sharples LD. Contemporary roles of registries in clinical cardiology: when do we need randomized trials? Taylor & Francis; 2014.
 42. Tian Q, Liu M, Min L, An J, Lu X, Duan H. An automated data verification approach for improving data quality in a clinical registry. *Comput Methods Programs Biomed.* 2019;181:104840.
 43. Korngut L, Genge A, Johnston M, Benstead T, Bourque P, Briemberg H, et al. Establishing a Canadian registry of patients with amyotrophic lateral sclerosis. *Can J Neurol Sci.* 2013;40(1):29-35.
 44. Williams A, Goedicke W, Tissera KA, Mankarious LA. Leveraging Existing Tools in Electronic Health Record Systems to Automate Clinical Registry Compilation. *Otolaryngol Head Neck Surg.* 2020;162(3):408-9.
 45. Bergin K, Moore E, McQuilten Z, Wood E, Augustson B, Blacklock H, et al. Design and development of the Australian and New Zealand (ANZ) myeloma and related diseases registry. *BMC.* 2016;16(1):151.
 46. Plotnicki L, Kohl C, Höcker B, Krupka K, Rahmel A, Pape L, et al., editors. The CERTAIN Registry: a novel, web-based registry and research platform for pediatric renal transplantation in Europe. *Transplantation proceedings*; 2013: Elsevier.
 47. Gerke AK, Tang F, Cozier YC, Lash MT, Schappet J, Phillips E, et al. A web-based registry for patients with sarcoidosis. *Sarcoidosis Vasc Diffuse Lung Dis.* 2017;34(1):26-34.
 48. Niazkhani Z, Cheshmekaboodi M, Pirnejad H, Makhdoomi K, Nikibakhsh AA, Abkhiz S, et al. Enabling informed policymaking for chronic kidney disease with a registry: Initiatory steps in Iran and the path forward. *Health Policy Technol.* 2018;7(1):73-80.