

# CORD-19-on-FHIR

Guoqian Jiang, Mayo Clinic  
Harold Solbrig, Johns-Hopkins University  
David Booth, Yosemite Project

7-Apr-2020

Github: <https://github.com/fhircat/CORD-19-on-FHIR>

<https://pages.semanticscholar.org/coronavirus-research>

## COVID-19 Open Research Dataset (CORD-19)

Access this dataset to help with the fight against COVID-19

### A Free, Open Resource for the Global Research Community

In response to the COVID-19 pandemic, the [Allen Institute for AI](#) has partnered with leading research groups to prepare and distribute the COVID-19 Open Research Dataset (CORD-19), a free resource of over 44,000 scholarly articles, including over 29,000 with full text, about COVID-19 and the coronavirus family of viruses for use by the global research community.

This dataset is intended to mobilize researchers to apply recent advances in natural language processing to generate new insights in support of the fight against this infectious disease. The corpus will be updated weekly as new research is published in peer-reviewed publications and archival services like [bioRxiv](#), [medRxiv](#), and others.

### Participate in the CORD-19 Challenge

Kaggle is hosting the [COVID-19 Open Research Dataset Challenge](#), a series of important questions designed to inspire the community to use CORD-19 to find new insights about the COVID-19 pandemic including the natural history, transmission, and diagnostics for the virus, management measures at the human-animal interface, lessons from previous epidemiological studies, and more.



# <https://pages.semanticscholar.org/coronavirus-research>

## Download CORD-19

By downloading this dataset you are agreeing to the [Dataset License](#). Specific licensing information for individual articles in the dataset is available in the metadata file.

Additional licensing information is available on the [PMC website](#), [medRxiv website](#) and [bioRxiv website](#).

### Download here:

*Latest release contains papers up until 2020-03-20 with over 29,000 full text articles. ([Changelog from previous release.](#))*

- [Commercial use subset](#) (includes PMC content) – 9118 full text (new: 128), 183Mb
- [Non-commercial use subset](#) (includes PMC content) – 2353 full text (new: 385), 41Mb
- [Custom license subset](#) – 16959 full text (new: 15533), 345Mb
- [bioRxiv/medRxiv subset](#) (pre-prints that are not peer reviewed) – 885 full text (new: 110), 14Mb
- [Metadata file](#) – 60Mb
- [Readme](#)

Each paper is represented as a single JSON object. The [schema](#) is available here.

### Description:

The dataset contains all COVID-19 and coronavirus-related research (e.g. SARS, MERS, etc.) from the following sources:

- PubMed's PMC open access corpus using this [query](#) (COVID-19 and coronavirus research)
- Additional COVID-19 research articles from a corpus maintained by the [WHO](#)
- bioRxiv and medRxiv pre-prints using the same query as PMC (COVID-19 and coronavirus research)

# <https://github.com/fhircat/CORD-19-on-FHIR>

## **CORD-19-on-FHIR -- Semantics for COVID-19 Discovery**

---

CORD-19-on-FHIR is a superset of the COVID-19 Open Research Dataset ([CORD-19](#)) data, provided by the [Allen Institute](#) to support research on COVID-19 / SARS-CoV-2 / Novel Coronavirus. It is represented in [FHIR RDF](#) and was produced by data mining the CORD-19 dataset and adding semantic annotations. The purpose is to facilitate linkage with other biomedical datasets and enable answering research questions.

### **Wiki**

---

<https://github.com/fhircat/CORD-19-on-FHIR/wiki>

### **Semantic annotations**

---

So far, CORD-19-on-FHIR adds the following semantic annotations to the CORD-19 dataset, based only on parsing the titles and abstracts:

- Conditions - 103,968 instances
- Medications - 16,406 instances
- Procedures - 54,720 instances

We plan to parse the full text articles soon, for those that we can access.

If you can generate other semantic linkages, please let us know. Collaboration is invited. Please issue a pull request or contact Guoqian Jiang [jiang.guoqian@mayo.edu](mailto:jiang.guoqian@mayo.edu) and Harold Solbrig [solbrig@jhu.edu](mailto:solbrig@jhu.edu).

# https://fhircat.org/cord-19/fhir/Commercial/Composition/00623bf2715e25d3

000cb2f210d6000ed040e2cb.json

JSON Raw Data Headers

Save Copy Collapse All Expand All Filter JSON

resourceType: "Composition" ← Composition

id: "00623bf2715e25d3acacb3f210d6888ed840e3cb.json"

identifier:

- system: "urn:ietf:rfc:3986"
- value: "urn:uuid:33a112c3-0131-3617-9cb1-765b5474a049"
- status: "final"
- confidentiality: "N"

section: ← Section

- 0:
- code:

  - coding:

    - 0:
    - system: "http://hl7.org/fhir/ValueSet/doc-section-codes"
    - code: "61150-9"

text:

- status: "additional"
- div: "<div xmlns=\"http://www.w3.org/1999/xhtml\">TITLE: Transmissible gastroenteritis virus infection decreases arginine uptake by downregulating CAT-1 expression \n</div>"
- mode: "working"

entry:

- 0: reference: "Condition/d207290f-d42b-35cd-a57c-db9feffc3f10"
- 1: reference: "Condition/ce256190-76b2-3fad-ae10-6060dee4c3cf"
- 2: reference: "Condition/dabd9625-c8d7-3df2-baf5-a1596fe36c8e"
- 3: reference: "MedicationStatement/30e2e66d-99da-348c-b242-e56ac7b7e6a6"
- 4: reference: "Medication/a56d2d61-ec6b-323e-b8c5-6cb11fdb735d"
- 5: reference: "Procedure/0b2d9261-9015-3dbb-9d1f-54fa5f1f4c14"

Semantic Annotations

JSON Raw Data Headers

Save Copy Collapse All Expand All Filter JSON

resourceType: "Condition"

id: "00006534-eee4-3d38-bbe9-45374ee0f3c7"

identifier:

- 0:
- system: "urn:ietf:rfc:3986"
- value: "urn:uuid:00006534-eee4-3d38-bbe9-45374ee0f3c7"

code:

- coding:

  - 0:
  - system: "http://snomed.info/sct"
  - code: "233604007"

text: "pneumonia"

abatementString: "positive"

<https://www.ncbi.nlm.nih.gov/research/pubtator/>



U.S. National Library of Medicine

# PubTator<sub>Central</sub>

Bioconcept

Optional comma-delimited list of the bioconcept types to include in the results, one or more of: **gene, disease, chemical, species, mutation and/or cellline.**  
If this parameter is not present, then results will contain all six bioconcepts.

**⚠** *This parameter is only compatible with **pubtator** format.*

# http://bioportal.bioontology.org/ontologies/CI DO



BioPortal

[Ontologies](#) [Search](#) [Annotator](#) [Recommender](#) [Mappings](#) [Resource Index](#)

Login

Support ▾

## Coronavirus Infectious Disease Ontology

Last updated: March 12, 2020



[Summary](#) [Classes](#) [Properties](#) [Notes](#) [Mappings](#) [Widgets](#)

### Details

Acronym	CIDO
Visibility	Public
Description	The Ontology of Coronavirus Infectious Disease (CIDO) is a community-driven open-source biomedical ontology in the area of coronavirus infectious disease. The CIDO is developed to provide standardized human- and computer-interpretable annotation and representation of various coronavirus infectious diseases, including their etiology, transmission, pathogenesis, diagnosis, prevention, and treatment.
Status	Alpha
Format	OWL
Contact	Yongqun Oliver He, <a href="mailto:yongqunh@med.umich.edu">yongqunh@med.umich.edu</a>
Categories	All Organisms, Health, Human

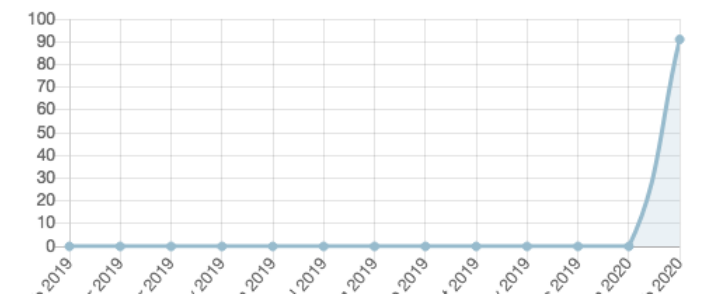
### Metrics ?

Classes	2,722
Individuals	336
Properties	72
Maximum depth	34
Maximum number of children	403
Average number of children	3
Classes with a single child	506
Classes with more than 25 children	9
Classes with no definition	2,559

### Submissions

Version	Released	Uploaded	Downloads
<a href="#">Vision Release: 1.0.87</a> (Parsed, Indexed, Metrics, Annotator)	03/14/2020	03/12/2020	<a href="#">OWL</a>   <a href="#">CSV</a>   <a href="#">RDF/XML</a>   <a href="#">Diff</a>
<a href="#">Vision Release: 1.0.85</a> (Archived)	02/02/2020	02/02/2020	<a href="#">OWL</a>   <a href="#">Diff</a>
<a href="#">Vision Release: 1.0.25</a> (Archived)	01/28/2020	01/28/2020	<a href="#">OWL</a>   <a href="#">Diff</a>
<a href="#">Vision Release: 1.0.06</a> (Archived)	01/26/2020	01/26/2020	<a href="#">OWL</a>

### Visits



In Progress



# Evidence Resource from

## EDM - FLUID

Structure

UML

XML




















JSON

Turtle

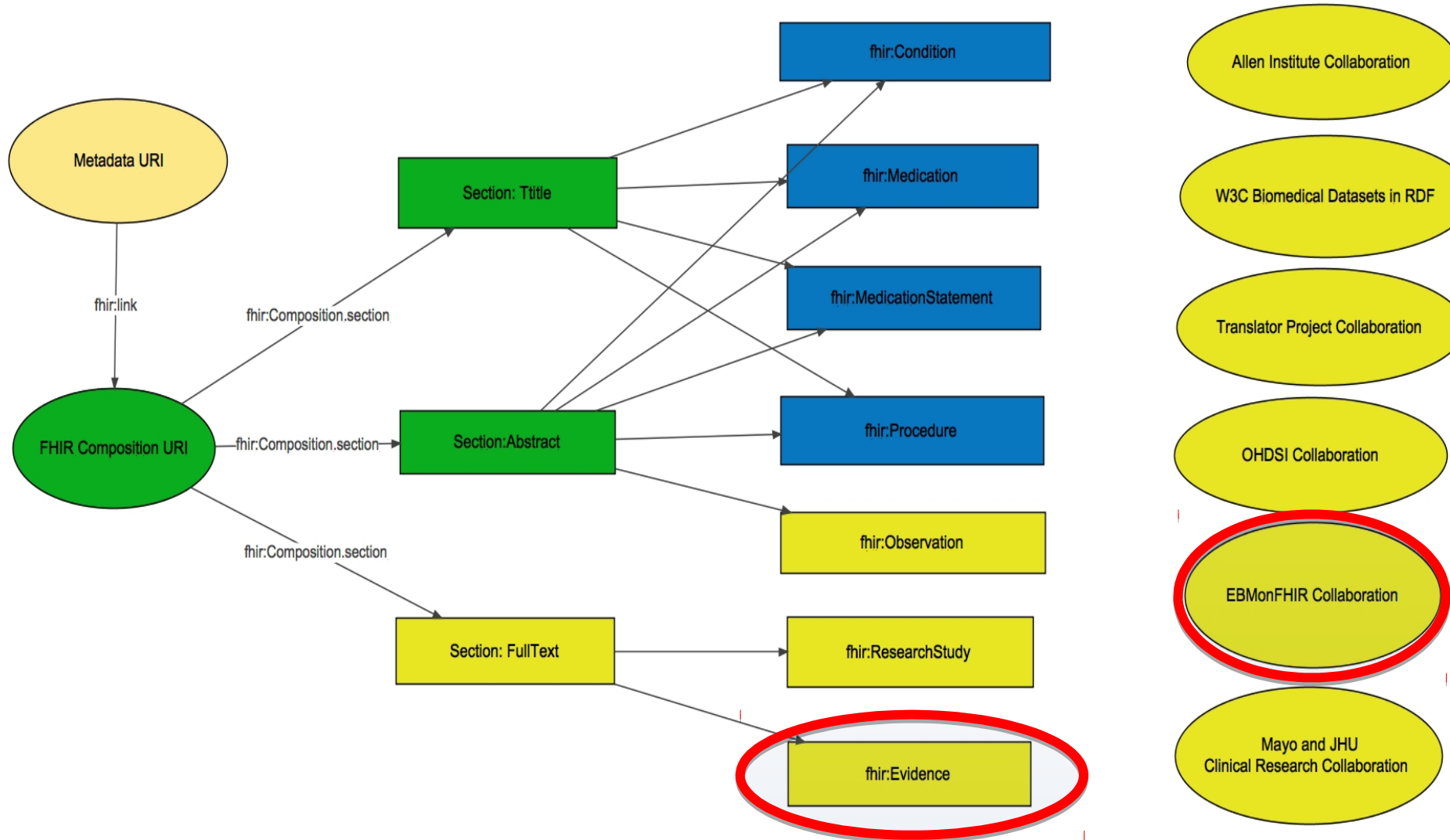
R3 Diff

All

### Structure

Name	Flags	Card.	Type	Description & Constraints
 Evidence	<b>TU</b>		MetadataResource	Single evidence bit + <i>Warning: Name should be usable as an identifier for the module by machine processing applications such as code generation</i> Elements defined in Ancestors: <a href="#">id</a> , <a href="#">meta</a> , <a href="#">implicitRules</a> , <a href="#">language</a> , <a href="#">text</a> , <a href="#">contained</a> , <a href="#">extension</a> , <a href="#">modifierExtension</a> , <a href="#">url</a> , <a href="#">identifier</a> , <a href="#">version</a> , <a href="#">name</a> , <a href="#">title</a> , <a href="#">status</a> , <a href="#">experimental</a> , <a href="#">date</a> , <a href="#">publisher</a> , <a href="#">contact</a> , <a href="#">description</a> , <a href="#">useContext</a> , <a href="#">jurisdiction</a> , <a href="#">purpose</a> , <a href="#">copyright</a> , <a href="#">approvalDate</a> , <a href="#">lastReviewDate</a> , <a href="#">effectivePeriod</a>
...  url	Σ	0..1	uri	Canonical identifier for this evidence, represented as a globally unique URI
...  identifier	Σ	0..*	Identifier	Additional identifier for the summary
...  version	Σ	0..1	string	Business version of this summary
...  title	Σ	0..1	string	Name for this summary (human friendly)
...  status	?! Σ	1..1	code	draft   active   retired   unknown <a href="#">PublicationStatus</a> (Required)
...  date	Σ	0..1	dateTime	Date last changed
...  useContext	Σ	0..*	UsageContext	Use context
...  approvalDate		0..1	date	When the summary was approved by publisher
...  lastReviewDate		0..1	date	When the summary was last reviewed
...  contributor		0..*	Contributor	Contact detail and role for contributor to summary
...  relatedArtifact		0..*	RelatedArtifact	Link or citation to artifact associated with the summary
...  description		0..1	markdown	Description of the particular summary
...  assertion		0..1	markdown	Declarative description of the Evidence
...  note		0..*	Annotation	Footnotes and/or explanatory notes
 variableDefinition		1..*	BackboneElement	Evidence variable
...  description		0..1	markdown	Description of the variable
...  note		0..*	Annotation	Footnotes and/or explanatory notes
...  variableRole	Σ	1..1	CodeableConcept	population   subpopulation   exposure   referenceExposure   measuredVariable   confounder <a href="#">EvidenceVariableRole</a> (Extensible)

# CORD-19-on-FHIR Roadmap



# Brain's Suggestions

We would be interested in coordination with COVID-19-on-FHIR in a few ways:

1. We created an EBSCO COVID-19 Portal at <https://covid-19.ebscomedical.com/>
  - a. For Researchers page <https://covid-19.ebscomedical.com/research> already includes COVID-19
  - b. We can add a link to COVID-19-on-FHIR there as well.
2. We can help by reviewing the actual Evidence and EvidenceVariable resources creating for the COVID-19-on-FHIR related Evidence sets. We will be in the best position to comment on how we conceptualized these resource structures and will get immediate real-world insight to its application.
3. We can help by further developing the Evidence Data Entry form. Note the current form was highly focused on evidence for effect estimates and risk estimates. Other developments will be needed for many other types of evidence that can be created.

# Breakthrough: Chloroquine phosphate has shown apparent efficacy in treatment of COVID-19 associated pneumonia in clinical studies

Jianjun Gao<sup>1,\*</sup>, Zhenxue Tian<sup>2</sup>, Xu Yang<sup>2</sup>

<sup>1</sup> Department of Pharmacology, School of Pharmacy, Qingdao University, Qingdao, China;

<sup>2</sup> Department of Pharmacy, Qingdao Municipal Hospital, Qingdao, China.

**SUMMARY** The coronavirus disease 2019 (COVID-19) virus is spreading rapidly, and scientists are endeavoring to discover drugs for its efficacious treatment in China. Chloroquine phosphate, an old drug for treatment of malaria, is shown to have apparent efficacy and acceptable safety against COVID-19 associated pneumonia in multicenter clinical trials conducted in China. The drug is recommended to be included in the next version of the Guidelines for the Prevention, Diagnosis, and Treatment of Pneumonia Caused by COVID-19 issued by the National Health Commission of the People's Republic of China for treatment of COVID-19 infection in larger populations in the future.

**Keywords** COVID-19, SARS-CoV-2, 2019-nCoV, pneumonia, chloroquine

**ARTICLE**      **OPEN**

# High expression of ACE2 receptor of 2019-nCoV on the epithelial cells of oral mucosa

Hao Xu<sup>1</sup>, Liang Zhong<sup>1</sup>, Jiaxin Deng<sup>1</sup>, Jiakuan Peng<sup>1</sup>, Hongxia Dan<sup>1</sup>, Xin Zeng<sup>1</sup>, Taiwen Li<sup>1</sup> and Qianming Chen<sup>1,2</sup>

It has been reported that ACE2 is the main host cell receptor of 2019-nCoV and plays a crucial role in the entry of virus into the cell to cause the final infection. To investigate the potential route of 2019-nCoV infection on the mucosa of oral cavity, bulk RNA-seq profiles from two public databases including The Cancer Genome Atlas (TCGA) and Functional Annotation of The Mammalian Genome Cap Analysis of Gene Expression (FANTOM5 CAGE) dataset were collected. RNA-seq profiling data of 13 organ types with para-carcinoma normal tissues from TCGA and 14 organ types with normal tissues from FANTOM5 CAGE were analyzed in order to explore and validate the expression of ACE2 on the mucosa of oral cavity. Further, single-cell transcriptomes from an independent data generated in-house were used to identify and confirm the ACE2-expressing cell composition and proportion in oral cavity. The results demonstrated that the ACE2 expressed on the mucosa of oral cavity. Interestingly, this receptor was highly enriched in epithelial cells of tongue. Preliminarily, those findings have explained the basic mechanism that the oral cavity is a potentially high risk for 2019-nCoV infectious susceptibility and provided a piece of evidence for the future prevention strategy in dental clinical practice as well as daily life.

# Github

<https://github.com/fhircat/CORD-19-on-FHIR>