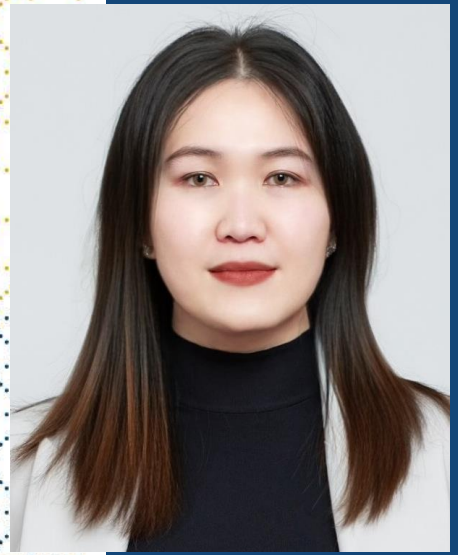
Facilitating Rapid and Accurate Meta-Analysis in Public Health Emergencies: The Role of CDISC Standards in Academic Research

Presented by Yen Phan, Founder and Sr. Clinical Data Scientist,
CodLad/University of Oxford



Meet the Speaker

Yen Phan

Title: Founder and Sr Clinical Data Scientist

Organization: CodLad/University of Oxford

Yen Phan is a Founder of CodLad and a Senior Data Scientist with over 10 years of experience in the industry. She is also a Lecturer at Technological University Dublin, Ireland on Pharma Science and Evidence based healthcare. She is a Doctoral Candidate at The George Washington University School of Medicine and Health Sciences. At the same time, she is finalizing her dissertation at University of Oxford with a focus of a meta-analysis.

Disclaimer and Disclosures

- *The views and opinions expressed in this presentation are those of the author(s) and do not necessarily reflect the official policy or position of CDISC.*
- *This presentation is intended for informational purposes only. While every attempt has been made to ensure the accuracy and reliability of the information provided, the author does not make any representations or warranties of any kind, express or implied, about the completeness, accuracy, reliability, suitability, or availability of the information contained in this presentation.*
- *The presenter declares no conflicts of interest in relation to this presentation. Any views or opinions presented are solely those of the presenter and do not necessarily represent those of the respective organizations mentioned within the presentation. Any examples of companies, studies, or technologies mentioned are for illustrative purposes only and do not constitute endorsements or recommendations.*

Agenda

1. Meta-Analysis
2. The importance of CDISC standards in Public Health Emergencies
3. Cases in point – experiences from selected EU projects
4. ISARIC-COVID-19 dataset

Meta-Analysis in Academic Research

- Meta-analysis is a statistical technique that combines the results of multiple studies to come up with a more comprehensive and reliable understanding of a research topic.
- It is a powerful tool for synthesizing research findings and drawing stronger conclusions than would be possible from any single study.

Steps in a Meta-analysis

Role of Meta-analysis in Synthesizing Research Findings

- Provides a more precise estimate of the effect of an intervention or treatment
- Helps to identify and explain heterogeneity
- Can be used to identify subgroup effects
- Can be used to inform clinical practice and policy decisions

The importance of CDISC standards in Public Health Emergencies

- Facilitates rapid data synthesis and analysis
- Enhances collaboration across institutions and countries
- Enables quicker decision-making processes
- Ensures data consistency and comparability

Cases in point – experiences from selected EU projects

Acronym (Project Title)	Start date	Duration	EU Funding	Approx. % dedicated to data harmonisation
ReCoDID (Reconciliation of Cohort Data for Infectious Diseases)	Jan 1st 2019	4 years	€7.760.021	60%
ORCHESTRA (Connecting European cohorts to increase common and effective response to SARS-CoV-2 pandemic)	Dec 1st 2020	3 years	€27.887.638	10%
uncover (Unravelling Data for Rapid Evidence-Based Response to COVID-19)	Nov 15th 2020	2 years	€2.997.440	70%
SYNCHROS (SYNergies for Cohorts in Health: integrating the ROle of all Stakeholders)	Jan 1st 2019	3,5 years	€ 1.991.812	N/A Data harmonisation not a direct project objective

ISARIC-COVID-19 dataset

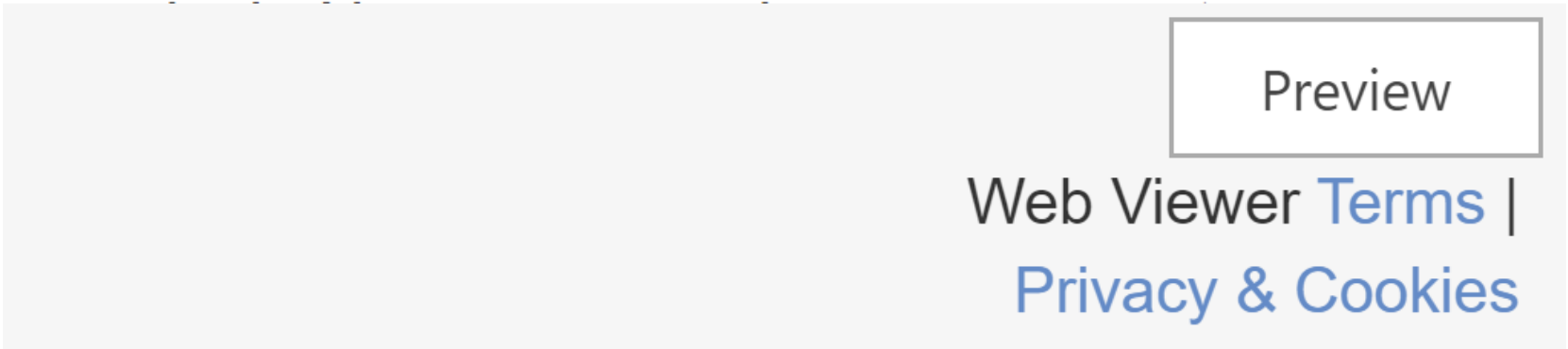
- The International Severe Acute Respiratory and Emerging Infection Consortium (ISARIC)
- January 2020, ISARIC launched a research response to the emergence of a COVID-19
- June 2022, 500 million reported cases and more than 6 million deaths.
- Despite unprecedented success in the rapid generation of vaccines and effective treatments, COVID-19 continues to cause severe and widespread health consequences.

ISARIC-COVID-19 dataset

- Between January 2020 and September 2021, the ISARIC-COVID-19 dataset was aggregated from the info of 705,000 patients with COVID-19, hospitalized across 62 countries.
- Data submitted to ISARIC by completing the CRF on REDCap hosted by the University of Oxford. Alternatively, data from different format/system can be shared to the ISARIC COVID-19 data platform, hosted by IDDO.

Insert Web Page

This app allows you to insert secure web pages starting with https:// into the slide



ISARIC-COVID-19 dataset
Data standardization

- The ISARIC COVID-19 dataset is a large, international, clinically comprehensive resource.
- Data was standardized using the CDISC SDTM model to harmonize diverse data structures and ontologies.
- CDISC SDTM was chosen for its flexibility and ability to accommodate varied data types, unlike the more rigid OMOP model.
- The SDTM model enhances interoperability, facilitating integration with future COVID-19 clinical trial data.

ISARIC COVID-19 DATA PLATFORM

Data are standardized to the Clinical Data Interchange Standards Consortium Study Data Tabulation Model (SDTM)

Data Collection

- Datasets collected on institution -specific systems
- Data collected using ISARIC REDCap data system

→ De-identification: direct identifiers are removed (i.e name, or identification number).

→ Pre-mapping: data and documentation revision by the data curator.

→ Mapping: Data transformation and assignation of data to the appropriate SDTM domain, variable, and controlled vocabulary.

→ Review and edit checks: review and checking of the different domains.

→ ISARIC COVID-19 DATASET

SDTM DOMAIN TYPES

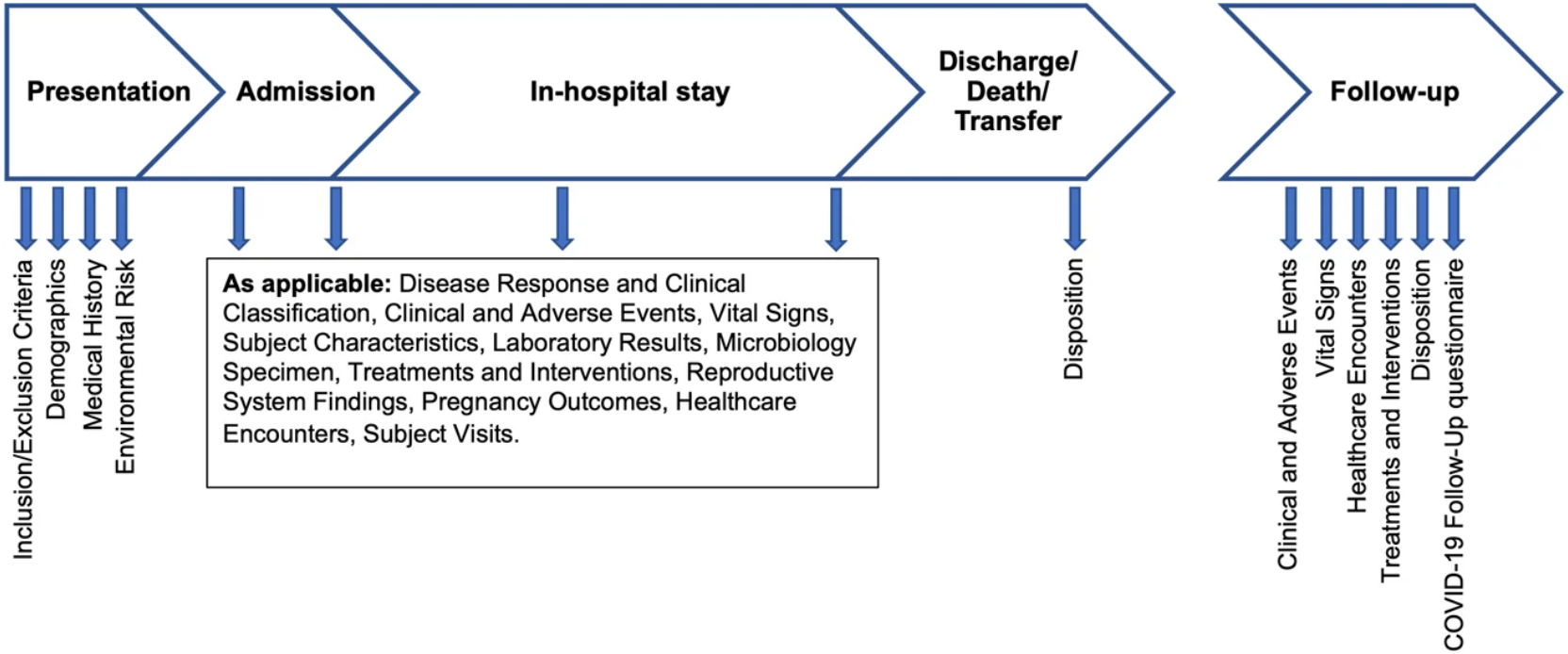
- Special- Purpose
  - Demographics (DM)
  - Subject Visits (SV)
- Trial Design
  - Trial Inclusion Exclusion Criteria (TI)
  - Trial Summary (TS)
- Findings
  - Vital Signs (VS)
  - Laboratory Results (LB)
- Events
  - Clinical and Adverse Events (SA)
  - Disposition (DS)
- Interventions
  - Treatments and Interventions (IN)

Variables

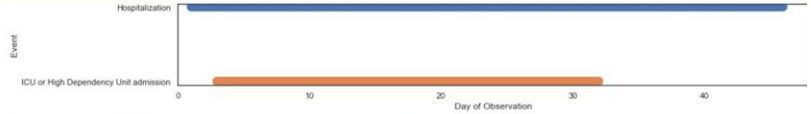
- **STUDYID:** Study Identifier
- **USUBJID:** Subject Identifier
- **DTHFL:** flag to indicate the death of the patient
- **AGE:** Age at the time of study enrolment
- **SEX:** patient's sex

Variables

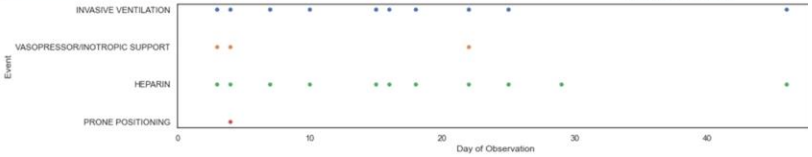
- **USUBJID:** Unique Subject Identifier
- **INTRT:** Reported Name of Drug or Intervention
- **INCAT:** Category for Observation
- **INOCCUR:** Observat-ion Occurrence
- **INDOSE:** Dose per Administration
- **INDY:** Study Day of Observation/Collection

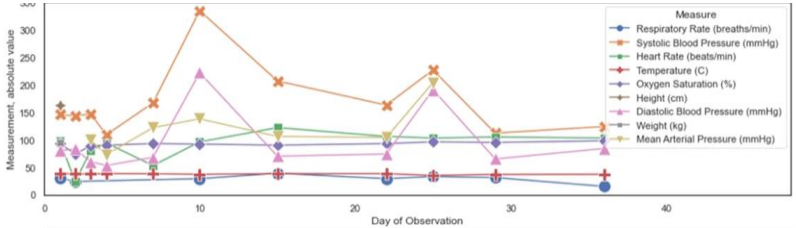

Presentation
Admission
In-hospital stay
Discharge/ Death/ Transfer
Follow-up
Inclusion/Exclusion Criteria
Demographics
Medical History
Environmental Risk
As applicable: Disease Response and Clinical Classification, Clinical and Adverse Events, Vital Signs, Subject Characteristics, Laboratory Results, Microbiology Specimen, Treatments and Interventions, Reproductive System Findings, Pregnancy Outcomes, Healthcare Encounters, Subject Visits.
Disposition
Clinical and Adverse Events
Vital Signs
Healthcare Encounters
Treatments and Interventions
Disposition
COVID-19 Follow-Up questionnaire

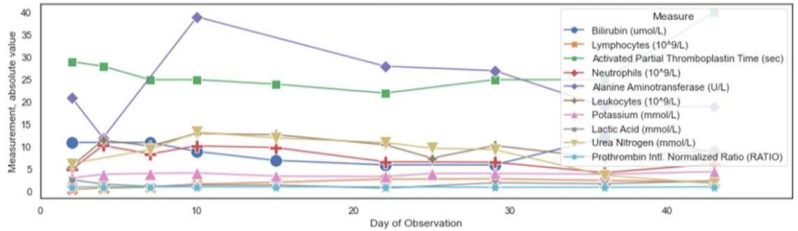
Variable	Domain	Value
Sex	Demographics	Female
Age (years)		66
Ethnic		White
Country		Norway
Comorbidities	Clinical and Adverse Events	Smoking
		Asthma
		Hypertension
		Obesity



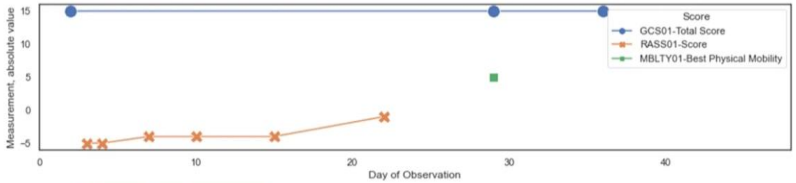
Variable	Domain	Value
Signs and symptoms at hospital admission	Clinical and Adverse Events	Fever/history of fever
		Lost/altered sense of taste
		Lost/altered sense of smell
		Cough - productive
		Vomiting/nausea
		Diarrhoea
		Muscle aches/joint pain
		Wheezing
		Shortness of breath

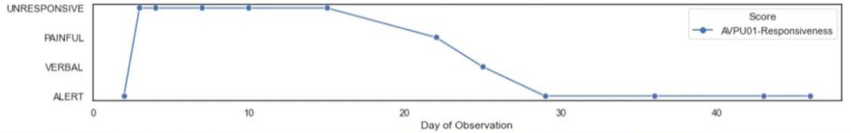




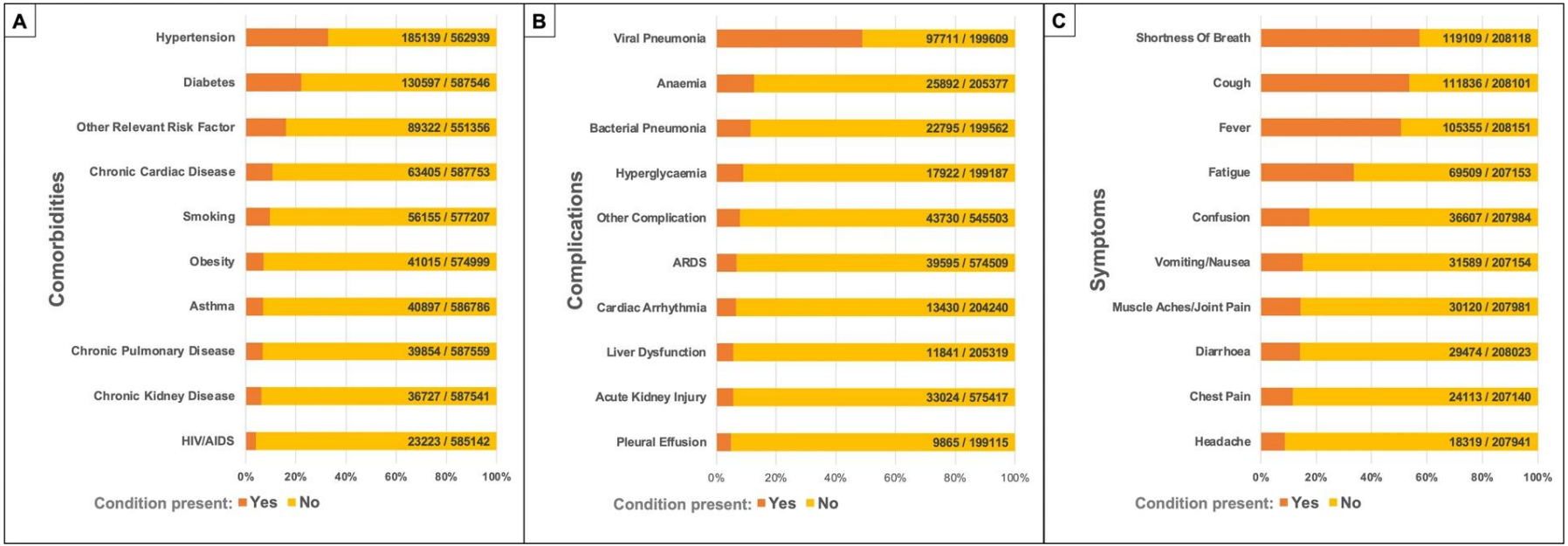


Variable	Domain	Value
Microorganism	Microbiology Specimen	ESCHERICHIA COLI
		HAEMOPHILUS INFLUENZAE
		COVID-2019/SARS-COV2





Variable	Domain	Value
Outcome	Disposition	Discharged alive


A
Comorbidities
Hypertension
185139 / 562939
Diabetes
130597 / 587546
Other Relevant Risk Factor
89322 / 551356
Chronic Cardiac Disease
63405 / 587753
Smoking
56155 / 577207
Obesity
41015 / 574999
Asthma
40897 / 586786
Chronic Pulmonary Disease
39854 / 587559
Chronic Kidney Disease
36727 / 587541
HIV/AIDS
23223 / 585142
0% 20% 40% 60% 80% 100%
Condition present: Yes No
B
Complications
Viral Pneumonia
97711 / 199609
Anaemia
25892 / 205377
Bacterial Pneumonia
22795 / 199562
Hyperglycaemia
17922 / 199187
Other Complication
43730 / 545503
ARDS
39595 / 574509
Cardiac Arrhythmia
13430 / 204240
Liver Dysfunction
11841 / 205319
Acute Kidney Injury
33024 / 575417
Pleural Effusion
9865 / 199115
0% 20% 40% 60% 80% 100%
Condition present: Yes No
C
Symptoms
Shortness Of Breath
119109 / 208118
Cough
111836 / 208101
Fever
105355 / 208151
Fatigue
69509 / 207153
Confusion
36607 / 207984
Vomiting/Nausea
31589 / 207154
Muscle Aches/Joint Pain
30120 / 207981
Diarrhoea
29474 / 208023
Chest Pain
24113 / 207140
Headache
18319 / 207941
0% 20% 40% 60% 80% 100%
Condition present: Yes No

The CDISC SDTM data model offers key advantages

- Adapts to any number of events, recording the sequence and timing of events using specific variables
- Captures whether a variable was collected and its outcome, allowing accurate denominator counts in diverse datasets
- Includes free-text entries for capturing fine-grained information, which can be analyzed using search functions or NLP

References

- https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9351292/
- https://www.nature.com/articles/s41597-022-01534-9
- https://livedataoxford.shinyapps.io/CovidClinicalDataDashboard/
- https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9351292/table/tbl0001/?report=objectonly

Thank You!