

OMOP Mapping of Real-World Data From Brazil & Pakistan Towards Management of COVID-19 In the Global South

Authors

Elzo Pereira Pinto Junior¹, Priscilla Normando¹, Renzo Flores-Ortiz¹, Muhammad Usman Afzal², Muhammad Asaad Jamil², Sergio Fernandez Bertolin³, Vinícius de Araújo Oliveira¹, Valentina Martufi¹, Edward Burn⁴, Maria Yury Ichihara¹, Maurício L. Barreto¹, Talita Duarte Salles³, Daniel Prieto-Alhambra⁴, Haroon Hafeez², Sara Khalid⁴

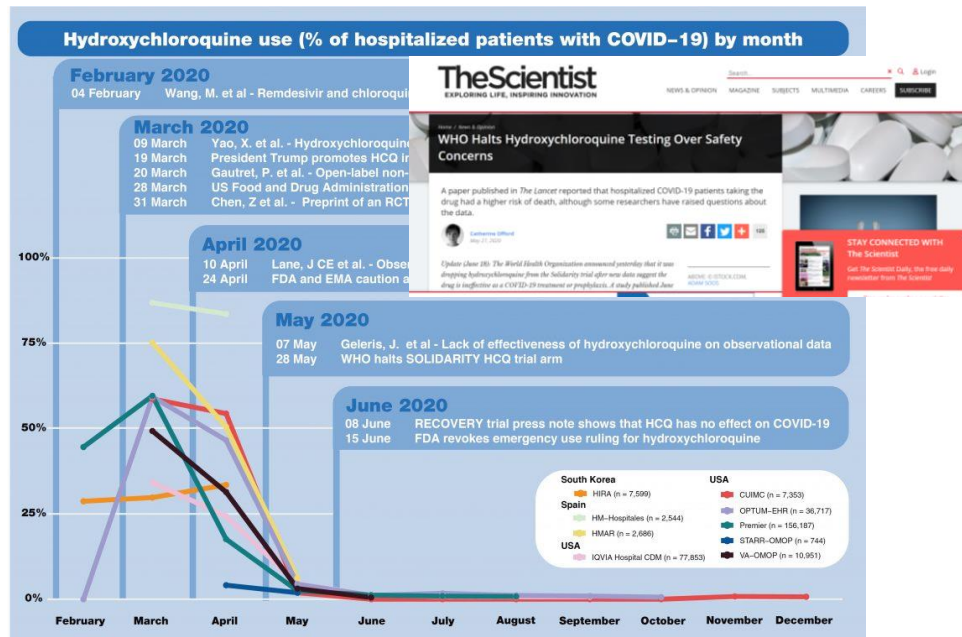
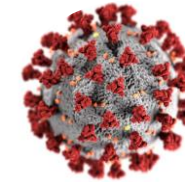
¹ The Center for Data and Knowledge Integration for Health (CIDACS – FIOCRUZ), Brazil

² Shaukat Khanum Memorial Cancer Hospital and Research (SKMHR&C), Pakistan

³ The Information System for Research in Primary Care (SIDIAP), Spain

⁴ Centre for Statistics in Medicine, University of Oxford, UK.

The COVID-19 pandemic highlighted need for rapid, reliable, **representative** evidence generation



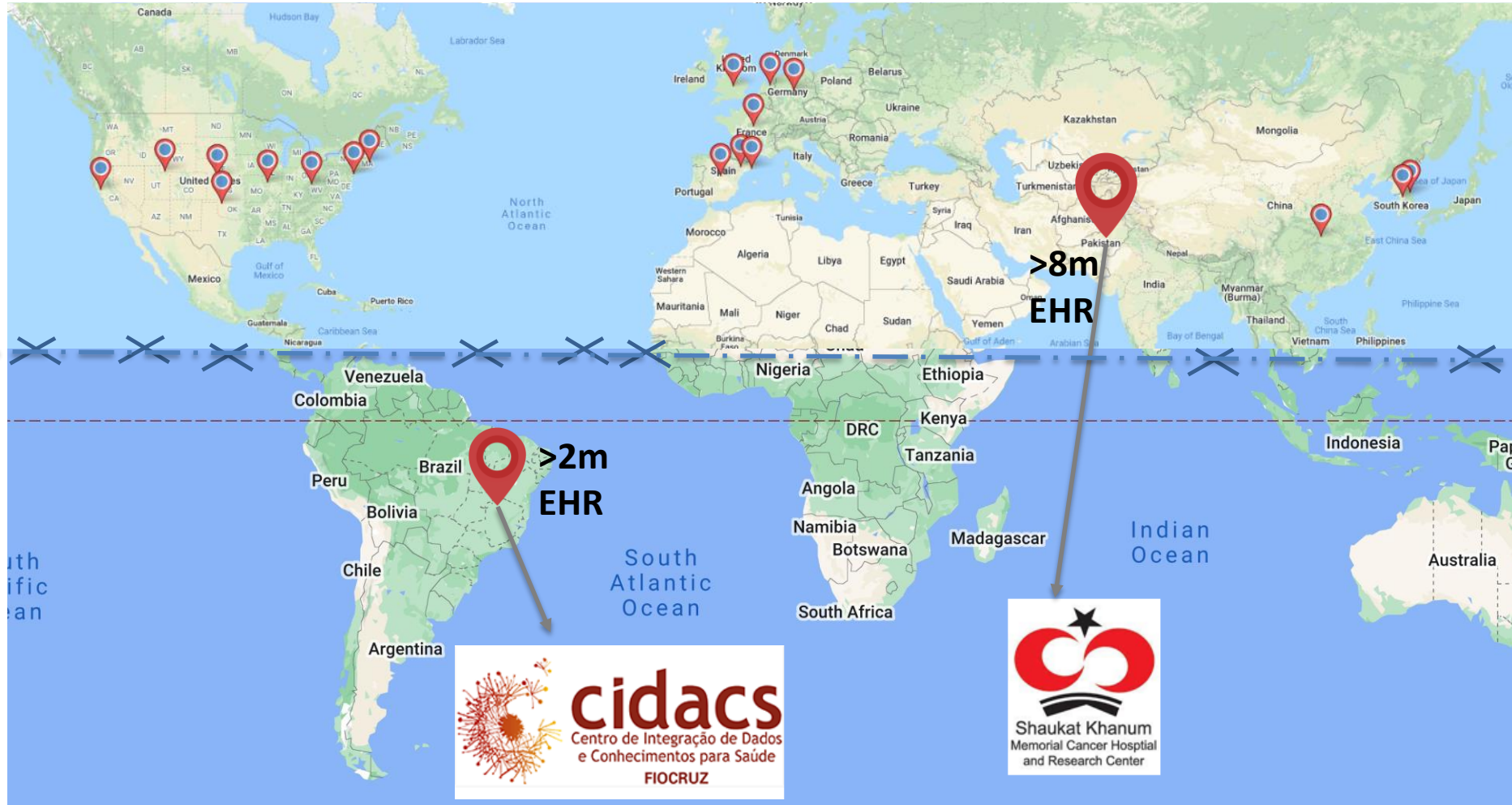
The OHDSI COVID-19 Data Network



- Rapid
- Reliable
- Representative?

USA (8)	EUROPE (7)	ASIA-PACIFIC (3)
Premier (National – Hospital Billing)	CPRD (UK – Electronic Health Records)	HIRA (South Korea – Administrative Claims)
HealthVerity (Claims linked to diagnostic testing)	SIDIAP (Spain – Electronic Health Records)	DCMC (South Korea – Electronic Health Records)
Optum EHR (National – Electronic Health Records)	SIDIAP-H (Spain – EHR hospital linkage)	Nanfeng Hospital (China – Electronic Medical Records)
IQVIA Open Claims (National – Administrative Claims)	HM Hospitales (Spain – Hospital Billing)	Together, OHDSI has studied (to date): <ul style="list-style-type: none"> • >4.5m patients tested for SAR-COV-2 • >1.2m patients diagnosed or tested positive for COVID-19 • >250k hospitalized for COVID-19
Department of Veterans Affairs (National – Electronic Health Records)	ICPI (Netherlands – Electronic Health Records)	
Stanford University (CA – Electronic Health Records)	LPD France (France – Electronic Health Records)	
Tufts University (MA – Electronic Health Records)	Germany DA (Germany – Electronic Health Records)	
Columbia University (NY – Electronic Health Records)		

The OHDSI COVID-19 Data Network



Data Partners – Pakistan

- Data source:
 - De-identified electronic health records
- Period:
 - 1994 – 2022 (ongoing)
- Unique records:
 - 8.3 million individuals
- Regional COVID-19 hub

SKMHR&C Hospital Network in South Asia



3 Hospitals



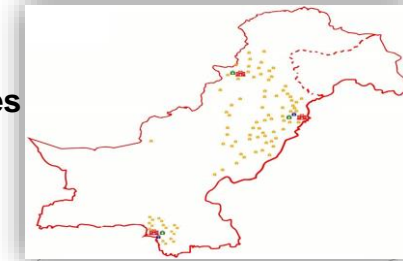
3 Diagnostic Centres



3 Walk-in Clinics



180 Labs in 70 cities



Country Facts

- Population: 220 million
- 5 Provinces

shutterstock.com · 234985177

Clinical activity snapshot (2021)



12,018

New Registrations



210,667

OPD Visits



14,840

Admissions



63,725

Chemotherapy Sessions



19,128

Surgical Procedures



77,809

Radiotherapy Sessions



198,393

Imaging Studies



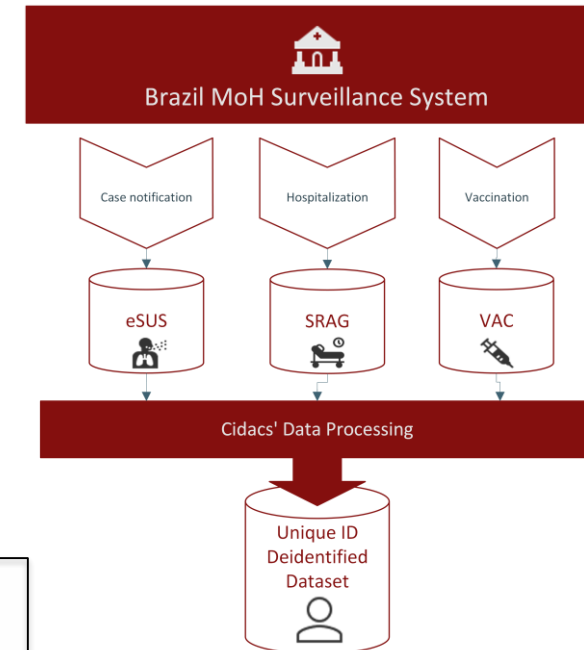
6,277,572

Pathology Tests

Data Partners – Brazil

- Data source/s:
 - Brazil Ministry of Health Influenza Surveillance System (SIVEP-Gripe)
- Period
 - 2020 – 2022 (ongoing)
- COVID-19 records:
 - 2.6 million individuals
 - ~67,000 hospitalisations
 - ~27,000 deaths

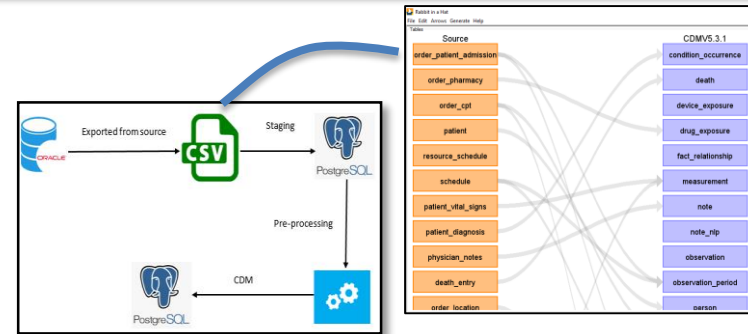
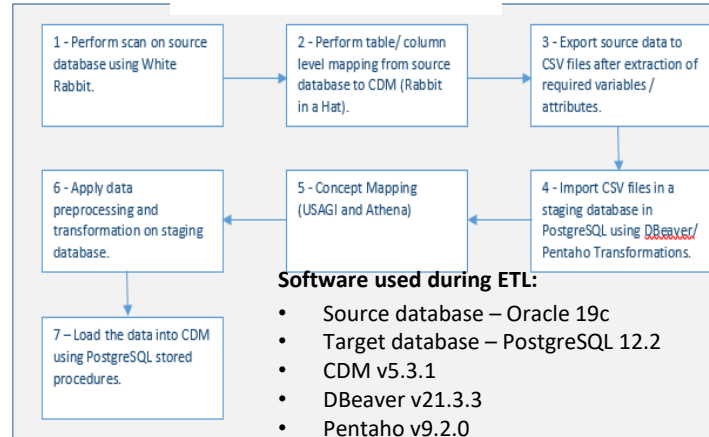
CIDACS-FIOCRUZ COVID-19 Data Platform



Harmonisation to OMOP

ETL Process

Pakistan



Mapping in Numbers

- >100K (source) to 108K (CDM) concepts
- >33M measurements
- >2M procedures
- >600K observations
- <1% missing matching concepts

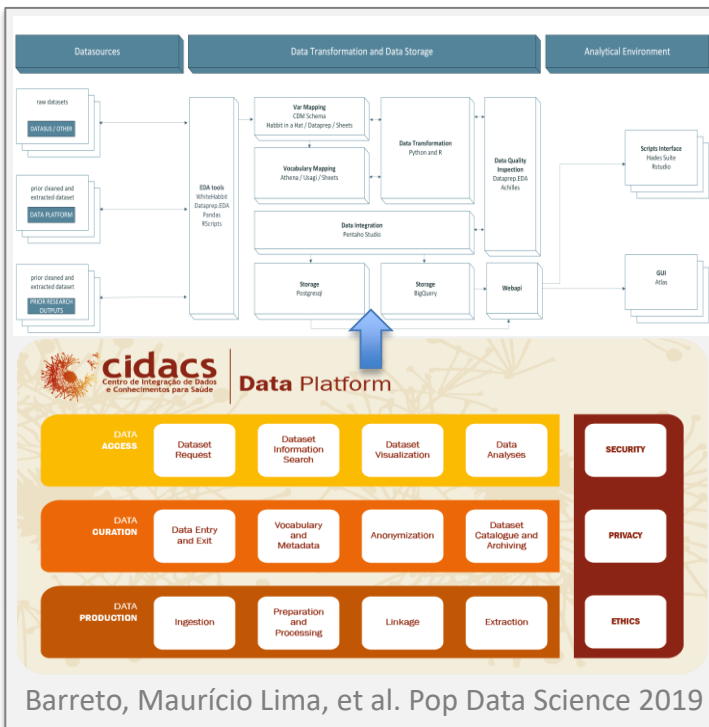
DATA QUALITY ASSESSMENT

HOSPITAL INFORMATION SYSTEM

DataQualityDashboard Version: 1.0.0
Results generated at 2022-02-15 12:05:02 in 1 hours

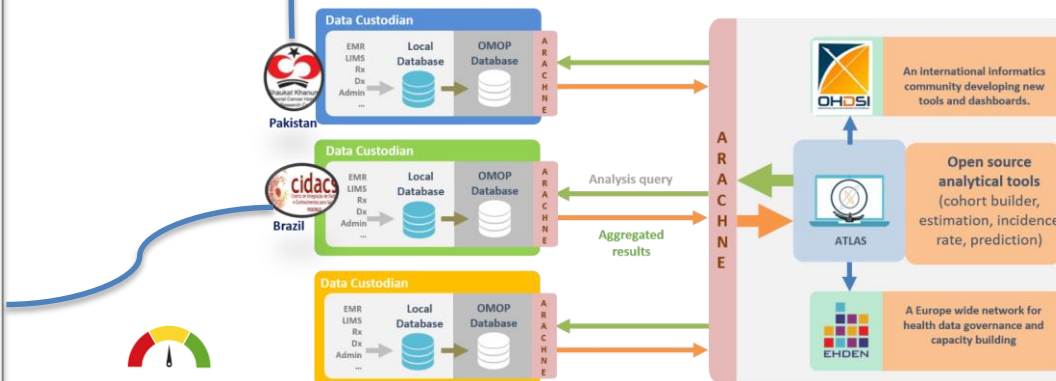
	Verification				Validation				Total			
	Pass	Fail	Total	% Pass	Pass	Fail	Total	% Pass	Pass	Fail	Total	% Pass
Possibility	2012	1	2013	100%	287	0	287	100%	2299	1	2300	99%
Conformance	675	6	681	99%	104	0	104	100%	779	6	785	99%
Completeness	380	6	386	98%	15	0	15	100%	395	6	401	99%
Total	3067	13	3080	100%	406	0	406	100%	3473	13	3486	100%

Brazil



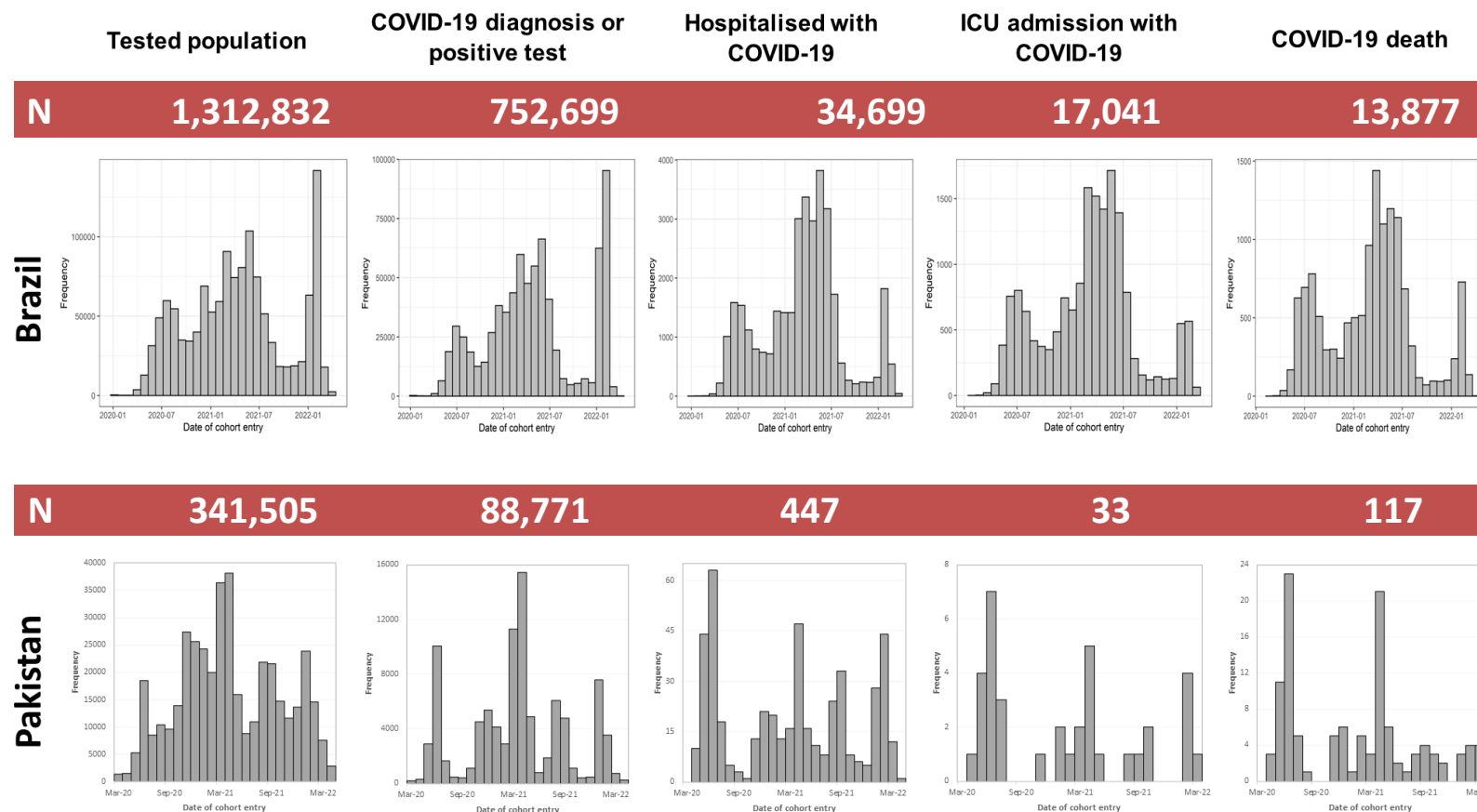
Mapping Complete

The OHDSI Federated Data Network Model



Mapping In Progress

COVID-19 cohorts – cases over time

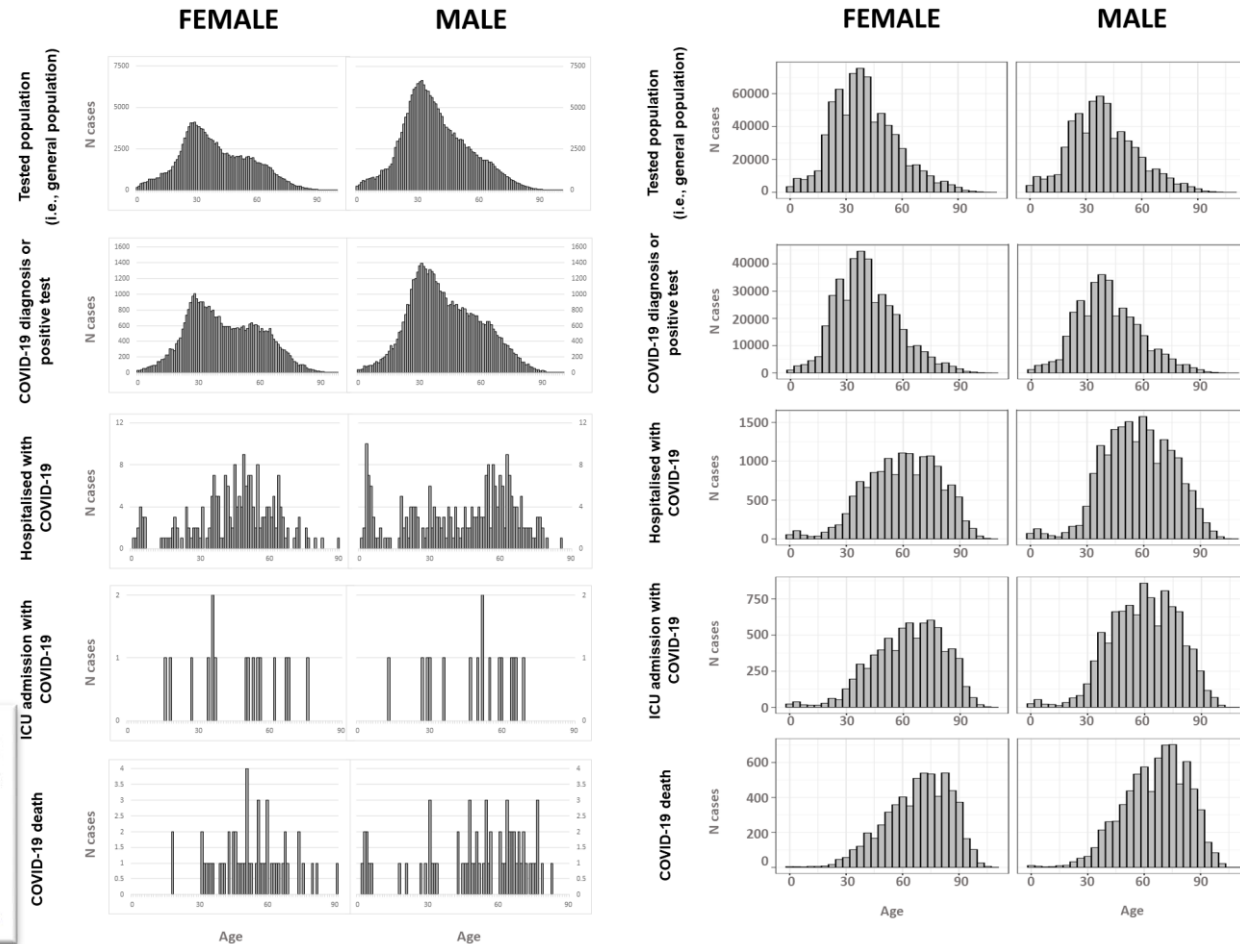


Distribution of cases over time (Jan/March 2020 – April 2022)



COVID-19 cohorts – baseline characteristics

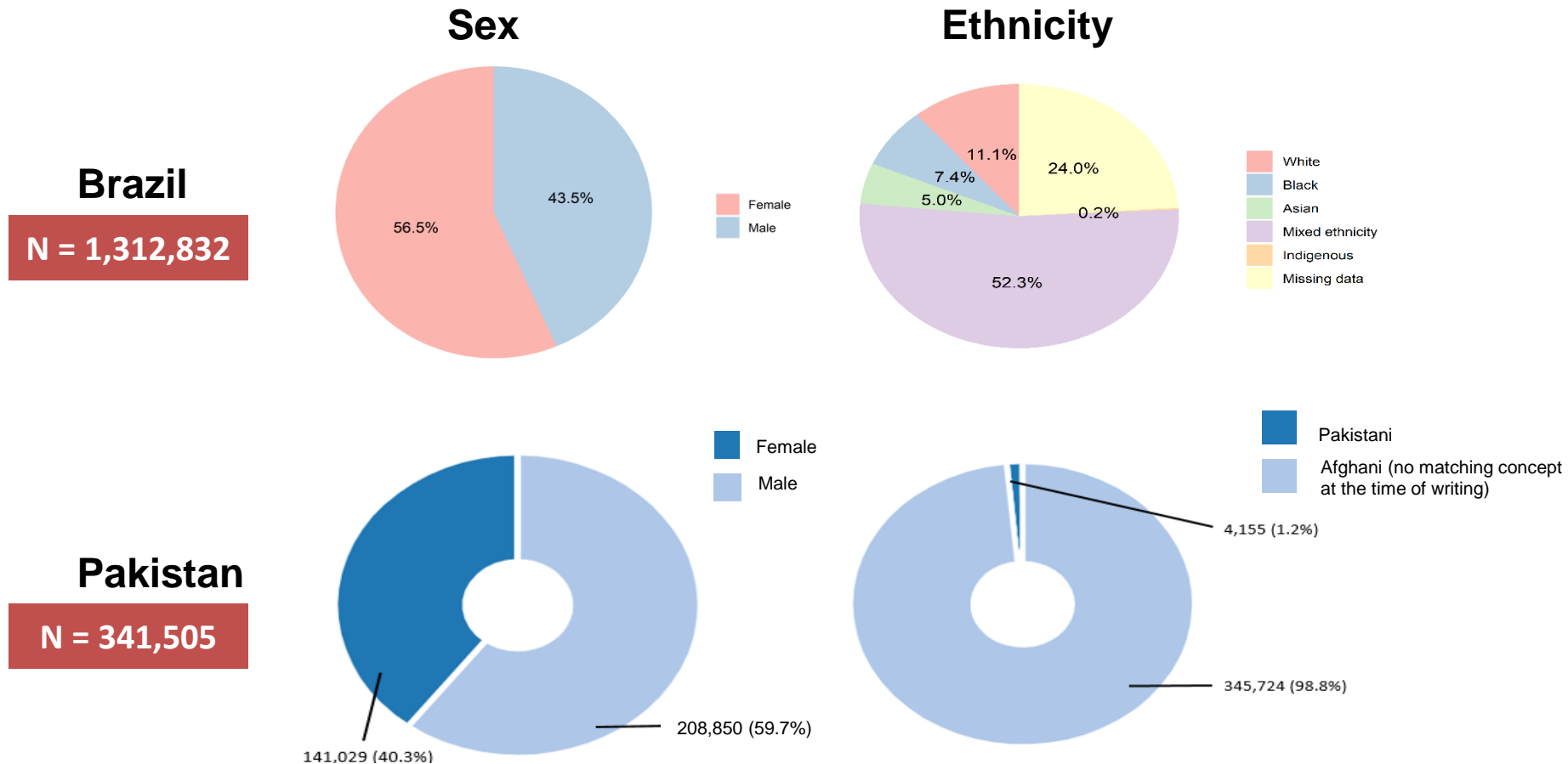
- COVID-19 outcomes were more severe in men, elderly, and those with co-morbidities



Distribution of cases by age and sex in each cohort



COVID-19 cohorts – baseline characteristics



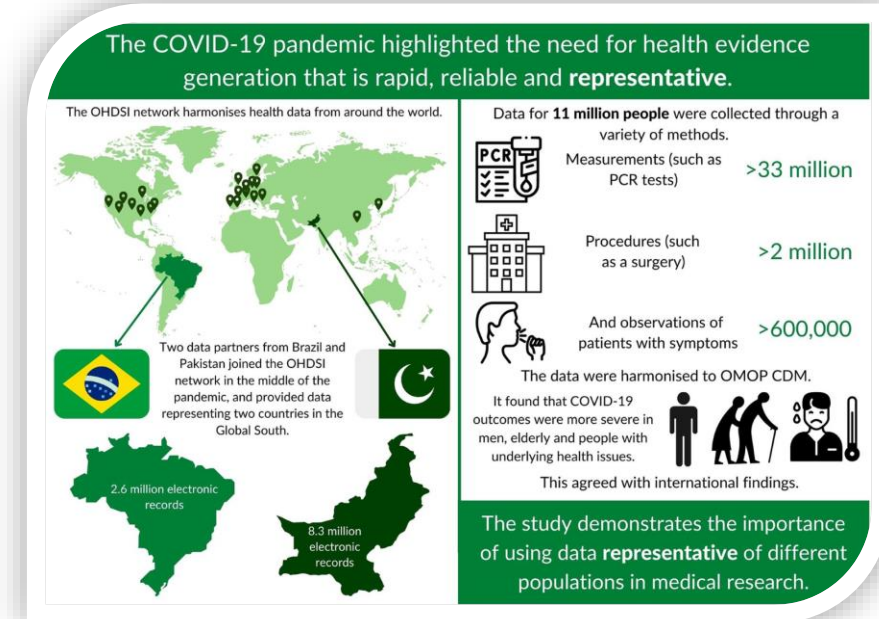
General population tested for COVID-19:

- Age: median (IQR) was 36 (25 -75) and 38 (27 - 50) for Pakistan and Brazil
- Sex: 45.5% and 55% were female in Pakistan and Brazil
- Ethnicity/race: 1.2% Pakistan individuals had “Afghan” ethnicity. In Brazil, 52.3% had “Mixed” ethnicity.




Summary

- Two health databases covering 8.3 million people from Pakistan and 2.6 million people from Bahia, Brazil were analysed.
- 109,504 (Pakistan) and 921 (Brazil) medical concepts were harmonised to OMOP CDM.
- 341,505 (4.1%) people in the Pakistan dataset and 1,312,832 (49.2%) people in the Brazilian dataset tested for COVID-19 between 1st Jan 2020 and 30 April 2022.
- In agreement with international findings, COVID-19 outcomes were more severe in men, elderly, and those with underlying health conditions.
- This proof-of-concept study demonstrates potential for OMOP-harmonised data from under-represented regions for global knowledge mobilisation and clinical translation for timely response to healthcare needs in pandemics and beyond.



JAMIA
A SCHOLARLY JOURNAL OF INFORMATICS IN HEALTH AND BIOMEDICINE

JOURNAL ARTICLE

Integrating real-world data from Brazil and Pakistan into the OMOP common data model and standardized health analytics framework to characterize COVID-19 in the Global South 

Elzo Pereira Pinto Junior, Priscilla Normando, Renzo Flores-Ortiz, Muhammad Usman Afzal, Muhammad Asaad Jamil, Sergio Fernandez Bertolin, Vinícius de Araújo Oliveira, Valentina Martufi, Fernanda de Sousa, Amir Bashir ... Show more



What's next

- COVID-19 – variant and vaccine surveillance study
- Communicable, NCDs
 - Cancer (**OHDSI Oncology WG**)
- Environment, equity, and artificial intelligence
 - Social deprivation dashboard (**OHDSI GIS WG, OHDSI Equity WG**)
- Data science ecosystem
 - Capacity building
 - Data re-use projects
 - Data governance



Thank You

Acknowledgements



OMOP Mapping of Real-World Data From Brazil & Pakistan Towards Management of COVID-19 In the Global South

Authors

Elzo Pereira Pinto Junior¹, Priscilla Normando¹, Renzo Flores-Ortiz¹, Muhammad Usman Afzal², Muhammad Asaad Jamil², Sergio Fernandez Bertolin³, Vinícius de Araújo Oliveira¹, Valentina Martufi¹, Edward Burn⁴, Maria Yury Ichihara¹, Maurício L. Barreto¹, Talita Duarte Salles³, Daniel Prieto-Alhambra⁴, Haroon Hafeez², Sara Khalid⁴

¹ The Center for Data and Knowledge Integration for Health (CIDACS – FIOCRUZ), Brazil

² Shaukat Khanum Memorial Cancer Hospital and Research (SKMHR&C), Pakistan

³ The Information System for Research in Primary Care (SIDIAP), Spain

⁴ Centre for Statistics in Medicine, University of Oxford, UK.