Harmonizing Multi-Site Electronic Health Records Data for Critical Care Comparative Effectiveness Studies:

How do we ensure that data quality is equivalent to traditional clinical trials?

Annie Simpson PhD, Daniel Brinton PhD, Kit Simpson DrPH,
Dee Ford MD, MSCR, Katie Kirchoff, MSHI, Andrew Goodwin, MD

MUSC Medical University of South Carolina

Changing What's Possible | MUSC.edu

Conflicts and Funding

- None of the authors has any conflicts of interest to disclose.

Work made possible by the following funding sources:

1. This project was supported, in part, by the National Center for Advancing Translational Sciences of the National Institutes of Health under Grant Number UL1 TR001450. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health, as a pilot award via the Carolina's Collaborative group of CTSA's (CCR31803).
2. Health Resources and Services Administration (HRSA) of the U.S. Department of Health and Human Services (HHS) as part of the National Telehealth Center of Excellence Award (U66 RH31458-01-00). The contents are those of the author(s) and do not necessarily represent the official views of, nor an endorsement, by HRSA, HHS or the U.S. Government.

Clinical Problem Statement

- Patients with ventilator dependent respiratory failure (VDRF) have the greatest mortality, morbidity, and cost of any sub-group of intensive care unit (ICU) patients
- Decades of research in this patient population have identified multiple strategies that can improve the quality of care, including optimizing 'settings' chosen for the ventilator (Lung Protective Ventilation-LPV)
- Recent research has found that LPV remains grossly underutilized internationally and thus patients with VDRF are needlessly experiencing low quality care
- GOALS: Improve patient survival and lower cost of care in this population

Motivating Early Work

- Spent significant time to learn how to pull EHR data on ventilated subjects in our single hospital with four different ICU's
  - Required multiple data pulls
  - Examination of the health record data entry pages
  - Consultation with attendings, fellows, and nurses to find flow sheet data

- Examination of raw data
  - To define the phenotype
  - To define "within ventilator guidelines"
  - To estimate if our definition of guideline-based care is associated with in-hospital mortality

- Wrote an R01 and then U01 for multi-site study to test if tele-ICU ventilator management program improves outcomes
  - Reviewers: You haven't proven your phenotype is valid or that you can do multi-site EHR data extraction/management

Technical/Research Problem Statement

- Increasingly, clinical trials in the ICU are incorporating pragmatic designs in which data capture occurs via EHR extraction as opposed to traditional real-time data extraction

- This enables data acquisition in shorter timeframes using less manpower with the net result of robust datasets at a fraction of the cost of traditional large prospective analyses.

- Pragmatic trials require that investigators:
  1. Define and extract valid populations
  2. Extract all key variables
  3. Properly harmonize data from multiple centers

Study Aims

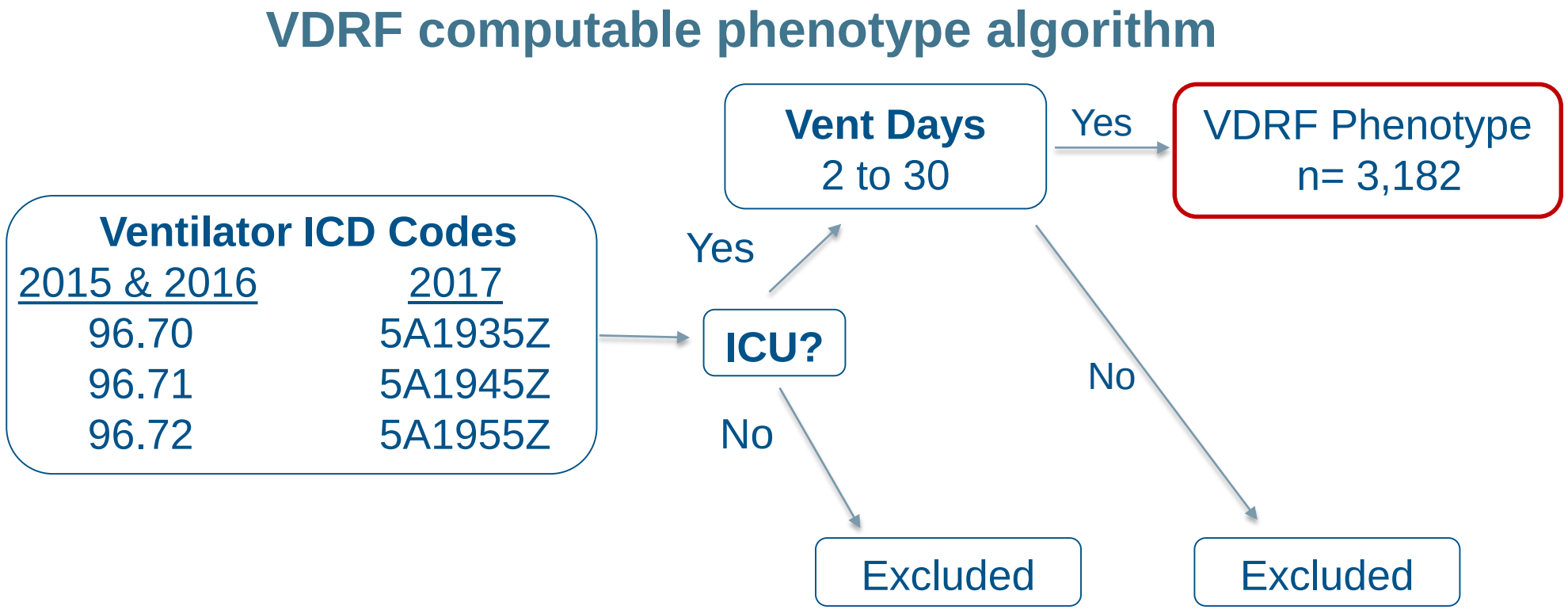
Aim 1: Construct and validate a VDRF computable phenotype at partner institutions in the Carolinas Collaborative (*CC*; MUSC, UNC, Wake Forest)

- Sub-aim 1a: Extract, merge, and clean data from the CC CDM and partner institutions' EHR to create a VDRF computable phenotype at each institution
- Sub-aim 1b: Validate each site's VDRF computable phenotype via chart abstraction

Aim 2: Harmonize the VDRF datasets from each institution and assess data quality for estimation of LPV adherence

Challenge: Do this in a 1-year pilot… for only $25k per site

Computable “e-phenotype” (preliminary data)



Data Sources and Elements

Data Sources

1. Carolina's Collaborative Clinical Data Warehouse (CDW, our Regional PCORNet)
2. Each institution's clinical Research Data Warehouse
3. Each institution's EHR (Clarity within EPIC EHR)

Key Elements for Valid Phenotype

ICU Indicator

Ventilator ICD Code(s)

Ventilator start and end date/time (including multiple episodes)

Patient Identification

- Planned to find via CDW and link to EHR data
- Actually had to find via EHR and link to CDQ

PCORNet CDW Planned Data Elements

Demographic
Age (at encounter)
Ethnicity
Race
Sex/Gender
Vital Status

Diagnosis
ICD-9-CM/ICD-10-CM
Diagnosis Date
Diagnosis Modifier (e.g., Admitting, Discharge)

Encounter
Discharge Disposition
Discharge Status
Encounter Type
Payor
Payor Modifier (i.e., primary, secondary)
Length of Stay
Admission Date
Discharge Date

Laboratory (Labs use a coding system called LOINC. A data analyst can help identify LOINC codes for needed labs at the data request stage)

Lab Tests (platelet, bilirubin, creatinine, 24hr Urine)
Lab Collection Date

Medications (Medications use a coding system called RxNorm. A data analyst can help identify RxNorm codes for needed medications at the data request stage)

Ordered Medications (vasopressors: i.e. dobutamine, dopamine, epinephrine, norepinephrine, phenylephrine)
Administration Start Date/time
Administration End Date/time

Procedure
CPT-4/HCPCS (i.e., HCPCS Level I, II)
ICD-9-PCS/ICD-10-PCS
Procedure Date

Vitals
Height
Weight
Vital Measurement Date/time

Additional Data to be added from local CDM or EHR

Ventilator Setting (Value, Type, Date/Time)
Ventilator Start Date/Time
Ventilator End Date/Time
ICU Admission Date/Time
ICU Discharge Date/Time
Glasgow Coma Scale

Actual Essential Data Elements

CDW | EHR | BOTH*

*Site Specific

SOFA Score

Labs

- Creatinine
- PaO2
- FiO2
- Platelets
- Bilirubin

Flowsheets

- MAP
- Glasgow Coma Scale score
- 24hr Urine

Medications (Vasopressors)

- Dobutamine
- Dopamine
- Epinephrine
- Norepenephrine
- Phenylephrine

Vitals (BMI)

- Height
- Weight

ICU Stay

- Vent times
- Vent Settings

Hospital Stay

- Admission Date
- Discharge Date
- Discharge Disposition
- Diagnosis Codes (ICD 9/10)
- Procedure Codes (HCPCS/CPT)

Demographics

- Age
- Sex
- Race
- Ethnicity

MUSC Validation Results

Table of EHR Inclusion_Met by RedCap Inclusion_Met			
EHR Inclusion Met	RedCap Inclusion Met		
Frequency Row Pct	Yes	No	Total
Yes	31 88.57	4 11.43	35
No	3 8.57	32 91.43	35
Total	34	36	70

Sensitivity and Specificity				
Statistic	Estimate	Standard Error	95% Confidence Limits	
Sensitivity	0.9118	0.0486	0.8164	1.0000
Specificity	0.8889	0.0524	0.7862	0.9915
Positive Predictive Value	0.8857	0.0538	0.7803	0.9911
Negative Predictive Value	0.9143	0.0473	0.8215	1.0000

5 Data Entry Errors on chart review (actions)

- 2 vent dates had wrong year entered (fix data entry error)
- 1 vent stop date was missing/not entered (patient discharged on vent, updated instructions discharge date=end date)
- 1 missing episode within the same ICU stay (Recap data entry missed second episode)
- 1 missing episodes - 3 discrete episodes on vent within the same ICU stay, all episodes not entered in Redcap

2 Rounding Errors – EHR vent date data includes date/hour/minute/second, Chart review only entered date/rounded hour

- 2 patient episodes rounded to more than two days on vent when they were just under two days when including minutes. (updated instructions to round to the nearest half hour, rather than hour, with also calculate this way during EHR data cleaning)

Other Data Point Validation Results

Measure	Fisher Transformed Pearson Correlation Coefficient for test/retest reliability	Fisher Transformed 95% CI and p-values
Lowest tidal volume	0.80	(0.69,0.87) p<.0001
Hospital LOS	0.998	(0.997,0.999) p<.0001
ICU LOS	0.99	(0.991,0.996) p<.0001
SOFA	0.94	(0.87,0.97) p<.0001

SOFA

Sequential Organ Failure Assessment

System	Assigned Score 0	1	2	3	4
Respiration PaO_2/FiO_2, mmHg	≥ 400	< 400	< 300	< 200[a]	< 100[a]
Coagulation Platelets x $10^3/mm^3$	≥ 150	< 150	< 100	< 50	< 20
Hepatic Bilirubin, mg/dl (µmol/l)	< 1.2 (< 20)	1.2 – 1.9 (20 – 32)	2.0 – 5.9 (33 – 101)	6.0 – 11.9 (102 – 204)	≥ 12.0 (> 204)
Cardiovascular Hypotension	MAP ≥ 70 mm Hg	MAP < 70 mm Hg	Dopamine ≤ 5 or Dobutamine (any dose)[b]	Dopamine > 5 Epinephrine ≤ 0.1 Norepinephrine ≤ 0.1 Phenylephrine[c] ≤ 0.22	Dopamine > 15 Epinephrine > 0.1 Norepinephrine > 0.1 Phenylephrine[c] > 0.22
Central Nervous System Glasgow Coma Scale Score	15	13 – 14	10 – 12	6 – 9	< 6
Renal Creatinine, mg/dl (µmol/l)	< 1.2 (< 110)	1.2 – 1.9 (110 – 170)	2.0 – 3.4 (171 – 299)	3.5 – 4.9 (300 – 440)	> 5.0 (> 440)
Urine output, ml/day				< 500	< 200

[a] With respiratory support
[b] Administered for at least 1 hour
[c] Phenylephrine added by Knox et al. to list of vasopressors according to standard equivalency

MAP = Mean Arterial Pressure
PaO_2/FiO_2 also known as the Horowitz Index

Vincent JL, Moreno R, Takala J, Willatts S, De Mendonca A, Bruining H, Reinhart CK, Suter PM, Thijs LG. (1996). The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure. On behalf of the Working Group on Sepsis-Related Problems of the European Society of Intensive Care Medicine. *Intensive Care Med, 22*(7):707-710.

MUSC Site Lessons Learned

Multiple data pulls required from multiple data sources

- 2 pulls even after preliminary study
- IT had to pull some items from the EHR that were supposed to be in the CC CDW because they were either had large percentage of missing and/or did not include time (for date/time values)

Age – found twice in data. Careful how defined. Pulled both as age at time of encounter and age at time of data extraction (error ranged: 0-4 years)

LOINC (Laboratory indicator codes) – SAS wants to import as times (e.g. 19255-9 becomes 19255:09).

FiO2 LOINC – initially extracted, and only (471/4663 = 10.1%) of patients had an FiO2. Only 1 LOINC was used for extraction, however there are 4 LOINCs for FiO2 and 10 for Horowitz index (PaO2/FiO2).

Platelets – Two units were presented, written as K/CUMM and X10E3/UL. These both meant 1000/*some unit of measure*. Determined these were *1,000/mm^3 and *1,000/µL*. Quick conversion showed these are equivalent units (1.0 mm^3 = 1.0 µL).

MAP – Values that made no sense (e.g. negative values [as low as -8.0] and really high values [109,109]). The extremes/nonsense were removed (deleted if < 0 or > 200 mm Hg).

Vasopressors (from CDW) – All med start times and stop times are the same moment. - needed to re-pull from EHR

Data Pull Lessons from Other Sites

Labs – Many times did not contain times (which are essential). Often missing labs essential to SOFA score creation (e.g. PaO2, FiO2)

Vasopressor administration – Many times did not contain times (only dates). Individual entries were listed, with an action (i.e. *Medication administered, Dose change, Dose verification, Medication stopped*)

Null values for height/weight – Likely a site-specific data extraction issue

Patient & Encounter IDs needed – Patients may have multiple ICU stays, so both are needed (but not always present)

Large CSVs may be corrupted in Microsoft Excel – When examining CSVs for data quality in Microsoft Excel prior to import into SAS, Excel corrupted date/time fields making them unreadable by SAS

Data Pull Lessons from Shared Research Warehousing

Largest Lesson Learned:

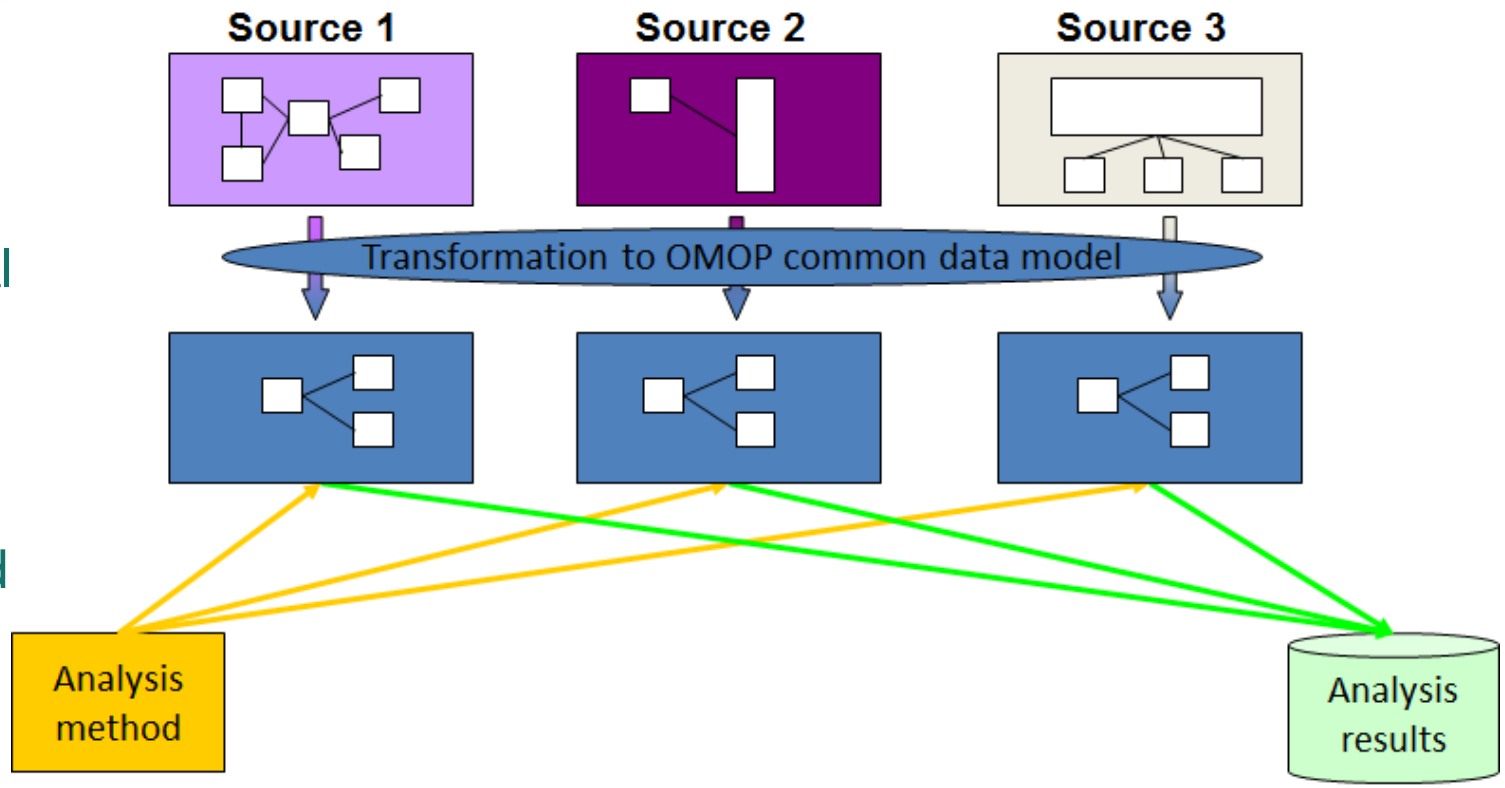
Data load from EHR to site-specific and CC-regional CDW, many Labs and Medications have not been mapped to shared coding standards.

Labs: LOINC

Meds: RXNorm and/or NDC

Examining Best Methods for Conversion of Data Elements to a Common Data Model

- Using OMOP (Observational Medical Outcomes Partnership) as a open source base common data model
- OMOP – very flexible with continual improvement by developers at UCSF (and international data consortium)
- Standardized structured query language (SQL) queries are shared in a common open-source repository and updated regularly. All data documentation is freely available online.
- PCORNet does not use OMOP



Source: https://www.ohdsi.org/data-standardization/the-common-data-model/

Aim 2 – (Not Completed)

- Each site's data will undergo a series of data quality checks similar to previously published studies of PEDSnet processes that will follow the "**Conformance, Completeness, and Plausibility**" theoretical framework set forth by Kahn and colleagues, based on the "fit for use" paradigm

- Conformance will include checks for consistency of cohort definition and source values
- Completeness will include missing data checks (including type of missingness, i.e. missing at random, not missing at random)
- Plausibility will include field checks for unexpected facts, events and date fields (temporal checks), numerical outliers, unexpected frequencies, and categories

- Our hypothesis is that the harmonized dataset can be analyzed with our previously developed multivariable model to assess the effect of LPV baseline adherence on survival, with less than 15% of patients from any site with missing essential values

Kahn, M.G., Callahan, T.J., Barnard, J., Bauck, A.E., Brown, J., Davidson, B.N., Estiri, H., Goerg, C., Holve, E., Johnson, S.G., & Liaw, S.T., (2016). A harmonized data quality assessment terminology and framework for the secondary use of electronic health record data. *Egems*, 4(1).

Contact Info:

Annie Simpson, PhD
simpsona@musc.edu

Thank You