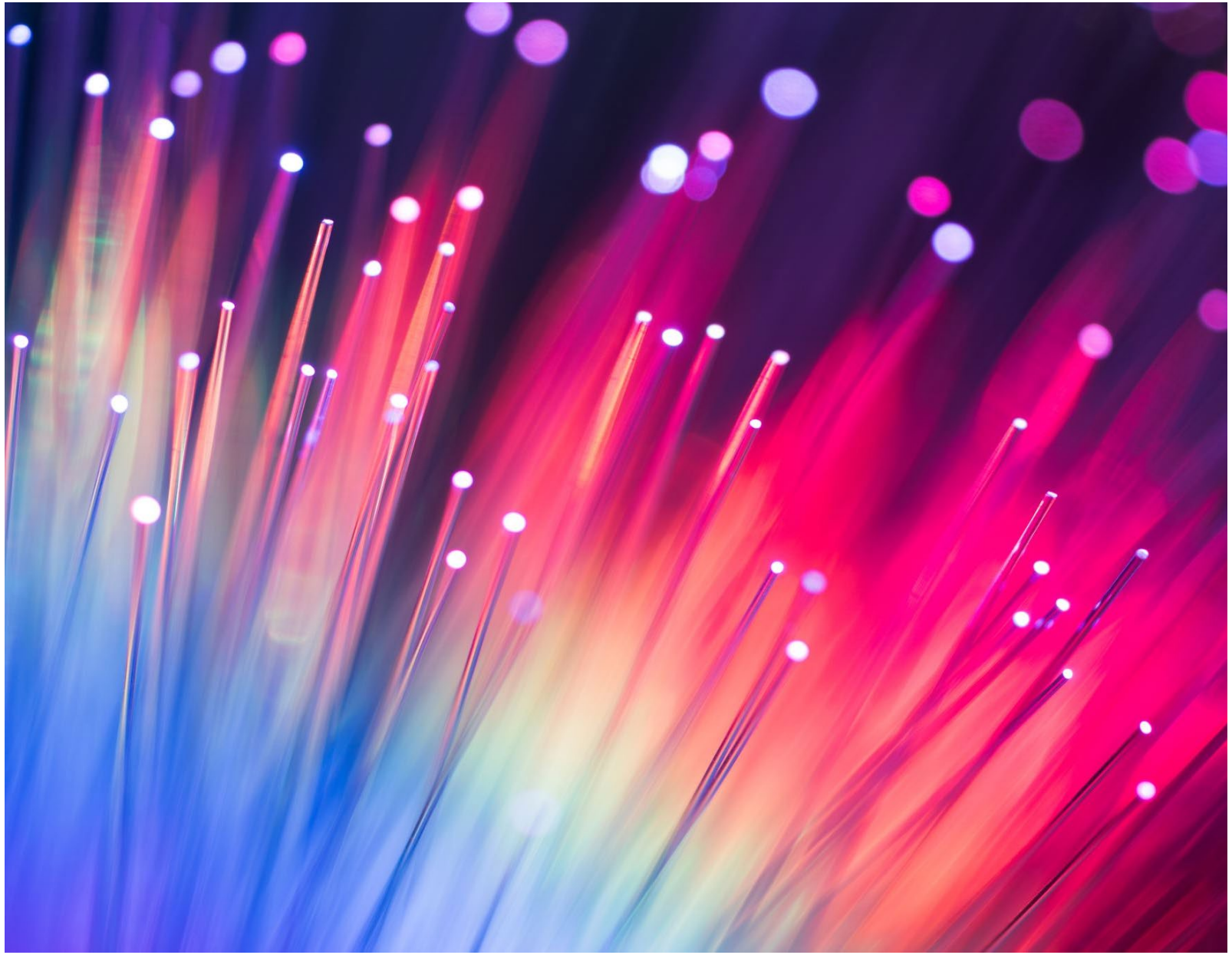


FHIR-BASED COMPUTATIONAL PHENOTYPING

ZAINAB APALARA

CHIP 490.297

APRIL 2022

AGENDA

- Problem
- Drivers of the Problem
- Solution
- Workflow
- Main Takeaway
- References

PROBLEM

- Maternal morbidity is associated with **worse health outcomes** of mothers and infants.
- Currently, the associations between maternal morbidity and some patient characteristics are **unknown** and/or **disputable**.
- The problem is **exacerbated** as a result of **racial disparities**.
- **We need more efficient methods to determine which patients are more susceptible.**

DRIVERS OF THE PROBLEM

- **Fragmented and missing data**
- The **lack of interoperability** between EHR systems

SOLUTION

Using **FHIR-based computational phenotyping** based on structured and unstructured **patient data** from **multiple healthcare systems**

WE CAN identify these high-risk patients.

Fragmented and missing data will be mitigated by using patient data from multiple healthcare systems

Interoperability will be improved by exporting and using patient data in FHIR format.

WORKFLOW: BIRD'S EYE VIEW

Different EHR System data can be exported in FHIR format

1

Patient information (structured & unstructured) must be organized into 6 designated folders with FHIR-organized and formatted .json files: 1. patient and their demographics, 2. patients with their classification (i.e., maternal morbidity present or absent), 3. lab data, 4. medication data, 5. vitals data, and 6. unstructured clinical notes.

2

The organized information is placed into CLARK which is a machine-learning classification tool that creates computable phenotypes from structured and unstructured data

3

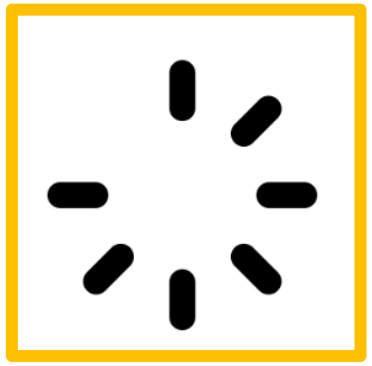
FHIR-based computational phenotypes are created and identify patients who will likely be susceptible to maternal morbidity based on their characteristics

WORKFLOW: CLARK

80:20

Create a training (labeled) and testing set (unlabeled)

a



Load training set into CLARK

b

For the **structured data**: simply select the features of interest (e.g., demographics, lab, vital)

For the **unstructured data**: create regular expressions to select the features of interest

(?i)bleed

c

Iteratively train the model (e.g., decision trees, support vector machine)

d

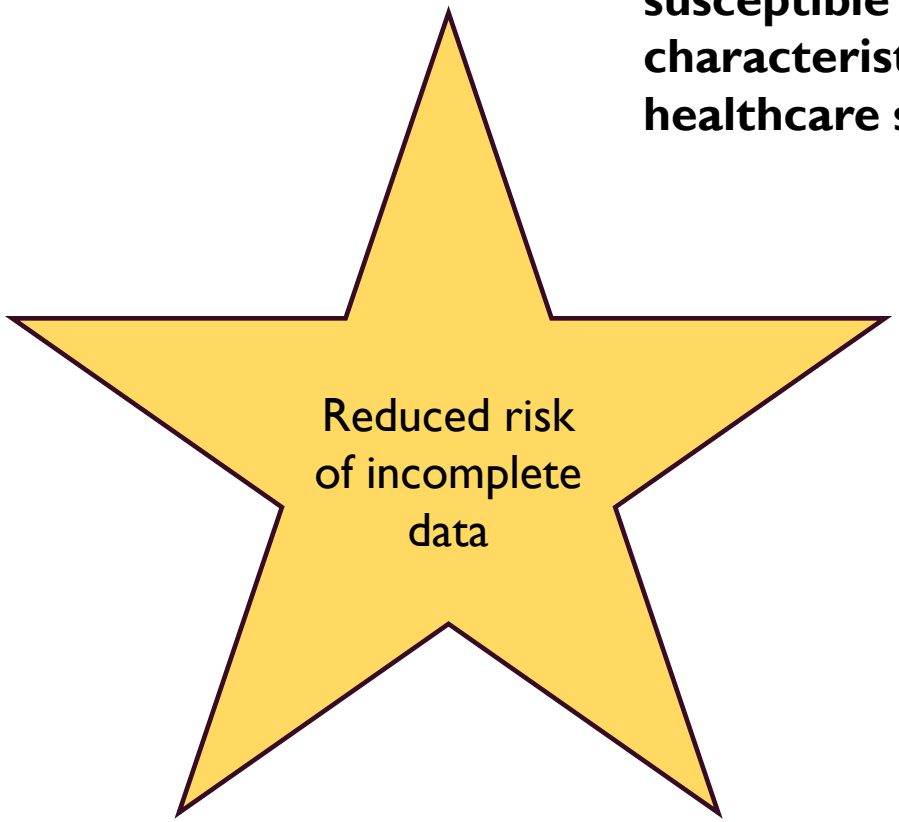
Evaluate the model's performance using the testing set

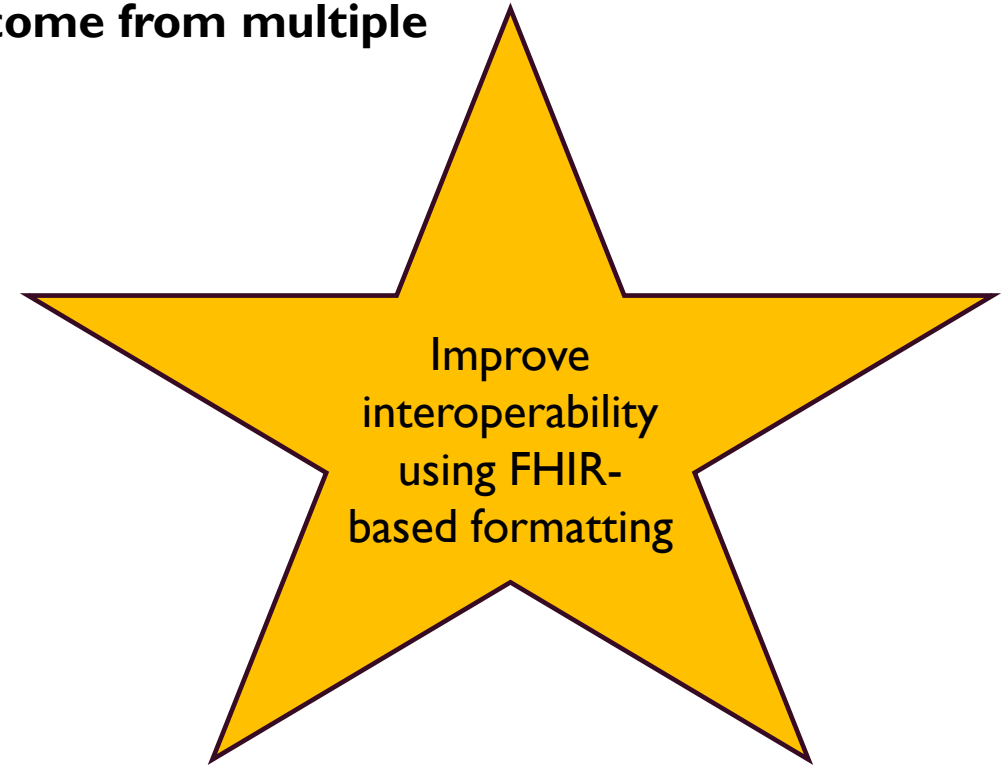
F1
Recall
Precision

MAIN TAKEAWAY

Problem: We need more **efficient methods** to determine which **patients are more susceptible**.

Solution: Create **FHIR-based computational phenotypes** (using **CLARK**) to identify **patients who will likely be susceptible** to maternal morbidity **based on their characteristics**. The **patient data will come from multiple healthcare systems.**





https://tracs.unc.edu/docs/clark-2-documentation.html

https://effectivehealthcare.ahrq.gov/products/maternal-morbidity-mortality/protocol

REFERENCES