

# An algorithm using electronic medical record data accurately identifies patients with unknown HIV status in a large urban healthcare system

Felsen UR<sup>1,2</sup>, Zingman BS<sup>1,2,3</sup>, Cunningham CO<sup>1,2,3</sup>, Bellin E<sup>1,2,3</sup>

<sup>1</sup>Albert Einstein College of Medicine; <sup>2</sup>Montefiore Medical Center; <sup>3</sup>Einstein-Montefiore Center for AIDS Research; Bronx, NY

## BACKGROUND

- Routine HIV testing is a critical strategy toward improving early HIV diagnosis and preventing transmission
- Effective HIV testing interventions require the ability to identify patients with unknown HIV status, however substantial logistical hurdles inhibit the identification of these patients
- Electronic medical records capture a wide variety of patient level data that may be useful for identifying patients with unknown HIV status
- There have been few large-scale efforts to operationalize EMR data to identify patients unknown HIV status

## OBJECTIVES

- To identify patients with unknown HIV status, we developed and validated an algorithm using commonly available EMR data

## METHODS

### Setting

- The Bronx, in New York City, is one of the epicenters of the domestic HIV epidemic and Montefiore Medical Center (MMC) is the largest healthcare provider in the Bronx
- MMC has an extensive electronic medical record (EMR) system that integrates clinical, laboratory, and administrative data from inpatient, outpatient, and Emergency Department settings beginning in 1997

### Development

- All HIV-related laboratory, billing, and problem list data available in the EMR were considered as candidate criteria for the algorithm. Criteria could comprise individual data elements or combine multiple elements (see Table 1)
- MMC's Clinical Looking Glass (CLG) software was used to query the EMR for patients fulfilling each of the candidate criteria
- The charts of random samples of patients fulfilling each criteria were reviewed for concordance with a gold-standard of chart review. The a priori definitions used for chart review are below:

Known HIV Status	HIV Infected	Any of the following: -WB + or VL+ -Documentation of HIV in chart -ART administered or prescribed (excluding postexposure prophylaxis)
	HIV Negative	Not HIV infected AND -Negative HIV screening test (ELISA or rapid)
Unknown HIV Status	?	Neither HIV infected or HIV negative

\*WB= western blot, VL=viral load, ART=antiretroviral therapy

- Final algorithm included criteria with greatest concordance

### Validation

- CLG was used to apply the final algorithm to the MMC population
- The charts of random samples of patient from each category of HIV status as identified by the algorithm were reviewed for concordance using the same gold-standard chart review process as above
- HIV infected category was oversampled for chart review
- Performance characteristics of the algorithm were calculated

## RESULTS

Table 1. Development: Included and Excluded candidate criteria

EMR Data	Included in Algorithm	Excluded from Algorithm
Lab	HIV Ab (rapid and ELISA) HIV VL HIV WB CD4 Count (concurrent w/ VL)	HIV Genotype HIV Phenotype HIV Tropism CD4 Count (alone)
Billing (ICD9) (Inpatient: x 1 or Outpatient: x 2)	042-044 (HIV or AIDS related) 079.53 (HIV, type 2) 795.78 (Serologic/Culture HIV) V08 (Asymptomatic HIV)	795.71 (Nonspecific evidence of HIV)
Problem List	Contains: "Human immunodeficiency Virus," "Acquired Immunodeficiency Syndrome," "HIV" or "AIDS"	"HIV counseling" "HIV exposure" "HIV infection in mother" "HIV complicating pregnancy"

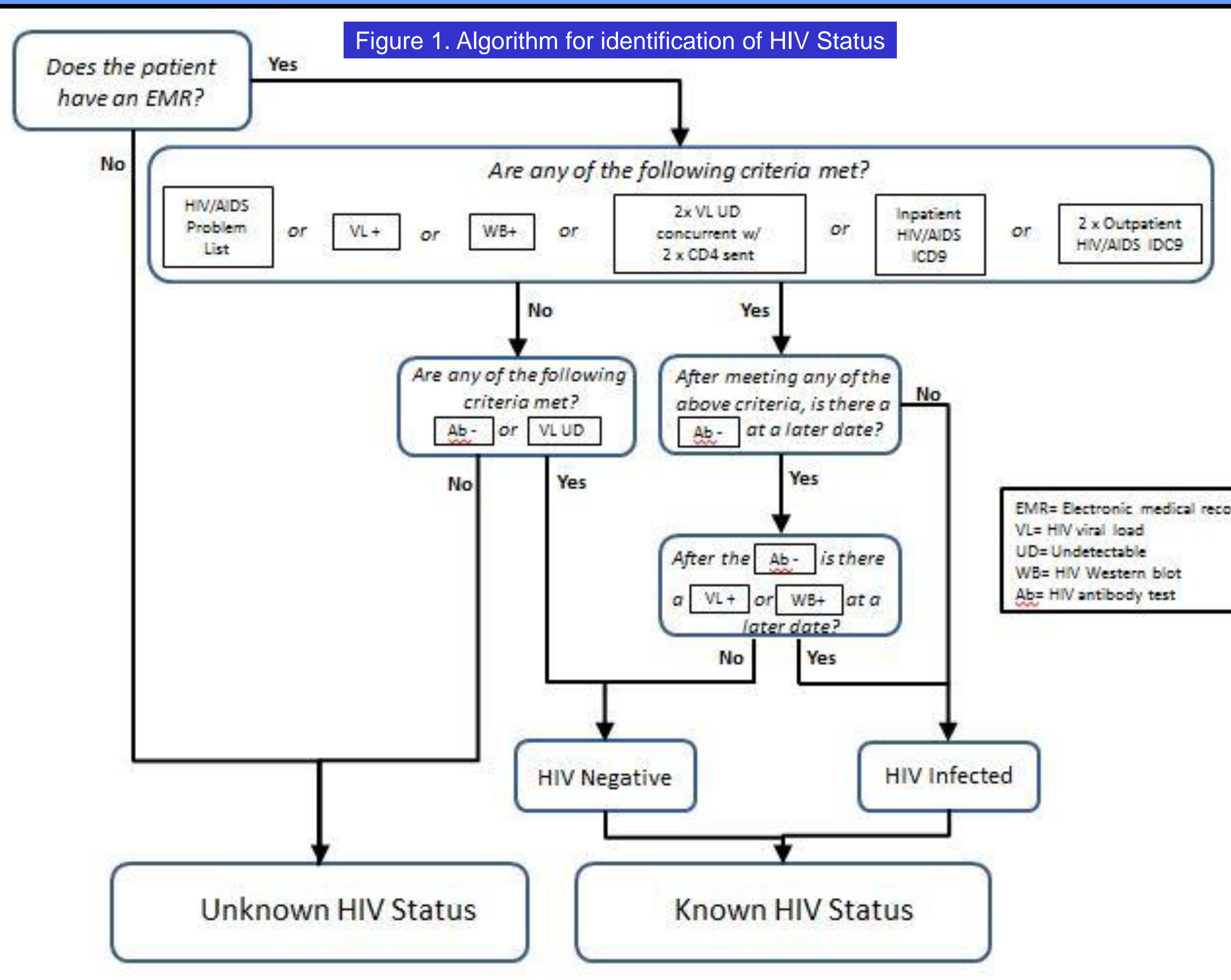


Table 2. Validation: Concordance of algorithm with chart review

Algorithm Results	Chart Review	
	Unknown HIV Status	Known HIV Status
Unknown HIV Status	50 ("True Unknown")	0 ("False Unknown")
Known HIV Status	15 ("False Known")	418 ("True Known")

Table 3. Performance characteristics of algorithm

Sensitivity	76.9% (64.8-86.5)
Specificity	100% (99.1-100)
Positive Predictive Value	100% (92.9-100)
Negative Predictive Value	96.5% (94.3-98.0)

## LIMITATIONS

- Antiretroviral data was not included in the algorithm
- "HIV negative" is a transient state
- Data for algorithm limited to a single healthcare system. Does not account for patient interactions outside of this system
- Results of this algorithm therefore represent **maximum** frequency of unknown HIV status

## CONCLUSIONS

- An algorithm using commonly available data from the EMR can accurately identify patients with an unknown HIV status in a large urban healthcare system
- The algorithm should be reproducible in other healthcare systems
- Potential application in diverse clinical and research settings:
  - Calculate baseline rates of unknown HIV status
  - Support planning of expanded HIV testing strategies
  - Monitor impact of new testing strategies over time
  - Integrate into EMR-based clinical decision support programs to identify patients that should be offered HIV testing

## NEXT STEPS

- Evaluation of antiretroviral data for inclusion in algorithm
- Comparison of algorithm results to self-report of HIV status
- Assessment of baseline prevalence and trends over time of unknown HIV status in inpatient, outpatient, and Emergency Department settings
- Monitor impact of expanded HIV testing strategy in MMC Emergency Departments on prevalence of unknown status
- Apply algorithm to identify risk factors for unknown HIV status
- Evaluate impact of integrating algorithm into automated, EMR-based prompt on rate of offer and uptake of HIV testing

## CONTACT

ufelsen@montefiore.org

This poster was supported in part by the Center for AIDS Research at Albert Einstein College of Medicine and Montefiore Medical Center funded by the National Institutes of Health (NIH AI-051519); and by the National Center for Research Resources (NCRR) and the National Center for Advancing Translational Sciences (NCATS), components of the NIH, through CTSA grant numbers UL1RR025750, KL2RR025749 and TL1RR025748. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the NIH.