



Published in final edited form as:

AIDS Care. 2014 October ; 26(10): 1318–1325. doi:10.1080/09540121.2014.911813.

Development of an electronic medical record-based algorithm to identify patients with unknown HIV status

Uriel R. Felsen, MD, MPH, MSc,

Division of Infectious Diseases, Montefiore Medical Center, 111 E 210th Street, Bronx, NY 10467, Tel: (718)920-8588, Fax: (718)405-0610

Eran Y. Bellin, MD,

Division of Infectious Diseases, Montefiore Medical Center, 3411 Wayne Avenue, Suite 4-H, Bronx, NY 10467, Tel: (917)371-4463, Fax: (718)920-2746

Chinazo O. Cunningham, MD, MSc, and

Division of General Internal Medicine, Montefiore Medical Center, 111 E. 210th Street, Bronx, NY 10467, Tel: (718)944-3860, Fax: (718)944-3841

Barry S. Zingman, MD

Division of Infectious Diseases, Montefiore Medical Center, 111 E 210th Street, Bronx, NY 10467, Tel: (718)920-2647, Fax: (718)405-0610

Uriel R. Felsen: ufelsen@gmail.com; Eran Y. Bellin: EBELLIN@emerginghealthit.com; Chinazo O. Cunningham: CCUNNING@montefiore.org; Barry S. Zingman: BZINGMAN@montefiore.org

Abstract

Individuals with unknown HIV status are at risk for undiagnosed HIV, but practical and reliable methods for identifying these individuals have not been described. We developed an algorithm to identify patients with unknown HIV status using data from the electronic medical record (EMR) of a large healthcare system. We developed EMR-based criteria to classify patients as having known status (HIV-positive or HIV-negative) or unknown status and applied these criteria to all patients seen in the affiliated healthcare system from 2008–2012. Performance characteristics of the algorithm for identifying patients with unknown HIV status were calculated by comparing a random sample of the algorithm's results to a reference standard medical record review. The algorithm classifies all patients as having either known or unknown HIV status. Its sensitivity and specificity for identifying patients with unknown status are 99.4% (95% CI: 96.5%–100%) and 95.2% (95% CI: 83.8%–99.4%), respectively; with positive and negative predictive values of 98.7% (95% CI: 95.5%–99.8%) and 97.6% (95% CI: 87.1%–99.1%), respectively. Using commonly available data from an EMR, our algorithm has high sensitivity and specificity for identifying patients with unknown HIV status. This algorithm may inform expanded HIV testing strategies aiming to test the untested.

Keywords

HIV/AIDS; Algorithm; Electronic Medical Record; Expanded HIV testing

INTRODUCTION

Fifty-five percent of Americans have never been tested for HIV (CDC, 2010), and individuals with unknown HIV status are at risk for undiagnosed HIV. Undiagnosed HIV, estimated as 18.1% of HIV infections in the US, remains a clinical and public health concern, and its prevalence has not substantially decreased in over a decade (CDC, 2012; Prejean et al., 2011). Individuals with undiagnosed HIV have increased risk for adverse clinical outcomes (Chadborn et al., 2006; Palella et al., 2003) and contribute disproportionately to HIV transmission (Marks et al., 2006). Risk-based HIV testing is linked to high rates of undiagnosed HIV and missed opportunities for earlier diagnosis (CDC, 2006; Jenkins et al., 2006). To address limitations of risk-based testing, in 2006 national recommendations and expert panels began calling for non-targeted expanded HIV testing for all individuals 13–64 years old at least once regardless of perceived risk, and annually if at high-risk (Branson et al., 2006; Chou et al., 2012; Moyer et al., 2013; Qaseem et al., 2009; White House Office of National AIDS Policy, 2010).

Identifying patients with unknown HIV status is critical for estimating the scope of undiagnosed HIV and estimating the impact of interventions. Unknown HIV status has primarily been ascertained using self-report of prior HIV testing (Merchant et al., 2009; Myers et al., 2012; Shuter et al., 1997). However, individuals frequently misunderstand whether they were HIV tested (Albrecht et al., 2012; Hutchinson et al., 2004; Jenness et al., 2009). Therefore, estimates of unknown HIV status based on self-report have questionable validity, and testing strategies relying on self-report may fail to reach a significant proportion of the untested.

Electronic medical records (EMRs) are conducive to automated queries and facilitate case-finding (Kawasumi et al., 2011; Navaneethan et al., 2011; Peissig et al., 2012; Szumski et al., 2009). Regarding HIV, EMRs have been used primarily to identify HIV-positive patients (Antoniou et al., 2011; Fasciano et al., 1998; Fultz et al., 2006). Whether EMR data can accurately identify patients with unknown HIV status remains an important question because unknown HIV status is not normally affirmed in EMRs as an explicit finding, but rather is inferred by the absence of data suggesting known HIV status either because the patient is HIV-positive or HIV-negative.

We describe the development of an EMR-based algorithm to identify patients with unknown HIV status in a large urban healthcare system and report the algorithm's performance characteristics. Elsewhere, we report the algorithm's application to ascertain prevalence of unknown HIV status and inform HIV testing strategies.

METHODS

Setting

Approximately 10% of those living with HIV/AIDS in the US live in New York City (NYC), and the Bronx is one of NYC's boroughs most profoundly impacted by HIV/AIDS (NYC Department of Health and Mental Hygiene [NYC DOHMH], 2012a). The Bronx has

an HIV prevalence of 1.7%, and has the highest HIV/AIDS-related mortality rate in NYC (NYC DOHMH, 2012b). Montefiore Medical Center (MMC) is the largest healthcare system in the Bronx and its patient population reflects that of the borough: 37% African American, 54% Hispanic/Latino, 51% have Medicaid, 13% uninsured, and 30% living below the poverty line. MMC has three adult hospitals and Emergency Departments (EDs) and over 50 outpatient sites. Together, these sites see 90,000 inpatient admissions, 300,000 ED visits, and 2.7 million outpatient visits annually (MMC, 2012). Clinical, laboratory, and administrative data from these sites are captured in an EMR that was introduced in 1997. We developed and applied our algorithm using health surveillance software that incorporates temporal and logical parameters to query a replicate of the MMC EMR (Clinical Looking Glass®; Emerging Health Information Technology, Yonkers, NY).

Algorithm Development and Assessment

To develop and assess our algorithm we defined a reference standard against which we compared the algorithm. We then selected EMR data to include in the algorithm, developed criteria to classify patients as known HIV status (HIV-positive or HIV-negative) or unknown HIV status, and finally applied a structure to the criteria. We calculated the algorithm's performance characteristics for identifying patients with unknown HIV status by comparing results of the algorithm to those of the reference standard.

Reference Standard—To evaluate the algorithm's accuracy for identifying patients with unknown status, we required a gold standard. Because no accepted gold standard for unknown HIV status exists, we defined a reference standard using medical record review. We designed the reference standard to replicate a review that clinicians would perform to ascertain whether patients have known or unknown HIV status using commonly accessible data in medical records. The medical record review included manual review of inpatient and outpatient clinical notes, laboratory results, radiology and pathology reports, and medications. Criteria not normally used by clinicians performing manual chart review (e.g. billing codes or criteria containing complex temporal rules) were not included in our reference standard. Our reference standard definitions are:

HIV-positive: Defined by the presence of a positive HIV Western blot (WB), detectable HIV viral load (VL), documentation of HIV infection by a provider, or prescription of antiretroviral therapy (ART) (excluding ART use as pre/post-exposure prophylaxis determined by documentation in the record).

HIV-negative: Defined by not fulfilling the definition of HIV-positive and by the presence of a negative HIV antibody (Ab) test (oral swab or blood test) or undetectable VL.

Unknown HIV status: Defined by neither fulfilling the definitions of HIV-positive nor HIV-negative.

In cases of conflicting medical record data (e.g. documentation of HIV infection followed by a negative HIV antibody), the most recent data were used to make the final determination of HIV status.

Selection of EMR Data—We considered laboratory, billing (*International Classification of Disease, 9th Revision, Clinical Modification [ICD9-CM]*), and problem list data for inclusion as criteria in the algorithm. Although we included prescription data in the reference standard definition of HIV status, we did not consider these data for use in the algorithm because of incomplete capture across our healthcare system.

To select laboratory data to include, we reviewed all HIV-related laboratory assays used in our healthcare system since EMR implementation. Table 1 includes the complete list of assays considered. We excluded HIV genotypes, phenotypes, and tropism assays because these results are sent to providers and not captured in the EMR.

To select billing and problem list data, we reviewed all billing and problem list entries that contained the terms “Human Immunodeficiency Virus,” “Acquired Immunodeficiency Syndrome,” “HIV,” or “AIDS.” Entries that contained these terms but were not specific for active HIV infection of the patient in whose record it appeared (e.g. “HIV counseling,” “HIV exposure,” and “HIV infection in mother”) were excluded. Table 1 reports all included and excluded billing codes and problem list entries.

Algorithm Criteria—Using EMR data, we developed criteria to classify patients as having known HIV status (HIV-positive or HIV-negative) or unknown HIV status. Attributes of data that we considered included whether data were present or absent in patients’ EMR, the frequency with which data appeared in patients’ EMR (e.g. number of HIV-related billing events or number of undetectable VLs), and values taken by the data (e.g. specific laboratory result, billing code, or problem list entry) in patients’ EMR. A criterion was considered fulfilled if the definition was met at any time in patients’ EMR.

Laboratory criteria we considered indicative of HIV infection included a positive HIV WB or detectable VL. To identify HIV-positive patients well controlled on ART, we also constructed a composite criterion that included at least two undetectable VLs from different times sent concurrently with at least two CD4 counts (regardless of the CD4 count value).

Laboratory criteria we considered indicative of HIV-negative status included a negative HIV Ab or undetectable VL among individuals who did not fulfill criteria as HIV-positive. Although maintaining HIV-negative status is dependent on risk-behaviors subsequent to one’s last negative HIV test, for the purposes of algorithm development, we considered any prior negative test as evidence of HIV-negative status.

Billing criteria we considered indicative of HIV infection included a single inpatient HIV-related billing event or at least two outpatient HIV-related billing events, which is consistent with a previously published algorithm (Fultz et al., 2006). Problem list criteria we considered indicative of HIV infection included the entry of an HIV-related problem. No billing or problem list criteria were considered indicative of HIV-negative status.

Patients who did not fulfill criteria as HIV-positive or HIV-negative were considered to have unknown HIV status. Because the only criterion for patients to be identified as having unknown HIV status was not fulfilling criteria for being HIV-positive or HIV-negative, the algorithm can process all patients regardless of the amount or type of EMR data. Our criteria

for classification of HIV-positive, HIV-negative, and unknown HIV status are summarized in Table 2.

Algorithm structure—Because patients' EMR can contain conflicting data regarding HIV status over time, we explored several algorithms allowing for different temporal and hierarchical relationships. For example, we developed an algorithm accounting for patients who erroneously fulfilled HIV-positive criteria but subsequently had a negative HIV WB. We created three different algorithms that included the same criteria, but allowed for different temporal associations among the criteria and gave certain criteria the ability to “override” others when determining HIV status. Of the 1,502,836 unique patients with an EMR, 299 (0.02%) were classified differently by the different algorithms.

We identified the algorithm most accurately classifying HIV status among these differentially classified patients using a random number generator to select a sample of 30 (10.0%) of these 299 patients to compare with the reference standard medical record review. A single researcher (UF), blinded to the algorithms' classifications, applied the reference standard to this sample. The algorithm with the greatest proportion of patients whose HIV status was concordant with the reference standard was chosen as the final algorithm.

Determination of performance characteristics—To determine the final algorithm's performance for identifying patients with unknown HIV status, we applied the algorithm to all unique patients with at least one inpatient admission, outpatient visit, or ED visit between 2008–2012. We chose this timeframe to ensure accessibility to all medical record elements for the reference standard review. Using the random sampling command in STATA (v.12, StataCorp; College Station, TX), we selected a sample of 200 patients to undergo classification of HIV status according to the reference standard medical record review. Estimating that between 1%–2% of patients are known to be HIV-positive, we chose this sample size to increase the likelihood of including patients in each category of HIV status (unknown HIV status, HIV-negative, and HIV-positive). A single researcher (UF), blinded to the algorithm's classification, performed medical record reviews. Based on concordance between the reference standard classification and that of the algorithm, we calculated the performance characteristics for how the final algorithm identifies unknown HIV status, including its sensitivity, specificity, and positive and negative predictive values. The 95% confidence intervals were calculated using binomial expansion.

RESULTS

Figure 1 displays the final algorithm for the determination of known and unknown HIV status. Between 2008–2012, 1,000,738 unique patients had at least one inpatient admission, outpatient visit, or ED visit. The algorithm identified 791,634 (79.1%) patients as having unknown HIV status. Of the remaining 209,104 patients identified as having known HIV status, 197,037 (19.7%) were identified as HIV-negative, and 12,067 (1.2%) were identified as HIV-positive.

The random sample of 200 patients selected to undergo the reference standard medical record review represents 0.02% of patients. The number of patients from the random sample

identified by the algorithm as having unknown status, HIV-negative, and HIV-positive were 159 (79.5%), 38 (19.0%), and 3 (1.5%), respectively. Among sampled patients, the reference standard identified 158 (79.0%), 39 (19.5%), and 3 (1.5%) as having unknown HIV status, HIV-negative, and HIV-positive, respectively. Concordance between classifications of HIV status according to the reference standard and the algorithm is shown in Table 3. The algorithm's sensitivity and specificity for identifying patients with unknown HIV status were 99.4% (95% CI: 96.5%–100%) and 95.2% (95% CI: 83.8%–99.4%), respectively. In the population sampled, the algorithm's positive and negative predictive values for identifying patients with unknown status were 98.7% (95% CI: 95.5%–99.8%) and 97.6% (95% CI: 87.1%–99.9%), respectively (Table 3).

DISCUSSION

We developed an algorithm using commonly available data from the EMR of a large healthcare system to identify patients with unknown HIV status. Our algorithm has high sensitivity and specificity for identifying patients with unknown HIV status, as well as high positive and negative predictive values. To our knowledge, this is the first description of an algorithm using EMR data to identify patients with unknown HIV status that has been assessed against a reference standard. Identifying patients with unknown HIV status has important individual and public health implications.

Several studies reported algorithms that identify patients' HIV status. These algorithms use billing data alone to identify HIV-positive patients and are primarily used to monitor epidemiologic trends and clinical outcomes (Antoniou et al., 2011; Fasciano et al., 1998; Fultz et al., 2006). Unlike these algorithms, by incorporating a wide variety of EMR data and a structure accounting for contradictory data regarding HIV status, our algorithm identifies not only HIV-positive patients, but also those who are HIV-negative and have unknown HIV status. We are aware of one other study that used an EMR-based tool to identify patients with unknown HIV status (Goetz et al., 2008). Unlike that study, we report which EMR data were used in the determination of HIV status, how conflicts among data elements were resolved, and the performance characteristics of our algorithm for identifying patients with unknown HIV status when compared to a reference standard.

Using an EMR-based algorithm to identify patients with unknown HIV status has several strengths. First, by including a wide range of data, such a tool addresses the many ways in which patients are identified as having a known HIV status, particularly those who are HIV-positive. By identifying and then excluding patients with known HIV status, an EMR-based tool enables a robust, objective, and reproducible process for identifying those with unknown HIV status. This has distinct advantages over patient self-report of prior testing as a method for ascertaining unknown HIV status, which has questionable validity (Albrecht et al., 2012; Hutchinson et al., 2004; Jenness et al., 2009). Second, by using an EMR-based tool to identify patients with unknown HIV status, automated clinical decision support in the EMR can be used to strategically offer HIV testing. Finally, because many EMRs contain similar data elements, an EMR-based algorithm should be transferrable across healthcare systems.

Despite our algorithm's strengths, it also has limitations. First, the algorithm is limited to data available in our EMR and cannot access data, including HIV tests, from outside our healthcare system. Because our EMR was launched in 1997 and captures data from a large, integrated healthcare system repeatedly accessed by many patients, we believe this limitation is attenuated. Two limitations due to the algorithm's design deserve comment. First, because of how prescription data are captured in our EMR, we did not include these data as criteria in the algorithm. However, when we applied the algorithm to a registry of 3,517 patients who received HIV care in our system from 2011–2012, our algorithm identified 3,505 (99.7%) as HIV-positive, suggesting the value added by addition of a medication criterion would be minimal. Second, according to the current criteria, the algorithm identifies patients as HIV-negative if they ever tested negative for HIV. This definition allows for a maximum estimate of patients known to be HIV-negative, and is consistent with the New York State Public Health law mandating the offer of HIV testing to patients in healthcare settings at least once (New York State Department of Health, 2010). However, patients who engage in HIV risk-behaviors after a negative HIV test remain at risk for HIV infection, and therefore, the accuracy of our "HIV-negative" classification depends on time since the last negative test and ongoing risk factors. In future iterations of the algorithm, we will examine how varying time since last negative HIV test impacts rates of unknown HIV status.

Our algorithm has numerous applications that can help understand expanded HIV testing strategies. First, our process for determining the algorithm's performance characteristics demonstrates its ability to produce a prevalence estimate of unknown HIV status. We found a prevalence of unknown HIV status of 79.1% among patients who had any inpatient admission, outpatient visit, or ED visit in our healthcare system from 2008–2012. While this demonstrates feasibility of using the algorithm for this purpose, it will be important to apply the algorithm to specific populations for whom the question of unknown HIV status is more relevant. Second, serial application of the algorithm to monitor trends in prevalence of unknown HIV status will enable us to measure the impact of expanded HIV testing strategies on the prevalence of unknown HIV status, a key outcome for expanded testing strategies aiming to "test the untested." Third, efficient implementation of expanded testing strategies in healthcare systems in which low-risk patients frequently have recurrent visits is challenging (Hudepohl et al., 2011; Lyons et al., 2009). Because our algorithm is EMR-based, it can be operationalized to facilitate strategic testing efforts directed to those patients who may be at greatest risk for undiagnosed HIV. Finally, while the algorithm will be most accurate and useful when applied to healthcare systems with robust EMRs that practice routine HIV testing, as use of EMRs and health information exchanges proliferate (Jha et al., 2010; Adler-Milstein et al., 2009), application of the algorithm in different healthcare systems may provide insight into different roles that HIV testing plays in local prevention efforts within an HIV epidemic notable for regional diversity.

In conclusion, we developed an algorithm with high sensitivity and specificity that can be used to identify individuals with unknown HIV status. This algorithm can advance our understanding of the proportion of individuals with unknown HIV status, measure the impact that expanded HIV testing strategies have on the proportion of individuals with unknown HIV status, allow for strategic testing of those with unknown status, and should be

transferrable across healthcare systems. Identifying individuals with unknown HIV status can inform expanded HIV testing strategies implemented as part of HIV prevention efforts.

Acknowledgments

This study was supported in part by the Center for AIDS Research at the Albert Einstein College of Medicine and Montefiore Medical Center (NIH AI-51519); HIV Prevention Trial Network study 065; NIH R25DA023021; NIH R01DA032110; NIH R34DA031066; and CTSA grants UL1RR025750, KL2RR025749, and TL1RR025748 from the NCCR, a component of the NIH. We thank Devin Thompson for his contribution to the development of the algorithm.

References

- Adler-Milstein J, Bates DW, Jha AK. U.S. Regional health information organizations: progress and challenges. *Health Aff (Millwood)*. 2009; 28(2):483–492.10.1377/hlthaff.28.2.483 [PubMed: 19276008]
- Albrecht E, Frascarolo P, Meystre-Agustoni G, Farron A, Gilliard N, Darling K, Cavassini M. An analysis of patients' understanding of 'routine' preoperative blood tests and HIV screening. Is no news really good news? *HIV Med*. 2012;10.1111/j.1468-1293.2012.00993.x
- Antoniou T, Zagorski B, Loutfy MR, Strike C, Glazier RH. Validation of case-finding algorithms derived from administrative data for identifying adults living with human immunodeficiency virus infection. *PLoS One*. 2011; 6(6):e21748.10.1371/journal.pone.0021748 [PubMed: 21738786]
- Branson BM, Handsfield HH, Lampe MA, Janssen RS, Taylor AW, Lyss SB, Clark JE. Revised recommendations for HIV testing of adults, adolescents, and pregnant women in health-care settings. *Morbidity and mortality weekly report Recommendations and reports / Centers for Disease Control*. 2006; 55(RR-14):1–17.
- Centers for Disease Control and Prevention (CDC). Missed opportunities for earlier diagnosis of HIV infection--South Carolina, 1997–2005. *MMWR Morb Mortal Wkly Rep*. 2006; 55(47):1269–1272. [PubMed: 17136020]
- Centers for Disease Control and Prevention (CDC). Vital signs: HIV testing and diagnosis among adults--United States, 2001–2009. *MMWR Morb Mortal Wkly Rep*. 2010; 59(47):1550–1555. [PubMed: 21124295]
- Centers for Disease Control and Prevention (CDC). Monitoring selected national HIV prevention and care objectives by using HIV surveillance data-- United States and 6 U.S. dependent areas--2010. *HIV Surveillance Supplemental Report*. 2012; 17(3 part A) Retrieved from <http://www.cdc.gov/hiv/topic/surveillance/resources/reports/>.
- Chadborn TR, Delpuch VC, Sabin CA, Sinka K, Evans BG. The late diagnosis and consequent short-term mortality of HIV-infected heterosexuals (England and Wales, 2000–2004). *AIDS*. 2006; 20(18):2371–2379.10.1097/QAD.0b013e32801138f7 [PubMed: 17117024]
- Chou R, Cantor AG, Zakher B, Bougatsos C. Screening for HIV in pregnant women: systematic review to update the 2005 U.S. Preventive Services Task Force recommendation. *Ann Intern Med*. 2012; 157(10):719–728.10.7326/0003-4819-157-10-201211200-00009 [PubMed: 23165663]
- Fasciano NJ, Cherlow AL, Turner BJ, Thornton CV. Profile of Medicare beneficiaries with AIDS: application of an AIDS casefinding algorithm. *Health Care Financ Rev*. 1998; 19(3):19–38. [PubMed: 10345411]
- Fultz SL, Skanderson M, Mole LA, Gandhi N, Bryant K, Crystal S, Justice AC. Development and verification of a "virtual" cohort using the National VA Health Information System. *Med Care*. 2006; 44(8 Suppl 2):S25–30.10.1097/01.mlr.0000223670.00890.74 [PubMed: 16849965]
- Goetz MB, Hoang T, Bowman C, Knapp H, Rossman B, Smith R, Group QUHHP. A system-wide intervention to improve HIV testing in the Veterans Health Administration. *J Gen Intern Med*. 2008; 23(8):1200–1207.10.1007/s11606-008-0637-6 [PubMed: 18452045]
- Hudepohl NJ, Lindsell CJ, Hart KW, Ruffner AH, Trott AT, Fichtenbaum CJ, Lyons MS. Effect of an Emergency Department HIV Testing Program on the Proportion of Emergency Department Patients Who Have Been Tested. *Annals of Emergency Medicine*. 2011; 58(1, Supplement 1):S140–S144.10.1016/j.annemergmed.2011.03.038 [PubMed: 21684393]

- Hutchinson AB, Corbie-Smith G, Thomas SB, Mohanan S, del Rio C. Understanding the patient's perspective on rapid and routine HIV testing in an inner-city urgent care center. *AIDS Educ Prev.* 2004; 16(2):101–114. [PubMed: 15134119]
- Jenkins TC, Gardner EM, Thrun MW, Cohn DL, Burman WJ. Risk-based human immunodeficiency virus (HIV) testing fails to detect the majority of HIV-infected persons in medical care Settings. *Sex Transm Dis.* 2006; 33(5):329–333.10.1097/01.olq.0000194617.91454.3f [PubMed: 16547450]
- Jeness SM, Murrill CS, Liu KL, Wendel T, Begier E, Hagan H. Missed opportunities for HIV testing among high-risk heterosexuals. *Sex Transm Dis.* 2009; 36(11):704–710.10.1097/OLQ.0b013e3181ab375d [PubMed: 19652632]
- Jha AK, DesRoches CM, Kralovec PD, Joshi MS. A progress report on electronic health records in U.S. hospitals. *Health Aff (Millwood).* 2010; 29(10):1951–1957.10.1377/hlthaff.2010.0502 [PubMed: 20798168]
- Kawasumi Y, Abrahamowicz M, Ernst P, Tamblyn R. Development and validation of a predictive algorithm to identify adult asthmatics from medical services and pharmacy claims databases. *Health Serv Res.* 2011; 46(3):939–963.10.1111/j.1475-6773.2010.01235.x [PubMed: 21275988]
- Lyons MS, Lindsell CJ, Raab DL, Ruffner AH, Trott AT, Fichtenbaum CJ. Comparison of emergency department HIV testing data with visit or patient as the unit of analysis. *J Med Screen.* 2009; 16(1):29–32.10.1258/jms.2009.008086 [PubMed: 19349528]
- Marks G, Crepaz N, Janssen RS. Estimating sexual transmission of HIV from persons aware and unaware that they are infected with the virus in the USA. *AIDS.* 2006; 20(10):1447–1450.10.1097/01.aids.0000233579.79714.8d [PubMed: 16791020]
- Merchant RC, Catanzaro BM, Seage GR 3rd, Mayer KH, Clark MA, Degruittola VG, Becker BM. Demographic variations in HIV testing history among emergency department patients: implications for HIV screening in US emergency departments. *J Med Screen.* 2009; 16(2):60–66.10.1258/jms.2009.008058 [PubMed: 19564517]
- Montefiore Medical Center. Annual Report 2011- Where We Stand. 2012. Retrieved from <http://www.montefiore.org/annualreport>
- Moyer VA. U.S. Preventive Services Task Force. Screening for HIV: U.S. Preventive Services Task Force Recommendation Statement. *Ann Intern Med.* 2013; 159(1):51–60.10.7326/0003-4819-159-1-201307020-00645 [PubMed: 23698354]
- Myers JE, Braunstein SL, Shepard CW, Cutler BH, Mantsios AR, Sweeney MM, Tsoi BW. Assessing the impact of a community-wide HIV testing scale-up initiative in a major urban epidemic. *J Acquir Immune Defic Syndr.* 2012; 61(1):23–31.10.1097/QAI.0b013e3182632960 [PubMed: 22739134]
- Navaneethan SD, Jolly SE, Schold JD, Arrigain S, Saupe W, Sharp J, Nally JV Jr. Development and validation of an electronic health record-based chronic kidney disease registry. *Clin J Am Soc Nephrol.* 2011; 6(1):40–49.10.2215/CJN.04230510 [PubMed: 21051745]
- New York City Department of Health and Mental Hygiene (NYC DOHMH). New York City HIV/AIDS Annual Surveillance Statistics, 2010. 2012a. Retrieved from <http://www.nyc.gov/html/doh/html/ah/hivtables.shtml>
- New York City Department of Health and Mental Hygiene (NYC DOHMH). The Bronx Knows. HIV Testing Initiative: Final Report. 2012b. Retrieved from <http://www.nyc.gov/html/doh/downloads/pdf/ah/bronx-knows-summary-report.pdf>
- New York State Department of Health (NYS DOH). Frequently asked questions regarding the HIV testing law. 2010. Retrieved from <http://www.health.ny.gov/diseases/aids/testing/law/faqs.htm>
- Palella FJ Jr, Deloria-Knoll M, Chmiel JS, Moonman AC, Wood KC, Greenberg AE, Investigators HIVOS. Survival benefit of initiating antiretroviral therapy in HIV-infected persons in different CD4+ cell strata. *Ann Intern Med.* 2003; 138(8):620–626. [PubMed: 12693883]
- Peissig PL, Rasmussen LV, Berg RL, Linneman JG, McCarty CA, Waudby C, Starren JB. Importance of multi-modal approaches to effectively identify cataract cases from electronic health records. *J Am Med Inform Assoc.* 2012; 19(2):225–234.10.1136/amiajnl-2011-000456 [PubMed: 22319176]
- Prejean J, Song R, Hernandez A, Ziebell R, Green T, Walker F, Group HIVIS. Estimated HIV incidence in the United States, 2006–2009. *PLoS One.* 2011; 6(8):e17502.10.1371/journal.pone.0017502 [PubMed: 21826193]

- Qaseem A, Snow V, Shekelle P, Hopkins R Jr, Owens DK. Clinical Efficacy Assessment Subcommittee A. C. o. P. Screening for HIV in health care settings: a guidance statement from the American College of Physicians and HIV Medicine Association. *Ann Intern Med.* 2009; 150(2): 125–131. [PubMed: 19047022]
- Shuter J, Alpert PL, DeShaw MG, Greenberg B, Klein RS. Rates of and factors associated with self-reported prior HIV testing among adult medical patients in an inner city emergency department in the Bronx, New York City. *J Acquir Immune Defic Syndr Hum Retroviro.* 1997; 14(1):61–66. [PubMed: 8989212]
- Szumski NR, Cheng EM. Optimizing algorithms to identify Parkinson's disease cases within an administrative database. *Mov Disord.* 2009; 24(1):51–56.10.1002/mds.22283 [PubMed: 18816696]
- White House Office of National AIDS Policy. National HIV/AIDS Strategy for the United States. 2010. Retrieved from <http://www.whitehouse.gov/sites/default/files/uploads/NHAS.pdf>

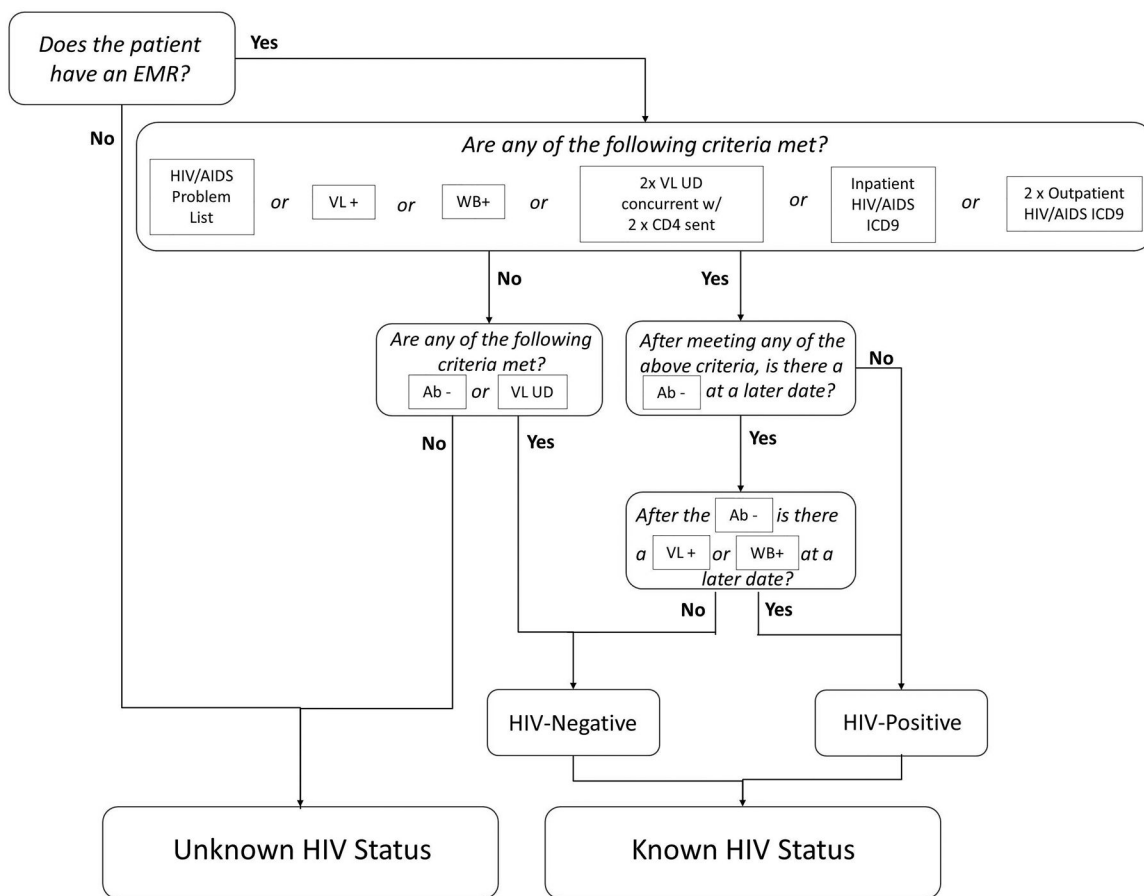


Figure 1. Algorithm for the Determination of HIV Status using EMR Data. The algorithm uses logical and temporal criteria to classify all patients as having known HIV status (HIV-positive or negative) or unknown HIV Status. EMR indicates electronic medical; ICD9, International Classification of Disease, 9th Revision; VL, viral load; UD, undetectable; WB, HIV Western blot; Ab, HIV antibody test.

Table 1

HIV-related laboratory, billing, and problem data considered for algorithm criteria

Category of Data	Laboratory*	Billing (ICD9)	Problem List
Included	HIV Ab (rapid) HIV Ab (standard) HIV viral load HIV western blot CD4 count (concurrent w/ VL [†])	042 HIV Disease 042.0 HIV & Specific Infection 042.1 HIV Causing Other Infection 042.2 HIV with Neoplasm 042.9 Unspecified AIDS 043 HIV Causing Condition NEC 043.0 HIV Lymphadenopathy 043.1 HIV causing CNS disease 043.2 HIV causing other disorders involving the immune mechanism 043.3 HIV causing disease NEC 043.9 AIDS Related Complex NOS 044 Other HIV Infection 044.0 HIV with Acute Infection 044.9 HIV infection NOS 079.53 HIV, Type 2 795.78 Positive Serologic Findings; HIV V08 Asymptomatic HIV Infection	AIDS AIDS due to HIV-1 AIDS due to HIV-II AIDS related dementia Asymptomatic HIV infection Cryptosporidiosis related to HIV HIV infection HIV positive HIV-1 AIDS HIV-1 infection HIV-2 AIDS HIV-2 infection HIV
Excluded	HIV Genotype HIV Phenotype HIV tropism	795.71 Nonspecific evidence of HIV	HIV counseling HIV exposure HIV infection in mother HIV complicating pregnancy

* Each category of laboratory test listed includes multiple different assays used in our system since implementation of the EMR in 1997

[†] Viral load

Table 2

Summary of criteria included in algorithm

Category of Data	HIV-positive	HIV-negative	Unknown HIV status
Laboratory	<ol style="list-style-type: none"> 1 Presence of a positive HIV western blot 2 Presence of a detectable HIV viral load 3 Presence of two undetectable HIV viral loads sent concurrently with two CD4 counts 	<ol style="list-style-type: none"> 1 Presence of a negative HIV antibody test (rapid or standard) * 2 Presence of an undetectable HIV viral load* 	Neither fulfilling criteria for HIV-positive or HIV-negative status
Billing	<ol style="list-style-type: none"> 1 Presence of a single inpatient HIV-related ICD9 event 2 Presence of two outpatient HIV-related ICD9 events 	None	
Problem List	<ol style="list-style-type: none"> 1 Presence of a single HIV-related problem list entry 	None	

* Among patients not fulfilling criteria for HIV-positive

Table 3

Performance characteristics of final algorithm

		Medical Records	
		Unknown HIV Status	Known HIV Status*
Algorithm	Unknown HIV Status	157 (“True Unknown”)	2 (“False Unknown”)
	Known HIV Status*	1 (“False Known”)	40 (“True Known”)

Sensitivity= 99.4% (96.5–100)	Positive Predictive Value = 98.7% (95.5–99.8)
Specificity= 95.2% (83.8–99.4)	Negative Predictive Value = 97.6% (87.1–99.9)

* Known HIV status includes those identified as HIV-positive or HIV-negative