HHS Public Access

Author manuscript

Int J Cardiol. Author manuscript; available in PMC 2023 February 01.

Published in final edited form as:

Int J Cardiol. 2022 February 01; 348: 152–156. doi:10.1016/j.ijcard.2021.12.022.

Use of troponins in the classification of myocardial infarction from electronic health records. The Atherosclerosis Risk in Communities (ARIC) Study

Anna M. Kucharska-Newton[1,2], **Matthew Shane Loop**[3], **Manuela Bullo**[4], **Carlton Moore**[5], **Stephanie W. Haas**[6], **Lynne Wagenknecht**[7], **Eric A. Whitsel**[1], **Gerardo Heiss**[1]

[1]Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC

[2]Department of Epidemiology, College of Public Health, University of Kentucky, Lexington, KY

[3]Division of Pharmacotherapy and Experimental Therapeutics, UNC Eshelman School of Pharmacy, University of North Carolina at Chapel Hill, Chapel Hill, NC

[4]Hospital Ramos Mejia, Buenos Aires, Argentina

[5]Department of Medicine, School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC

[6]School of Information and Library Science, University of North Carolina at Chapel Hill, Chapel Hill, NC

[7]Wake Forest University Population Health Sciences, Winston-Salem, NC

Abstract

Objective: Electronic health record (EHR) data are underutilized for abstracting classification criteria for heart disease. We compared extraction of EHR data on troponin I and T levels with human abstraction.

Methods: Using EHR for hospitalizations identified through the Atherosclerosis Risk in Communities (ARIC) Study in four US hospitals, we compared blood levels of troponins I and T extracted from EHR structured data elements with levels obtained through data abstraction by human abstractors to 3 decimal places. Observations were divided randomly 50/50 into training and validation sets. Bayesian multilevel logistic regression models were used to estimate agreement by hospital in first and maximum troponin levels, troponin assessment date, troponin

Corresponding author: Anna Kucharska-Newton, PhD MPH, Department of Epidemiology, University of North Carolina at Chapel Hill, 123 W. Franklin St, Suite 401, Chapel Hill, NC 27514, anna_newton@unc.edu.

Author statement

Anna Kucharska-Newton: Conceptualization, Methodology, Data curation, Investigation, Formal analysis, Writing-Original draft preparation **Matthew Loop**: Formal analysis, Writing-Review and Editing **Manuela Bullo:** Data curation, Formal analysis, Writing-Review and Editing **Carlton Moore**: Writing-Review and Editing **Stephanie Haas**: Data curation, Writing-Review and Editing. **Lynne Wagenknecht**: Writing – Review and Editing. **Eric Whitsel**: Writing-Review and Editing **Gerardo Heiss:** Funding Acquisition, Supervision, Conceptualization, Writing-Reviewing and Editing

upper limit of normal (ULN), and classification of troponin levels as normal (< ULN), equivocal (1 – 2x ULN), abnormal (>2x ULN), or missing.

Results: Estimated overall agreement in first measured troponin level in the validation data was 88.2% (95% credible interval: 65.0% – 97.5%) and 95.5% (91.2 – 98.2%) for the maximum troponin level observed during hospitalization. The largest variation in probability of agreement was for first troponin measured, which ranged from 66.4% to 95.8% among hospitals.

Conclusion: Extraction of maximum troponin values during a hospitalization from EHR structured data is feasible and accurate.

1. Introduction

Algorithmic identification and classification of cardiovascular events from information contained in electronic health records (EHRs) is of increasing interest to clinicians and researchers. EHR-derived event classification has the potential to advance population approaches to healthcare by providing relatively low cost and large-scale outcome assessment including performance measures which can be embedded in patient medical records.[1] A recent systematic review of studies examining the use of EHR in the diagnosis of acute myocardial infarction (MI) underscored the extent to which EHR data are underutilized in the abstraction of cornerstone MI classification criteria, including information on patient symptoms such as chest pain, cardiac biomarkers, and electrocardiographic (ECG) findings.[2]

In this report, we leverage data from the Atherosclerosis Risk in Communities (ARIC) Study community surveillance of cardiovascular events[3] to examine the performance of protocols for algorithmic abstraction of troponin T and I levels directly from patient EHR.

2. Materials and Methods

2.1. Study population

Analyses were performed using EHR from hospitalizations identified through the Atherosclerosis Risk in Communities (ARIC) Study. From 1987 through 2014, the ARIC study conducted surveillance of cardiovascular hospitalizations in four geographic regions of the United States, Jackson, MS, Forsyth County, NC, Washington County, MD, and the suburbs of Minneapolis, MN.[3] Six hospitals in the ARIC study geographic areas were contacted with a request to provide EHR versions of medical records for the year 2015, which had already been submitted to the ARIC Study Coordinating Center as part of the ARIC community surveillance of cardiovascular events. One hospital was in transition to a different EHR platform, thus unable to provide data. Remaining hospitals had functional data warehouses and were able to provide copies of patient medical records. A templated request for structured data elements was submitted to each hospital's data analysts. The template was in the form of an Excel document and included spreadsheet tabs for the following structured data elements: Vital Signs, Laboratory, Medications, Procedures, Imaging, Diagnoses, and Demographics. Medical record number, encounter ID and date of discharge served as unique identifiers which were included in all requested components of the medical records and were used to link the data across record categories. Data fields

within each spreadsheet tab were structured using the common data model[4] to assure data harmonization across hospitals. Also requested were provider notes including Emergency Department notes, History and Physical notes, Progress notes, and Discharge Summaries. All notes were requested as delimited documents.

2.2. Data preparation

To protect patient confidentiality, all patient and provider names were removed from the records upon receipt of the data at the ARIC Coordinating Center using the MIT deID program, an open-source, Perl-based software package.[5] Additionally, a crosswalk file between medical record numbers and an encrypted ID was created for each hospital to allow for continued linkage of the data across categories. Thereafter, the medical record numbers were removed from the records and the crosswalk stored on a secure, encrypted server.

Structured data elements were converted to comma delimited files and processed using standard statistical software (SAS version 9.4, SAS Institute, Cary, NC) and STATA version 15.1 (StataCorp LP, College Station, TX). Python programs were used to identify pre-specified negation terms. All data extraction protocols included date and time stamps to ascertain the temporal evolution of clinical manifestations and biomarkers.

2.3. Classification of biomarker levels

Troponin I and troponin T data were abstracted from the Laboratory portion of the structured data elements. We created a stratified random sample of records for algorithm training within each of four hospitals, with a sampling probability of 0.5, reserving the other records in those hospitals for validation. The number of records obtained from the fifth hospital was so small that the records were not used in data analysis. Algorithms for data extraction, developed as SAS programs, were iteratively derived using the training data. SAS programs were also used to examine the agreement on dates of a troponin assay, its upper limit of normal, and troponin levels to 3 decimal places against data for the same records abstracted by human ARIC Study abstractors.[6] The highest troponin level recorded during the first 4 days of the hospitalization was used to arrive at its EHR classification. Although all biomarker levels were used to arrive at the corresponding ARIC classification, it was usually based on troponin I or T. However, it also could have been based on creatine kinase (CK) and / or its isoenzyme CK-MB if they led to a more abnormal overall classification. We used the following classification scheme: abnormal (>2x the upper limit normal), equivocal (1 – 2x the upper limit of normal), normal (< upper limit of normal), and missing (no biomarkers available).

To “downgrade” troponin levels, which may have been elevated due to non-cardiac events, we abstracted from EHR structured data elements International Classification of Diseases, 9th Revisions (ICD-9) procedure and diagnosis codes suggesting evidence of trauma or a cardiac procedure preceding the time of the troponin measurement. The purpose of the downgraded classification was to remove falsely abnormal troponin levels that were likely not due to a cardiac event.

Algorithms derived using training data were then applied to the validation data. We present descriptive results from the training and validation datasets, with final modeling performed

on the validation datasets for inference. Descriptive statistics are counts or percentages and counts, unless otherwise noted.

2.4. Statistical analyses

In order to assess the probability of agreement across all hospitals, we estimated agreement (yes or no) on the following outcomes: troponin levels, troponin assessment date, troponin upper limit of normal, troponin classification (abnormal, equivocal, normal or incomplete), and a binary classification of troponins as abnormal versus normal equivocal or incomplete. We only used observations where there was a non-missing value for the outcome of interest in either the EHR algorithmic or ARIC manual abstraction. This choice was made in order to remove observations that would have been missing in both, and thus counted as agreement, but were “false agreements” that would have inflated the estimated probability of agreement (e.g., the hospitalization lasted only 2 days, so agreement on biomarker levels on day 4 was nonsensical).

We used Bayesian multilevel models to estimate the probability of agreement on our outcomes among the four hospitals. All models included random intercepts for hospital. Models for agreement on date of biomarker collection also included random intercepts for participant and fixed effects for day of assessment. We fit these models using the Stan programming language,[7] which used Hamiltonian Monte Carlo methods in order to estimate the posterior of the probability. We used the following priors in the model: normal(0, 5) for the intercept, normal(0, 2.5) for the effects of day of hospitalization, and a student half-t(3, 0, 1) for all random intercept standard deviations.[8] Models were fit using the brms[9,10] and rstanarm [11] packages in R,[12] which provided a high-level interface to the Stan modeling language.

To summarize the results of the models for each outcome, we presented the median of the posterior for the probability of agreement, along with 95% credible intervals based upon the quantiles of the posterior distribution. We provided both overall probabilities for the “average” hospital, as well as the hospital-specific probabilities of agreement. Probabilities for agreement on dates also assumed an “average” participant.

3. Results

Of a total 1,460 training records and 1,515 validation records, 879 and 916 records, respectively, were eligible for abstraction as possible coronary heart disease events according to the ARIC Study protocol. Records not eligible for full abstraction were hospitalizations where the participant was deceased upon arrival to the hospital, or there was very little evidence of a potential cardiac cause of the hospitalization. A total 384 training and 411 validation records were eligible for complete biomarker abstraction according to the ARIC Study protocol. Eligibility for biomarker abstraction was based on the presence of more than one ECG, any cardiac enzymes above the normal limit, hospital transfer, and/or whether the participant was a part of the ARIC cohort. Finally, we excluded 46 records from the training set and 36 records from the validation set that did not contain at least one non-missing troponin value measured, leading to final sample sizes for statistical modeling

of 338 in the training set and 375 in the validation set. Results presented in the manuscript pertain to the validation set. Training set data are presented as Supplementary material.

Characteristics of the four hospitals which contributed data to this study are presented in Table 1. The hospitals were based in the four ARIC geographic areas, Forsyth County, NC; Jackson, MS; Washington County, MD; and the suburbs of Minneapolis, MN. Three of the hospitals were affiliated with an academic institution and used the EPIC EHR platform, while one used Allscripts. Hospital bed size varied from 247 to 873.

Descriptive statistics on the number of validation records and agreement on troponin measurements, dates of measurement, upper limit normal of assays, and ARIC biomarker classifications for the validation records are presented in Table 2. The same statistics for the training records are presented in Supplemental Table 2.

We observed excellent agreement between troponin values abstracted using algorithms applied to the EHR structured data elements and values obtained by ARIC abstractors. The Cohen's kappa for agreement in troponin values across all four hospitals and across all days of hospitalization (up to 4 days) was on average 93.5% (95% CI 91.2%, 95.8%) for troponin I and 99.7% (95% CI 98.7%, 100%) for troponin T. It is important to note that troponin I was measured in three out of the four hospitals in this study, while only one hospital (hospital B) measured troponin T. The agreement in the dates of biomarker abstraction was on average 90.6% (95% CI 87.4%, 93.3%).

Figure 1 shows the expected probability of agreement in dates of biomarker assessment with 95% Bayesian credible intervals (the probability is 0.95 that the true probability lies within those intervals) in the validation data. The estimated probability of agreement for date of biomarker assessment for all days and hospitals was approximately 1, after accounting for the large variation among individuals ($\widehat{\sigma_{Id}} = 10.2$). Further evaluation of the median posterior random intercepts for each participant revealed three distinct populations: (1) those who had all four dates correct; (2) those who had one date wrong due to a typing error in the ARIC database; and (3) those for whom the first date was different by one day between the EHR and ARIC abstraction, leading to disagreement on all 4 dates (data not shown). Probabilities of agreement were relatively consistent across days, since once the first hospitalization date was selected, the rest of the days were determined.

Figure 2 shows the posterior probabilities of agreement on the first and maximum troponin values, ARIC troponin classification variables, and the upper limit normal of the troponin assays. The largest variation in agreement among hospitals occurred for the first troponin measurement, with the estimated probabilities ranging from 66.4% (95% credible interval: 56.2 – 75.5%) in hospital C to 95.8% (83.6% – 99.9%) in hospital B. The estimated agreement on the maximum troponin values varied less among hospitals than for the first troponin value, with an estimated overall agreement for the average hospital (i.e., random intercept of 0 for hospital) being 95.5% (91.2 – 98.2%). Agreement on the non-downgraded biomarker classifications was high and showed little variability. Agreement for the downgraded biomarker classifications was noticeable worse than the non-downgraded ones with more variability among hospitals, with the downgraded 4-level biomarker

classification ranging from 77.2% in hospital C to 85.8% in hospital B. The downgraded abnormal biomarker classification ranged from 80.5% in hospital C to 86.6% in hospital B. Cross-tabulation for the classification of biomarkers and the downgraded biomarkers is presented in Supplemental Tables 3a/b and 4a/b, respectively.

4. Discussion

In this study, which was based on EHR data from four distinct hospitals, we observed an excellent agreement between algorithmically obtained EHR data and manual protocol-guided ARIC gold standard in the final classification of biomarker levels relative to the upper limit normal. We observed a modestly good agreement in dates of troponin assessment and in troponin I and T values between these two abstraction modalities. Discrepancies in the dates were small and pertained mostly to differences in time and date stamps occurring close to midnight, presumably related to lags in ordering tests, drawing blood, assaying troponin, and reporting levels of this biomarker. Those discrepancies did not impact the assessment of the temporal sequence of troponin values, allowing for a comparable review of the clinically significant pattern of troponin elevation and decrease occurring during the course of the hospitalization

Cardiac troponins, regulatory proteins which are released from the myocardiocytes in response to injury[13], are highly specific to the myocardium and are therefore considered the biomarker of choice in the identification and classification of acute coronary syndromes.[14] Patients experiencing an acute MI event will present with an initial increase in blood troponin levels, which is followed by a more gradual decline. In this dynamic pattern, both the biomarker levels as well as their trajectory over time help differentiate MI events from troponin elevations associated with chronic diseases.[15] Therefore, serial troponin assessments over the course of the clinical encounter are needed for optimal MI diagnosis and management. In our study we demonstrated the ability to algorithmically capture from EHR the date and time of biomarker assessment, thus allowing for the real-time ascertainment of patterns of troponin elevation. As stated above, a proportion of the records were offset by one calendar day relative to the ARIC standard. Although this modestly impacted our level of agreement with manually abstracted dates, the temporal sequence of troponin levels was maintained.

Acute MI diagnoses are based on information regarding serial cardiac biomarker levels, electrocardiographic (ECG) evidence of active ischemia, and evidence of the presence of chest pain of cardiac origin. Although none of these diagnostic criteria can stand alone in the final MI diagnosis, cardiac troponin is a class I risk stratification criterion that can help guide the management of MI patients.[16] Recent studies of high sensitivity troponin thresholds for the triage of patients presenting to the emergency departments with suspected acute coronary syndromes and identification of low risk subgroups point to the independent utility of troponins in clinical practice.[17,18] In this study, we did not examine high sensitivity troponins, as at the time of data collection those were not yet in routine use in the United States. However, our findings suggest that blood troponin levels algorithmically assessed from EHR are highly informative with respect to the classification of biomarker levels as abnormal, equivocal, or normal in relation to threshold criteria. We found this

classification robust to the algorithmic EHR-based identification of conditions influential in the “downgrading” of biomarker levels, including trauma and cardiac procedures.

This study makes an important contribution to the rapidly increasing interest in the use of EHR data in population studies. Our focus on the identification and classification of troponin levels from structured EHR data has independent value in the assessment of patients presenting with acute coronary syndromes. In our analysis we used EHR from four separate hospitals, thus we were able to observe differences in how information from patients’ encounter is entered into the EHR by different providers. We used a detailed structured data request form that was reviewed by hospitals’ data warehouse analysts for ease of use prior to application. Data review protocols were conducted by a data coordinating center using, as the gold standard, abstraction and classification protocols from a highly established cardiovascular study.

This study has several limitations. As a feasibility project, it was based on a small number of hospitals, three of which used EPIC as their EHR platform, with only one hospital using a different platform. This limited our generalizability. We used the upper limit normal of troponin detection as the threshold for the classification of troponin levels, whereas many clinical settings rely on the sex-specific 99^{th} percentile of the standard population distribution. This study did not address EHR-based protocols for the abstraction and classification of high-sensitivity troponins. Such analyses will be possible as those biomarkers become more prevalent in clinical settings. Finally, because of our inclusion/exclusion criteria and use of ARIC Study surveillance data, our conclusions apply to a population of patients hospitalized with some evidence of coronary heart disease (i.e., by ECG or biomarkers).

5. Conclusion

In this feasibility study we established standards for the use of EHR in the abstraction and classification of troponin levels. Our results suggest that algorithmic extraction protocols, based on structured EHR data elements, accurately identify cardiac troponin levels to 3 decimal places, allowing for their diagnostic classification.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

The Atherosclerosis Risk in Communities study has been funded in whole or in part with Federal funds from the National Heart, Lung, and Blood Institute, National Institutes of Health, Department of Health and Human Services, under Contract nos. (HHSN268201700001I, HHSN268201700002I, HHSN268201700003I, HHSN268201700005I, HHSN268201700004I). The authors thank the staff and participants of the ARIC study for their important contributions.

7. References

1. Casey JA, Schwartz BS, Stewart WF, Adler NE. Using Electronic Health Records for Population Health Research: A Review of Methods and Applications. Annu Rev Public Health. 2016;37:61–81. [PubMed: 26667605]
2. Rubbo B, Fitzpatrick NK, Denaxas S, et al. Use of electronic health records to ascertain, validate and phenotype acute myocardial infarction: A systematic review and recommendations. Int J Cardiol. 2015;187:705–711. [PubMed: 25966015]
3. White AD, Folsom AR, Chambless LE, et al. Community surveillance of coronary heart disease in the Atherosclerosis Risk in Communities (ARIC) Study: methods and initial two years' experience. J Clin Epidemiol. 1996;49(2):223–233. [PubMed: 8606324]
4. Overhage JM, Ryan PB, Reich CG, Hartzema AG, Stang PE. Validation of a common data model for active safety surveillance research. J Am Med Inform Assoc. 2012;19(1):54–60. [PubMed: 22037893]
5. Neamatullah I, Douglass MM, Lehman LW, et al. Automated de-identification of free-text medical records. BMC Med Inform Decis Mak. 2008;8:32. [PubMed: 18652655]
6. Rosamond WD, Chambless LE, Folsom AR, et al. Trends in the incidence of myocardial infarction and in mortality due to coronary heart disease, 1987 to 1994. N Engl J Med. 1998;339(13):861–867. [PubMed: 9744969]
7. Gelman SA, Roberts SO. How language shapes the cultural inheritance of categories. Proc Natl Acad Sci U S A. 2017;114(30):7900–7907. [PubMed: 28739931]
8. Gelman SA. Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). Anal International Society for Bayesian Analysis 2006;1:515–534.
9. Burkner PC. Advanced Bayesian Multilevel Modeling with the R Package brms. The R Journal. 2018;10(1):395–411.
10. Burkner P brms: An R Package for Bayesian Multilevel Models Using Stan Journal of Statistical Software. 2017;80(1):1–28.
11. Goodrich B, Gabry J, Ali I, rstanarm SB: Bayesian applied regression modeling via Stan. R package version 2.21.1. https://mc-stan.org/rstanarm. Published 2020. Accessed.
12. Team RC. A language and environment for statistical computing. R Foundation for Statistical Computing, https://www.R-project.org/. Published 2019. Accessed2019.
13. Daubert MA, Jeremias A. The utility of troponin measurement to detect myocardial infarction: review of the current findings. Vasc Health Risk Manag. 2010;6:691–699. [PubMed: 20859540]
14. Thygesen K, Alpert JS, Jaffe AS, et al. Fourth Universal Definition of Myocardial Infarction (2018). Circulation. 2018;138(20):e618–e651. [PubMed: 30571511]
15. Newby LK, Jesse RL, Babb JD, et al. ACCF 2012 expert consensus document on practical clinical considerations in the interpretation of troponin elevations: a report of the American College of Cardiology Foundation task force on Clinical Expert Consensus Documents. J Am Coll Cardiol. 2012;60(23):2427–2463. [PubMed: 23154053]
16. Amsterdam EA, Wenger NK, Brindis RG, et al. 2014 AHA/ACC Guideline for the Management of Patients with Non-ST-Elevation Acute Coronary Syndromes: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. J Am Coll Cardiol. 2014;64(24):e139–e228. [PubMed: 25260718]
17. Bularga A, Lee KK, Stewart S, et al. High-Sensitivity Troponin and the Application of Risk Stratification Thresholds in Patients With Suspected Acute Coronary Syndrome. Circulation. 2019;140(19):1557–1568. [PubMed: 31475856]
18. Chapman AR, Hesse K, Andrews J, et al. High-Sensitivity Cardiac Troponin I and Clinical Risk Scores in Patients With Suspected Acute Coronary Syndrome. Circulation. 2018;138(16):1654–1665. [PubMed: 30354460]

1. Troponin values can be accurately extracted algorithmically from EHR
2. Such information can be easily applied in diagnostic classification

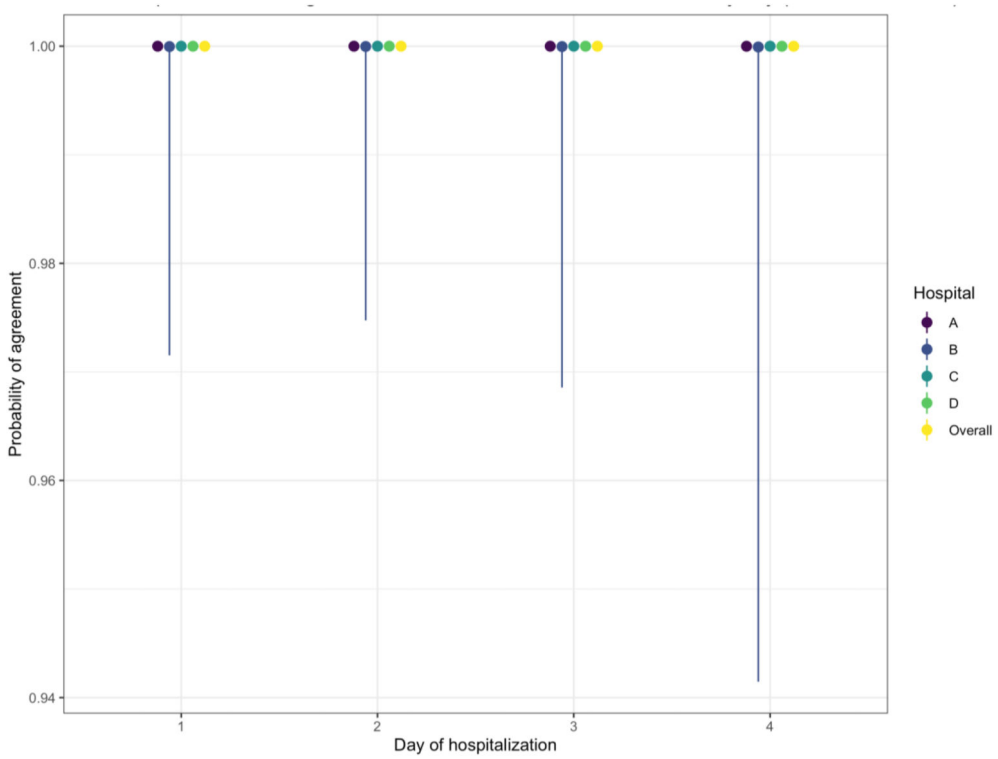

Figure 1. Probability of agreement between algorithmic and human medical record abstraction of information on day of hospitalization. The ARIC Study.

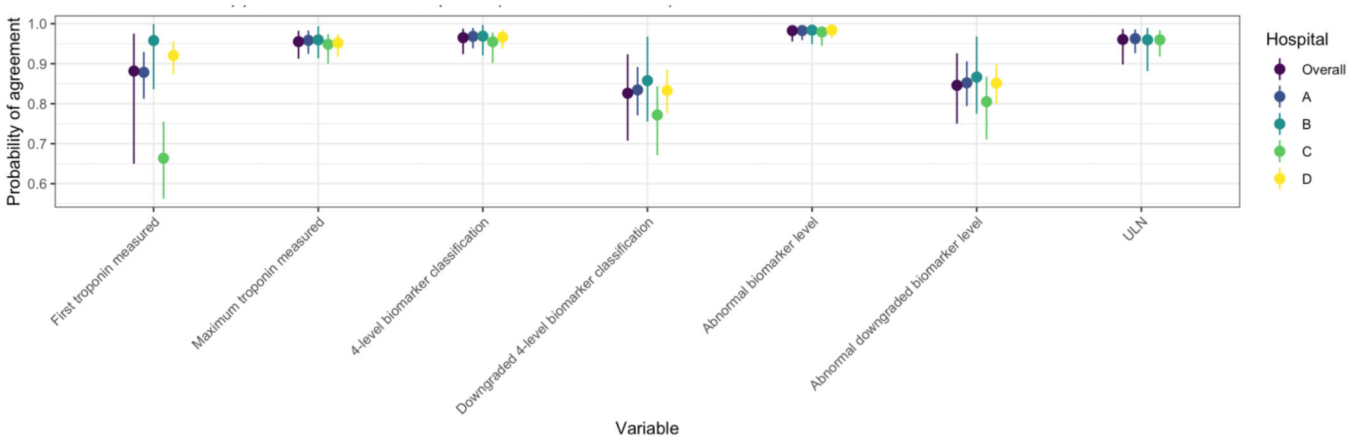

Figure 2. Probability of agreement between algorithmic and human medical record abstraction of troponin classification parameters. The ARIC Study.

Table 1.

Hospital characteristics

	Hospital			
Characteristic	**A**	**B**	**C**	**D**
Type of EHR	EPIC	EPIC	EPIC	Allscripts
Status	Academic	Academic	Academic	Non-academic
Hospital bedsize	873	700	385	247
Abstracted records (N)	232	73	199	334
Records with complete biomarker abstraction	130	25	98	158

Table 2.

Agreement on troponin measurements, % (n)

	Hospital				Overall
	A	B	C	D	
Agreement on troponin levels *, % (n)					
First measured troponin level	87.6% (113)	100.0% (18)	64.8% (91)	92.2% (153)	84.5% (375)
Maximum troponin level	96.5% (113)	100.0% (18)	93.4% (91)	94.8% (153)	95.2% (375)
Agreement on date of biomarker collection, % (n)					
Day 1	88.6% (114)	68.4% (19)	94.6% (92)	89.5% (153)	89.4% (378)
Day 2	88.5% (113)	73.7% (19)	93.5% (92)	90.2% (153)	89.7% (377)
Day 3	88.2% (102)	66.7% (18)	92.9% (84)	88.7% (142)	88.4% (346)
Day 4	87.5% (80)	64.3% (14)	91.1% (56)	84.8% (105)	85.9% (255)
Agreement on upper limit normal of troponin assay, % (n)	96.9% (130)	96.0% (25)	95.9% (98)	-**	96.4% (253)
Agreement on 4-level biomarker classification, % (n)	84.6% (130)	72.0% (25)	86.7% (98)	93.7% (158)	87.8% (411)
Agreement on abnormal vs. not abnormal biomarker classification, % (n)	98.5% (130)	100.0% (25)	96.9% (98)	98.7% (158)	98.3% (411)
Agreement on downgraded 4-level biomarker classification, % (n)	73.1% (130)	68.0% (25)	68.4% (98)	81.0% (158)	74.7% (411)
Agreement on downgraded abnormal vs. not abnormal enzyme classification, % (n)	87.7% (130)	96.0% (25)	78.6% (98)	86.1% (158)	85.4% (411)

* Agreement occurred when values from EHR and human abstraction were identical to 3 decimal places, or when both values were missing.

** Information on the upper limit normal value for troponin I was missing in data received from hospital C.