



Observational study of sudden cardiac arrest risk (OSCAR): Rationale and design of an electronic health records cohort

Kyndaron Reinier^{a,*}, Harpriya S. Chugh^a, Audrey Uy-Evanado^a, Elizabeth Heckard^a, Marco Mathias^a, Nichole Bosson^{b,c,d}, Vinicius F. Calsavara^e, Piotr J. Slomka^f, David A. Elashoff^g, Alex A.T. Bui^h, Sumeet S Chugh^{a,f,*}

^a Center for Cardiac Arrest Prevention, Department of Cardiology, Smidt Heart Institute, Cedars-Sinai Medical Center Los Angeles CA USA

^b Los Angeles County EMS Agency Los Angeles CA USA

^c Harbor-UCLA Medical Center, Department of Emergency Medicine, and The Lundquist Institute Torrance CA USA

^d David Geffen School of Medicine at UCLA Los Angeles CA USA

^e Department of Computational Biomedicine, Biostatistics Shared Resource, Cedars-Sinai Medical Center Los Angeles CA USA

^f Division of Artificial Intelligence in Medicine, Department of Medicine, Cedars-Sinai Medical Center Los Angeles CA USA

^g Department of Medicine, Statistics Core, University of California Los Angeles CA USA

^h Medical & Imaging Informatics Group, University of California, Los Angeles 90095 CA, USA

ARTICLE INFO

Keywords:

Sudden cardiac death
Out-of-hospital cardiac arrest
Electronic health records cohort
Prediction

ABSTRACT

Background: Out-of-hospital sudden cardiac arrest (SCA) is a major cause of mortality and improved risk prediction is needed. The Observational Study of Sudden Cardiac Arrest Risk (OSCAR) is an electronic health records (EHR)-based cohort study of patients receiving routine medical care in the Cedars-Sinai Health System (CSHS) in Los Angeles County, CA designed to evaluate predictors of SCA. This paper describes the rationale, objectives, and study design for the OSCAR cohort.

Methods and Results: The OSCAR cohort includes 379,833 Los Angeles County residents with at least one patient encounter at CSHS in each of two consecutive calendar years from 2016 to 2020. We obtained baseline cohort characteristics from the EHR from 2012 until the start of follow-up, including demographics, vital signs, clinical diagnoses, cardiac tests and imaging, procedures, laboratory results, and medications. Follow-up will continue until Dec. 31, 2025, with an expected median follow-up time of ~ 7 years. The primary outcome is out-of-hospital SCA of likely cardiac etiology attended by Los Angeles County Emergency Medical Services (LAC-EMS). The secondary outcome is total mortality identified using California Department of Public Health – Vital Records death certificates. We will use conventional approaches (diagnosis code algorithms) and artificial intelligence (natural language processing, deep learning) to define patient phenotypes and biostatistical and machine learning approaches for analysis.

Conclusions: The OSCAR cohort will provide a large, diverse dataset and adjudicated SCA outcomes to facilitate the derivation and testing of risk prediction models for incident SCA.

1. Introduction

1.1. Background and rationale

Out-of-hospital sudden cardiac arrest (SCA) is a sudden, unexpected loss of pulse causing an abrupt collapse with loss of circulation, breathing, and consciousness. [1] SCA is fatal without immediate cardiopulmonary resuscitation (CPR) and/or defibrillation. Survival following SCA remains low at ~ 9% in the US; [2,3] thus > 90% of SCA

results in sudden cardiac death (SCD) which accounts for more years of life lost than any single cancer. [4] Prevention of SCA would be the most effective means to reduce the population SCD burden, yet prediction and prevention of SCA remain challenging.

Currently, identification of persons at high risk for SCA is mostly based on a single marker – reduced left ventricular ejection fraction (LVEF). Individuals with LVEF ≤ 35–40%, indicating poor cardiac pumping efficiency, are candidates for primary prevention of SCD with the implantable cardioverter defibrillator (ICD), [5] which can avert an

* Corresponding authors.

E-mail addresses: kyndaron.reinier@cshs.org (K. Reinier), sumeet.chugh@cshs.org (S.S. Chugh).

<https://doi.org/10.1016/j.ijcha.2025.101614>

Received 10 January 2025; Accepted 12 January 2025

2352-9067/© 2025 Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

SCD by defibrillating the heart back into a normal sinus rhythm. While the ICD does improve survival, [6,7] it has become less effective over time because incidence of SCA among individuals with LVEF $\leq 35\%$ has declined (in one study by 44% from 1995 to 2014), [8] at least partly due to improvements in treatment for heart failure. [9] There is also increased recognition that LVEF $\leq 35\%$ alone is not sufficient to identify the majority of individuals at high risk of SCA; at least 70% of SCA occurs among individuals with LVEF $> 35\%$, [10,11] for whom there is no effective population-based risk prediction. Thus, SCA remains a significant public health problem and improved prediction is needed.

Prior research has identified many risk factors for SCA including cardiovascular risk factors; [12] established cardiac disease; [13] echocardiogram (ECG) [14] and echocardiographic risk markers besides LVEF, [15,16] and non-cardiac conditions including chronic obstructive pulmonary disease, [17] epilepsy, [18] and chronic kidney disease. [19] While a combination of these risk factors could be expected to identify individuals at elevated risk of SCA, no prediction models are currently used in clinical practice. Furthermore, many individuals at risk of SCA are also at high risk of other causes of mortality. To be most clinically useful, a risk score for SCA should specifically predict risk of SCA as opposed to other causes of death, as individuals at highest specific risk of arrhythmic death will theoretically benefit most from the ICD or other interventions targeted to reduce arrhythmia risk.

Our research group developed and validated a risk prediction model (VFRisk) for shockable SCA in two population-based case-control studies [20] that had good discrimination (AUC = 0.81 [95 % CI 0.77–0.84]) and was successfully validated in internal and external validation datasets (AUC = 0.78 in each). [20] However, VFRisk has not yet been evaluated in a prospective study that can account for competing risk of non-SCD events. An earlier population-based score (the ARIC SCD risk score) was developed for use in the general population, included widely available traditional cardiovascular risk factors, [12,21] and was tested in external populations. [22] While this model performed reasonably well in predicting SCA, [12,21,22] it did not distinguish individuals at risk of SCA from those at risk for overall cardiovascular death.

We have designed the Observational Study of Cardiovascular Arrest Risk (OSCAR), an electronic health records (EHR) cohort of 379,833 individuals, to allow a large enough sample size to evaluate risk of SCA in a general patient population, to develop SCA risk prediction models for potential clinical use, and to evaluate SCA risk vs competing risk of other mortality.

1.2. Objectives

The overall goal of the OSCAR EHR cohort is to provide a large, diverse dataset of individuals with carefully defined clinical phenotypes in which to develop, test, and refine models to identify individuals at high risk of SCA, including among population subgroups by sex, race, ethnicity, and clinical profile.

Specifically, our objectives are to:

- (1) Characterize each individual's baseline clinical profile, including information from clinical diagnoses, procedures, tests, clinical labs, electrocardiograms, echocardiograms, and medications using existing electronic health records (EHR).
- (2) Follow the cohort for incident SCA (primary outcome) or other mortality (secondary outcome).
- (3) Construct detailed and carefully defined datasets that will allow for construction of risk prediction models for SCA.
- (4) Identify predictors for SCA risk specifically, separate from other causes of mortality.
- (5) Evaluate risk predictors among population subgroups by sex, race, and ethnicity.

2. Methods

2.1. Cohort design / cohort construction

Overview of cohort design: The Observational Study of Cardiac Arrest Risk (OSCAR) is an electronic health records (EHR) cohort study of 379,833 individuals (274,007 aged ≥ 35) without a history of SCA who will be followed for SCA outcomes and total mortality. The OSCAR cohort is a subset of the broader Observational Study of Cardiovascular Event Risk (OSKER, $n = 382,121$), which has the same inclusion criteria, but that additionally includes individuals with a history of SCA at baseline.

Source of cohort data: The OSCAR cohort, within the larger OSCER cohort, is based in the Cedars-Sinai Health System (CSHS), a nonprofit healthcare organization including a large academic 915-bed hospital and > 30 ambulatory clinics in Los Angeles County (LAC), CA with a linked EHR. Baseline data for this cohort has been obtained from the Clinical Data Warehouse (CDW), a data repository from the EHR established by the Data Intelligence Team at CSHS. No contact will be made with cohort members; all analyses will be carried out in a de-identified manner. The study protocol conforms to the ethical guidelines of the 1975 Declaration of Helsinki as reflected in a priori approval by the CSHS institutional review board.

Cohort inclusion and exclusion criteria: Individuals from the CDW were included in the OSCER cohort if they: (1) had at least one patient encounter recorded in each of two consecutive calendar years during the 5-year capture period (2016–2020); (2) were a LAC resident; (3) were alive at the second qualifying patient encounter. No other exclusions were applied for the OSCER cohort based on patients' age, race, ethnicity, or baseline clinical profile (Fig. 1). The OSCAR cohort is restricted to individuals with no history of SCA prior to the start of follow-up.

We required encounters in two consecutive years to increase the likelihood that patients were receiving regular care at CSHS and to allow a more complete assessment of medical history. [23] Preliminary cohort data indicate that 63% of encounters were ambulatory (outpatient) visits. We required residence in LAC because SCA outcomes will be obtained from Los Angeles County Emergency Medical Services (LAC-

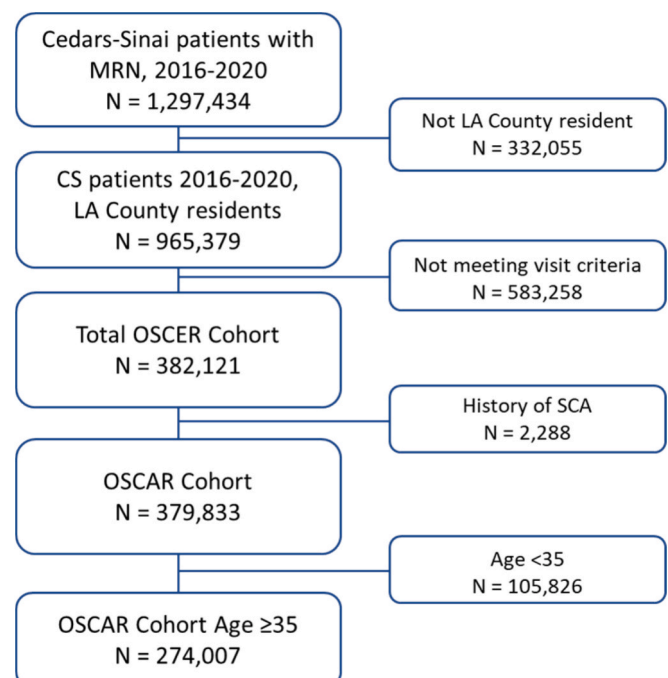


Fig. 1. Flow chart: OSCAR Cohort.

EMS) data (see below), facilitating data linkage for outcome events. Individuals with a history of SCA before or during the exposure run-in period will be excluded from the OSCAR cohort, which is focused on identifying risk factors for first SCA. [24] As risk predictors in younger SCA are unique, [25] most predictive modeling will be performed in the subset of the OSCAR cohort ages ≥ 35 .

2.2. Data retrieval and processing

Data sources and types: EHR data has been obtained from routine, urgent, emergency, and inpatient encounters at CSHS. Structured data includes: patient demographics; vital signs; anthropometrics; risk behaviors (smoking, drinking, drug use); International Classification of Diseases (ICD-10) clinical diagnostic codes from all outpatient, inpatient, and emergency department encounters; Current Procedural Terminology (CPT) codes for cardiac and selected non-cardiac procedures; outpatient medications; and clinical laboratory results represented using Logical Observation Identifiers Names and Codes (LOINC). [23] We have also obtained quantitative and qualitative data from 12-lead ECGs stored in the MUSE system and results from echocardiograms stored in the Syngo system. Unstructured data includes free-text clinical notes.

Time period for data retrieval: Prevalent clinical variables at baseline were obtained from the CDW from 2012 until the start of follow-up for each cohort participant (Fig. 2).

Data management: EHR data received from the CDW was processed using SQL to create pre-specified data tables and to visualize and check data quality. Data will be stored as SAS datasets in wide or long format as appropriate, linked by study ID.

2.3. Variable definitions

Definition of clinical variables: Clinical phenotypes at baseline will be defined by ICD-10 diagnosis codes with additional data as appropriate (e.g., laboratory, medication, procedures). Whenever possible, standard and published clinical phenotype definitions will be used. [26] Additional phenotype algorithms will be considered when appropriate, such as those in the Phenotype KnowledgeBase (PheKB), [27] Human Phenotype Ontology (HPO) for encoding rare disease phenotypes; [28] and PheNorm algorithms, an automatic unsupervised algorithm for identifying patient phenotypes (e.g., coronary artery disease) from EHR without requiring expert-labeled samples for training. [29]

Manual review of selected records for validation of phenotype algorithms:

Because reliance on ICD-10 diagnosis codes and/or algorithms alone can lead to errors in clinical phenotyping, [30] validation studies are planned for selected clinical phenotypes to compare sensitivity and positive predictive value (PPV) of these phenotype algorithms to manual review of medical records (gold standard).

2.4. Missing data

Approach to missing data: As expected, clinical tests such as ECGs, echocardiograms, and laboratory measurements are not available for all cohort subjects at baseline, since our cohort uses data obtained during routine medical care. Our approach to missingness will depend on the proportion of missing values and assumptions regarding missingness at random. When data is not missing at random (e.g., cohort members with ECG and echo data available are expected to differ from those without these tests), missingness will be handled through the design process by constructing prediction models in cohort subsets with this data available. These specific models will be generalizable to individuals in other health systems with comparable data available. Strategies for other missing data will include treating missing as a category for categorical variables; assigning a mean or median value for a continuous variable; or multiple imputation approaches.

Variability in number of health care encounters: Patients in the OSCAR cohort with more frequent encounters are expected to have a more complete medical record and thus more diagnoses than patients with fewer encounters. [31] Both conventional and machine learning techniques will be used to evaluate potential bias or added information created by these patterns, e.g., multiple imputation models including variables such as number of healthcare encounters and overall health status. [32]

2.5. Outcome ascertainment

Longitudinal follow-up: All individuals in the OSCAR cohort will be followed from their entry into the cohort until they experience the primary outcome (SCA), secondary outcome (death from other cause), or until Dec. 31, 2025 (Fig. 2).

Challenges of ascertaining out-of-hospital SCA events: Given its sudden onset, high fatality rate, and setting outside the hospital, SCA is a difficult outcome to study. SCA is not a reportable condition, [33] and a small minority have autopsies conducted. Data from regional EMS systems can provide the most complete ascertainment of out-of-hospital

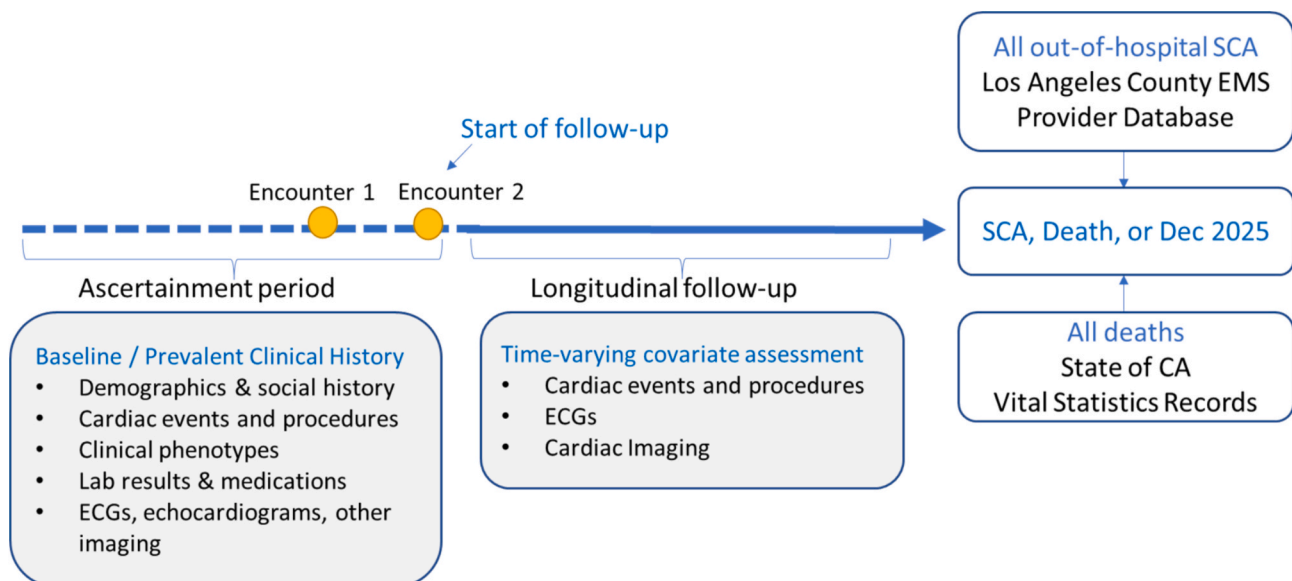


Fig. 2. OSCAR Cohort Design.

cardiac arrest because 9–1–1 is nearly always called, regardless of the arrest location (home, public indoor or outdoor, nursing home), or whether the patient had a witnessed collapse or was found unconscious or obviously dead. In contrast, the use of hospital records would significantly under-ascertain out-of-hospital SCA that occur in the community since ~ 38% of individuals with SCA are declared dead in the field despite attempted resuscitation, [34] while just under half (45.2%) of cases do not have resuscitation initiated due to obvious death or presumed futility. [35]

Primary outcome ascertainment: To achieve the most complete ascertainment possible, we have partnered with LAC-EMS to identify all individuals in the OSCAR cohort who have SCA during follow-up (Fig. 2). LAC-EMS, with > 4,200 paramedics, serves the LA County population of 10.1 million with 29 provider agencies and 70 911-receiving hospitals. [36] LAC-EMS maintains an EMS Provider Agency Database for all field encounters that includes the chief complaint for each 9–1–1 call as recorded by the medical dispatcher and paramedic-recorded provider impressions for each encounter. Presumed SCA cases from the LAC-EMS database will be the subset of EMS field encounters with: (a) chief complaint of “cardiac arrest” and/or (b) EMS provider impression of non-traumatic out-of-hospital cardiac arrest (OHCA) with or without resuscitation.

Data linkage between OSCAR cohort and LAC-EMS: Presumed SCA cases from the LAC-EMS Provider Database from Jan 1, 2017 through Dec. 31, 2025 will be linked to the OSCAR cohort by first, middle, and last name, date of birth, and date of OHCA using exact and pattern (probabilistic) matching (Fig. 2).

Secondary outcome: Data linkage for overall mortality: California Department of Public Health – Vital Records death certificates and the Social Security Death Index will be linked to the OSCAR cohort using exact and pattern matching to determine deaths from all causes during follow-up (Fig. 2). CSHS EHR data on mortality will also be obtained.

Adjudication of SCA for OSCAR: Following record linkage, for the subset with death certificates (expected for ~ 85% of cases), we will use data on immediate and contributing cause of death to exclude individuals whose sudden death occurred due to drug overdose (accidental or intentional), trauma (e.g., blunt or penetrating), hanging, drowning, and non-cardiac natural causes of death (e.g., metastatic cancer, other terminal illness, stroke, or pulmonary embolism). For the subset who survived SCA (expected ~ 10–15% of total SCA outcomes), pre-arrest and peri-arrest EHR encounters will be reviewed and those who meet inclusion criteria will be included in the SCA outcome definition. For both survivors and non-survivors, adjudicated SCA will be defined as a sudden and unexpected loss of pulse due to a likely cardiac etiology.

Expected number of SCA outcomes: Incidence of out-of-hospital SCA with attempted resuscitation across US states ranges from 49 to 137 per 100,000 per year, as reported by the CARES registry. [3] Adjudicated SCA for OSCAR may be larger, since the above estimate does not include SCA cases without attempted resuscitation. [35] Based on earlier estimates, 95–98% of SCA outcomes will occur among OSCAR cohort members aged ≥ 35. Assuming an annual SCA incidence of ~ 120 per 100,000 person-years among 274,007 cohort members aged ≥ 35 followed a median of 7 years, we expect ~ 2,300 SCA cases during follow-up. Table 1 provides estimates of SCA events by age based on age-specific rates from a Danish registry. [37] Actual number of outcomes will be affected by several factors, including how the relative age and health of the CSHS patient population compares to that of the populations that generated published estimates of SCA incidence.

2.6. Analytical approach

Descriptive analysis: We have performed descriptive statistics for baseline cohort characteristics, including demographics, clinical comorbidities, history of cardiac conditions and events, and ECG and echocardiogram availability. This longitudinal cohort study will allow

Table 1

Expected SCA events in OSCAR cohort aged ≥ 35 over expected median 7 years of follow-up (2017–2025).

Age group	Population annual event rate per 100,000*	Age group size (n)	Expected total events in OSCAR
35–44	14	58,283	58
45–54	43	57,536	173
55–64	86	57,455	345
65–74	159	56,406	627
75–84	310	30,282	657
≥85	465	14,045	457
Entire cohort	121	274,007	2317

*Estimated incidence rates based on reported age-specific event rates for OHCA cases aged ≥ 35 of likely cardiac etiology with resuscitation attempted [37], assuming those event rates represent ~ 70% of total event rates, since OHCA with attempted resuscitation is a subset of total OHCA. Total expected events in OSCAR ages < 35 not shown in table; estimated total n < 50 is based on OSCAR cohort size ages < 35 and published annual incidence rates per 100,000 of 2.1 (ages 0–2), 0.61 (ages 3–13), 1.44 (ages 14–24), and 4.4 (ages 25–35). [25].

calculation of SCA incidence by patient demographics (age, sex, race, ethnicity, insurance status) and clinical characteristics (history of coronary disease, heart failure, etc.).

Planned predictive analyses: Planned analyses include conventional statistical modeling of hazard ratios or other risk ratios for SCA associated with specific risk factors, and prediction models, including competing risk models to distinguish risk predictors for SCA from predictors of total mortality risk. Modeling will follow TRIPOD guidelines (Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis) for prediction of future outcomes. [38] As both EMS and EHR data are available, we will also be able to evaluate patient-level factors associated with survival from SCA.

We evaluated power based on cohort sample size, assuming most predictive analyses will include be conducted in 70% derivation and 30% internal validation datasets. For inferences regarding individual risk factors for SCA with a prevalence of ~ 25% we will have ≥ 80% power at a 5% significance level to detect hazard ratios of 1.21 to 1.42, respectively, in datasets requiring variables from clinical history only or clinical, ECG, and echocardiogram data, respectively. [39] In addition, minimum detectable differences (MDD) in the AUC curves of prediction models will range from 0.02 to 0.07, based on the c-statistic between pairs of models based on DeLong’s test. [20]

Artificial intelligence / machine learning approaches: We plan to use artificial intelligence (AI) and machine learning (ML) to refine variable (feature) definitions and to construct prediction models. Use of natural language processing (NLP) can reduce missingness and improve accuracy of clinical phenotype data, reducing misclassification and theoretically leading to better prediction models. [40] For selected predictor variables (features), we will derive NLP-based algorithms using diverse structured and unstructured data inputs for feature engineering. [41]

We plan to use deep learning-derived features from whole waveform ECG data, which contains richer information than discrete ECG markers and has been shown to predict SCD more accurately than the individual ECG variables. [42] A similar approach is planned to use ML-analyzed echocardiogram images rather than discrete variables from echocardiogram reports. Finally, risk modeling using ML (with either conventional or AI-derived variables/features) has the potential to identify relationships between the SCD outcome and complex data available in the OSCAR EHR cohort that conventional statistical analysis may not identify. [43]

2.7. Study Governance and oversight

A steering committee of principal investigators and advisors will provide scientific oversight and guidance. The OSCAR study has been reviewed and approved by the institutional review board at CSHS and

approved by the Honest Enterprise Research Broker (HERB) Committee of CSHS’s Enterprise Information Systems (EIS) Research Informatics and Scientific Computing Core (RISCC), which must approve all studies proposing the use of CSHS electronic medical records.

3. Initial results

Description of cohort: We identified 382,121 individuals meeting the residency and regular visit inclusion criteria for the full OSCER cohort (Fig. 1). The OSCAR cohort (n = 379,833) is the subset of the broader OSCER cohort excluding 2,288 individuals with a history of SCA. Most predictive modeling will be carried out in the OSCAR subset who are aged ≥ 35 at the start of follow-up (n = 274,007) (Fig. 1). Among these individuals, mean age was 59.3 ± 14.9 years, and 57% were female. White non-Hispanic individuals constitute 65% of the cohort, while Black (11%), Hispanic (11%), and Asian (8%) individuals constitute 30% (Table 2). Baseline prevalence of selected clinical phenotypes was similar to those reported in a contemporary EHR cohort (C3PO) in Massachusetts which had similar inclusion criteria. [40] Baseline prevalence of the clinical components of the VFRisk score for SCA [20] are presented in Table 2.

4. Discussion

We have constructed an EHR-based cohort of 379,833 individuals receiving routine care at a large health system who will be followed for 5 to 9 years (expected median follow-up ~ 7 years), with an expected 2,300 incident SCA events. The OSCAR cohort includes innovations in approach and planned analyses that are designed to address challenges in SCA research (Table 3). Importantly, traditional cardiovascular cohort studies with 5,000 to 20,000 individuals have required decades to accrue several hundred SCA cases for analysis (given an annual incidence of 50–100 SCA cases per 100,000 individuals), resulting in smaller numbers of outcomes and baseline risk predictors quite remote from the SCA event. OSCAR is significantly larger than traditional cardiovascular cohort studies, and thus will allow more timely outcome assessment relative to baseline, with < 10 years accrued between baseline and SCA.

4.1. Innovation and strengths

The large, diverse study cohort will allow development and

Table 2
Baseline demographics and clinical profile of the OSCAR cohort aged ≥ 35.

Characteristics	OSCAR Aged ≥ 35 N = 274,007
Age, mean ± SD	59.3 ± 14.9
Male	118,364 (43 %)
Female	155,643 (57 %)
Race and ethnicity*	
White	170,228 (65 %)
Black	29,109 (11 %)
Hispanic	27,619 (11 %)
Asian	20,924 (8 %)
Other	12,141 (5 %)
VFRisk Phenotypes	
Diabetes	40,581 (15 %)
Atrial fibrillation	18,593 (7 %)
Stroke	14,895 (5 %)
Heart failure	21,880 (8 %)
Myocardial infarction	11,318 (4 %)
COPD	20,936 (8 %)
Seizure Disorder	3,576 (1 %)
Syncope	14,652 (5 %)

*Race and ethnicity missing for ~ 5 % of cohort.

Table 3
Innovations in OSCAR.

INNOVATION	Description
Large, diverse cohort Larger number of SCA outcomes than prior cohorts	The OSCAR study will follow 379,833 individuals (274,007 aged ≥ 35) with expected accrual of ~ 2,300 SCA events, more than the combined total SCAs ascertained during prior US cohort studies of SCA.
Evaluate sex-, race-, and ethnicity-specific risk factors for SCA	To date, population-based cohort studies have accrued too few SCA events in US Hispanic and Asian individuals to allow evaluation of risk in these groups. Cedars-Sinai, based in Los Angeles County, serves a demographically diverse population. The OSCAR cohort aged ≥ 35 is ~ 58% female, ~11% Black race, ~11% Hispanic ethnicity, and ~ 8% Asian, allowing evaluation of SCA risk in these subgroups.
Cohort represents general patient population Includes “apparently low risk” individuals who account for the majority of SCA	Our planned analyses will specifically address prediction among individuals at apparently low risk who have SCD with LVEF > 35% and individuals without LVEF assessment (the majority), for whom there is currently no SCD risk prediction available.
Innovative ascertainment of SCA outcomes Collaboration with county EMS system for ascertainment of SCA	To obtain data for all out-of-hospital SCA cases among OSCAR cohort participants, we have established an innovative, unique collaboration with the Los Angeles County EMS system (LAC-EMS). SCA case ascertainment will be more complete than hospital-based approaches. We will also have access to arrest circumstances data, such as arrest location, initial rhythm, bystander CPR, and survival.
Less than 10 years from baseline to outcome Risk prediction for actionable time frames	Baseline predictors will be measured in relatively close proximity to the outcome (median follow-up ~ 7 years), resulting in prediction models that can be interpreted as indicating clinically actionable risk.
Machine learning and AI approaches Natural language processing for unstructured EHR data	We plan to apply natural language processing (NLP) to extract useable data from text fields in the EHR and to reduce missingness.
Machine learning models to improve prediction	Machine learning models have the potential to identify relationships between the SCA outcome and the complex data available in the OSCAR EHR cohort that conventional statistical analysis may not identify.
Multiple clinical applications Mechanistic, prediction, and prevention insights	Analyses may reveal novel associations highlighting potential mechanisms involved in SCA risk. New results may provide information useful for designing upstream prevention for SCD and for identifying and treating individuals at high risk for SCD.

evaluation of robust prediction models overall and in important patient subgroups by sex, race, and ethnicity. Most individuals in the cohort do not have established heart disease or left ventricular dysfunction. Nonetheless, most SCA events are expected in these apparently “low risk” individuals (and in whom no prediction models are in clinical use). The broad range of EHR-derived variables in the cohort will allow for a more complete evaluation of predictors of SCA, including non-cardiac factors. Finally, the OSCAR cohort has the potential to produce

prediction models to identify high-risk individuals across the LVEF spectrum for staged screening in a general patient population to identify individuals needing more work-up or those who could potentially benefit from primary prevention ICD. [43]

Also, US cohorts with SCA outcomes have included at most two racial groups. The OSCAR cohort has been designed to facilitate evaluation of risk factors for SCA in women and non-White populations. Most prior research indicates that, compared to men, women are less likely to have had LVEF \leq 35% or coronary artery disease (CAD) prior to SCD, [44–46] and present with lower prevalence of other risk factors. [47–50] Thus, there is a particular need for improved SCA prediction and prevention among women currently miss-classified as low risk. [51] Further, higher SCD incidence has been well-documented among Black individuals compared to their White counterparts. [52–54] Adjustment for CVD risk factors, comorbidities, and socio-demographics has explained only part of the increased risk, highlighting the need to identify new predictors of SCD risk in Black individuals. [53,54] Less is known regarding SCD incidence and risk factors in US Hispanic, Asian, and other non-White groups. Our group recently reported that CKD may play a large role in SCD risk among Hispanic individuals. [55] The racially and ethnically diverse OSCAR cohort will allow modeling of subgroup-specific risk prediction to build upon prior studies' results regarding risk factors for SCA that may differ by sex [56] and race. [52,57]

The unique linkage between our EHR-based cohort and the EMS system will facilitate more complete SCA case ascertainment than hospital-based approaches and more carefully defined SCA outcomes because both EMS and medical history data are available for adjudication.

Application of NLP techniques to semi-structured and free-text EHR data will improve the accuracy of clinical phenotyping and address potential limitations of code-based algorithms for clinical phenotypes. [23] In addition to conventional time-to-event models accounting for competing risk of mortality, we will also use machine learning approaches to identify relationships between SCA and the complex data available in the OSCAR EHR cohort that conventional statistical analysis may not identify. [43] This multipronged approach to data analysis is designed to improve model generalizability.

4.2. Limitations and potential challenges

Representativeness: EHR-based cohorts are naturally limited to individuals who seek regular healthcare, who are older and less healthy than those who do not regularly seek care. [23] While this limitation remains, our proposed risk scores for predicting SCA risk are intended for use in a patient population.

Incomplete data: EHR data has known limitations regarding missingness and potential bias. We are addressing these limitations through study design approaches, such as inclusion criteria requiring at least one patient encounter in each of two consecutive years (63% of baseline encounters were ambulatory visits). We are also extending back five years to obtain more complete clinical comorbidity, ECG and echocardiogram data. We will validate code-based algorithms against gold-standard phenotypes based on manual review, use NLP to capture unstructured EHR data and potentially improve accuracy of clinical diagnoses, and use advanced methods to handle missing variables.

Restriction to variables with adequate non-missing data for prediction modeling: Some prediction models may include variables obtained from ECG or echocardiogram evaluation. In clinical practice, these models will only be applicable to patient populations with these tests available.

4.3. Conclusions

The OSCAR EHR cohort is large, racially and ethnically diverse, with carefully defined clinical phenotypes and SCA outcomes. This cohort will facilitate descriptive analysis of SCA incidence in a general patient population and the development of prediction models to identify

individuals at high risk of SCA who could benefit from additional screening and treatment to prevent SCA.

CRedit authorship contribution statement

Kyndaron Reinier: Conceptualization, Formal analysis, Funding acquisition, Methodology, Writing – original draft. **Harpriya S. Chugh:** Data curation, Formal analysis, Methodology, Writing – review & editing. **Audrey Uy-Evanado:** Data curation, Investigation, Writing – review & editing. **Elizabeth Heckard:** Formal analysis, Visualization, Writing – review & editing. **Marco Mathias:** Data curation, Writing – review & editing. **Nichole Bosson:** Data curation, Methodology, Writing – review & editing. **Vinicius F. Calsavara:** Formal analysis, Writing – review & editing. **Piotr J. Slomka:** Methodology, Resources, Writing – review & editing. **David A. Elashoff:** Conceptualization, Formal analysis, Supervision, Writing – review & editing. **Alex A.T. Bui:** Conceptualization, Methodology, Writing – review & editing. **Sumeet S Chugh:** Conceptualization, Funding acquisition, Methodology, Project administration, Resources, Supervision, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- [1] G.I. Fishman, S.S. Chugh, J.P. Dimarco, C.M. Albert, M.E. Anderson, R.O. Bonow, A.E. Buxton, P.S. Chen, M. Estes, X. Jouven, et al., Sudden cardiac death prediction and prevention: report from a National Heart, Lung, and Blood Institute and Heart Rhythm Society Workshop, *Circulation*. 122 (2010) 2335–2348, <https://doi.org/10.1161/CIRCULATIONAHA.110.976092>.
- [2] S. Girotra, S. van Diepen, B.K. Nallamothu, M. Carrel, K. Vellano, M.L. Anderson, B. McNally, B.S. Abella, C. Sasson, P.S. Chan, et al., Regional Variation in Out-of-Hospital Cardiac Arrest Survival in the United States, *Circulation*. 133 (2016) 2159–2168, <https://doi.org/10.1161/CIRCULATIONAHA.115.018175>.
- [3] S.S. Martin, A.W. Aday, Z.I. Almarzooq, C.A.M. Anderson, P. Arora, C.L. Avery, C. M. Baker-Smith, B. Barone Gibbs, A.Z. Beaton, A.K. Boehme, et al., 2024 Heart Disease and Stroke Statistics: A Report of US and Global Data From the American Heart Association, *Circulation*. 149 (2024) e347–e913, <https://doi.org/10.1161/CIR.0000000000001209>.
- [4] E.C. Stecker, K. Reinier, E. Marjion, K. Narayanan, C. Teodorescu, A. Uy-Evanado, K. Gunson, J. Jui, S.S. Chugh, Public health burden of sudden cardiac death in the United States, *Circ. Arrhythm. Electrophysiol.* 7 (2014) 212–217, <https://doi.org/10.1161/CIRCEP.113.001034>.
- [5] A.E. Epstein, J.P. DiMarco, K.A. Ellenbogen, N.A. Estes 3rd, R.A. Freedman, L. S. Gettes, A.M. Gillinov, G. Gregoratos, S.C. Hammill, D.L. Hayes, et al., ACC/AHA/HRS 2008 Guidelines for Device-Based Therapy of Cardiac Rhythm Abnormalities: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Writing Committee to Revise the ACC/AHA/NASPE 2002 Guideline Update for Implantation of Cardiac Pacemakers and Antiarrhythmia Devices): developed in collaboration with the American Association for Thoracic Surgery and Society of Thoracic Surgeons, *Circulation*. 117 (2008) e350–e408, <https://doi.org/10.1161/CIRCULATIONAHA.108.189742>.
- [6] G.H. Bardy, K.L. Lee, D.B. Mark, J.E. Poole, D.L. Packer, R. Boineau, M. Domanski, C. Troutman, J. Anderson, G. Johnson, et al., Amiodarone or an implantable cardioverter-defibrillator for congestive heart failure, *N Engl J Med*. 352 (2005) 225–237.
- [7] A.J. Moss, W. Zareba, W.J. Hall, H. Klein, D.J. Wilber, D.S. Cannom, J.P. Daubert, S.L. Higgins, M.W. Brown, M.L. Andrews, Prophylactic implantation of a defibrillator in patients with myocardial infarction and reduced ejection fraction, *N Engl J Med*. 346 (2002) 877–883.
- [8] L. Shen, P.S. Jhund, M.C. Petrie, B.L. Claggett, S. Barlera, J.G.F. Cleland, H. J. Dargie, C.B. Granger, J. Kjekshus, L. Kober, et al., Declining Risk of Sudden Death in Heart Failure, *N Engl J Med*. 377 (2017) 41–51, <https://doi.org/10.1056/NEJMoa1609758>.
- [9] Butler J, Talha KM, Aktas MK, Zareba W, Goldenberg I. Role of Implantable Cardioverter Defibrillator in Heart Failure With Contemporary Medical Therapy. *Circ Heart Fail*. 2022;15:e009634. doi: 10.1161/CIRCHEARTFAILURE.122.009634.
- [10] A.P. Gorgels, C. Gijssbers, J. de Vreede-Swagemakers, A. Lousberg, H.J. Wellens, *Eur. Heart J.* 24 (2003) 1204–1209, [https://doi.org/10.1016/s0195-668x\(03\)00191-x](https://doi.org/10.1016/s0195-668x(03)00191-x).
- [11] Stecker EC, Vickers C, Waltz J, Socoteanu C, John BT, Mariani R, McNulty JH, Gunson K, Jui J, Chugh SS. Population-based analysis of sudden cardiac death with and without left ventricular systolic dysfunction: two-year findings from the

- Oregon Sudden Unexpected Death Study. *Journal of the American College of Cardiology*. 2006;47:1161-1166. doi: S0735-1097(05)03107-4 [pii] 10.1016/j.jacc.2005.11.045.
- [12] B.M. Bogle, H. Ning, J.J. Goldberger, S. Mehrotra, D.M. Lloyd-Jones, A Simple Community-Based Risk-Prediction Score for Sudden Cardiac Death, *Am. J. Med.* 131 (532-539) (2018) e535.
- [13] R.J. Myerburg, M.J. Junttila, Sudden cardiac death caused by coronary heart disease, *Circulation*. 125 (2012) 1043-1052, <https://doi.org/10.1161/CIRCULATIONAHA.111.023846>.
- [14] A.L. Aro, K. Reinier, C. Rusinaru, A. Uy-Evanado, N. Darouian, D. Phan, W.J. Mack, J. Jui, E.Z. Soliman, L.G. Tereshchenko, et al., Electrical risk score beyond the left ventricular ejection fraction: prediction of sudden cardiac death in the Oregon Sudden Unexpected Death Study and the Atherosclerosis Risk in Communities Study, *Eur. Heart J.* 38 (2017) 3017-3025, <https://doi.org/10.1093/eurheartj/ehx331>.
- [15] Narayanan K, Reinier K, Teodorescu C, Uy-Evanado A, Aleong R, Chugh H, Nichols GA, Gunson K, London B, Jui J, et al. Left ventricular diameter and risk stratification for sudden cardiac death. *J Am Heart Assoc*. 2014;3:e001193. doi: 10.1161/jaha.114.001193.
- [16] K. Reinier, C. Dervan, T. Singh, A. Uy-Evanado, S. Lai, K. Gunson, J. Jui, S. S. Chugh, Increased left ventricular mass and decreased left ventricular systolic function have independent pathways to ventricular arrhythmogenesis in coronary artery disease, *Heart Rhythm*. 8 (2011) 1177-1182, <https://doi.org/10.1016/j.hrthm.2011.02.037>.
- [17] M.E. van den Berg, B.H. Stricker, G.G. Brusselle, L. Lahousse, Chronic obstructive pulmonary disease and sudden cardiac death: A systematic review, *Trends Cardiovasc Med*. 26 (2016) 606-613, <https://doi.org/10.1016/j.tcm.2016.04.001>.
- [18] Bardai A, Lamberts RJ, Blom MT, Spanjaart AM, Berdowski J, van der Staal SR, Brouwer HJ, Koster RW, Sander JW, Thijs RD, et al. Epilepsy is a risk factor for sudden cardiac arrest in the general population. *PLoS one*. 2012;7:e42749. doi: 10.1371/journal.pone.0042749.
- [19] M.K. Shamseddin, P.S. Parfrey, Sudden cardiac death in chronic kidney disease: epidemiology and prevention, *Nat Rev Nephrol*. 7 (2011) 145-154, <https://doi.org/10.1038/nrneph.2010.191>.
- [20] S.S. Chugh, K. Reinier, A. Uy-Evanado, H.S. Chugh, D. Elashoff, C. Young, A. Salvucci, J. Jui, Prediction of Sudden Cardiac Death Manifesting With Documented Ventricular Fibrillation or Pulseless Ventricular Tachycardia, *JACC Clin Electrophysiol*. 8 (2022) 411-423, <https://doi.org/10.1016/j.jacep.2022.02.004>.
- [21] R. Deo, F.L. Norby, R. Katz, N. Sotoodehnia, S. Adabag, C.R. DeFilippi, B. Kestenbaum, L.Y. Chen, S.R. Heckbert, A.R. Folsom, et al., Development and Validation of a Sudden Cardiac Death Prediction Model for the General Population, *Circulation*. 134 (2016) 806-816, <https://doi.org/10.1161/CIRCULATIONAHA.116.023042>.
- [22] S. Welten, S. Remmelzwaal, M.T. Blom, A.A. van der Heijden, G. Nijpels, H.L. Tan, I. van Valkengoed, J.P. Empana, X. Jouven, F.N. Agesen, et al., Validation of the ARIC prediction model for sudden cardiac death in the European population: The ESCAPE-NET project, *Am. Heart J.* 262 (2023) 55-65, <https://doi.org/10.1016/j.ahj.2023.03.018>.
- [23] Williams BA, Voyce S, Sidney S, Roger VL, Plante TB, Larson S, LaMonte MJ, Labarthe DR, DeBarmore BM, Chang AR, et al. Establishing a National Cardiovascular Disease Surveillance System in the United States Using Electronic Health Record Data: Key Strengths and Limitations. *J Am Heart Assoc*. 2022;11:e024409. doi: 10.1161/JAHA.121.024409.
- [24] R.I. Griffiths, C.D. O'Malley, R.J. Herbert, M.D. Danese, Misclassification of incident conditions using claims data: impact of varying the period used to exclude pre-existing disease, *BMC Med Res Methodol*. 13 (2013) 32, <https://doi.org/10.1186/1471-2288-13-32>.
- [25] L. Meyer, B. Stubbs, C. Fahrenbruch, C. Maeda, K. Harmon, M. Eisenberg, J. Drezner, Incidence, causes, and survival trends from cardiovascular-related sudden cardiac arrest in children and young adults 0 to 35 years of age: a 30-year review, *Circulation*. 126 (2012) 1363-1372, <https://doi.org/10.1161/CIRCULATIONAHA.111.076810>.
- [26] R. Williams, E. Kontopantelis, I. Buchan, N. Peek, Clinical code set engineering for reusing EHR data for research: A review, *J Biomed Inform*. 70 (2017) 1-13, <https://doi.org/10.1016/j.jbi.2017.04.010>.
- [27] J.C. Kirby, P. Speltz, L.V. Rasmussen, M. Basford, O. Gottesman, P.L. Peissig, J. A. Pacheco, G. Tromp, J. Pathak, D.S. Carrell, et al., PheKB: a catalog and workflow for creating electronic phenotype algorithms for transportability, *J Am Med Inform Assoc*. 23 (2016) 1046-1052, <https://doi.org/10.1093/jamia/ocv202>.
- [28] Gargano MA, Matentzoglou N, Coleman B, Addo-Lartey EB, Anagnostopoulos AV, Anderson J, Avillach P, Bagley AM, Bakstein E, Balhoff JP, et al. The Human Phenotype Ontology in 2024: phenotypes around the world. *Nucleic Acids Res*. 2024;52:D1333-D1346. doi: 10.1093/nar/gkad1005.
- [29] S. Yu, Y. Ma, J. Gronsbell, T. Cai, A.N. Ananthakrishnan, V.S. Gainer, S. E. Churchill, P. Szolovits, S.N. Murphy, I.S. Kohane, et al., Enabling phenotypic big data with PheNorm, *J Am Med Inform Assoc*. 25 (2018) 54-60, <https://doi.org/10.1093/jamia/ocx111>.
- [30] J. Davidson, A. Banerjee, R. Muzambi, L. Smeeth, C. Warren-Gash, Validity of Acute Cardiovascular Outcome Diagnoses Recorded in European Electronic Health Records: A Systematic Review, *Clin Epidemiol*. 12 (2020) 1095-1111, <https://doi.org/10.2147/CLEP.S265619>.
- [31] D. Agniel, I.S. Kohane, G.M. Weber, Biases in electronic health record data due to processes within the healthcare system: retrospective observational study, *BMJ*. 361 (2018) k1479, <https://doi.org/10.1136/bmj.k1479>.
- [32] B.J. Wells, K.M. Chagin, A.S. Nowacki, M.W. Kattan, Strategies for handling missing data in electronic health record derived data, *EGEMS (wash DC)*. 1 (2013) 1035, <https://doi.org/10.13063/2327-9214.1035>.
- [33] G. Nichol, J. Rumsfeld, B. Eigel, B.S. Abella, D. Labarthe, Y. Hong, R.E. O'Connor, V.N. Mosesso, R.A. Berg, B.B. Leeper, et al., Essential features of designating out-of-hospital cardiac arrest as a reportable event: a scientific statement from the American Heart Association Emergency Cardiovascular Care Committee; Council on Cardiopulmonary, Perioperative, and Critical Care; Council on Cardiovascular Nursing; Council on Clinical Cardiology; and Quality of Care and Outcomes Research Interdisciplinary Working Group, *Circulation*. 117 (2008) 2299-2308, <https://doi.org/10.1161/CIRCULATIONAHA.107.189472>.
- [34] Garcia RA, Girotra S, Jones PG, McNally B, Spertus JA, Chan PS, Group CS. Variation in Out-of-Hospital Cardiac Arrest Survival Across Emergency Medical Service Agencies. *Circulation Cardiovascular quality and outcomes*. 2022;15:e008755. doi: 10.1161/CIRCOUTCOMES.121.008755.
- [35] S.C. Brooks, R.H. Schmicker, S. Cheskes, J. Christenson, A. Craig, M. Daya, P. J. Kudenchuk, G. Nichol, D. Zive, L.J. Morrison, et al., Variability in the initiation of resuscitation attempts by emergency medical services personnel during out-of-hospital cardiac arrest, *Resuscitation*. 117 (2017) 102-108, <https://doi.org/10.1016/j.resuscitation.2017.06.009>.
- [36] Rollman JE, Kloner RA, Bosson N, Niemann JT, Gausche-Hill M, Williams M, Clare C, Tan W, Wang X, Shavelle DM, et al. Emergency Medical Services Responses to Out-of-Hospital Cardiac Arrest and Suspected ST-Segment-Elevation Myocardial Infarction During the COVID-19 Pandemic in Los Angeles County. *J Am Heart Assoc*. 2021;10:e019635. doi: 10.1161/JAHA.120.019635.
- [37] M. Winther-Jensen, M.N. Christiansen, C. Hassager, L. Kober, C. Torp-Pedersen, S. M. Hansen, F. Lippert, E.F. Christensen, J. Kjaergaard, C. Andersson, Age-specific trends in incidence and survival of out-of-hospital cardiac arrest from presumed cardiac cause in Denmark 2002-2014, *Resuscitation*. 152 (2020) 77-85, <https://doi.org/10.1016/j.resuscitation.2020.05.005>.
- [38] Collins GS, Reitsma JB, Altman DG, Moons KG, Group T, Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement, The TRIPOD Group. *Circulation*. 131 (2015) 211-219, <https://doi.org/10.1161/CIRCULATIONAHA.114.014508>.
- [39] D.A. Schoenfeld, Sample-size formula for the proportional-hazards regression model, *Biometrics*. 39 (1983) 499-503.
- [40] S. Khurshid, C. Reeder, L.X. Harrington, P. Singh, G. Sarma, S.F. Friedman, P. Di Achille, N. Diamant, J.W. Cunningham, A.C. Turner, et al., Cohort design and natural language processing to reduce bias in electronic health records research, *NPJ Digit Med*. 5 (2022) 47, <https://doi.org/10.1038/s41746-022-00590-0>.
- [41] D.S. Carrell, J.S. Floyd, S. Gruber, B.L. Hazlehurst, P.J. Heagerty, J.C. Nelson, B. D. Williamson, R. Ball, A general framework for developing computable clinical phenotype algorithms, *J Am Med Inform Assoc*. 31 (2024) 1785-1796, <https://doi.org/10.1093/jamia/ocae121>.
- [42] L. Holmstrom, H. Chugh, K. Nakamura, Z. Bhanji, M. Seifer, A. Uy-Evanado, K. Reinier, D. Ouyang, S.S. Chugh, An ECG-based artificial intelligence model for assessment of sudden cardiac death risk, *Commun Med (lond)*. 4 (2024) 17, <https://doi.org/10.1038/s43856-024-00451-9>.
- [43] M.Z. Kolk, S. Ruiperez-Campillo, A.A. Wilde, R.E. Knops, S.M. Narayan, F.V. Tjong, Prediction of sudden cardiac death using artificial intelligence: Current status and future directions, *Heart Rhythm*. (2024), <https://doi.org/10.1016/j.hrthm.2024.09.003>.
- [44] S.S. Chugh, A. Uy-Evanado, C. Teodorescu, K. Reinier, R. Mariani, K. Gunson, J. Jui, Women have a lower prevalence of structural heart disease as a precursor to sudden cardiac arrest: The Ore-SUDS (Oregon Sudden Unexpected Death Study), *J. Am. Coll. Cardiol.* 54 (2009) 2006-2011, <https://doi.org/10.1016/j.jacc.2009.07.038>.
- [45] Tseng ZH, Ramakrishna S, Salazar JW, Vittinghoff E, Olgin JE, Moffatt E. Sex and Racial Differences in Autopsy-Defined Causes of Presumed Sudden Cardiac Death. *Circ Arrhythm Electrophysiol*. 2021;14:e009393. doi: 10.1161/CIRCEP.120.009393.
- [46] T. Skjelbred, D. Rajan, J. Svane, T.H. Lyng, J. Tfelt-Hansen, Sex differences in sudden cardiac death in a nationwide study of 54 028 deaths, *Heart*. 108 (2022) 1012-1018, <https://doi.org/10.1136/heartjnl-2021-320300>.
- [47] A.C.T. Ha, B.S. Doumouras, C.N. Wang, J. Tranmer, D.S. Lee, Prediction of Sudden Cardiac Arrest in the General Population: Review of Traditional and Emerging Risk Factors, *Can J Cardiol*. 38 (2022) 465-478, <https://doi.org/10.1016/j.cjca.2022.01.007>.
- [48] M.A.E. Haukilahti, T.V. Kentta, J.T. Tikkanen, O. Anttonen, A.L. Aro, T. Kerola, A. Eranti, A. Holkeri, H. Rissanen, M. Heliovaara, et al., Electrocardiographic Risk Markers of Cardiac Death: Gender Differences in the General Population, *Front Physiol*. 11 (2020) 578059, <https://doi.org/10.3389/fphys.2020.578059>.
- [49] S.J. Howell, D. German, A. Bender, F. Phan, S.V. Mukundan, E.A. Perez-Alday, N. M. Rogovoy, K.T. Haq, K. Yang, A. Wirth, et al., Does Sex Modify an Association of Electrophysiological Substrate with Sudden Cardiac Death? The Atherosclerosis Risk in Communities (ARIC) Study, *Cardiovasc Digit Health J*. 1 (2020) 80-88, <https://doi.org/10.1016/j.cvdhj.2020.08.003>.
- [50] S. Zaman, T. Deshmukh, A. Aslam, C. Martin, P. Kovoov, Sex Differences in Electrophysiology, Ventricular Tachyarrhythmia, Cardiac Arrest and Sudden Cardiac Death Following Acute Myocardial Infarction, *Heart Lung Circ*. 29 (2020) 1025-1031, <https://doi.org/10.1016/j.hlc.2019.07.017>.
- [51] K. Reinier, E.C. Stecker, A. Uy-Evanado, H.S. Chugh, A. Binz, K. Nakamura, A. Sargsyan, J. Jui, S.S. Chugh, Sudden Cardiac Death as First Manifestation of Heart Disease in Women: The Oregon Sudden Unexpected Death Study, 2004-2016, *Circulation*. 141 (2020) 606-608, <https://doi.org/10.1161/circulationaha.119.044169>.

- [52] K. Reinier, G.A. Nichols, A. Huertas-Vazquez, A. Uy-Evanado, C. Teodorescu, E. C. Stecker, K. Gunson, J. Jui, S.S. Chugh, Distinctive Clinical Profile of Blacks Versus Whites Presenting With Sudden Cardiac Arrest, *Circulation*. 132 (2015) 380–387, <https://doi.org/10.1161/CIRCULATIONAHA.115.015673>.
- [53] R. Deo, M.M. Safford, Y.A. Khodneva, D.P. Jannat-Khah, T.M. Brown, S.E. Judd, W. M. McClellan, J.D. Rhodes, M.G. Shlipak, E.Z. Soliman, et al., Differences in Risk of Sudden Cardiac Death Between Blacks and Whites, *J. Am. Coll. Cardiol.* 72 (2018) 2431–2439, <https://doi.org/10.1016/j.jacc.2018.08.2173>.
- [54] D. Zhao, W.S. Post, E. Blasco-Colmenares, A. Cheng, Y. Zhang, R. Deo, R. Pastor-Barriuso, E.D. Michos, N. Sotoodehnia, E. Guallar, Racial Differences in Sudden Cardiac Death, *Circulation*. 139 (2019) 1688–1697, <https://doi.org/10.1161/CIRCULATIONAHA.118.036553>.
- [55] Reinier K, Moon JY, Chugh HS, Sargsyan A, Nakamura K, Norby FL, Uy-Evanado A, Talavera GA, Gallo LC, Daviglius ML, et al. Risk Factors for Sudden Cardiac Arrest Among Hispanic or Latino Adults in Southern California: Ventura PRESTO and HCHS/SOL. *J Am Heart Assoc.* 2023;12:e030062. doi: 10.1161/JAHA.123.030062.
- [56] Butters A, Arnott C, Sweeting J, Winkel BG, Semsarian C, Ingles J. Sex Disparities in Sudden Cardiac Death. *Circ Arrhythm Electrophysiol.* 2021;14:e009834. doi: 10.1161/CIRCEP.121.009834.
- [57] Reinier K, Sargsyan A, Chugh HS, Nakamura K, Uy-Evanado A, Klebe D, Kaplan R, Haddock K, Shepherd D, Young C, et al. Evaluation of Sudden Cardiac Arrest by Race/Ethnicity Among Residents of Ventura County, California, 2015-2020. *JAMA Netw Open.* 2021;4:e2118537. doi: 10.1001/jamanetworkopen.2021.18537.