# ARIC Manuscript Proposal #2532

**PC Reviewed:  4/14/15**  **Status: <u>A</u>**  **Priority: <u>2</u>**
**SC Reviewed: _____**  **Status: _____**  **Priority: ____**

**1.a.**  **Full Title**:  "Multimodality Cardiovascular Risk Prediction; Implications for Astronaut Selection and Monitoring During Prolonged Spaceflight, And For Risk Prediction in the General Population"

  **b.**  **Abbreviated Title (Length 26 characters)**: Multimodale CVRisk Predict

**2.**  **Writing Group**: James de Lemos, MD; Christie M. Ballantyne, MD; Thomas Wang, MD; Matthew Budoff, MD; Jarett Berry, MD; William G. Hundley, MD; Amit Khera, MD; Benjamin Levine, MD

I, the first author, confirm that all the coauthors have given their approval for this manuscript proposal. __JAD__ **[please confirm with your initials electronically or in writing]**

F**irst author**:  Name:  **James Andrew de Lemos, MD**
          Address:      UT Southwestern Medical Center
                        5323 Harry Hines Blvd
                        Dallas, TX 75390
          Phone:        214-645-7528 or 214-648-7610  Fax:  214-645-7501
          E-mail:       James.delemos@UTSouthwestern.edu

**ARIC author** to be contacted if there are questions about the manuscript and the first author does not respond or cannot be located (this must be an ARIC investigator).
          Name:  **Christie M. Ballantyne, MD, FACP, FACC**
          Address:      Baylor College of Medicine
                        6565 Fannin, Suite B157
                        Houston, TX  77030
          Phone:        713-798-5034   Fax:  713-798-7885
          E-mail:       cmb@bcm.edu

**3.**  **Timeline**:        5/1/2014 to 4/30/2017

**4.**  **Rationale**:      The most likely cause of a non-traumatic life- or mission-threatening medical event in astronauts would be from acute cardiovascular disease (CVD). Risk prediction models in current use—both for astronauts as well as the general population-- utilize only traditional atherosclerosis risk factors and focus narrowly on atherosclerotic events rather than global cardiovascular risk, ignoring outcomes such as heart failure or atrial fibrillation that could also be potentially mission-threatening. To date, only a single global CVD risk model has been developed, and this studied only traditional ASCVD risk factors. (1)

A second issue of particular relevance for NASA is the need for risk prediction models with different time horizons. Existing risk prediction models focus on a 10-year or longer time horizon, but screening prior to a manned Mars mission requires a focus on a shorter time window. Addressing short-term risk prediction will require collaborative studies, due to the small numbers

of CVD events in this time window within individual population studies. In addition, it is likely that different risk markers will emerge in short-term compared with longer-term risk models. For example, while traditional risk factors provide stable risk estimates over time, markers of existing CVD may demonstrate decrementing relative risk over time,(2) suggesting that "near term" screening for subclinical CVD with imaging studies prior to mission approval may be an effective strategy.

New risk prediction strategies are clearly needed, because traditional risk factor scores discriminate CV risk only moderately well,(3) and large numbers of cardiovascular events occur among individuals predicted to be at low risk by these risk calculators. Numerous studies have evaluated novel risk markers in an attempt to improve CVD risk prediction, with several promising imaging and blood-based biomarkers identified. These include: a) specific protein biomarkers such as high sensitivity troponin and NT-proBNP;(4-6)  b) combinations of protein biomarkers that reflect relevant ongoing biologic process such as hemodynamic stress, inflammation, or vascular injury;(7,8)  c) specific genetic variants;(9) and d) high resolution imaging techniques  such as coronary artery calcium (10)  and cardiovascular MRI which can very accurately assess atherosclerosis burden, cardiac structure, scar, and fibrosis.(11)

Most of the studies evaluating these technologies have investigated the incremental predictive value of a single biomarker added to a traditional risk factor model, with only a few reporting combinations of biomarkers. Those that have evaluated biomarker combinations have focused on biomarkers within the same testing modality, such as combinations of protein markers.(7,8) Few studies have evaluated strategies for risk prediction that **cross testing modalities**. Such a multi-modality approach has the potential to markedly improve CVD risk prediction among potential and existing astronauts, and would have direct relevance to the general population.

## 5.    Main Hypothesis/Study Questions:

To develop two distinct multi-modality risk prediction tools for prediction of cardiovascular risk in astronauts and astronaut candidates, one based on 10-year global CVD risk (the career duration of the astronaut) and one based on 3-year CVD risk (the likely duration of a manned spaceflight to Mars). These models will sequentially evaluate novel testing modalities on top of standard risk factors, including coronary calcium and/or carotid IMT (measures of atherosclerosis burden), blood based protein biomarkers that reflect inflammation, cardiac injury and cardiac stress, ECG measurements of left ventricular hypertrophy, as well as imaging-based assessments of cardiac structure and function.

**Aim 1:** To develop a multi-modality risk prediction tool for 10 year global CVD events that combines imaging and novel blood-based biomarkers.

**Hypothesis 1:** Risk prediction strategies that <u>cross testing modalities</u> and include a hierarchy of blood-based and imaging biomarkers will improve prediction of 10-year global cardiovascular event rates in middle aged men and women**.

**Aim 2:** To develop a risk prediction tool for short-term (3 year) global CVD events.

**Hypothesis 2:** Distinct risk markers will identify individuals at increased risk for *short-term* global CVD events. Markers of existing subclinical disease, including imaging-based measures of atherosclerosis and cardiac structure and function, and biomarkers of subclinical myocardial injury and neurohormonal activation will enhance short-term risk prediction.

## 6. Design and analysis (study design, inclusion/exclusion, outcome and other variables of interest with specific reference to the time of their collection, summary

**of data analysis, and any anticipated methodologic limitations or challenges if present).**

**CONCISE SUMMARY OF PROJECT:** The team of collaborative investigators will pool data from multiple existing cohort studies (MESA, ARIC, Framingham Heart Study, Dallas Heart Study) to develop two distinct multi-modality risk prediction tools, one based on 10-year global CVD risk and one based on 3-year CVD risk. These models will sequentially evaluate novel testing modalities on top of standard risk factors, including coronary calcium (a measure of the extent of coronary atherosclerosis), blood based protein biomarkers that reflect inflammation, cardiac injury and cardiac stress (hs-cTnT, NT-proBNP, and hs-CRP), 12-lead ECG assessment of left ventricular hypertrophy (LVH), as well as imaging-based assessments of cardiac structure and function by MRI or echo. No new data will be collected for the study, as this will represent a pooling project of existing data across different cohorts. The primary endpoint will be global CVD events, including cardiovascular death, nonfatal MI, stroke, coronary or peripheral revascularization, and congestive heart failure. We will also evaluate a secondary endpoint that includes the primary endpoint plus atrial fibrillation. As tertiary endpoints, we will consider individual endpoint components as well as "hard" CHD (coronary death or MI) and "hard" CVD (cardiovascular death, MI, stroke, heart failure). These "hard" endpoints will be of interest to the non-astronaut community, as the implications of revascularization events may be less in the general population than the astronaut corps.

**CRITERIA FOR INCLUSION OF SUBJECTS:** The study will include all asymptomatic individuals with biomarker data available from ARIC visit 4.

**CRITERIA FOR EXCLUSION OF SUBJECTS:** Participants with prevalent cardiovascular disease (prior MI, stroke, coronary or extra-coronary revascularization, heart failure, or atrial fibrillation) will be excluded.

**SOURCES OF RESEARCH MATERIAL:** De-identified data present in the ARIC Database.

**RECRUITMENT OF SUBJECTS:** Not applicable, no subject recruitment will be required.

**POTENTIAL RISKS:** There are no risks to any individuals because all data collected for this database is either anonymous or de-identified.

**SPECIAL PRECAUTIONS:** No special precautions are required because all data is de-identified; however only members of the research team will have access to this data.

**PROCEDURES TO MAINTAIN CONFIDENTIALITY:** All information will be kept electronically, and password protected. Only de-identified data will be used. Only members of the research team will have access to the data.

**POTENTIAL BENEFITS:** By combining data from multiple well defined non-astronaut populations, we will establish the utility of novel biomarkers for predicting long term (10yr) and short term (2-5 year) risk in an "astronaut like" demographic; this effort will allow mission managers to weigh precisely the relative risk of mission threatening cardiovascular events in the context of other non-medical mission concerns. This will also inform cardiovascular risk assessment for non-astronauts.

**FUNDING** Investigator and statistical support is provided by grant 13-13NSBRI2-0016 from the National Space Biomedical Research Institute (de Lemos, PI).

**DATA REQUESTED (visit 4)**

Demographics and Risk Factors: Age, sex, race/ethnicity; hypertension, diabetes, smoking status, family history of CAD; continuous measures of total, HDL, LDL cholesterol and triglycerides, systolic and diastolic BP, serum creatinine and eGFR, BMI.

<u>ECG</u> LVH by Cornell criteria

<u>Biomarkers</u>: hs-CRP, hs-cTnT, NT-proBNP, cystatin C.  If available, GDF-15, Gal-3 and sST2.

<u>Genetics:</u> 9P21 variant status

<u>Carotid IMT: common carotid IMT, carotid plaque from visit 4 when available, and if not visit 3</u>

<u>Echo  (nearest to visit 4, among subgroup with echo data):</u>  Categorical LVH; concentric LVH, eccentric LVH, Continuous measures of LV diameter, LV mass, LV mass index, LVEF, LV posterior wall thickness, relative wall thicknessmean mid-wall circumferential strain, E/A ratio

<u>Cardiovascular Endpoints (with dates):</u> All cause, CV, and CHD death, MI, PCI or CABG, stroke, heart failure hospitalization, atrial fibrillation.


**STATISTICAL APPROACH:**  Primary participant level data for demographic information, risk factors, CAC scores, biomarkers, imaging data, and CV outcomes will be obtained from the individual studies and the data will be pooled.  Cox proportional hazard models will be created to estimate the risk of global CVD events with standard traditional risk factors included as independent variables (age, sex, total and HDL cholesterol, current smoking, systolic blood pressure, and antihypertensive therapy), and family history of MI .  We will also include estimated creatinine clearance in the base model given its strong associations with many of the candidate markers and with global CVD outcomes.  Discrimination of the model will be assessed by calculating the time-dependent Harrell C statistic, and calibration will be assessed using the modified Hosmer Lemeshow test.  To evaluate the additional prognostic value of the candidate biomarkers and imaging studies, each biomarker will initially be tested individually and then in combinations to determine the association of the novel biomarker, and combinations of markers, with the CVD endpoint after adjusting for the baseline variables.  NT-proBNP will be evaluated as a continuous variable, and hs-cTnT as both a categorical (detectable/undetectable) and as a semi-continuous variable using a piecewise linear model.  ECG-LVH will be characterized as a dichotomous variable (present/absent) using Cornell criteria. The metrics for evaluating performance of the new markers will include improvement in the c-statistic, goodness of fit and calibration, as well as the Bayes Information Criteria (BIC), and overall model log-likelihood test. We will also evaluate net reclassification improvement and the integrated discrimination index when the new markers are added, using methods described by Pencina et al.  To correct discrimination and calibration for the over-optimism inherent in self-testing, we will use bootstrap sampling (n=1,000).   The potential role of hs-CRP will be evaluated as well, using similar strategies as described for the other biomarkers, although we do not anticipate that hs-CRP will improve risk prediction.

Validation will be performed using bootstrapping, as opposed to data splitting or cross-validation, as bootstrapping has the advantage that model development will be carried out on the entire dataset. In fact, fewer model fits are required than with cross-validation. Furthermore, the estimates of predictive accuracy will be unbiased by utilizing the bootstrapping technique. Bootstrap corrected performance characteristics will be derived by subtracting the optimism estimate from the apparent estimate. These measures of discrimination, calibration, and beta estimates will be honest estimates of internal validity, penalizing for overfitting.

All analyses will be conducted primarily using a complete-case method. However, we will also consider multiple imputation strategies which result in valid statistical inferences that properly reflect the uncertainty due to missing values. We propose to generate $m$ $(10{\leq}m)$ complete data sets, with results from the $m$ complete data sets combined for the inference. We will assume that the data are missing at random. In particular, we will use the Markov Chain Monte Carlo method,

which will simulate the entire joint posterior distribution of the unknown quantities and obtain simulation-based estimates of posterior parameters that are of interest. SAS (SAS Institute, Cary, NC) software will be used for all analyses.

**7.a.  Will the data be used for non-CVD analysis in this manuscript?**
**\_\_\_\_ Yes    \_\_X\_\_ No**

  **b. If Yes, is the author aware that the file ICTDER03 must be used to exclude persons with a value RES_OTH = "CVD Research" for non-DNA analysis, and for DNA analysis RES_DNA = "CVD Research" would be used?**
    **\_\_X\_\_ Yes    \_\_\_\_ No**

    (This file ICTDER has been distributed to ARIC PIs, and contains the responses to consent updates related to stored sample use for research.)

**8.a.  Will the DNA data be used in this manuscript? \_\_X\_\_ Yes    \_\_\_\_ No**

**8.b.  If yes, is the author aware that either DNA data distributed by the Coordinating Center must be used, or the file ICTDER03 must be used to exclude those with value RES_DNA = "No use/storage DNA"? \_\_X\_\_ Yes \_\_\_\_ No**

**9.  The lead author of this manuscript proposal has reviewed the list of existing ARIC Study manuscript proposals and has found no overlap between this proposal and previously approved manuscript proposals either published or still in active status.**  ARIC Investigators have access to the publications lists under the Study Members Area of the web site at:  http://www.cscc.unc.edu/ARIC/search.php

    \_\_X\_\_ Yes    _____ No

**10. What are the most related manuscript proposals in ARIC (authors are encouraged to contact lead authors of these proposals for comments on the new proposal or collaboration)?**
The related proposals and papers include those reporting associations of hs-cTnT and NT-proBNP with outcomes. The senior author on these papers was Christie Ballantyne, who is a co-PI on this NSBRI grant. James de Lemos (the PI on the NSBRI grant and of this paper request) was a coauthor on many of these papers. Thus, ARIC investigators who contributed the data are represented in this writing group.

**11.a. Is this manuscript proposal associated with any ARIC ancillary studies or use any ancillary study data? \_\_\_\_ Yes    \_\_X\_\_ No**

**11.b. If yes, is the proposal**
        **\_\_\_    A. primarily the result of an ancillary study (list number\* _____)**
        **\_\_\_    B. primarily based on ARIC data with ancillary data playing a minor role (usually control variables; list number(s)\* _____)**
\*ancillary studies are listed by number at http://www.cscc.unc.edu/aric/forms/

**12a. Manuscript preparation is expected to be completed in one to three years.  If a manuscript is not submitted for ARIC review at the end of the 3-years from the date of the approval, the manuscript proposal will expire.**

**12b. The NIH instituted a Public Access Policy in April, 2008** which ensures that the public has access to the published results of NIH funded research.  It is **your responsibility to upload manuscripts to PUBMED Central** whenever the journal does not and be in compliance with this policy.  Four files about the public access policy from http://publicaccess.nih.gov/ are posted in http://www.cscc.unc.edu/aric/index.php, under Publications, Policies & Forms. http://publicaccess.nih.gov/submit_process_journals.htm shows you which journals automatically upload articles to Pubmed central.