Sources of Variability in Measurements of Cardiac Troponin T in a Community-Based Sample: The Atherosclerosis Risk in Communities Study

Sunil K. Agarwal,1,2 Christy L. Avery,1 Christie M. Ballantyne,3 Diane Catellier,4 Vijay Nambi,3 Justin Saunders,3 A. Richey Sharrett,5 Josef Coresh,5 Gerardo Heiss,1* and Ron C. Hoogeveen2

BACKGROUND: Application of cardiac troponin T (cTnT) as a marker of myocyte damage requires knowledge of its measurement variability. Using a highly sensitive assay for measurement, we evaluated the long-term storage stability of plasma cTnT at −70 °C and the sources of cTnT variability.

METHODS: Samples from the Atherosclerosis Risk in Communities study collected in 1996–1998 and 2005–2006 were assayed centrally to quantify variability in cTnT attributable to processing (replicates from same blood draw, n = 87), laboratory (replicates after freeze thaw, n = 29), short-term (n = 40) and long-term biological variation (repeat visit, n = 38), and degradation in frozen storage (n = 7677).

RESULTS: Approximately 30% of this population-based cohort had cTnT concentrations below the detection limit (3 ng/L). Reliability coefficients for all paired comparisons exceeded 0.93 except for samples drawn 8 years apart (r = 0.36). Sources of cTnT variation (as CVs) were: laboratory, 2.1% and 11.2% in those with and without heart failure, respectively; processing, 18.3%; biological, 16.6% at 6 weeks and 48.4% at 8 years. The reference change value at 6 weeks (68.5%) indicated that 4 samples are needed to determine a homeostatic set point within ±25%. The estimated cTnT degradation rate over the first year in long-term frozen storage was 0.36 ng/L per year.

CONCLUSIONS: cTnT was detectable in approximately 70% of community-dwelling middle-aged study participants and stable in −70 °C storage. The variability in cTnT attributable to 1 freeze–thaw cycle is of small magnitude. The observed high laboratory and intra-individual (biological) reliability of cTnT support its use for population-based research, and in clinical settings that rely on classification and serial measurements.

Cardiac troponins are markers of myocardial injury commonly used in the diagnosis of acute coronary events (1–4). Increased concentrations of circulating troponins are also detectable in patients with acute decompensated heart failure (HF) (5) and chronic HF and have been found to be associated with poor patient outcomes (5–12). Troponin concentrations are generally lower in patients with chronic HF than in patients with acute coronary syndromes (13), and therefore sensitive troponin assays are necessary. For example, the highly sensitive cardiac troponin T (cTnT) assay can detect troponin T concentrations more than 10-fold lower than those detected with traditional assays (14).

There is limited information on the variability of cTnT in individuals not examined in an acute care context, although a 1999 report from the National Academy of Clinical Biochemistry Standards of Laboratory Practice emphasized a need for such studies (15). Such information is important to determine cutoffs to suggest an acute event in individuals with chronic troponin T elevation. The variability of cTnT was recently reported to be higher than that of the commonly used troponin I (16).

We examined the variability of cTnT in individuals who participated in the Atherosclerosis Risk in Communities (ARIC) study and quantified the components of variability attributed to laboratory, processing, and biological variation. We also estimated the impact on cTnT concentrations of specimen storage at −70 °C over a 3-year period.

1 Departments of Epidemiology and 4 Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, NC; 2 Department of Internal Medicine, University of Pittsburgh, Pittsburgh, PA; 3 Department of Medicine, Baylor College of Medicine, Houston, TX; 5 Department of Epidemiology, Johns Hopkins School of Public Health, Baltimore, MD.
* Address correspondence to this author at: University of North Carolina School of Public Health, Department of Epidemiology, 137 E. Franklin St., Suite 306, Chapel Hill, NC 27514-3628. E-mail gerardo_heiss@unc.edu.

Received November 30, 2010; accepted March 17, 2011. Previously published online at DOI: 10.1373/clinchem.2010.159350

© 2011 American Association for Clinical Chemistry
Materials and Methods

STUDY SAMPLE
The ARIC study cohort of African American and white males and females aged 45–64 years was sampled from 4 US communities and enrolled in 1987–1989. The baseline examination was followed by 3 triennial reexamination visits. Subsequently, novel cellular, metabolic, and genomic correlates of carotid atherosclerosis and early pathologic changes in the carotid artery walls were measured in 2066 cohort members in the ARIC Carotid MRI (Car-MRI) study (2005–2006). The present study assayed stored samples from the 2005–2006 examination to quantify the components of variability, and specimens from the 1996–1998 examination to estimate the stability of stored frozen cTnT over 3 years of sample collection. The study was approved by the institutional review boards of the 4 participating centers.

BLOOD SAMPLING AND PROCESSING
Participants were requested to report after an overnight fast. Venipuncture and sample processing were performed by centrally trained technicians, using standard protocols, at the 4 field centers. After centrifugation at 3000g for 10 min at 4 °C, the supernatant plasma was stored at −70 °C, and shipped weekly to the ARIC Central Laboratory (17). The specimens were stored at −70 °C until assayed for this study in 2010.

BIOMARKER ASSAYS
Plasma samples were stored centrally at −70 °C and used for measurement of the biomarkers. cTnT concentrations were measured with a novel precommercial highly sensitive assay, Elecsys Troponin T (Roche Diagnostics), on an automated Cobas e411 analyzer with a lower limit of detection (LOD) of 3 ng/L. The 99th percentile value for the cTnT measured by the new assay in a healthy subpopulation is 14 ng/L (Roche Diagnostics, data on file). The between-assay CVs were 2.6% and 6.9% for control materials, with mean cTnT concentrations of 2378 ng/L and 29 ng/L, respectively [based on 103 separate (daily) runs for each control level].

STUDY DESIGN
A total of 5 substudies (I to V) were conducted, 4 of them to estimate the components of variability, i.e., laboratory, processing, short-term biological, and long-term biological, and the fifth to estimate degradation over 3 years. A schematic diagram showing various components of variability in the measurement of an assay is shown in Fig. 1. In substudy I examining laboratory variability, 30 samples (15 each from individuals with and without HF) from the Car-MRI study were split into 2 aliquots just before the first aliquot was assayed; the second aliquot was stored for 2 days before being assayed. The paired assays were compared and measures of variation estimated. One participant without HF had cTnT concentrations below the LOD in the split assay; thus information on 14 participants was available for estimation. Measurement variation estimated from these data cannot be attributed to variation in blood drawing, local processing, shipment procedures, or within-subject variation over time.

In substudy II involving 120 Car-MRI study participants, each field center drew duplicate samples into blood tubes using a single venipuncture during the
same visit, but the tubes were processed (shipping and other details) separately. These duplicate samples were sent to the central laboratory under a blinded QC ID that was indistinguishable from other IDs. These paired samples (n = 87 pairs with concentrations >LOD) were used to estimate the processing variability (i.e., variability in blood processing, shipping, and laboratory handling and analysis).

In substudy III involving 60 Car-MRI study participants, each field center was asked to recruit 15 volunteers to repeat the entire clinic visit within 4–8 weeks of their original visit. Volunteers generally reflected the age, sex, and racial composition of the overall study population (18). Again, duplicate samples were submitted to the central laboratory and stored under a blinded QC ID. Results from the paired samples (n = 40 pairs with values >LOD) were compared to give an estimate of short-term biological variability. It is to be noted that in this and all similar studies paired samples from participants over time will also include laboratory and processing variability.

In substudy IV (n = 161 participants), samples from the Car-MRI study visit (substudies I–III) and corresponding assays from samples obtained during visit 4 (approximately 8 years before) were assayed, and n = 38 (those with a value > LOD at both visits) were used to evaluate long-term biological variability. No differences in the results were seen after exclusion of study participants with coronary heart disease (CHD), HF, or stroke.

In substudy V, stored samples from all the participants in the ARIC cohort field center visit 4 (n = 11220) that had cTnT above the LOD were used (n = 7677) to evaluate analytic stability when stored at −70 °C over approximately 36 months.

**COVARIATES**

Smoking status, race, and sex were self-reported. Estimated glomerular filtration rate was calculated using the Modification of Diet in Renal Disease formula (19). Prevalent CHD was defined as either self-reported history of CHD at the baseline visit or an adjudicated CHD event before the fourth examination visit (for the visit 4 analyses) or December 31, 2004, the last date of a CHD event before the fourth examination visit (for the Car-MRI study visit). Participants with hospitalizations listing the International Classification of Diseases, Ninth Revision discharge code 428 in any position were classified as having prevalent HF.

**STATISTICAL METHODS**

Descriptive statistics are presented for participants in each of the substudies (I through V) and participants in the Car-MRI study. By treating paired measurements as a random effect in a linear mixed effects model, we partitioned the total variance (σ²TOT) into a between-pair (or between-person; σ²BP) and within-pair component of variance. The within-pair component of variance derived from within-visit reliability substudy II, in which duplicate samples were obtained from participants on the same day, corresponds to an estimate of variation due to method variation: a combination of variation in blood collection, processing, and laboratory analysis (σ²). In contrast, the within-pair component of variance derived from between-visit reliability substudy III and substudy IV, in which duplicate samples were obtained from participants at 2 separate visits, corresponds to an estimate of the within-person (biological) variation over time plus method variation (σ²BP + σ²). The within-pair variance derived from the split-samples after 1 freeze–thaw cycle (substudy I) corresponds to the laboratory variability plus analyte stability. The proportion of the total variance attributable to between-person variability, or the reliability coefficient (r = σ²BP/σ²TOT) can be interpreted as the correlation between paired measurements. The following benchmarks were used for characterization of the adequacy of reliability (20): slight reliability, 0–0.2; fair reliability, 0.21–0.4; moderate reliability, 0.41–0.6; substantial reliability, 0.61–0.8; almost perfect reliability, 0.81–1.0. Based on our sample sizes of 87 and 40 for the 2 substudies, the 95% CIs, assuming a moderate reliability of 0.60, will have lower limits of 0.45 and 0.38, respectively. If the estimated reliability coefficient is greater than the lower limit of the 95% CI, we expect that it lies within the bounds of moderate reliability.

The CV was derived as the SD of the within-pair differences divided by the mean of the paired observations multiplied by 100. CV values >10% for laboratory or processing variability were considered to be a cause for concern.

A reference change values (RCV) was defined as a difference between 2 consecutive test results in an individual that is statistically significant in a given proportion of all similar persons (21). For estimating RCVs, those below the LOD were not included. RCVs were calculated as: \(2^{1/2} \times \sqrt{Z^2 (CV_A^2 + CV_I^2)}\), where \(CV_A\) denotes analytical CV, \(CV_I\) denotes within-subject CV, and Z is the multiplier for the SD required for a stated probability under the normal curve. The number of samples required to produce a precise homeostatic set point estimate is given as:

\[
n = \left( \frac{Z \times CV_A + CV_I}{D} \right)^2,
\]

where D is the desired percentage closeness to the homeostatic set point (22). Bland–Altman plots (substudy III and IV) were used to assess the degree of disagreement (including systematic differences) and
identify outliers and trends between the 2 repeat measures at 4–8 weeks apart (23).

Although our study design precluded performing cTnT assays at prespecified time intervals to evaluate sample degradation at −70 °C storage over time, the timing of the cohort examination visits in ARIC was assigned at random, resulting in a uniform distribution of participants by study month. Degradation in stored frozen storage was therefore examined by regressing analyte concentration on time since specimen collection. Specifically, the ln (cTnT) concentration was regressed on time in storage. A statistically significant slope (i.e., \( P < 0.05 \)) would suggest evidence of degradation. All analyses were performed using SAS software version 9.2.

Results

Demographic and clinical characteristics of study participants for substudies I through V are shown in Table 1. Median age at sample collection was 63 years at visit 4 (long-term variability and degradation study) and older than 70 years for substudies I through III, using only Car-MRI visit data. The mean prevalence of co-morbidities among participants for various substudies during visit 4 vs Car-MRI visit ranged from 6% to 8% vs 15% to 18% for CHD and 2.4% to 2.5% vs 5.0% to 6.0% for HF, respectively.

Estimates of the 95th and 99th percentiles of cTnT in the ARIC visit 4 cohort (n = 10870) were 18 ng/L and 40 ng/L, with an interquartile range (IQR) of 3–8 ng/L. Also, 31.6% of visit 4 and 26.9% of Car-MRI cTnT samples were below the LOD.

High reliability (\( r > 0.93 \)) was seen for cTnT measures repeated after a single freeze–thaw cycle, between specimens split in the course of a single venous stick, and also between specimens collected at repeat visits 4–8 weeks apart (Table 2). However, reliability after about 8 years was low: \( r = 0.36 \) (Table 2).

The mean of cTnT was 24.5 ng/L in those with HF vs 8.6 ng/L in those without HF. The CVs in split replicates after a freeze–thaw cycle in those with and without HF were 2.1% and 11.2%, respectively. There was high processing variability with CV of 18.3% (though reliability was high, \( r = 0.93 \)).

The variability between repeated visits 4 to 8 weeks apart was not large (CV, 16.6%) compared to processing variability (CV, 18.3%). The mean of the differences in cTnT concentrations between these repeat visits was 0 ng/L, with a range from −5 to 6 ng/L and IQR from −1 ng/L to 1 ng/L. Bland–Altman plots (Fig. 2) did not support a relationship between differences in cTnT measures and the mean cTnT concentration. The RCV, calculated by using results for samples from repeat visits spaced approximately 6 weeks apart, was 68.5%, suggesting that cTnT changes over a 6-week interval.

| Table 1. Demographic and clinical characteristics of ARIC participants who contributed to the within-visit and between-visit replicate samples.\(^a\) |
|---|---|---|---|---|---|---|
| | Substudy I | Substudy II | Substudy III | Substudy IV | Substudy V |
| **Laboratory variability, HF** | 15 | 15 | 60 | 161 | 11,220 |
| **N** | 15 | 15 | 120 | 60 | 161 |
| **Age, median (range), years** | 77 (69–81) | 73 (69–81) | 71 (60–82) | 71 (61–82) | 63 (53–75) |
| **Male, %** | 53.3 | 46.7 | 47.9 | 57.4 | 50.3 |
| **African American, %** | 26.7 | 33.3 | 21.0 | 29.5 | 23.0 |
| **BMI, median (range)** | 31.0 (23.7–43.3) | 27.6 (21.8–45.5) | 28.6 (19.7–57.1) | 26.5 (19.8–46.2) | 27.3 (16.8–56.4) |
| **eGFR, median (range)** | 59.4 (35.3–86.9) | 80.6 (65.9–143.1) | 78.2 (29.6–189.5) | 82.0 (29.6–228.8) | 104.4 (57.4–239.6) |
| **Prevalent CHD, %** | 66.7 | 0 | 14.4 | 18.0 | 6.3 |
| **Prevalent HF, %** | 100 | 0 | 5.9 | 4.9 | 2.5 |
| **Current smokers, %** | 7.1 | 0 | 6.0 | 3.3 | 11.8 |

\(^a\) Presented as percentages for dichotomous variables and median (range) for continuous variables.

\(^b\) Calculated during the 4th ARIC clinic visit. ARIC, Atherosclerosis Risk in Communities Study; BMI, body mass index; CHD, coronary heart disease; eGFR, estimated glomerular filtration rate; IMT, intima medial thickness.

\(^c\) BMI, body mass index; eGFR, estimated glomerular filtration rate.
interval must be >67% of the baseline value to be considered biologically significant (Table 3). Correspondingly, for a 6-week interval, 23 and 4 samples would be needed to identify an individual’s cTnT (i.e., homeostatic set point) concentration with 95% confidence within ±10% and ±25%, respectively (Table 3).

Examination of long-term (i.e., 8-year) biological variability suggested an SD of 4.3 ng/L at a mean concentration of 9 ng/L (CV, 48.4%). The mean of differences between these visits 8 years apart was 1.6 ng/L with range from −16 ng/L to 14 ng/L, and an IQR from −4 ng/L to 1 ng/L. There was evidence of late degradation when stored samples collected over a period of 36 months were regressed on time since collection. A mean degradation of 0.36 (95% CI, 0.19 – 0.53) ng/L in cTnT per year was estimated during year 1 (Fig. 3). This degradation was not evident when we examined proportions of samples below the LOD collected during a quarter year with varying duration of storage since collection during visit 4 to measurement in 2009–2010 (online Supplementary Fig. 1).

Discussion

In this study we comprehensively investigated various components of variability in the measurement of high-sensitivity troponin T and on the stability of troponin T

---

**Table 2. Sample means, SDs, reliability coefficients, and CVs for cTnT replicate samples in the ARIC (1996–1998) and ARIC Car-MRI (2005–2006) studies.**

| Substudy (number) | QC pairs, n | Mean, ng/L | SD | CV | r
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Laboratory reliability (I)</td>
<td>15</td>
<td>24.5</td>
<td>0.51</td>
<td>0.99</td>
<td>2.1</td>
</tr>
<tr>
<td>Without HF</td>
<td>14</td>
<td>8.6</td>
<td>0.96</td>
<td>0.94</td>
<td>11.2</td>
</tr>
<tr>
<td>Processing reliability (II)</td>
<td>87</td>
<td>9.4</td>
<td>1.7</td>
<td>0.93</td>
<td>18.3</td>
</tr>
<tr>
<td>Intraindividual variability (III and IV)</td>
<td>6 weeks</td>
<td>40</td>
<td>9.7</td>
<td>1.6</td>
<td>0.94</td>
</tr>
<tr>
<td>8 years</td>
<td>38</td>
<td>9.0</td>
<td>4.3</td>
<td>0.36</td>
<td>48.4</td>
</tr>
</tbody>
</table>

a SD = square root (within-subject variance).

b The reliability coefficient (r) serves as an estimate of correlation between repeated measurements.

c Laboratory SD expressed as a percentage of the mean of QC pairs. Laboratory reliability is estimated from split aliquots, the first assayed immediately and the second after a 2-day freeze–thaw cycle. The variation can be attributed to measurement error and a freeze–thaw cycle. Processing reliability is estimated by assaying duplicate samples collected from a single venipuncture but shipped separately to the central laboratory. Differences in shipping, handling, and laboratory variability contribute to this variability.

**Table 3. RCVs and number of samples needed to determine the homeostatic set point for cTnT over 6-week measurement intervals in the ARIC Car-MRI Study, 2005–2006.**

<table>
<thead>
<tr>
<th>RCV, %</th>
<th>No. samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>68.5</td>
<td>23</td>
</tr>
<tr>
<td>(±10; CI 95%)</td>
<td>4</td>
</tr>
<tr>
<td>(±25; CI 95%)</td>
<td>1</td>
</tr>
</tbody>
</table>

---

**Fig. 2. Bland–Altman plot showing bias against the mean of cTnT measurements obtained 6 weeks (A) and 8 years (B) apart with 95% levels of agreement (broken lines).** The differences (repeat visit – initial visit) and means are expressed in nanograms per liter.
in plasma samples stored at −70 °C. The laboratory reliability of the cTnT assay was high: the reliability coefficient \( r \) was 0.99 for samples from patients with HF and 0.94 for samples from patients without HF. The intraindividual (biological) reliability was similarly high after 6 weeks (\( r = 0.94 \)), but was lower after 8 years (\( r = 0.36 \)). After we excluded samples from patients with events (CHD, HF, and stroke) between the repeat visits, \( r \) changed to 0.28 and the CV was 54.9% (compared to 48.4% in the complete sample).

In another study in which the precommercial electrochemiluminescence assay (Elecsys 2010 analyzer, Roche Diagnostics) was used, the LOD was 1 ng/L \((14)\). A significant correlation was reported between the highly sensitive cTnT assay and the traditional cTnT assay in samples with detectable concentrations \((r = 0.84)\). The intraassay CVs were 5% at 10 ng/L and 1% at 100 ng/L \((14)\).

In a recent study performed to examine components of variability in 20 healthy and young volunteers (age range 25–56 years), very short-term (0–4 h) and short-term (0–8 weeks) intraindividual CVs were 48% and 85%, respectively \((16)\). The CV at short-term (85%) was higher than that seen in the study reported here (16.6%), a result that was possibly a function of the lower mean value of cTnT in a study of healthy young volunteers. In our study there were no appreciable changes in the intraindividual CV (18.5%) after exclusion of samples from study participants with CHD, HF, or AF.

We are not aware of published studies performed to investigate the stability of cTnT in stored specimens. The use of an indirect approach in this study indicates that there is little late-phase degradation; however, degradation in the early period of storage could not be ruled out. There was no indication of an increase in the proportion of the population below the lower LOD that was associated with longer freezer storage time (online Supplementary Fig. 1), although the study power for detecting such a change is limited.

It has been reported that the use of a sensitive assay for troponin I improves the early diagnosis of acute myocardial infarction and aids in risk stratification, regardless of the time of chest-pain onset \((27)\), whereas cTnT did not have higher diagnostic accuracy for acute myocardial infarction \((24)\). In general, the use of high-sensitivity troponin assays improves the information at the lower tail of the distribution of cardiac troponin concentrations, which is usually left-censored and assumed to be either 0 or set at the lower LOD. As shown by results of several studies \((5–12)\), such a lower range of values, when detectable, helps in prognostication. Whether a change in this low range can help in early detection and prevention of myocardial injury, such as occurs in drug-induced cardiomyopathy or ischemia not amounting to infarction, requires additional study. Given the favorable measurement properties of cTnT reported here and elsewhere \((14)\), the ability to detect very low concentrations of plasma cTnT should allow for a wider range of applications in clinical and research contexts.

Our results indicate that the laboratory reliability of the cTnT assay is high, as is the intraindividual (biological) reliability estimated at 6 weeks. These characteristics recommend the use of the cTnT assay in clinical and population-based research, such as for the estimation of associations and prediction of outcomes. The combination of low laboratory and biological, intraindividual variability makes the cTnT assays quite suitable for use in clinical settings, in which diagnostic classification and reliance on serial changes in repeat measurements are valuable.

**Author Contributions:** All authors confirmed they have contributed to the intellectual content of this paper and have met the following 3 requirements: (a) significant contributions to the conception and design, acquisition of data, or analysis and interpretation of data; (b) drafting or revising the article for intellectual content; and (c) final approval of the published article.

**Authors’ Disclosures or Potential Conflicts of Interest:** Upon manuscript submission, all authors completed the Disclosures of Potential Conflict of Interest form. Potential conflicts of interest:
Variability in cTnT

Employment or Leadership: V. Nambi, Southwest Lipid Association, American Society of Echocardiography, and Vascular Council.

Consultant or Advisory Role: C. Ballantyne, Roche.

Stock Ownership: None declared.

Honoria: None declared.

Research Funding: C. Ballantyne, Roche; V. Nambi, Abbott. C. Avery, grant K99-HL-098458 from the National Heart, Lung, and Blood Institute. The Atherosclerosis Risk in Communities Study is carried out as a collaborative study supported by National Heart, Lung, and Blood Institute contracts N01-HC-55015, N01-HC-55016, N01-HC-55018, N01-HC-55019, N01-HC-55020, N01-HC-55021, and N01-HC-55022.

Expert Testimony: None declared.

Other: Roche Diagnostics provided reagents and the loan of an instrument to conduct the high-sensitivity assays of cTnT.

Role of Sponsor: The funding organizations played no role in the design of study, choice of enrolled patients, review and interpretation of data, or preparation or approval of manuscript.

Acknowledgments: The authors thank the staff and participants of the ARIC study for their important contributions.

References