Intraindividual variability over time in plasma biomarkers of inflammation and effects of long-term storage

Sheetal Hardikar · Xiaoling Song · Mario Kratz · Garnet L. Anderson · Patricia L. Blount · Brian J. Reid · Thomas L. Vaughan · Emily White

Received: 22 January 2014 / Accepted: 7 May 2014 / Published online: 17 May 2014
© Springer International Publishing Switzerland 2014

Abstract

Purpose Systemic measures of chronic inflammation, often based on a single blood draw, are frequently used to study the associations between inflammation and chronic diseases such as cancer. However, more information is needed on the measurement error in these markers due to laboratory error, within-person variation over time, and long-term storage.

S. Hardikar · X. Song · M. Kratz · G. L. Anderson · P. L. Blount · B. J. Reid · T. L. Vaughan · E. White
Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, WA, USA

S. Hardikar · M. Kratz · T. L. Vaughan · E. White
Department of Epidemiology, University of Washington, Seattle, WA, USA

S. Hardikar (✉)
1100 Fairview Ave. N., M4-B402, PO Box 19024, Seattle, WA 98109-1024, USA
e-mail: shardika@fhcrc.org

M. Kratz · B. J. Reid
Department of Genome Sciences, University of Washington, Seattle, WA, USA

G. L. Anderson
Department of Biostatistics, University of Washington, Seattle, WA, USA

P. L. Blount · B. J. Reid
Human Biology Division, Fred Hutchinson Cancer Research Center, Seattle, WA, USA

P. L. Blount · B. J. Reid
Department of Medicine, University of Washington, Seattle, WA, USA

Methods We investigated the intraindividual variability of inflammation markers C-reactive protein (CRP), interleukin-6 (IL-6), and soluble tumor necrosis factor receptors I and II (sTNFRI and II) in a subsample of the Seattle Barrett's esophagus study cohort. Two fasting blood samples were collected between 1995 and 2009 from 360 participants on average 1.8 years apart. CRP, IL-6, and sTNF receptor levels were measured by immunonephelometry, ELISA, and multiplex assays, respectively. Intra- and inter-batch coefficients of variation (CV) were estimated using blinded pooled samples within each batch. Intraclass correlations (ICCs) were computed using random effects ANOVA.

Results Intra- and inter-batch CVs for the pooled plasma aliquots were low (2.4–8.9 %), suggesting little laboratory variability. Reliability over time was excellent for sTNF receptors ($ICC_{sTNF-RI} = 0.89$, $ICC_{sTNF-II} = 0.85$) and fair-to-good for CRP and IL-6 ($ICC_{CRP} = 0.55$, $ICC_{IL-6} = 0.57$). For samples stored for over 13 years, the ICCs for CRP and IL-6 were decreased but those for sTNF receptors were unaffected.

Conclusion sTNF receptor levels are more stable within person over time than CRP or IL-6. Long-term storage of samples appears to increase the variability of CRP and IL-6 measures, while the reliability of soluble TNF receptor measures was not affected by storage time.

Keywords ICC · CV · C-reactive protein · Interleukin-6 · Soluble tumor necrosis factor receptors

Introduction

C-reactive protein (CRP), an acute-phase reactant, and pro-inflammatory cytokines such as interleukin-6 (IL-6) and

tumor necrosis factor- α (TNF- α) have been implicated in the development as well as progression of a variety of human diseases, such as cancers [1–3], cardiovascular disease [4, 5], and diabetes [6]. Typically, studies investigating the association of inflammatory biomarkers on subsequent risk of chronic diseases utilize established cohorts with repositories of biological specimens collected prospectively, since biomarker levels will often be modified after disease diagnosis. Therefore, specimens will often be stored for 10 years or more before laboratory analysis. Additionally, in most studies, biological samples collected for analyses are limited in quantities, and only a single biospecimen per participant may be available for analysis. Although numerous epidemiological studies have used cytokine measurements from a single blood draw as their primary exposure, it is uncertain whether a single measure of an inflammatory biomarker is reflective of an individual's average level over time. The current evidence surrounding the reliability and intraindividual variability of inflammatory biomarkers is primarily restricted to studies involving CRP and IL-6 [4, 7–22]. The existing studies are limited by small sample sizes [7, 9, 10, 12–16, 18, 20, 22] and most evaluate the within-person variability over a relatively short period of few weeks to a year [8, 9, 13, 15, 18–20, 22]. Moreover, they typically involve healthy participants with considerably lower (and sometimes undetectable) blood levels of cytokines thus having to deal with missing data. None of these existing studies have evaluated the effect of long-term storage of blood samples on their reliability.

In this report, we analyzed the intraindividual variability of biomarkers of inflammation including CRP, IL-6, soluble tumor necrosis factor receptor I (sTNF-RI), and soluble tumor necrosis factor receptor II (sTNF-RII) in a subset of the Seattle Barrett's esophagus study (SBES), a large prospective cohort of Barrett's esophagus patients being followed to development of esophageal adenocarcinoma. In addition, we assessed the laboratory error and the effects of long-term storage of blood samples of these biomarkers. The biomarkers studied in this report are those that were simultaneously being evaluated for their association with esophageal adenocarcinoma, as part of separate analyses.

Materials and methods

Study population

The parent study, the Seattle Barrett's esophagus study (SBES), is an ongoing prospective cohort study designed to identify predictors of neoplastic progression in Barrett's esophagus that may assist in risk stratification, screening, and surveillance. It is comprised of persons with an

established histological diagnosis of Barrett's esophagus with no history of esophageal cancer. The details involving cohort recruitment, baseline protocol, and follow-up procedures have been described previously [23, 24]. Briefly, at the baseline visit, participants underwent an extensive personal interview, dietary and anthropometric assessments, and provided blood samples in addition to an endoscopy with biopsy as per a standard cancer surveillance protocol [25, 26]. At subsequent follow-up visits, shorter personal interviews were completed by the participants, and blood and biopsy samples were collected. Biomarkers of interest were measured at the first two time points with available blood samples (baseline and first follow-up visits for most). The median duration between the two visits was 1.8 years (IQR 1.3–2.0 years). The Institutional Review Boards at the University of Washington and Fred Hutchinson Cancer Research Center approved the study.

A total of 427 participants were enrolled in the SBES cohort between February 1, 1995 and September 30, 2009, 411 of whom had at least one follow-up visit completed at the time of present data analysis. Of the 411, 51 participants had either exhausted their stored blood or had blood specimens available only at a single time point. Hence, for the purposes of the current report, we used plasma measurements of inflammation markers from the remaining 360 participants. In addition, 31 individuals had one or both of their plasma CRP levels greater than 10 mg/l and hence were omitted from the CRP analysis due to an *a priori* hypothesis that levels over 10 mg/l are indicative of acute inflammation and should be excluded from the reliability study, as they would be from the parent study. Thus, we used plasma CRP measures from 329 participants while all 360 participants' biomarker measurements for IL-6, sTNF-RI and sTNF-RII were utilized for data analysis.

Laboratory methods for biomarker assessment

Fasting venous blood samples were obtained from each participant, processed within 2 h after collection and stored at -80°C until analysis. Plasma levels of CRP were measured at the Research Testing Services laboratories at the University of Washington Medical Center, Seattle, WA; while all other cytokines were measured at the PHS Biomarker lab at the Fred Hutchinson Cancer Research Center, Seattle, WA. CRP and IL-6 assays were conducted on never-thawed plasma samples, while once-thawed samples were used for the sTNF receptors measurement. Plasma CRP concentrations were determined with a high-sensitivity assay using immunonephelometry (Dade Behring Inc, Deerfield, IL, USA) as per the manufacturer's instructions. The minimum detectable limit for this assay was 0.2 mg/l for CRP; a mid-range value of 0.1 mg/l was

assigned to values below the detection limit (<1 % samples). Plasma levels of IL-6 were assayed using the Quantikine HS human IL-6 Elisa kit (R&D Systems Inc, HS600B) according to the manufacturer's instructions. The limit of quantification for this assay was 0.156 pg/ml, and none of the samples were below this level. For every participant, IL-6 samples were run in duplicate and averaged, with a median duplicate coefficient of variation (CV) of 2.7 %; samples with CVs greater than 12 % were re-run in duplicate, and the two new measurement values were used for analysis after averaging the duplicate values. sTNF-RI and sTNF-RII were measured using the MILLIPLEX MAP Human Soluble Cytokine Receptor Panel (Millipore, Billerica, MA, HSCR-32K) according to the manufacturer's instructions. The limit of quantification for the Millipore assay was 12.2 pg/ml, and all the samples were over this limit. Similar to IL-6, samples were run in duplicate with a median duplicate CV of 4.11 % for sTNF-RI and 3.37 % for sTNF-RII; samples with CVs greater than 12 % were re-run, and the re-run values were used for analyses. For every individual, the two samples from different time points were run in the same batch for laboratory analysis thus limiting our capability to estimate between-batch error in the ICCs.

In addition to the quality control procedures conducted by the laboratories, we included aliquots of a blinded pooled plasma sample in every batch of study samples analyzed (two aliquots per IL-6 and sTNF receptor batches, one aliquot with every CRP batch). Intra- and inter-batch CVs were calculated based on the measurements from these blinded aliquots.

Statistical analyses

Median levels of the various biomarkers studied were computed at each of the two time points along with their interquartile ranges. To assess the laboratory component of error, we estimated the intra- and inter-batch CVs for IL-6 and sTNF receptors and inter-batch CVs for CRP using the blinded pooled plasma specimens included with each batch of study samples. CVs were calculated as 100 times the ratio of the SD (σ) to the mean (μ) and expressed as a percentage; $CV = 100 \times \sigma/\mu$. ICCs and their 95 % confidence intervals (95 % CIs) were calculated to evaluate the total measurement error including variation within individuals over time, using random effects analysis of variance (ANOVA) models. The within (σ_w^2)- and between (σ_B^2)-subject variance from the ANOVA models was used to compute the ICCs using the formula $ICC = \sigma_B^2 / (\sigma_B^2 + \sigma_w^2)$ [27]. Following Rosner's definitions of reliability, ICCs ≥ 0.75 were considered to indicate excellent correlation, $0.4 \leq ICCs < 0.75$ were considered to be indicative of fair-to-good correlation, and ICCs < 0.4 were

Table 1 Characteristics of the study participants in the Seattle Barrett's esophagus study ($n = 360$)

Characteristic	%	Median (IQR)
Age (years)		61 (53–69)
Gender		
Male	82.2	
Female	17.8	
Body Mass Index (kg/m ²)		
≤ 25	12.5	
25.1 to ≤ 30	47.5	
30.1 to ≤ 35	30.6	
> 35	9.4	

IQR interquartile range

indicative of poor reproducibility [28]. Additionally, the ICC analyses were repeated after stratifying by duration between the two samples (≤ 2 years, > 2 years) as well as by the storage time of the earlier of the two samples (i.e., longest stored sample) used for ICC calculation (≤ 13 years, > 13 years) to evaluate whether long-term storage affected the repeatability of these inflammation markers. The cut-points for the stratified analyses were based on the mean values for the duration between the two samples and the mean storage time of the earlier of the two samples, respectively. Finally, we evaluated whether the actual median biomarker values of the longest stored of the two samples varied with storage time by conducting regression analyses and likelihood-ratio tests for each biomarker, and estimating the two-sided p value associated with the biomarker when treated as a grouped linear variable. The ICC computations were completed using the STATA version 12.0 statistical software (StataCorp LP, College Station, Texas, USA).

Results

The baseline-descriptive characteristics of our study sample are summarized in Table 1. The median age of the study participants was 61 years (IQR 53–69 years) and majority were males (82.2 %). A large proportion (40 %) of the study participants were obese, i.e., body mass index (BMI) greater than 30 kg/m², with 10 % having a BMI over 35 kg/m². The median waist–hip ratio for the study participants was 0.95 (IQR 0.91–0.99).

The intra- and inter-batch CVs estimating assay precisions based on the blinded pooled plasma samples are reported in Table 2. As we included only one blinded pooled sample with every CRP batch, we were unable to report the intra-batch variability in CRP measurements. CVs for intra-batch variation for the different assays

Table 2 Intra- and inter-batch coefficients of variations for blinded pooled plasma samples

Analyte	<i>n</i>	Intra-batch CV for pooled plasma samples (%)	Inter-batch CV for pooled plasma samples (%)
CRP	21	–	2.88
IL-6	22	4.13	4.35
sTNF-RI	22	5.87	8.93
sTNF-RII	22	2.39	6.09

CV coefficient of variation, CRP C-reactive protein, IL-6 interleukin-6, sTNF-RI surface tumor necrosis factor receptor I, sTNF-RII surface tumor necrosis factor receptor II

ranged from 2.4 to 5.9 % while those for inter-batch variation were between 2.9 and 8.9 %. Low CVs are indicative of less dispersion from the mean and therefore better reproducibility.

The medians and the interquartile ranges of the various cytokines at the two time points are reported in Table 3 along with the ICCs and their 95 % CI. The overall ICCs for sTNF-RI and sTNF-RII were excellent [$ICC_{sTNF-RI} = 0.89$, 95 % CI 0.87,0.91; $ICC_{sTNF-RII} = 0.85$, 95 % CI 0.82,0.88] while those for CRP and IL-6 were fair-to-good [$ICC_{CRP} = 0.55$, 95 % CI 0.47,0.62; $ICC_{IL-6} = 0.57$, 95 % CI 0.50,0.64]. The higher the ICC for an inflammation marker, the greater the proportion of total variation in that marker can be attributed to actual differences between individuals, rather than intraindividual variability or laboratory error. We further evaluated the reliability of these cytokines after stratification by the duration between two time points at which blood samples were collected as well as by the duration the first (earliest) of the two samples was stored (Table 3). The mean duration between the two blood samples used for biomarker measurements was 1.8 years (range 0.3–8.8 years). Hence, we used a cutoff of 2 years to stratify the duration between two samples. The ICC values for the two subsets stratified by duration between the two samples were similar for most of the biomarkers (The ICCs for sTNF-RII appeared to differ by duration between samples, but it is unlikely that reliability can be better over longer duration). The average storage time for the earlier of the two samples analyzed was 12.5 years (range 2.3–15.4 years) for CRP and 12.6 years (range 2.5–15.5 years) for IL-6 and sTNF receptors. We used a cutoff of 13 years to stratify storage time for the longest stored samples used for analyses. The average storage time for samples stored ≤ 13 years was 9.4 years (range 2.3–13 years) while that for samples stored >13 years was 14.2 years (range 13–15.5 years). The ICCs for CRP and IL-6 appeared to differ with storage time such that ICCs were lower when at least one of the two samples was stored for more than 13 years [$ICC_{CRP, \text{ storage } >13 \text{ years}} = 0.50$

(95 % CI 0.40, 0.60) and $ICC_{IL-6, \text{ storage } >13 \text{ years}} = 0.51$ (95 % CI 0.42, 0.61)] than for samples stored for <13 years [$ICC_{CRP, \text{ storage } \leq 13 \text{ years}} = 0.64$ (95 % CI 0.54, 0.75) and $ICC_{IL-6, \text{ storage } \leq 13 \text{ years}} = 0.67$ (95 % CI 0.57, 0.76)]. In contrast, the ICCs for sTNF receptors did not vary by the storage time for the longest stored sample.

We also evaluated whether the median biomarker levels for the longest stored sample varied by storage duration. The median values of all the biomarkers at four separate storage time categories (≤ 10 , >10 to ≤ 13 , >13 to ≤ 14 , and >14 years) are displayed in Table 4. As seen in Table 4, there was no apparent trend in biomarker levels with increasing storage time, suggesting that the median values of the biomarkers were stable across storage times (at -80°C), even though there was increased within-person variability with long-term storage as shown in Table 3.

Discussion

In this methodologic study, we evaluated the intraindividual variability of inflammation markers CRP, IL-6, sTNF-RI, and sTNF-RII over 1.8 years on average, in a subsample of the SBES cohort. Comparing the biomarker concentrations at the first two available plasma samples during follow-up (usually baseline and first follow-up), we observed that the ICCs for CRP and IL-6 were fair-to-good while those for sTNF-RI and TNF-RII were excellent, as per the criteria described by Rosner [28]. We also observed that the ICCs for sTNF-RI and sTNF-RII did not differ by storage time, while those for CRP and IL-6 were lower among those participants that had at least one of the two samples stored for over 13 years, as compared to study participants with both samples stored for less than 13 years. To our knowledge, this is the first study to examine the effects of long-term storage of blood samples on the intraindividual variability and reliability of inflammation markers. Overall, these data suggest that long-term storage of blood samples may impact the reliability of some biomarkers as shown by the lower ICCs for CRP and IL-6 with storage over 13 years, while other biomarkers may be largely unaffected by prolonged storage, e.g., sTNF receptors. Hence, in epidemiological studies of longer duration, it is essential to understand the variability of the biomarkers under consideration before drawing conclusions.

The overall ICCs reported in our study are fairly consistent with those reported by previous studies. Numerous studies have examined the intraindividual variability of CRP measurements, with reported ICCs between 0.58 and 0.76 over a wide range of time between blood draws (2 weeks to 12 years) [4, 7, 8, 10, 11, 14, 18–21]. The ICC for CRP in our study (0.55), although slightly lower than those reported in the past, is still indicative of good

Table 3 Temporal reliability of biomarkers of inflammation overall, by duration between the two samples and by storage time for the longest stored sample

Analyte	<i>n</i>	First sample		Second sample		ICC	95 % CI
		Median	IQR	Median	IQR		
CRP (mg/l) ^a							
Overall	329	1.8	0.9–3.5	1.8	0.9–3.6	0.55	0.47,0.62
Duration between two samples (years)							
≤2	237	1.9	1.0–3.5	1.8	0.9–3.5	0.52	0.43,0.62
>2	92	1.8	0.7–3.5	1.8	0.7–4.25	0.60	0.47,0.73
Storage time for the first sample ^b (years)							
≤13	118	1.75	0.9–3.2	1.7	0.8–2.9	0.64	0.54,0.75
>13	211	1.9	1.0–3.8	1.9	0.9–4.0	0.50	0.40,0.60
IL-6 (pg/ml)							
Overall	360	1.79	1.15–3.04	1.87	1.25–2.98	0.57	0.50,0.64
Duration between two samples (years)							
≤2	256	1.80	1.19–2.88	1.83	1.28–2.93	0.54	0.45,0.62
>2	104	1.65	1.12–3.16	1.99	1.17–3.42	0.62	0.50,0.74
Storage time for the first sample ^b (years)							
≤13	129	1.58	1.12–2.74	1.82	1.13–2.98	0.67	0.57,0.76
>13	231	1.85	1.20–3.17	1.88	1.29–2.94	0.51	0.42,0.61
sTNF-RI (pg/ml)							
Overall	360	1,370	1,125–1,650	1,366	1,116–1,672	0.89	0.87,0.91
Duration between two samples (years)							
≤2	256	1,383	1,132–1,660	1,382	1,132–1,672	0.90	0.88,0.92
>2	104	1,330	1,116–1,601	1,349	1,078–1,688	0.88	0.84,0.92
Storage time for the first sample ^b (years)							
≤13	129	1,343	1,119–1,609	1,344	1,108–1,608	0.90	0.87,0.94
>13	231	1,375	1,125–1,682	1,375	1,120–1,714	0.89	0.86,0.92
sTNF-RII (pg/ml)							
Overall	360	5,366	4,615–5,360	5,361	4,619–6,505	0.85	0.82,0.88
Duration between two samples (years)							
≤2	256	5,355	4,639–6,385	5,329	4,627–6,471	0.83	0.80,0.87
>2	104	5,429	4,549–6,373	5,455	4,588–6,585	0.88	0.84,0.93
Storage time for the first sample ^b (years)							
≤13	129	5,485	4,599–6,451	5,435	4,689–6,410	0.87	0.83,0.91
>13	231	5,350	4,627–6,384	5,307	4,609–6,547	0.84	0.80,0.88

CRP C-reactive protein, IL-6 interleukin-6, sTNF-RI surface tumor necrosis factor receptor I, sTNF-RII surface tumor necrosis factor receptor II, ICC intraclass correlation coefficient, IQR interquartile range, 95 % CI 95 % confidence interval

^a Thirty-one participants had one or both of their CRP values greater than 10 mg/l, and were omitted from analysis

^b Average storage time for samples stored ≤ 13 years was 9.4 years (range 2.3–13 years), for samples stored > 13 years is 14.2 years (range 13–15.5 years)

reproducibility. The ICC values for IL-6 reported in the literature are highly variable with studies reporting values as low as 0.37 [9] to as high as 0.92 [14], with most values in the range of 0.5–0.8 [10, 12, 15, 16, 18, 20, 22]. We reported an ICC value of 0.57 for IL-6 which is suggestive of good reproducibility. Only four studies have reported on the reliability of sTNF receptors [10, 12, 14, 20], and the sample sizes of three of them were much smaller ($n \leq 65$) as compared to our study. Our findings regarding ICCs of sTNF receptors I and II (0.89 and 0.85, respectively) are similar to those by Navarro et al. (ICCs of 0.92 and 0.90 for sTNF-RI and sTNF-RII, respectively) and suggest excellent reproducibility of sTNF receptors. Our estimates for sTNF receptors are higher than those reported by the other three studies [10, 12, 14]. However, they used a kit

different from the one we used for the estimation of sTNF receptor status, which may have contributed to the difference.

In most large-scale prospective epidemiological studies, enrollment into the study, data collection, and follow-up may span over several years; in some studies, like ours, it may even span over decades. Such studies rely on their biological banks where samples are stored for multiple years before analysis. It is, therefore, important to understand the effects of long-term storage of biological samples on their reliability, a question that is, to our knowledge, as yet unexplored with respect to markers of inflammation. We examined ICCs of the inflammation markers after stratifying on the duration of the longest stored sample (i.e., the first of the two samples analyzed for every

Table 4 Median Values of Biomarkers of Inflammation by Storage Time of the First Sample

Biomarker	Median and IQR by storage time of first sample (years)												<i>p</i> trend ^a
	≤10			>10 to ≤13			>13 to ≤14			>14			
	<i>n</i>	Med	IQR	<i>n</i>	Med	IQR	<i>n</i>	Med	IQR	<i>n</i>	Med	IQR	
CRP (mg/l)	47	1.11	0.7–2.9	71	2.1	1.1–3.4	77	1.8	0.9–3.6	134	2.05	1.0–3.8	0.58
IL-6 (pg/ml)	54	1.21	1.0–2.92	75	1.94	1.29–2.73	76	1.88	1.31–3.26	155	1.85	1.11–3.14	0.63
sTNF-RI (pg/ml)	54	1,322	1,116–1,503	75	1,439	1,131–1,672	73	1,467	1,196–1,803	158	1,354	1,114–1,565	0.09
sTNF-RII (pg/ml)	54	5,413	4,485–6,603	75	5,485	4,599–6,296	73	5,522	4,961–6,741	158	5,289	4,532–6,073	0.20

CRP C-reactive protein, IL-6 interleukin-6, sTNF-RI surface tumor necrosis factor receptor I, sTNF-RII surface tumor necrosis factor receptor II, *Med* median, *IQR* interquartile range

^a Test for trend was based on the likelihood-ratio test associated with the biomarker under consideration treated as a grouped linear variable, adjusted for confounding effects of age, gender and obesity as captured by the baseline waist–hip ratio

individual) and found that even after more than 13 years of storage, the ICCs suggested good-to-excellent reproducibility. There was no difference in the ICCs for the sTNF receptors between the two groups, suggesting that these markers were largely unaffected by the duration of storage. ICCs for CRP and IL-6 decreased somewhat for the group with over 13 years of sample storage (CRP 0.50, IL-6 0.51) as compared to the group with less than 13 years of storage (CRP 0.64, IL-6 0.67).

To assess whether there was any systematic degradation in actual serum biomarker levels over storage time, we examined the median biomarker levels for the earlier of the two samples collected. We did not find any evidence of a statistical trend in the biomarker levels with increasing storage time. Our results suggest that there is some concern regarding increased variability (i.e., lower repeatability as measured by ICCs) for CRP and IL-6 over prolonged storage; however, there does not seem to be any systematic increase or decrease in the actual biomarker levels over storage time.

In addition to estimating the ICCs, we also evaluated the amount of laboratory measurement error or batch-to-batch variability in inflammation marker measurements by including blinded pooled plasma samples with every batch of biomarker analyzed. Little variability was found in the inflammation marker measurements in pooled aliquots within as well as between the different batches. The intra- and inter-batch CVs for the blinded pooled plasma samples were low (range 2.4–8.9 %), indicative of good reproducibility of the assays with low batch-to-batch variation.

One strength of our study is that the ICCs computed in our study were able to capture most sources of error that affect epidemiologic studies using these biomarkers: the laboratory error, the field procedures error, e.g., due to variations in how specimens were collected, some amount of storage error as well as individual variation over time. The only error that we were unable to capture as a part of our ICCs is the batch-to-batch error, as samples from the same individual were tested in the same batch; however, that error is measured by the inter-batch CV.

Our study has a number of other strengths, the most important one being its large sample size. To date, this is the largest study assessing the intraindividual variability of the sTNF receptors. Additionally, it is the first study to evaluate the effects of long-term storage of blood specimens on the reliability of inflammation markers. Another strength of our study is the use of never-thawed samples for most biomarker measurements and minimal processing times (<2 h) after blood collection, both of which minimize the coefficient of variation and the laboratory error. This is particularly important for cytokines due to the concern of ex vivo expression of cytokines from leukocytes due to prolonged blood-processing times after collection [29]. Both strength as well as a limitation of our study was a large proportion of

obese individuals (40 %) in the cohort. The study participants might have had elevated levels of inflammation markers as compared to the general population due to their obesity [30]. On the one hand, this was beneficial as, unlike many of the previous studies, we did not encounter the problem of a large proportion of biomarkers being below detection limit. On the other hand, the high ICCs observed in this report may have been the result of a higher variation in the biomarker levels of the cohort, due to a higher proportion of obese individuals [31, 32]. Another limitation of our study was that all the participants in our cohort had Barrett's esophagus and hence our results may not be readily generalizable to other populations.

In addition, we were unable to assess the effects of shorter-term storage on markers of inflammation. Some reports have suggested a possible degradation of samples within 2–5 years of storage [33–35], even with storage temperatures of -70°C or below, for various cytokines. However, in our study, only 15 % of samples were stored for less than 10 years, limiting our ability to assess either the within-person variability of samples stored for less than 10 years or any trend in median values due to degradation of samples over the first 10 years. Further studies are needed to evaluate the stability of measures of inflammation over the first decade of storage, ideally comparing measures on the same individuals from fresh blood and periodically over the first 10 years of storage.

In conclusion, we observed that biomarkers of CRP and IL-6 have moderate variability over time within person, and appear to be affected by prolonged storage of blood samples, while biomarkers of sTNF receptors vary little within person over time and are largely unaffected by long-term blood storage. Hence, in epidemiological studies of prolonged duration involving CRP or IL-6, methods to reduce this variability should be considered such as averaging multiple measures over time.

Acknowledgments We thank Tricia Christopherson for project management; Terri Watson for database management; and Christine Karlsen for coordination of patient care.

Financial support This work was supported by United States National Institutes of Health (Grants P01CA091955, K05CA124911 and R25CA094880). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Balkwill F, Mantovani A (2001) Inflammation and cancer: back to Virchow? *Lancet* 357:539–545
2. Il'yasova D, Colbert LH, Harris TB, Newman AB, Bauer DC, Satterfield S, Kritchevsky SB (2005) Circulating levels of inflammatory markers and cancer risk in the health aging and body composition cohort. *Cancer Epidemiol Biomarkers Prev* 14:2413–2418
3. Grote VA, Kaaks R, Nieters A, Tjonneland A, Halkjaer J, Overvad K, Skjelbo Nielsen MR, Boutron-Ruault MC, Clavel-Chapelon F, Racine A, Teucher B, Becker S, Pischon T, Boeing H, Trichopoulou A, Cassapa C, Stratigakou V, Palli D, Krogh V, Tumino R, Vineis P, Panico S, Rodriguez L, Duell EJ, Sanchez MJ, Dorronsoro M, Navarro C, Gurtea AB, Siersema PD, Peeters PH, Ye W, Sund M, Lindkvist B, Johansen D, Khaw KT, Wareham N, Allen NE, Travis RC, Fedirko V, Jenab M, Michaud DS, Chuang SC, Romaguera D, Bueno-de-Mesquita HB, Rohrmann S (2012) Inflammation marker and risk of pancreatic cancer: a nested case-control study within the EPIC cohort. *Br J Cancer* 106(11):1866–1874. doi:10.1038/bjc.2012.172
4. Danesh J, Wheeler JG, Hirschfield GM, Eda S, Eiriksdottir G, Rumley A, Lowe GD, Pepys MB, Gudnason V (2004) C-reactive protein and other circulating markers of inflammation in the prediction of coronary heart disease. *N Engl J Med* 350:1387–1397
5. Emerging Risk Factors Collaboration, Kaptoge S, Di Angelantonio E, Pennells L, Wood AM, White IR, Gao P, Walker M, Thompson A, Sarwar N, Caslake M, Butterworth AS, Amouyel P, Assmann G, Bakker SJ, Barr EL, Barrett-Connor E, Benjamin EJ, Bjorkelund C, Brenner H, Brunner E, Clarke R, Cooper JA, Cremer P, Cushman M, Dagenais GR, D'Agostino RB Sr, Dankner R, Davey-Smith G, Deeg D, Dekker JM, Engstrom G, Folsom AR, Fowkes FG, Gallacher J, Gaziano JM, Giampaoli S, Gillum RF, Hofman A, Howard BV, Ingelsson E, Iso H, Jorgensen T, Kiechl S, Kitamura A, Kiyohara Y, Koenig W, Kromhout D, Kuller LH, Lawlor DA, Meade TW, Nissinen A, Nordestgaard BG, Onat A, Panagiotakos DB, Psaty BM, Rodriguez B, Rosengren A, Salomaa V, Kauhanen J, Salonen JT, Shaffer JA, Shea S, Ford I, Stehouwer CD, Strandberg TE, Tipping RW, Tosetto A, Wassertheil-Smoller S, Wennberg P, Westendorp RG, Whincup PH, Wilhelmsen L, Woodward M, Lowe GD, Wareham NJ, Khaw KT, Sattar N, Packard CJ, Gudnason V, Ridker PM, Pepys MB, Thompson SG, Danesh J (2012) C-reactive protein, fibrinogen, and cardiovascular disease prediction. *The New England journal of medicine* 367(14):1310–1320. doi:10.1056/NEJMoa1107477
6. Pradhan AD, Manson JE, Rifai N, Buring JE, Ridker PM (2001) C-reactive protein, interleukin 6, and risk of developing type 2 diabetes mellitus. *JAMA* 286(3):327–334
7. Al-Delaimy WK, Jansen EH, Peeters PH, van der Laan JD, van Noord PA, Boshuizen HC, van der Schouw YT, Jenab M, Ferrari P, Bueno-de-Mesquita HB (2006) Reliability of biomarkers of iron status, blood lipids, oxidative stress, vitamin D, C-reactive protein and fructosamine in two Dutch cohorts. *Biomarkers* 11:370–382
8. Block G, Dietrich M, Norkus E, Jensen C, Benowitz NL, Morrow JD, Hudes M, Packer L (2006) Intraindividual variability of plasma antioxidants, markers of oxidative stress, C-reactive protein, cotinine, and other biomarkers. *Epidemiology* 17:404–412
9. Cava F, Gonzalez C, Pascual MJ, Navajo JA, Gonzalez-Buitrago JM (2000) Biological variation of interleukin 6 (IL-6) and soluble interleukin 2 receptor (sIL2R) in serum of healthy individuals. *Cytokine* 12(9):1423–1425. doi:10.1006/cyto.2000.0714
10. Clendenen TV, Arslan AA, Lokshin AE, Idahl A, Hallmans G, Koenig KL, Marrangoni AM, Nolen BM, Ohlson N, Zeleniuch-Jacquotte A, Lundin E (2010) Temporal reliability of cytokines and growth factors in EDTA plasma. *BMC Res Notes* 3:302. doi:10.1186/1756-0500-3-302

11. DeGoma EM, French B, Dunbar RL, Allison MA, Mohler ER 3rd, Budoff MJ (2012) Intraindividual variability of C-reactive protein: the Multi-Ethnic Study of Atherosclerosis. *Atherosclerosis* 224(1):274–279. doi:[10.1016/j.atherosclerosis.2012.07.017](https://doi.org/10.1016/j.atherosclerosis.2012.07.017)
12. Epstein MM, Breen EC, Magpantay L, Detels R, Lepone L, Penugonda S, Bream JH, Jacobson LP, Martinez-Maza O, Birman BM (2013) Temporal stability of serum concentrations of cytokines and soluble receptors measured across two years in low-risk HIV-seronegative men. *Cancer Epidemiol Biomarkers Prev* 22(11):2009–2015. doi:[10.1158/1055-9965.EPI-13-0379](https://doi.org/10.1158/1055-9965.EPI-13-0379)
13. Gonzalez C, Cava F, Ayllon A, Guevara P, Navajo JA, Gonzalez-Buitrago JM (2001) Biological variation of interleukin-1beta, interleukin-8 and tumor necrosis factor-alpha in serum of healthy individuals. *Clinical Chem Lab Med* 39(9):836–841. doi:[10.1515/CCLM.2001.139](https://doi.org/10.1515/CCLM.2001.139)
14. Gu Y, Zeleniuch-Jacquotte A, Linkov F, Koenig KL, Liu M, Velikokhatnaya L, Shore RE, Marrangoni A, Toniolo P, Lokshin AE, Arslan AA (2009) Reproducibility of serum cytokines and growth factors. *Cytokine* 45:44–49
15. Ho GY, Xue XN, Burk RD, Kaplan RC, Cornell E, Cushman M (2005) Variability of serum levels of tumor necrosis factor-alpha, interleukin 6, and soluble interleukin 6 receptor over 2 years in young women. *Cytokine* 30:1–6
16. Hofmann JN, Yu K, Bagni RK, Lan Q, Rothman N, Purdue MP (2011) Intra-individual variability over time in serum cytokine levels among participants in the prostate, lung, colorectal, and ovarian cancer screening trial. *Cytokine* 56(2):145–148. doi:[10.1016/j.cyto.2011.06.012](https://doi.org/10.1016/j.cyto.2011.06.012)
17. Kayaba K, Ishikawa S, Gotoh T, Nago N, Kajii E, Nakamura Y, Kario K (2000) Five-year intra-individual variability in C-reactive protein levels in a Japanese population-based study: the Jichi Medical School Cohort Study at Yamato, 1993–1998. *Jpn Circ J* 64(4):303–308
18. Lee SA, Kallianpur A, Xiang YB, Wen W, Cai Q, Liu D, Fazio S, Linton MF, Zheng W, Shu XO (2007) Intra-individual variation of plasma adipokine levels and utility of single measurement of these biomarkers in population-based studies. *Cancer Epidemiol Biomarkers Prev* 16:2464–2470
19. Nasermoaddeli A, Sekine M, Kagamimori S (2006) Intra-individual variability of high-sensitivity C-reactive protein: age-related variations over time in Japanese subjects. *Circ J* 70:559–563
20. Navarro SL, Brasky TM, Schwarz Y, Song X, Wang CY, Kristal AR, Kratz M, White E, Lampe JW (2012) Reliability of serum biomarkers of inflammation from repeated measures in healthy individuals. *Cancer Epidemiol Biomarkers Prev* 21(7):1167–1170. doi:[10.1158/1055-9965.EPI-12-0110](https://doi.org/10.1158/1055-9965.EPI-12-0110)
21. Ockene IS, Matthews CE, Rifai N, Ridker PM, Reed G, Stanek E (2001) Variability and classification accuracy of serial high-sensitivity C-reactive protein measurements in healthy adults. *Clin Chem* 47:444–450
22. Rao KM, Pieper CS, Currie MS, Cohen HJ (1994) Variability of plasma IL-6 and crosslinked fibrin dimers over time in community dwelling elderly subjects. *Am J Clin Pathol* 102:802–805
23. Hardikar S, Onstad L, Blount PL, Odze RD, Reid BJ, Vaughan TL (2013) The role of tobacco, alcohol, and obesity in neoplastic progression to esophageal adenocarcinoma: a prospective study of Barrett's esophagus. *PLoS ONE* 8(1):e52192. doi:[10.1371/journal.pone.0052192](https://doi.org/10.1371/journal.pone.0052192)
24. Vaughan TL, Kristal AR, Blount PL, Levine DS, Galipeau PC, Prevo LJ, Sanchez CA, Rabinovitch PS, Reid BJ (2002) Non-steroidal anti-inflammatory drug use, body mass index, and anthropometry in relation to genetic and flow cytometric abnormalities in Barrett's esophagus. *Cancer Epidemiol Biomarkers Prev* 11(8):745–752
25. Ertan A, Younes M (2000) Barrett's esophagus. *Dig Dis Sci* 45:1670–1673
26. Oh DS, Demeester SR (2010) Pathophysiology and treatment of Barrett's esophagus. *World J Gastroenterol* 16:3762–3772
27. Shrout PE, Fleiss JL (1979) Intraclass correlations: uses in assessing rater reliability. *Psychol Bull* 86(2):420–428
28. Rosner B (2000) Fundamentals of biostatistics, vol 5th. Vol book, whole. Duxbury, Pacific Grove
29. Thavasu PW, Longhurst S, Joel SP, Slevin ML, Balkwill FR (1992) Measuring cytokine levels in blood. Importance of anti-coagulants, processing, and storage conditions. *J Immunol Methods* 153(1–2):115–124
30. Yudkin JS, Stehouwer CD, Emeis JJ, Coppack SW (1999) C-Reactive protein in healthy subjects: associations with obesity, insulin resistance, and endothelial dysfunction: a potential role for cytokines originating from adipose tissue? *Arterioscler Thromb Vasc Biol* 19:972–978
31. White E (2011) Measurement error in biomarkers: sources, assessment, and impact on studies. *IARC Sci Publ* 163:143–161
32. White E, Armstrong BK, Saracci R (2008) Principles of exposure measurement in epidemiology, 2nd edn, vol book, whole. Oxford University Press, USA
33. de Jager W, Bourcier K, Rijkers GT, Prakken BJ, Seyfert-Margolis V (2009) Prerequisites for cytokine measurements in clinical trials with multiplex immunoassays. *BMC Immunol* 10:52
34. Tworoger SS, Hankinson SE (2006) Use of biomarkers in epidemiologic studies: minimizing the influence of measurement error in the study design and analysis. *Cancer Causes Control* 17(7):889–899. doi:[10.1007/s10552-006-0035-5](https://doi.org/10.1007/s10552-006-0035-5)
35. House RV (2001) Cytokine measurement techniques for assessing hypersensitivity. *Toxicology* 158(1–2):51–58