

Influence of assay variability, location and storage time on cytokine measurements in population-based biobanking

Cytokine measurements in biobanking

Robert P. van Waateringe MD ¹

Anneke C. Muller Kobold ²

Jan Koerts ²

Jana V. van Vliet-Ostaptchouk ¹

Gabriele Anton ³

Annette Peters ³

Gerlinde Trischler ⁴

Kirsti Kvaloy MSc, PhD ⁵

Marit Naess ⁵

Vibeke Videm ⁶

Melanie M. van der Klauw ¹

Kristian Hveem ⁵

Melanie Waldenberger ³

Wolfgang Koenig MD, PhD ⁴

Bruce H.R. Wolffenbuttel MD, PhD ¹

Affiliations

1. Dept. Of Endocrinology, University of Groningen, University Medical Center Groningen, P.O. Box 30001, 9700 RB Groningen, The Netherlands.
2. Dept of Laboratory Medicine, University of Groningen, University Medical Center Groningen, P.O. Box 30001, 9700 RB Groningen, The Netherlands
3. Research Unit of Molecular Epidemiology and Institute of Epidemiology II, Helmholtz Zentrum München, Munich, Germany.
4. Department of Internal Medicine II – Cardiology, University of Ulm Medical Centre, D-89081 Ulm, Germany.
5. Department of Public Health and General Practice, Faculty of Medicine, NTNU-Norwegian University of Science and Technology, Trondheim, Norway.
6. Department of Laboratory Medicine, Children's and Women's Health, Faculty of Medicine, NTNU-Norwegian University of Science and Technology, and Department of Immunology and Transfusion Medicine, St. Olavs Hospital, Trondheim, Norway.

Keywords: cytokines, ELISA, variability, storage

Address for correspondence:

Bruce HR Wolffenbuttel, MD PhD

Department of Endocrinology

University of Groningen, University Medical Center Groningen

HPC AA31, PO Box 30001

9700 RB Groningen, The Netherlands

Phone +31-50-3613962, Fax +31-50-3619392

Editor-in-Chief: Jim Vaught

Robert P. van Waateringe MD, Dept of Endocrinology, University of Groningen, University

Medical Center Groningen, P.O. Box 30001, 9700 RB Groningen, The Netherlands. E-mail:

r.p.van.waateringe@umcg.nl

Anneke C. Muller Kobold, PhD, Dept of Laboratory Medicine, University of Groningen,

University Medical Center Groningen, P.O. Box 30001, 9700 RB Groningen, The

Netherlands. E-mail: a.c.muller@umcg.nl

Jan Koerts, Dept of Laboratory Medicine, University of Groningen, University Medical

Center Groningen, P.O. Box 30001, 9700 RB Groningen, The Netherlands. E-mail:

j.koerts01@umcg.nl

J.V. van Vliet-Ostapchouk, PhD, Dept. of Endocrinology, University of Groningen,

University Medical Center Groningen, HPC AA31, P.O. Box 30001, 9700 RB Groningen,

The Netherlands, E-mail: j.v.van.vliet@umcg.nl

Gabriele Anton, PhD, Research Unit of Molecular Epidemiology and Institute of

Epidemiology II, Helmholtz Zentrum München, Munich, Germany, E-mail:

gabriele.anton@helmholtz-muenchen.de

Annette Peters, PhD, Research Unit of Molecular Epidemiology and Institute of

Epidemiology II, Helmholtz Zentrum München, Munich, Germany, E-mail:

peters@helmholtz-muenchen.de

Gerlinde Trischler, Department of Internal Medicine II – Cardiology, University of Ulm Medical Centre, D-89081 Ulm, Germany, E-mail: Gerlinde.Trischler@uniklinik-ulm.de

Kirsti Kvaloy MSc, PhD, Dept of Public Health and General Practice, NTNU-Norwegian Norwegian University of Science and Technology, Trondheim, Norway. E-mail: kirsti.kvaloy@ntnu.no

Marit Naess, HUNT Research Centre, Department of Public Health and General Practice, Faculty of Medicine, NTNU-Norwegian University of Science and Technology, Trondheim, Norway. E-mail: marit.nass@ntnu.no

Vibeke Videm, Department of Laboratory Medicine, Children's and Women's Health, Faculty of Medicine, NTNU-Norwegian University of Science and Technology, and Department of Immunology and Transfusion Medicine, St. Olavs Hospital, Trondheim, Norway. E-mail: vibeke.videm@ntnu.no

Melanie M. van der Klauw, MD, PhD, Dept of Endocrinology, University of Groningen, University Medical Center Groningen, P.O. Box 30001, 9700 RB Groningen, The Netherlands. E-mail: m.m.van.der.klauw@umcg.nl

Kristian Hveem, MD PhD, Dept of Public Health and General Practice, NTNU-Norwegian Norwegian University of Science and Technology, Trondheim, Norway. E-mail: kristian.hveem@ntnu.no

Melanie Waldenberger, PhD, Research Unit of Molecular Epidemiology and Institute of Epidemiology II, Helmholtz Zentrum München, Munich, Germany. E-mail: waldenberger@helmholtz-muenchen.de

Wolfgang Koenig MD, PhD, Department of Internal Medicine II – Cardiology, University of Ulm Medical Centre, D-89081 Ulm, Germany, Email: wolfgang.koenig@uniklinik-ulm.de

Bruce H.R. Wolffenbuttel MD, PhD, Dept of Endocrinology, University of Groningen, University Medical Center Groningen, P.O. Box 30001, 9700 RB Groningen, The Netherlands. E-mail: bwo@umcg.nl

Abstract

Introduction In biobanking, harmonization and standardization of high-throughput assays to measure biomarkers, such as inflammatory biomarkers, is crucial. We assessed the reproducibility and standardization of inflammatory biomarker measurements of high sensitive CRP (hsCRP), high sensitive Interleukin-6 (hsIL6) and high sensitive Tumor Necrosis Factor Alpha (hsTNF α) with different assays, and analyzed the reproducibility of a specific measurement between laboratories as well as the influence of sample storage time.

Methods In total, 240 fasting plasma samples were obtained from the LifeLines biobank. Samples had been stored at -80°C for less than 2, or over 4 years. hsCRP was measured by nephelometry and ELISA, and hsIL6 and hsTNF α were measured by two different ELISA assays in three separate laboratories. For confirmation, a similar analysis was performed in samples obtained from 80 obese subjects participating in a weight maintenance program. Passing Bablok regression analysis was used to compare results.

Results The measurements suggested good stability of samples stored at -80°C. hsCRP measured at the day of blood draw was similar to levels measured after > 4 years of storage. There were small inter-laboratory differences with the R&D ELISA's for hsIL6 and hsTNF α . There was a linear relationship between the Bender and R&D ELISA for hsIL6, with statistically significant higher levels measured with the R&D assay. Over 90% of samples measured with the IBL hsTNF α ELISA were below the detection limit of 0.13 ng/l, rendering this assay unsuitable for large scale analysis. Similar results were obtained in the confirmation study.

Conclusion Plasma samples stored over a period of 2-4 years at -80 °C showed good stability to perform measurements of inflammatory biomarkers. Even when the same ELISA method was used, there were small variations in results reported by different laboratories. Cytokine assays should be rigorously tested before large sample sets are measured.

Introduction

Population-based biobanks are involved in processing and storage of biospecimens for future studies and additionally accumulate epidemiological data, sometimes over decades. . Several new biobanking initiatives have been launched building on these strengths. The Biobank Standardization and Harmonization for Research Excellence in the European Union (BioSHaRE-EU) project is based on international collaborative projects between European and Canadian institutes and European cohort studies. The project has developed and applied several methods and tools for harmonization and standardization in European biobanks and major biomedical studies ¹. If an efficient organization of these existing resources is implemented, rapid progress in the biomedical field can be achieved. This has been impressively demonstrated by the success of genome-wide association (GWAS) studies and the combined analysis of these data in large meta-analyses ².

For other new techniques like metabolomics ³ and epigenomics ⁴, the availability of samples of high quality is crucial. Standardization of sample collection, pre-analytics, harmonization and standardization of high-throughput assays to measure biomarkers, such as inflammatory biomarkers is crucial ^{5,6}. In general, biomarkers are defined as objectively measurable indicators for biological or pathobiological processes or pharmacological responses towards medical treatment ⁷. Biomarkers may serve as surrogate endpoints, which correlate with clinical endpoints, indicate disease progression and regression under therapy, and may allow outcome prediction. For optimal collaboration between cohort studies or biobanks, harmonization and standardization of analytical procedures of biomarker measurements is essential. Analytical results may be affected by pre-analytical conditions and analytical variability ^{6,8}. For example, different types of samples may be available for analyses, but may yield different results upon measurement (i.e. serum versus plasma).

Furthermore, measurements may need to be performed on stored samples, which may have been influenced by different temperatures and different duration of storage⁹⁻¹¹. Finally, results may be based on the use of different assays, techniques or equipment.

Studies investigating the variability of sample processing, different assays, the use of different sample types and the reproducibility of archived samples are scarce, particularly with regards to measurement of inflammatory biomarkers like cytokines¹². Aziz et al. examined pre-analytical variables on high sensitive C-reactive protein (hsCRP) and found that hsCRP levels in serum were not significantly different from plasma samples¹³. In addition, storage of samples at -70 °C for 3 weeks had no effect on hsCRP concentrations. However, some contradictory data regarding long-term storage of hsCRP exists^{14, 15}. Only a few studies have evaluated and compared different assays to measure other inflammatory markers. López-Campos et al. compared enzyme-linked immunosorbent assay (ELISA) with immunonephelometry for the measurement of hsCRP in patients with stable COPD¹⁶. Although the serum hsCRP concentrations measured by ELISA and nephelometry correlated well, concentrations measured using ELISA tended to be lower.

The present study aimed to assess the reproducibility and standardization of several inflammatory biomarker measurements, including high sensitive CRP (hsCRP), high sensitive Interleukin-6 (hsIL6) and high sensitive Tumor Necrosis Factor Alpha (hsTNF α), taking into account the influence of assay variability, the reproducibility of a specific measurement between laboratories as well as the influence of sample storage time and conditions.

Materials and Methods

Participants and sample collections

Subjects included were participants from the LifeLines Cohort Study¹⁷. Lifelines is a multi-disciplinary prospective population-based cohort study examining in a unique three-generation design the health and health-related behaviours of persons living in the North of The Netherlands. It started in 2007, and employs a broad range of investigative procedures in assessing the biomedical, socio-demographic, behavioural, physical and psychological factors which contribute to the health and disease of the general population, with a special focus on multi-morbidity and complex genetics. The methodology has been described previously¹⁸. All participants were between 18 and 90 years old at the time of enrolment. They provided written informed consent before participating in the study. The study protocol was approved by the medical ethical review committee of the University Medical Center Groningen.

Blood samples were drawn by venipuncture in the fasting state, between 8 and 10 a.m. After blood withdrawal, tubes were transported (2 hours) to the LifeLines laboratory either at room temperature, or at 4 °C, depending on the required procedure for the specific tube. Subsequently, samples for direct analysis were handed to technicians of the department of Laboratory Medicine at the University Medical Center of Groningen. All other samples were centrifuged and aliquoted into smaller tubes and stored in freezers at -80 °C until the measurements were performed. For reliable and secure storage of all samples, LifeLines has used the Micronic 0.9 ml screw-cap tubes.

For the current study, a 900 µL plasma sample from each participant was to be selected by the LifeLines Scientific Bureau according to the study protocol. Half the samples had been stored for less than 2 years (n=120) and the other half for more than 4 years (n=120). Samples were thawed once, aliquoted into 0.9 ml shipment tubes, and stored again at -80 °C. Aliquots representing all the samples were shipped on dry ice to the laboratories of Trondheim in Norway, Ulm in Germany and Groningen in The Netherlands, where they remained in storage at -80 °C until use for the specific measurements. The inflammatory

biomarkers were measured in all 240 participants using assays available at the given locations (Tables 1 and 2).

For confirmation of the initial results, additional analysis of IL6 and TNF α were performed by one dedicated analyst in a set of 80 samples obtained from obese individuals who participated in a weight-reduction program in The Netherlands (the LOWER study, www.clinicaltrials.gov, NCT00862953).

Laboratory measurements

Details of the assays including detection ranges are summarized in Table 2. HsCRP was measured in 120 fresh serum samples and re-measured in 4-year old samples (n=120) using latex enhanced immunonephelometry (Siemens Healthcare Diagnostics). Standardization is based on protein reference ERM DA 470 (CRM 470). Results of hsCRP measurements were compared using either the R&D ELISA (location A) or nephelometry (location C). At each location, measurements were done according to the instructions of the manufacturer. The minimum detection limit of the assay was 0.010 ng/mL.

Both at location A and location C, 240 stored serum samples were measured using the hsIL6 ELISA R&D, whereas location B used the hsIL6 ELISA from Bender MedSystems. Both kits were used according to the instructions of the manufacturer. The R&D ELISA used 100 μ l of sample for analysis. The Bender MedSystems ELISA kit was specifically chosen as it used a smaller amount of sample (50 μ l) for analysis.

hsTNF α was measured with the R&D ELISA kit (HSTA00D) at location A and C and with the IBL ELISA at location B. Sample volume needed was 200 μ l. wa

The IBL ELISA was specifically chosen as it uses a lower amount of sample (50 μ l) compared to the R&D ELISA.

Statistical analyses

Statistical analyses were performed using SPSS version 22. Passing-Bablok regression analysis and Bland-Altman plots were created in order to evaluate the degree of agreement between (1) different types of assays and (2) between fresh and stored samples for the measurement of the specific inflammatory biomarkers.

Results

Of the 240 participants, 137 (57%) were males. Mean (\pm SD) age was 60 ± 11 years, and mean body mass index was 27.9 ± 3.8 kg/m². The samples were obtained from equal numbers of healthy individuals, individuals with type 2 diabetes mellitus, and individuals who had previously suffered from a cardiovascular event.

Measurement of hsCRP with nephelometry in 4-year old stored samples yielded identical results as was obtained when these samples (n=116) were measured at the day of blood drawing (Figure 1 and Supplementary Figure 1). For comparison of nephelometry and ELISA (R&D), 200 of the 240 samples were available for analysis; 3 samples were below the detection limit of the ELISA, while 37 samples were above the detection limit. These samples were not re-measured in dilution due to insufficient sample material. Measurement of hsCRP by ELISA yielded different results compared to the Gold Standard method nephelometry (Figures 2 and Supplementary Figure 2). Specifically, hsCRP concentrations measured by ELISA tended to be lower than concentrations measured with nephelometry.

For hsIL6, there were significant differences in the results, as well as differences regarding the detection limit. The R&D ELISA yielded similar results in both laboratories (location A and C) except for some outliers (Figure 3). There was a linear relationship between the Bender MedSystems (location B) and R&D ELISA (location C), but with

statistically significant higher levels measured by the R&D assay (Figure 4 and Supplementary Figure 3). This was similar for sample stored up to 2 years, vs. samples stored for more than 4 years. It was observed that 38 samples were below the limit of detection of the R&D ELISA used at location A whereas at location C, using the same ELISA, we found 2 samples with concentrations above 5 ng/l. With the Bender MedSystems ELISA used at location B, two samples could not be measured due to insufficient sample material, while three samples were below the detection limit of the assay.

The results for the measurement of hsTNF α samples are shown in Figure 5. At location C, where the R&D ELISA assay was used, one serum sample concentration was higher than 16 ng/l, while location A -using the same assay- reported that 7 samples were below the detection limit. One serum sample yielded results above the detection limit. Unfortunately, 177 samples were lost for evaluation due to a technical error (Location A). Hence, only 55 samples could be compared with the results in location C (Figure 5). These figures show that there was reasonable agreement between the two locations (both using the R&D ELISA) although there were a few outliers. To assess any influence of storage time on measurement of TNF α samples, we were only able to examine samples which had been stored for more than 4 years. Location B used the IBL ELISA and reported that the majority of measurements were below the detection limit of 0.13 ng/l; only 18 samples yielded feasible and measurable results. Therefore, comparison of the IBL and R&D ELISA (Figure 6) demonstrated poor agreement between the assays.

For confirmation and replication, we performed an additional evaluation of assays for the measurement of hsTNF α and hsIL6 in stored samples obtained from a population of 80 obese individuals (50% males) participating in a weight-reduction program.

Their mean age was 52 ± 12 years, BMI 38.0 ± 6.2 kg/m². Their samples had been stored at -80 °C for an average period of 4 - 7 years. An excellent correlation between the

hsIL6 results using either R&D ELISA and Bender MedSystems ELISA was observed (Figure 7), which was significantly better than the results depicted in Figure 4. Regarding the hsTNF α measurements, the IBL ELISA yielded very poor results compared to the results from the R&D ELISA: 66 out of 80 samples yielded results below the detection limit as set by the manufacturer (Figure 8) and thereby confirmed the former findings with this assay.

Discussion

In the present study we compared different assays for measurement of the inflammatory biomarkers hsCRP, hsIL6 and hsTNF α , and assessed the effect of storage time on the reproducibility of the measurement of hsCRP. Our data showed that short- to medium-term storage (less than 2 years, more than 4 years) did not influence the plasma levels of hsCRP and hsIL6 measured with nephelometry and with ELISA, respectively. However, there were small differences between two hsIL6 ELISA methods. With the hsTNF α ELISA of IBL, the majority of samples remained below the detection limit of the assay, whereas the R&D ELISA yielded stable and reproducible results.

hsCRP is frequently measured in clinical and epidemiological studies. We observed only a moderate agreement between results of nephelometry and ELISA. Similar findings have been reported in a study by López-Campos et al. ¹⁶. They reported higher serum hsCRP concentrations when measured by nephelometry compared to measurement by ELISA. We observed similar results by the ELISA method clearly showing an upper limit of detection. For hsCRP concentrations above 25 mg/l, a sample needs to be diluted and re-analyzed, which is not the case with nephelometry. Nevertheless, our studies in stored samples indicated excellent sample stability after > 4 years of storage at -80 °C. Other studies regarding long-term storage of hsCRP have reported contradictory results ¹⁴.

Doumatey et al showed that serum hsCRP concentrations remained stable with storage for up to 11 years at $-80\text{ }^{\circ}\text{C}$ ¹⁴. This was in contrast to a study from Japan reporting that hsCRP levels increased in samples stored at $-80\text{ }^{\circ}\text{C}$ for 13.8 years¹⁵.

Measurement of hsIL6 with the same R&D ELISA method in different laboratories showed good agreement. However, the fact that 38 samples gave results below the detection limit in one laboratory should be taken into account. In addition, no influence of storage time on the measurement of IL6 samples was observed when samples were stored between 2 and 4 years at $-80\text{ }^{\circ}\text{C}$, although it should be noted that we have no data on fresh measurements or samples stored for less than 2 years. There is remarkable little information on studies investigating different assays or the influence of storage time on measurement of IL6. A recent study by Hardikar et al. showed moderate stability of IL6 in samples that were stored at $-80\text{ }^{\circ}\text{C}$ for less than 13 years¹⁹.

A previous study examining the influence of short-term storage of several biomarkers showed excellent stability for TNF α in samples stored at $-80\text{ }^{\circ}\text{C}$ for 90 days²⁰. Although this is encouraging, the relevance of these data for biobanking, where samples have been or will be stored for many years, is limited. As was the case for hsIL6, between-laboratory variation of the R&D ELISA was very small. When biomarker measurements are performed in stored serum or plasma samples of biobanks, the amount of sample needed for a specific measurement is of great importance. This was the main reason why we chose the Bender MedSystems hsIL6 assay and the IBL hsTNF α assay for our comparison studies, as both assays required only 50 μl of sample. In contrast, the R&D assays required 100 μl and 200 μl , respectively. A head-to-head comparison between the Bender MedSystems and the R&D ELISA demonstrated reasonable agreement, although the Bender MedSystems assay gave significantly lower plasma levels of hsIL6. However, we were very surprised to observe that the IBL assay for measurement of TNF α was not suited for the measurement of plasma

samples because of problems with the limit of detection. The majority of concentrations were below the detection limit of the assay, despite meticulously following the specific instructions. Our replication study confirmed that this was not an incidental finding. Despite the low amount of sample needed, we can currently not recommend the use of this specific assay.

Our results show that when designing a study, it is advisable to thoroughly test all assays needed for the study before measuring samples obtained from long-term storage in biobanks. In addition, the traceability of its standardization is also very important. These samples are usually 'expensive' samples, with limited amounts of sample material available in storage. For testing purposes, we therefore recommend use of sample material obtained in daily practice, samples obtained from (paid) volunteers, or anonymized left-over material from a laboratory or blood bank facility.

In summary, plasma samples stored over a period of 2-4 years at -80 °C showed good stability for measurements of inflammatory biomarkers. Even when the same ELISA method was used, there were small variations in results reported by different laboratories. Although it appears attractive to utilize assays which need only small amounts of sample, such assays should be rigorously tested before large sample sets are measured.

Acknowledgements

The authors wish to acknowledge all participants of the LifeLines Cohort Study and everybody involved in the set-up and implementation of the study. The authors wish to thank Bettine Haandrikman, Lucy Wagenmakers and Eduard Heine, Dept of Laboratory Medicine of the UMCG, and Oddrun K Storrø, Department of Immunology and Transfusion Medicine, St. Olavs Hospital, for performing several of the assays.

Funding

Lifelines has been funded by a number of public sources, notably the Dutch Government, The Netherlands Organization of Scientific Research NOW [grant 175.010.2007.006], the Northern Netherlands Collaboration of Provinces (SNN), the European fund for regional development, Dutch Ministry of Economie Affairs, Pieken in de Delta, Provinces of Groningen and Drenthe, the Target project, BBMRI-NL, the University of Groningen, and the University Medical Center Groningen, The Netherlands. This work was supported by the National Consortium for Healthy Ageing, and funds from the European Union's Seventh Framework program (FP7/2007-2013) through the BioSHaRE-EU (Biobank Standardisation and Harmonisation for Research Excellence in the European Union) project, grant agreement 261433. LifeLines (BRIF4568) is engaged in a Bioresource research impact factor (BRIF) policy pilot study, details of which can be found at:

<https://www.bioshare.eu/content/bioresource-impact-factor>

Competing Interests

No competing financial interests exist.

References

1. Doiron D, Burton P, Marcon Y, et al. Data harmonization and federated analysis of population-based studies: the BioSHaRE project. *Emerg. Themes Epidemiol.* 2013;10: 12-7622-10-12
2. Berndt SI, Gustafsson S, Magi R, et al. Genome-wide meta-analysis identifies 11 new loci for anthropometric traits and provides insights into genetic architecture. *Nat. Genet.* 2013;45: 501-512
3. Inouye M, Kettunen J, Soininen P, et al. Metabonomic, transcriptomic, and genomic variation of a population cohort. *Mol. Syst. Biol.* 2010;6: 441
4. Yang J, Loos RJ, Powell JE, et al. FTO genotype is associated with phenotypic variability of body mass index. *Nature* 2012;490: 267-272
5. Grizzle WE, Gunter EW, Sexton KC, et al. Quality management of biorepositories. *Biopreserv Biobank* 2015;13: 183-194
6. Betsou F, Bulla A, Cho SY, et al. Assays for Qualification and Quality Stratification of Clinical Biospecimens Used in Research: A Technical Report from the ISBER Biospecimen Science Working Group. *Biopreserv Biobank* 2016;14: 398-409
7. Biomarkers Definitions Working Group. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. *Clin. Pharmacol. Ther.* 2001;69: 89-95
8. Ellervik C, Vaught J. Preanalytical Variables Affecting the Integrity of Human Biospecimens in Biobanking. *Clin. Chem.* 2015;61: 914-934
9. Gislefoss RE, Grimsrud TK, Morkrid L. Stability of selected serum hormones and lipids after long-term storage in the Janus Serum Bank. *Clin. Biochem.* 2015;48: 364-369

10. Ikeda K, Ichihara K, Hashiguchi T, et al. Evaluation of the short-term stability of specimens for clinical laboratory testing. *Biopreserv Biobank* 2015;13: 135-143
11. Anton G, Wilson R, Yu ZH, et al. Pre-analytical sample quality: metabolite ratios as an intrinsic marker for prolonged room temperature exposure of serum samples. *PLoS One* 2015;10: e0121495
12. Lee JE, Kim JW, Han BG, et al. Impact of Whole-Blood Processing Conditions on Plasma and Serum Concentrations of Cytokines. *Biopreserv Biobank* 2016;14: 51-55
13. Aziz N, Fahey JL, Detels R, et al. Analytical performance of a highly sensitive C-reactive protein-based immunoassay and the effects of laboratory variables on levels of protein in blood. *Clin.Diagn.Lab.Immunol.* 2003;10: 652-657
14. Doumatey AP, Zhou J, Adeyemo A, et al. High sensitivity C-reactive protein (Hs-CRP) remains highly stable in long-term archived human serum. *Clin.Biochem.* 2014;47: 315-318
15. Ishikawa S, Kayaba K, Gotoh T, et al. Comparison of C-reactive protein levels between serum and plasma samples on long-term frozen storage after a 13.8 year interval: the JMS Cohort Study. *J.Epidemiol.* 2007;17: 120-124
16. Lopez-Campos JL, Arellano E, Calero C, et al. Determination of inflammatory biomarkers in patients with COPD: a comparison of different assays. *BMC Med.Res.Methodol.* 2012;12: 40-2288-12-40
17. Stolck RP, Rosmalen JG, Postma DS, et al. Universal risk factors for multifactorial diseases: LifeLines: a three-generation population-based study. *Eur.J.Epidemiol.* 2008;23: 67-74
18. Scholtens S, Smidt N, Swertz MA, et al. Cohort Profile: LifeLines, a three-generation cohort study and biobank. *Int.J.Epidemiol.* 2015;44: 1172-1180

19. Hardikar S, Song X, Kratz M, et al. Intraindividual variability over time in plasma biomarkers of inflammation and effects of long-term storage. *Cancer Causes Control* 2014;25: 969-976
20. Zander J, Bruegel M, Kleinhempel A, et al. Effect of biobanking conditions on short-term stability of biomarkers in human serum and plasma. *Clin.Chem.Lab.Med.* 2014;52: 629-639

Table 1. Sample analysis scheme for the inflammatory cytokines

	LifeLines	Location A	Location B	Location C
hsCRP	Fresh	ELISA R&D 200 of 240	N.A.	Nephelometry 116 of 240
hsIL6	N.A.	ELISA R&D 200 of 240	ELISA Bender MedSystems 233 of 240	ELISA R&D 238 of 240
hsTNF α	N.A.	ELISA R&D 54 of 240	ELISA IBL 18 of 240	ELISA R&D 239 of 240

* Location B: hsCRP measurement was not performed due to limited amount of sample (0.9 ml) available for this study

Table 2. Assay overview for different inflammatory cytokines

	Producer	Cytokine	Standardization	Min. detection limit	Lowest vs. highest standard
ELISA	R&D systems	hsCRP	NIBSC 85/506	0.010 mg/L	0.78-50 mg/L
		hsIL6	NIBSC 89/548	0.039 ng/L	0.156-10 ng/L
		hsTNF α	NIBSC 88/786	0.106 ng/L	0.5-32.0 ng/L
ELISA	Bender MedSystems	hsIL6	NIBSC 89/548	0.030 ng/L	0.08-5.0 ng/L
ELISA	IBL International	hsTNF α	NIBSC 87/650	0.13 ng/L	0.31-20.00 ng/L
Nephelometry	Siemens Healthcare Diagnostics	hsCRP	CRM 470	0.175 mg/L	0.175-11.00 mg/L

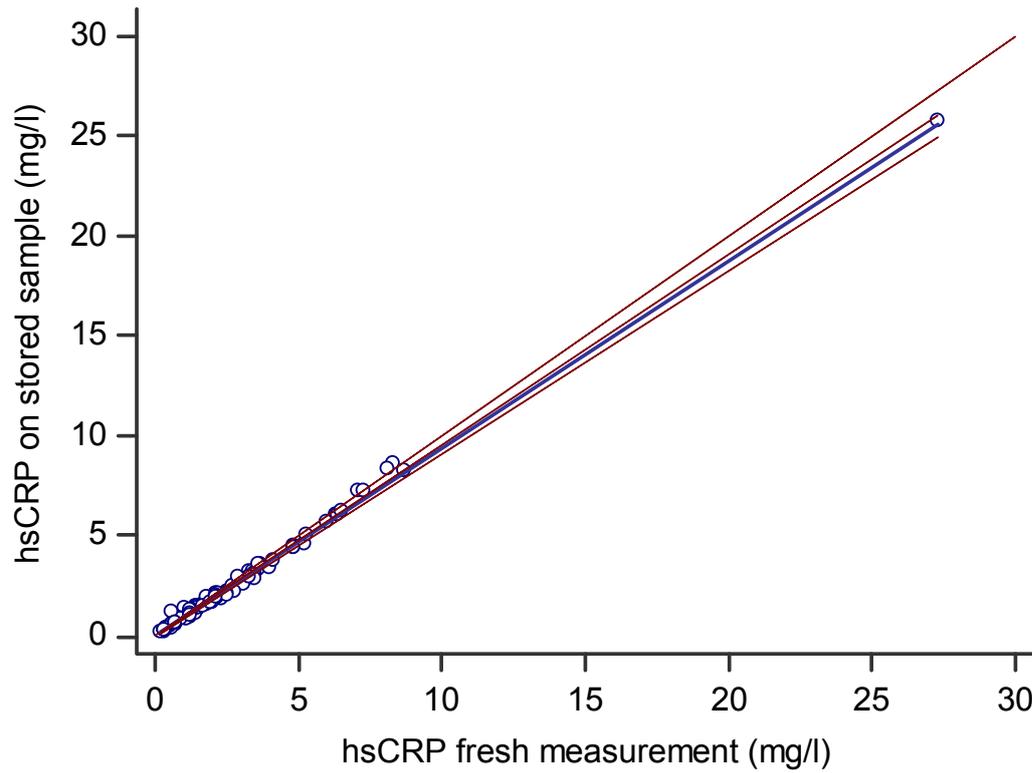


Figure 1. Passing - Bablok analysis of measurement of hsCRP with nephelometry in 4-years stored samples compared to the results obtained in fresh samples, measured on the same day as the blood was drawn.

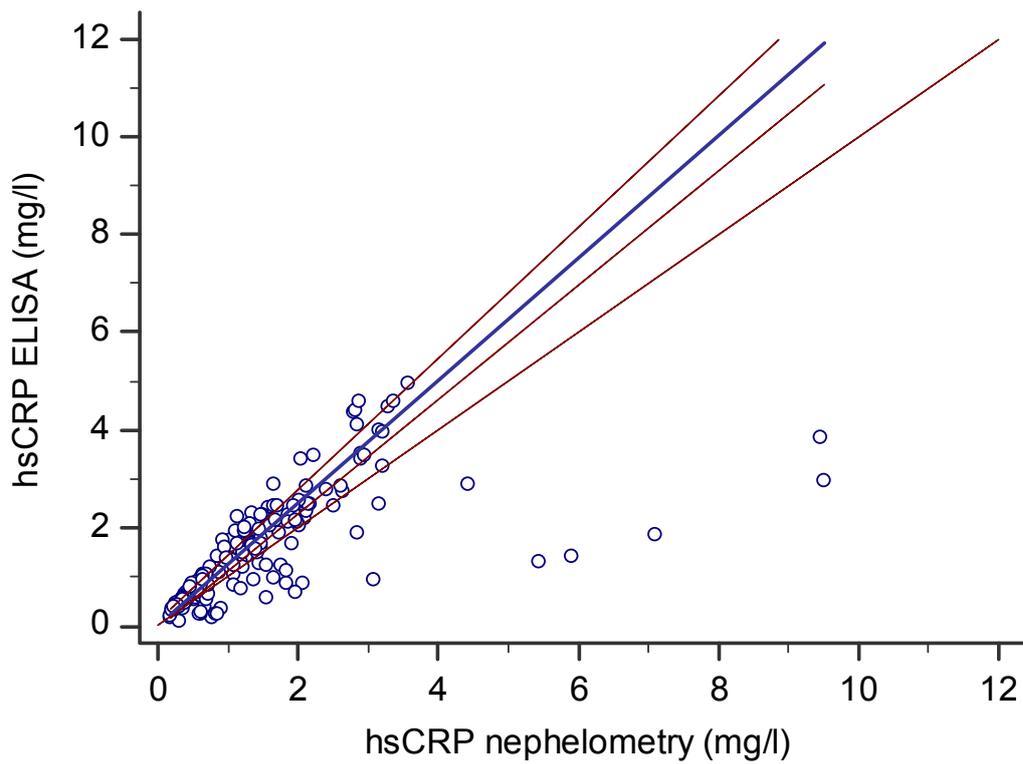


Figure 2. Passing Bablok regression and the corresponding regression characteristics for hsCRP measured with nephelometry (location C) versus R&D ELISA (location A).

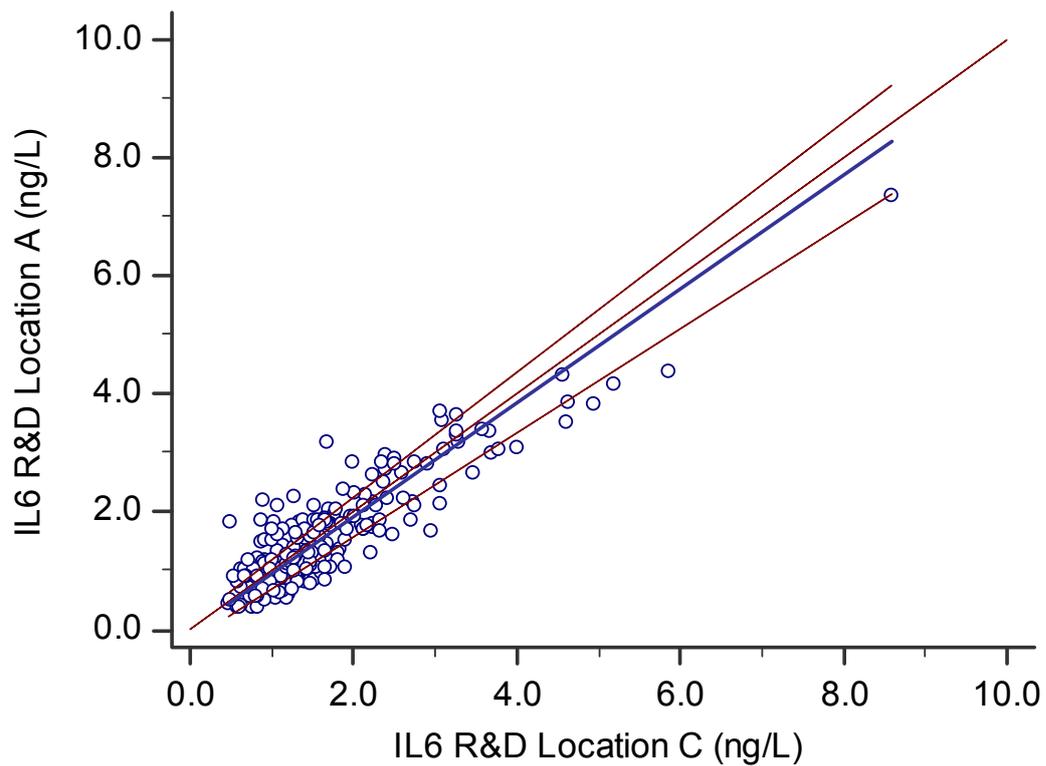


Figure 3. Passing-Bablok regression for IL6 (in ng/L) with R&D ELISA assay on location C versus location A.

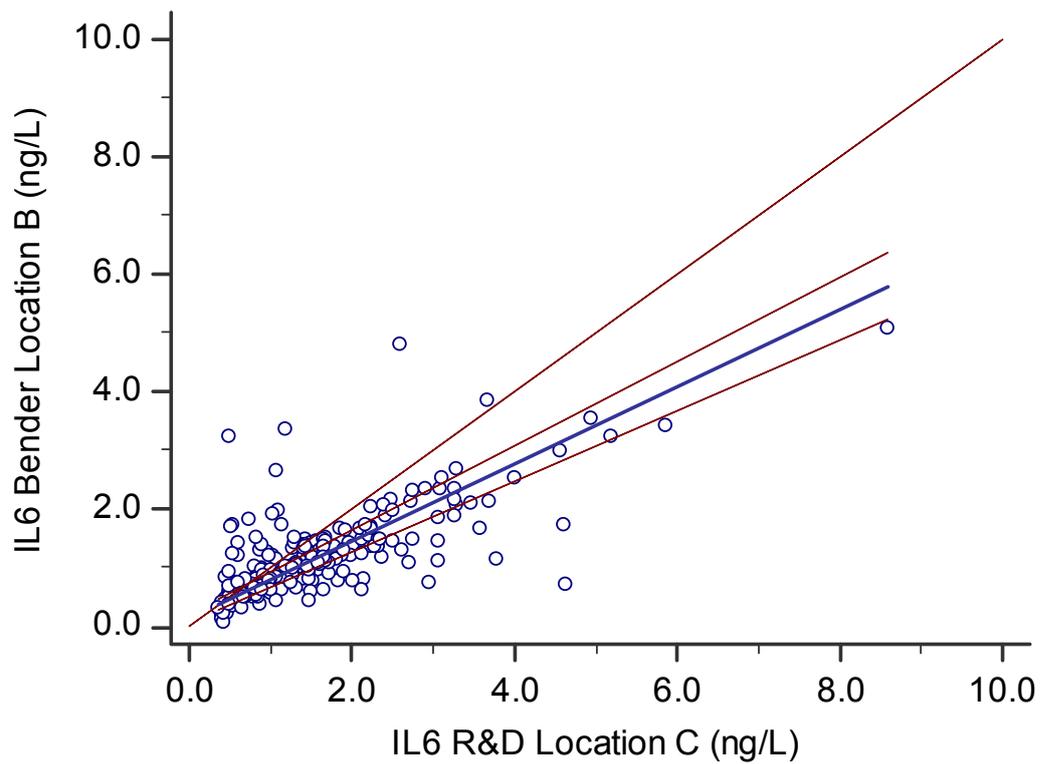


Figure 4. Passing-Bablok regression for IL6 (in ng/l) measured with R&D ELISA assay in location C and Bender MedSystems in location B.

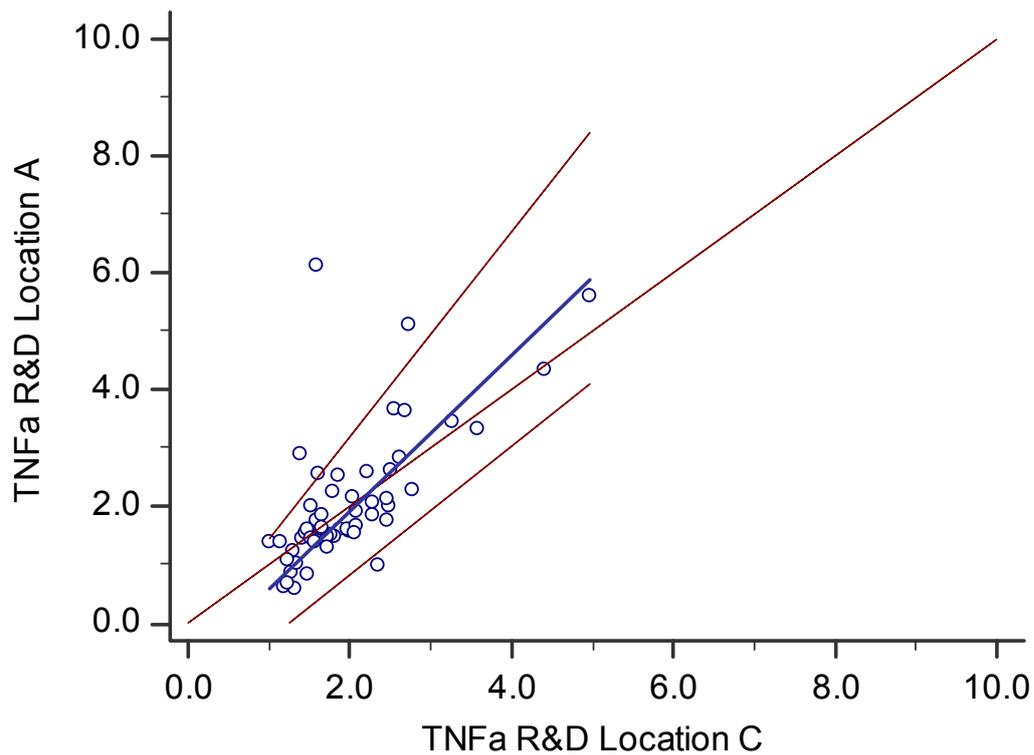


Figure 5. Passing-Bablok regression for TNFa measured with R&D ELISA assay both on location C and location A. Only n=54 samples available for comparison (for explanation see text).

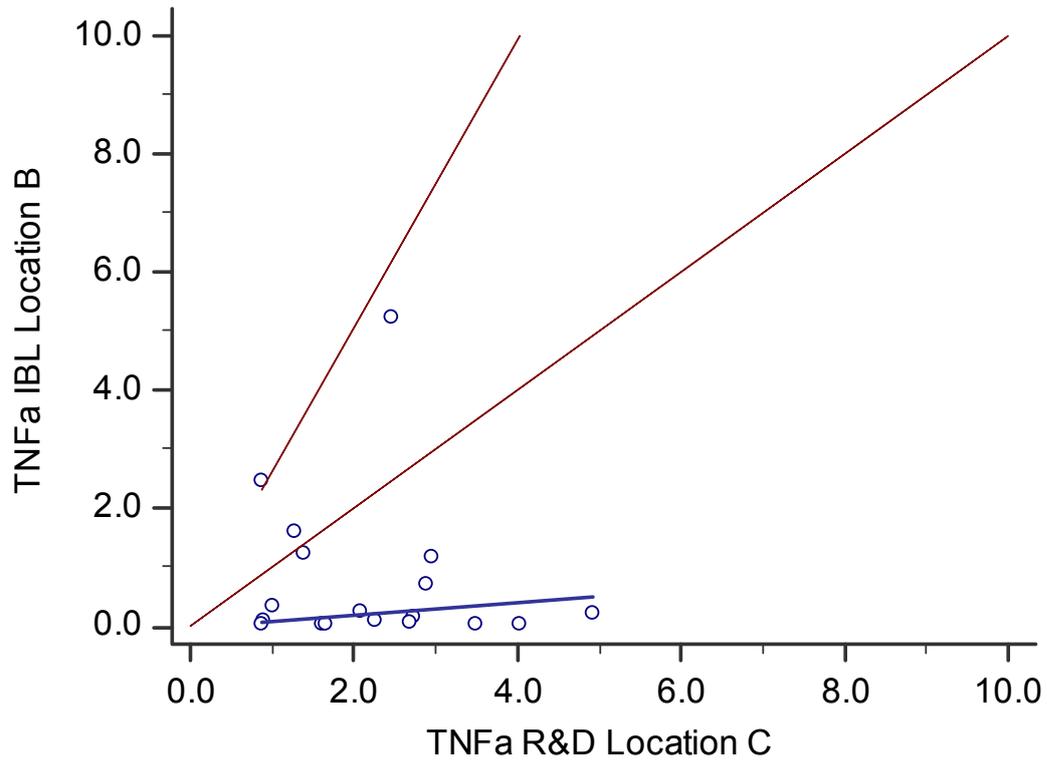


Figure 6. Passing-Bablok regression for TNFa (in ng/l) measured with R&D ELISA assay in Location C and IBL ELISA in location B. Only 18 samples available for comparison, all other samples yielded results with the IBL assay which were below the detection limit.

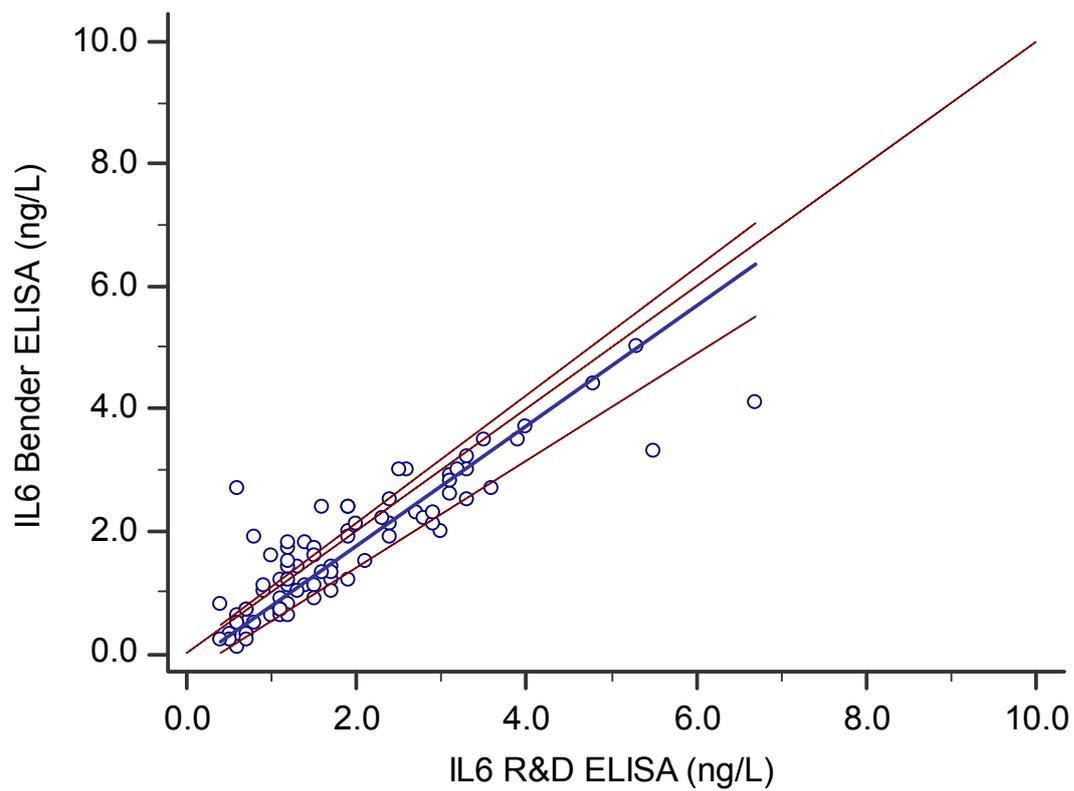


Figure 7. Passing Bablok regression for IL6 (in ng/L) with R&D ELISA versus Bender MedSystems ELISA in 80 obese participants in a weight-reduction program

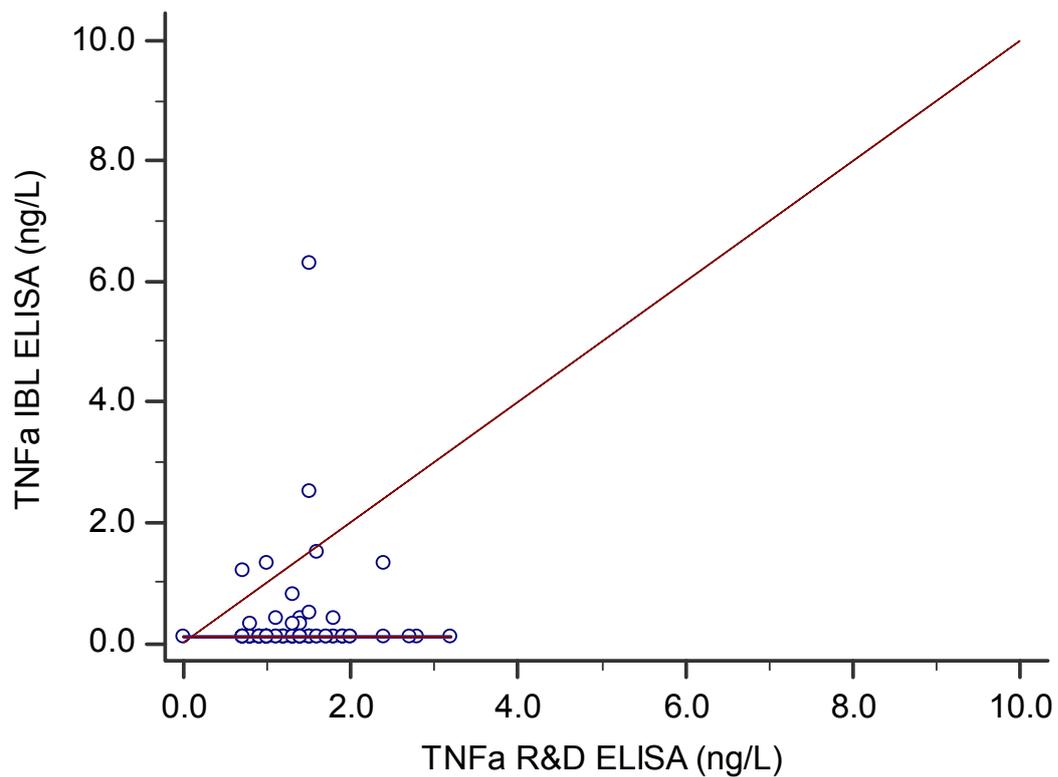
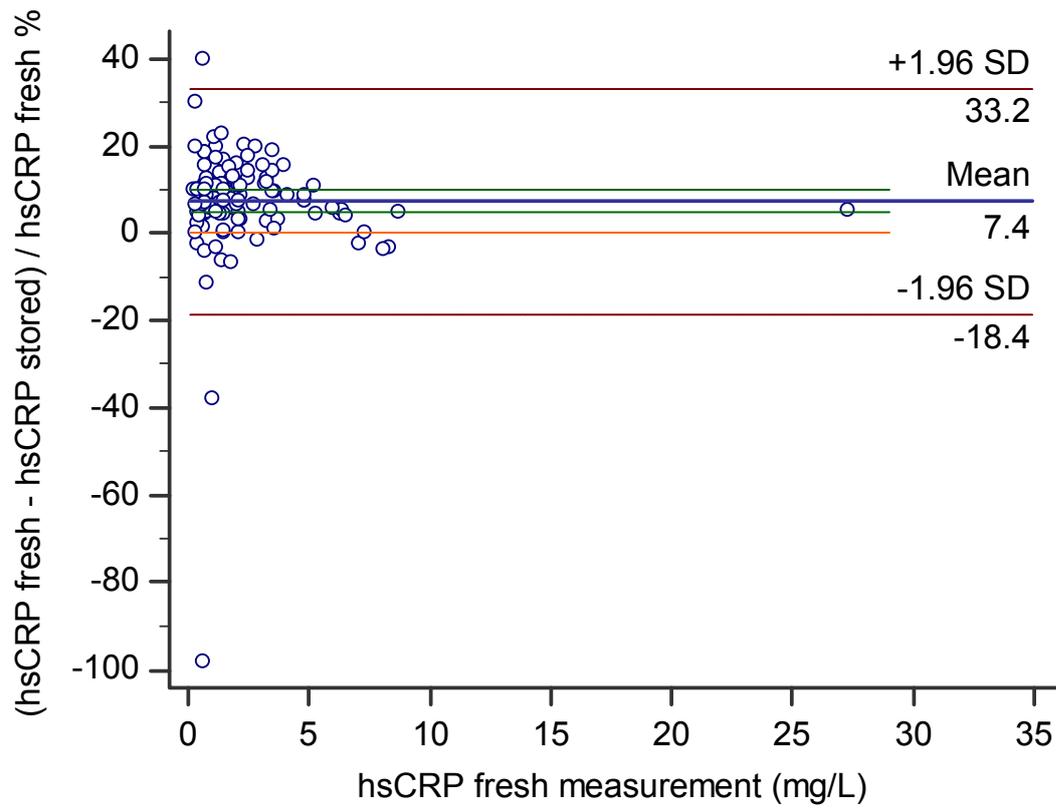
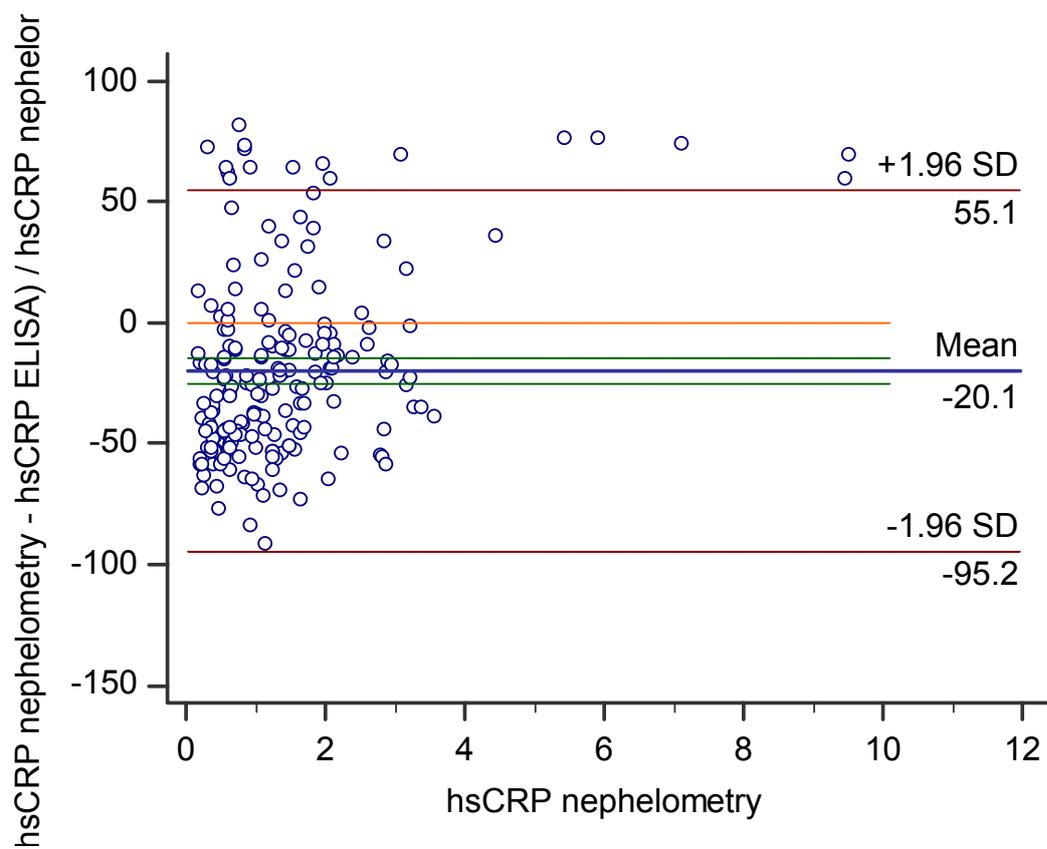


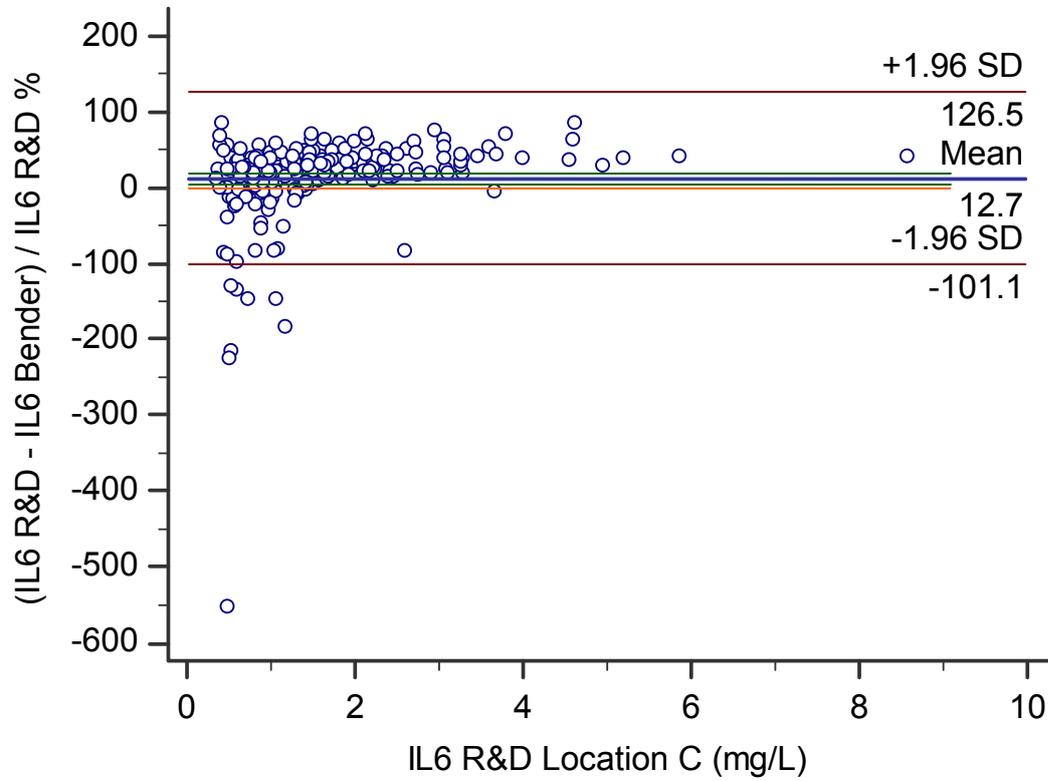
Figure 8. Measurement of TNF α (in ng/L) with R&D ELISA versus IBL ELISA in 80 obese participants in a weight-reduction program



Supplementary Figure 1. Bland-Altman plot of measurement of hsCRP with nephelometry in the 4-years stored samples compared to the results obtained in fresh samples.



Supplementary Figure 2. Bland-Altman plot for hsCRP measured with nephelometry (location C) versus R&D ELISA (location A).



Supplementary Figure 3. Bland-Altman plot for IL6 measured with R&D ELISA assay in location C and Bender MedSystems in location B.

