Reproducibility of measurement techniques used for creatine kinase, interleukin-6 and high-sensitivity c-reactive protein determination over a 48 h period in males and females

Journal:	*Measurement in Physical Education and Exercise Science*
Manuscript ID	HMPE-2017-0060.R2
Manuscript Type:	Original Article
Keywords:	exercise induced muscle damage, biomarkers, cytokines, inflammation, Reliability

Reproducibility of measurement techniques used for creatine kinase, interleukin-6 and high-sensitivity c-reactive protein determination over a 48 h period in males and females

Reproducibility of measurement techniques for blood biomarker determination

Abstract

To examine the reproducibility of three measurement techniques used to determine creatine kinase, interleukin-6, and high-sensitivity C-reactive protein, fifty participants had blood samples taken on two occasions. Fingertip plasma samples were analysed using the Reflotron for CK determination. Venous blood samples collected into serum separator tubes were used for IL-6 and hs-CRP analyses. IL-6 was measured using an enzyme linked immune assay development kit. The hs-CRP was measured by an in-house ELISA method. Dependent t-tests showed no systematic bias between samples. The interdian CV was 20.0% for CK, 15.3% for IL-6 and 44.2% for hs-CRP. The intraclass correlation coefficient was 0.90 for CK, 0.98 for IL-6 and 0.70 for hs-CRP. The 95% limits of agreement were -69.7 to 63.5 IU/L for CK, -1.48 to 1.80 pg/ml for IL-6 and -1.10 to 0.91 µg/L for hs-CRP. The results demonstrate low absolute reproducibility, which may obscure a true experimental effect.

Key words: Exercise induced muscle damage; Reliability; Biomarkers; Cytokines; Inflammation

Introduction

Circulating blood biomarkers of muscle damage [e.g. creatine kinase (CK)] and inflammation [e.g. interleukin 6 (IL-6)] are ubiquitous in exercise induced muscle damage (EIMD) research (Clifford, Bell, West, Howatson, & Stevenson, 2016; McLeay, Stannard, Mundel, Foskett, & Barnes, 2016; Tseng et al., 2016; Vieira et al., 2016), prolonged endurance research (Gill et al., 2015; Nielsen, Oktedalen, Opstad, & Lyberg, 2016; Niemela, Kangastupa, Niemela, Bloigu, & Juvonen, 2016; Scherr et al., 2011), and recovery related research within team sports (Coelho, Morandi, de Melo, & Silami-Garcia, 2011; Fullagar, Skorski, Duffield, & Meyer, 2016; Harper et al., 2016; Romagnoli et al., 2016). Such markers are frequently employed to determine the magnitude of EIMD, and monitor readiness to train, despite their inherent limitations. For example, there is large inter-individual variability of circulating cytokines (Paulsen, Mikkelsen, Raastad, & Peake, 2012), and CK (Kraemer et al., 2013) following EIMD. Furthermore, basal levels of these markers may also depend on a number of individual factors such as age (Horska, Fishbein, Fleg, & Spencer, 2000), body composition (Salvadori, Fanari, Ruga, Brunani, & Longhini, 1992) and training status (Vincent & Vincent, 1997), although these findings are equivocal. Nevertheless, when used in combination with other measurement tools (e.g. maximum voluntary contraction, perceived soreness), they may provide researchers and practitioners with important information regarding the muscle damage and inflammatory response to various exercise paradigms, in addition to the recovery profile across many sports. Therefore, if these blood markers are to be successfully employed within a research context, and utilised by practitioners, it is essential that research investigates the intra-individual variability (i.e. reproducibility) of the measurement techniques utilised for their determination. This information is integral to interpreting a change from baseline that is attributed to muscle damage and/or inflammation, and not

simply due to random variability. For example, if the magnitude of change from baseline was within the test-retest error, then the measurement techniques employed to determine these biomarkers would not be considered appropriate.

Recent research measuring baseline CK and IL-6 in male soccer players across a 7 d period reported a CV of 18.5% for both markers (Harper et al., 2016). Whilst such data is useful for researchers and practitioners employing these measurements, in practice, such measures may be collected on a more frequent basis. For example, in football a period of 72 h is considered sufficient to achieve recovery to pre-match performance values, although in reality players have much less time between fixtures and training (Owen et al., 2015). CK values typically peak 24 – 48 h post exercise (Brancaccio, Maffulli, & Limongelli, 2007; Meister, Aus der Funten, & Meyer, 2014; Young, Hepner, & Robbins, 2012), and therefore, measurements of these markers within a 24 – 72 h period are likely to be employed (Howatson et al., 2010; Niemela et al., 2016). Therefore, the aim of the present study was to examine the reproducibility of three measurement tools used to detect the circulating blood biomarkers (CK, IL-6 and hs-CRP) in healthy males and females, over a 48 h period. For CK fingertip blood samples were obtained and analysed using the Reflotron. Venous blood samples were collected for IL-6 and hs-CRP analyses using serum separator vacuette tubes. A R&D DuoSet (DY206) enzyme linked immune assay (ELISA) development kit (R&D Systems, Abingdon, UK) was used for IL-6 determination, and for hs-CRP an in-house ELISA method using anti-human CRP antibodies, calibrators and controls from Abcam (Abcam®, Cambridge, UK) was employed.

Methods

Participants

The 50 (14 female), apparently healthy participants who volunteered for this study had the following characteristics: median (min - max) age = 26 (18 – 49) y; mean (SD) height = 176 (9) cm and mean (SD) mass = 70.7 (11.8) kg. Participants were free from musculoskeletal injury, non-smokers, and engaged in regular physical activity (> 30 min, three times a week for at least 6 months). There was no control for the use of any oral or transdermal contraceptive medication. However, all females were eumenorrheic and completed the testing in the luteal phase of menstruation. Participants provided written informed consent, and were asked to adhere to written pre-measurement procedures for the duration of the study. These pre-measurement procedures stipulated that participants did not engage in any exercise for 7 d prior to commencing the study, that no large meals or stimulants were consumed within 4 h of each measurement, that the participants followed and replicated the same diet during the testing period, and that at least 500 ml of fluid was consumed 2 h prior to each measurement. There was no exercise allowed in between the trials. Adherence to these procedures was monitored using a pre-measurement procedure checklist, which participants completed and signed prior to the commencement of each measurement. The apparent adherence was 100% in all instances. Additionally, a pre-test health and medical questionnaire was completed and checked prior to each trial, in order to monitor the participants sleep, diet, and any health conditions (e.g. the common cold). There were no reported cases of changes in health (i.e. the common cold) or diet during the testing period. Participants were free to leave the study at any point without reason, and anonymity, and confidentiality was ensured. Ethical approval was granted by the University of Hull, Department of Sport & Exercise Science Ethics Committee.

Experimental approach to the problem

Participants visited the laboratory on two separate occasions with each visit separated by 48 h. Visit times were held constant within individuals (± 0.5 h) to negate any effects of circadian variation (Drust, Waterhouse, Atkinson, Edwards, & Reilly, 2005). Blood samples were collected as described below.

Blood collection and analyses

Collection and storage of blood samples

Prior to any blood collection, participants were asked to lay in a supine position for 15 min. Fingertip capillary blood samples were collected using standard techniques. Firstly, the finger was cleaned using a sterile alcohol wipe and an Accucheck softclicks lancet device (Roche Diagnostics, Manheim, Germany) was used to puncture the area. The initial drop of blood was wiped away and the second drop used for analysis.

Venous blood samples were drawn from a superficial vein in the antecubital fossa of the forearm using standard venepuncture techniques. Samples were collected into serum separator vacuette tubes (Greiner Bio-one, Kremsmunster, Austria) and left to clot for 30 min at room temperature before being centrifuged in a Heraues Labofuge 400R (Kendro Labatory products, Bishops Stortford, UK) at 1509 g and 19°C for 15 min. Subsequently, serum was pipetted and stored in several 1.0 ml CryoPure tubes (SARSTEDT Ltd., Beaumont Leys, Leicestershire, UK) in a -80°C freezer until transportation to the specialist assay laboratory at Manchester Royal Infirmary.

Haemoglobin (Hb) and haematocrit (Hct) concentrations were obtained to measure changes in plasma volume according to the Dill and Costill method (Dill & Costill, 1974). Blood measures were subsequently adjusted to account for shifts in plasma volume. The percentage change in plasma volume was either added or subtracted from the concentration of the blood biomarker as required.

Blood analyses

Creatine kinase

Plasma CK was determined from a 32 µl fingertip capillary blood sample obtained using standard techniques while participants were semi-recumbent on a treatment couch. The sample of whole blood was immediately pipetted to a test strip and analysed for CK using a colorimetric assay procedure (Reflotron, Boehringer Mannheim, Germany). Measurements were obtained and analysed in triplicate and the average used. The Reflotron was used due to ecological validity as several studies employ this measurement tool (Coelho et al., 2011; Howatson, Goodall, & van Someren, 2009; Howatson et al., 2012; McLellan, Lovell, & Gass, 2010; Owen et al., 2015). This system uses a plasma separation principle which is incorporated in the reagent carrier on the test strip. Briefly, following application of the whole blood to the test strip the sample flows into the reaction zone where erythrocytes are separated from the plasma. Subsequently, the formation of dye is measured kinetically at 642 nm and at 37°C. According to manufacturer guidelines, the 'normal' range of CK activity is 24 – 195 IU/L and 24 – 170 IU/L for males and females respectively. The intra-assay CV was 9.0%.

The analysis of IL-6 and hs-CRP was performed using serum in accordance with previously published methods (Pemberton, Aboutwerat, Smith, & Warnes, 2006). Paired samples were always analysed on the same plate with the same assay kit. If an absorbance value for a sample was greater than the highest standard (15µg/L for hs-CRP, 100pg/ml for IL-6) then the analysis was repeated using sample at a higher dilution.

Interleukin-6 (IL-6)

IL-6 was measured using a R&D DuoSet (DY206) enzyme linked immune assay (ELISA) development kit (R&D Systems, Abingdon, UK). The range of the assay is up to 100 pg/ml, and the minimum detection limit calculated from the mean plus two standard deviations of 12 single replicate analyses of reagent blank was found to be 0.6 pg/ml. Intra and inter-assay CV was 5.9% and 17.2%, respectively.

High-sensitivity C-reactive protein (hs-CRP)

The hs-CRP was measured by an in-house ELISA method using anti-human CRP antibodies, calibrators and controls from Abcam (Abcam®, Cambridge, UK). The ELISA technique is based on the antibody sandwich principle. First, the capture antibody (rabbit anti-human CRP antibody) was bound to a microtitre plate to create a solid phase. A blocking buffer containing BSA (Sigma-Aldrich®, Poole. UK) was then added. Following a wash, samples, standards and controls were incubated with the solid phase antibody that captures the CRP. After further washing, the conjugated detection antibody (HRP-labelled goat anti-human CRP antibody) was added. This detection antibody binds to a different epitope of CRP thus

completing the sandwich. Subsequently the samples were washed to remove unbound detection antibody, the substrate O-phenylenediamine (Sigma-Aldrich®, Poole, UK) was added and colour developed in proportion to the amount of bound HRP. Colour development was stopped by addition of strong acid and the intensity of colour then measured at $\lambda = 490$ nm. The content of each sample was then calculated from the standard curve. The range of the assay is up to 15 µg/L and the minimum detection limit, calculated from the mean plus two standard deviations of 8 replicate analyses of reagent blank, was 0.1 µg/L**.** Intra and inter-assay CV was 4.7% and 5.3%, respectively.

Statistical analyses

The number of participants required for this study was determined *a priori* with an alpha level of 0.05 using a 2 tailed t-test for the main outcome measures for a follow up study (data not shown) using Power Analysis and Sample Size Software (PASS) version 13.0 (NCSS, LLC, Utah, USA). A sample size of n = 50 achieved 99% and 94% power to detect the required minimum worthwhile effects. Analyses were completed using the statistical software package IBM SPSS Statistics version 19.0 (SPSS Inc, Chicago, IL, USA) and graphs created using SigmaPlot version 12.3 (Systat Software Inc, CA, USA). Standard graphical methods were preferred over null hypothesis significance testing to check statistical assumptions (Grafen & Hails, 2002). For descriptive purposes the mean and standard deviation have been used to report the central tendency and dispersion of the observed data. Combinations of statistical methods were chosen in order to compare reliability between different measures and different studies.

Systematic bias was tested using two-tailed dependent t-tests, and also examined graphically with Bland-Altman plots. Absolute measurement error was determined using repeated measures CV and 95% limits of agreement (LoA). The CV (expressed as a percentage) was calculated by dividing the standard deviation of the differences by the square root of two and dividing the answer by the grand mean (Hopkins, 2000). The intra assay CV for CK, IL-6 and hs-CRP was measured in duplicate for all participants (n = 50). The % CV for each sample was calculated by finding the standard deviation of results 1 and 2, dividing that by the duplicate mean and multiplying by 100. The average of the individual CV's was reported as the intra-assay CV. Inter-assay CV for IL-6 and hs-CRP were measured in duplicate for both high and low controls on ten different plates. The plate means for high and low were calculated and used to obtain the overall mean, standard deviation and % CV. Overall % CV was calculated by dividing the standard deviation of the plate means by the mean of the plate means and multiplying this by 100. The average of the high and low % CV is reported as the inter-assay CV. Relative reliability was determined using a two-way random model intraclass correlation coefficient (ICC), which is a measure of the ratio of between-subject variance to within-subject variance. A combination of reproducibility statistics were employed in order to allow comparison with published research. Additionally, data were analysed separately by sex. There were no significant differences between males and females ($p \geq 0.23$) and therefore, the data set was collapsed for the final analyses to increase statistical power and satisfy the principle of parsimony. The two-tailed alpha level for significance testing was set as $p < 0.05$.

Results

The mean (SD) values of the circulating blood biomarkers obtained from measurement one and two are shown in Table 1. Dependent t-tests indicated there was no systematic bias ($p \geq 0.24$) for any of the variables (Table 1).

***Insert Table 1 here ***

The CV, ICC and 95% LoA for CK, IL-6 and hs-CRP are reported in Table 2. The individual mean circulating blood biomarker results plotted against their individual differences (Bland-Altman plots) are shown in Figure 1.

***Insert Table 2 here ***

***Insert Figure 1 here ***

Discussion

The aim of the present study was to examine the reproducibility of three measurement tools used to assess CK, IL-6 and hs-CRP. The main findings in the present study were that although there was no systemic bias between trials for CK, IL-6 and hs-CRP (Table 1), the CV results show that all the blood biomarkers demonstrated high variability (i.e. low reproducibility) (Table 2).

In support of the findings in the present study, CV's of 19% (Nicholson, Morgan, Meerkin, Strauss, & McLeod, 1985) and 18.5% (Harper et al., 2016) for CK have been previously reported. Conversely, lower CV's of 10.5% (Chen, Lin, Chen, Lin, & Nosaka, 2011) and 3.4% (Horder et al., 1991) for CK have also been shown. Discrepancy between these findings

may be due to random biological variation, technical error, differences in timings between measurements, and measurement technique. For example, Chen and colleagues used spectrophotometry to measure CK compared to the Reflotron used in the present study. The intra-assay CV for the analysis of CK in the present study was 9%, supporting the premise that some of the variability resides within the measurement technique itself (i.e. the Reflotron). Nevertheless, the CV in the present study was comparable to Harper and colleagues, despite the different timings between measurements (i.e. 48 h compared to 7 d), and measurement technique employed (i.e. Reflotron versus Cobas 8000). A Reflotron was employed in the present study as this is typically utilised in research, and by practitioners (Coelho et al., 2011; Howatson et al., 2012). However, both the Howatson and Horder studies utilised a Reflotron and reported similar and more reliable CV's; < 3% and 3.4% respectively. Therefore, differences in the interdian CV between the Howatson study and the present study may be due to the population employed, as opposed to the measurement technique. Howatson and colleagues used only male participants in their study (n = 12) who were trained in the competitive national league for football (homogenous population), compared to recreationally active males and females (heterogeneous population) in the present study. We acknowledge that despite no significant differences between males and females in the present study, the effects of menstruation, oral and transdermal contraceptives were not measured. Subsequently, this could explain the higher interdian variability observed. However, Harper and colleagues also used a homogenous population (i.e. males, n = 7) and reported a similar interdian CV to the present study (i.e. 18.5%). Previous research has suggested that high variability in baseline CK values could be due to minor injury, genetic factors, physical activity status, and medication (Prelle et al., 2002). Subsequently, strict pre-test procedures were employed in the present study with apparent 100% adherence. Despite this, low reproducibility between testes was still observed. Therefore, researchers and

practitioners must account for this error when utilising CK values to determine the presence of muscle damage, as changes in CK values could be used to prescribe recovery techniques (e.g. pharmacological aids such as ibuprofen) and/or modify the training load, otherwise incorrect application of both recovery techniques and changes to training load could be performed. Additionally, this highlights the importance of not utilising just one measurement (e.g. CK) when making such decisions.

Basal levels of CK are considered to be 24 – 195 IU/L, which is in agreement with the findings in the present study (Table 1). Similarly, Clarkson and colleagues reported average baseline CK values of 118 IU/L (range 33 – 481 IU/L) in 203 participants, although no reproducibility statistics were reported in their study (Clarkson, Kearns, Rouzier, Rubin, & Thompson, 2006). Post exercise concentrations of CK in athletes are thought to range between 300 – 500 IU/L (Mougios, 2007). Furthermore, Coelho, Morandi, de Melo, & Silami-Garcia (2011) observed average CK values (measured using a Reflotron) of 786 IU/L (min – max; 96 – 1580 IU/L), 388 IU/L (38 – 749 IU/L), 299 IU/L (31 – 595 IU/L) and 317 IU/L (197 – 654 IU/L) 12 – 20 h, 36 – 48 h, 60 – 65 h and 90 – 110 h respectively following a soccer game. Similarly, Howatson et al. (2012) observed a 3-4 fold increase in CK, which peaked at 24 h post eccentric exercise. Subsequently, it appears that on average the magnitude of changes in CK following exercise, is typically outside the CV reported in the present study (i.e. 20%). Similarly, the magnitude of change in CK far exceeds the 95% LoA reported in the present study (Figure 1), suggesting the test-retest error of this measurement technique is low enough not to obscure a true experimental effect. However, Clarkson and colleagues also measured CK 4, 7 and 10 days following a bout of eccentric exercise, and the authors reported an increase in CK to 7713 IU/L (range 55 – 80, 550 IU/L), 2603 IU/L (range 49 – 21, 675 IU/L) and 486 IU (45 – 7034 IU/L) respectively, highlighting high variability in

the CK response. The average percentage increase in CK was 6420, 2100 and 311% across the 4, 7 and 10 d respectively. Subsequently, the minimum change in CK reported in the aforementioned studies means that some individuals would fall within the test-retest error value and 95% LoA reported in the present study (Figure 1). However, as the authors did not conduct their own reproducibility statistics, it cannot be assumed that the test-retest error would be comparable to the present study. In order for practitioners and researchers to accurately distinguish between a 'true' experimental effect and test-retest error, it is vital that the interdian and intra assay CV are reported, and compared to the experimental effect.

Previous research has reported a CV of 27% for IL-6 (Knudsen et al., 2008), which is slightly higher than that found in the present study (15.3%). The average IL-6 values reported in the present study (Table 1) appear to be higher than those reported in previous research (Abedelmalek et al., 2013; Conceicao et al., 2012; Nieman et al., 2005; Scherr et al., 2011), whilst the hs-CRP values are comparable to previous studies (Peake, Nosaka, Muthalib, & Suzuki, 2006; Robson-Ansley et al., 2009; Scherr et al., 2011). The reasons behind the higher baseline values for IL-6 in the present study are not clear. Factors such as upper respiratory tract infections (Martin, Pence, & Woods, 2009), training status (Fischer, 2006), and time of day (Abedelmalek et al., 2013) are known to affect IL-6 values. Time of day was standardised within participants in the present study, however, between participant testing time was not constant. Moreover, although all participants met the inclusion criteria regarding regular physical activity, the population sample was heterogeneous. Therefore, the combination of between participant differences in time of day, and the heterogeneity of training status may have contributed to the variability in IL-6, and could explain the high values observed (See Table 2, and Figure 1) compared to previous research. Furthermore, the intra and inter-assay CV for IL-6 and hs-CRP were 5.9% and 17.2%, and 4.7% and 5.3%

respectively. This suggests that the technical error component is somewhat large for the inter-assay CV for IL-6 (17.2%). Conceicao et al. (2011) reported intra and inter assay variability of 7.8% and 7.2% respectively for IL-6. Similarly, Peake, Nosaka, Muthalib, & Suzuki (2005) reported an inter assay CV of 5.4% for IL-6. Furthermore, Robson-Ansley et al. (2009) reported intra and inter assay CV's of < 4% and, 5% for hs-CRP respectively, which is similar to the intra and inter-assay CV in the present study (4.7% and 5.3%). The high inter-assay CV for IL-6 in the present study may account for the higher baseline values of IL-6, which could have been caused by technical error in the measurement technique. However, the five participants who had a baseline level of IL-6 > 10 pg/ml in trial 1 also had an almost identical value for trial 2. Therefore, it is more plausible that these five participants had higher IL-6 values due to the time of day, and training status, as previously discussed. Nevertheless, the large interdian CV (Table 2), and low reproducibility reported in the present study for IL-6 is clearly a combination of both technical and biological variation. The high interdian CV for hs-CRP in the present study (Table 2), appears to support previous findings demonstrating considerable intra-individual variability in males (n = 100) with either previous history of coronary artery disease or no history (Bogaty et al., 2013), and healthy older men (n = 50) measured over a four year period (Platz et al., 2010). Moreover, in a mixed sex (n = 541) study the interdian CV for hs-CRP measured on average 18.9 d apart was 46.2% (Bower, Lazo, Juraschek, & Selvin, 2012), which is almost identical to the 44.2% reported in the present study, despite the fact that we utilised an in house assay for hs-CRP. However, even when employing the same assay, reproducibility may still differ as it is dependent on many factors (i.e. laboratory techniques, sample population, etc.). Subsequently, it is important that researchers and practitioners conduct their own reproducibility using their own measurement tools, and replicate this within populations, to distinguish between the true test-retest error and experimental effect. Overall, these factors

highlight the importance of all studies reporting reproducibility statistics for the measurement tools employed and ensuring adherence to pre-test procedures to control for any confounding factors, which may affect baseline levels. Again, this highlights the potential problem of using these bio-markers in isolation, and it is recommended that multiple measurement techniques (e.g. wellness questionnaires, perceived exertion, maximal voluntary contraction) are also obtained to ensure robust and reliable information is gathered.

The ICC's reported in the present study (Table 2) suggest 'high', 'high' and 'questionable' reproducibility for CK, IL-6 and hs-CRP respectively. Therefore, if we had simply used the ICC in isolation within our experimental design to determine whether these measurement tools are reproducible, it would indeed suggest that CK and IL-6 are reproducible markers. These findings appear to contradict the values observed for the CV's. High ranges observed for the blood biomarkers (Figure 1) could have increased the size of the correlation. Therefore, 'high' ICC's may be a reflection of more heterogeneous results rather than high reproducibility. This highlights the potential problem when researchers purely employ a correlation (e.g. Pearson's r or ICC) to determine the reproducibility of a measurement tool. Consequently, it has been suggested that the CV may be a more accurate representation of the reproducibility of a particular measure, and therefore, it is suggested that these results are used to make recommendations in the present study. Nevertheless, previous research has reported wide variation in ICC's for the blood biomarkers, ranging from questionable to moderate. For example ICC'S of 0.48 (Navarro et al., 2012), 0.66 (Karakas et al., 2010) and 0.47 – 0.80 (Walshe et al., 2010) have all been reported for IL-6. Additionally, poor reproducibility has been shown for hs-CRP with ICC'S of 0.62 (Navarro et al., 2012) and 0.66 (Karakas et al., 2010).

Conclusion

Due to the high interdian CV of CK, IL-6 and hs-CRP, it could be recommended that these variables are not used in isolation within research and by practitioners, as the high test-retest error associated with these markers may obscure a true experimental effect. However, whilst the CV values observed in the present study are outside the recommended threshold of 10%, this does not necessarily imply that these biomarkers are not appropriate measurement tools. To assess the usefulness of a measurement tool, it is critical to compare the CV with the minimum worthwhile effect. As long as the CV is smaller than the minimum worthwhile difference, then the measurement tool can be deemed appropriate.

Further research is required in order to investigate the biological variation and error of these biomarkers over different periods of time. For example, it would be useful to quantify the intra-individual variation and error over a more acute (1 – 12 h) and chronic (24 h – 7 d) period as several blood samples are often obtained within research in order to investigate the recovery pattern following different types of exercise. It is advised that researchers and practitioners include a reproducibility trial within their research design in order to account for potential error within their measurements, and allow for greater accuracy when reporting changes in these biomarkers. Furthermore, future work should seek to perform repeated reproducibility in the same population to provide an accurate profile of an individual's variability in these markers.

Acknowledgements

The authors would like to thank the University of Hull, Department of Sport, Health and Exercise Science for the support and opportunity to conduct this study. The authors would also like to thank the volunteers for participation in this study.

Conflicts of interest

The authors declare that this research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Funding

None

References

Abedelmalek, S., Chtourou, H., Aloui, A., Aouichaoui, C., Souissi, N., & Tabka, Z. (2013). Effect of time of day and partial sleep deprivation on plasma concentrations of IL-6 during a short-term maximal performance. *European Journal of Applied Physiology, 113*(1), 241-248. doi: 10.1007/s00421-012-2432-7

Bogaty, P., Dagenais, G. R., Joseph, L., Boyer, L., Leblanc, A., Bélisle, P., & Brophy, J. M. (2013). Time Variability of C-Reactive Protein: Implications for Clinical Risk Stratification. *Plos One, 8*(4), e60759. doi: 10.1371/journal.pone.0060759

Bower, J. K., Lazo, M., Juraschek, S. P., & Selvin, E. (2012). Within-person variability in high-sensitivity C-reactive protein. *Arch Intern Med, 172*(19), 1519-1521. doi: 10.1001/archinternmed.2012.3712

Brancaccio, P., Maffulli, N., & Limongelli, F. M. (2007). Creatine kinase monitoring in sport medicine. *British Medical Bulletin, 81-82*, 209-230. doi: 10.1093/bmb/ldm014

Chen, T. C., Lin, K. Y., Chen, H. L., Lin, M. J., & Nosaka, K. (2011). Comparison in eccentric exercise-induced muscle damage among four limb muscles. *European Journal of Applied Physiology, 111*(2), 211-223. doi: 10.1007/s00421-010-1648-7

Clarkson, P. M., Kearns, A. K., Rouzier, P., Rubin, R., & Thompson, P. D. (2006). Serum creatine kinase levels and renal function measures in exertional muscle damage. *Med Sci Sports Exerc, 38*(4), 623-627. doi: 10.1249/01.mss.0000210192.49210.fc

Clifford, T., Bell, O., West, D. J., Howatson, G., & Stevenson, E. J. (2016). The effects of beetroot juice supplementation on indices of muscle damage following eccentric exercise. *Eur J Appl Physiol, 116*(2), 353-362. doi: 10.1007/s00421-015-3290-x

Coelho, D. B., Morandi, R. F., de Melo, M. A. A., & Silami-Garcia, E. (2011). Creatine kinase kinetics in professional soccer players during a competitive season. *Brazilian Journal of Kineanthropometry and Human Performance, 13*(3), 189.

Conceicao, M. S., Libardi, C. A., Nogueira, F. R. D., Bonganha, V., Gaspari, A. F., Chacon-Mikahil, M. P. T., . . . Madruga, V. A. (2012). Effects of eccentric exercise on systemic concentrations of pro- and anti-inflammatory cytokines and prostaglandin (E2): comparison between young and postmenopausal women. *European Journal of Applied Physiology, 112*(9), 3205-3213. doi: 10.1007/s00421-011-2292-6.

Dill, D. B., & Costill, D. L. (1974). Calculation of percentage changes in volumes of blood, plasma, and red-cells in dehydration. *Journal of Applied Physiology, 37*(2), 247-248.

Drust, B., Waterhouse, J., Atkinson, G., Edwards, B., & Reilly, T. (2005). Circadian rhythms in sports performance - An update. *Chronobiology International, 22*(1), 21-44. doi: 10.1081/cbi-200041039

Fischer, C. P. (2006). Interleukin-6 in acute exercise and training: what is the biological relevance? *Exerc Immunol Rev, 12*, 6-33.

Fullagar, H., Skorski, S., Duffield, R., & Meyer, T. (2016). The effect of an acute sleep hygiene strategy following a late-night soccer match on recovery of players. *Chronobiol Int, 33*(5), 490-505. doi: 10.3109/07420528.2016.1149190

Gill, S. K., Hankey, J., Wright, A., Marczak, S., Hemming, K., Allerton, D. M., . . . Costa, R. J. (2015). The Impact of a 24-h Ultra-Marathon on Circulatory Endotoxin and Cytokine Profile. *Int J Sports Med, 36*(8), 688-695. doi: 10.1055/s-0034-1398535

Grafen, G., & Hails, R. (2002) *Modern Statistics for the Life Sciences* (pp. 153 - 184). New York, USA: Oxford University Press.

Harper, L. D., Hunter, R., Parker, P., Goodall, S., Thomas, K., Howatson, G., . . . Russell, M. (2016). Test-Retest Reliability of Physiological and Performance Responses to 120 Minutes of Simulated Soccer Match Play. *J Strength Cond Res, 30*(11), 3178-3186. doi: 10.1519/jsc.0000000000001400

Hopkins, W. G. (2000). Measures of reliability in sports medicine and science. *Sports Medicine, 30*(1), 1-15. doi: 10.2165/00007256-200030010-00001

Horder, M., Jorgensen, P. J., Hafkenscheid, J. C. M., Carstensen, C. A., Bachmann, C., Bauer, K., . . . Vogt, W. (1991). Creatine kinase determination - a european evaluation of the creatine kinase determination in serum, plasma and whole blood with the Reflotron system *European Journal of Clinical Chemistry and Clinical Biochemistry, 29*(10), 691-696.

Horska, A., Fishbein, K. W., Fleg, J. L., & Spencer, R. G. S. (2000). The relationship between creatine kinase kinetics and exercise intensity in human forearm is unchanged by age. *American Journal of Physiology-Endocrinology and Metabolism, 279*(2), E333-E339.

Howatson, G., Goodall, S., & van Someren, K. A. (2009). The influence of cold water immersions on adaptation following a single bout of damaging exercise. *Eur J Appl Physiol, 105*(4), 615-621. doi: 10.1007/s00421-008-0941-1

Howatson, G., Hoad, M., Goodall, S., Tallent, J., Bell, P. G., & French, D. N. (2012). Exercise-induced muscle damage is reduced in resistance-trained males by branched chain amino acids: a randomised, double-blind, placebo controlled study. *Journal of International Society of Sports Nutrition, 12*(9), 20. doi: 10.1186/1550-2783-9-20

Howatson, G., McHugh, M. P., Hill, J. A., Brouner, J., Jewell, A. P., van Someren, K. A., . . . Howatson, S. A. (2010). Influence of tart cherry juice on indices of recovery following marathon running. *Scandinavian Journal of Medicine & Science in Sports, 20*(6), 843-852. doi: 10.1111/j.1600-0838.2009.01005.x

Karakas, M., Baumert, J., Greven, S., Ruckerl, R., Peters, A., & Koenig, W. (2010). Reproducibility in Serial C-Reactive Protein and Interleukin-6 Measurements in Post-Myocardial Infarction Patients: Results from the AIRGENE Study. *Clinical Chemistry, 56*(5), 861-864. doi: 10.1373/clinchem.2010.143719

Knudsen, L. S., Christensen, I. B. J., Lottenburger, T., Svendsen, M. N., Nielsen, H. J., Nielsen, L., . . . Johansen, J. S. (2008). Pre-analytical and biological variability in circulating interleukin 6 in healthy subjects and patients with rheumatoid arthritis. *Biomarkers, 13*(1), 59-78. doi: 10.1080/13547500701615017

Kraemer, W. J., Looney, D. P., Martin, G. J., Ratamess, N. A., Vingren, J. L., French, D. N., . . . Fleck, S. J. (2013). Changes in creatine kinase and cortisol in national collegiate athletic association division i american football players during a season. *Journal of Strength and Conditioning Research, 27*(2), 434-441. doi: 10.1519/JSC.0b013e318281d1b0

Martin, S. A., Pence, B. D., & Woods, J. A. (2009). Exercise and Respiratory Tract Viral Infections. *Exercise and sport sciences reviews, 37*(4), 157-164. doi: 10.1097/JES.0b013e3181b7b57b

McLeay, Y., Stannard, S. R., Mundel, T., Foskett, A., & Barnes, M. (2016). Effect of Alcohol Consumption on Recovery From Eccentric Exercise Induced Muscle Damage in Females. *Int J Sport Nutr Exerc Metab*, 1-20. doi: 10.1123/ijsnem.2016-0171

McLellan, C. P., Lovell, D. I., & Gass, G. C. (2010). Creatine kinase and endocrine responses of elite players pre, during, and post rugby league match play. *J Strength Cond Res, 24*(11), 2908-2919. doi: 10.1519/JSC.0b013e3181c1fcb1

Meister, S., Aus der Funten, K., & Meyer, T. (2014). Repeated monitoring of blood parameters for evaluating strain and overload in elite football players: is it justified? *J Sports Sci, 32*(13), 1328-1331. doi: 10.1080/02640414.2014.927070

Mougios, V. (2007). Reference intervals for serum creatine kinase in athletes. *British Journal of Sports Medicine, 41*(10), 674-678. doi: 10.1136/bjsm.2006.034041

Navarro, S. L., Brasky, T. M., Schwarz, Y., Song, X., Wang, C. Y., Kristal, A. R., . . . Lampe, J. W. (2012). Reliability of serum biomarkers of inflammation from repeated measures in healthy individuals. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology, 21*(7), 1167-1170. doi: 10.1158/1055-9965

Nicholson, G. A., Morgan, G., Meerkin, M., Strauss, E., & McLeod, J. G. (1985). The creatine-kinase reference interval - an assessment of intraindividual and interindividual variation. *Journal of the Neurological Sciences, 71*(2-3), 225-231. doi: 10.1016/0022-510x(85)90061-9

Nielsen, H. G., Oktedalen, O., Opstad, P. K., & Lyberg, T. (2016). Plasma Cytokine Profiles in Long-Term Strenuous Exercise. *J Sports Med (Hindawi Publ Corp), 2016*, 7186137. doi: 10.1155/2016/7186137

Nieman, D. C., Henson, D. A., Dumke, C. L., McAnulty, S. R., McAnulty, L. S., Gross, S. J., & Lind, R. H. (2005). Muscle Damage Is Linked To Cytokine Changes Following A 160-km Race. *Medicine and Science in Sports and Exercise, 37*, S336-S336. doi: 10.1097/00005768-200505001-01740

Niemela, M., Kangastupa, P., Niemela, O., Bloigu, R., & Juvonen, T. (2016). Acute Changes in Inflammatory Biomarker Levels in Recreational Runners Participating in a Marathon or Half-Marathon. *Sports Med Open, 2*(1), 21. doi: 10.1186/s40798-016-0045-0

Owen, A., Dunlop, G., Rouissi, M., Chtara, M., Paul, D., Zouhal, H., & Wong del, P. (2015). The relationship between lower-limb strength and match-related muscle damage in

elite level professional European soccer players. *J Sports Sci, 33*(20), 2100-2105. doi: 10.1080/02640414.2015.1064155

Paulsen, G., Mikkelsen, U. R., Raastad, T., & Peake, J. M. (2012). Leucocytes, cytokines and satellite cells: what role do they play in muscle damage and regeneration following eccentric exercise? *Exercise Immunology Review, 18*, 42-97.

Peake, J. M., Nosaka, K., Muthalib, M., & Suzuki, K. (2006). Systemic inflammatory responses to maximal versus submaximal lengthening contractions of the elbow flexors. *Exercise Immunology Review, 12*, 72-85.

Pemberton, P. W., Aboutwerat, A., Smith, A., & Warnes, T. W. (2006). Ursodeoxycholic acid in primary biliary cirrhosis improves glutathione status but fails to reduce lipid peroxidation. *Redox Report, 11*(3), 117-123. doi: 10.1179/135100006x116600

Platz, E. A., Sutcliffe, S., De Marzo, A. M., Drake, C. G., Rifai, N., Hsing, A. W., . . . Kristal, A. R. (2010). Intra-individual variation in serum C-reactive protein over four years: An implication for epidemiologic studies. *Cancer Causes Control, 21*(6), 847-851. doi: 10.1007/s10552-010-9511-z

Prelle, A., Tancredi, L., Sciacco, M., Chiveri, L., Comi, G. P., Battistel, A., . . . Moggio, M. (2002). Retrospective study of a large population of patients with asymptomatic or minimally symptomatic raised serum creatine kinase levels. *J Neurol, 249*(3), 305-311.

Robson-Ansley, P., Barwood, M., Canavan, J., Hack, S., Eglin, C., Davey, S., . . . Ansley, L. (2009). The effect of repeated endurance exercise on IL-6 and sIL-6R and their relationship with sensations of fatigue at rest. *Cytokine, 45*(2), 111-116. doi: 10.1016/j.cyto.2008.11.006

Romagnoli, M., Sanchis-Gomar, F., Alis, R., Risso-Ballester, J., Bosio, A., Graziani, R. L., & Rampinini, E. (2016). Changes in muscle damage, inflammation, and fatigue-related parameters in young elite soccer players after a match. *J Sports Med Phys Fitness, 56*(10), 1198-1205.

Salvadori, A., Fanari, P., Ruga, S., Brunani, A., & Longhini, E. (1992). Creatine-kinase and creatine kinase-mb isoenzyme during and after exercise testing in normal and obese young-people. *Chest, 102*(6), 1687-1689. doi: 10.1378/chest.102.6.1687

Scherr, J., Braun, S., Schuster, T., Hartmann, C., Moehlenkamp, S., Wolfarth, B., . . . Halle, M. (2011). 72-h Kinetics of High-Sensitive Troponin T and Inflammatory Markers after Marathon. *Medicine and Science in Sports and Exercise, 43*(10), 1819-1827. doi: 10.1249/MSS.0b013e31821b12eb

Tseng, K. W., Tseng, W. C., Lin, M. J., Chen, H. L., Nosaka, K., & Chen, T. C. (2016). Protective effect by maximal isometric contractions against maximal eccentric exercise-induced muscle damage of the knee extensors. *Res Sports Med, 24*(3), 243-256. doi: 10.1080/15438627.2016.1202826

Vieira, A., Siqueira, A. F., Ferreira-Junior, J. B., do Carmo, J., Durigan, J. L., Blazevich, A., & Bottaro, M. (2016). The Effect of Water Temperature during Cold-Water Immersion on Recovery from Exercise-Induced Muscle Damage. *Int J Sports Med, 37*(12), 937-943. doi: 10.1055/s-0042-111438

Vincent, H. K., & Vincent, K. R. (1997). The effect of training status on the serum creatine kinase response, soreness and muscle function following resistance exercise. *International Journal of Sports Medicine, 18*(6), 431-437.

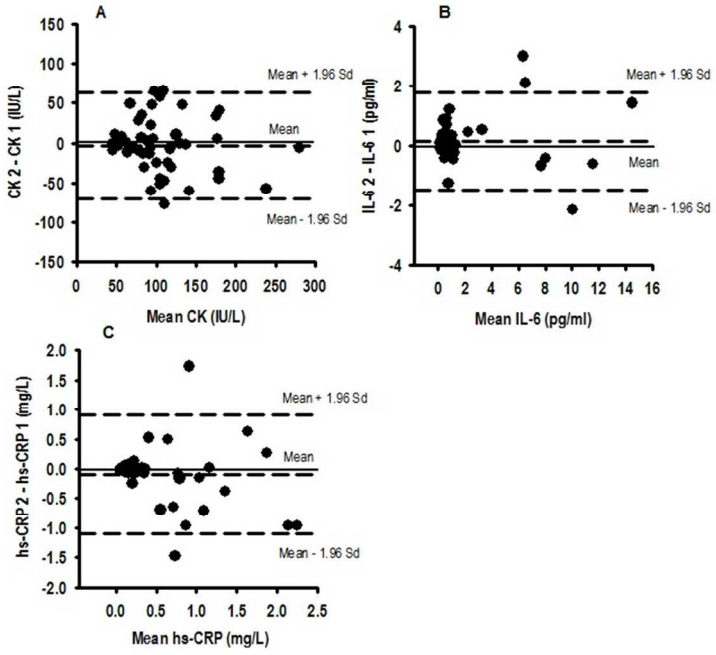
Walshe, I., Robson-Ansley, P., Gibson, A. S. C., Lawrence, C., Thompson, K. G., & Ansley, L. (2010). The reliability of the IL-6, sIL-6R and sgp130 response to a preloaded time trial. *European Journal of Applied Physiology, 110*(3), 619-625. doi: 10.1007/s00421-010-1548-x

Young, W. B., Hepner, J., & Robbins, D. W. (2012). Movement demands in Australian rules football as indicators of muscle damage. *J Strength Cond Res, 26*(2), 492-496. doi: 10.1519/JSC.0b013e318225a1c4

Figure caption

Figure 1. Bland-Altman plots showing individual differences between the absolute values plotted against their individual means (n = 50) for **A)** CK = creatine kinase, **B)** IL-6 = interleukin-6 and **C)** hs-CRP = high sensitivity C-reactive protein. Horizontal dashed lines = 95% limits of agreement; solid line = zero reference line; S_d = within-subject standard deviation

Reproducibility of measurement techniques for blood biomarker determination



254x190mm (300 x 300 DPI)

Table 1. Circulating blood biomarker values obtained during trial 1 and trial 2 (n = 50)

Measure	Trial 1			Trial 2			Trial 1 – Trial 2 differences			
	Mean	SD	Range	Mean	SD	Range	Mean diff	95% CI	S_d	p value
CK (IU/L)	109.8	40.4	42.6 – 283.0	108.7	42.7	41.3 – 210.0	-1.0	-7.9, 10.0	31.6	0.82
IL-6 (pg/ml)	3.96	6.60	0.02 – 33.02	4.18	7.25	0.10 – 31.80	0.21	-0.31, 0.74	1.24	0.41
hs-CRP (mg/L)	0.63	0.70	0.04 – 2.73	0.53	0.58	0.02 – 2.00	-0.10	-0.26, 0.07	0.51	0.24

Any discrepancies between the mean diff and means is due to rounding error; SD = between-subject standard deviation; mean diff = mean difference, 95% CI = lower and upper bounds of the 95% confidence interval for the mean difference, S_d = within-subject standard deviation CK = creatine kinase; IL-6 = interleukin 6, hs-CRP = high sensitivity C reactive protein

Table 2. Reproducibility statistics for the circulating blood biomarkers (n = 50)

Measure	CV (%)	ICC	95% LoA
CK (IU/L)	20.0	0.90	-69.7, 63.5
IL-6 (pg/ml)	15.3	0.98	-1.48, 1.80
hs-CRP (mg/L)	44.2	0.70	-1.10, 0.91

See footnotes of Table 1 for an explanation of the abbreviations; CV = coefficient of variation for repeated measures; ICC = intraclass correlation coefficient (two-way random model for a single rater); 95% LoA = 95% limits of agreement.