

RESEARCH ARTICLE

Reliability and agreement of the IsoKai isokinetic lift test – A test used for admission to the Swedish Armed Forces

Tony Bohman^{1*}, Matthias Tegern^{1,2}, Alexandra Halvarsson^{1,3}, Lisbet Broman¹, Helena Larsson^{1,4}

1 Department of Neurobiology, Care Sciences and Society, Division of Physiotherapy, Karolinska Institutet, Stockholm, Sweden, **2** Department of Community Medicine and Rehabilitation, Umeå University, Umeå, Sweden, **3** Allied Health Professionals Function, Karolinska University Hospital, Stockholm, Sweden, **4** Swedish Armed Forces, Headquarters, Medical Services, Stockholm, Sweden

* tony.bohman@ki.se



Abstract

This study was performed to evaluate the reliability and agreement of the IsoKai isokinetic lift test as it is currently administered in admission to the Swedish Armed Forces. The study included an intrarater ($n = 534$) and interrater reliability sample ($n = 137$), of Swedish male conscripts who performed the test on two test occasions about two hours apart. Two-to-four lifts were performed at each occasion, and the highest mean (IsoKai_{MF}) and peak force (IsoKai_{PF}) produced (N) were used for evaluation. All intraclass coefficients showed excellent reliability. The interrater analyses resulted in intraclass coefficients of 0.942 (95% CI; 0.920–0.959) and 0.858 (95% CI; 0.806–0.896) for the IsoKai_{MF} and IsoKai_{PF}, respectively, while the corresponding coefficients for the intrarater analyses were 0.935 (95% CI; 0.923–0.946) and 0.865 (95% CI; 0.842–0.886). Agreement, the capability of a test to detect changes, was assessed by the standard error of measurement (SEM/SEM%) and the smallest real difference (SRD/SRD%). These estimate indicated that it is possible to achieve measurements relevant to use in real practice with the IsoKai isokinetic lift test. Bland and Altman analyses revealed no systematic errors in either sample. Based on these findings, the IsoKai isokinetic lift test is suggested to be a highly reliable test for maximal dynamic muscular strength. The test could be of use in selection procedures in order to accurately evaluate maximal dynamic muscular strength, and for evaluating longitudinal changes in strength.

OPEN ACCESS

Citation: Bohman T, Tegern M, Halvarsson A, Broman L, Larsson H (2018) Reliability and agreement of the IsoKai isokinetic lift test – A test used for admission to the Swedish Armed Forces. PLoS ONE 13(12): e0209419. <https://doi.org/10.1371/journal.pone.0209419>

Editor: Riccardo Di Giminiani, University of L'Aquila, ITALY

Received: November 10, 2017

Accepted: December 5, 2018

Published: December 19, 2018

Copyright: © 2018 Bohman et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the manuscript and its Supporting Information files.

Funding: The author(s) received no specific funding for this work.

Competing interests: The authors have declared that no competing interests exist.

Introduction

Measuring muscular strength is often of interest and importance in many areas such as sports, rehabilitation, military settings and research [1–3]. Measures of muscular strength could either be isometric or dynamic [2]. Standard methods of measuring maximal dynamic strength include the isoinertial one-repetition maximum test (1-RM test), using external weight loading, and isokinetic strength testing performed with various commercial devices [2, 4]. In an isokinetic test, the movement velocity is held constant while the resistance adapts to the muscle force

produced in every part of the movement. This is in contrast to an isoinertial strength test where the resistance is constant throughout the full movement [4, 5]. An isokinetic test is considered to be a safer alternative to an isoinertial test since the risk of overload is lower due to the adaptation of the resistance to the muscle force [2, 5]. Other benefits of isokinetic tests are the fast and easy administration of tests and their often good reliability [2, 5]. The purpose of a strength test could be to define individual physical fitness level, distinguish individuals with regard to strength capacity, or to evaluate change in strength over time in individuals or in groups [1–3]. This could, however, only be accomplished if the measures of strength are valid and reliable. The validity of a test concerns the degree to which it measures what it is intended to measure [6]. To be reliable, tests need to have a high degree of reliability and agreement, properties that set the standard for validity, and provide quantifications of the test measurement [6, 7]. Moreover, the term reliability refers to the ability of a test to differentiate between individuals, while agreement is a measure of the capability to detect changes [6, 7]. Reliability and agreement of a measure could be assessed using an inter- as well as intrarater reliability study design [6, 7]. Interrater reliability evaluates the consistency of test results at two test occasions administered by two different raters, while intrarater reliability evaluates the consistency between two test occasions administered by the same rater [6, 7]. In some physically demanding occupations, muscular strength tests are used to evaluate work capacity in connection to admission for certain programs and for job selection during employment, e.g. firefighters, policemen and soldiers [8–10]. It is important that these strength tests are related to the work tasks commonly performed in the specific occupations in order to minimize the risk for disability, time off-duty and earlier retirement [8, 11]. Preferably, such tests should be safe, fast and easy to administer.

One example of such a test is the IsoKai isokinetic lift test that has, since 1995 been used to assess muscular strength on admission to service in the Swedish Armed Forces (SwAF). The IsoKai lift test was introduced as a safe and better test to predict dynamic strength needed in military work tasks than the former used isometric strength test [12, 13]. The IsoKai isokinetic lift test have excellent content validity for lifting, carrying with hands and digging, which are work tasks commonly performed during military service [14]. The test is performed in a device measuring the isokinetic mean force (IsoKai_{MF}) and peak force (IsoKai_{PF}), with values in Newton (N), during a maximal two-handed lift of a weight-lifting bar from “floor” to shoulder level (Fig 1).

The IsoKai_{MF} measure represents the average force produced during one lift, while the IsoKai_{PF} measure represents the maximum force produced during a lift. The IsoKai_{PF} is produced and registered when the bar reaches about hip level. Since 2013, an admission test for the SwAF and the Swedish Police includes two to four IsoKai lift attempts. The highest IsoKai_{MF} and IsoKai_{PF} registered from the lifts are used as test measures [15]. A study from 2009 showed the IsoKai test to have excellent interrater reliability when using the mean value of the IsoKai_{MF} from two lifts as outcome in a sample of Swedish conscripts [16]. However, the IsoKai test has not been assessed for reliability when using the highest IsoKai_{MF} and IsoKai_{PF} as test measures, i.e. the way it is currently used in the SwAF and the Swedish Police [15]. At this backdrop, the aim of this study was to examine the intra- and interrater reliability and agreement of the IsoKai isokinetic lift test in Swedish male conscripts using the highest registered IsoKai_{MF} and IsoKai_{PF} as measures.

Material and methods

Study design and study sample

The study was designed to include both an intra- and interrater reliability evaluation using two different samples. The data used to assess the intrarater reliability part of the study were collected from IsoKai isokinetic tests on male conscripts at the time of entering military service



Fig 1. The IsoKai isokinetic lift test.

<https://doi.org/10.1371/journal.pone.0209419.g001>

in 2002. Out of 578 eligible for test, 44 were absent due to sick-leave or other assignments, resulting in a sample of 534 conscripts. Data for the interrater reliability part were collected from another sample of 137 male conscripts, randomly selected from 601 eligible conscript at the end of their 10 months military service in 2001. At the time of data collection, the Swedish military system was based on compulsory military service for males. To be included in the study, conscripts had to be healthy and without any pain which could influence the test procedure. The IsoKai lift tests were administered by personnel from the Swedish Defence Recruitment Agency, which were well-educated in test procedures, and all had several years of experience in physical testing. Participants signed a written informed consent after receiving written information. The individual in this manuscript (Fig 1) has given written informed consent, as outlined in PLOS consent form, to publish these case details. The study was approved by the regional ethic committee in Stockholm, Gothenburg and Orebro, Sweden, Dnr 500:16 307/01. Table 1 presents the characteristics of the two study samples.

Data collection and test procedure

In the intrarater part of the study, the same rater administered the test on every participant on two occasions. In the interrater part, two raters, independent of each other, administered the test on every participant once only. In both the intra- and interrater part, all participants conducted the two tests occasions on the same day with about two hours between tests. Raters and participants were blinded to the result from the first test when the second test was conducted.

The IsoKai lift test procedure. IsoKai is a device used for measuring isokinetic muscular performance during a vertical lifting procedure. The device consists of a frame holding a

Table 1. Characteristics of the intrarater and interrater study samples.

	Intrarater reliability (n = 534)		Interrater reliability (n = 137)	
	Mean	SD	Mean	SD
Age (years)	19	0.6	19	0.4
Height (m)	1.81	0.07	1.81	0.06
Weight (kg)	77	14	78	10
BMI (kg/m ²)	23.6	3.7	23.7	2.7

SD, standard deviation; BMI, body mass index.

<https://doi.org/10.1371/journal.pone.0209419.t001>

vertical lifting bar that, via two wires, is connected to a hydraulic system regulating the speed of the lift at 0.30 m/sec. The IsoKai device allows a small horizontal movement (6–7 cm) of the vertical bar during the lift. The muscular force (N) is, for each cm of the lift, registered by a computer connected to a force plate on which the participant was instructed to stand with feet separated by shoulder width. The test was carried out from a start position where the back was straight while being forwardly inclined and knees bent. Thereafter, a two-handed lift of a vertical bar from 30 cm above the force plate was carried out until in upright standing with the bar at shoulder level (Fig 1). The participants were instructed to use maximal effort during the lift. No shoes were allowed during test. Before the test, participants performed a 10 minutes warm-up session using a cycle ergometer. Length was registered by an electronic measure instrument connected to the IsoKai device, and weight was registered by the force plate. This was done in order to adapt the IsoKai device calculations to each participant. The rater then gave oral instructions on how to execute the lift, and, in addition, demonstrated the lift to secure a safe test. One sub-maximal practice test lift was allowed for the participant to get used to the testing procedure. Each test occasion consisted of two to four maximal lifts. For each lift, a graph in which the force was plotted against each cm of the lift was generated (Fig 2).

The IsoKai mean force (IsoKai_{MF}) and peak force (IsoKai_{PF}) measures generated at each lift were registered. The IsoKai_{MF} represents the mean of all the registered forces (plots in the force curve) during one lift, while the IsoKai_{PF} is the maximum force produced during the lift (the top of the force curve). Only the highest IsoKai_{MF} and IsoKai_{PF} measures produced during the

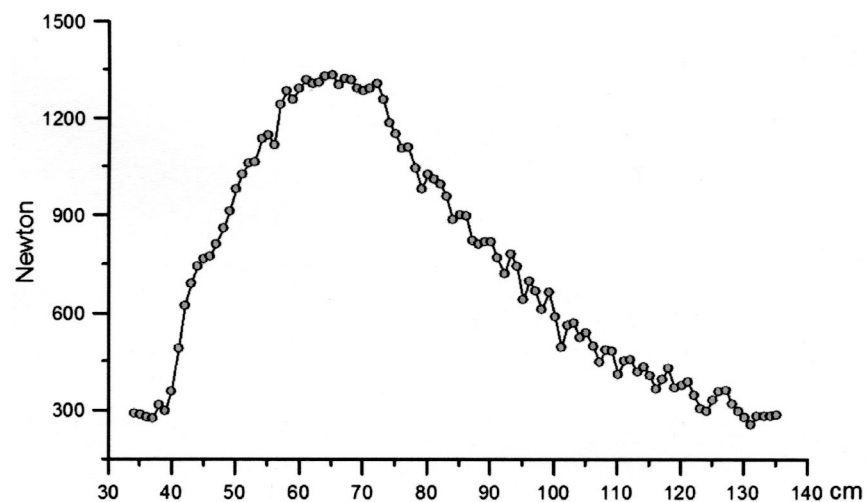


Fig 2. The force curve generated from an IsoKai isokinetic lift test.

<https://doi.org/10.1371/journal.pone.0209419.g002>

test occasion were used as estimates in the analyses. Prior to the test occasion, the IsoKai device was calibrated by experienced personnel from the Swedish Defence Recruitment Agency.

Statistical analyses

Descriptive data were presented with means and standard deviations (SD). The level of significance was set to 95%. Statistical analyses were performed using IBM SPSS Statistics version 23 (IBM Corporation, USA). See the [S1 Appendix](#) for the ANOVA results, and some example ICC calculations.

Reliability. The intraclass correlation coefficient (ICC) was calculated together with its corresponding 95% confidence interval (95% CI) [17–19]. The ICC_{1,1} equation was used for the intrarater reliability calculations, and the ICC_{2,1} equation for the interrater reliability calculations [17–19].

$$ICC_{1,1} = (BMS - WMS) / (BMS + WMS)$$

$$ICC_{2,1} = (BMS - EMS) / (BMS + (k - 1)EMS + \frac{k}{n}(JMS - EMS))$$

The equation variables were estimated using repeated measures analysis of variance (ANOVA). BMS represents the between-people variability of the measurement, WMS the within-people variability, JMS the between-rater variability, EMS the residual mean square (error) variability, *k* the number of raters, and *n* the number of participants ([S1 Appendix](#)) [17–19]. An ICC coefficient of 1.0 indicates perfect reliability of a measurement while zero indicates a totally unreliable measurement [6]. Evaluation of the ICC coefficients was interpreted using the classification suggested by Cicchetti, with reliability coefficients categorised as: 0.75 to 1.0 = excellent, 0.60 to 0.74 = good, 0.40 to 0.59 = fair and less than 0.40 = poor [20].

Agreement. Agreement was assessed using the standard error of measurement (SEM) and the smallest real difference (SRD), along with their 95% CI [21–23]. The SRD shows the limit for the smallest difference (N) that indicates a real change for an individual. Further, the SEM% and SRD% were calculated to be able to evaluate the measurement error and individual changes independently from the units of measurement (N) [23]. SEM% and SRD% are recommended to be used for comparison with other studies, especially to account for the fact that different units of measurement are used [23]. Formulas used to assess agreement:

$$SEM = \sqrt{WMS}$$

$$SRD = \sqrt{2} \times 1.96 \times SEM$$

SRD 95% CI = $\bar{d} \pm SRD$, where \bar{d} represent the overall mean difference between two tests.

$$SEM\% = (SEM / Grand\ mean) \times 100$$

$$SRD\% = (SRD / Grand\ mean) \times 100$$

Grand mean = (the sample mean force of the IsoKai_{MF} or IsoKai_{PF} from test occasion 1 + the sample mean force of the IsoKai_{MF} or IsoKai_{PF} from test occasion 2) / 2.

Bland and Altman methods. Bland and Altman methods were used to assess the dispersion of data, measurement error and possible systematic bias of the test [7, 24]. These methods were based on the analysis of differences between measurements from the repeated test occasions. Using Bland and Altman plots, the mean values from two tests were plotted against the difference between the tests for each participant. The Bland and Altman plots visualised the

distribution of the individual test differences around the overall mean difference between tests (\bar{d}) together with the limits of agreement (LOA).

The LOA which represented the precision of the measurement was calculated by the formula: $\bar{d} \pm 2 \text{ standard deviations (SD)}$ [24]. Further, as an overall mean difference between the tests occasions (\bar{d}) significantly different from zero indicates a systematic bias in the test, a possible bias was formally assessed by estimating \bar{d} together with the 95% CI [18, 24, 25]. The 95% CI of \bar{d} were calculated using the formula; $\bar{d} \pm t_{n-1}$, where t_{n-1} represents the probability point of the t distribution on n-1 degrees of freedom, and SE the standard error of \bar{d} [18].

Results

Table 2 presents the sample mean force and the “Grand mean” for the IsoKai_{MF} and IsoKai_{PF} measures in the intra- and interrater samples respectively.

Reliability

All results showed that the IsoKai isokinetic lift test had excellent reliability, with ICC coefficients ranging from 0.858 to 0.942 (Table 3).

Agreement

The agreement estimates based on the ANOVA analyses (SEM, SEM%, SRD and SRD%) indicated a higher degree of agreement for the IsoKai_{MF} than IsoKai_{PF}, and a higher degree of agreement for the interrater sample compared to the intrarater sample (Table 4).

Bland and Altman methods

Bland and Altman results for the intra- and interrater samples are presented in Table 5 and Figs 3 and 4. The overall mean difference between tests (\bar{d}) in the intrarater sample was significantly different from zero, with -6 N (95% CI; -10 to -3) for the IsoKai_{MF} and -14 N (95% CI; -23 to -4) for the IsoKai_{PF}.

Discussion

This study revealed that the IsoKai isokinetic lift test is a highly reliable test for evaluating maximal dynamic muscular strength related to lifting. To establish if a measurement gives reliable information, assessing both measures of reliability and agreement is recommended [6, 7, 22]. Our analyses demonstrated that the IsoKai isokinetic lift test had excellent reliability for measuring maximal strength irrespective of whether the tests were performed by the same rater (intrarater reliability) or by two different raters (interrater reliability). Overall, the ICC

Table 2. The sample mean force for the IsoKai_{MF} and IsoKai_{PF} at the different test occasions and the “Grand mean” in the intra- and interrater samples.

	Intrarater reliability (n = 534)			Interrater reliability (n = 137)		
	Test 1	Test 2	G-mean (SD)	Rater 1	Rater 2	G-mean (SD)
	Mean (SD)	Mean (SD)		Mean (SD)	Mean (SD)	
IsoKai _{MF} (N)	711 (114)	717 (111)	714 (111)	757 (103)	752 (104)	754 (102)
IsoKai _{PF} (N)	1246 (216)	1260 (211)	1253 (206)	1368 (194)	1356 (192)	1362 (186)

IsoKai_{MF}, IsoKai mean force; IsoKai_{PF}, IsoKai peak force; N, Newton; G-mean = Grand mean = (sample mean force of the IsoKai_{MF}/IsoKai_{PF} from test 1/rater 1 + sample mean force of the IsoKai_{MF}/IsoKai_{PF} from test 2/rater 2) / 2.

<https://doi.org/10.1371/journal.pone.0209419.t002>

Table 3. Intraclass correlation estimates (ICC) for the IsoKai_{MF} and IsoKai_{PF} measures of the IsoKai isokinetic lift test for the intra- and interrater reliability analyses.

	Intrarater reliability (n = 534)		Interrater reliability (n = 137)	
	ICC _{1,1}	95% CI	ICC _{2,1}	95% CI
IsoKai _{MF}	0.935	0.923 to 0.946	0.942	0.920 to 0.959
IsoKai _{PF}	0.865	0.842 to 0.886	0.858	0.806 to 0.896

IsoKai_{MF}, IsoKai mean force; IsoKai_{PF}, IsoKai peak force; ICC_{1,1} and ICC_{2,1}, intraclass correlation coefficient; 95% CI, 95% confidence interval.

<https://doi.org/10.1371/journal.pone.0209419.t003>

coefficients showed that the test is an excellent test to use in order to rank male conscripts with regard to maximal muscular strength needed for lifting and similar tasks. Further, the IsoKai test was found to be an excellent test to distinguish between individuals assessing maximal strength when lifting from “floor” level to hip level (IsoKai_{PF}; ICC_{1,1} = 0.865, ICC_{2,1} = 0.858), as well as from “floor” level to shoulder level (IsoKai_{MF}; ICC_{1,1} = 0.935, ICC_{2,1} = 0.942). Even if only the IsoKai_{MF} is tested for content validity, we believe that the IsoKai_{PF} evaluates mainly low body and back strength, while the IsoKai_{MF} is a measure of whole body strength [14]. All the lower limits of the ICC confidence intervals exceeded 0.8 which further support the excellent reliability of the IsoKai isokinetic lift test.

There are no general recommendations set for how to evaluate agreement estimates as the evaluation depends on the context in which the test is used [7]. Therefore, the results in the present study could be evaluated in the light of finding potential changes in maximal muscular strength due to training or de-training [26]. The SEM and SEM% represent the error of the measurement. The SEM in the two samples were 29 and 25 N for the IsoKai_{MF}, and 78 and 73 for the IsoKai_{PF}. If taking the SEM of 29 N IsoKai_{MF} in the interrater sample as an example, this corresponds to a measurement error of only 3 kg (29N*0.102). We believe this supports the IsoKai test to be a rather precise measurement of changes in dynamic muscular strength. The precision of measuring individual changes represented by the SRD and SRD% was less pronounced, with SRDs between 69 to 217 N (Table 4), or 7 to 22 kg.

The Bland and Altman analyses showed no major signs of systematic errors in either the intrarater sample or the interrater sample. The plotted differences between two tests were evenly distributed about the overall mean difference (\bar{d}) in all analyses. In the intrarrater sample a statistically significant difference from zero was found specifically for the mean difference between tests, with values of -6 N and -14 N for both the IsoKai_{MF} and IsoKai_{PF}, respectively (Table 5). This corresponded to about 1% of the “Grand mean” force values; a minor

Table 4. Agreement estimates for the IsoKai_{MF} and IsoKai_{PF} of the IsoKai isokinetic lift test for the intra- and interrater agreement part of the study.

	Intrarater agreement (n = 534)					Interrater agreement (n = 137)				
	G-mean N	SEM N	SEM% %	SRD (95% CI) N	SRD% %	G-mean N	SEM N	SEM% %	SRD (95% CI) N	SRD% %
IsoKai _{MF}	714	29	4.0	79 (73 to 86)	11.1	754	25	3.3	69 (64 to 74)	9.2
IsoKai _{PF}	1253	78	6.2	217 (203 to 231)	17.3	1362	73	5.4	202 (191 to 214)	14.8

IsoKai_{MF}, IsoKai mean force; IsoKai_{PF}, IsoKai peak force; N, Newton; G-mean = Grand mean = (sample mean force of the IsoKai_{MF}/IsoKai_{PF} from test 1/rater 1 + sample mean force of the IsoKai_{MF}/IsoKai_{PF} from test 2/rater 2) / 2; SEM, standard error of measurement; SEM%, (SEM/Grand mean) x 100; SRD, smallest real difference; 95% CI, 95% confidence interval; SRD%, (SRD/Grand mean) x 100.

<https://doi.org/10.1371/journal.pone.0209419.t004>

Table 5. Bland and Altman measure for the intra- and interrater sample IsoKai_{MF} and IsoKai_{PF}.

	Intrarater sample (n = 534)				Interrater sample (n = 137)			
	\bar{d}	95% CI	SE	SD	\bar{d}	95% CI	SE	SD
IsoKai _{MF} (N)	-6	-10 to -3	1.7	40	5	-1 to 11	3.0	35
IsoKai _{PF} (N)	-14	-23 to -4	4.8	110	12	-6 to 29	8.8	103

IsoKai_{MF}, IsoKai mean force; IsoKai_{PF}, IsoKai peak force; N, Newton; \bar{d} , overall mean difference between test/rater; 95% CI, 95% confidence interval; SE, standard error; SD, standard deviation.

<https://doi.org/10.1371/journal.pone.0209419.t005>

systematic error which could be negligible in this case. Concerning the interrater sample, about the same differences from zero as in the intrarater sample were found for the overall mean difference (\bar{d}), however, these were not statistically significant.

Comparison with other studies

Isokinetic strength measures have been criticised as they mostly engage only one or two joints and their associated muscle groups which bear little relationship to the multi-joint and multi-muscle actions that take place during movements in real practice [4]. As such, the IsoKai lift isokinetic lift test is unique among isokinetic tests as it engages almost all muscle groups and joints in the body in a movement imitating a normal lifting procedure. As mentioned in the introduction, Larsson et al. (2009) assessed the reliability of the IsoKai test as it was performed in the SwAF before 2013 [16]. By using the mean value of two IsoKai_{MF} registrations as a measure of muscle strength capacity in a sample of Swedish conscripts (n = 427), they found an ICC_{3,1} of 0.94, a SEM of 30 and a SEM% of 4.3, findings very similar to ours. The IsoKai_{PF} was not evaluated. In addition to this study, we found two reliability evaluations of multi-joint isokinetic tests similar to the IsoKai isokinetic lift test [27, 28]. Bridgeman et al. found an isokinetic multi-joint squat devise to have good test-retest reliability, which was evaluated within 3 test sessions over a 3-week period in a sample of 10 strength trained male athletes. The concentric peak force (N) outcome revealed ICCs ranging from 0.87 to 0.98, and coefficients of variation (CV) from 7.6 to 15.4. The CV was calculated by dividing the method error (ME) by the

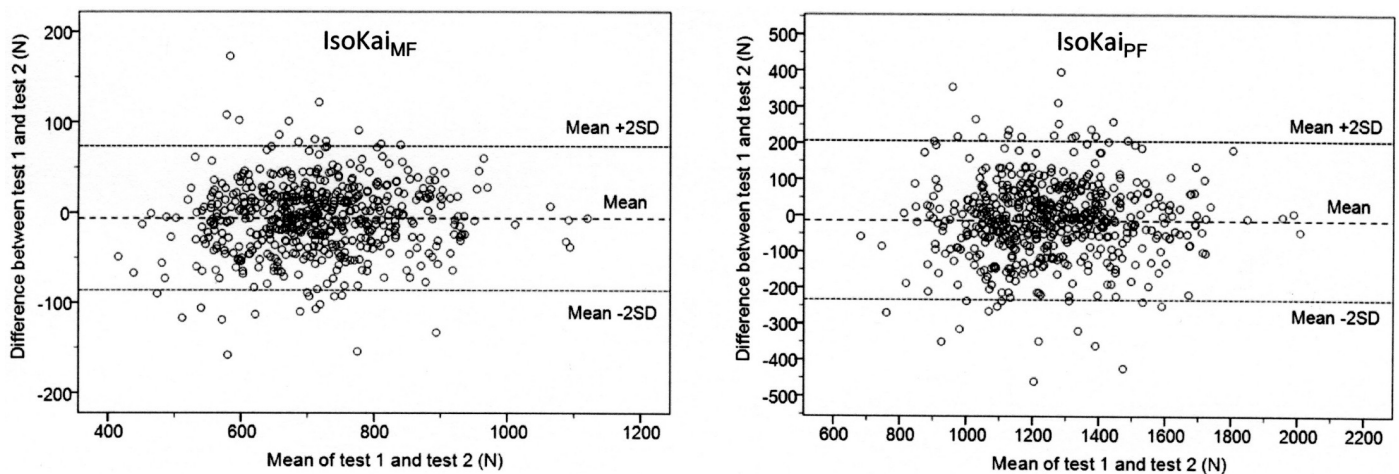


Fig 3. The mean of test 1 force and test 2 force plotted against the difference between the test 1 force and test 2 force in the intrarater sample. IsoKai_{MF}, IsoKai mean force (N); IsoKai_{PF}, IsoKai peak force (N); LOA, limits of agreement.

<https://doi.org/10.1371/journal.pone.0209419.g003>

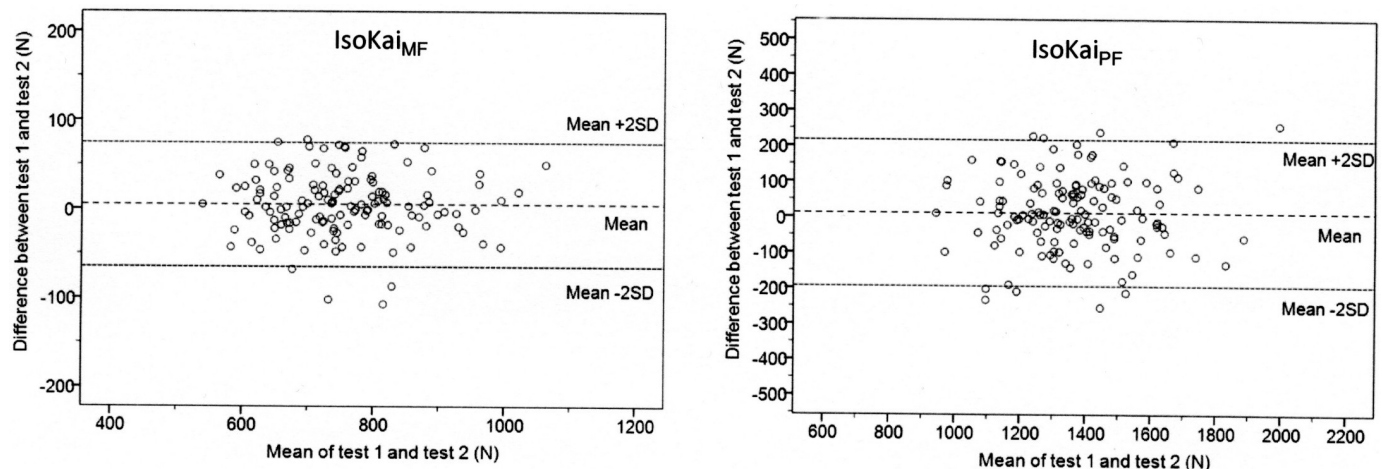


Fig 4. The mean of test 1 force (rater 1) and test 2 force (rater 2) plotted against the difference between test 1 force (rater 1) and test 2 force (rater 2) in the interrater sample. IsoKai_{MF}, IsoKai mean force (N); IsoKai_{PF}, IsoKai peak force (N); LOA, limits of agreement.

<https://doi.org/10.1371/journal.pone.0209419.g004>

overall mean difference of two tests (\bar{d}) and then multiply by 100 [23]. Lexell and Downham states that “if the sample size is sufficiently large and the mean difference small, both highly likely conditions, ME and SEM take similar values”, indicating that CV and SEM% could be comparable between studies [23]. In 1997, Wilson et al. examined another isokinetic squat device by measuring concentric peak force (N) in 29 athletic male subjects performing two tests with 3 minutes rest in-between, and found an ICC of 0.89 and a CV of 8.7 [27]. The IsoKai_{PF} measures in the current study estimated a muscular force comparable with the outcome in the above mentioned squat tests, and resulted in ICCs of the same magnitude (0.858 and 0.865, Table 3). The IsoKai_{PK} SEM% of 5.4 and 6.2 indicate better precision of the IsoKai lift test compared to the CVs from the squat tests evaluations; however, such a comparison should be done with caution since the small sample sizes might have influenced the results.

Methodological considerations

Today many women apply for enrolment in physically demanding occupations such as in the police force and military service. As such, the lack of women in our samples could be regarded as a limitation. However, unpublished data from the SwAF indicates excellent ICC values for women ($n = 18$) regarding the IsoKai_{MF} (ICC_{1,1} and ICC_{2,1} of 0.811) and IsoKai_{PF} (ICC_{1,1} and ICC_{2,1} of 0.767). Given these preliminary results, we find no reason to believe that the IsoKai isokinetic lift test would not be a reliable test among women. The data in the present study might be regarded as old since it dates back to 2001 and 2002. Still, as the IsoKai device used at time of data collection was exactly the same type as used today, we believe that the results are still applicable. There are several suggested ICC equations recommended in the literature depending on the context of the test procedures, but no standardised system on how to choose the most appropriate ICC exists [7, 17, 18]. We decided to use the ICC_{1,1} equation (intrarater reliability) and the ICC_{2,1} equation (interrater reliability) in the analyses, as they corresponded well to recommendations when compared to the test proceedings in which the IsoKai lift test was performed, and how it is used in the Swedish Defence Recruitment Agency [17–19, 29]. A number of arbitrary recommendations exist regarding the interpretation of satisfactory levels of reliability of a test when using the ICC. Resultantly, we chose the recommendations by Cicchetti [20]. Portney and Watkins suggest that an ICC above 0.75 indicate good reliability, while

Currier et al suggest coefficients ranging from 0.80 to 0.90 indicate good reliability and above 0.90 indicate high reliability [6, 30]. No matter which of these recommendations used, the IsoKai isokinetic lift test would be regarded as a test with good-to-high levels of reliability even if the lower limit of the CI is considered, as is recommended [7, 29]. The overall lower ICCs and larger agreement measures for the IsoKai_{PF} compared to the IsoKai_{MF} might be explained by the fact that the IsoKai_{PF} is registered momentarily while the IsoKai_{MF} is averaged during the completed lift. Therefore, the measurement error inherent to the device itself, or differences in the momentarily force produced by the participant, might result in a larger WMS and EMS estimates for the IsoKai_{PF} than the IsoKai_{MF}, causing these differences [22]. This might also affect the relationship between the BMS and the WMS/EMS, as reflected in the F statistics of the ANOVA (S1 Appendix). A high F statistics indicates a large discrepancy between the BMS and WMS/EMS, as in the IsoKai_{MF}, which results in high ICC measures, while a lower F statistics, as in the IsoKai_{PK}, results in lower ICC measures [6, 22, 31]. We believe that our sample sizes are considered a strength as it widely exceeds recommendations for reliability studies. Bonett et al. suggests a sample size of 108 subjects in order to detect an ICC of 0.70 or higher with a power of 0.80 and an alpha level set to 0.05 [32]. With the same power and alpha level, Walter et al. estimated a sample of 117 subjects is required to find an expected reliability of 0.80 [33]. Bonnet et al. further emphasise that it is unnecessary to include more subjects than needed for the desired power, as it increases the costs of research, and do not necessarily improve the results [32]. We fully agree, but as the data in our study already were collected as a part of the SwAF admission procedure, we found no reason not to use the entire sample. Another strength is that our data were collected in a way that Carter et al. call a “partially standardised approach” which has the purpose of describing reliability with the levels of standardisation that could be achieved in real settings [31]. We believe this to further supports the findings of the IsoKai lift test as a reliable measurement when implemented into real practice. Finally, we have fulfilled all the requirements in the GRRAS checklist for reporting of studies of reliability and agreement [7].

Implications of the results

Military personnel are often exposed to high physical demands in their work tasks, especially during military missions [34–36]. Common tasks during military service include marching, digging, carrying and lifting, with lifting to be the most frequent work task [11, 34, 37]. In 2015, Larsson et al. found that the IsoKai isokinetic lift test had excellent content validity with respect to digging, carrying and lifting tasks [14]. Hence, to use it as an admission test in military settings and other physically demanding occupations with similar work tasks would be of great benefit. Further, as we evaluated the IsoKai isokinetic test using a representative population in comparison with the target population, the results support the use of the test in military settings as well as other similar settings e.g. the police force. The IsoKai lift test is fast, easy to assess and relatively safe as the resistance in the isokinetic test adapts to muscle force [2, 5]. This makes it an excellent test in selection procedures targeting large groups especially where time constraints exist. The test also has some drawbacks that have to be mentioned. The device is difficult to transport, needs experienced personnel for calibration, and is quite costly. It is important to notice that the reliability estimates are dependent of the variability of the outcome in the population measured, which might limit the external validity [7, 22]. The mean age in our samples was 19 years, which could be argued to restrict our findings to a young population. However, even if our samples are homogenous regarding age, the between subject variability (BMS) of the outcome showed large heterogeneity in all our analyses (S1 Appendix). Therefore, our results could probably be valid even for other age categories as it is not the actual age that matters but

the variance of the measures when evaluating reliability [6]. The IsoKai isokinetic lift test could also be used to assess clinical important changes in muscular strength in individuals following an intervention, for example in sport or sports medicine. Measures of agreement, such as the SEM and the SRD, are not dependent on the variability between subjects, since only the measurement error is of importance [22]. Therefore, measures of agreement could to a greater extent be used in various populations other than reliability measures.

Recommendations for future research

The constant speed and varying resistant in an isokinetic strength test have advantages, as discussed earlier, but could also be a disadvantage in relation to muscle performance in real practice where speed typically varies and load remains constant. This could challenge the validity of the IsoKai isokinetic lift test in comparison to isoinertial lifting capacity. One possible solution to this could be to investigate the criterion or concurrent validity of the IsoKai lift test in relation to maximal strength measured with an isoinertial lift test, for example a one repetition maximum lift test. Moreover, we believe that examining the reliability of the IsoKai isokinetic lift test in other populations would be beneficial in order to increase the external validity of the test.

Conclusion

The IsoKai isokinetic lift test was found to be a highly reliable measurement of maximal dynamic muscular strength. The test could be used to monitor dynamic muscular strength for the purpose of distinguishing between individuals. As such, the IsoKai isokinetic lift test could be recommended for use as a test for selection based on capacity levels regarding maximal dynamic muscular strength in military settings as well as other physically demanding occupations. In addition, the test is useful for evaluating changes in maximal muscle strength in individuals following interventions.

Supporting information

S1 Appendix.

(DOCX)

S1 Data. IsoKai intrarater reliability.

(XLSX)

S2 Data. IsoKai interrater reliability.

(XLSX)

Acknowledgments

The authors thank all the participants and personnel who took part in the study.

Author Contributions

Conceptualization: Tony Bohman, Matthias Tegern, Alexandra Halvarsson, Lisbet Broman, Helena Larsson.

Data curation: Lisbet Broman, Helena Larsson.

Formal analysis: Tony Bohman, Lisbet Broman, Helena Larsson.

Funding acquisition: Helena Larsson.

Investigation: Tony Bohman, Matthias Tegern, Alexandra Halvarsson, Lisbet Broman, Helena Larsson.

Methodology: Tony Bohman.

Project administration: Tony Bohman, Helena Larsson.

Software: Tony Bohman.

Supervision: Helena Larsson.

Validation: Tony Bohman, Matthias Tegern, Alexandra Halvarsson, Lisbet Broman, Helena Larsson.

Visualization: Tony Bohman, Lisbet Broman.

Writing – original draft: Tony Bohman.

Writing – review & editing: Tony Bohman, Matthias Tegern, Alexandra Halvarsson, Lisbet Broman, Helena Larsson.

References

1. Nindl BC, Alvar BA, J RD, Favre MW, Martin GJ, Sharp MA, et al. Executive Summary From the National Strength and Conditioning Association's Second Blue Ribbon Panel on Military Physical Readiness: Military Physical Performance Testing. *J Strength Cond Res.* 2015; 29 Suppl 11:S216–20.
2. Swain DP, American College of Sports M. ACSM's resource manual for Guidelines for exercise testing and prescription. 7th ed. Philadelphia: Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins; 2014.
3. Paul DJ, Nassis GP. Testing strength and power in soccer players: the application of conventional and traditional methods of assessment. *J Strength Cond Res.* 2015; 29(6):1748–58. <https://doi.org/10.1519/JSC.0000000000000807> PMID: 25546446
4. Abernethy P, Wilson G, Logan P. Strength and power assessment. Issues, controversies and challenges. *Sports Med.* 1995; 19(6):401–17. <https://doi.org/10.2165/00007256-199519060-00004> PMID: 7676101
5. Baltzopoulos V, Brodie DA. Isokinetic dynamometry. Applications and limitations. *Sports Med.* 1989; 8(2):101–16. <https://doi.org/10.2165/00007256-198908020-00003> PMID: 2675256
6. Portney LG, Watkins MP. Foundations of clinical research: applications to practice. 3rd edition, revised. Upper Saddle River, N.J.: Pearson/Prentice Hall; 2015.
7. Kottner J, Audige L, Brorson S, Donner A, Gajewski BJ, Hrobjartsson A, et al. Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed. *J Clin Epidemiol.* 2011; 64(1):96–106. <https://doi.org/10.1016/j.jclinepi.2010.03.002> PMID: 21130355
8. Hauschild VD, DeGroot DW, Hall SM, Grier TL, Deaver KD, Hauret KG, et al. Fitness tests and occupational tasks of military interest: a systematic review of correlations. *Occup Environ Med.* 2017; 74(2):144–53. <https://doi.org/10.1136/oemed-2016-103684> PMID: 27810940
9. Lindberg A-S, Malm C, Oksa J, Gavhed D. Self-Rated Physical Loads of Work Tasks Among Firefighters. *International Journal of Occupational Safety and Ergonomics.* 2014; 20(2):309–21. <https://doi.org/10.1080/10803548.2014.11077042> PMID: 24934427
10. Strating M, Bakker RH, Dijkstra GJ, Lemmink KAPM, Groothoff JW. A job-related fitness test for the Dutch police. *Occupational Medicine.* 2010; 60(4):255–60. <https://doi.org/10.1093/occmed/kqq060> PMID: 20511266
11. Hydren JR, Borges AS, Sharp MA. Systematic Review and Meta-Analysis of Predictors of Military Task Performance: Maximal Lift Capacity. *J Strength Cond Res.* 2017; 31(4):1142–64. <https://doi.org/10.1519/JSC.0000000000001790> PMID: 28135227
12. Nordesjö L, Schéle R. Validity of an ergometer cycle test and measures of isometric muscle strength when prediction some aspects of military performance. *Swedish J Defence Med.* 1974; 10:11–23.
13. Feeler L, St James JD, Schapmire DW. Isometric strength assessment, part I: static testing does not accurately predict dynamic lifting capacity. *Work.* 2010; 37(3):301–8. <https://doi.org/10.3233/WOR-2010-1082> PMID: 20978337

14. Larsson H, Tegern M, Monnier A, Skoglund J, Helander C, Persson E, et al. Content Validity Index and Intra- and Inter-Rater Reliability of a New Muscle Strength/Endurance Test Battery for Swedish Soldiers. *PLoS One*. 2015; 10(7):e0132185. <https://doi.org/10.1371/journal.pone.0132185> PMID: [26177030](https://pubmed.ncbi.nlm.nih.gov/26177030/)
15. The Swedish Defence Recruitment Agency. Available from: <http://www.rekryteringsmyndigheten.se/english/>.
16. Larsson H, Broman L, Harms-Ringdahl K. Individual risk factors associated with premature discharge from military service. *Military medicine*. 2009; 174(1):9–20. PMID: [19216293](https://pubmed.ncbi.nlm.nih.gov/19216293/)
17. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychological bulletin*. 1979; 86(2):420–8. PMID: [18839484](https://pubmed.ncbi.nlm.nih.gov/18839484/)
18. Rankin G, Stokes M. Reliability of assessment tools in rehabilitation: an illustration of appropriate statistical analyses. *Clin Rehabil*. 1998; 12(3):187–99. <https://doi.org/10.1191/026921598672178340> PMID: [9688034](https://pubmed.ncbi.nlm.nih.gov/9688034/)
19. McGraw K, Wong SP. Forming Inferences About Some Intraclass Correlation Coefficients. *Psychological Methods*. 1996; 1(1):30–46.
20. Cicchetti DV. The precision of reliability and validity estimates re-visited: distinguishing between clinical and statistical significance of sample size requirements. *J Clin Exp Neuropsychol*. 2001; 23(5):695–700. <https://doi.org/10.1076/jcen.23.5.695.1249> PMID: [11778646](https://pubmed.ncbi.nlm.nih.gov/11778646/)
21. Bland JM, Altman DG. Measurement error. *Brit Med J*. 1996; 312(7047):1654. PMID: [8664723](https://pubmed.ncbi.nlm.nih.gov/8664723/)
22. de Vet HC, Terwee CB, Knol DL, Bouter LM. When to use agreement versus reliability measures. *J Clin Epidemiol*. 2006; 59(10):1033–9. <https://doi.org/10.1016/j.jclinepi.2005.10.015> PMID: [16980142](https://pubmed.ncbi.nlm.nih.gov/16980142/)
23. Lexell JE, Downham DY. How to assess the reliability of measurements in rehabilitation. *Am J Phys Med Rehab*. 2005; 84(9):719–23.
24. Bland JM, Altman DG. Statistical Methods for Assessing Agreement between Two Methods of Clinical Measurement. *Lancet*. 1986; 1(8476):307–10. PMID: [2868172](https://pubmed.ncbi.nlm.nih.gov/2868172/)
25. Brennan P, Silman A. Statistical methods for assessing observer variability in clinical measures. *BMJ*. 1992; 304(6840):1491–4. PMID: [1611375](https://pubmed.ncbi.nlm.nih.gov/1611375/)
26. Wernbom M, Augustsson J, Thomee R. The influence of frequency, intensity, volume and mode of strength training on whole muscle cross-sectional area in humans. *Sports Med*. 2007; 37(3):225–64. <https://doi.org/10.2165/00007256-200737030-00004> PMID: [17326698](https://pubmed.ncbi.nlm.nih.gov/17326698/)
27. Wilson GJ, Walshe AD, Fisher MR. The development of an isokinetic squat device: reliability and relationship to functional performance. *Eur J Appl Physiol Occup Physiol*. 1997; 75(5):455–61. <https://doi.org/10.1007/s004210050188> PMID: [9189735](https://pubmed.ncbi.nlm.nih.gov/9189735/)
28. Bridgeman LA, McGuigan MR, Gill ND, Dulson DK. Test-Retest Reliability of a Novel Isokinetic Squat Device With Strength-Trained Athletes. *J Strength Cond Res*. 2016; 30(11):3261–5. <https://doi.org/10.1519/JSC.0000000000001410> PMID: [26950357](https://pubmed.ncbi.nlm.nih.gov/26950357/)
29. Fleiss JL. The design and analysis of clinical experiments. New York: Wiley & Sons, Inc; 1986.
30. Currier DP. Elements of research in physical therapy. 3 ed. Baltimore: Williams & Wilkins; 1990.
31. Carter R, Lubinsky J, Domholdt E. Rehabilitation research: principles and applications. 4 ed. St.Louis, Missouri: Elsevier/Saunders; 2011.
32. Bonett DG. Sample size requirements for estimating intraclass correlations with desired precision. *Stat Med*. 2002; 21(9):1331–5. <https://doi.org/10.1002/sim.1108> PMID: [12111881](https://pubmed.ncbi.nlm.nih.gov/12111881/)
33. Walter SD, Eliasziw M, Donner A. Sample size and optimal designs for reliability studies. *Stat Med*. 1998; 17(1):101–10. PMID: [9463853](https://pubmed.ncbi.nlm.nih.gov/9463853/)
34. Jaenen S. Identification of Common Military Tasks. In: Optimizing Operational Physical Fitness, Final Report of Task Group 019, The Research and Technology Organisation (RTO) of Nato, Report TR-HFM-080. 2009.
35. Knapik JJ, Reynolds KL, Harman E. Soldier load carriage: historical, physiological, biomechanical, and medical aspects. *Military medicine*. 2004; 169(1):45–56. PMID: [14964502](https://pubmed.ncbi.nlm.nih.gov/14964502/)
36. Roy TC, Lopez HP, Piva SR. Loads worn by soldiers predict episodes of low back pain during deployment to Afghanistan. *Spine (Phila Pa 1976)*. 2013; 38(15):1310–7.
37. Middleton KJ, Carstairs GL, Caldwell JN, Billing DC, Beck B. The sensitivity of a military-based occupational fitness test of muscular strength. *Appl Ergon*. 2017; 60:255–9. <https://doi.org/10.1016/j.apergo.2016.12.004> PMID: [28166884](https://pubmed.ncbi.nlm.nih.gov/28166884/)