# *Fitness tests and match performance in a male ice hockey national league*



Haugen, Thomas A, Hopkins, William, Breitschädel, Felix, Paulsen, Gøran and Solberg, PA (2021) Fitness tests and match performance in a male ice hockey national league. International Journal of Sports Physiology and Performance, 16 (9). pp. 1303-1310. ISSN 1555-0265

Haugen, Thomas A, Hopkins, William, Breitschädel, Felix, Paulsen, Gøran and Solberg, PA (2021) *Fitness tests and match performance in a male ice hockey national league.* International Journal of Sports Physiology and Performance, 16 (9). pp. 1303-1310. ISSN 1555-0265

# Fitness tests and match performance in male Norwegian upper-league ice-hockey players

**Physical fitness in ice-hockey players**

Thomas Haugen[1,2], Will Hopkins[3], Felix Breitschädel[2,4], Gøran Paulsen[2,5], Paul Solberg[2]

[1]School of Health Sciences, Kristiania University College
[2]Norwegian Olympic Federation, Oslo, Norway
[3]Institute for Health and Sport, Victoria University, Melbourne, Australia
[4]Norwegian University of Science and Technology, Department of Civil and Transport Engineering, Trondheim, Norway
[5]Norwegian School of Sport Sciences, Norway

"This manuscript has been read and approved by all the listed co-authors and meets the requirements of co-authorship as specified in the Human Kinetics Authorship Guidelines. This manuscript is original and not previously published, nor is it being considered elsewhere until a decision is made as to its acceptability by the IJSPP Editorial Review Board."

This is an original investigation with 3 tables and 2 figures. The abstract consists of 221 words and the text 3591 words.

**Corresponding author:**

Thomas Haugen
School of Health Sciences, Kristiania University College, PB 1190 Sentrum, 0107 Oslo, Norway. E-mail: thomas.haugen@kristiania.no.

# Abstract

**Purpose**: To determine if generic off-ice physical fitness tests can provide useful predictions of ice-hockey players' match performance. **Methods:** Approximately 40-60 defenders and 70-100 forwards from the Norwegian male upper ice-hockey league were tested for strength (one repetition maximum [1RM] in squat and bench-press), power (40-m sprint and countermovement jump) and endurance (hanging sit-ups, chins and 3000-m run) annually at the end of every pre-season period between 2008 and 2017. Measures of match performance were each player's season mean counts per match of assists, points, goals, penalty minutes, and plus-minus score. **Results**: Overall, match performance measures displayed trivial to small correlations with the fitness tests. More specifically, points per game had at most small correlations with measures of strength (range, approximately -0.2 to 0.3), speed (approximately -0.2 to 0.3) and endurance (approximately -0.1 to 0.3). After adjustments for age (which showed moderate-to-large correlations with player match performance, multiple regression analyses of each test measure still provided some predictability amongst players of the same age. However, players selected for the national team had substantially better mean scores for most test and match performance measures than those not selected, with a moderate-large difference for age, 1RM squat, and 1RM bench-press. **Conclusions**: Fitness tests had only marginal utility for predicting match performance in Norwegian hockey players, but those selected into the national team had better general fitness.

**Keywords:** Physical capacity; test battery; high-level athletes; hockey game.

# Introduction

Ice-hockey is an intermittent, high-intensive and body-contact team sport where total body fitness is considered compulsory.[1] Beside specific technical and tactical skills, physical characteristics such as speed, strength, muscular power and aerobic capacity must allegedly reach a certain level for players to be successful.[1-4] Accordingly, players in the National Hockey League (NHL) perform 2-4 weekly off-ice conditioning sessions throughout the training year.[5] In order to monitor training status, NHL strength and conditioning coaches assess their players regularly on a broad range of physical and physiological parameters.[5]

The validity and usefulness of physical testing have been the subject of much debate among practitioners and scientists. Some argue that only sport-specific tests are useful,[6,7] while others suggest that generic tests provide a valuable understanding of the underlying physical resources of performance factors.[8] This information can in turn be used as framework for individual and collective training prescriptions, informing recovery strategies and load management. Previous ice-hockey related studies have typically focused on tests such as $VO_2max$, Wingate, vertical or horizontal jumps, sprint, abdominal endurance, leg press, squat, chins, bench-press and their isolated association to one or two of the following on-ice performance parameters: entry draft selection order,[9-11] cross-sections of playing standards (elite vs. non-elite, division I vs. division III, seniors vs. juniors),[12,13] individual skating performance,[13,14] play-off success or number of points acquired by the team during the season.[15] Peyer et al.[16] reported that leg press, pull-ups, bench press, and repeat sprint performance were significantly correlated with plus-minus score (every player on the ice for the scoring team gets a "plus" point, while every player on the ice for the team scored against gets a "minus" point; each player's score for the match is the sum of the points). Other authors have reported that athletes of higher playing standards or draft status achieve superior test scores in one or several fitness tests than players at lower performance levels.[10-13] However, inconsistencies among these studies are present in terms of which physical measures are associated with which performance variables. In contrast, Quinney et al.[15] reported that pre-season fitness was not related to team success, and Vescovi et al.[9] concluded that off-ice testing did not predict playing ability in terms of draft order. The utility of physical fitness tests for predicting match performance in ice hockey therefore remains uncertain. Performance tests and measures of match performance vary between studies and may partly explain inconsistencies in outcomes. Moreover, commonly used performance metrics among practitioners such as numbers of scored goals, assists, total points (the sum of goals and assists) and +/- score are rarely included in scientific studies. Consequently, more comprehensive and longitudinal studies are needed.

The Norwegian Olympic training center is a standard testing facility for a large number of teams, including the entire national upper league in ice-hockey, as well as the national team players. A broad range of physical test results and match performance data for the corresponding seasons have been performed and collected over 10 years, and this database provides the potential for addressing several different questions related to the role of physical fitness in ice-hockey. Therefore, the aim of this study was to determine which tests provide useful predictions of players' match performance in the subsequent season.

# Methods

### Subjects

The present correlational study included 848 male ice-hockey players (age 23 ± 4 y, body mass 84 ± 7 kg, body height 182 ± 5 cm) from 14 different clubs in the Norwegian upper league. Player positions were identified for each athlete by their coaches or by self-report as: goalkeepers, forwards or defenders. National team athletes were defined as players who represented Norway in senior World Championships or corresponding qualifying matches. All players had at least two years of strength and conditioning experience.

This study was based on pre-existing data from annually testing that these athletes performed for training purposes, and thus, no informed consent was obtained. The Regional Committee for Medical and Health Research Ethics waives the requirement for ethical approval for this study. Therefore, the ethics of the study is done according to the institutional requirements at School of Health Sciences, Kristiania University College. Approval for data security and handling was obtained from the Norwegian Centre for Research Data (reference number 292977). Because of the newly implemented General Data Protection Regulations (GDPR) by the European Union, all de-identified data stored at Kristiania University College were deleted after the project completion.

**Methodology**

All included players were tested at the Norwegian Olympic Training Centre in the time period 2008-2017. The participants were tested at the end of the pre-season period, 1-2 wk prior to season start. Regarding nutrition, hydration, sleep and physical activity, the athletes were instructed to prepare themselves as they would for a regular match, including no high intensity training the last two days before testing. The tests were performed between 9 AM and 6 PM. The warm-up typically consisted of 10-15 minutes easy jogging, followed by 5-6 min with sprint specific drill exercises and 2-3 strides with increasing speed.

*Sprint and vertical jump testing:* The fitness testing was initiated with a sprint test on a rubberized indoor track. The athletes started from a standing, split-stance position with the tiptoe of the front foot placed at the starting line. Two to three trials were performed every 3–5 min, and their best 40-m sprint time was retained for analysis as sprint speed. Procedures and apparatus are described in Haugen et al.[17] Immediately after the sprint test, the athletes performed three trials of countermovement jumps (CMJ). Procedures and apparatus are described in Haugen et al.[18] The best out of three trials was retained for analysis.

*3000-m run*: Immediately after the CMJ test, all participants performed a brief re-warm up consisting of 5-6 minutes easy jog. In the 3000-m running test, the athletes ran 7.5 laps on an outdoor 400-m track with rubberized surface, and they were timed with hand-held stopwatches. Time was converted to speed. Outside temperature was between 15 and 24 °C at the time of testing.

*1RM squat and bench-press*: About 30 min after the 3000-m test, the athletes performed a squat specific warm up consisting of 3-4 sets with increasing sub-maximal loads (40-80 % of 1RM), followed by 1-2 single lifts at 85-90 % of 1RM. Up to three single test lifts were allowed. Best performance approved by the referee was retained for analysis. Between 5 and 10 min after the squat, the athletes performed a bench press specific warm-up similar to the squat warm-up protocol. Best performance of the three attempts approved by the referee was retained for analysis. Strong and laud verbal encouragement was provided. For more information regarding procedures and apparatus, we refer to Haugen et al.[19]

*Pull-ups*: Maximum number of bodyweight pull-ups was tested with pronated grip on a bar (Figure 1, panel A and B). A full arm extension in the lower position was required for each attempt, and the athletes pulled up until the chin was above the bar. No kipping was allowed during the pull.

*Hanging sit-ups*: This exercise was used to assess abdominal muscle capacity (Figure 1, panel C and D). Ankles were secured and attached to a bracket on top of the apparatus. In the lower position, the athlete suspended alongside a vertically oriented bench, holding a circular rope

(10 cm diameter) behind his neck with elbows and shoulders flexed. In the upper position of the sit ups, the elbows had to touch the knees while the glutes were in contact with the bench. The knee joints were fixed at 90° during the entire test. The athletes were allowed to take one break (≤ 3 s) during the entire test, and maximum number of repetitions was registered.

*Box jumps*: To assess anaerobic capacity, the athletes performed a 90-s box-jump test (Figure 1, panel E and F). The athlete started beside a 40-cm tall box and jumped up and down alternately from each side. Number of box jumps attained during the 90-s period was registered.

***Figure 1 about here***

Match performance measures were downloaded from the Norwegian Ice-Hockey Association's web site (https://www.hockey.no/live) and included number of games played during the season, goals, assists, total points, plus-minus score, and whether the player was selected for the national team.

**Statistical analysis**

A correlation matrix for all fitness measures (including age and body mass index) for all players' data over all years of competition was structured to reveal clusters of measures with higher correlations between measures among clusters than between clusters. The relationships between match-performance measures were also investigated with correlation matrices for defenders and forwards separately, but not for goalies, whose match roles and performance scores differed markedly from those of defenders and forwards. Uncertainty in the correlations was expressed as 90% compatibility limits, which were derived conservatively by assuming a sample size given by the number of different players contributing to the correlation.

Points per game was chosen as the measure of match performance for further investigation of the relationship between fitness and match performance. Correlations were derived for each fitness-test measure with points per game for each year and presented as a time course for defenders and forwards. Age was the strongest predictor of match performance, so it was included in multiple linear regressions analyses with each fitness measure, to determine the extent to which the predictability of age could be improved by including fitness and the extent to which fitness retained predictability amongst players of the same age. The highest correlation between age the fitness measures was 0.32, so collinearity was not an issue.
The analyses were performed with Proc Reg in the Statistical Analysis System (Version 9.4, University Studio Edition, SAS Institute, Cary NC). Separate analyses were performed for each year; the multiple correlation was expressed as the square root of the R-squared adjusted for degrees of freedom, and partial correlations of age and the fitness measure were expressed as the square root of the respective Type-II partial R-squareds. Compatibility limits were derived by assuming the Fisher z transformation of the correlation was normally distributed.
A comparison of the mean fitness and match-performance scores of players selected for the national team with those not selected was performed with Proc Mixed in SAS. Fixed effects in the linear mixed model were nominal variables for year of competition (10 levels, to adjust for annual changes) and a binary variable indicating whether the player was selected. Least-squares means of the binary variable provided the comparison of means. Random effects were player identity (to account for repeated measurement from year to year) and the residual variance (representing within-player variability from year to year). A separate variance was estimated for selected and non-selected players and for their residual variances; the variances were combined to give between-player standard deviations for selected and non-selected players. The standard deviation of non-selected players was used to standardize and thereby assess the magnitude of the difference between the means of selected and non-selected players, evaluated

with the following scale for trivial, small, moderate, large, very large and extremely large respectively: <0.2, ≥0.2, ≥0.6, ≥1.2, ≥2.0, and ≥4.0 times the standard deviation.[20] Uncertainty in the difference in means is presented as 90% compatibility limits. Decisions about magnitudes accounting for the uncertainty were based on a reference-Bayesian analysis with a minimally informative prior,[21-23] which provided estimates of chances that the true magnitude was a substantial decrease, a trivial value, and a substantial increase. Effects had adequate precision if the chances of one or other substantial true value were <5% (the 90% compatibility interval did not include a substantial increase and decrease). Effects with adequate precision are reported with a qualitative descriptor for the magnitudes with chances that were >25% using the following scale: 25-75%, possibly; 75-95%, likely; 95-99.5%, very likely; >99.5%, most likely.[20] When the chances of a substantial magnitude were >95%, the magnitude is also described as clear. The chances of substantial and trivial magnitudes of the true effect were the percent areas of the sampling distribution of the effect statistic to the left or right of the smallest important value (the trivial-small threshold). The sampling distribution was assumed to be a *t* distribution. To account for inflation of error with the large number of comparisons, effects with precision deemed adequate with 99% compatibility intervals (chances of one or other substantial true value <0.5%) are highlighted.

# Results

Table 1 shows the correlation matrix for age, body mass index and fitness-test measures for all positions and years. The variables have been ordered in the matrix to reveal three clusters (strength, power, endurance) with generally higher values for correlations within clusters than between clusters. Age had higher correlations with the strength measures but was not included with this cluster.

***Table 1 about here***

Table 1 also shows games played, assists per game and points per game had generally higher correlations than the other three measures of match performance with anthropometric and fitness-test measures Games played, assists, points and goals per game had the strongest correlations to age and the fitness variables. The correlations between assists per game and total points per game (defenders, 0.97; forwards, 0.96) showed that these measures effectively represented the same construct. Further analyses for the relationships of match performance with fitness were therefore performed with points per game (including both assists and goals).

Figure 2 shows the correlations of points per game with the other measures for defenders and forwards in each year. Age had higher correlations than any of the test measures, with correlations being moderate-high (forwards, range ~0.50 to 0.60, 90% compatibility limits ~±0.15; defenders ~0.40 to 0.65, ±0.20). Correlations with the other test measures were at most borderline low-moderate for strength (~-0.2 to 0.3), speed (~-0.2 to 0.3), power (~-0.1 to 0.3) and endurance (~ -0.3 to 0.2). In the five most recent years (2013-2017) points per game had generally low correlations with measures of strength for forwards (squat and bench press) and trivial correlations for defenders (except pull-ups), correlations with measures of power were trivial-low for both positions, and measures of endurance were generally trivial-small, with the exception of box jumps, which showed reasonably consistent small negative correlations for both positions.

***Figure 2 about here***

Multiple linear correlations in which points per game was predicted by age and each of the fitness-test measures showed that none of the measures provided a substantial improvement in the prediction by age alone. Table 2 shows the results of these analyses for representative measures of strength, power and endurance for the year 2017. The multiple correlation adjusted for degrees of freedom and the partial correlation for age were practically identical to the simple correlation with age, showing that adjustment for fitness did not substantially affect the predictability of age. For fitness-test measures with low simple correlations, the partial correlations generally remained trivial or small, showing that these measures still provided some predictability amongst players of the same age.

***Table 2 about here***

The extent to which fitness and match-performance measures were related to selection of players for the national team is evident in Table 3. Selected defenders and forwards had clearly higher moderate-large mean scores for age, 1RM squat and 1RM bench-press, and there were similar clear differences for defenders with pull-ups and forwards with the 3000-m run. All these differences had adequate precision at the 99% level. Body mass index and box jumps were likely higher in selected vs non-selected forwards (small differences at the 99% level), but other measures showed trivial-moderate differences that were either unclear or likely higher only at the 90% level for selected defenders and forwards. Selected goalies had small-large higher test scores, but owing to the small sample size, most were unclear. Selected defenders and forwards had small-moderate clearly higher mean scores for most match performance measures, and selected goalies clearly played moderately more games; most of these differences had adequate precision at the 99% level.

***Table 3 about here***

# Discussion

The purpose of the present study was to determine whether physical fitness tests could provide useful predictions of players' match performance in the subsequent season. Games played, assists, points and goals per game displayed the strongest correlations to the fitness variables. However, the correlations with the physical tests were small at most, and they retained small predictability among players of the same age. Based on these observations, the usefulness of physical fitness tests for match performance predictions can be questioned. Still, players selected for the national team displayed substantially better mean scores for most tests and match performance metrics than those not selected.

The test battery in this study was generic in nature, representing measures of strength, power and endurance. The tests were designed to ensure that a large number of players could perform all tests within one day each year at the same location. Similar tests are commonly used by practitioners[5] and scientists.[6] Indeed, the selection of tests to model strength, power and endurance will affect predictive abilities. For example, 3000-m running performance was used in the present study, while $VO_2$max on a cycle ergometer is used in the NHL combine. Although both tests can be used to encapsulate endurance, it is reasonable to assume that motivation, pacing and skills may affect the outcomes of the tests. Moreover, while several authors have emphasized the importance of strength, power and endurance for performance in ice-hockey and similar sports,[16,19,24,25] others have questioned the application of such tests due to lack of specificity. [6,7] It is therefore possible that more specific tests (e.g., individual skating performance) have stronger relationships to performance. Previous studies of this issue have shown inconsistent results, possibly explained by differences in fitness test batteries, testing protocols, performance outcomes and sample size. Interpreting and comparing results must therefore be done with caution.

Draft success has been used as the dependent variable in several studies,[9-11] but the draft process is influenced by confounding factors such as reports of past performance, professional scouting intuition, game observation, player aggression, psychological factors, team/coaching philosophy and playing position. For example, based on a very large sample size ($n$=853), Burr et al.[11] observed that several fitness test results from the National Hockey League Entry Draft (NHLED) combine significantly predicted draft success, but the variance explained by the best model was only 7% (a multiple correlation of $\sqrt{0.07} = 0.26$). Overall, most performance measures have inherent advantages and limitations. By including a broad range of fitness tests and match performance measures, as performed in the present study, a clearer understanding of the relationship between on-ice and off-ice performance can be provided.

Although age is not a performance measure, it is essential for practitioners and scientists to possess knowledge regarding how generic physical skills develop as a function of age. Data obtained from world-class athletics, weightlifting and powerlifting contestants show that performance mainly peaks in the age 25-28 y on average.[25-27] However, in sports with higher maximal force demands, peak age is even higher.[27] The younger and poorly trained the athletes, the larger the annual change scores. Quantification of peak age for the varying fitness tests was outside the scope of this study. However, because most players were in the age range 18-30 y, and most of them likely achieve peak performance in most fitness tests in their late 20s, it is not surprising to observe high correlations among age and fitness test scores. Inclusion of age does not detract from the present analysis. If inclusion of a fitness measure with age does not substantially improve the prediction of performance, one can conclude that there is no point in measuring this fitness measure. However, some small effects on performance was still evident for squat among forwards, and CMJ and pull-ups for backwards.

The national team players in the present study were generally more fit compared to the other players, in line with observations from several previous cross-sectional studies.[12,13] Direct comparisons can be made with national team handball players, as they have tested sprint, CMJ, 3000-m running and 1RM bench press and squat with identical apparatus and procedures.[19] These comparisons show that national team ice-hockey players outperform handball players in all these tests, particularly in 1RM squat. We cannot disregard the possibility that the national team ice-hockey players were selected partly on the basis of the fitness tests. One of the incentives for conducting this yearly testing was to monitor the players' physical fitness development and create a strong training culture. Players who considered themselves in contention for selection might have trained harder, thinking that fitness would be taken into account in the selection process. Hence, it is possible that the national team players were disproportionally fit for their ice-hockey performance level.

# Practical applications

Although the fitness tests in this study displayed only marginal utility for predicting match performance, this does not imply that off-ice conditioning for ice-hockey players is a waste of time. In most team sports, the key technical-tactical skills must be maximized, while other capabilities merely need to meet a minimum requirement. Based on the results presented in Table 3, it is reasonable to argue that most of the players in this study are well-trained and above the minimum requirements. Hence, the annual testing concept has been a success, as it has inspired athletes to improve their generic fitness. Well-developed physical skills make players prepared for duels and reduce relative match intensity, thereby avoiding negative effects associated with fatigue (technical performance, decision making, injuries, etc.).[24,29]

# Conclusions

This study showed that generic fitness testing had limited value for predicting match performance in Norwegian ice-hockey players. However, players selected into the national team exhibited better physical fitness than the remaining upper league players.


# Acknowledgements

The authors would like to thank the Norwegian Ice-Hockey Association, the participating players and the following organizers and test personnel: Per Egil Refsnes, Annbjørg Engeseth, Kristoffer Holm, Jon Haukeland, Petter Saltsten, Magnus Midttun, Christian Helland, Sindre Madsgaard, Eirik Hole and Benjamin Jensen.

# References


1. Cox MH, Miles DS, Verde TJ, Rhodes EC. Applied physiology of ice-hockey. *Sports Med*. 1995;19:184-201.
2. Montgomery DL. Physiological profile of professional hockey players - a longitudinal comparison. *Appl Physiol Nutr Metab*. 2006;31:181-185.
3. Stanula A, Roczniok R. Game intensity analysis of elite adolescent ice-hockey players. *J Hum Kinet*. 2014;44:211-221.
4. Stanula AJ, Gabryś TT, Roczniok RK, Szmatlan-Gabryś UB, Ozimek MJ, Mostowik AJ. Quantification of the demands during an ice-hockey game based on intensity zones determined from the incremental test outcomes. *J Strength Cond Res*. 2016;30:176-183.
5. Ebben WP, Carroll RM, Simenz CJ. Strength and conditioning practices of National Hockey League strength and conditioning coaches. *J Strength Cond Res*. 2004;18:889-897.
6. Nightingale SC, Miller S, Turner A. The usefulness and reliability of fitness testing protocols for ice-hockey players: a literature review. *J Strength Cond Res*. 2013;27:1742-1748.
7. Runner AR, Lehnhard RA, Butterfield SA, Tu S, O'Neill T. Predictors of speed using off-ice measures of college hockey players. *J Strength Cond Res*. 2016;30:1626-1632.
8. Mendez-Villanueva A, Buchheit M. Football-specific fitness testing: adding value or confirming the evidence? *J Sports Sci*. 2013;31:1503-1508.
9. Vescovi JD, Murray TM, Fiala KA, VanHeest JL. Off-ice performance and draft status of elite ice-hockey players. Int J Sports Physiol Perform. 2006;1:207-221.
10. Burr JF, Jamnik VK, Dogra S, Gledhill N. Evaluation of jump protocols to assess leg power and predict hockey playing potential. *J Strength Cond Res*. 2007;21:1139-1145.
11. Burr JF, Jamnik RK, Baker J, Macpherson A, Gledhill N, McGuire EJ. Relationship of physical fitness test results and hockey playing potential in elite-level ice-hockey players. *J Strength Cond Res*. 2008;22:1535-1543.
12. Hoff J, Kemi OJ, Helgerud J. Strength and endurance differences between elite and junior elite ice-hockey players. The importance of allometric scaling. *Int J Sports Med*. 2005;26:537-541.
13. Peterson BJ, Fitzgerald JS, Dietz CC, Ziegler KS, Ingraham SJ, Baker SE, Snyder EM. Division I hockey players generate more power than Division III players during on- and off-ice performance tests. *J Strength Cond Res*. 2015;29:1191-1196.
14. Krause DA, Smith AM, Holmes LC, Klebe CR, Lee JB, Lundquist KM, Eischen JJ, Hollman JH. Relationship of off-ice and on-ice performance measures in high school male hockey players. J Strength Cond Res. 2012;26:1423-1430.
15. Quinney HA, Dewart R, Game A, Snydmiller G, Warburton D, Bell G. A 26 year physiological description of a National Hockey League team. *Appl Physiol Nutr Metab*. 2008;33:753-760.
16. Peyer KL, Pivarnik JM, Eisenmann JC, Vorkapich M. Physiological characteristics of National Collegiate Athletic Association Division I ice-hockey players and their relation to game performance. *J Strength Cond Res*. 2011;25:1183-1192.
17. Haugen T, Breitschädel F, Seiler S. Sprint mechanical variables in elite athletes: Are force-velocity profiles sport specific or individual? *PLoS ONE* 2019;14(7):e0215551.

18. Haugen T, Breitschädel F, Seiler S. Countermovement jump height in national team athletes of various sports: a framework for practitioners and scientists. *Int J Sports Physiol Perf*. 2020 (in press).
19. Haugen T, Tønnessen E, Seiler S. Physical and physiological characteristics of male handball players: influence of playing position and competitive level. *J Sports Med Phys Fitness* 2016;56:19-26.
20. Hopkins WG, Marshall SW, Batterham AM, Hanin J. Progressive statistics for studies in sports medicine and exercise science. *Med Sci Sports Exerc*. 2009;41:3-12.
21. Hopkins WG. Magnitude-based decisions. *Sportscience* 2019;23:i-iii.
22. Hopkins WG, Batterham AM. Error rates, decisive outcomes and publication bias with several inferential methods. *Sports Med*. 2016;46:1563-1573.
23. Hopkins WG, Batterham AM. The vindication of magnitude-based inference. *Sportscience* 2018;22:19-29.
24. Suchomel TJ, Nimphius S, Stone MH. The importance of muscular strength in athletic performance. *Sports Med*. 2016;46:1419-1449.
25. Haugen T, Tønnessen E, Seiler S. Anaerobic performance testing of professional soccer players 1995-2010. *Int J Sport Physiol Perform*. 2013;8:148-156.
26. Allen SV, Hopkins WG. Age of peak competitive performance of elite athletes: a systematic review. *Sports Med*. 2015;45:1431-1441.
27. Haugen T, Solberg PA, Morán-Navarro R, Breitschädel F, Hopkins W, Foster C. Peak age and performance progression in world-class track-and-field athletes. *Int J Sports Physiol Perform*. 2018; 13(9):1122-1129.
28. Solberg PA, Hopkins WG, Paulsen G, Haugen TA. Peak age and performance progression in world-class weightlifting and powerlifting athletes. *Int J Sports Physiol Perf*. 2019. [E-pub ahead of print]
29. Haugen T, Seiler S. Physical and physiological testing of soccer players: Why, what and how should we measure? *Sportsscience* 2015;19:10-26.

# Figure legends

**Figure 1.** Upper and lower positions of the pull-ups (Panel A and B), hanging sit-ups (Panel C and D) and box jumps (Panel E and F).

**Figure 2.** Correlations of points per game with age, body mass index and each of fitness-test measures for defenders (•) and forwards (o) in each year 2008-2017. Shading indicates trivial correlations. Dashed lines indicate thresholds for low (±0.10), moderate (±0.30), high (±0.50) and very high correlations (±0.70). Bars are 90% compatibility intervals.

**Table 1**. Correlations between all test measures, using data for all players for all years 2008-2017. Measures have been ordered and outlined to show higher correlations among those representing generally strength (body mass index, 1RM squat, 1RM bench-press), power (40-m sprint speed, countermovement jump height) and endurance (sit-ups, pull-ups, box jumps, 3000-m running speed). Correlations of each test measure with each game measure are also shown, excluding goalies.

| | **Age** | **BMI** | **Squat** | **Bench press** | **40-m sprint** | **CMJ** | **Sit-ups** | **Chin-ups** | **Box jumps** | **3000-m run** |
|---|---|---|---|---|---|---|---|---|---|---|
| **Age** | | 0.32 | 0.19 | 0.31 | -0.04 | 0.08 | -0.06 | -0.02 | -0.22 | 0.05 |
| **BMI** | 0.32 | | 0.41 | 0.46 | -0.10 | -0.02 | -0.03 | -0.15 | -0.17 | -0.08 |
| **Squat** | 0.19 | 0.41 | | 0.61 | 0.24 | 0.32 | 0.11 | 0.13 | 0.15 | 0.03 |
| **Bench press** | 0.31 | 0.46 | 0.61 | | 0.19 | 0.22 | 0.13 | 0.28 | 0.06 | 0.04 |
| **40-m sprint** | -0.04 | -0.1 | 0.24 | 0.19 | | 0.60 | 0.19 | 0.31 | 0.31 | 0.15 |
| **CMJ** | 0.08 | -0.02 | 0.32 | 0.22 | 0.60 | | 0.28 | 0.25 | 0.16 | 0.03 |
| **Sit-ups** | -0.06 | -0.03 | 0.11 | 0.13 | 0.19 | 0.28 | | 0.35 | 0.31 | 0.21 |
| **Pull-ups** | -0.02 | -0.15 | 0.13 | 0.28 | 0.31 | 0.25 | 0.35 | | 0.32 | 0.26 |
| **Box jumps** | -0.22 | -0.17 | 0.15 | 0.06 | 0.31 | 0.16 | 0.31 | 0.32 | | 0.41 |
| **3000-m run** | 0.05 | -0.08 | 0.03 | 0.04 | 0.15 | 0.03 | 0.21 | 0.26 | 0.41 | |
| **Games played** | 0.36 | 0.28 | 0.20 | 0.28 | 0.06 | 0.14 | 0.05 | 0.00 | -0.01 | 0.15 |
| **Assists/game** | 0.50 | 0.05 | 0.09 | 0.20 | 0.11 | 0.14 | -0.07 | 0.12 | -0.14 | 0.08 |
| **Points/game** | 0.50 | 0.06 | 0.10 | 0.21 | 0.12 | 0.16 | -0.06 | 0.12 | -0.12 | 0.07 |
| **Goals/game** | 0.37 | 0.06 | 0.11 | 0.19 | 0.11 | 0.17 | -0.04 | 0.10 | -0.05 | 0.04 |
| **Plus-minus/game** | 0.22 | -0.03 | 0.03 | 0.06 | 0.12 | 0.18 | 0.06 | 0.06 | 0.05 | 0.08 |
| **Penalty min./game** | 0.25 | 0.16 | 0.14 | 0.23 | -0.04 | 0.05 | 0.07 | 0.00 | -0.09 | -0.01 |

BMI = body mass index, CMJ = countermovement jump. Number of players: 677 (3000-m run) through 848 (age). Total number of observations per correlation: 1029-2020 (test measures); 378-681(game with test measures). 90% compatibility limits < ~±0.07 (test measures); <~±0.11 (game with test measures).

**Table 2**. Simple, multiple and partial correlations of points per game with age and each of three other measures of fitness exemplifying strength (squat), power (countermovement jump) and endurance (pull-ups) for defenders and forwards in 2017.

| **Fitness predictor** | *n* | **Simple correlation** | | **Multiple correlation**[a] | **Partial correlation** | |
|---|---|---|---|---|---|---|
| | | *Age* | *Fitness* | | *Age* | *Fitness* |
| *Defenders* | | | | | | |
| Squat | 43 | 0.50, ±0.19 | -0.05, ±0.25 | 0.47, ±0.21 | 0.51, ±0.18 | -0.07, ±0.25 |
| CMJ | 52 | 0.48, ±0.18 | 0.23, ±0.22 | 0.48, ±0.18 | 0.47, ±0.17 | 0.19, ±0.22 |
| Chin-ups | 47 | 0.49, ±0.19 | 0.26, ±0.23 | 0.49. ±0.19 | 0.47, ±0.18 | 0.19, ±0.23 |
| *Forwards* | | | | | | |
| Squat | 90 | 0.44, ±0.14 | 0.21, ±0.17 | 0.45, ±0.14 | 0.43, ±0.14 | 0.17, ±0.17 |
| CMJ | 105 | 0.45, ±0.13 | 0.03, ±0.16 | 0.43, ±0.13 | 0.45, ±0.12 | 0.06, ±0.16 |
| Pull-ups | 94 | 0.43, ±0.14 | 0.08, ±0.17 | 0.42, ±0.14 | 0.43, ±0.13 | 0.09, ±0.17 |

Correlations are shown with ±90% compatibility limits.
*n* = sample size, CMJ = countermovement jump.
[a]Square root of the R-squared adjusted for degrees of freedom.

**Table 3**. Age, body mass index (BMI), fitness and match-performance measures for players not selected and selected for the national team in all years 2008-2017.

| | | Not selected | Selected | Selected minus not selected | |
|---|---|---|---|---|---|
| | **Position** | *Mean ± SD (n)* | *Mean ± SD (n)* | *Mean, ±90%CI* | *Decision* |
| Age (y) | Defenders | 23.2 ± 5.3 (295) | 28.7 ± 6.3 (10) | 5.5, ±3.7 | **mod.↑*** ** |
| | Forwards | 23.2 ± 5.5 (467) | 27.3 ± 4.2 (18) | 4.1, ±1.8 | **mod.↑******** |
| | Goalies | 22.6 ± 5.2 (86) | 26.0 ± 6.0 (5) | 3.4, ±5.8 | mod.↑ |
| BMI ($kg·m^{-2}$) | Defenders | 25.4 ± 1.6 (295) | 25.8 ± 1.3 (10) | 0.4, ±0.7 | small↑ |
| | Forwards | 25.3 ± 1.6 (467) | 26.0 ± 1.1 (18) | 0.7, ±0.5 | **small↑**** |
| | Goalies | 24.5 ± 1.6 (86) | 26.0 ± 2.8 (5) | 1.5, ±2.7 | mod.↑ |
| 1RM squat (kg) | Defenders | 151 ± 22 (240) | 185 ± 20 (7) | 34, ±12 | **large↑******** |
| | Forwards | 151 ± 24 (404) | 168 ± 18 (13) | 17, ±9 | **mod.↑***** |
| | Goalies | 141 ± 20 (65) | 157 ± 17 (4) | 16, ±19 | mod.↑** |
| 1RM bench-press (kg) | Defenders | 107 ± 14 (276) | 121 ± 11 (9) | 14, ±6 | **mod.↑******** |
| | Forwards | 106 ± 15 (440) | 115 ± 12 (15) | 10, ±6 | **mod.↑***** |
| | Goalies | 94 ± 13 (74) | 108 ± 14 (4) | 14, ±16 | mod.↑** |
| 40-m sprint ($m·s^{-1}$) | Defenders | 7.61 ± 0.27 (274) | 7.60 ± 0.43 (8) | -0.01, ±0.21 | trivial↓ |
| | Forwards | 7.66 ± 0.26 (437) | 7.71 ± 0.31 (14) | 0.05, ±0.15 | small↑ |
| | Goalies | 7.39 ± 0.28 (78) | 7.56 ± 0.33 (5) | 0.17, ±0.31 | mod.↑ |
| CMJ (cm) | Defenders | 39.8 ± 5.0 (282) | 43.1 ± 5.9 (10) | 3.3, ±3.1 | mod.↑** |
| | Forwards | 39.7 ± 5.1 (450) | 40.3 ± 4.5 (16) | 0.6, ±1.9 | trivial↑ |
| | Goalies | 37.5 ± 4.8 (83) | 40.1 ± 4.4 (5) | 2.6, ±4.2 | small↑ |
| Hanging sit-ups (n) | Defenders | 19.8 ± 5.3 (277) | 23.0 ± 7.2 (8) | 3.2, ±4.7 | mod.↑ |
| | Forwards | 20.5 ± 4.9 (430) | 21.7 ± 5.1 (15) | 1.2, ±2.3 | small↑ |
| | Goalies | 19.0 ± 4.7 (76) | 22.8 ± 2.0 (3) | 3.8, ±1.8 | **mod.↑***** |
| Pull-ups (n) | Defenders | 12.5 ± 4.2 (272) | 16.2 ± 3.8 (9) | 3.7, ±2.3 | **mod.↑***** |
| | Forwards | 13.1 ± 4.7 (431) | 12.9 ± 4.4 (13) | -0.2, ±2.2 | trivial↓ |
| | Goalies | 12.1 ± 5.1 (73) | 14.4 ± 5.7 (4) | 2.3, ±6.5 | small↑ |
| Box jumps (n) | Defenders | 88.7 ± 9.6 (240) | 90.7 ± 6.3 (7) | 2.0, ±4.8 | small↑ |
| | Forwards | 91.3 ± 8.9 (398) | 95.3 ± 5.8 (14) | 4.0, ±2.8 | **small↑**** |
| | Goalies | 84.1 ± 6.3 (67) | 93.8 ± 3.2 (2) | 9.7, ±11.2 | large↑** |
| 3000-m run ($m·s^{-1}$) | Defenders | 4.23 ± 0.27 (235) | 4.26 ± 0.35 (6) | 0.03, ±0.29 | trivial↑ |
| | Forwards | 4.27 ± 0.30 (374) | 4.50 ± 0.17 (13) | 0.23, ±0.09 | **mod.↑******** |
| | Goalies | 4.14 ± 0.38 (68) | 4.38 ± 0.10 (2) | 0.24, ±0.23 | mod.↑** |
| Games played | Defenders | 29 ± 16 (295) | 36 ± 12 (10) | 7, ±6 | **small↑**** |
| | Forwards | 30 ± 16 (467) | 37 ± 11 (18) | 7, ±5 | **small↑**** |
| | Goalies | 14 ± 15 (86) | 26 ± 12 (5) | 12, ±7 | **mod.↑***** |
| Assists per game | Defenders | 0.21 ± 0.21 (263) | 0.42 ± 0.20 (10) | 0.21, ±0.11 | **mod.↑***** |
| | Forwards | 0.26 ± 0.25 (423) | 0.55 ± 0.18 (18) | 0.28, ±0.07 | **mod.↑******** |
| | Goalies | 0.04 ± 0.05 (61) | 0.06 ± 0.03 (5) | 0.01, ±0.03 | small↑ |
| Points per game | Defenders | 0.27 ± 0.28 (263) | 0.57 ± 0.24 (10) | 0.30, ±0.14 | **mod.↑***** |
| | Forwards | 0.45 ± 0.41 (423) | 0.83 ± 0.25 (18) | 0.38, ±0.11 | **mod.↑******** |
| | Goalies | 0.06 ± 0.05 (61) | 0.07 ± 0.03 (5) | 0.01, ±0.03 | small↑ |
| Goals per game | Defenders | 0.07 ± 0.08 (263) | 0.14 ± 0.08 (10) | 0.07, ±0.04 | **mod.↑***** |
| | Forwards | 0.18 ± 0.18 (423) | 0.28 ± 0.11 (18) | 0.10, ±0.05 | **small↑***** |
| | Goalies | - | - | - | - |
| Plus-minus per game | Defenders | -0.03 ± 0.43 (263) | 0.19 ± 0.56 (10) | 0.22, ±0.32 | small↑ |
| | Forwards | -0.05 ± 0.35 (423) | 0.28 ± 0.26 (18) | 0.33, ±0.09 | **mod.↑******** |

| | | | | |
|---|---|---|---|---|
| Goalies | -0.20 ± 1.23 (61) | -0.23 ± 0.59 (5) | -0.02, ±0.49 | trivial↓ |

*n* = number of players, CI = compatibility interval, mod. = moderate, ↑ = increase, ↓ = decrease, CMJ = countermovement jump. Asterisks indicate substantial effects with adequate precision as follows: *possibly, **likely, ***very likely, ****most likely. Decisions in bold have adequate precision with 99% compatibility intervals.