Self-reported Measures of Strength and Sport-Specific Skills Distinguish Ranking in an International Online Fitness Competition

Paul R. Serafini, Yuri Feito, and Gerald T. Mangine

Department of Exercise Science and Sport Management, Kennesaw State University, Kennesaw, Georgia

Abstract

Serafini, PR, Feito, Y, and Mangine, GT. Self-reported measures of strength and sport-specific skills distinguish ranking in an international online fitness competition. *J Strength Cond Res* 32(12): 3483–3493, 2018—To determine if self-reported performance measures could distinguish ranking during the 2016 CrossFit Open, data from 3,000 male (n = 1,500; 27.2 ± 8.4 years; 85.2 ± 7.9 kg; 177.0 ± 6.5 cm) and women (n = 1500, 28.7 ± 4.9 years; 63.7 ± 5.8 kg; 163.7 ± 6.6 cm) competitors was used for this study. Competitors were split by gender and grouped into quintiles (Q1–Q5) based upon their final ranking. Quintiles were compared for one-repetition maximum (1RM) squat, deadlift, clean and jerk (CJ), snatch, 400 m sprint, 5,000 m run, and benchmark workouts (Fran, Helen, Grace, Filthy-50, and Fight-Gone-Bad). Separate one-way analyses of variance revealed that all competitors in Q1 reported greater ($p \leq 0.05$) 1RM loads for squat (men: 201.6 ± 19.1 kg; women: 126.1 ± 13.0 kg), deadlift (men: 232.4 ± 20.5 kg; women: 148.3 ± 14.5 kg), CJ (men: 148.9 ± 12.1 kg; women: 95.7 ± 8.4 kg), and snatch (men: 119.4 ± 10.9 kg; women 76.5 ± 7.6 kg) compared with other quintiles. In addition, men in Q1 (59.3 ± 5.9 seconds) reported faster ($p \leq$ 0.05) 400 m times than Q3 only (62.6 ± 7.3 seconds), but were not different from any group in the 5,000 m run. Women in Q2 (67.5 ± 8.8 seconds) reported faster ($p \leq 0.05$) 400 m times than Q3–Q5 (73.5–74.8 seconds), and faster (21.3 ± 1.8 minutes, $p < 0.02$) 5,000 m times than Q4 (22.6 ± 2.2 minutes) and Q5 (22.6 ± 1.9 minutes). Faster ($p \leq 0.05$) Fran times were reported by Q1 (men: 138.2 ± 13.3 seconds; women: 159.4 ± 28.3 seconds) compared with other groups, while the results of other workouts were variable. These data indicate that the most successful athletes excel in all areas of fitness/skill, while lower-ranking athletes should focus on developing strength and power after achieving sufficient proficiency in sport-specific skills.

Key Words CrossFit, muscle strength and power, anaerobic performance, WOD

Introduction

Winners of the Reebok CrossFit Games (i.e., the sport of fitness) are awarded the title "Fittest on Earth." The format of this competition is designed to narrow the initial participant pool (>200,000 athletes) down to the top CrossFit athletes within each category (i.e., 40 adult men, 40 adult women, 40 teenagers, and 200 Masters). The first stage of this competition is the CrossFit Open (CFO), a 5-week, 5-workout, online competition, which is open to anyone in the world. Its purpose is to identify the top 10–30 individuals from each of the 17 worldwide regions who will then vie for a final position in The Games. To accomplish this goal, the competitors are presented with a workout challenge each week. They are given 4 days to complete the challenge to the best of their ability before submitting their results. Although individual CFO workouts have been repeated from year to year, generally each week's workout is unique and unknown to the participants until its release. Nevertheless, it is expected that each workout will require some combination of strength, power, endurance, and/or sport-specific skill (13). Athletes who possess a strong foundation in each of these abilities would appear to be best equipped to rank high each week of the CFO and progress to the regional qualifier. However, it is not clear whether competitors should focus on all physiological components related to the CFO challenges, or if proficiency in specific components is more common in higher-ranking competitors.

Physiological adaptations and improvements in performance are thought to be specific to the characteristics of training (i.e., modality, volume load, intensity load, rest intervals, etc.) (14,15). Although leading strength and conditioning authorities such as the National Strength and Conditioning Association recommend a linear approach (i.e., focusing on a specific goal during each training phase) to maximize adaptations (14), evidence suggests that targeting

Address correspondence to Dr. Gerald T. Mangine, gmangine@kennesaw.edu.

32(12)/3474–3484

Journal of Strength and Conditioning Research

several physiological systems within the same phase of training (i.e., nonlinear periodization model) may be equally (or more) beneficial for trained athletes (4,17,24). Regardless of training status, CrossFit training follows a nonlinear model by combining traditional strength/power movements (e.g., squats, deadlifts, cleans, presses, and snatch), continuous exercise (e.g., running and rowing), and CrossFit-specific skills (e.g., rope jumping and climbing, gymnastic movements, and the use of nontraditional implements). Workouts are commonly structured so that the prescribed exercises are completed (or as many repetitions as possible (AMRAP) are performed) within a certain time frame, while the specific characteristics of each workout (i.e., volume and intensity load, exercise duration, and tempo, etc.) vary from day to day to shift the emphasis between anaerobic and aerobic energy system pathways. Interestingly, anecdotal evidence suggests that elite CrossFit competitors design their training regimen instead of following what is planned for the general public (20,25).

The details regarding elite-competitor programing are not publicly available, and typically vary between athletes. However, it may be possible to identify and target important training outcomes based upon the performance measures in which they excel. CrossFit athletes monitor their training progress via traditional measures of strength and power, muscular endurance, anaerobic endurance, aerobic endurance, and several benchmark workouts. Proficiency in these measures appears to be related (5,10), and may be useful for distinguishing competitive level within this sport. For instance, aerobic capacity and anaerobic power have been observed to influence total repetitions completed during a 12-minute AMRAP (i.e., 12-minute AMRAP that included medicine ball throws, kettlebell swings, and burpee pullups) (5), while strength in the CrossFit Total (overhead press, squat, deadlift) and aerobic fitness (i.e., $\dot{V}O_2$ at anaerobic threshold) have been related ($r = -0.53$ to -0.65) to specific benchmark workouts (i.e., Fran and Grace) (10). Though these data appear to suggest that CFO success might be explained by performance in these measures, a study of the top 30 competitors of the 2013 CrossFit did not observe differences amongst these competitors in several performance measures (6). Nevertheless, these findings only appear to be applicable to a limited sample of elite CrossFit athletes (6) or they are indicative of performance in a predetermined, controlled setting (5,10). Since CFO workouts are unknown prior to the contest, athletes have little time to prepare for a specific event. Thus, it is unclear how proficiency in traditional measures of athletic fitness or common benchmark workouts is different across a larger spectrum of CrossFit athletes competing in the CFO.

Identifying the key physiological and athletic characteristics that distinguish performance in more traditional sports (e.g., football, baseball, basketball, etc.) is a common endeavor (16,18,19,21). Such information allows strength and conditioning professionals to develop more effective and specific training programs that may better translate to optimal in-game performance. However, the sport of fitness is relatively new and few studies have attempted to describe the characteristics that distinguish performance (5,6,10). While those studies have focused on small random samples or elite athletes, no study has attempted to examine these characteristics in a larger sample of elite CFO competitors. On the CrossFit website, CFO athletes have the option of creating a profile where they can upload individual scores and monitor their progress throughout the competition, which includes several traditional and sport-specific performance measures. Therefore, the purpose of this study was to determine if these self-reported performance measures could distinguish ranking during the 2016 CFO. Additionally, our secondary goal was to determine if these performance measures (i.e., muscular strength and power, muscular endurance, aerobic and anaerobic endurance, sport-specific measures) were greater in higher-ranking competitors. Since workouts are unknown before competition, and the sport of fitness requires proficiency in several fitness areas, we hypothesize that the highest-ranking competitors will report the best performance scores in all self-reported measures, and that performance scores will decrease along with ranking.

Methods

Experimental Approach to the Problem

Self-reported performance scores, located on a publicly available website (1), were collected from the individual competitor profiles of the top 1,500 male and top 1,500 female competitors of the 2016 CFO. Performance scores included measures of muscle strength, power and endurance, aerobic and anaerobic endurance, and sport-specific skill. To answer our research questions, competitors were split into quintiles (Q1–Q5) based upon their final ranking within each sex, and comparisons in each performance score measure were made between quintiles. All data were collected immediately following the conclusion of the 2016 CFO competition (28 March 2016–18 April 2016).

Subjects

Men ($n = 1{,}500$; 27.3 ± 4.3 years (range: 17.0–46.0 years); 85.2 ± 7.9 kg; 177.0 ± 6.5 cm), and women ($n = 1500$, 28.7 ± 4.9 years (range: 17.0–46.0 years); 63.7 ± 5.8 kg; 163.7 ± 6.6 cm) competitors of the 2016 CFO were included in this study. All characteristic were expressed as mean ± *SD*. All participants were selected in ranked order and possessed a profile on the CrossFit Games website (1) where their self-reported performance data was located. Since these data were preexisting and publicly available, the Kennesaw State University's Institutional Review Board classified this study as exempt, and participants did not have to provide their informed consent. All data were downloaded from The Games website and decoded so that no identifiable information (i.e., name) was available from any of the athletes. Descriptive characteristics of competitors within each quintile are presented in Table 1.

Table 1. Individual quintile characteristics for male and female competitors.

	Q1	Q2	Q3	Q4	Q5
Men					
Age (y)	25.9 ± 3.2	27.2 ± 4.4	27.5 ± 4.6	27.3 ± 4.1	27.9 ± 4.8
Height (cm)	176.5 ± 5.6	176.8 ± 6.1	177.3 ± 7.1	177.3 ± 6.6	177.3 ± 7.0
Mass (kg)	86.3 ± 8.0	86.0 ± 7.7	85.5 ± 7.5	84.9 ± 8.1	84.9 ± 8.0
Women					
Age (y)	28.3 ± 4.3	28.5 ± 4.7	28.8 ± 5.1	28.9 ± 5.0	29.1 ± 5.2
Height (cm)	163.2 ± 6.2	163.8 ± 6.9	164.0 ± 6.9	163.7 ± 6.6	164.0 ± 6.4
Mass (kg)	63.7 ± 5.1	64.1 ± 5.5	63.6 ± 6.5	63.3 ± 5.9	63.5 ± 5.8

Performance Ranking. The CFO consisted of 5, previously unknown (to the competitors) challenges that were individually released on each Thursday at 5:00 PM pacific time (PT), of the 5-week competition. Competitors were not limited by the number of attempts they could take to achieve the best score possible, but scores had to be submitted to competition officials by the following Monday at 5:00 PM PT. To progress to the regional competition, all submitted attempts had to have been completed at a CrossFit affiliate in front of a judge who passed the online Judges Course. Alternatively, an attempt could be filmed using standardized filming criteria and submitted online to competition officials for review. All workout descriptions, standards, and scoring criteria are located on the competition website (20). At each week's conclusion, competitors with valid attempts were ranked within their sex and competitive class by competition officials. A competitor's final CFO ranking was calculated as the sum of his/her rank during each week. In this way, a competitor who ranks well each week will have a lower final ranking score than competitors who do not rank well. The specific events of the 2016 CFO are presented in Table 2.

Procedures

Performance Measures. The CrossFit Games website provides an athlete profile section where competitors can upload their individual performance scores for measures of muscle strength, power and endurance, anaerobic and aerobic endurance, and sport-specific skill. All available and continuous data from each profile were included in the present analysis. For this investigation, it was assumed that all performance scores were completed as per the criteria described by the competitive organization (7), and that all scores were current.

Muscle Strength, Power and Endurance. Absolute (±0.1 kg) and relative (±0.1 kg·body mass^{-1}) performance using commonly accepted technical requirements (7) were recorded for all measures of muscle strength and power. One-repetition maximum (1RM) in the back squat (SQ) and the deadlift (DL) exercises were recorded for muscle strength, while 1RM performance in the clean and jerk (CJ) and the snatch (SN) exercises were recorded for muscle power. The total number of continuous pull-ups (PU), defined as hanging from a bar with both arms extended and pulling the body up to a point where the chin is higher than the bar, and then lowering the body back to full arm extension, was recorded for muscle endurance.

Anaerobic and Aerobic Endurance. The fastest time running 400 m (±0.01 seconds) and 5,000 m (±0.1 minutes) were recorded for anaerobic and aerobic endurance, respectively.

Sport-Specific Skills. Details of the common benchmark workouts used as measures of sport-specific skills are described in Table 3. Briefly, 4 of the recorded events (i.e., Fran, Filthy 50, Helen, and Grace) were scored by time-to-completion (±0.1 minutes), while Fight-Gone-Bad was scored as the total number of repetitions completed within the set time frame.

Statistical Analyses

To control for accuracy, all performance scores exceeding 6 standard deviations from the mean were removed prior to any statistical analysis (27). Subsequently, statistical comparisons between quintiles were accomplished using separate one-way analyses of variance on each measure of performance. In the event of a significant F-ratio, Tukey post-hoc tests were used for pair-wise comparisons. All data are reported as mean ± standard deviations (*SD*) with significance set at ($p \leq 0.05$). SPSS statistical software (V. 22.0, SPSS Inc., Chicago, IL, USA) was used for all analyses.

Results

Muscular Strength, Power, and Endurance

Quintile comparisons for all absolute measures of muscle strength, power, and endurance for male and female athletes are illustrated in Tables 4 and 5, respectively. Significant

Table 2. Events of the 2016 CrossFit open.*

	16.1 (Week 1)	16.2 (Week 2)	16.3 (Week 3)	16.4 (Week 4)	16.5 (Week 5)
Scoring	20-min AMRAP	Total repetitions and time-to-completion	7-min AMRAP	13-min AMRAP	Time-to-completion
Workout	Walking overhead lunge for 25 ft	5 consecutive and continuous 4 min rounds.	10 Power snatches followed by 3 bar muscle-ups	55 Dead lifts	Thrusters (i.e., squat to shoulder press) followed by bar-over burpees for 21, 18, 15, 12, 9, 6, and 3 repetitions.
	8 bar-over burpees	Each round included: 25 Toes-to-bar, followed by 50 double-unders, and then squat cleans		55 wall-ball shots	
	Walking overhead lunge for 25 ft	Competitors were required to complete all repetitions within a round with 4 min in order to progress to the next round.		55 calorie row	
	8 chest-to-bar pull-ups			55 handstand push-ups	
Resistance load	Overhead lunge: men = 43.1 kg; women = 29.5 kg	Squat clean: R-1 = 15 repetitions at (men = 61.2 kg; women = 38.6 kg): R-2 = 13 repetitions at (men = 83.9 kg; women = 52.2 kg): R-3 = 11 repetitions at (men = 102.1 kg; women = 65.8): R-4 = 9 repetitions at (men = 124.7 kg; women = 79.4 kg): R-5 = 7 repetitions at (men = 142.9 kg; women = 93.0 kg)	Power snatch: men = 34.0 kg; women = 24.9 kg	Deadlift: men = 102.1 kg; women = 70.3 kg; wall-ball shots: men = 9.1 kg to 3 m height; women: 6.4 kg to 2.4 m height	Thrusters: men = 43.1 kg; women = 29.5 kg

*AMRAP = as many repetitions as possible; R = round.

($p < 0.001$) quintile differences were observed for all absolute and relative measures of muscle strength, power, and endurance for men (Table 4) and women (Table 5). With some exceptions, competitors in Q1 reported greater loads ($p \leq 0.05$) compared with all other successive quintiles. Beyond Q1, no differences were observed between Q2 and Q3 in any of these measures. Both Q2 and Q3 reported strength and power scores that exceeded ($p \leq 0.05$) those reported by Q4 and Q5. No differences between Q2–Q5 were observed in reported PU repetitions except for women in Q2 and Q5 ($p < 0.001$).

Aerobic and Anaerobic Endurance

Significant quintile differences existed in the reported 400 m sprint time for both men ($p < 0.001$) and women ($p = 0.003$). For men, similar sprint times ($p = 0.998$) were reported by Q1 (59.3 ± 5.9 seconds) and Q2 (59.0 ± 6.1 seconds), who both reported faster times ($p \leq 0.05$) compared with Q3 (62.6 ± 7.3 seconds). Additionally, men in Q2 reported

Table 3. Common benchmark workouts-of-the-day reported by competitors.

	Fran	Filthy 50	Helen	Grace	Fight-Gone-Bad
Scoring	Three rounds scored by time to completion.	Scored by time to completion	Three rounds scored by time to completion	Scored by time to completion	Three rounds of five 1-min stations scored by points. 1 rep = 1 point. As many reps as possible
Workout	21/15/9	50 × box jumps (0.6 m box)	400 m run	30 clean and jerks (men: 61.2 kg; women: 43.1 kg)	Wall-ball (men: 9.1 kg, 3 m target; women: 6.4 kg, 2.7 m target)
	Thrusters (men: 43.1 kg	50 × jumping pull-ups	21 kettlebell swings (men and women: 24.0 kg)		Sumo deadlift high-pull (men: 34.0 kg; women: 24.9 kg)
	women: 24.5 kg)	50 × kettlebell swings (16.4 kg)	12 pull-ups		Box jump (0.5 m box)
	Pull-ups	50 (steps) × walking lunges			Push-press (men: 34.0 kg; women: 24.9 kg)
		50 × kees-to-elbows			Calorie row
		50 × push press (men: 20.4 kg; women: 13.6 kg)			1-min rest
		50 × back extensions			
		50 × wall-ball shots (men: 9.1 kg, 3 m target; women: 6.4 kg, 2.7 m Target)			
		50 × burpees			
		50 × double-unders			

significantly faster times than Q5 (62.3 ± 7.0 seconds, $p = 0.033$). No differences were observed between men in Q4 (61.8 ± 10.2 seconds) and any other quintile. For women, no differences existed between Q1 (71.0 ± 9.2 seconds) and Q2 (67.5 ± 8.8 seconds). However, compared with Q2, significantly slower 400 m sprint times were reported in Q3 (74.8 ± 10.1 seconds, $p = 0.004$), Q4 (73.5 ± 9.7 seconds, $p = 0.033$), and Q5 (74.2 ± 11.3 seconds, $p = 0.022$). No other differences were observed in 400 m sprint performance.

For 5,000 m run performance, significant quintile differences were observed for women ($p = 0.004$) but not men ($p = 0.372$). Compared with women in Q1 (21.3 ± 1.8 minutes), significantly slower run times were reported by Q4 (22.6 ± 2.2 minutes, $p = 0.008$) and Q5 (22.6 ± 1.9 minutes, $p = 0.016$). No significant performance differences were observed between all other groups and women in Q2 (22.1 ± 2.6 minutes) and Q3 (22.3 ± 2.1 minutes).

Sport-Specific Skills

The performance differences between quintiles for common benchmark workouts are illustrated in Figures 1 and 2 for men and women, respectively.

Fran. Significant ($p < 0.001$) quintile differences were observed for men and women in their reported Fran time (Figures 1A and 2A). Specifically, Q1 (males: 2.3 ± 0.2 minutes; women: 2.7 ± 0.5 minutes) reported faster times compared with all subsequent quintiles for both men ($p \leq 0.05$) and women ($p < 0.001$), while Q2 reported times (men: 2.4 ± 0.3 minutes; women: 3.1 ± 0.7 minutes) that were faster than Q4 (men: 2.6 ± 0.4 minutes, $p < 0.001$; women: 3.3 ± 0.8 minutes, $p = 0.013$) and Q5 (men: 2.5 ± 0.3 minutes, $p = 0.003$; women: 3.5 ± 0.7 minutes, $p < 0.001$) but similar to Q3 (men: 2.5 ± 0.3 minutes, $p = 0.460$; women: 3.1 ± 0.6 minutes, $p = 0.979$). Differences also existed amongst Q3–Q5, but were less orderly.

Filthy 50. Although a significant F-ratio was observed between male quintiles in time to complete Filthy 50 ($p = 0.026$), no significant pairwise differences were found (Figure 1B). Men in Q1 (17.0 ± 3.4 minutes) tended to report faster times than those in Q3 (18.4 ± 2.7 minutes, $p = 0.060$) and Q5 (18.4 ± 1.9 minutes, $p = 0.062$). For women (Figure 2B), Q1 (19.6 ± 3.0 minutes) reported times were faster than those for Q3 (21.5 ± 3.5 minutes, $p = 0.044$) and Q5

Table 4. Means and standard deviations for muscular strength and endurance measures in men.

	Q1	Q2	Q3	Q4	Q5
Clean and jerk					
Absolute (kg)	148.9 ± 12.1*†‡§	142.9 ± 13.1‖‡§	139.9 ± 12.6 ‖§	138.0 ± 13.1‖*	135.2 ± 14.7‖*†
Relative ($kg \cdot kg^{-1}$)	1.72 ± 0.15*†‡§	1.66 ± 0.15‖§	1.64 ± 0.15‖	1.63 ± 0.23‖	1.59 ± 0.18‖*
Snatch					
Absolute (kg)	119.4 ± 10.9*†‡§	114.5 ± 10.7‖‡§	112.3 ± 12.1‖§	110.2 ± 12.0‖*	108.11 ± 12.3‖*†
Relative ($kg \cdot kg^{-1}$)	1.39 ± 0.12*†‡§	1.33 ± 0.12‖§	1.31 ± 0.14‖§	1.30 ± 0.17‖	1.27 ± 0.14‖*†
Deadlift					
Absolute (kg)	232.4 ± 20.5*†‡§	225.1 ± 21.7‖‡§	219.7 ± 25.0‖	217.4 ± 22.1‖*	217.6 ± 25.7‖*
Relative ($kg \cdot kg^{-1}$)	2.69 ± 0.23†‡§	2.63 ± 0.26	2.57 ± 0.28‖	2.56 ± 0.35‖	2.56 ± 0.31‖
Squat					
Absolute (kg)	201.6 ± 19.1*†‡§	195.2 ± 21.5‖†‡§	188.6 ± 22.4‖*	186.9 ± 20.4‖*	185.01 ± 21.8‖*
Relative ($kg \cdot kg^{-1}$)	2.33 ± 0.24†‡§	2.28 ± 0.25‡§	2.21 ± 0.26‖	2.21 ± 0.32‖*	2.18 ± 0.27‖*
Pull-ups (reps)	60.1 ± 13.3†‡§	56.5 ± 11.0	55.0 ± 13.0‖	52.9 ± 11.9‖	53.1 ± 10.0‖

*Significantly ($p \leq 0.05$) different from Q2.
†Significantly ($p \leq 0.05$) different from Q3.
‡Significantly ($p \leq 0.05$) different from Q4.
§Significantly ($p \leq 0.05$) different from Q5.
‖Significantly ($p \leq 0.05$) different from Q1.

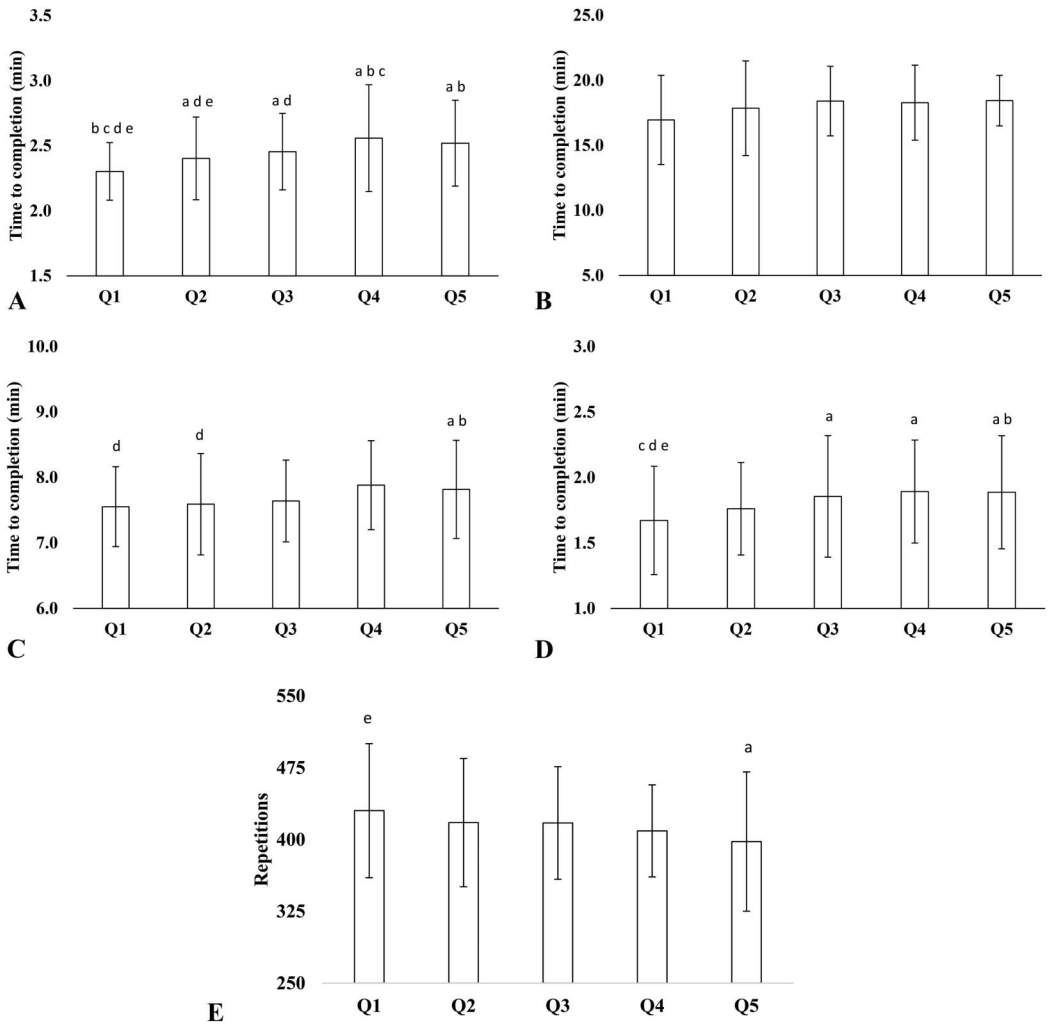
(21.9 ± 3.0 minutes, $p = 0.010$), but not Q2 (21.1 ± 3.7 minutes, $p = 0.255$) or Q4 (21.3 ± 3.0 minutes, $p = 0.106$). No other differences were observed between quintiles.

Helen. Significant quintile differences in time to complete Helen were observed for men ($p = 0.002$) and women ($p < 0.001$), but specific differences varied by gender. For

Table 5. Means and standard deviations for muscular strength and endurance measures in women.

	Q1	Q2	Q3	Q4	Q5
Clean and jerk					
Absolute (kg)	95.7 ± 8.4*†‡§	91.2 ± 10.2‖‡§	88.9 ± 8.5‖‡§	85.2 ± 9.6‖*†	85.4 ± 9.5‖*†
Relative ($kg \cdot kg^{-1}$)	1.51 ± 0.14*†‡§	1.43 ± 0.14‖‡§	1.40 ± 0.13‖‡§	1.35 ± 0.15‖*†	1.34 ± 0.14‖*†
Snatch					
Absolute (kg)	76.5 ± 7.6*†‡§	72.5 ± 8.5‖‡§	71.0 ± 8.0‖‡§	67.5 ± 9.0‖*†	66.6 ± 8.2‖*†
Relative ($kg \cdot kg^{-1}$)	1.21 ± 0.12 *†‡§	1.13 ± 0.12‖‡§	1.11 ± 0.12‖‡§	1.07 ± 0.13‖*†	1.04 ± 0.12‖*†
Deadlift					
Absolute (kg)	148.3 ± 14.5*†‡§	143.6 ± 16.7‖§	139.7 ± 21.3‖	139.3 ± 16.6‖	137.4 ± 16.6‖*
Relative ($kg \cdot kg^{-1}$)	2.35 ± 0.25 *†‡§	2.25 ± 0.25‖§	2.19 ± 0.31‖	2.20 ± 0.23‖	2.16 ± 0.23‖*
Squat					
Absolute (kg)	126.1 ± 13.0*†‡§	122.1 ± 15.4‖‡§	119.5 ± 15.8‖	115.8 ± 13.0‖*	116.4 ± 14.6‖*
Relative ($kg \cdot kg^{-1}$)	2.00 ± 0.22 *†‡§	1.91 ± 0.23‖‡§	1.88 ± 0.24‖	1.83 ± 0.22‖*	1.83 ± 0.21‖*
Pull-ups (reps)	46.1 ± 12.4 *†‡§	39.7 ± 10.5‖§	35.9 ± 11.4‖	37.4 ± 11.2‖	32.3 ± 10.1‖§

*Significantly ($p \leq 0.05$) different from Q2.
†Significantly ($p \leq 0.05$) different from Q3.
‡Significantly ($p \leq 0.05$) different from Q4.
§Significantly ($p \leq 0.05$) different from Q5.
‖Significantly ($p \leq 0.05$) different from Q1.



a. Significantly ($p < 0.05$) different from Q1.
b. Significantly ($p < 0.05$) different from Q2.
c. Significantly ($p < 0.05$) different from Q3.
d. Significantly ($p < 0.05$) different from Q4.
e. Significantly ($p < 0.05$) different from Q5.

Figure 1. Quintile performance comparisons for men in measures of sport-specific skill (A. Fran, (B) Filthy 50, (C) Helen, (D) Grace, (E) Fight-Gone-Bad). a, Significantly ($p \leq 0.05$) different from Q1. b, Significantly ($p \leq 0.05$) different from Q2. c, Significantly ($p \leq 0.05$) different from Q3. d, Significantly ($p \leq 0.05$) different from Q4. e, Significantly ($p \leq 0.05$) different from Q5.

men, similar completion times were reported by Q1 (7.6 ± 0.6 minutes), Q2 (7.6 ± 0.8 minutes), Q3 (7.6 ± 0.6 minutes), and Q5 (7.8 ± 0.7 minutes). However, only Q1 ($p = 0.007$) and Q2 ($p = 0.023$) reported faster times than competitors in Q4 (7.9 ± 0.7 minutes). For women, competitors in Q1 (8.6 ± 0.8 minutes, $p < 0.001$), Q2 (8.8 ± 0.8 minutes, $p = 0.015$), and Q3 (8.9 ± 1.5 minutes, $p = 0.039$) reported faster times than Q5 (9.4 ± 1.2 minutes). Although competitors in Q4 (9.0 ± 0.8 minutes, $p = 0.374$) reported similar times to Q5 ($p = 0.374$), their reported completion times were only significantly slower than Q1 ($p = 0.047$). No other quintile differences were observed for men (Figure 1C) or women (Figure 2C) in Helen time.

Grace. Significant ($p < 0.001$) differences were observed in time to complete Grace for men and women. For men, Q1 (1.7 ± 0.4 minutes) reported faster ($p < 0.001$) times than Q3 (1.9 ± 0.5 minutes), Q4 (1.9 ± 0.4 minutes), and Q5 (1.9 ± 0.4 minutes), but did not differ from Q2 (1.8 ± 0.4 minutes,

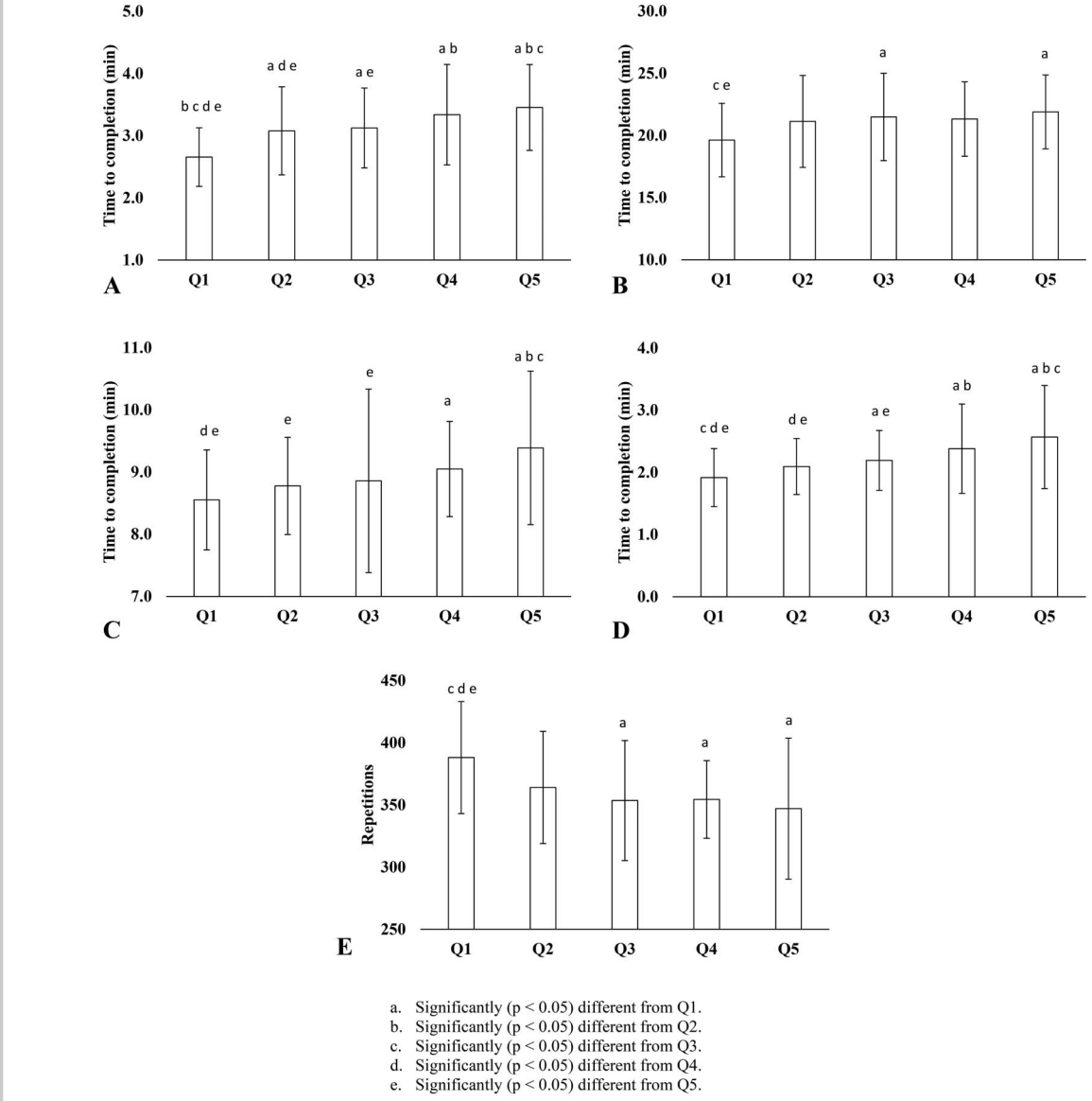


Figure 2. Quintile performance comparisons for women in measures of sport-specific skill (A. Fran, (B) Filthy 50, (C) Helen, (D) Grace, (E) Fight-Gone-Bad). a, Significantly ($p \leq 0.05$) different from Q1. b, Significantly ($p \leq 0.05$) different from Q2. c, Significantly ($p \leq 0.05$) different from Q3. d, Significantly ($p \leq 0.05$) different from Q4. e, Significantly ($p \leq 0.05$) different from Q5.

$p = 0.285$). The times reported by Q2 only tended to differ from those reported by Q4 ($p = 0.052$). For women, Q1 (1.9 ± 0.5 minutes) also reported faster times than Q3 (2.2 ± 0.5 minutes, $p = 0.005$), Q4 (2.4 ± 0.7 minutes, $p < 0.001$), and Q5 (2.6 ± 0.8 minutes, $p < 0.001$) and were similar to Q2 (2.1 ± 0.5 minutes, $p = 0.179$). Q2 reported faster times than Q4 ($p = 0.006$) and Q5 ($p < 0.001$), while Q3 reported faster times than Q5 ($p < 0.001$). No other quintile differences were observed for men (Figure 1D) or women (Figure 2D) in Grace time.

Fight-Gone-Bad. Significant quintile differences were observed in the number of completed repetitions during Fight-Gone-Bad by men ($p = 0.021$) and women ($p < 0.001$). For men, only participants in Q1 (430.3 ± 70.1 reps) reported completing more repetitions than those in Q5 (398 ± 72.8 reps, $p = 0.010$). For women, Q1 (388.0 ± 45.0 reps) reported completing more ($p < 0.001$) repetitions than Q3 (353.5 ± 48.2 reps), Q4 (354.4 ± 31.2 reps), and Q5 (347.0 ± 56.6 reps), but only tended to be more than Q2 (364.0 ± 45.1 reps, $p = 0.057$). There were no

other quintile differences for men (Figure 1E) or women (Figure 2E).

Discussion

Understanding the physical traits (e.g., muscle strength, power, endurance, etc.) and sport-specific skills that influence in-game performance is important for evaluating athletic potential, areas in need of improvement, and program development (14,15). In many traditional sports, measurable physical traits are often associated with in-game performance statistics (16,18,19), but are rarely a perfect representation of sport. In contrast, the sport of fitness incorporates several identifiable and measurable benchmarks into a single competition. However, because the actual competition consists of several random factors (e.g., exercise combinations and order, set/round/repetition scheme, time limits, etc.), a perfect portrayal of this sport is not possible in a controlled setting. Nevertheless, the inclusion of these measurable components into competition improves the likelihood that proficiency in these measures translates well to athletic success. Interestingly, and perhaps due to the variability in actual competitive event requirements (i.e., tasks are randomly designed each year), few studies have attempted to evaluate which traits identify success (5,6,10). Even though these studies identified certain indicators of success (i.e., strength, aerobic fitness, anaerobic power, CrossFit training experience), they did not consider whether these traits affected actual competitive rankings (i.e., CFO) in a large sample of competitors. To our knowledge, this study appears to be the first to investigate whether self-reported performance scores could distinguish actual placement in the CFO competition. Our data suggests that the highest-ranking competitors excel in most measures of athletic fitness and sport-specific skill. Beyond the highest tier, most competitors possess comparable levels of proficiency in anaerobic and aerobic endurance, as well as in sport-specific skill.

Previously, performance during CrosssFit training workouts have been positively related to strength (10) and peak power (5). Our data supports these findings. Regardless of sex, the most successful competitors (i.e., Q1) reported greater values for all measures of strength (DL and SQ) and power (CJ and SN) compared with lower-ranking competitors. This was also true with relative measures of strength and power, though second-tier competitors (i.e., Q2) reported similar scores as Q1 for DL (men only), SQ (men only), and SN (women only). Although CFO events rarely include single maximal repetitions (2), possessing sufficient strength and power is still necessary to complete traditional lifts for several repetitions. For instance, 16.2 (week 2) required participants to perform 5 rounds of a squat clean ladder (Table 2) that culminated with 7 repetitions at 142.9 kg for men and 93.0 kg for women. Being capable of completing all repetitions at this final stage would roughly equate to a 1RM squat clean of 172 kg for men and 112 kg for women (9), which would have been greater than the reported CJ values for competitors in Q1 (men: +23.1 kg; women: +19 kg). Nevertheless, these competitors would have been most capable of completing this task compared with lower-ranking competitors. Similarly, the first stage of 16.4 (week 4) required competitors to perform 55 DL repetitions at 102.1 kg for men or at 70.3 kg for women. While it would appear to be unlikely that any competitor could have completed all repetitions consecutively (i.e., without breaks), based upon their reported 1RM DL, competitors in Q1 (men and women) would have been capable of performing the most consecutive repetitions (9) resulting in less time spent within this stage. Although no differences in self-reported strength/power measures were previously reported within the top 30 competitors (of 2013) (6), our data suggests that greater CFO success occurs with stronger and more powerful competitors when considering a much larger sample population.

Even though many CFO events typically require anaerobic fitness (1,5), differences between higher- and lower-ranking competitors were less clear for 400 m sprint times compared with other anaerobic measures (i.e., strength and power). For 400 m sprint time, higher-ranking competitors also reported faster sprint times, but being elite in this task encompassed only the top 2 quintiles (i.e., Q1 and Q2) for both genders. Further, similar performances were typically observed in competitors within the bottom 3 quintiles (i.e., Q3–Q5). The lack of clear differences between quintile groups may be related to the indirect nature of a 400 m sprint time as an anaerobic performance measure (28). Although it is well accepted that the 400 m sprint requires a large contribution of energy production to be derived via anaerobic pathways (11), it is not clear from sprint time alone whether the event is truly representative of anaerobic capacity. Its duration (<2 minutes) is also much shorter than the duration of most benchmark and CFO workouts (17). It is possible that anaerobic capacity, and thus its importance to CFO performance, may be more clearly identified by longer (~3 minutes) maximal assessments (7,26). Another possibility may be related to the lack of running/sprinting events in the CFO competition (2) due to standardization concerns across testing locations. Consequently, developing proficiency in the 400 m sprint may be important for improving competitive rank, but only in a much broader sense compared with strength and power development.

CrossFit training has been reported to improve aerobic capacity in participants of all fitness levels (24). Nevertheless, aerobic performance (i.e., 5,000 m run) was only important in women, and could only distinguish between higher-ranking competitors and those ranking lower than Q3. These findings are supported by those reported by Butcher and colleagues (10), who did not observe an association between $\dot{V}O_2$max and performance in a CrossFit style workout. It is possible that the redundant nature of a 5,000 m run is not an accurate portrayal of the energy demands of CFO

events. Few CFO events or benchmark workouts last more than 20 minutes, and they require several transitions between exercises (27), which places a greater demand upon the cardiovascular and aerobic systems (12). Thus, it is possible that by including CrossFit training into a regimen, an athlete will experience sufficient aerobic adaptation to be competitive in the CFO.

Self-reported proficiency in CrossFit-specific workouts also appeared to distinguish ranking with the clearest differences between quintiles occurring in Fran and Grace. In these workouts, the reported completion times progressively increased as rank decreased. Though simple in their design (Table 3), these short (~1.5–3.5 minutes), moderate, heavy-intensity workouts require sufficient muscle strength, muscle endurance, anaerobic fitness, and technical skill to be competitive (3,13). Competency in these events would appear to be descriptive of an athlete who is proficient in several areas important to CFO performance. Interestingly, the distinctions between quintiles became less clear in longer (>3.5 minutes), more complicated benchmark workouts (i.e., Helen, Filthy 50, and Fight-Gone-Bad), despite all CFO workouts being similar or longer in duration (6–20 minutes) (27). Although higher-ranking competitors generally performed better, no differences were observed among Q1 through Q3 in any of these measures apart from female Filthy 50 scores. Meanwhile, compared with the lowest-ranking competitors (Q5), significantly better performances were limited to Q1, except for female Helen scores. It is possible that, as explained above, the limited differences observed in Helen (~7.5–9.5 minutes) scores were due to its running component (400 m sprint) not being specific to CFO events. For Filthy 50 and Fight-Gone-Bad, it is not clear why more specific differences were not observed, but it could be related to their greater complexity compared with any CFO workout of 2016. Therefore, it is advisable for competitors to focus on developing proficiency in short, simple, benchmark workouts rather than those that are longer and more complicated.

This is the second study to examine self-reported performance measures in relation to competitive ranking. Previously, Bellovary et al. (2014) examined the same measures of performance, apart from examining relative strength/power measures and benchmark workouts, in relation to the ranking of the top 30 athletes of the 2013 CrossFit Games. The limitation in that study, as well as the present investigation, is the usage of self-reported performance measures. These data were collected from user profiles located on a publicly available website where participants upload their information at any time. Consequently, the accuracy and report dates of these data are unknown. While verifying the accuracy of these data would be difficult and time-consuming, the variability within such data might be overcome by a large sample and removing unrealistic scores (27); here, we removed data scores that exceeded 6 standard deviations from the mean. Although these precautions do not eliminate the possibility that rankings in the 2016 CFO and self-reported performance measures were not chronologically analogous, it is generally accepted that performance in elite competitors is not likely to change dramatically from year to year (8,22).

Practical Applications

The CFO is an increasingly popular competition that in 2016 began with over 200,000 athletes. To identify which competitors will progress to the regional qualifier, 5 unique workouts are released each week that challenge several areas of physical fitness. While success in this competition would appear to require proficiency in each area, it was previously unknown which are most important. The self-reported performance scores of the most successful competitors indicate that they possess the greatest strength and power and were most proficient in shorter (<3.5 minutes) CrossFit-specific benchmark workouts. Similarly, higher-ranking competitors report faster 400 m sprint times and better performances during longer benchmark workouts, but the differences in these measures are less clear. In contrast, differences amongst competitors in their reported performance in a 5,000 m run were limited to women. Thus, when training to optimize performance in the CFO, it would be advisable for lower-ranking competitors to place an emphasis on developing their strength and power once a sufficient anaerobic/aerobic endurance-base and sport-specific skills have been established.

Acknowledgments

The authors would like to thank Allyson Box, Emily Bechke, Wade Hoffstetter, and Cassie Williamson for their assistance in data collection.

References

1. Available at: http://games.crossfit.com/leaderboard. Accessed May 1, 2016.
2. Available at: http://games.crossfit.com/workouts/the-open. Accessed May 1, 2016.
3. Babiash, PE. *Determining the Energy Expenditure and Relative Intensity of Two Crossfit Workouts*. La Crosse, WI: University of Wisconsin-La Crosse, 2013.
4. Baker, D, Wilson, G, and Carlyon, R. Periodization: The effect on strength of manipulating volume and intensity. *J Strength Cond Res* 8: 235–242, 1994.
5. Bellar, D, Hatchett, A, Judge, L, Breaux, M, and Marcus, L. The relationship of aerobic capacity, anaerobic peak power and experience to performance in CrossFit exercise. *Biol Sport* 32: 315–320, 2015.
6. Bellovary, B and Drum, SN. *A Performance Profile Related to Building Elite Fitness in Male Competitors*. Presented at Medicine and Science in Sports and Exercise, American College of Sports Medicine, Annual Meeting 2014; Orlando, FL; May 27–31, 2014.
7. Bergstrom, HC, Housh, TJ, Zuniga, JM, Traylor, DA, Lewis, RW Jr, Camic, CL, Schmidt, RJ, and Johnson, GO. Differences among estimates of critical power and anaerobic work capacity derived from five mathematical models and the three-minute all-out test. *J Strength Cond Res* 28: 592–600, 2014.

8. Berthelot, G, Sedeaud, A, Marck, A, Antero-Jacquemin, J, Schipman, J, Saulière, G, Marc, A, Desgorces, FD, and Toussaint, JF. Has athletic performance reached its peak? *Sports Med* 45: 1263–1271, 2015.
9. Brzycki, M. Strength testing–predicting a one-rep max from reps-to-fatigue. *J Phys Educ Recreat Dance* 64: 88–90, 1993.
10. Butcher, SJ, Neyedly, TJ, Horvey, KJ, and Benko, CR. Do physiological measures predict selected crossFit benchmark performance? *Open Access J Sports Med* 6: 241, 2015.
11. Gastin, PB. Energy system interaction and relative contribution during maximal exercise. *Sports Med* 31: 725–741, 2001.
12. Ghannem, M, Lepretre, P-M, Feron, R, and Ghannem, L. The benefits of using short combined aerobic-resistance training versus load match aerobic training for cardiorespiratory responses in coronary heart disease. *Annals of Cardiology and Angiology* 64: 419, 2015.
13. Glassman, G. CrossFit level 1 training guide. *CrossFit J* 6–15, 2011.
14. Haff, GG. Periodization. In: 4th, ed. GG Haff, NT Triplett, eds. *Essentials of Strength Training and Conditioning*. Champaign, IL: Human Kinetics, 2015. pp. 583–604.
15. Hoffman, JR. Principles of Training. In: *Physiological Aspects of Sport Training and Performance*. Champaign, IL: Human Kinetics, 2014. pp. 93–102.
16. Hoffman, JR, Vazquez, J, Pichardo, N, and Tenenbaum, G. Anthropometric and performance comparisons in professional baseball players. *J Strength Cond Res* 23: 2173–2178, 2009.
17. Kraemer, WJ and Fleck, SJ. Periodization of Resistance Training. In: *Optimizing Strength Training: Designing Nonlinear Periodization Workouts*. Champaign, IL: Human Kinetics, 2007. pp. 1–26.
18. Mangine, GT, Hoffman, JR, Vazquez, J, Pichardo, N, Fragala, MS, and Stout, JR. Predictors of fielding performance in professional baseball players. *Int J Sports Physiol Perform* 8: 510–516, 2013.
19. Mangine, GT, Hoffman, JR, Wells, AJ, Gonzalez, AM, Rogowski, JP, Townsend, JR, Jajtner, AR, Beyer, KS, Bohner, JD, and Pruna, GJ. Visual tracking speed is related to basketball-specific measures of performance in NBA players. *J Strength Cond Res* 28: 2406–2414, 2014.
20. Matossian, S. Number 1 & number 2: A day of training. *CrossFit J* 2011.
21. McGee, KJ and Burkett, LN. The national football league combine: A reliable predictor of draft status? *J Strength Cond Res* 17: 6–11, 2003.
22. Ratamess, N, Alvar, B, Evetovich, T, Housh, T, Kibler, W, and Kraemer, W. Progression models in resistance training for healthy adults [ACSM position stand]. *Med Sci Sports Exerc* 41: 687–708, 2009.
23. Rhea, MR, Ball, SD, Phillips, WT, and Burkett, LN. A comparison of linear and daily undulating periodized programs with equated volume and intensity for strength. *J Strength Cond Res* 16: 250–255, 2002.
24. Smith, MM, Sommer, AJ, Starkoff, BE, and Devor, ST. Crossfit-based high-intensity power training improves maximal aerobic fitness and body composition. *J Strength Cond Res* 27: 3159–3172, 2013.
25. Team Series CrossFit. Team Series. Available at: https://athlete.crossfit.com/profile. Accessed May 1, 2016.
26. Vanhatalo, A, Doust, JH, and Burnley, M. Determination of critical power using a 3-min all-out cycling test. *Med Sci Sports Exerc* 39: 548, 2007.
27. Vincent, W and Weir, J. Advanced Statistical Procedures. In: *Statistics in Kinesiology*. Champaign, IL: Human Kinetics, 1999. pp. 301–312.
28. Zagatto, AM, Beck, WR, and Gobatto, CA. Validity of the running anaerobic sprint test for assessing anaerobic power and predicting short-distance performances. *J Strength Cond Res* 23: 1820–1827, 2009.