European Journal of Applied Physiology (2019) 119:1633–1648
https://doi.org/10.1007/s00421-019-04152-3

ORIGINAL ARTICLE

A novel approach for establishing fitness standards for occupational task performance

Jeffrey W. Ryder[1] · Paul Fullmer[2] · Roxanne E. Buxton[3] · J. Brent Crowell[4] · Elizabeth Goetchius[3] · Omar Bekdash[1] · John K. DeWitt[1] · Emma Y. Hwang[1] · Alan Feiveson[5] · Kirk L. English[2,6] · Lori L. Ploutz-Snyder[7]

Received: 21 December 2018 / Accepted: 24 April 2019 / Published online: 8 May 2019

Abstract

Purpose To identify strength and performance thresholds below which task performance is impaired.

Methods A new weighted suit system was used to manipulate strength-to-body-weight ratio during the performance of simulated space explorations tasks. Statistical models were used to evaluate various measures of muscle strength and performance on their ability to predict the probability that subjects could complete the tasks in an acceptable amount of time. Thresholds were defined as the point of greatest change in probability per change in the predictor variable. For each task, median time was used to define the boundary between "acceptable" and "unacceptable" completion times.

Results Fitness thresholds for four space explorations tasks were identified using 23 physiological input variables. Area under receiver operator characteristic curves varied from a low of 0.68 to a high of 0.92.

Conclusion An experimental analog for altering strength-to-body weight combined with a probability-based statistical model for success was suitable for identifying thresholds for task performance below which tasks could either not be completed or time to completion was unacceptably high. These results provide data for strength recommendations for exploration mission ambulatory task performance. Furthermore, the approach can be used to identify thresholds for other areas where occupationally relevant tasks vary considerably.

Keywords Strength · Power · Work · Fitness · Threshold · Task

Communicated by William J. Kraemer.

✉ Jeffrey W. Ryder
jeffrey.ryder-1@nasa.gov

1 KBRwyle Science and Space, 2400 East NASA Parkway, Houston, TX, USA

2 Department of Clinical Health and Applied Sciences, University of Houston–Clear Lake, Houston, TX, USA

3 Department of Health and Human Performance, University of Houston, Houston, TX, USA

4 MEI Technologies, Houston, TX, USA

5 Biomedical Research and Environmental Sciences, National Aeronautics and Space Administration, Johnson Space Center, 2101 NASA Road 1, Houston, TX, USA

6 Exercise and Nutritional Health Institute, University of Houston–Clear Lake, Houston, TX, USA

7 School of Kinesiology, University of Michigan, Ann Arbor, MI, USA

Introduction

Occupations that involve physical work often have minimal fitness requirements; examples of such occupations include police, fire fighters, military personnel, baggage handlers, and construction and warehouse workers. Astronauts also benefit from minimal fitness requirements that can preserve their performance and safety. These requirements are especially important given clear evidence that astronauts lose muscle strength (Day et al. 1995; English et al. 2015; Gopalakrishnan et al. 2010; Tesch et al. 2005; Trappe et al. 2009; Mulavara et al. 2018) and aerobic power (Levine et al. 1996; Moore et al. 2014; Trappe et al. 2006) during spaceflight. The National Aeronautics and Space Administration (NASA) has standard recommendations for astronauts' health and performance that are reevaluated as additional data becomes available to warrant updates. To this end, we performed a ground-based study to determine if various measures of muscle strength can predict whether subjects could complete a variety of tasks in an acceptable amount of time. All the tasks we evaluated in the present study were simulations of tasks that astronauts will have to perform during exploration missions. The greater challenge in establishing fitness standards is often not simply characterizing the relationship between fitness and performance, but rather determining the acceptable cut-off times (performance) for completion of the tasks. We previously determined how a wide range of normalized strength affects performance of simple tasks, with normalized strength being defined as strength per body weight (Ryder et al. 2013). Using a two-slope spline regression analysis, we showed that numerous tasks have a natural physiological breakpoint. That is, when normalized strength is below the breakpoint threshold, the change in task performance time as a factor of strength is greater than when one's normalized strength is above the strength threshold. Presumably, this phenomenon exists because there is a functional reserve that is not challenged when strength is above the threshold value. The practical application of an identified threshold is that the performance time at this threshold can be considered "an acceptable amount of time" in many occupational settings. The rationale here is that being stronger than the threshold is of marginal to no benefit to being able to do the job. The identified human health and performance risk tied to muscle strength at NASA is entitled "risk of impaired performance due to reduced muscle mass, strength, and endurance". In this spirit, a breakpoint whereby task completion time more markedly increases with a decrease in strength can be considered as a threshold for "impaired performance". Here we describe both empirical and statistical models for the establishment of fitness standards for task performance. We implemented this process using a variety of strength and fitness metrics.

Methods

Subjects

Sixty recreationally active, non-astronaut subjects (32 males and 28 females) were recruited through the NASA Johnson Space Center Test Subject Screening Facility. We enrolled subjects with a range in age, weight, height, and fitness that represented the astronaut corps, a diverse population in respect to all of these parameters (English et al. 2015; Hackney et al. 2015). All study procedures were approved by the NASA Johnson Space Center Institutional Review Board, and signed informed consent was obtained from each subject prior to enrollment and testing.

Strategic approach

We sought to identify physiological performance thresholds below which task performance degrades. Unless noted otherwise, each subject completed 9 test and familiarization sessions. Subjects underwent physiological baseline testing, and also performed 4 space exploration-related tasks. Session 1 included isokinetic testing and measurement of VO_{2peak}. During session 2, measures of isometric strength, and isotonic power and total work were assessed. Session 3 included a Wingate anaerobic cycle test and the subjects were familiarized with the task simulations. Session 4 was a second task familiarization with the subject externally loaded with 40% of their body weight. Task performance was assessed in sessions 5 through 9. Each task was performed under 5 different external loading conditions (each on separate days) using a weighted suit (Fig. 1) specifically designed for this investigation (Terrazign, Inc, Portland, OR, USA). Ergonomically shaped weights were applied to the torso, hips, upper and lower arms, and upper and lower legs of the suit. The weighted suit conditions were sham, (suit fabric alone without additional weight added, ~ 10 kg), 20%, 40%, 60%, and 80% of the subject's body weight added as an external load. Weights were distributed anthropometrically across body segments (Barter 1957; Ryder et al. 2013). The first weighted test session never included the 80% body weight loading, but the order of loads was otherwise randomized in a balanced manner. The reported physiological metrics when compared versus the timed tasks are expressed per fully loaded body weight for analysis (i.e., including the weight of the suit). Our goal was to simulate reduced

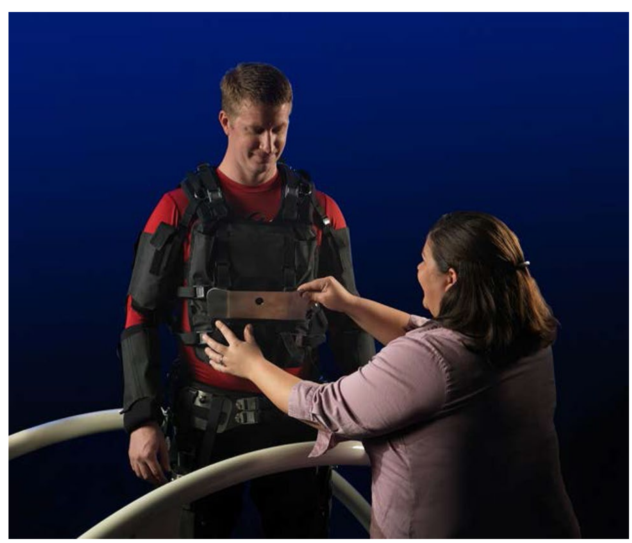

Fig. 1 Image of the weighted suit worn by subjects during the completion of the four exploration tasks

levels of strength-to-body weight, and provide up to five strength-to-body weight observations per subject to feed the statistical model for each task. Subject testing was scheduled in 2–4 h blocks of time within the normal work hours of 08:00–17:00. Subjects were instructed to report to testing in the non-fasted, euhydrated state. Additionally, subjects were asked to abstain from physical exercise for 12 h before testing, and from strenuous or maximal exercise for 24 h before testing. Outside of these restrictions, subjects were instructed to follow their normal schedule with respect to sleep, diet, hydration, and exercise.

Physiological measures (predictor variables)

Isokinetic strength

We obtained all isokinetic measures using a Biodex System 4 Pro isokinetic dynamometer (Shirley, NY, USA). Data collection commenced after the subjects completed a series of warm-up attempts. Three maximal efforts were attempted to determine peak torque; the greatest peak torque is reported. For total work, twenty consecutive maximal efforts were performed. During knee extension and flexion measures, subjects were secured in the upright seated position using shoulder straps. Concentric and eccentric knee extension and flexion measurements were determined at $60°s^{-1}$. Concentric knee extension and flexion total work were determined at $180°s^{-1}$. During calf strength measurements, subjects were secured in a prone position with their knee fully extended. Concentric and eccentric plantar flexion and dorsiflexion was determined at $30°s^{-1}$. During trunk strength measurements, subjects were strapped into a seated upright position in the trunk adaptor and concentric trunk flexion and extension peak torque was measured at $60°s^{-1}$. Isokinetic testing is generally highly reliable. Concentric test–retest intraclass coorelation coefficients (ICC) for each of these movements have been reported to be between 0.89 and 0.99 (Alvares et al. 2015; Feiring et al. 1990; Laughlin et al. 2009; Pinheiro and Chao 2006), with doriflexion (0.67) being an exception (Laughlin et al. 2009).

Isometric strength

Isometric strength was measured as previously described (Ryder et al. 2013; Spiering et al. 2011; Mulavara et al. 2018; Ploutz-Snyder et al. 2014, 2018). After the subject completed 2 submaximal warm-ups, they completed 3 maximal attempts with one minute of rest between each effort. If the 2 highest recorded force measurements differed by more than 5%, a fourth attempt was made. The greatest force recording was used for analysis. Leg-press maximal isometric force was performed at 90° knee flexion on a customized 35° leg press machine (600a, Nebula Fitness Equipment, Versailles, OH, USA) instrumented with a force plate (Type 9281C, Kistler Instrument Corp., Amherst, NY, USA). Custom pin stops were used to secure the position. Knee extension maximal isometric force was measured with the knee at 70° of flexion on a modified Nautilus NT-1220 knee extension device (Nautilus Inc., Vancouver, WA, USA) instrumented with a load cell (Transducer Techniques, Temecula, CA, USA). Bench press maximal isometric force was measured with the bar secured by means of turnbuckles to the safety rails of the bench press cage (Isotronic, Fitness Technologies, Skye, SA, Australia) at the position nearest (≤ 7.5 cm) to the subject's chest. Force was measured by a custom force plate equipped with four load cells (Interface, Scottsdale, AZ, USA) that was positioned beneath the bench. The isometric mid-thigh pull was performed in the bench press cage with the bench removed. Subjects adopted a flat back position with ~ 145° of knee extension as previously described (De Witt et al. 2018). The bar was secured within 2.5 cm of the midpoint between the knee and hip joint. Force data were sampled at 1000 Hz and low-pass filtered at 4 Hz for leg press, bench press and mid-thigh pull. For knee extension, data were collected at 5000 Hz and low-pass filtered at 220 Hz. Results were calculated using a customized LabVIEW (National Instruments, Austin, TX, USA) script. Test–retest ICC for leg press, bench press, and mid-thigh pull are 0.99, 0.99, and 0.89, respectively (De Witt et al. 2018; Spiering et al. 2011).

Isotonic power and work

For leg press (Mulavara et al. 2018; Ploutz-Snyder et al. 2014, 2018; Ryder et al. 2013; Spiering et al. 2011) and bench press (Mulavara et al. 2018; Ploutz-Snyder et al. 2018, 2014; Ryder et al. 2013; Spiering et al. 2011) peak power production and total work, the pin stops were removed from the leg press sled and the turnbuckles were removed from the bench press bar to allow freedom of movement of the load. Free weights were added to achieve loads corresponding to 40% and 30% of maximal leg press and bench press isometric force, respectively. Systems were instrumented with a position transducer (PT5A, Fitness Technologies, Skye, SA, Australia) and outfitted with a magnetic brake (Fitness Technologies, Skye, SA, Australia) that offloaded the return so that only concentric actions could be performed. Subjects were instructed to perform 21 maximal repetitions as rapidly as possible with no rest between attempts. Force and position data were sampled at 300 Hz and low-pass filtered at 4 Hz. Peak instantaneous power was determined from the single best repetition, and total work was calculated as the integral of all 21 repetitions in the set; each test utilized a customized LabVIEW script. Test–retest reliability ICC for leg press power and work are 0.99 each, whereas bench press power and work ICC are 0.97, and 0.99, respectively (Spiering et al. 2011).

Countermovement vertical jump power

Lower-extremity power was determined by a countermovement jump test (Ploutz-Snyder et al. 2018; Rittweger et al. 2007). Subjects stepped onto a force plate with their hands positioned on their hips and remained still. On command from a test operator, the subjects jumped vertically as high as possible off of the force plate while keeping their hands on their hips. Force data were collected at 1000 Hz and peak power was calculated using a customized LabVIEW script. The countermovement vertical jump power test–retest correlation coefficient has been reported to be 0.99 (Rittweger et al. 2004).

Wingate power

Anaerobic power was determined by Wingate test on a Monark Ergomedic 894 E cycle ergometer (Monark Exercise AB, Vansbro, SE). Subjects warmed up for approximately 5 min at a self-selected submaximal load. Once warmed up, subjects stopped pedaling and the ergometer flywheel was brought to a complete stop. A load equal to 7.5% bodyweight was added to the load basket, and the basket was raised to the unloaded position. Subjects pedaled to maximal cadence with no load on the flywheel, at which point the test load was applied to the flywheel. Subjects pedaled as fast as possible for 30 s. Five second peak power was determined using the manufacturer's software. Wingate test–retest peak power ICC is reported to be 0.91 (Ozkaya 2013).

Aerobic capacity

Peak aerobic capacity was determined during a graded exercise test on a Lode Excalibur cycle ergometer (Lode, Goningen, The Netherlands). Subjects began cycling for 3 min at 50 W. Thereafter, the load was increased by 25 W each min until volitional fatigue. Metabolic gas analysis was determined with the Oxycon mobile metabolic gas analysis system (Carefusion, Houten, The Netherlands).

Exploration task simulations

Tasks were performed in the weighted suit as described above. The order of the tasks was kept constant across sessions/loading conditions: (1) capsule egress, (2) ambulation and supply transfer, (3) rescue drag, (4) and hill climb and descent. During the capsule egress task, the subject wore an additional safety harness (3M Protecta PRO full-body, St. Paul, MN, USA) that was connected to a self-retracting lifeline (3M Protecta SRL Rebel, St. Paul, MN, USA) by a test operator during the egress task as a fall safety measure.

Capsule egress

An unaided top-hatch emergency egress task (Fig. 2a, b) was performed in a conical space capsule mock-up. For the purposes here, the general dimensions of the capsule and internal components such as the habitable space, hatches, and top-hatch ladder were similar to those for spacecraft under development for future space exploration (floor, 3.6 m diameter; ceiling, 2.5 m diameter; ceiling, 1.5 m height from floor; top hatch tunnel, 83 cm length, 81 cm diameter at narrowest point). The mock-up contained four seats 20 cm off of the floor; two upper seats positioned next to each other under a control console mock-up, and two positioned next to each other at the foot of the other seats. Each subject was positioned in a supine seated position in the upper left-hand seat and secured in a 5-point seat belt system. The display console was 61 cm from the seat back at its closest point. The task was performed in the hatch open position.

On command, the subject unbuckled the seat harness and egressed the capsule through the top hatch as quickly as possible. After rising from the seat, the subject navigated to the opposite side of the capsule where mock ups for an un-deployed life raft (13.6 kg) and two crew survival packs (6.4 kg each) were stowed. The raft and survival packs were moved to the center of the capsule below the edge of the top-hatch/escape tunnel. Subjects next deployed the egress ladder that was mounted, rolled up, and stowed midway up the top hatch tunnel. The ladder was made of six metal rungs connected by webbing. Deployment consisted of removing a fabric cover, unrolling the ladder, securing the feet of the ladder to eyebolts in the capsule floor, and cinching the ladder tight. While the subject deployed and secured the ladder, a test operator attached the SRL clasp to the subjects safety harness in a non-interfering manner. Once the ladder was secure, subjects lifted and pushed the life raft and survival packs out the top hatch and climbed out. The task ended when the subject had both feet on the safety platform that surrounded the outside of the top hatch. Timing was triggered manually by test operators using a button system with timing buttons inside of the capsule and on the top hatch platform.

Ambulation and supply transfer

The ambulation and supply transfer task (Fig. 2c) simulated traversing a planetary surface from a crew lander to a supply lander followed by moving supplies in simulated crew transfer bags (CTB) to a rover. The task was performed on a 50 m quasi-teardrop shaped track that was built to simulate a regolith-like surface. The track was

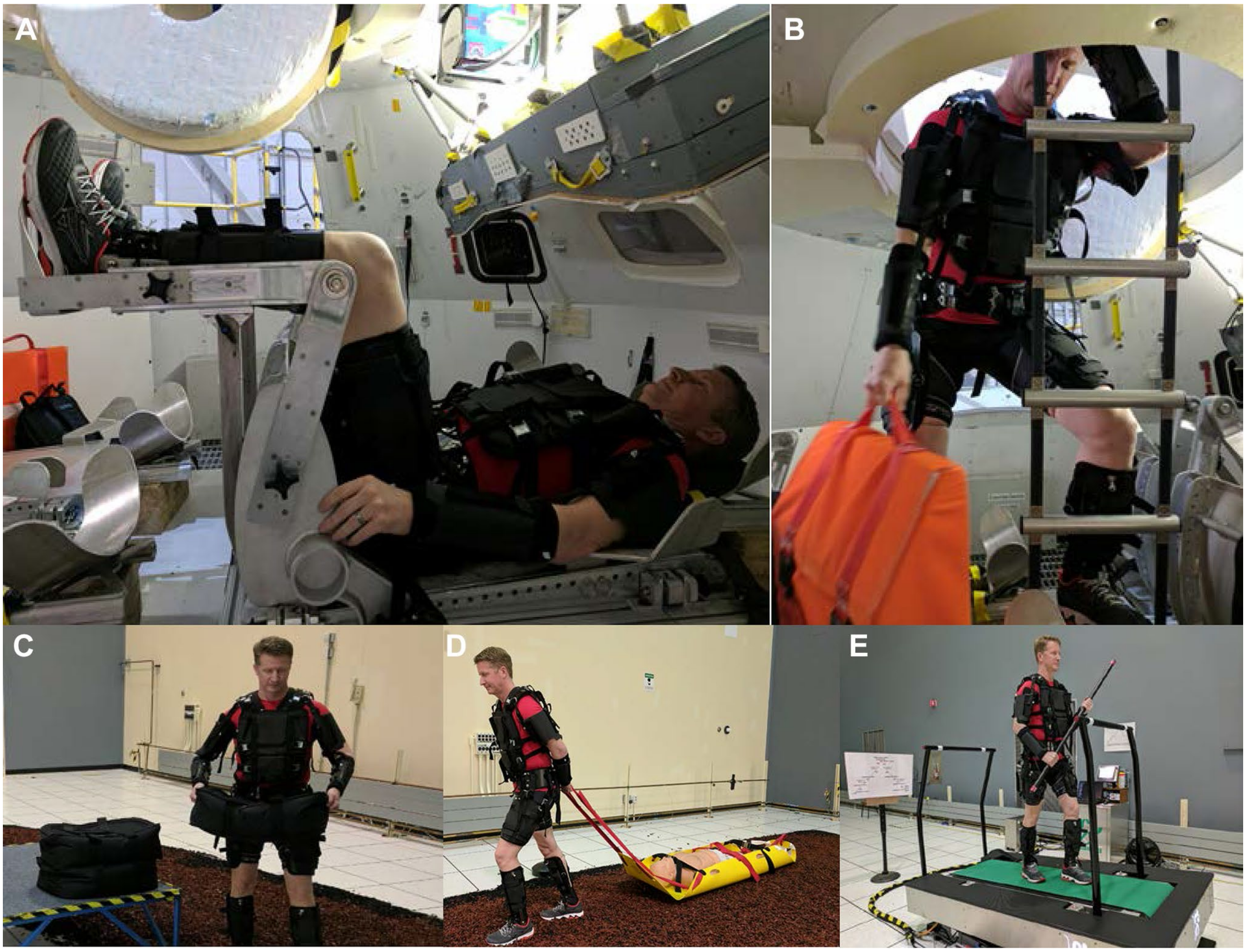

Fig. 2 Images of the tasks performed by the subjects **a, b** capsule egress, **c** ambulation and supply transfer, **d** rescue drag, **e** hill climb and descent

constructed on plastic sheets placed on the indoor test facility floor. The track include three small undulations (up to 15 cm) constructed from building lumber and plywood. The wooden structures were placed on the floor under the plastic sheets. A layer of "shredded" landscaping rubber mulch was glued to the plastic sheets, topped by a 2–3 cm layer of loose "nugget" -shaped landscaping rubber mulch.

The ambulation and supply task started with 30 laps around the track (i.e., 1.5 km) followed by transferring 30 CTB a distance of 5 m from one platform to another using the two CTB side handles. The CTB were transferred one at a time. The bags weighed 10.9 kg each, with dimensions of 60 × 36 × 20 cm. Following bag transfers the subject completed the task by crossing the start/finish line. The subject was instructed to complete the task as quickly as possible without running. Timing was measured using photo sensors placed along the track.

Rescue drag

The rescue drag task (Fig. 2d) involved the simulation of securing an incapacitated crewmember onto a rescue sled and dragging him/her a distance of 50 m to safety. The task was performed on the same track that was used for the ambulation and supply transfer task. A 75 kg rescue mannequin ("Rescue Randy"; Sauregerties, NY, USA) was placed in a prone position next to a plastic rescue sled (Skedco, Tualatin, OR, USA). Subjects were required to roll the mannequin over onto the sled into a supine position, fasten the buckles on the three safety restraints, cinch the restraints tight, and drag the mannequin 50 m around the track. Timing was measured using photo sensors placed along the track.

Hill climb and descent

The hill climb and descent (Fig. 2e) simulated a round-trip excursion to deploy a line-of-sight communications antenna. The trip profile included a 1000 m horizontal distance with a variable rise totaling 40 m (2–8%, 4% average grade) followed by a 10 m flat summit. This task utilized a Treadmetrix 3DI treadmill (Treadmetrix, Park City, UT, USA) which facilitated the simulation of an uphill trek. Subjects walked "uphill" carrying a 4.1 kg antenna mock up (Bodybar, Boulder, CO, USA). Upon reaching the top of the hill the treadmill was stopped and the subject handed the bar to a test operator. The subject then turned around and completed the downhill portion of the task (total round trip distance of 2020 m). The treadmill was instrumented with force plates, and custom software was installed by the manufacturer that allowed for treadmill belt velocity to be controlled by the walking speed of the subject via changes in location on the treadmill belt and center of pressure. Grade changes were controlled manually by a test operator and followed a prescribed profile that was the same for all subjects and sessions. Timing was triggered manually by a test operator using a button system.

Statistical modeling

Data

The data for this study consisted of (a) total completion times for 4 main types of simulated exploration mission tasks and (b) muscle-strength and fitness metrics for 23 targeted variables. In the original experiment design, each of 60 subjects was to perform all tasks in five different loading conditions and each of the baseline fitness measures. Between 273 and 290 of the possible 300 (60 subjects × 5 sessions) sets of measurements were actually available for analysis, depending on the particular combination of task/sub-task and strength/fitness variables being modeled (see analysis). Dropouts from the maximal possible are accounted by: three subjects did not complete all task performance sessions, one subject was excluded from the hill climb and descent task related to walking on the self-paced treadmill safely, one subject did not complete one hill climb and descent task for reasons unrelated to muscle performance, two subjects did not complete the isometric knee extension test due to limb length, one subject did not complete the Wingate test (technical issue, not rescheduled), one subject did not complete the eccentric isokinetic dorsiflexion (technical issue, not rescheduled), timing system malfunctions (one loss each for total task time for capsule egress, ambulation and supply transfer, and hill climb and descent).

Analysis approach

The general analysis approach was to first define criteria for successful completion of each task or sub-task and then model the probability of success as a function $P(x)$ of each exercise metric x. Given this probability model, we then defined exercise performance thresholds as x_{M}, the value of x such that $P'(x)$ is maximum. Without loss of generality, assume that larger values of x correspond to higher strength. Then the rationale for choosing the threshold in this way is that for $x > x_{\mathrm{M}}$ there is a diminishing return in terms of increased probability of success per unit increase in x.

Success criteria

Performance metrics for all tasks are in terms of time to complete. Recognizing that these tasks are meant to be typical of what astronauts might have to do in exploration missions, we defined "success" for each task (or sub-task) as being able to complete the task (or sub-task) with a time no greater than the median completion time as observed in this study. In calculating median completion time, we took into account that in some cases subjects failed to complete a task. For purposes of calculating the median, these failures were regarded as arbitrarily large completion times, so that the median is a value, say y_0, such that approximately 50% of the trials for a task had completion times less than or equal to y_0, and 50% of trials either had completion times greater than y_0 or the task was not completed.

Probability model

Let x_{ij} denote the observed exercise metric with the j-th weight handicap for one of the exercises in Table 1 performed by the i-th subject. Additionally, let y_{ij} be the task or sub-task completion time for the same subject with the same weighted condition. Then using the above definition of success, we defined the indicator of success S_{ij}, by $S_{ij} = 1$ if $y_{ij} \le y_0$; otherwise $S_{ij} = 0$. Note $S_{ij} = 0$ includes the event that the subject fails to complete the task as well as the event $y_{ij} > y_0$. We then used mixed-effects logistic regression (Pinheiro and Chao 2006) to model $P\left(S_{ij} = 1 | x_{ij} = x\right)$. This model takes into account the repeated-measures design, allowing for correlation between the trials (j) for the same subject. More specifically, the model is

$$P\left(S_{ij} = 1 | x_{ij} = x\right) = \frac{\exp\left(\beta_0 + u_{0i} + \beta_1 x\right)}{1 + \exp\left(\beta_0 + u_{0i} + \beta_1 x\right)}, \tag{1}$$

Table 1 Characteristics of subjects' physiological performance

Physiological measure	Mean ± SD	Range
Age (years)	37 ± 7	24–52
Weight (kg)	75.0 ± 13.3	52.7–112.2
Height (m)	1.73 ± 0.10	1.52–1.93
Isokinetic Knee extension—Con (Nm kg^{-1})	2.56 ± 0.47	1.66–3.53
Isokinetic Knee extension—Ecc (Nm kg^{-1})	3.35 ± 0.72	1.52–5.01
Isokinetic Knee flexion—Con (Nm kg^{-1})	1.29 ± 0.26	0.86–2.01
Isokinetic Knee flexion—Ecc (Nm kg^{-1})	1.95 ± 0.35	1.15–2.98
Isokinetic knee extension Work (J kg^{-1})	32.09 ± 5.78	20.88–44.61
Isokinetic knee flexion work (J kg^{-1})	17.85 ± 3.50	10.73–26.09
Isokinetic plantar flexion—Con (Nm kg^{-1})	1.68 ± 0.30	1.05–2.19
Isokinetic plantar flexion—Ecc (Nm kg^{-1})	2.45 ± 0.55	1.12–3.58
Isokinetic dorsiflexion—Con (Nm kg^{-1})	0.45 ± 0.08	0.32–0.75
Isokinetic dorsiflexion—Ecc (Nm kg^{-1})	0.80 ± 0.11	0.56–1.06
Isokinetic trunk extension (Nm kg^{-1})	5.27 ± 1.04	3.04–8.03
Isokinetic trunk flexion (Nm kg^{-1})	2.60 ± 0.51	1.53–3.51
Isometric leg press (N kg^{-1})	24.12 ± 4.56	15.84–33.31
Leg press power (W kg^{-1})	20.61 ± 4.35	10.67–28.23
Leg press work (J kg^{-1})	104.56 ± 24.25	58.28–161.97
Isometric bench press (N kg^{-1})	8.98 ± 3.04	4.24–16.34
Bench press power (W kg^{-1})	8.71 ± 2.78	3.94–15.30
Bench press work (J kg^{-1})	35.61 ± 12.61	17.82–57.73
Isometric mid thigh pull (N kg^{-1})	25.50 ± 5.51	14.00–37.79
Isometric knee extension (N kg^{-1})	7.68 ± 1.78	4.84–12.67
Vertical jump power (W kg^{-1})	43.50 ± 8.57	27.80–61.85
Wingate power (W kg^{-1})	8.03 ± 1.62	4.47–11.16
Aerobic capacity (ml kg^{-1} min^{-1})	38.37 ± 7.44	25.80–59.70

where β_0 and β_1 are logistic regression coefficients, and $u_{0i} \sim N\left(0, \sigma_u^2\right)$ is a normally distributed perturbation to β_0 for the i-th subject. The model given by Eq. (1) is specific to the i-th subject, but to make it general, we have to average over the possible values of the u_{0i}:

$$P(x) = \int_{-\infty}^{\infty} \frac{1}{\sigma_u} \phi\left(u/\sigma_u\right) \frac{\exp\left(\beta_0 + u + \beta_1 x\right)}{1 + \exp\left(\beta_0 + u + \beta_1 x\right)} \mathrm{d}u \tag{2}$$

where $\phi(\cdot)$ is the standard normal probability density function. In Eq (2). $P(x)$ is the probability that a subject chosen at random is able to complete the task in a time $\leq y_0$ given the exercise metric for that subject is x.

Exercise threshold

After fitting the model given by Eq (1), we found the threshold as x_M, as the value of x such that $|P'(x)|$ is maximum. For the special case of the logistic conditional model in Eq (1), the maximum of $|P'(x)|$ occurs when $P(x) = 1/2$. Furthermore, because the distribution of u_{0i} is symmetric, this occurs when $\beta_0 + \beta_1 x = 0$. As a result, the threshold is

$$x_M = -\beta_0/\beta_1, \tag{3}$$

however, for non-linear models in the logistic metric or for non-logistic models, this may not be the case. The standard error of x_M was obtained using the delta method (Greene 2000) applied to the variance-covariance matrix of the estimates of β_0 and β_1.

Comparing predictors

An important objective of this study was to compare or rank the thresholds as potential standards of fitness to perform the tasks. To do this we calculated three criteria: (1) non-parametric receiver operator characteristic (ROC) area, (2) the normalized magnitude of $P'\left(x_M\right)$, and (3) $\frac{x_M}{SE(x_M)}$, the relative accuracy of the estimate of x_M. The ROC area (Hanley and McNeil 1982) is the area under the curve of sensitivity vs. $1 -$ specificity, where for a given value of x, say X, the sensitivity is the fraction of observations with $x > X$ out of those in which the task was successfully performed ($y < y_0$). By contrast, the specificity is the fraction of observations with $x < X$ out of those in which the task was not successfully performed ($y > y_0$ or failure to complete). The closer the ROC area is to 1.0, the better the predictor. Normalization of $|P'(x_M)|$ was made by multiplying by the interquartile range of x resulting in an index of average steepness of $P(x)$ over this range. The concept here being that better predictors should have greater average steepness. The third criterion (relative accuracy) reflects variability and/or a lack of a useful relationship between y and x. The smaller this criterion is, the better the predictor.

Results

Physiological characteristics for the 60 subjects in this study are shown in Table 1. Included are measures of strength and performance normalized to body weight under shirt-sleeve conditions (i.e., wearing no special suit or equipment). Threshold values for muscle strength, power, and work capacity performance metrics (predictor variables) versus total completion time are shown in Tables 2, 3, 4, 5 for capsule egress, ambulation and supply transfer, rescue drag, and hill climb and descent tasks, respectively. Plots for select predictor variables versus the time to compete capsule egress are shown in Fig. 3. No meaningful differences between threshold values based on total time (Tables 2, 3, 4, 5) and those based on sub-task (intermediate-task) times were observed, therefore, the subtask thresholds are not shown. The number of unsuccessful attempts for each of the task were as follows: capsule egress 2 of 289 trials, ambulation and supply transfer 7 of 290 trials, rescue drag 2 of 290 trials, and hill climb and descent 28 of 283 trials. Indices describing the quality of the predictive nature of each of the variables for the four tasks are presented in Tables 2, 3, 4, 5. These indices include ROC area under the curve, the percent relative accuracy of the estimate of x_M (SDrat), and the normalized magnitude of the probability slope at the threshold $P'(x_M)$ (PpNorm). Each of these indices are designed to compare predictabilities across the various strength measures independent of differences in units. ROC plots for the scatter plot comparisons presented in Fig. 3 are shown in Fig. 4.

Table 2 Physiological performance thresholds for capsule egress

Predictor variable	Threshold	ROC	SDrat	PpNorm
Isokinetic knee extension—Con (Nm kg^{-1})	1.83	0.87	8.01	0.70
Isokinetic knee extension—Ecc (Nm kg^{-1})	2.38	0.84	9.35	0.59
Isokinetic knee flexion—Con (Nm kg^{-1})	0.93	0.83	9.91	0.57
Isokinetic knee flexion—Ecc (Nm kg^{-1})	1.40	0.80	11.62	0.48
Isokinetic knee extension work (J kg^{-1})	22.90	0.87	8.12	0.68
Isokinetic knee flexion work (J kg^{-1})	12.78	0.84	9.66	0.58
Isokinetic plantar flexion—Con (Nm kg^{-1})	1.21	0.83	9.40	0.59
Isokinetic plantar flexion—Ecc (Nm kg^{-1})	1.76	0.83	9.59	0.58
Isokinetic dorsiflexion—Con (Nm kg^{-1})	0.33	0.72	15.42	0.37
Isokinetic dorsiflexion – Ecc (Nm kg^{-1})	0.58	0.76	12.71	0.44
Isokinetic trunk extension (Nm kg^{-1})	3.79	0.83	10.39	0.54
Isokinetic trunk flexion (Nm kg^{-1})	1.88	0.78	10.84	0.52
Isometric leg press (N kg^{-1})	17.36	0.80	11.10	0.50
Leg press power (W kg^{-1})	14.82	0.86	9.21	0.61
Leg press work (J kg^{-1})	74.73	0.85	9.05	0.62
Isometric bench press (N kg^{-1})	6.40	0.84	8.62	0.66
Bench press power (W kg^{-1})	6.22	0.82	9.36	0.61
Bench press work (J kg^{-1})	25.26	0.85	7.66	0.76
Isometric mid thigh pull (N kg^{-1})	18.12	0.88	7.89	0.70
Isometric knee extension (N kg^{-1})	5.44	0.85	9.08	0.63
Vertical jump power (W kg^{-1})	31.16	0.84	9.36	0.60
Wingate power (W kg^{-1})	5.74	0.82	10.64	0.53
Aerobic capacity (ml kg^{-1} min^{-1})	27.57	0.84	9.82	0.57

Con concentric, *Ecc* eccentric, *ROC* receiver operator characteristic area under the curve, *SDrat* percent relative accuracy of the estimate of x_M, *PpNorm* the normalized magnitude of the probability slope at the threshold $P'(x_M)$

Table 3 Physiological performance thresholds for ambulation and supply transfer

Predictor variable	Threshold	ROC	SDrat	PpNorm
Isokinetic knee extension—Con (Nm kg^{-1})	1.83	0.88	6.41	0.84
Isokinetic knee extension—Ecc (Nm kg^{-1})	2.41	0.82	8.09	0.67
Isokinetic knee flexion—Con (Nm kg^{-1})	0.93	0.84	8.22	0.65
Isokinetic knee flexion—Ecc (Nm kg^{-1})	1.41	0.83	8.91	0.60
Isokinetic knee extension Work (J kg^{-1})	22.98	0.89	6.51	0.83
Isokinetic knee flexion work (J kg^{-1})	12.85	0.85	7.84	0.69
Isokinetic plantar flexion—Con (Nm kg^{-1})	1.21	0.84	7.71	0.69
Isokinetic plantar flexion—Ecc (Nm kg^{-1})	1.77	0.85	7.43	0.73
Isokinetic dorsiflexion—Con (Nm kg^{-1})	0.33	0.80	10.23	0.52
Isokinetic dorsiflexion—Ecc (Nm kg^{-1})	0.57	0.83	8.39	0.64
Isokinetic trunk extension (Nm kg^{-1})	3.80	0.84	8.28	0.65
Isokinetic trunk flexion (Nm kg^{-1})	1.88	0.81	8.23	0.65
Isometric leg press (N kg^{-1})	17.51	0.78	9.64	0.56
Leg press power (W kg^{-1})	14.91	0.84	8.28	0.65
Leg press work (J kg^{-1})	75.55	0.82	8.49	0.63
Isometric bench press (N kg^{-1})	6.57	0.75	8.66	0.65
Bench press power (W kg^{-1})	6.41	0.75	8.87	0.63
Bench press work (J kg^{-1})	26.11	0.76	7.95	0.71
Isometric mid thigh pull (N kg^{-1})	18.36	0.84	7.82	0.69
Isometric knee extension (N kg^{-1})	5.43	0.85	7.92	0.69
Vertical jump power (W kg^{-1})	31.46	0.80	8.61	0.62
Wingate power (W kg^{-1})	5.78	0.83	8.54	0.64
Aerobic capacity (ml kg^{-1} min^{-1})	27.40	0.88	6.78	0.79

Con concentric, *Ecc* eccentric, *ROC* receiver operator characteristic area under the curve, *SDrat* percent relative accuracy of the estimate of x_M, *PpNorm* the normalized magnitude of the probability slope at the threshold $P'(x_M)$

Discussion

For many tasks, specific strength (i.e., strength/body weight) is of greater relevance to performance time than is absolute strength. This is especially true of tasks and activities that involve ambulation or moving oneself about in opposition to gravity; e.g., standing up, walking, and climbing. We, therefore, chose to characterize task performance time as a function of relative strength capabilities for four space-exploration relevant tasks. The goal here was to identify thresholds that could be used as physiological capability standards that are linked to task performance. The threshold would demarcate a point whereby dropping further below the threshold results in markedly longer task performance times, while increasing beyond the threshold provides only marginal gains. Two challenges to characterizing such relationships are recruiting enough subjects to generate an adequate number of data points for curve generations, and having a subject pool that includes a wide strength spectrum. To this end we implemented an experimental weighted-suit model approach in which we applied five excess weight conditions under which each subject would implement the task performance portion of the study. This approach helped to generate a wider spread of specific strength values than what would otherwise have been possible to obtain. For example, when a subject was wearing the suit with 80% of his/her body weight applied as the external load, the resulting specific strength metric dropped to 45% of the original shirt-sleeve specific strength. This gave us reasonable assurance that the specific-strength threshold values that delineate the boundaries between acceptable an unacceptable task performance would be captured within the range of the observed values. The present weighted suit approach also served to increase the number of observations by up to five times the number of subjects that enrolled in the study.

Occupational task performance is undoubtedly influenced by a multitude of factors such as age, height, nutrition, hydration status, and sleep. To this end, we used the current weighted suit approach, which utilizes repeated test trials for each subject at different strength-to-body weight ratios. This control mechanism within subjects along with instructions to subjects to report in a non-fasted, euhydrated state with no prior exercise within the stipulated time periods, helps give confidence that observed differences in task performance times are mainly attributable to differences in muscle performance capability. The approach we describe

Table 4 Physiological performance thresholds for crewmember rescue

Predictor variable	Threshold	ROC	SDrat	PpNorm
Isokinetic knee extension—Con (Nm kg^{-1})	1.83	0.79	17.87	0.32
Isokinetic knee extension—Ecc (Nm kg^{-1})	2.40	0.78	17.40	0.33
Isokinetic knee flexion—Con (Nm kg^{-1})	0.93	0.78	19.56	0.30
Isokinetic knee flexion—Ecc (Nm kg^{-1})	1.40	0.76	21.16	0.28
Isokinetic knee extension Work (J kg^{-1})	22.84	0.82	15.64	0.37
Isokinetic knee flexion work (J kg^{-1})	12.75	0.81	16.76	0.34
Isokinetic plantar flexion—Con (Nm kg^{-1})	1.21	0.73	21.69	0.27
Isokinetic plantar flexion—Ecc (Nm kg^{-1})	1.76	0.78	18.26	0.32
Isokinetic dorsiflexion—Con (Nm kg^{-1})	0.33	0.68	29.69	0.20
Isokinetic dorsiflexion—Ecc (Nm kg^{-1})	0.58	0.72	23.02	0.26
Isokinetic trunk extension (Nm kg^{-1})	3.78	0.79	19.22	0.30
Isokinetic trunk flexion (Nm kg^{-1})	1.90	0.67	24.61	0.24
Isometric leg press (N kg^{-1})	17.30	0.74	21.40	0.27
Leg press power (W kg^{-1})	14.80	0.82	17.02	0.34
Leg press work (J kg^{-1})	73.81	0.87	11.79	0.49
Isometric bench press (N kg^{-1})	6.39	0.88	10.79	0.54
Bench press power (W kg^{-1})	6.26	0.83	13.37	0.44
Bench press work (J kg^{-1})	25.08	0.90	8.56	0.69
Isometric mid thigh pull (N kg^{-1})	18.20	0.82	15.48	0.37
Isometric knee extension (N kg^{-1})	5.34	0.79	16.97	0.35
Vertical jump power (W kg^{-1})	31.12	0.81	16.55	0.35
Wingate power (W kg^{-1})	5.68	0.84	15.74	0.37
Aerobic capacity (ml kg^{-1} min^{-1})	27.41	0.75	20.44	0.28

Con concentric, *Ecc* eccentric, *ROC* receiver operator characteristic area under the curve, *SDrat* percent relative accuracy of the estimate of x_M, *PpNorm* the normalized magnitude of the probability slope at the threshold $P'(x_M)$

here also provides a methodological path forward for future determinations of occupational strength standards where critical factors need to be included into the testing design to be most relevant to occupations with highly specific physical, physiological, or environmental constraints. An example may be (de)hydration status when developing strength standards for individuals deployed to desert regions.

The selection of benchmark tasks was driven by the aim of providing recommendations for minimum strength requirements for space explorations missions. Upcoming missions will involve capsule-like crewed vehicles that land in the water upon return to Earth. Some of these missions will be shorter "shake out" missions, while the longest could include multi-year missions to Mars and back. While a full complement of landing support crews are expected to be available for splashdown returns on Earth, off-nominal landing scenarios are always a possibility and characterizing human strength needs for unassisted egress is a vital part of an overall plan to ensure crew safety and mission success. For this reason, we included simulated capsule egress as part of the task battery. The other three tasks that were selected focused on planetary surface extravehicular activities (EVA) for Mars or the moon. Currently, EVA operations are only notional, therefore, our EVA-like tasks were designed to capture the most essential and probable components of nominal and contingency scenarios. Most planetary EVA will involve basic ambulation. Furthermore, lifting and carrying of objects at some point in time is a reasonable assumption. The ambulation and supply transfer task incorporated these activities based on a scenario whereby Mars astronauts have to traverse by foot from their crew lander to a pre-deployed supply lander where they retrieve supplies and load them onto a nearby rover delivered on the supply lander. The rescue drag task was a mock-up based on a need to be able to move an incapacitated crewmember to a rover from a location that the rover could not access. Lastly, communication will be essential to the success of planetary exploration and deployment of line-of-sight communications hardware on high ground will help enable this capability. Our hill climb and descent task characterized the physical capability requirements of traversing a hill with a communications antenna and returning down the hill.

Perhaps the most striking finding here is the uniformity of the threshold values across the four tasks. This similarity is systematic throughout the entire list of predictor variables. Some example ranges of the

Table 5 Physiological performance thresholds for hill climb and descent

Predictor variable	Threshold	ROC	SDrat	PpNorm
Isokinetic knee extension—Con (Nm kg^{-1})	1.86	0.87	6.68	0.80
Isokinetic knee extension—Ecc (Nm kg^{-1})	2.44	0.80	8.40	0.64
Isokinetic knee flexion—Con (Nm kg^{-1})	0.94	0.86	7.26	0.74
Isokinetic knee flexion—Ecc (Nm kg^{-1})	1.43	0.83	8.30	0.64
Isokinetic knee extension work (J kg^{-1})	23.34	0.86	6.97	0.77
Isokinetic knee flexion work (J kg^{-1})	12.98	0.87	6.97	0.77
Isokinetic plantar flexion—Con (Nm kg^{-1})	1.22	0.88	6.44	0.83
Isokinetic plantar flexion—Ecc (Nm kg^{-1})	1.78	0.85	7.24	0.74
Isokinetic dorsiflexion—Con (Nm kg^{-1})	0.33	0.85	8.39	0.64
Isokinetic dorsiflexion—Ecc (Nm kg^{-1})	0.58	0.88	6.53	0.82
Isokinetic trunk extension (Nm kg^{-1})	3.84	0.85	7.74	0.69
Isokinetic trunk flexion (Nm kg^{-1})	1.89	0.84	7.02	0.76
Isometric leg press (N kg^{-1})	17.65	0.81	8.55	0.63
Leg press power (W kg^{-1})	15.10	0.84	7.73	0.70
Leg press work (J kg^{-1})	76.76	0.82	8.02	0.67
Isometric bench press (N kg^{-1})	6.67	0.74	8.47	0.69
Bench press power (W kg^{-1})	6.49	0.74	8.81	0.66
Bench press work (J kg^{-1})	26.50	0.76	7.77	0.76
Isometric mid thigh pull (N kg^{-1})	18.53	0.85	7.25	0.74
Isometric knee extension (N kg^{-1})	5.54	0.82	8.30	0.65
Vertical jump power (W kg^{-1})	31.86	0.80	8.13	0.66
Wingate power (W kg^{-1})	5.83	0.84	7.53	0.72
Aerobic capacity (ml kg^{-1} min^{-1})	27.51	0.92	5.26	1.03

Con concentric, Ecc eccentric, *ROC* receiver operator characteristic area under the curve, *SDrat* percent relative accuracy of the estimate of x_M, *PpNorm* the normalized magnitude of the probability slope at the threshold $P'(x_M)$

thresholds over the four tasks (Tables 2, 3, 4, 5) are concentric isokinetic knee extension (1.83–1.86 Nm kg^{-1}), isometric leg press (17.30–17.65 N kg^{-1}), isometric bench press (6.39–6.67 N kg^{-1}), and isometric mid-thigh pull (18.12–18.53 N kg^{-1}). Similar consistency was also observed for Wingate anaerobic power (5.68–5.83 W kg^{-1}) and aerobic capacity (27.40–27.57 ml kg^{-1} min^{-1}). Our previous work identified natural breakpoints for strength versus time during shorter tasks using a two-slope spline regression approach (Ryder et al. 2013). Interestingly, thresholds that were obtained for tasks most similar to those in the present study were close to the current ones despite a different mathematical approach and a less matured weighted suit (i.e., a less ergonomic, piecemeal suit, comprised of commercially available fitness equipment) than used previously (Ryder et al. 2013). For example, for the (previous) supine and upright egress and walk tasks, threshold values for isokinetic knee extension were 1.8 and 1.9 Nm kg^{-1}, respectively (Ryder et al. 2013). These values compare closely to the capsule-egress threshold of 1.83 Nm kg^{-1} established in the present investigation (Table 2). This held true for most predictor measures that were assessed in both studies including isometric knee extension (5.7 N kg^{-1} supine, 5.9 N kg^{-1} upright, 5.44 N kg^{-1} capsule egress), leg press strength (13.7 N kg^{-1} supine, 17.8 N kg^{-1} upright, 17.36 N kg^{-1} capsule egress), leg press power (14.6 W kg^{-1} supine, 17.6 W kg^{-1} upright, 14.82 W kg^{-1} capsule egress), and leg press work (72.4 J kg^{-1} supine, 78.8 J kg^{-1} upright, 74.73 J kg^{-1} capsule egress). Collectively, results from these studies provide confidence in the identified thresholds as physiological breakpoints below which performance time is impaired for tasks that involve moving oneself through space working against gravity. While our aim was to determine thresholds for exploration tasks for astronauts, the uniformity of the thresholds across tasks suggests that these strength thresholds can apply more broadly to other occupational areas as well.

We included a wide variety of physiological predictor variables into the study design. The purpose of this was to include the types of measures that have been previously used in spaceflight research and operations (Day et al. 1995; English et al. 2015; Gopalakrishnan et al. 2010; Mulavara et al. 2018; Tesch et al. 2005; Trappe et al. 2009), or could be in the future. Our goal was to identify which metrics best (or satisfactorily) predict task success and could thereby reasonably be used as recommendations for fitness thresholds

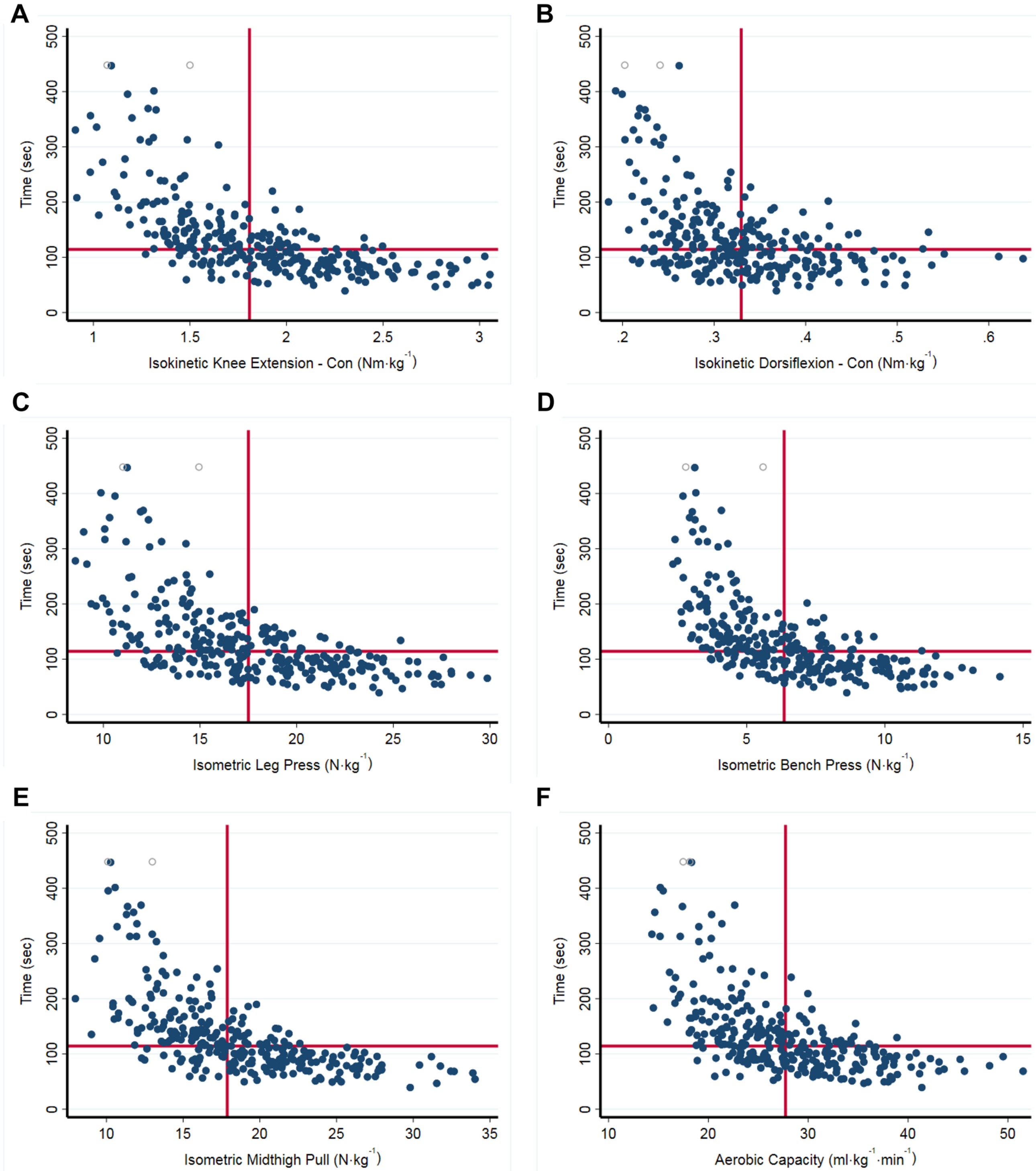

Fig. 3 Scatter plots of **a** concentric isokinetic knee extension strength, **b** concentric isokinetic dorsiflexion strength, **c** isometric leg press strength, **d** isometric bench press strength, **e** isometric mid thigh pull strength, and **f** aerobic capacity versus completion times for the capsule egress task. Horizontal lines represent the median completion time. Vertical lines represent the threshold strength or fitness values. Filled circles represent data for a completed task. Open circles represent an uncompleted task and are plotted equal to the slowest completed time

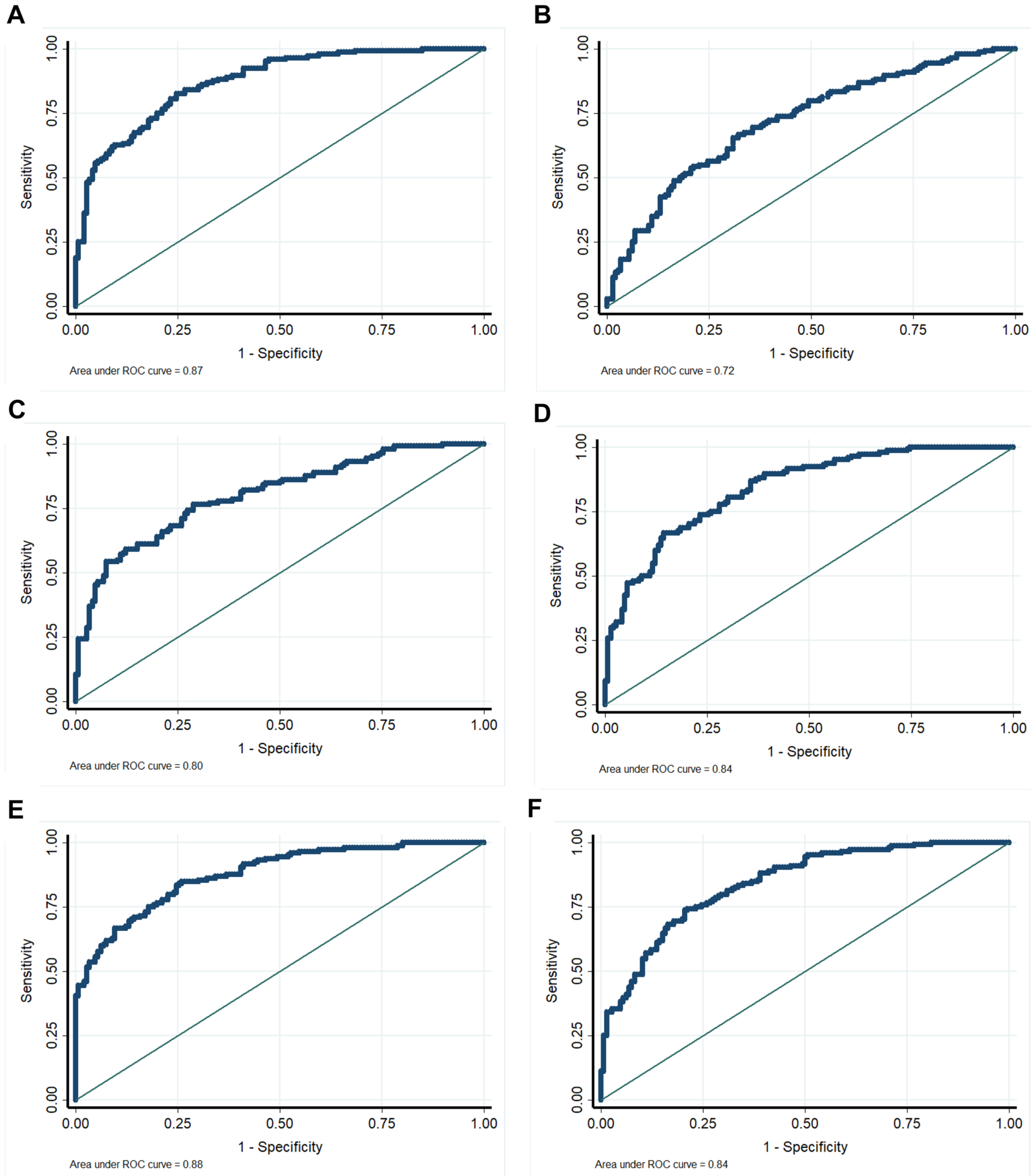

Fig. 4 Receiver operator characteristic curves for **a** concentric isokinetic knee extension strength, **b** concentric isokinetic dorsiflexion strength, **c** isometric leg press strength, **d** isometric bench press strength, **e** isometric mid thigh pull strength, and **f** aerobic capacity for the capsule egress task

that relate to exploration tasks. Including a wide variety of predictor options also allows for situational selection of predictor values. In an ideal situation, the best predictor would be used under all scenarios. However, availability of the necessary equipment to measure a given predictor variable may not be feasible for spaceflight or other remote locations. In this case, measures that give acceptable results, but require less equipment and are more portable become more useful. If one accepts the following ratings for ROC; excellent [1.00–0.91], good [0.90–0.81], fair [0.80–0.71], poor [0.70–0.61], then most of the physiological metrics that we measured would be considered good or better with a few exceptions. For example, isokinetic dorsiflexion (concentric and eccentric) and trunk flexion rate fair to poor for capsule egress and crew rescue. The SDrat and PpNorm values follow suit with the ROC, with SDrat and PpNorm values higher and lower, respectively, for those physiological measures. In general, ROC and PpNorm scores were slightly lower, and SDrat higher, for all measures during the crew rescue drag compared to the other three tasks. The reason for this is not fully understood, but could be an artifact of the weighted suit model. Overall, the added external weight makes it more difficult to complete tasks as the subject is always working against gravity. Conversely, it is possible that the weight of the suit provided some assistance during the forward lean that partially aids in the dragging of the rescue dummy. A partial reduction in the handicap of the suit may have impacted the predictability of the drag task performance. While threshold values did not vary to meaningful degrees across tasks, specificity of the physiological predictors to the tasks can be seen in the ROC, SDrat, and PpNorm values. The ROC for aerobic capacity was 0.88 and 0.92 for the longer duration ambulation and supply transfer and hill climb and descent tasks, Tables 3 and 5 respectively. This is in contrast to 0.84 and 0.75 for the shorter duration capsule egress and rescue drag tasks, Tables 2 and 4. Bench press ROCs were 0.82–0.85 for the capsule egress, and 0.83–0.90 for the rescue drag. These tasks involved pressing motions to lift and push the life raft out the top hatch during the capsule egress task, and rolling the rescue dummy over onto the sled during the rescue drag task. The ambulation and supply transfer task and the hill climb and descent task did not have specific pressing motions involved. The bench press ROC values for these tasks were 0.75–0.76 and 0.74–0.76, respectively. Overall, the results show that many different muscle performance metrics (single and multi-joint) could predict success well. This provides flexibility in test choice where availability of the various test capabilities may differ.

Strength-to-weight thresholds, especially when used for standards or mission readiness recommendations, need to be used within the context of the scenario. For example, if the thresholds are applied to daily living requirements, then the strength based on an individual's true "shirt-sleeve" weight can be used in the calculation of the strength-to-weight ratio by simply dividing the absolute strength of the individual by their body weight. The same can be said of many other occupations whereby there is no meaningful added weight burden beyond their body weight. Examples may include utility crew workers or search crews. However, for occupations that involve a significant weight burden to be borne by the individual, the external weight should be added to the individual body weight in the calculation of required strength. For example, strength thresholds for firefighters should include the weight of the turnout gear in the calculation of their strength-to-weight ratio. Thus, the need for task-specific gear, or weight differences between gear of similar purpose all affect the strength-to-weight ratio. Depending on whether these changes occur in proximity to the threshold breakpoints determines the degree of change in performance (Ryder et al. 2013).

The concept of strength-to-weight is highly relevant for space exploration tasks. In the example of the capsule egress assessment in this study, a 75 kg person would need an absolute leg press of 1,302 N to be above the 17.36 N kg^{-1} leg press threshold (Table 2) in shirt-sleeve, but would need 1,519 N if a 12.5 kg launch and entry suit was worn. These numbers assume a 1-G return to Earth. In relative strength terms the individual would need to have a shirt-sleeve relative strength of 20.25 N kg^{-1} to be above the fully loaded 17.36 N kg^{-1} threshold. For comparison, the mean shirt-sleeve leg press isometric force observed here was 24.12 ± 5.46 N kg^{-1} (Table 1). This is not to assume that individuals below the threshold value are incapable of completing the task, as there were only two observed trials (out of 289) in which subjects were unable to complete the entire capsule egress task. Rather, their strength would be such that the probability of completing the tasks faster than the median time is less than 0.5, which in practical terms separates acceptable from impaired performance.

Boots on the ground EVA is undoubtedly the crowning achievement of any planetary exploration mission. To this end, protecting human physiological capabilities for surface EVA is essential. Multiple factors go into defining strength needs for EVA on different planetary surfaces, even when defining strength normalized to the weight of the individual. One of these is differences in gravitation pull compared to that of the Earth. As discussed above, similar to how the launch and entry suit needs to be taken into account for return to Earth capsule egress, the weight of a planetary exploration suit also needs to be accounted for during surface EVA. While thresholds were quite similar across tasks, we observed the highest leg press threshold to be associated with the hill climb and descent task, a value of 17.65 N kg^{-1} body weight. On Earth, a 75 kg individual's isometric leg press strength would need to be 1,324 N to meet the threshold mark in shirt-sleeve. After adjusting for Mars 3/8-G, the 17.65 N kg^{-1} threshold value translates to 6.62 N Earth-equivalent kg body $weight^{-1}$. The absolute isometric leg press strength needed would then become only 497 N in an (unrealistic) unsuited condition. However, with the inclusion of a surface EVA suit that weighs 65 kg (on Earth), the 75 kg individual would need an absolute leg press strength of 927 N to be at the threshold level. In this case, their suited strength would need to be only 70% of their Earth shirt-sleeve equivalent value to meet the leg press strength threshold. In the case of 1/6 lunar-G, the equivalent absolute strength needed to meet the threshold drops to 411 N even in a 65 kg exploration suit.

While the weighted suit experimental model coupled with weight-normalized strength thresholds allows for simple mathematical accounting for partial-G environments, suit weight, and equipment weight, there are some limitations to this mathematical approach, especially as it applies to pressurized suits. For instance, torque about the joints that is independent of moving the mass of the limbs is not accounted for here and would need to be addressed directly in prototype suits and those selected for missions. Nevertheless, the results of this study highlight that strength demands will be higher for return to Earth unaided egress, than for planetary surface operations. Optimization of suit ergonomics combined with protecting strength for return to Earth contingencies will best enable overall mission success.

Conclusion

We have described a weighted suit model that is useful in generating a wide spectrum of strength-to-weight ratio data for which to benchmark various physiological fitness markers versus specific occupational tasks. Using a probability-based model, rather than a direct regression model to obtain thresholds, makes it more likely that such thresholds would be robust to differences between true operational tasks and the simulated ones used here. In addition, the probability-based model accommodates cases where subjects did not complete the task. Our breakpoint fitness thresholds should provide good starting points for predicting whether operational tasks in future space missions can be expected to be completed in an acceptable time.

Acknowledgements The authors thank Tinh Trinh for engineering support, and Nichole Gadd and Galen Kreutzberg for test operator support. This work was supported by the National Aeronautics and Space Administration Human Research Program NNJ15HK11B.

Compliance with ethical standards

Conflict of interest The authors report that they have no conflicts of interest.

Ethical standards All experiments comply with the current laws of the United States, the country in which the experiments were performed.

References

Alvares JB, Rodrigues R, de Azevedo Franke R, da Silva BG, Pinto RS, Vaz MA, Baroni BM (2015) Inter-machine reliability of the Biodex and Cybex isokinetic dynamometers for knee flexor/extensor isometric, concentric and eccentric tests. Phys Ther Sport 16(1):59–65. https://doi.org/10.1016/j.ptsp.2014.04.004

Barter JT (1957) Estimation of body mass segments. WADC technical report 57-260 (ASTIA 118222) Wright Air Development Center, Wright Patterson Airforce Base, Dayton, Ohio

Day MK, Allen DL, Mohajerani L, Greenisen MC, Roy RR, Edgerton VR (1995) Adaptations of human skeletal muscle fibers to spaceflight. J Gravit Physiol 2(1):P47–P50

De Witt JK, English KL, Crowell JB, Kalogera KL, Guilliams ME, Nieschwitz BE, Hanson AM, Ploutz-Snyder LL (2018) Isometric midthigh pull reliability and relationship to deadlift one repetition maximum. J Strength Cond Res 32(2):528–533. https://doi.org/10.1519/JSC.0000000000001605

English KL, Lee SMC, Loehr JA, Ploutz-Snyder RJ, Ploutz-Snyder LL (2015) Isokinetic strength changes following long-duration spaceflight on the ISS. Aerosp Med Hum Perform 86(12 Suppl):68–77. https://doi.org/10.3357/AMHP.EC09.2015

Feiring DC, Ellenbecker TS, Derscheid GL (1990) Test–retest reliability of the biodex isokinetic dynamometer. J Orthop Sports Phys Ther 11(7):298–300

Gopalakrishnan R, Genc KO, Rice AJ, Lee SM, Evans HJ, Maender CC, Ilaslan H, Cavanagh PR (2010) Muscle volume, strength, endurance, and exercise loads during 6-month missions in space. Aviat Space Environ Med 81(2):91–102

Greene WH (2000) Econometric analysis, 4th edn. Prentice Hall, Upper Saddle River

Hackney KJ, Scott JM, Hanson AM, English KL, Downs ME, Ploutz-Snyder LL (2015) The Astronaut-Athlete: optimizing human performance in space. J Strength Cond Res 29(12):3531–3545. https://doi.org/10.1519/JSC.0000000000001191

Hanley JA, McNeil BJ (1982) The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 143(1):29–36. https://doi.org/10.1148/radiology.143.1.7063747

Laughlin MS, Lee SMC, Loehr JA, Amonette W (2009) Isokinetic strength and endurance tests pre- and post-spaceflight: test–retest reliability, NASA TM-2009-214787. National Aeronautics and Space Administration, Hanover, MD 21076-1320

Levine BD, Lane LD, Watenpaugh DE, Gaffney FA, Buckey JC, Blomqvist CG (1996) Maximal exercise performance after adaptation to microgravity. J Appl Physiol 81(2):686–694. https://doi.org/10.1152/jappl.1996.81.2.686

Moore AD Jr, Downs ME, Lee SM, Feiveson AH, Knudsen P, Ploutz-Snyder L (2014) Peak exercise oxygen uptake during and following long-duration spaceflight. J Appl Physiol 117(3):231–238. https://doi.org/10.1152/japplphysiol.01251.2013

Mulavara AP, Peters BT, Miller CA, Kofman IS, Reschke MF, Taylor LC, Lawrence EL, Wood SJ, Laurie SS, Lee SMC, Buxton RE, May-Phillips TR, Stenger MB, Ploutz-Snyder LL, Ryder JW, Feiveson AH, Bloomberg JJ (2018) Physiological and functional alterations after spaceflight and bed rest. Med Sci Sports Exerc. https://doi.org/10.1249/MSS.0000000000001615

Ozkaya O (2013) Familiarization effects of an elliptical all-out test and the wingate test based on mechanical power indices. J Sports Sci Med 12(3):521–525

Pinheiro JC, Chao EC (2006) Efficient Laplacian and Adaptive Gaussian Quadrature algorithms for multilevel generalized linear mixed models. J Comput Graph Stat 15(1):58–81. https://doi.org/10.1198/106186006X96962

Ploutz-Snyder LL, Downs M, Ryder J, Hackney K, Scott J, Buxton R, Goetchius E, Crowell B (2014) Integrated resistance and aerobic exercise protects fitness during bed rest. Med Sci Sports Exerc 46(2):358–368. https://doi.org/10.1249/MSS.0b013e3182a62f85

Ploutz-Snyder LL, Downs M, Goetchius E, Crowell B, English KL, Ploutz-Snyder R, Ryder J, Dillon EL, Sheffield-Moore M, Scott JM (2018) Exercise training mitigates multi-system deconditioning during bed rest. Med Sci Sports Exerc. https://doi.org/10.1249/MSS.0000000000001618

Rittweger J, Schiessl H, Felsenberg D, Runge M (2004) Reproducibility of the jumping mechanography as a test of mechanical power output in physically competent adult and elderly subjects. J Am Geriatr Soc 52(1):128–131

Rittweger J, Felsenberg D, Maganaris C, Ferretti JL (2007) Vertical jump performance after 90 days bed rest with and without flywheel resistive exercise, including a 180 days follow-up. Eur J Appl Physiol 100(4):427–436. https://doi.org/10.1007/s00421-007-0443-6

Ryder JW, Buxton RE, Goetchius E, Scott-Pandorf M, Hackney KJ, Fiedler J, Ploutz-Snyder RJ, Bloomberg JJ, Ploutz-Snyder LL (2013) Influence of muscle strength to weight ratio on functional task performance. Eur J Appl Physiol 113(4):911–921. https://doi.org/10.1007/s00421-012-2500-z

Spiering BA, Lee SM, Mulavara AP, Bentley JR, Buxton RE, Lawrence EL, Sinka J, Guilliams ME, Ploutz-Snyder LL, Bloomberg JJ (2011) Test battery designed to quickly and safely assess diverse indices of neuromuscular function after unweighting. J Strength Cond Res 25(2):545–555. https://doi.org/10.1519/JSC.0b013e3181f56780

Tesch PA, Berg HE, Bring D, Evans HJ, LeBlanc AD (2005) Effects of 17-day spaceflight on knee extensor muscle function and size. Eur J Appl Physiol 93(4):463–468. https://doi.org/10.1007/s00421-004-1236-9

Trappe T, Trappe S, Lee G, Widrick J, Fitts R, Costill D (2006) Cardiorespiratory responses to physical work during and following 17 days of bed rest and spaceflight. J Appl Physiol 100(3):951–957. https://doi.org/10.1152/japplphysiol.01083.2005

Trappe S, Costill D, Gallagher P, Creer A, Peters JR, Evans H, Riley DA, Fitts RH (2009) Exercise in space: human skeletal muscle after 6 months aboard the International Space Station. J Appl Physiol 106(4):1159–1168. https://doi.org/10.1152/japplphysiol.91578.2008