ERS technical standard on reference values for cardiopulmonary exercise testing: summary report and a call for action

Thomas Radtke[1,2], Luis Chávez[3], Lauren Duggan[3], Sanja Stanojevic[3], Piergiuseppe Agostoni[4,5], Paul Burns[6], Brenda Button[7], Christopher B. Cooper[8], Jana De Brandt[9,10], Luiza Helena Degani-Costa[11,12], Monika Franczuk[13], Aisling McGowan[14,15], Jana Kivastik[16], Pierantonio Laveneziana[17,18], Zoe L. Saynor[19,20], Irene Steenbruggen[21], Helge Hebestreit[22], Karl P. Sylvester[23,24] and the contributing GLI CPET task force members[25]

[1]Department of Epidemiology, Epidemiology, Biostatistics and Prevention Institute, University of Zurich, Zurich, Switzerland. [2]Population Research Center, University of Zurich, Zurich, Switzerland. [3]Department of Community Health and Epidemiology, Dalhousie University, Halifax, NS, Canada. [4]Heart Failure Unit, Centro Cardiologico Monzino IRCCS, Milan, Italy. [5]Department of Clinical Sciences and Community Health, Cardiovascular Section, University of Milan, Milan, Italy. [6]Respiratory and Sleep Physiology Department, Royal Hospital for Children, Glasgow, UK. [7]Departments of Physiotherapy and Respiratory Medicine Alfred Health and Department of Medicine, Nursing and Health Sciences, Monash University, Melbourne, Australia. [8]Departments of Medicine and Physiology, University of California Los Angeles, Los Angeles, CA, USA. [9]Faculty of Medicine, Department of Community Medicine and Rehabilitation, Section of Physiotherapy, Umeå University, Umeå, Sweden. [10]Faculty of Rehabilitation Sciences, REVAL – Rehabilitation Research Center, BIOMED – Biomedical Research Institute, Hasselt University, Diepenbeek, Belgium. [11]Faculdade Israelita de Ciências da Saúde Albert Einstein – Hospital Israelita Albert Einstein, São Paulo, Brazil. [12]Laboratory of Exercise and Quality of Life, Centro Universitário São Camilo, São Paulo, Brazil. [13]Lung Pathophysiology Department, National Tuberculosis and Lung Diseases Research Institute, Warsaw, Poland. [14]Department of Respiratory and Sleep Diagnostics, Connolly Hospital, Dublin, Ireland. [15]School of Physics, Clinical and Optometric Sciences, Technological University Dublin, Dublin, Ireland. [16]Department of Physiology, University of Tartu, Tartu, Estonia. [17]Sorbonne Université, INSERM, UMRS1158 Neurophysiologie Respiratoire Expérimentale et Clinique, Paris, France. [18]AP-HP, Groupe Hospitalier Universitaire APHP–Sorbonne Université, Hôpitaux Pitié-Salpêtrière et Tenon, Service des Explorations Fonctionnelles de la Respiration, de l'Exercice et de la Dyspnée (Département R3S), Paris, France. [19]School of Health Sciences, Faculty of Environmental and Life Sciences, University of Southampton, Southampton, UK. [20]NIHR Southampton Biomedical Research Centre, University Hospital Southampton NHS Foundation Trust, Southampton, UK. [21]Lung Function Department, Isala Hospital, Zwolle, The Netherlands. [22]Paediatric Department, University Hospitals Würzburg, Würzburg, Germany. [23]Respiratory Physiology, Papworth Hospital NHS Foundation Trust, Cambridge, UK. [24]Respiratory Physiology, Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK. [25]Contributors to the working group are listed at the end of this document.

Corresponding authors: Karl P. Sylvester (karl.sylvester@nhs.net) and Thomas Radtke (thomas.radtke@uzh.ch)

Shareable abstract (@ERSpublications)
Cardiopulmonary exercise testing data were too heterogeneous to generate Global Lung Function Initiative reference equations. Standardised protocols, quality control, and centralised analysis may improve reliability to support its clinical and research use. https://bit.ly/46gBzGH

Cite this article as: Radtke T, Chávez L, Duggan L, *et al*. ERS technical standard on reference values for cardiopulmonary exercise testing: summary report and a call for action. *Eur Respir J* 2026; 67: 2502463 [DOI: 10.1183/13993003.02463-2025].

Received: 04 April 2025
Accepted: 23 Dec 2025

Abstract

Background Cardiopulmonary exercise testing (CPET) assesses physiological responses to incremental exercise and identifies potential causes of exercise limitation. There have been several population-specific reference equations published, but none that span the human age range. It would be advantageous to have all-age global reference ranges. This task force aimed to derive Global Lung Function Initiative reference equations for peak oxygen uptake ($\dot{V}_{O_2peak}$) and peak work rate (W_{peak}) in healthy individuals.

Methods CPET data were collected retrospectively. Generalised additive models of location, shape and scale were used to develop reference ranges, including age, sex, height and weight as explanatory variables. The influence of geographic region, equipment, testing protocols and averaging methods for peak exercise data were also examined.

Results Data from 5956 healthy individuals aged between 6 and 83 years across 17 sites in Europe, North and South America and Asia were analysed. There was substantial between-subject variability in both $\dot{V}_{O_2peak}$ and W_{peak}, with wide confidence intervals across age groups. Heterogeneity in $\dot{V}_{O_2peak}$ was related

to geographic region, metabolic cart type, and averaging methods for peak exercise values. Controlling for these variables improved model fit, but not sufficiently to be reliable predictors for reference ranges.
Conclusion Significant heterogeneity in CPET testing methodology and outcomes between sites precluded the development of reference ranges for $\dot{V}_{O_2peak}$ and W_{peak}. This task force has developed a framework for prospective data collection with strictly standardised protocols and centralised data analysis to reduce variability and establish robust, clinically meaningful reference ranges for CPET outcomes.

Background

Cardiopulmonary exercise testing (CPET) is a commonly used method to characterise individual physiological responses to exercise and to assess cardiorespiratory fitness [1–3]. It is used to determine the potential causes of exercise limitation such as shortness of breath on exertion. CPET is also utilised to risk stratify individuals who need surgical interventions and to monitor the response to pharmacological, nonpharmacological and surgical therapies [1, 4–7]. Cardiorespiratory fitness is considered a strong prognostic indicator of respiratory and all-cause morbidity and mortality, comparable in magnitude to traditional risk factors such as smoking, hypertension and type 2 diabetes [8].

One aspect of CPET interpretation includes the comparison of measured values to what is expected in an otherwise healthy population. Numerous different reference equations for both adult and paediatric populations are available, but none that spans the entire age range [9–21]. To date, the largest data source providing reference values for submaximal and maximal CPET outcomes for both treadmill and cycle ergometry is the FRIEND (Fitness Registry and the Importance of Exercise National Database) registry [9]. This United States-based registry contains tests from >20 000 adults with measurements collected between 1960 and 2021. The FRIEND registry is a valuable resource for the interpretation of cardiorespiratory fitness for the USA population; however, it is probably not applicable to other countries due to differences in population characteristics, such as the high prevalence of obesity and those with cardiometabolic disorders in the USA. There are numerous published reference equations for CPET, again with notable differences between them [12, 16, 18, 20, 22]. Furthermore, there is evidence that clinical interpretation of measured values differs based on which reference equation is applied [23].

A recent European Respiratory Society (ERS) technical standards document summarised procedures for CPET as part of clinical and/or research-related investigations as a prerequisite for between-centre comparisons [1]. The task force report [1] highlighted the need for the collection of globally applicable reference data for CPET outcomes. The aim of this study was to collate CPET data from healthy individuals to derive Global Lung Function Initiative (GLI) reference equations for CPET outcomes.

Methods

The ERS approved the application to develop GLI all-age reference values for CPET outcomes and initiated the task force (ERS TF-2021-09). The task force comprised researchers and healthcare professionals with expertise in CPET, the development of international standards, lung physiology, epidemiology and biostatistics, including three early-career members (L. Duggan, L. Chávez, J. De Brandt).

Data sources

We conducted a systematic literature review, supported by an information specialist, to identify published studies that included CPET in healthy individuals spanning from the year 2000 and beyond. Details of the search strategy are provided in the supplementary material (supplementary tables S1–S5). We contacted the authors of studies with ⩾50 healthy participants who had undergone a CPET on either a cycle ergometer or treadmill, employing either a ramp or minute-by-minute incremental exercise protocol and invited them to share their data with the task force. This restriction was applied to ensure that data were collected only from centres with significant experience conducting such assessments. Invitations were also circulated through international and local respiratory societies including the American Thoracic Society (ATS), Asia Pacific Society of Respirology, ERS, Latin American Chest Association, Pan African Thoracic Society, Thoracic Society of Australia and New Zealand to solicit unpublished data.

Data collection, management and processing

Contributors were required to obtain approval from the local ethics committee to contribute data to the GLI Network. All data were pseudo-anonymised before submission and entered into a standard data template. Study data were collected and managed using Research Electronic Data Capture (REDCap) hosted at Dalhousie University, Canada [24]. Prior to submission of individual CPET data, contributors signed a memorandum of understanding and completed a survey regarding the meta-data including details of the study population, CPET equipment type and manufacturer (and any modification of commercial equipment), as well as testing protocols that had been applied (supplementary table S6). Submitted data

underwent checks for consistency and quality including review of unplausible data (K.P. Sylvester, L. Chávez, T. Radtke).

All submitted data were reviewed to identify missing data and implausible outliers; contributors were contacted directly to clarify discrepancies. Suspected outliers were confirmed with the study sites and corrected as appropriate, if needed. Weight and height values with z-scores <−5 or >5 were classified as outliers and subsequently excluded from the analysis. If multiple assessments were available for the same individual (*i.e.* serial assessments), one CPET assessment was randomly selected.

Minimum data requirements for inclusion in the final analyses included the origin of the data (continent), CPET testing protocol, participant age, biological sex, year of birth, height, weight, measured peak oxygen uptake ($\dot{V}_{O_2peak}$), peak work rate (W_{peak}), respiratory exchange ratio (RER_{peak}), and heart rate at peak exercise (HR_{peak}).

Maximal effort during CPET was defined using RER_{peak} and HR_{peak}. A test was classified as maximal if RER_{peak} was ⩾1.1 or if HR_{peak} exceeded the lower limit of normal for their age predicted maximal heart rate (LLN HR_{max}) [3, 25, 26]. Thus, subjects neither reaching a predicted HR_{peak} greater than the LLN (HR_{peak} >LLN HR_{max}) or a RER_{peak} ⩾1.1 at peak exercise, were excluded.

Healthy individuals without known cardiopulmonary or cardiometabolic disease were defined as those who were not current smokers and had smoked ⩽100 cigarettes in their lifetime. This definition includes individuals with missing information on their smoking or ever-smoking status (n=1729 subjects). Obese and underweight individuals were excluded from the analysis due to the potential impact on cardiorespiratory fitness. The Centers for Disease Control and Prevention growth charts were used for children aged 2–18 years with obesity defined as body mass index (BMI) >95th percentile and underweight defined as <5th percentile [27]. For adults, obesity was defined as BMI >30 $kg \cdot m^{-2}$; underweight was defined as BMI <18.5 $kg \cdot m^{-2}$. Furthermore, contributors were instructed not to share data from elite athletes to maintain the study's focus on the general healthy population. We did not accept data collected from one site, a sports science institute, due to suspected inclusion of elite athletes. We cannot be certain that the dataset does not include recreational and or competitive athletes that met all other conditions of health but were not indicated as being elite athletes.

Statistical analysis

The generalised additive models of location shape and scale (GAMLSS) technique, previously used for other GLI projects, was used to develop the reference equations for $\dot{V}_{O_2peak}$ and W_{peak} on cycle ergometer. Briefly, the GAMLSS technique allows the median value to be summarised as a function of multiple explanatory variables (*e.g.* height, age), the spread of values around the median value to be constant or vary by a function of explanatory variables, and any departure from a normal distribution (skewness, kurtosis) to be transformed to normal using a Box–Cox transformation. Thus, the resulting model residuals are normally distributed. The Box–Cox Cole and Green family distribution and a log transformation of the response variable was applied. All analyses were performed using the GAMLSS package in the statistical programme R (The R Project for Statistical Computing, www.r-project.org; version 4.4.2).

Two primary CPET outcomes were collected, $\dot{V}_{O_2peak}$ and W_{peak}. Age, sex, height and weight were treated as explanatory variables and modelled in a multivariable framework (herein referred to as operational model). To better understand the contribution of each predictor and to rank factors influencing variability, we first performed univariable analysis and then applied a stepwise approach to the multivariable regression models. We did not investigate race and ethnicity as a predictor of either $\dot{V}_{O_2peak}$ or W_{peak} because race and ethnicity can be seen as social constructs, without a consistent definition globally, and recent statements endorsed by both the ATS and the ERS have recommended against its continued use in reference equations [28–31]. The goodness of fit of each model was assessed using Akaike Information Criteria (AIC), the coefficient of determination (R^2), residual Q–Q plots, and residual plots. Variables that improved overall model fit were retained in the final model. Differences in $\dot{V}_{O_2peak}$ and W_{peak} by site, continent, exercise testing protocol and averaging method of peak exercise data were investigated after the operational models were derived. To assess the magnitude of between-subject variability in $\dot{V}_{O_2peak}$, we calculated the coefficient of variation and compared it with existing GLI reference equations for forced vital capacity [32]. The coefficient of variation is a standardised measure of dispersion that expresses the standard deviation of a measurement as a percentage of the mean. A lower coefficient of variation indicates less relative variability, whereas a higher coefficient of variation reflects greater spread in the data.

Results

We contacted 320 study authors, of whom 23 accepted our invitation to contribute, resulting in the submission of 11 966 individual participants' data. Of these, only three sites submitted CPET data collected by treadmill assessments (n=1101). We restricted the analysis to cycle ergometry data due to the relatively low sample size of treadmill data. After exclusions (figure 1), data from 5956 healthy participants (50.5% female) across 17 sites from Europe (n=8), North and South America (n=7), and Asia (n=2) were included.

Characteristics of the study population by site are summarised in table 1. Nine sites contributed data from adults; four sites contributed data from children, adolescents and adults; and four sites contributed data from children and adolescents. Overall, data were collated from healthy individuals between 6 and 83 years (supplementary figure S1), with relatively fewer observations in the 20–40-years age range as well as among individuals aged ⩾70 years. Details regarding metabolic carts, software, cycle ergometers, and testing protocols are summarised in supplementary table S6. 14 sites used a ramp incremental protocol, while three sites used a minute-by-minute incremental protocol (supplementary table S6). The number of participants meeting maximal effort criteria is provided in supplementary table S7. Raw data showing subjects with maximal *versus* submaximal effort and the heterogeneity in achieved $\dot{V}_{O_2peak}$ and W_{peak} are visualised in supplementary figure S2.

Both $\dot{V}_{O_2peak}$ and W_{peak} were higher in males compared to females and showed a wide range of values across age groups (figure 2, supplementary figure S3). Univariable and multivariable analyses of predictors

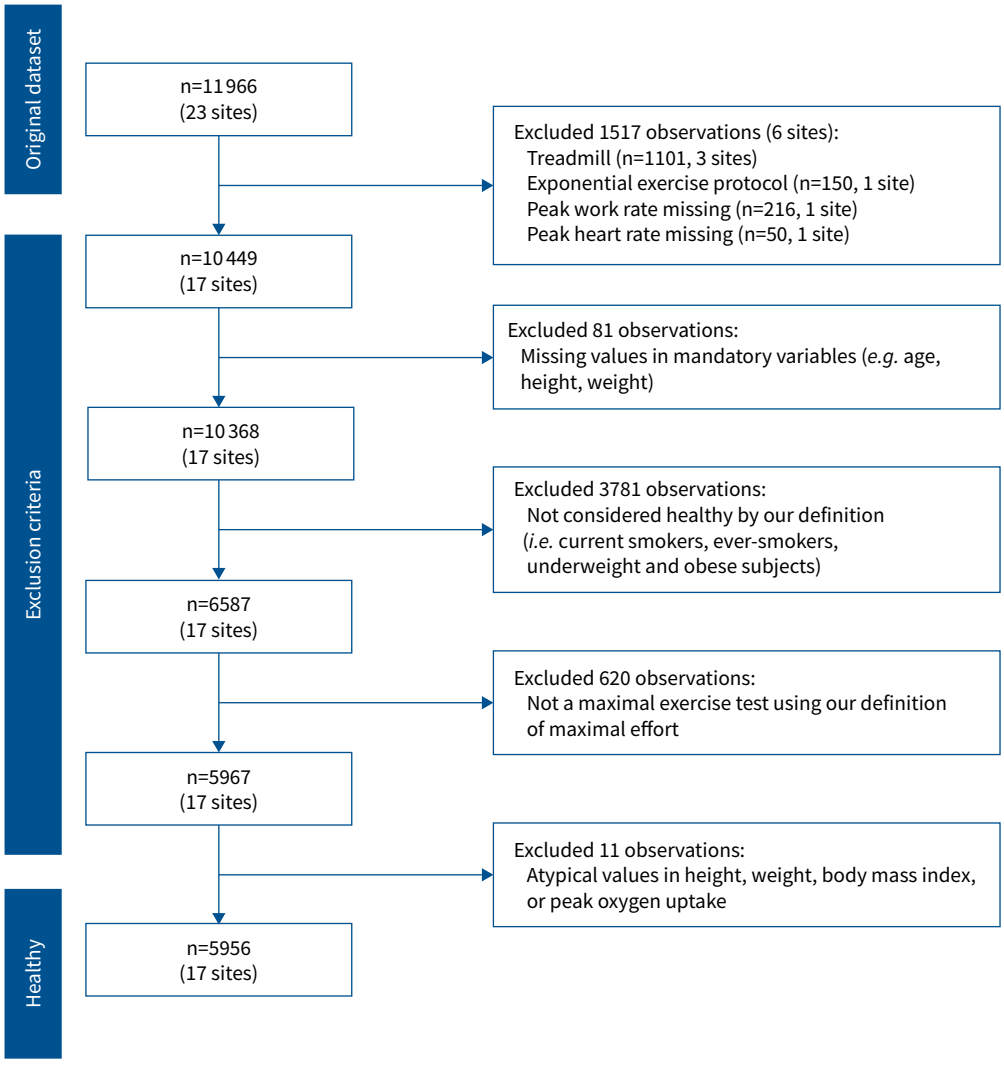

FIGURE 1 Exclusions flow chart.

TABLE 1 Summary of data included in final models by study site

ID	Country[#]	Subjects[¶]	Age years	Female %	Overweight[+] %	W_{peak} $W \cdot kg^{-1}$	$\dot{V}_{O_2peak}$ $mL \cdot kg^{-1} \cdot min^{-1}$
#4	Brazil	76	19–69	53.9	28.9	2.4 (2.0–2.7)	28.4 (24.9–32.0)
#6	Brazil	28	19–72	67.9	39.3	2.2 (1.7–2.5)	25.4 (23.6–32.6)
#7	Brazil	7	25–67	71.4	71.4	1.6 (1.4–2.0)	21.7 (20.1–24.1)
#8	Canada	221	12–17	52.5	10.9	3.5 (3.0–4.0)	42.2 (36.5–50.2)
#1	China	831	18–79	40.6	34.3	1.7 (1.5–1.9)	22.3 (20.1–25.4)
#16	Finland	513	58–78	54.0	59.5	2.2 (1.9–2.6)	25.9 (22.1–30.9)
#17	Germany	866	21–83	61.2	58.0	2.1 (1.7–2.5)	25.4 (21.0–31.0)
#14	Greece	112	14–74	45.5	56.2	2.0 (1.6–2.5)	25.9 (21.7–32.4)
#15	Greece	27	16–62	37.0	44.4	2.4 (2.0–3.0)	25.5 (21.1–32.6)
#13	Italy	314	21–77	41.1	29.6	2.3 (1.8–2.8)	29.2 (23.5–35.0)
#3	Mexico	236	9–81	51.7	39.0	2.2 (1.7–2.5)	29.2 (22.8–34.8)
#2	Spain	213	18–83	50.2	36.6	2.5 (2.0–3.1)	33.0 (27.8–39.3)
#5	Switzerland	70	13–16	55.7	5.7	3.6 (3.2–4.2)	46.0 (39.0–53.7)
#10	Thailand	388	21–78	62.4	26.0	1.7 (1.4–2.2)	24.0 (20.2–28.8)
#9	UK	70	7–17	51.4	14.3	2.7 (2.5–3.6)	40.1 (34.1–45.1)
#11	USA	1809	6–18	47.8	18.4	3.1 (2.7–3.5)	39.6 (34.0–45.2)
#12	USA	175	14–65	45.1	57.7	2.2 (1.8–2.6)	25.3 (21.3–30.4)

Data are presented as n, range or median (interquartile range), unless otherwise stated. W_{peak}: peak work rate; $\dot{V}_{O_2peak}$: peak oxygen uptake. [#]: countries are displayed in alphabetical order; [¶]: we requested datasets from sites with a minimum of 50 healthy subjects. Due to exclusions (figure 1), <50 subjects from three sites could be included in the final analysis; [+]: overweight was defined as a body mass index >25 $kg \cdot m^{-2}$ for adults and >85th percentile for children aged 2–18 years using the Centers for Disease Control and Prevention growth charts [27].

of $\dot{V}_{O_2peak}$ ($mL \cdot min^{-1}$) and W_{peak} (Watts) and their corresponding model goodness of fit are provided in tables 2 and 3, respectively. The operational models, incorporating age, sex, height and weight as predictors, explained 68.8% of the variability of $\dot{V}_{O_2peak}$ and 67.5% of the variability of W_{peak} (tables 2 and 3, supplementary figure S4). In the operational models, the 80% predicted value for $\dot{V}_{O_2peak}$, often used as a proxy of the LLN, was consistently higher than the LLN across the age range of 20–80 years for both males and females (figure 3, supplementary table S8). For example, in a 30-year-old male with an average height (176 cm) and weight (75 kg), the limits of normal for $\dot{V}_{O_2peak}$ ranged from 1843 to 3463 $mL \cdot min^{-1}$ or 24.6 to 46.2 $mL \cdot kg^{-1} \cdot min^{-1}$, respectively (supplementary table S8). Furthermore, a substantial proportion of participants, particularly females aged ⩾50 years, had $\dot{V}_{O_2peak}$ values <17.5 $mL \cdot kg^{-1} \cdot min^{-1}$, corresponding to less than five metabolic equivalents of task (METs) (figure 4). This cut-off crossed the LLN at approximately age 40 years in females and 60 years in males, considering average height. Moreover, the coefficient of variation, a measure of relative variability between subjects, $\dot{V}_{O_2peak}$ was high and ranged between 16.5% and 26.3%, indicating substantial variability around predicted values (figure 5).

Differences in $\dot{V}_{O_2peak}$ and W_{peak} values by site persisted after accounting for age, sex, height and weight (figure 6). We further investigated if geographical region (*i.e.* continent), metabolic cart type, exercise protocol and averaging method for determining $\dot{V}_{O_2peak}$ contributed to the between-subject variability in $\dot{V}_{O_2peak}$ (table 2), and whether continent, cycle ergometer type and exercise protocol contributed to the variability in W_{peak} (table 3). Each of the investigated covariates (which were added to the operational model including age, sex, height and weight) improved model fit (*i.e.* lower AIC) and increased the explained between-subject variability (*i.e.* higher R^2 value) in $\dot{V}_{O_2peak}$ and W_{peak}. The final models including all covariates explained 75.3% and 77.8% of the variability in $\dot{V}_{O_2peak}$ and W_{peak}, respectively. Compared with the operational models, the inclusion of all covariates increased the R^2 value by 6.5 percentage points for $\dot{V}_{O_2peak}$ and 10 percentage points for W_{peak}, with continent and metabolic cart type contributing most to the additional explained variability (tables 2 and 3). In the final models, residual differences between geographical region, metabolic cart type, and averaging methods for determining $\dot{V}_{O_2peak}$ remained (figure 7).

Notably, the shape of the curves for both $\dot{V}_{O_2peak}$ and W_{peak} as a function of age (supplementary figure S4) and the between-subject variability were probably driven by site-specific effects. These differences in $\dot{V}_{O_2peak}$ values between sites persisted and could not be explained after adjusting for basic demographic

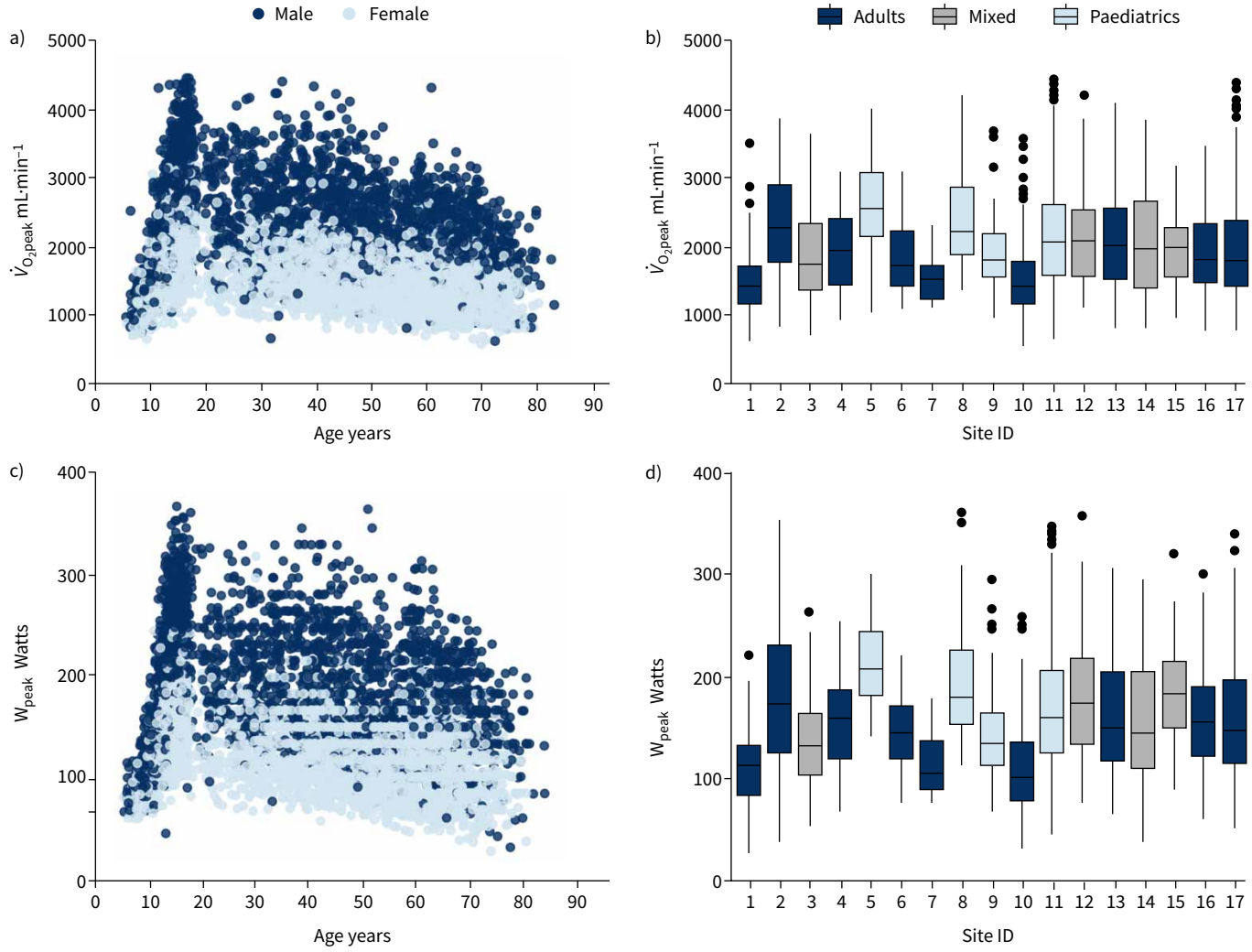

FIGURE 2 a) Peak oxygen uptake ($\dot{V}_{O_2peak}$) *versus* age, stratified by sex; b) $\dot{V}_{O_2peak}$ *versus* site, stratified by age range of site; c) peak work rate (W_{peak}) *versus* age, stratified by sex; and d) W_{peak} *versus* site, stratified by age range of site.

variables (age, sex, height, weight), as shown in figure 6. These unexplained site differences create an artificial spread in $\dot{V}_{O_2peak}$ values, which limits the validity of a single generalised equation. Consequently, it was not possible to define robust reference ranges and to establish a single reference equation that can be generalised beyond the sites included in this study.

Finally, we compared predicted values for $\dot{V}_{O_2peak}$ with frequently used reference equations (supplementary figures S5–S9) [9, 16, 20, 22, 47]. Overall, there was considerable heterogeneity in predicted values and percentiles between the different equations, but the wide range of $\dot{V}_{O_2peak}$ values in both males and females and across age groups was consistent across all equations.

Here, we propose a framework for a prospective collection of CPET outcomes to establish age-specific reference ranges in healthy individuals, using population-based sampling, strictly standardised testing protocols, quality control and centralised data analysis (table 4).

Discussion

This ERS task force aimed to establish all age reference equations for multiple CPET parameters utilised in the assessment of individual's exercise capacity. The large variability identified in $\dot{V}_{O_2peak}$ and W_{peak} from 5956 healthy individuals across 17 sites in Asia, Europe, and North and South America, with the underrepresentation of certain age groups, precluded generation of generalisable reference ranges. Further

TABLE 2 Univariable and multivariable analyses of predictors of peak oxygen uptake ($mL \cdot min^{-1}$) and corresponding model goodness of fit

	Akaike information criterion	R^2
Univariable models		
Height	90342	0.456
Sex	91990	0.281
Age	92360	0.236
Weight	92468	0.222
Continent[#]	92905	0.163
Metabolic cart type[¶]	93315	0.105
Exercise protocol[+]	93887	0.012
Averaging method[§]	93922	0.009
Multivariable models		
Basic model (age ~ sex ~ height)	87422	0.666
Operational model (age ~ sex ~ height ~ weight)	87036	0.688
Operational model *plus* exercise protocol	86948	0.693
Operational model *plus* averaging method	86752	0.704
Operational model *plus* metabolic cart type	86460	0.718
Operational model *plus* continent	86122	0.733
Final model including all covariates	85720	0.753

[#]: Asia, Europe, North America, South America. [¶]: Cardiovit, Oxycon Pro, Ultima CPX, Vmax Encore, Vyntus CPX, Quark CPET, miscellaneous; the miscellaneous category contains a mix of different metabolic carts used by study sites where it was not clear which cart was used for which study participant. [+]: ramp incremental protocol, minute-by-minute incremental protocol. [§]: time averaged 10 s, 20 s or 30 s; breath averaged (mid 5–7 breaths); breath averaged (five rolling averages); three highest peak values averaged (30 s).

prospective studies with strict standardisation of CPET protocols are needed to reduce variability and to establish robust and clinically meaningful interpretation strategies for CPET outcomes [48].

Early in the process, it became evident that significant variability existed in how CPET was conducted between sites, including differences in ergometers, testing protocols, and data averaging methods, which might impact the accuracy of submaximal parameters such as anaerobic threshold or important slopes. We therefore focussed on maximal parameters ($\dot{V}_{O_2peak}$ and W_{peak}), which are key performance and health indicators, and would be less likely to be impacted by the variations in how the test was conducted.

TABLE 3 Univariable and multivariable analyses of predictors of peak work rate (Watts) and corresponding model goodness of fit

	Akaike information criterion	R^2
Univariable models		
Height	60595	0.474
Sex	62695	0.251
Age	62715	0.249
Weight	62886	0.227
Continent[#]	63095	0.200
Cycle ergometer type[¶]	63717	0.113
Exercise protocol[+]	64394	0.003
Multivariable models		
Basic model (age ~ sex ~ height)	58023	0.659
Operational model (age ~ sex ~ height ~ weight)	57735	0.675
Operational model *plus* exercise protocol	57545	0.686
Operational model *plus* cycle ergometer type	56442	0.740
Operational model *plus* continent	56151	0.752
Final model including all covariates	55520	0.778

[#]: Asia, Europe, North America, South America. [¶]: cycle ergometer type: Ergometry ERG 911 S Plus, Ergoline 900, Ergoselect 200, Lode Corival, VIAsprint 150P, miscellaneous; the miscellaneous category contains a mix of different ergometers used by study sites where it was not clear which ergometer was used for which study participant. [+]: ramp incremental protocol, minute-by-minute incremental protocol.

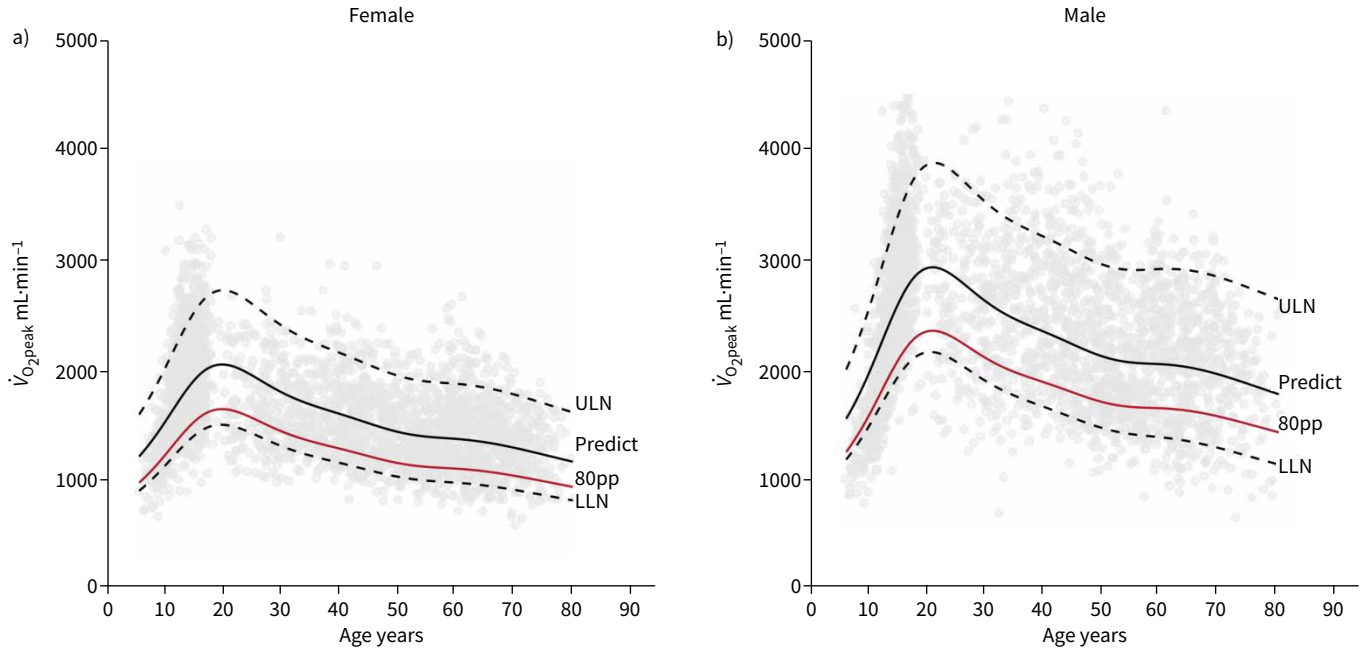

FIGURE 3 Model predictions of peak oxygen uptake ($\dot{V}_{O_2peak}$) with 90% prediction intervals for a) females and b) males considering average height and weight. LLN: lower limit of normal (5th percentile); ULN: upper limit of normal (95% percentile); 80pp: 80% of the predicted value of $\dot{V}_{O_2peak}$.

Cardiorespiratory fitness exhibits substantial variability within a given age and sex group [9, 16, 20, 22] and this is supported by our data. The range of values observed in otherwise healthy individuals is very wide and extends beyond traditionally used "fixed cut-offs" (*e.g.* 80% predicted or 17.5 mL·kg^{-1}·min^{-1}. Importantly, this wide range is not unique to our dataset; it is also evident in many other commonly used reference equations (supplementary figures S5–S9), yet rarely acknowledged as a limitation. Such

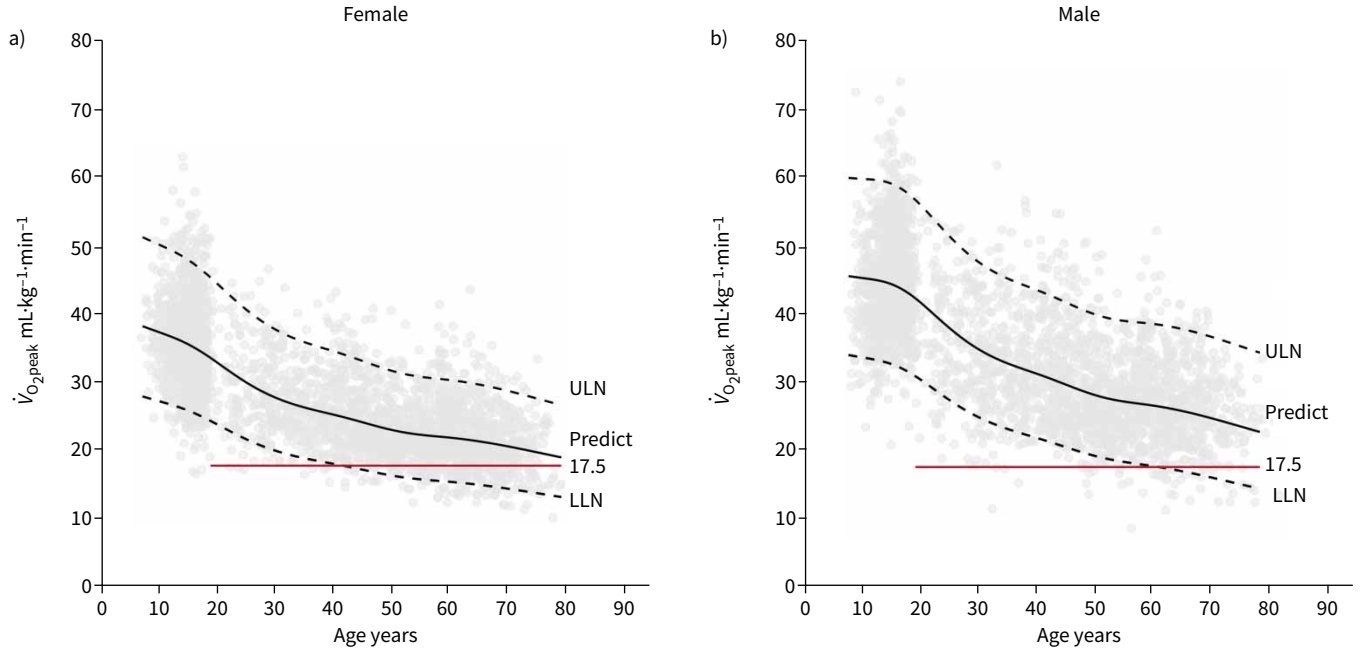

FIGURE 4 Model predictions of peak oxygen uptake ($\dot{V}_{O_2peak}$) with 90% prediction intervals for a) females and b) males considering average height. The red line displays a $\dot{V}_{O_2peak}$ value of 17.5 mL·kg^{-1}·min^{-1}, corresponding to 5 metabolic equivalents of task. LLN: lower limit of normal (5th percentile); ULN: upper limit of normal (95th percentile).

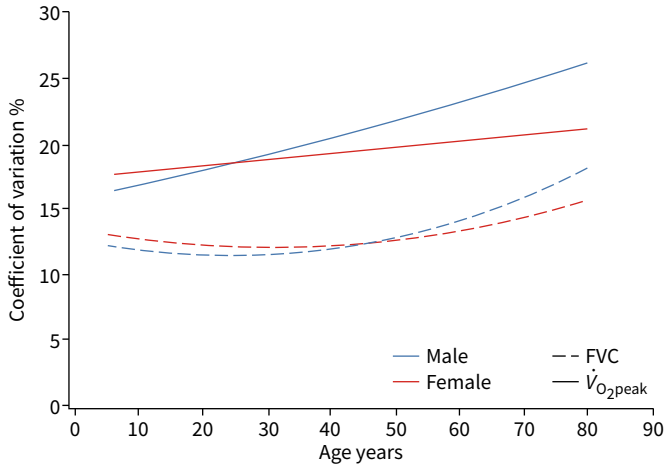

FIGURE 5 Relationship between age and the coefficient of variation for Global Lung Function Initiative reference equations for peak oxygen uptake ($\dot{V}_{O_2peak}$) and forced vital capacity (FVC) among males and females.

variability may lead to misinterpretation at the individual level, particularly in defining low cardiorespiratory fitness.

Between-subject variability can be expressed using the coefficient of variation, a relative measure of dispersion, and ranged between 16.5% and 26.3% for $\dot{V}_{O_2peak}$. This exceeds the variability typically observed in resting pulmonary function tests, such as spirometry or pulmonary diffusing capacity for carbon monoxide [49]. While forced expiratory manoeuvres primarily depend on an individual's cooperation and reflect pulmonary limitations, CPET performance depends on maximal effort (or symptom limitation in clinical populations), and reflects the integrated function of the lungs, heart and skeletal muscles. Therefore, greater variability is to be expected but it poses challenges for interpretation at the individual level.

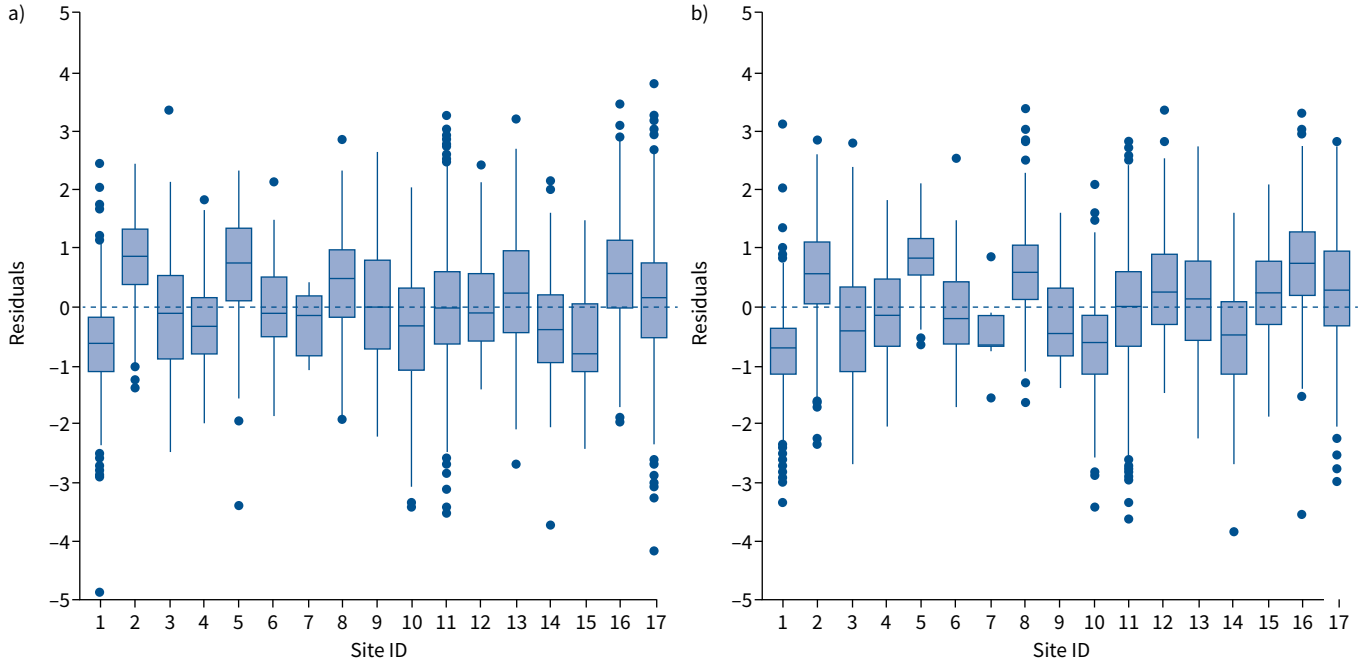

FIGURE 6 Standardised residuals of a) peak oxygen uptake *versus* site, and b) peak work rate *versus* site.

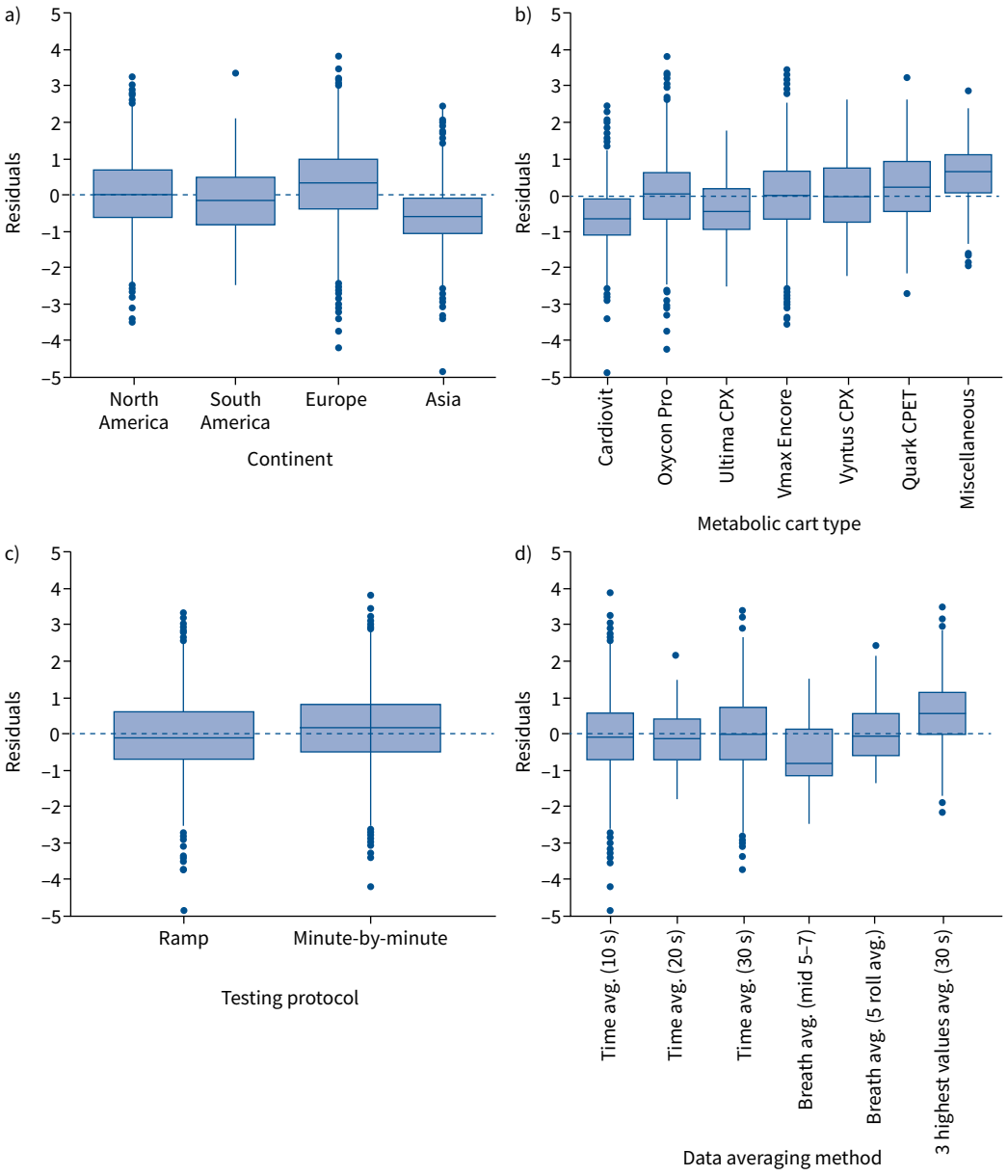

FIGURE 7 Standardised residuals of peak oxygen uptake *versus* a) continent, b) metabolic cart type, c) exercise testing protocol, and d) averaging method of peak exercise values. The "miscellaneous" category contains a mix of different metabolic carts used by study sites where it was not clear which cart was used for which study participant. Avg: average.

Not surprisingly, there was substantial disagreement between the $\dot{V}_{O_2peak}$ values predicted from our dataset and those derived from other reference equations [9, 16, 20, 22, 47]. Global variation in data sources, sampling strategies, the inclusion of subjects with and without cardiovascular and/or cardiometabolic risk factors (*e.g.* obesity), the inclusion or exclusion of current smokers, and the use of different modelling techniques including predictor variables probably contribute to the observed discrepancies. A major challenge common to all reference value datasets is the definition of "healthy" and the population to which the equations are intended to apply. Reference values aim to identify individuals at risk, for example with low cardiorespiratory fitness. Including people with cardiovascular or cardiometabolic conditions may bias these ranges and reduce sensitivity for detecting those truly at risk.

Data from the FRIEND registry [9], which includes data from 34 laboratories across the United States of America, support this observation, as their apparently healthy adults show relatively low $\dot{V}_{O_2peak}$ values.

TABLE 4 Proposed framework for the prospective collection of cardiopulmonary exercise testing (CPET) outcomes to establish all age reference ranges in healthy individuals

	Explanation	Justification	Task force suggestions
CPET-specific considerations	Competent practitioners[#]	Data contributors to a prospective study should have demonstrated experience in conducting CPETs to ensure high-quality data.	Practitioners must complete proficiency testing and certification before participation.
	Metabolic cart	The measurement accuracy between commercially available devices varies substantially [33]. The use of devices from the same manufacturer will not be feasible in a large prospective study and limits the generalisability of normal references ranges. Measurement error can be reduced by strict calibration and verification procedures. Our analysis showed that the type of metabolic cart improved the explained between-subject variability in $\dot{V}_{O_2peak}$ (table 2).	Equipment should be rigorously calibrated before every test. This task force suggests the use of devices with a measurement error <5% for key gas exchange variables, as suggested previously [33].
	Metabolic simulator/biological control[#]	Regular quality control using metabolic simulators or healthy subjects experienced with CPET is warranted to ensure equipment stability and detection of system malfunction to generate valid data [2, 33, 34].	This task force suggests monthly quality control to ensure collection of valid data. Metabolic simulators are preferable.
	Testing protocols[#]	CPET protocols should include the following phases: minimum 3 min of rest, strictly 3 min of unloaded pedalling[¶], 8–12 min of exercise, and ⩾2 min of recovery [1–3, 35]. The use of different exercise protocols may influence submaximal and maximal CPET outcomes [36]. Minute-by-minute incremental exercise protocols or ramp incremental protocols yield comparable results for submaximal and maximal CPET outcomes, provided the exercise phase duration is similar [37, 38]. This is supported by our analysis, which showed that the explained between-subject variability of $\dot{V}_{O_2peak}$ was unaffected when the type of testing protocol (*i.e.* minute-by-minute incremental protocol *versus* ramp incremental protocol) was added as a covariate to the operational model (R^2=0.688 *versus* 0.693; table 2).	To support stricter standardisation of CPET and harmonise data generation, this task force suggests the use of continuous ramp incremental protocols, as they facilitate clearer interpretation and analysis of physiological responses during incremental exercise by minimising the impact of work rate transitions.
	Criteria for defining a maximal test[#]	Numerous criteria are used to define maximal effort including a plateau in $\dot{V}_{O_2peak}$, peak heart rate relative to age-predicted maximum values, various thresholds for the peak respiratory exchange ratio, blood lactate concentration at peak exercise, and the subject's perceived exertion. These criteria were applied inconsistently across published reference value studies, which may influence peak exercise outcomes and limit comparability between reference value datasets. This is due to the considerable heterogeneity across standardisation documents, which must be resolved. This task force found that neither peak heart rate nor respiratory exchange ratio discriminate well between people with higher *versus* lower cardiorespiratory fitness. This is probably a consequence of sites using different criteria to judge maximal effort. This is supported by data from the FRIEND registry showing negligible differences in $\dot{V}_{O_2peak}$ between exercise tests when maximal effort was defined as a respiratory exchange ratio ⩾1.0 or ⩾1.1 [9].	Tests should be terminated based on individual symptom limitation or if there are safety concerns. They must not be stopped based on a perceived maximal effort. Criteria to define maximal effort need to be pre-defined and reported to facilitate centralised data analysis. This task force suggests stricter standardisation in the definition of maximal exercise tests with the support of future research. This includes the development of a standardised encouragement tool [39].

Continued

TABLE 4 Continued

	Explanation	Justification	Task force suggestions
	Sampling strategies for CPET[#]	Previous statements on cardiopulmonary exercise testing suggested interval averaging of 30–60 s [2, 3], or 10–15 s for tabular data, and 20–60 s for the final CPET report [1]. Different post-test analysis strategies result in differences in peak exercise values [40–42] and potentially increase between-subject variability. This task force found large between-site heterogeneity in analytical approaches to determine peak exercise values (supplementary table S6, figure 7). To improve comparability of CPET outcomes across sites, sampling strategies to determine peak exercise values must be standardised.	Although previous recommendations have been to use a 30-s time-based average to provide peak exercise values, this probably underestimates true $\dot{V}_{O_2peak}$. This task force suggests a five-breath rolling average, as this might represent 5–10 s towards maximal exercise. Future research is warranted to determine the impact of data averaging on an individual's maximum achieved exercise capacity. This task force suggests determining peak exercise values based on a rolling average of five breaths in a prospective data collection.
	Interpretation of CPET data[#]	Centralised interpretation of submaximal and maximal CPET data is critical to ensure high-quality data and consistency across study sites. Subjectivity may explain some of the variability in the interpretation of CPET between sites and within site practitioners. Automated interpretation algorithms would support the standardisation of data interrogation.	This task force suggests use of automated centralised CPET data interpretation.
Population and sampling	Representative samples including people from the general population	Previous reference value studies often include apparently healthy subjects, where smokers and obese subjects are excluded. This selection limits the generalisability of the data to the broader population to whom it is intended to apply. Representative samples with respect to age and sex distribution, socioeconomic status, habitual physical activity levels are crucial to minimise selection bias.	Future research in this area should focus on the impact of conditions such as obesity and diseases like diabetes and hypertension on exercise outcomes. This would provide clarification on how to use CPET data to differentiate health from disease.
Assessment of potential influencing factors	In addition to basic demographic data (age, biological sex, height, weight, smoking status), medical history, geographical region, and socioeconomic status, assessment of habitual physical activity) is needed	Habitual physical activity, in particular of vigorous intensity, is associated with higher $\dot{V}_{O_2peak}$ [18]. The impact of physical activity on CPET outcomes has not been studied systematically in population-based studies. Socioeconomic status [43] is often ignored as an important influencing factor of cardiorespiratory fitness and should be assessed in future prospective studies. Neither physical activity nor socioeconomic status was available due to the retrospective nature of the study, we could not assess their impact on cardiorespiratory fitness. This task force demonstrated that geographical region improved the explained between-subject variability in $\dot{V}_{O_2peak}$ by ~5%. In univariable analysis, geographical region explained 16.3% and 20.0% of the variability in $\dot{V}_{O_2peak}$ and W_{peak}, respectively. A standardised protocol in prospective studies is needed to understand the factors that explained these differences.	This task force suggests assessing physical activity using validated instruments such as the short form of the International Physical Activity Questionnaire or Global Physical Activity Questionnaire which are feasible for large-scale studies and allow comparisons across populations [44–46]. If possible, accelerometry is preferred to measure physical activity. In addition, assessing potential sociodemographic factors, as suggested, is essential for understanding population differences in outcomes.

$\dot{V}_{O_2peak}$: peak oxygen uptake; FRIEND: Fitness Registry and the Importance of Exercise National Database; W_{peak}: peak work rate. [#]: domains considered critical and compulsory for prospective data collection; [¶]: manufacturers should notify users about the wattage during unloaded pedalling.

https://doi.org/10.1183/13993003.02463-2025

For instance, males and females aged 30–39 years 10th-percentile $\dot{V}_{O_2peak}$ values were as low as 19.1 and 15.0 mL·kg^{-1}·min^{-1}, respectively, while 5th percentile (LLN) values would yield even lower values (data not reported). Such low fitness (∼5 METs, equivalent to ∼17.5 mL·kg^{-1}·min^{-1}) is associated with a high risk of cardiovascular disease and all-cause mortality [8]. Similarly, in our dataset, a large proportion of females aged ⩾45 years had $\dot{V}_{O_2peak}$ values below this threshold, despite remaining above the LLN. Furthermore, when using 80% of the predicted value, which is often applied as a proxy for the LLN [48], we found that corresponding $\dot{V}_{O_2peak}$ values were consistently above the LLN in both sexes across the adult age range.

Altogether, these data demonstrate the limited precision of percentile-based thresholds (LLN) due to wide interindividual variability in apparently healthy populations, questioning their usefulness for defining low cardiorespiratory fitness. Novel approaches for estimating cardiorespiratory fitness, such as the first percentile value for $\dot{V}_{O_2peak}$ ($\dot{V}_{O_2}$Q), show strong associations with all-cause mortality, do not rely on reference equations, and warrant further investigation across different patient populations [50]. As with other physiological assessments, CPET results should be interpreted as an adjunct to, rather than a substitute for, comprehensive clinical evaluation. Clinicians should integrate CPET findings with the patient's symptoms, medical history, and other diagnostic information, rather than relying on CPET outcomes in isolation.

Overall, the operational models that included age, sex, weight and height explained 68.8% and 67.5% of the variability in $\dot{V}_{O_2peak}$ and W_{peak}, respectively. Although, this presents a relatively high proportion of explained variance, significant heterogeneity between sites remained. Geographical region, equipment, and averaging method for peak exercise values contributed to this variability. Including these variables in the models improved the overall fit, increasing the explained variability by 6.5% for $\dot{V}_{O_2peak}$ and 10% for W_{peak}, respectively. Among these factors, equipment (*i.e.* metabolic cart for $\dot{V}_{O_2peak}$ and ergometer for W_{peak}) and geographical region had the strongest effects. Notably, the influence of different exercise testing protocols on model fit was negligible, aligning with previous research showing no differences in peak exercise outcomes between ramp and minute-by-minute incremental cycling protocols [43, 51]. This suggests that other factors not captured in our analysis contribute to the observed variability. Potential contributors include differences in participant sampling and recruitment (supplementary table S6), socioeconomic status [43], habitual physical activity levels [18], environmental factors like air pollution [51, 52], and genetic influences. Notably, ∼50% of the variance in cardiorespiratory fitness is thought to be heritable [53], which is challenging to measure.

Our retrospective data collection revealed several limitations and methodological inconsistencies across sites in Asia, Europe, North and South America despite the availability of CPET standardisation documents [1–3]. These issues prevented us from combining the different datasets to develop reference equations for $\dot{V}_{O_2peak}$ and W_{peak}. Still, our finding of large and not thoroughly explained variability of $\dot{V}_{O_2peak}$ and W_{peak} serves to shed light on the inconsistencies of current definitions of low, moderate and high cardiorespiratory fitness [8]. An in-depth understanding of the factors driving between-subject variability in CPET outcomes can only be achieved through prospective, multicountry population-based research.

This task force proposes a framework for the prospective collection of CPET outcomes to establish age-specific reference ranges in healthy individuals. It incorporates population-based sampling, strictly standardised testing protocols, rigorous quality control, and centralised data overread and analysis to overcome the limitations of retrospective research, and to minimise between-subject variability in cardiorespiratory fitness. The extent to which between-subject variability can be reduced through strict standardisation of testing protocols, equipment, and research methodologies remains unclear.

Limitations

This study has several limitations inherent to its retrospective design. Despite a growing number of CPET reference datasets, including >80 000 individuals [12], our sample size was restricted to datasets voluntarily submitted to the GLI task force. Data governance regulations and authorship policies prevented contributions from several sites, restricting the comprehensiveness and applicability of this dataset; a common challenge of ERS task forces [29]. However, given the magnitude of between-site variability, we do not believe a larger sample size would lower the observed variability in cardiorespiratory fitness outcomes. Substantial between-site variability was probably influenced by differences in participant recruitment procedures in the original studies, geographical region, and testing equipment. Unmeasured confounders, as well as residual confounding of the included variable (limited sample size in some age ranges), may also have influenced the results.

Furthermore, our goal was to include healthy, nonelite athletes; however, given the retrospective approach employed, we could not use any validated instruments or criteria to identify or exclude endurance athletes, nor can we be certain that individuals with chronic conditions were fully excluded. Finally, we did not have access to the raw CPET data to verify its quality and accuracy. One centre, for instance, provided data with a high percentage of individuals exhibiting high RERs at peak exercise; however, $\dot{V}_{O_2peak}$ values were within the expected range and appeared unaffected. The elevated RERs may have led to an overestimation of the number of subjects with maximal effort based on our pre-specified criteria. Finally, as part of the methods of this task force, we defined maximal effort using the RER and heart rate at peak exercise, which has resulted in uncertainties about true individual maximal effort from the retrospective data. It appeared that neither of these criteria effectively discriminates between subjects with higher and lower cardiorespiratory fitness (supplementary figure S2).

Future directions

Given the substantial variability in CPET data collection practices, further retrospective efforts to establish universally applicable all-age reference equations for cardiorespiratory fitness are unlikely to succeed. Therefore, we propose a methodological framework for establishing high-quality, all-age, multicountry reference ranges for cardiorespiratory fitness. This can only be achieved through prospective data collection within the general population using strictly standardised CPET protocols, calibration and verification procedures, that includes other quality assurance methods such as metabolic carts, uniform analytical approaches and central data overread. Further standardisation of CPET analysis and reporting strategies is necessary to overcome these limitations [48].

Conclusions

A standardised approach in measuring and interpreting CPET outcomes is needed to facilitate prospective data collection of representative populations using standardised CPET protocols, rigorous quality control, and centralised data processing and analysis to generate clinically meaningful reference ranges. Our findings from this ERS task force suggest that current practices of comparing measured CPET values to general population norms has significant limitations. Instead, we advocate for clinically driven models of cardiorespiratory fitness interpretation, focusing on longitudinal tracking and outcome-based risk stratification.

Acknowledgements: We kindly thank David Yabar, Olivier Laflamme, and Noah Johnson (Dalhousie University, Halifax, NS, Canada) for their support during the data collection process.

This document was endorsed by the ERS Executive Committee on 15 January 2026, and by the ATS on 20 February 2026.

Contributors to the working group (listed in alphabetical order): Piergiuseppe Agostoni, Heart Failure Unit, Centro Cardiologico Monzino IRCCS, Milan, Italy and Department of Clinical Sciences and Community Health, Cardiovascular Section, University of Milan, Milan, Italy; Santiago C. Arce, Instituto de Investigaciones Médicas Alfredo Lanari, Universidad de Buenos Aires, Argentina; Dimitrios Benas, 2nd Department of Cardiology, Medical School, National and Kapodistrian University of Athens, Attikon Hospital, Athens, Greece; Danilo C. Berton, Unidade de Fisiologia Pulmonar, Hospital de Clínicas de Porto Alegre, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil; Joel Blanchard, Department of Pediatrics, Faculty of Medicine and Health Sciences, University of Sherbrooke, and Sherbrooke University Hospital Research Center, Sherbrooke, Canada; Faculty of Physical Activity Sciences, Université de Sherbrooke, Sherbrooke, Canada; Paul Burns, Respiratory and Sleep Physiology Dept, Royal Hospital for Children, Glasgow, UK; Danielle S. Burstein, Division of Pediatric Cardiology, The University of Vermont Medical Center, Burlington, VT, USA; Virginie Carrier, Centre de recherche du Centre hospitalier universitaire de Sherbrooke, Sherbrooke, QC, Canada; Christopher B. Cooper, Departments of Medicine and Physiology, David Geffen School of Medicine, University of California Los Angeles, Los Angeles, CA, USA; Frederic Dallaire, Department of Pediatrics, Faculty of Medicine and Health Sciences, Université de Sherbrooke, and Centre de recherche du Centre hospitalier universitaire de Sherbrooke, Sherbrooke, Sherbrooke, QC, Canada; Yaoshan Dun, Division of Cardiac Rehabilitation, Department of Physical Medicine and Rehabilitation, Xiangya Hospital Central South University, Changsha, China; National Clinical Research Center for Geriatric Disorders, Xiangya Hospital of Central South University, Changsha, China; Elisabeth Edvardsen, Department of Sports Medicine, Norwegian School of Sport Sciences, Oslo, Norway; Department of Pulmonary Medicine, Oslo University Hospital, Ullevål, Oslo, Norway; Ralf Ewert, Cardiology, Pneumology, Infectious Diseases and Intensive Care Medicine, Dept of Internal Medicine B, University Hospital Greifswald, Greifswald, Germany; Eloara V.M. Ferreira, Pulmonary Function and Clinical Exercise Physiology Unit (SEFICE), Division of Respiratory Diseases, Federal University of Sao Paulo (EPM-Unifesp), Sao Paulo, Brazil; Laura Gochicoa-Rangel, Instituto Nacional de Enfermedades Respiratorias Ismael Cosío Villegas, Mexico City, Mexico; Till Ittermann, Institute for Community Medicine, SHIP/

Clinical-Epidemiological Research, University of Greifswald, Greifswald, Germany; Djordje G. Jakovljevic, Research Centre for Health and Life Sciences, Coventry University, Coventry, UK; University Hospital Coventry and Warwickshire NHS Trust, Coventry, UK; Faculty of Medical Sciences, Newcastle University, Newcastle upon Tyne, UK; Silvia Cid-Juárez, Instituto Nacional de Enfermedades Respiratorias Ismael Cosío Villegas, Mexico City, Mexico; Pirjo Komulainen, Kuopio Research Institute of Exercise Medicine, Kuopio, Finland; Suixin Liu, Division of Cardiac Rehabilitation, Department of Physical Medicine and Rehabilitation, Xiangya Hospital Central South University, Changsha, China; National Clinical Research Center for Geriatric Disorders, Xiangya Hospital of Central South University, Changsha, China; Guy MacGowan, Faculty of Medical Sciences, Newcastle University, Newcastle upon Tyne, UK; Freeman Hospital, Newcastle upon Tyne Hospital NHS Foundation Trust, Newcastle upon Tyne, UK; Michael J. Morris, Pulmonary/Critical Care Service, Department of Medicine, Brooke Army Medical Center, JBSA Fort Sam, Houston, TX, USA; Stephen M. Paridon, Division of Cardiology, Department of Pediatrics, The Children's Hospital of Philadelphia, PA, USA; Heikki Pentikainen, Kuopio Research Institute of Exercise Medicine, Kuopio, Finland; Prapaporn Pornsuriyasak, Department of Clinical Epidemiology and Biostatistics and Department of Medicine, Faculty of Medine, Ramathibodi Hospital, Mahidol University, Bangkok, Thailand; Luis Puente-Maestu, Respiratory Department, Hospital General Universitario Gregorio Marañón, Madrid, Spain; Universidad Complutense de Madrid, Medical School, Madrid, Spain; Elisabetta Salvioni, Heart Failure Unit, Centro Cardiologico Monzino IRCCS, Milan, Italy; Kai Savonen, Kuopio Research Institute of Exercise Medicine, Kuopio, Finland; School of Medicine, Institute of Public Health and Clinical Nutrition, University of Eastern Finland, Kuopio, Finland; Beate Stubbe, Cardiology, Pneumology, Infectious Diseases and Intensive Care Medicine, Dept of Internal Medicine B, University Hospital Greifswald, Greifswald, Germany; Helen Triantafyllidi, 2nd Department of Cardiology, Medical School, National and Kapodistrian University of Athens, Attikon Hospital, Athens, Greece; Ryan Welch, Respiratory Services, Te Whatu Ora Te Toka Tumai Auckland, Auckland, New Zealand; Medicine, The University of Auckland, Auckland, New Zealand; Christina White, Pulmonary/Critical Care Service, Department of Medicine, Brooke Army Medical Center, JBSA Fort Sam, Houston, TX, USA; Matthias Wilhelm, Medical Division Rehabilitation and Sports Medicine, Inselspital University Hospital, University of Bern, Bern, Switzerland.

Ethics statement: Contributors were required to obtain approval from the local ethics committee to contribute 224 data to the GLI Network. All data were pseudo-anonymised.

Conflict of interest: T. Radtke reports support for attending meetings from the European Cystic Fibrosis Society, and is co-chair of the exercise working group of the European Cystic Fibrosis Society. S. Stanojevic reports support for the present study from CIHR, consultancy fees from BiomX, payment or honoraria for lectures, presentations, manuscript writing or educational events from Vyaire Medicine and GOLD, support for attending meetings from GOLD, participation on a data safety monitoring board or advisory board with Ndd technologies, and is a member of the American Thoracic Society Pulmonary Function Testing Committee, chair of the European Respiratory Society Global Lung Function Initiative Clinical Research Collaboration, statistical editor for *Thorax*, associate editor for the *European Respiratory Journal* and a junior associate editor for the *Canadian Journal of Respiratory, Critical Care and Sleep Medicine*. P. Agostoni reports grants from European Union – Next Generation EU (NRRP M6C2 Investment 2.1), consultancy fees from Schiller, payment or honoraria for lectures, presentations, manuscript writing or educational events from CPX International, payment for expert testimony from AstraZeneca, and support for attending meetings from Società Italiana di Cardiologia and CPX International. B. Button reports consultancy fees from Vitalograph, payment or honoraria for lectures, presentations, manuscript writing or educational events from Vyaire, and support for attending meetings from Vyaire and Vitalograph. C.B. Cooper reports support for the present study from NIH/NHLBI, Foundation NIH and COPD Foundation, royalties or licenses from Cambridge University Press, consultancy fees from AstraZeneca, payment or honoraria for lectures, presentations, manuscript writing or educational events from GlaxoSmithKline and Chulalongkorn University, payment for expert testimony from various law firms, support for attending meetings from AstraZeneca, patents planned, issued or pending (Systems and methods for determining time constant for oxygen uptake kinetics, United States Provisional Patent Application 63/409,865), and participation on a data safety monitoring board or advisory board with Nuvaira, Horizon Therapeutics, MGC Diagnostics, Chiesi, Herbalife Nutrition Institute, Respiree, Aer Therapeutics, Genentec, RS Biotherapeutics and Verona. L.H. Degani–Costa reports payment or honoraria for lectures, presentations, manuscript writing or educational events from Chiesi, Trudell Medical International, AstraZeneca and GSK, and support for attending meetings from Oxyvie, Chiesi and Elivie. P. Laveneziana reports payment or honoraria for lectures, presentations, manuscript writing or educational events from Chiesi, Trudell Medical International, AstraZeneca and GSK, and support for attending meetings from Oxyvie, Chiesi and Elivie. Z.L. Saynor reports participation on a data safety monitoring board or advisory board with Rugby Players Association England, leadership or fiduciary role with Clinical Exercise Physiology UK, and receipt of equipment, materials, drugs, medical writing, gifts or other services from TidalSense Ltd. H. Hebestreit reports grants from Vertex Pharmaceuticals, payment or honoraria for lectures, presentations, manuscript writing or educational events from Chiesi GmbH, support for attending meetings from German Paediatric Society DGKJ, European Respiratory Society and European Reference Network LUNG, and leadership or fiduciary role with German Pediatric Society

(Commission on Rare Diseases) and Working Group of Centers for Rare Diseases in Germany. The remaining authors have no potential conflicts of interest to disclose.

Support statement: Funding was received from the European Respiratory Society, task force TF-2021-09. Funding information for this article has been deposited with the Open Funder Registry.

References

1 Radtke T, Crook S, Kaltsakas G, *et al.* ERS statement on standardisation of cardiopulmonary exercise testing in chronic lung diseases. *Eur Respir Rev* 2019; 28: 180101.
2 Pritchard A, Burns P, Correia J, *et al.* ARTP statement on cardiopulmonary exercise testing 2021. *BMJ Open Respir Res* 2021; 8: e001121.
3 American Thoracic Society; American College of Chest Physicians. ATS/ACCP statement on cardiopulmonary exercise testing. *Am J Respir Crit Care Med* 2003; 167: 211–277.
4 Neder JA. Cardiopulmonary exercise testing applied to respiratory medicine: myths and facts. *Respir Med* 2023; 214: 107249.
5 Neder JA, Berton DC, Rocha A, *et al.* Abnormal pattern of response to incremental exercise. *In*: Palange P, Laveneziana P, Neder JA, *et al.*, eds. Clinical Exercise Testing (ERS Monograph). Sheffield, European Respiratory Society, 2018; pp. 34–58.
6 Harvie D, Levett DZH. Exercise testing for pre-operative evaluation. *In:* Palange P, Laveneziana P, Neder JA, *et al.*, eds. Clinical Exercise Testing (ERS Monograph). Sheffield, European Respiratory Society, 2018; pp. 251–279.
7 Laveneziana P, Di Paolo M, Palange P. The clinical value of cardiopulmonary exercise testing in the modern era. *Eur Respir Rev* 2021; 30: 200187.
8 Ross R, Blair SN, Arena R, *et al.* Importance of assessing cardiorespiratory fitness in clinical practice: a case for fitness as a clinical vital sign: a scientific statement from the American Heart Association. *Circulation* 2016; 134: e653–e699.
9 Kaminsky LA, Arena R, Myers J, *et al.* Updated reference standards for cardiorespiratory fitness measured with cardiopulmonary exercise testing: data from the Fitness Registry and the Importance of Exercise National Database (FRIEND). *Mayo Clin Proc* 2022; 97: 285–293.
10 Blanchard J, Blais S, Chetaille P, *et al.* New reference values for cardiopulmonary exercise testing in children. *Med Sci Sports Exerc* 2018; 50: 1125–1133.
11 Mylius CF, Krijnen WP, van der Schans CP, *et al.* Peak oxygen uptake reference values for cycle ergometry for the healthy Dutch population: data from the LowLands Fitness Registry. *ERJ Open Res* 2019; 5: 00056-2018.
12 Takken T, Mylius CF, Paap D, *et al.* Reference values for cardiopulmonary exercise testing in healthy subjects – an updated systematic review. *Expert Rev Cardiovasc Ther* 2019; 17: 413–426.
13 Kokkinos P, Kaminsky LA, Arena R, *et al.* A new generalized cycle ergometry equation for predicting maximal oxygen uptake: the Fitness Registry and the Importance of Exercise National Database (FRIEND). *Eur J Prev Cardiol* 2018; 25: 1077–1082.
14 Vainshelboim B, Arena R, Kaminsky LA, *et al.* Reference standards for ventilatory threshold measured with cardiopulmonary exercise testing: the Fitness Registry and the Importance of Exercise: a National Database. *Chest* 2020; 157: 1531–1537.
15 Rossi Neto JM, Tebexreni AS, Alves ANF, *et al.* Cardiorespiratory fitness data from 18,189 participants who underwent treadmill cardiopulmonary exercise testing in a Brazilian population. *PLoS One* 2019; 14: e0209897.
16 Rapp D, Scharhag J, Wagenpfeil S, *et al.* Reference values for peak oxygen uptake: cross-sectional analysis of cycle ergometry-based cardiopulmonary exercise tests of 10 090 adult German volunteers from the Prevention First Registry. *BMJ Open* 2018; 8: e018697.
17 Loe H, Nes BM, Wisløff U. Predicting VO_{2peak} from submaximal- and peak exercise models: the HUNT 3 fitness study, Norway. *PLos One* 2016; 11: e0144873.
18 Wagner J, Knaier R, Infanger D, *et al.* Novel CPET reference values in healthy adults: associations with physical activity. *Med Sci Sports Exerc* 2021; 53: 26–37.
19 Hakola L, Komulainen P, Hassinen M, *et al.* Cardiorespiratory fitness in aging men and women: the DR's EXTRA study. *Scand J Med Sci Sports* 2011; 21: 679–687.
20 Koch B, Schäper C, Ittermann T, *et al.* Reference values for cardiopulmonary exercise testing in healthy volunteers: the SHIP study. *Eur Respir J* 2009; 33: 389–397.
21 Hansen JE, Sue DY, Wasserman K. Predicted values for clinical exercise testing. *Am Rev Respir Dis* 1984; 129: S49–S55.
22 Lewthwaite H, Benedetti A, Stickland MK, *et al.* Normative peak cardiopulmonary exercise test responses in Canadian adults aged ⩾40 years. *Chest* 2020; 158: 2532–2545.
23 Waterfall JL, Burns P, Shackell D, *et al.* The risks of applying normative values in paediatric cardiopulmonary exercise testing: a case report. *ERJ Open Res* 2020; 6: 00333-2020.

24 Harris PA, Taylor R, Minor BL, *et al.* The REDCap consortium: building an international community of software platform partners. *J Biomed Inform* 2019; 95: 103208.
25 Tanaka H, Monahan KD, Seals DR. Age-predicted maximal heart rate revisited. *J Am Coll Cardiol* 2001; 37: 153–156.
26 Bongers BC, Hulzebos HJ, van Brussel M, *et al.* Pediatric Norms for Cardiopulmonary Exercise Testing: In Relation to Gender and Age. 's Hertogenbosch, Uitgeverij BOXpress, 2014.
27 Kuczmarski RJ, Ogden CL, Guo SS, *et al.* 2000 CDC growth charts for the United States: methods and development. *Vital Health Stat 11* 2002; 246: 1–190.
28 Ramsey KA, Stanojevic S, Chavez L, *et al.* ERS technical standard: Global Lung Function Initiative reference values for multiple breath washout indices. *Eur Respir J* 2024; 64: 2400524.
29 Högman M, Bowerman C, Chavez L, *et al.* ERS technical standard: Global Lung Function Initiative reference values for exhaled nitric oxide fraction (F_{ENO50}). *Eur Respir J* 2024; 63: 2300370.
30 Bhakta NR, Bime C, Kaminsky DA, *et al.* Race and ethnicity in pulmonary function test interpretation: an official American Thoracic Society statement. *Am J Respir Crit Care Med* 2023; 207: 978–995.
31 Bowerman C, Bhakta NR, Brazzale D, *et al.* A race-neutral approach to the interpretation of lung function measurements. *Am J Respir Crit Care Med* 2023; 207: 768–774.
32 Quanjer PH, Stanojevic S, Cole TJ, *et al.* Multi-ethnic reference values for spirometry for the 3-95-yr age range: the global lung function 2012 equations. *Eur Respir J* 2012; 40: 1324–1343.
33 Van Hooren B, Souren T, Bongers BC. Accuracy of respiratory gas variables, substrate, and energy use from 15 CPET systems during simulated and human exercise. *Scand J Med Sci Sports* 2024; 34: e14490.
34 Revill SM, Morgan MD. Biological quality control for exercise testing. *Thorax* 2000; 55: 63–66.
35 Clinical exercise testing with reference to lung diseases: indications, standardization and interpretation strategies. ERS Task Force on Standardization of Clinical Exercise Testing. European Respiratory Society. *Eur Respir J* 1997; 10: 2662–2689.
36 Beltz NM, Gibson AL, Janot JM, *et al.* Graded exercise testing protocols for the determination of V_{O2max}: historical perspectives, progress, and future considerations. *J Sports Med* 2016; 2016: 3968393.
37 Whipp BJ, Davis JA, Torres F, *et al.* A test to determine parameters of aerobic function during exercise. *J Appl Physiol Respir Environ Exerc Physiol* 1981; 50: 217–221.
38 Myers J, Buchanan N, Walsh D, *et al.* Comparison of the ramp *versus* standard exercise protocols. *J Am Coll Cardiol* 1991; 17: 1334–1342.
39 Cooper CB, Abrazado M, Legg D, *et al.* Development and implementation of treadmill exercise testing protocols in COPD. *Int J Chron Obstruct Pulmon Dis* 2010; 5: 375–385.
40 Myers J, Walsh D, Sullivan M, *et al.* Effect of sampling on variability and plateau in oxygen uptake. *J Appl Physiol* 1990; 68: 404–410.
41 Smart NA, Jeffriess L, Giallauria F, *et al.* Effect of duration of data averaging interval on reported peak VO_2 in patients with heart failure. *Int J Cardiol* 2015; 182: 530–533.
42 Hill DW, Stephens LP, Blumoff-Ross SA, *et al.* Effect of sampling strategy on measures of VO_{2peak} obtained using commercial breath-by-breath systems. *Eur J Appl Physiol* 2003; 89: 564–569.
43 Lindgren M, Börjesson M, Ekblom Ö, *et al.* Physical activity pattern, cardiorespiratory fitness, and socioeconomic status in the SCAPIS pilot trial – a cross-sectional study. *Prev Med Rep* 2016; 4: 44–49.
44 Craig CL, Marshall AL, Sjöström M, *et al.* International physical activity questionnaire: 12-country reliability and validity. *Med Sci Sports Exerc* 2003; 35: 1381–1395.
45 Armstrong T, Bull F. Development of the World Health Organization Global Physical Activity Questionnaire (GPAQ). *J Public Health* 2006; 14: 66–70.
46 Whiting S, Mendes R, Abu-Omar K, *et al.* Physical inactivity in nine European and Central Asian countries: an analysis of national population-based survey results. *Eur J Public Health* 2021; 31: 846–853.
47 Cooper CB, Storer TW. Exercise Testing and Interpretation. A Practical Approach. 5th edn. Cambridge, Cambridge University Press, 2001.
48 Staes M, Gyselinck I, Goetschalckx K, *et al.* Identifying limitations to exercise with incremental cardiopulmonary exercise testing: a scoping review. *Eur Respir Rev* 2024; 33: 240010.
49 Stanojevic S, Graham BL, Cooper BG, *et al.* Official ERS technical standards: Global Lung Function Initiative reference values for the carbon monoxide transfer factor for Caucasians. *Eur Respir J* 2017; 50: 1700010.
50 Knox-Brown B, Barnes J, Harding C, *et al.* Peak $\dot{V}_{O_2}Q$: a new approach for the interpretation of cardiorespiratory fitness estimates. *Exp Physiol* 2026; 111: 1454–1465.
51 Cipryan L, Litschmannova M, Barot T, *et al.* Air pollution, cardiorespiratory fitness and biomarkers of oxidative status and inflammation in the 4HAIE study. *Sci Rep* 2024; 14: 9620.
52 Gao Y, Chan EYY, Zhu Y, *et al.* Adverse effect of outdoor air pollution on cardiorespiratory fitness in Chinese children. *Atmos Environ* 2013; 64: 10–17.
53 Bouchard C, Daw EW, Rice T, *et al.* Familial resemblance for V_{O2max} in the sedentary state: the HERITAGE family study. *Med Sci Sports Exerc* 1998; 30: 252–258.