Article

Criterion Validity and Responsiveness of Estimated Cardiorespiratory Fitness Models in Patients with Inflammatory Joint Disease

Kristine Røren Nordén [1,2,3,*], Hanne Dagfinrud [1,2], Anne Grete Semb [4], Jonny Hisdal [5,6], George S. Metsios [7], Joseph Sexton [8], Camilla Fongen [1], Emilie Andrea Bakke [1] and Anne Therese Tveter [1,9]

1   Center for Treatment of Rheumatic and Musculoskeletal Diseases (REMEDY), Norwegian National Advisory Unit on Rehabilitation in Rheumatology, Diakonhjemmet Hospital, 0319 Oslo, Norway
2   Institute of Health and Society, Faculty of Medicine, University of Oslo, 0316 Oslo, Norway
3   Norwegian National Unit for Rehabilitation for Rheumatic Patients with Special Needs, Diakonhjemmet Hospital, 0319 Oslo, Norway
4   Preventive Cardio-Rheuma Clinic, Center for Treatment of Rheumatic and Musculoskeletal Diseases (REMEDY), Division of Rheumatology and Research, Diakonhjemmet Hospital, 0319 Oslo, Norway
5   Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, 0318 Oslo, Norway; jonny.hisdal@medisin.uio.no
6   Department of Vascular Surgery, Oslo University Hospital—Aker, 0586 Oslo, Norway
7   Department of Nutrition and Dietetics, University of Thessaly, 42132 Trikala, Greece
8   Center for Treatment of Rheumatic and Musculoskeletal Diseases (REMEDY), Division of Rheumatology and Research, Diakonhjemmet Hospital, 0319 Oslo, Norway
9   Faculty of Health Sciences, Institute of Rehabilitation Science and Health Technology, Oslo Metropolitan University, 0166 Oslo, Norway
*   Correspondence: kristineroren.norden@diakonsyk.no

Abstract: Cardiorespiratory fitness (CRF) is an excellent marker of overall health. This study aimed to assess criterion validity and responsiveness of estimated CRF models (eCRF) in patients with inflammatory joint disease (IJD). CRF was measured directly as peak oxygen uptake (VO_{2peak}) by a Cardiopulmonary Exercise Test (CPET), while one generic eCRF model ($eCRF_{GEN}$) and two disease-specific eCRF models ($eCRF_{ALT}$ and $eCRF_{PGA}$) were used to estimate CRF at baseline and after 3 months in 55 Norwegian patients with IJD. Moderate correlations were observed between $eCRF_{GEN}$, $eCRF_{ALT}$, $eCRF_{PGA}$, and VO_{2peak} at baseline (ICC 0.60, 0.64 and 0.62, respectively) and 3 months (ICC 0.62, 0.65 and 0.57, respectively). All eCRF models overestimated measured VO_{2peak}, and there was large variability in agreement of individual measurements at baseline and at 3 months. Weak correlations were observed for responsiveness of $eCRF_{GEN}$ (ICC 0.39), $eCRF_{ALT}$ (ICC 0.40) and $eCRF_{PGA}$ (ICC 0.39). Mean differences between change in eCRF models and change in VO_{2peak} were small, but the wide limits of agreement exceeded the pre-defined clinically acceptable margins. The eCRF models possessed adequate ability to detect ≥3.5 mL/kg/min improvement in VO_{2peak}. eCRF may suffice for group-level assessment, but caution is advised when applying eCRF to individual patients with IJD.

Keywords: cardiorespiratory fitness; inflammatory joint disease; rheumatoid arthritis; spondyloarthritis; validity; responsiveness

1. Introduction

Cardiorespiratory fitness (CRF) reflects the functional capacity of the cardiovascular and respiratory systems to transport and deliver oxygen to working muscles during physical activity [1]. Comprehensive epidemiological data have shown consistent inverse associations between CRF and the risk of cardiovascular disease, cancer mortality, and all-cause mortality [2–5]. Furthermore, measures of CRF can offer valuable insights into the effectiveness of various interventions aimed at improving health outcomes. In light of this,

CRF has gained recognition as an important indicator of overall health, and the American Heart Association advocates measures of CRF as a clinical vital sign [6].

Inflammatory joint diseases (IJD), including rheumatic diseases such as rheumatoid arthritis, spondyloarthritis, and psoriatic arthritis are characterized by joint inflammation, pain, fatigue, and varying levels of physical disability [7–9]. Additionally, IJD associates with an increased risk of CVD, influenced by systemic inflammation and a higher prevalence of classic CVD risk factors [10–12]. There are multiple observations of inferior levels of CRF in patients with rheumatoid arthritis [13,14], spondyloarthritis [15,16], and combined diagnoses [17], and low levels of CRF may contribute to the elevated CVD and mortality risk [18,19]. Accurate assessment of CRF is therefore of clinical value in IJD care, as it can shed light on CVD risk and measure the effects of interventions designed to enhance CRF.

The criterion method to assess CRF is the Cardiopulmonary Exercise Test (CPET), which measures peak oxygen uptake (VO_{2peak}) by indirect calorimetry during a progressive exercise test to maximal exertion [1,6]. However, CPET is time-consuming and requires specialized equipment and trained personnel, thereby restricting its applicability to large-scale studies and primary care settings [6]. To overcome these limitations and facilitate broader implementation of CRF assessment, researchers have explored the potential of non-exercise estimated CRF (eCRF) algorithms. These models often apply age, gender, and a measure of body composition, combined with self-reported habitual exercise and/or physical activity, to estimate an individual's CRF without the need for an exercise test [6]. Among the available eCRF algorithms, a generic model developed in a Norwegian population [20,21] has been recommended for use [22]. In an attempt to address characteristics of patients with rheumatoid arthritis that may associate with CRF, the original eCRF model [20] has been adapted in two rheumatoid arthritis-specific models [23].

A trend to overestimate CRF in individuals at the lower end of the fitness spectrum is reported across various eCRF models [22]. As the accuracy of eCRF may vary depending on the characteristics of the population studied, it is essential to verify the validity of eCRF models in target populations with underlying health conditions. Before incorporating eCRF in IJD settings, it is equally important to assess the ability of eCRF models to detect positive changes in CRF resulting from a more active lifestyle, as well as the ability to capture declining levels of CRF that may call for additional assessments and referral to exercise interventions. Despite the simplicity of eCRF, the responsiveness, i.e., longitudinal validity [24], of these models remains largely unknown [21], and requires further investigation. Accordingly, the purpose of the current study was to evaluate the criterion validity and responsiveness of eCRF compared to the gold standard VO_{2peak} among patients with IJD. We hypothesized that, at the group level, eCRF models would demonstrate a tendency to overestimate measured VO_{2peak} and change in VO_{2peak}.

2. Materials and Methods

2.1. Patients and Study Setting

Data from 55 Norwegian patients who attended baseline and 3-month study visits in the ExeHeart randomized controlled trial (RCT) (ClinicalTrials.gov NCT04922840) were included in the present analyses. Approval of the ExeHeart trial, including the current study, was granted by the Regional Committee for Medical and Health Research Ethics (201227) and the Data Protection Officer at Diakonhjemmet Hospital (reg.no. 00397). Study procedures adhered to the principles of the Helsinki Declaration and all patients provided informed written consent.

Details regarding the study sample, patient enrollment and outcome assessment are fully described in the ExeHeart study protocol [25]. In short, the aim of the ExeHeart trial was to assess the effect of 12 weeks of high-intensity interval training on CRF, classic cardiovascular disease risk factors and disease activity in patients with IJD. Trial results are being published elsewhere. The study enrolled individuals that met the following criteria: aged 18–70 years, with a body mass index (BMI) ranging from 18.5 to 40, IJD diagnosis (rheumatoid arthritis, spondyloarthritis or psoriatic arthritis), capable of walking unaided

for ≥15 min, and proficient in Norwegian and/or English. Patients with lower-extremity injury or surgery within the last 12 months, primary neurological disease, cognitive impairment, engagement in high intensity exercise in the three months prior to study inclusion and/or contraindications to maximal exercise testing [1] were excluded from study participation. Although no formal power calculation was conducted specifically for the current study, a sample size of 55 patients aligns with the recommended number for criterion validity approach in COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) guidelines [26]. The study was reported according to Guidelines for Reporting Reliability and Agreement Studies [27] and the COSMIN Reporting guideline for studies on measurement properties of patient reported outcome measures [28].

2.2. Outcome Assessment and Interim Period

Study visits took place at Diakonhjemmet Hospital, Oslo, Norway, from August 2021 to November 2022. During the interim period from baseline assessment to the 3-month session, 27 out of the 55 patients participated in a 12-week high-intensity interval training program (exercise group), while the remaining 28 patients formed the control group and did not receive any targeted exercise intervention. The 3-month sessions were scheduled at the same time of day as the corresponding baseline sessions. Outcome assessors (KRN, CF, and EAB) were physiotherapists with extensive experience in rheumatology and exercise testing, and were blinded to group allocation.

2.3. Demographic Variables

In the week leading up to study visits, patients answered a digital questionnaire [25], and items regarding education level, medication usage, and pain and fatigue over the past week (numerical rating scale 0–10, 0 = best) were included to describe the study sample, along with the following: IJD diagnosis and comorbidities were extracted from the patient's medical record. IJD-specific composite measures were applied to measure clinical disease activity, and disease activity was further categorized as remission, low, moderate or high using instrument-specific cut-off values. Further examinations included blood chemistry (lipid profile, C-reactive protein and erythrocyte sedimentation rate) and blood pressure measurements [25].

2.4. Criterion Measurement of CRF

CRF was measured directly as VO_{2peak} in mL/kg/min by the criterion method CPET [6,29]. Test equipment was calibrated according to the manufacturer's specifications with gas calibration every third hour, and automatic calibration of volume sensors between each test. Accuracy of the volume sensors was confirmed weekly through manual calibration with a 3 L syringe (Hans Rudolph Inc., Shawnee, KS, USA).

Patients were instructed to have a two-hour fast since their last meal, avoid the use of nicotine and caffeine for at least four hours prior to the test, and refrain from participating in vigorous exercise within the 24 h preceding the CPET. Pretest spirometry and maximal voluntary ventilation were conducted in accordance with recommendations [29,30]. The CPET was performed on a treadmill (PPS 55 Woodway, Würzburg, Germany), with 12-lead electrocardiogram (Customed cardio 300 BT_A, CareFusion, Ottobrunn, Germany), blood pressure monitor (Suntech Tango M2, SunTech Medical, Morrisville, NC, USA), and earlobe pulse oximetry. Ventilatory parameters and gas exchange data were collected breath-by-breath using a Hans Rudolph two-way mask (7450 series, Hans Rudolph Inc., Shawnee, KS, USA), analyzed by a metabolic cart (Vyntus CPX, Vyaire Medial, Hoechenberg, Germany), and balanced over 30 s intervals. A modified Balke ramp protocol [31] was applied, starting with a short warm-up period to familiarize the patient with the treadmill and determine an appropriate initial walking speed. Following warm-up, inclination was increased by 2% every minute up to 15 or 20%. If the patient was able to continue, speed was increased by 0.5 km/h per minute until volitional exhaustion. Ventilatory reserve was calculated as (maximal minute ventilation-peak minute ventilation/maximal minute

ventilation) $\times$ 100 [29]. The criteria for VO_{2peak} was a visible plateau in oxygen uptake despite an increase in work load. Notably, such plateaus are not always observed in patient populations [32], and in the absence of a VO_2 plateau, acceptable criteria for VO_{2peak} by CPET were defined as meeting $\geq$2 of the following criteria [30]: (a) peak heart rate within 90% of predicted (220—age), (b) Borg rating of perceived exertion $\geq$9 (0–10, 10 = maximal exertion), (c) ventilatory reserve less than 15% of maximal minute ventilation, and/or (d) adherence to age- and gender-specific cut-off values for respiratory exchange ratio and post-exercise blood lactate [32].

2.5. Estimated CRF

The generic eCRF (eCRF$_{GEN}$) model was used in all patients and includes the variables age, BMI, resting heart rate, and physical activity index [20,21].

Body mass to the nearest 0.1 kg (Tanita MC-780MA Tanita Corporation, Tokyo, Japan) and height to the nearest cm (KaWe Person Check, Kirchner & Wilhelm GmbH + Co. KG, Asperg, Germany) were used to calculate BMI as kg/m^2.

Resting heart rate (beats/min) was measured after 10 min of rest in a supine position using a mobile blood pressure monitor (Mobil-o-graph PWA, I.E.M. GmbH, Stolberg, Germany) and recorded as the mean of two measurements.

From the digital questionnaire, three items regarding exercise frequency, duration and intensity were used to calculate a physical activity index ranging from 0 to 45, with higher scores indicating better physical activity levels [20,33].

In addition to the variables included in eCRF$_{GEN}$, there are two rheumatoid arthritis-specific models [23]. The first model, eCRF$_{ALT}$, includes self-reported smoking status and was explored in all patients regardless of IJD entity. The second model, eCRF$_{PGA}$, further incorporates Patient Global Assessment of disease activity (PGA), and was answered by patients with rheumatoid arthritis in our study sample. For the purpose of scrutinizing these models, smoking status was coded as 1 for "ever smoked" or "current smoker" and 0 for "never smoked" [23]. PGA was phrased as "Considering all the symptoms from your rheumatic disease during the last week, how do you think your state is?" and answered on a 0–100 mm visual analog scale with 0 anchored as "good, no symptoms" and 100 anchored as "very severe" [23]. Equations for eCRF$_{GEN}$, eCRF$_{ALT}$ and eCRF$_{PGA}$ are provided in Table 1.

Table 1. Equations for the generic and RA-specific eCRF models.

Model	eCRF Equation
eCRF$_{GEN}$ [21]	
Female	70.77 $-$ (0.244 $\times$ age) $-$ (0.749 $\times$ BMI) $-$ (0.107 $\times$ resting heart rate) + (0.213 $\times$ physical activity index)
Male	92.05 $-$ (0.327 $\times$ age) $-$ (0.933 $\times$ BMI) $-$ (0.167 $\times$ resting heart rate) + (0.257 $\times$ physical activity index)
	(female = 0, male = 1; never smoked = 0, ever smoked = 1)
eCRF$_{ALT}$ [23]	77.851 + (gender $\times$ 25.460) $-$ (age $\times$ 0.381) $-$ (age–gender interaction $\times$ 0.254) $-$ (BMI $\times$ 0.743) $-$ (resting heart rate $\times$ 0.115) $-$ (smoking $\times$ 2.154) + (physical activity index $\times$ 0.209)
eCRF$_{PGA}$ [23]	77.961 + (gender $\times$ 28.791) $-$ (age $\times$ 0.358) $-$ (age–gender interaction $\times$ 0.326) $-$ (BMI $\times$ 0.700) $-$ (resting heart rate $\times$ 0.125) $-$ (smoking $\times$ 1.854) + (physical activity index $\times$ 0.211) $-$ (PGA $\times$ 0.071)

BMI: Body mass index, eCRF$_{ALT}$: Alternative rheumatoid arthritis-specific eCRF model, eCRF$_{GEN}$: Estimated cardiorespiratory fitness, generic model, eCRF$_{PGA}$: Rheumatoid arthritis-specific eCRF model, PGA: Patient global assessment of disease activity.

2.6. Statistical Analysis

Continuous data are presented as mean with standard deviation (SD) or median with interquartile range (IQR) for skewed data, whereas counts with percentages are provided for categorical data. Paired sample *t*-tests were used to calculate differences between corresponding measures of CRF at baseline and at 3 months, as well as between the individual eCRF models and VO_{2peak}. Intraclass correlation coefficient (ICC) two-way mixed-effects models with 95% confidence interval (CI) were used to assess absolute agreement between eCRF models and VO_{2peak} at baseline and 3 months, and change from baseline to 3 months. An ICC value <0.5 indicates poor agreement, 0.5–0.75 moderate agreement, 0.75–0.9 good agreement and values >0.9 excellent agreement [34]. A limitation of using correlation to assess responsiveness involves narrowing of the variable range, as this will often lead to lower correlation coefficients [35,36]. Therefore, we conducted supplementary univariate linear regression analyses with change in VO_{2peak} as the dependent variable and change in the respective eCRF model as the independent variable. Resulting regression coefficients offer insight into predicting change in VO_{2peak} following a one-unit increase in eCRF, while the R^2 statistic gives the fraction of change in VO_{2peak} explained by a unit change in eCRF. Model assumptions were assessed graphically.

Bland–Altman plots were generated to visualize the relationship between eCRF models and VO_{2peak} at baseline and 3 months, as well as the change in eCRF models and VO_{2peak} from baseline to 3 months. Plots include the average of eCRF models and VO_{2peak} on the *x*-axis, difference between the two methods on the *y*-axis, mean bias and 95% limits of agreement [37]. Assumptions of normal distribution of differences were checked by histograms and Shapiro–Wilk tests [38]. A 3.5 mL/kg/min increase in CRF associates with a significant reduction in cardiovascular disease and mortality risks [39,40], and a range of ±3.5 mL/kg/min was defined as the clinically acceptable difference between measurement methods.

We further investigated the capability of eCRF models to detect improvement in CRF using area under the curve (AUC) obtained from empirical receiver operating curve (ROC) analysis. A threshold of ≥3.5 mL/kg/min was used to dichotomize a substantial improvement in VO_{2peak} from baseline to 3 months, and AUC ≥ 0.80 was deemed acceptable [35,41]. As there were only two observations of a CRF decline ≥ 3.5 mL/kg/min within the patient cohort, the available dataset did not facilitate ROC analysis of the ability of eCRF models to accurately identify substantial deteriorations in VO_{2peak}.

A significance threshold of *p* < 0.05 was applied for all analyses, and STATA v. 17 was used for all statistical computations.

3. Results

3.1. Demographic and Clinical Characteristics

A total of 55 patients (31 women and 24 men) with IJD, with a median age of 59 years (IQR 51–63), and a median BMI of 25 (IQR 22–31) were included. Clinical characteristics from baseline sessions are provided in Table 2 and Supplementary Files, Table S1. Mean (range) VO_{2peak} was 30.3 (17.3, 47.9) mL/kg/min at baseline and 31.6 (18.9, 49.3) mL/kg/min at 3 months. Mean (range) change in VO_{2peak} from baseline to 3 months was 1.3 (−5.3, 7.8) mL/kg/min, and 13 (24%) patients exhibited a ≥3.5 mL/kg/min improvement in VO_{2peak} (Table 3 and Supplementary Files, Table S2).

3.2. ICC Analysis

As detailed in Table 3, measured VO_{2peak} and eCRF models were moderately correlated, with ICC values ranging from 0.60 to 0.64 at baseline and 0.57 to 0.65 at 3 months. Regarding change from baseline to 3 months, correlation coefficients were lower, ranging from 0.39 to 0.40 between changes in eCRF models and changes in VO_{2peak}. In the complemental regression analysis, individual regression coefficients were statistically significant. An increment of one unit in $eCRF_{GEN}$ associated with a change of 0.56 mL/kg/min in VO_{2peak}. Corresponding coefficients for change in $eCRF_{ALT}$ and $eCRF_{PGA}$ indicated an

associated change of 0.58 and 0.68 mL/kg/min, respectively, in VO_{2peak}. The proportion of variation in VO_{2peak} change explained by change in eCRF (R^2) varied from 17 to 24%.

Table 2. Clinical characteristics from baseline sessions ($n = 55$).

Variable	
Age, years, median (IQR)	59 (51–63)
Gender, female, *n* (%)	31 (56)
Education > 12 years, *n* (%)	42 (76)
Anthropometrics	
Height, cm, mean (SD)	173 (9)
Weight, kg, median (IQR)	74 (66–91)
BMI, median (IQR)	25 (22–31)
Diagnosis	
Rheumatoid arthritis, *n* (%)	26 (47)
Spondyloarthritis, *n* (%)	16 (29)
Psoriatic arthritis, *n* (%)	13 (24)
IJD disease duration, years, median (IQR)	15 (7–30)
Disease activity categorized	
Remission, *n* (%)	21 (38)
Low, *n* (%)	15 (27)
Moderate, *n* (%)	12 (22)
High, *n* (%)	7 (13)
Comorbidity	
Diabetes, *n* (%)	3 (5)
Chronic obstructive pulmonary disease, *n* (%)	4 (7)
Inflammatory bowel disease, *n* (%)	4 (7)
Ever smoker, yes, *n* (%)	34 (62)
Resting heart rate, beats/min, mean (SD)	68 (11)
Cardiopulmonary Exercise Test	
VO_{2peak}, mL/kg/min, mean (SD)	30.3 (6.9)
VO_{2peak}, L/min, mean (SD)	2.4 (0.7)
VO_2 plateau at peak exercise, yes, *n* (%)	29 (53)
Respiratory exchange ratio, VCO_2/VO_2, mean (SD)	1.16 (0.1)
Borg RPE 0–10 (10 = maximal), median (IQR)	10 (9–10)
Peak heart rate, beats/min, mean (SD)	165 (14)
Percent of predicted peak heart rate (220-age), mean (SD)	101 (7)
Post-exercise blood lactate, mmol/L, median (IQR)	9.2 (7.3–11.7) §
Ventilatory reserve, %, mean (SD)	25 (13) ^
Physical activity index (0–45, 45 = best), median (IQR)	0 (0–15)
Numerical Rating Scale (0–10), 0 = best	
Pain, median (IQR)	2 (1–4)
Fatigue, median (IQR)	3 (1–5)
Patient Global Assessment, 0–100 mm, 0 = best, mean (SD)	29 (22) $

§ $n = 52$; values ≥ 20 mmol/L omitted. ^ $n = 54$ patients. $ $n = 24$ patients with rheumatoid arthritis. BMI: Body Mass Index, IJD: Inflammatory Joint Disease, RPE: Rating of Perceived Exertion. VCO_2: Volume of carbon dioxide production. VO_2: Volume of oxygen uptake. VO_{2peak}: Peak oxygen uptake.

3.3. Bland–Altman Analysis

At baseline, eCRF overestimated VO_{2peak} with a significant mean bias of 4.2 mL/kg/min for $eCRF_{GEN}$ and 3.1 mL/kg/min for $eCRF_{ALT}$. Conversely, the 1.6 mL/kg/min difference observed in the subsample analyzed using $eCRF_{PGA}$ was not statistically significant (Table 3). The 95% limits of agreement showed large variability of agreement between eCRF models and VO_{2peak}, and exceeded the clinical threshold of ± 3.5 mL/kg/min (Figure 1a–c).

For the 3-month measurements, all eCRF models overestimated VO_{2peak}, with a significant mean bias of 4.9 mL/kg/min for $eCRF_{GEN}$, 3.8 mL/kg/min for $eCRF_{ALT}$, and 3.1 mL/kg/min for the subsample scrutinized using $eCRF_{PGA}$ (Table 3). As reflected in Figure 2a–c, the 95% limits of agreement surpassed the clinical threshold.

Regarding change scores from baseline to 3 months, both $eCRF_{GEN}$ and $eCRF_{ALT}$ revealed non-significant mean biases of 0.7 mL/kg/min, while the subsample analyzed using $eCRF_{PGA}$ demonstrated a significant mean bias of 1.5 mL/kg/min (Table 3). The 95% limits of agreement exceeded the clinically acceptable difference (Figure 3a–c).

Table 3. Agreement between VO_{2peak} and eCRF models at baseline, 3 months, and change from baseline to 3 months. Values are presented as mean (SD) unless otherwise indicated.

Variable	Baseline (*n* = 55)	3 Months (*n* = 55)	ΔBaseline to 3 Months (*n* = 55)
VO_{2peak}, mL/kg/min	30.3 (6.9)	31.6 (7.0)	1.3 (0.5 to 2.1) [a]
$eCRF_{GEN}$, mL/kg/min	34.5 (6.5)	36.4 (6.7)	2.0 (1.4 to 2.6) [a]
ICC (95% CI)	0.60 (0.15 to 0.80)	0.62 (0.03 to 0.84)	0.39 (0.15 to 0.59)
Linear regression, regression coefficient, (95% CI)			0.56 (0.22 to 0.90)
R2			0.17
Difference $eCRF_{GEN}$—VO_{2peak}, mL/kg/min. (95% CI)	4.2 (2.8 to 5.6) [a]	4.9 (3.6 to 6.1) [a]	0.7 (−0.1 to 1.4) [a]
95% Limits of agreement	−5.7 to 14.1	−4.2 to 13.9	−5.1 to 6.4
$eCRF_{ALT}$, mL/kg/min	33.6 (8.3)	35.5 (8.4) §	1.8 (1.2 to 2.4) §[a]
ICC (95% CI)	0.64 (0.38 to 0.79)	0.65 (0.31 to 0.82) §	0.40 (0.16 to 0.60) §
Linear regression, regression coefficient, (95% CI)			0.58 (0.24 to 0.91) §
R2			0.19 §
Difference $eCRF_{ALT}$—VO_{2peak}, mL/kg/min, (95% CI)	3.1 (1.5 to 4.7) [a]	3.8 (2.2 to 5.3) §[a]	0.7 (−0.1 to 1.4) §[a]
95% Limits of agreement	−8.4 to 14.7	−7.5 to 15.0 §	−4.8 to 6.1 §

Variable	Baseline (*n* = 24)	3 months (*n* = 24)	ΔBaseline to 3 months (*n* = 24)
VO_{2peak}, mL/kg/min	27.9 (5.2)	28.6 (5.8)	0.6 (−0.5 to 1.8) [a]
$eCRF_{PGA}$, mL/kg/min	29.6 (4.6)	31.7 (4.8)	2.1 (1.3 to 3.0) [a]
ICC (95% CI)	0.62 (0.30 to 0.81)	0.57 (0.13 to 0.81)	0.39 (0.01 to 0.68)
Linear regression, regression coefficient, (95% CI)			0.68 (0.06 to 1.29) [b]
R2			0.24
Difference $eCRF_{PGA}$—VO_{2peak}, mL/kg/min, (95% CI)	1.6 (−0.1 to 3.4) [a]	3.1 (1.3 to 5.0) [a]	1.5 (0.5 to 2.5) [a]
95% Limits of agreement	−6.6 to 9.8	−5.6 to 11.9	−3.2 to 6.2

[a] Analyzed by paired samples *t*-test. [b] Bootstrap confidence intervals, 50 replications. Δ = Change. § *n* = 54 patients (smoking status missing at 3 months; *n* = 1). $eCRF_{ALT}$: Alternative rheumatoid arthritis-specific eCRF model. $eCRF_{GEN}$: Estimated cardiorespiratory fitness, generic model. $eCRF_{PGA}$: Rheumatoid arthritis-specific eCRF model, assessed in patients with rheumatoid arthritis. ICC: Intraclass Correlation Coefficient. VO_{2peak}: Peak oxygen uptake.

3.4. Area-under-the-Curve Analysis

In order to assess the performance of eCRF models in predicting a threshold improvement in VO_{2peak}, patients with a decline in VO_{2peak} of ≥3.5 mL/kg/min from baseline to 3 months were excluded. ROC curves are presented in Supplementary Files, Figure S1. Among the 53 patients analyzed using $eCRF_{GEN}$, 13 patients had an improvement in VO_{2peak} ≥ 3.5 mL/kg/min, and AUC for $eCRF_{GEN}$ was 0.82 (95% CI 0.71 to 0.93). In the 52 patients assessed using $eCRF_{ALT}$ (one excluded due to incomplete eCRF data at 3 months), 12 patients demonstrated an improvement ≥3.5 mL/kg/min in VO_{2peak}, and AUC for $eCRF_{ALT}$ was 0.83 (95% CI 0.72 to 0.94). Within the subsample of 23 patients evaluated using $eCRF_{PGA}$, 4 patients had a ≥3.5 mL/kg/min improvement in VO_{2peak}, and $eCRF_{PGA}$ yielded an AUC of 0.97 (95% CI 0.91 to 1.00).
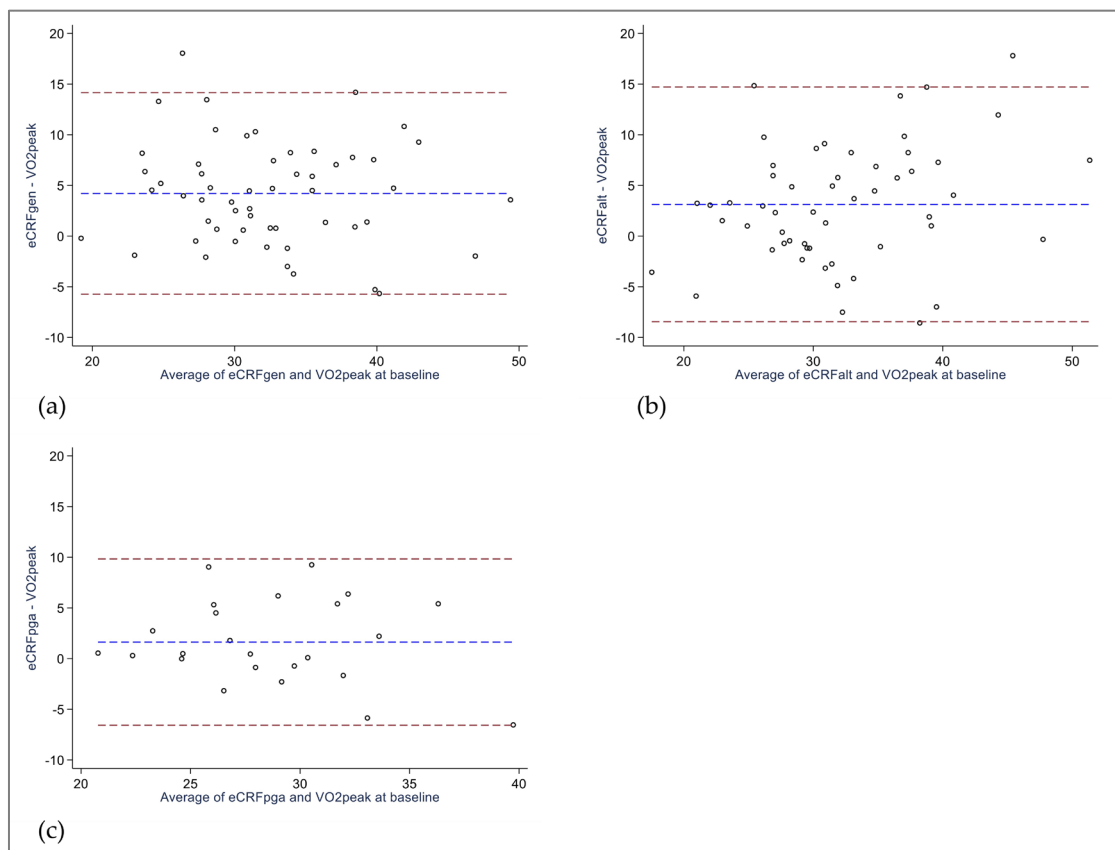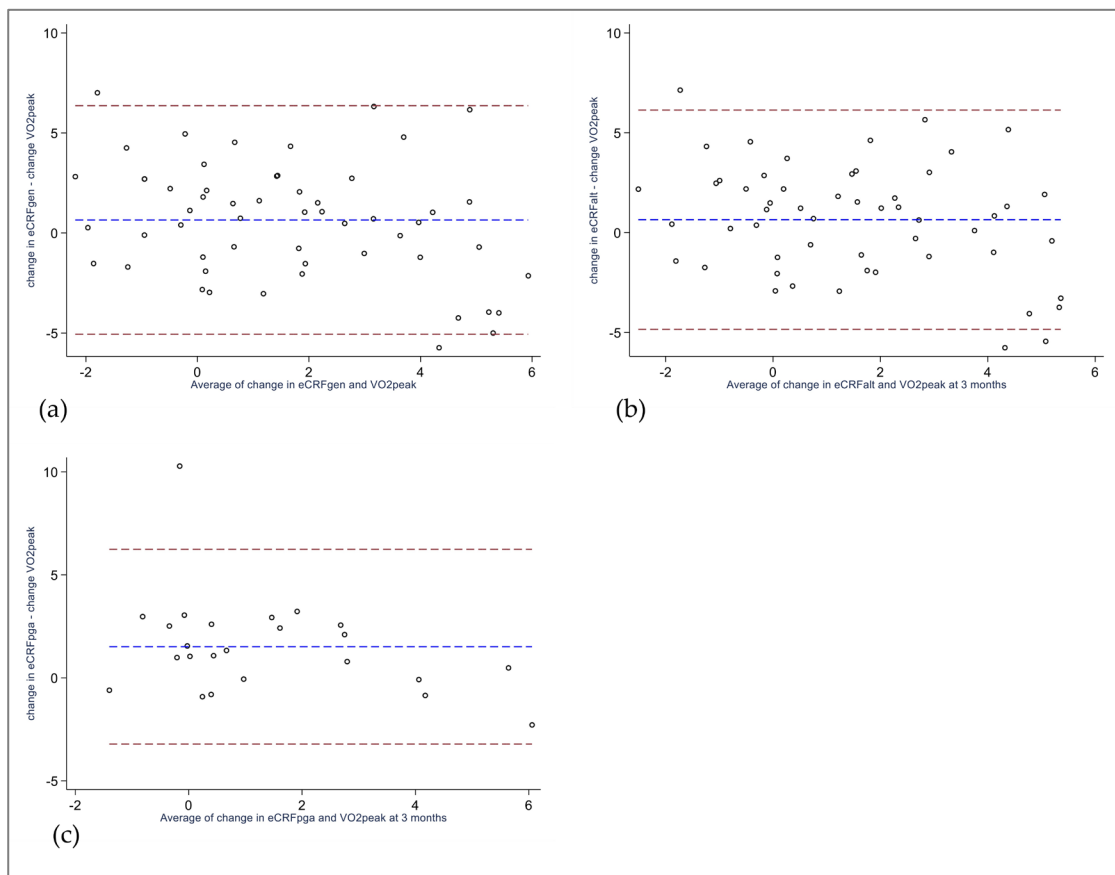
Figure 1. BlandAltman plots of eCRF versus VO$_{2peak}$ at **baseline** (**a**) eCRF$_{GEN}$ vs. VO$_{2peak}$, *n* = 55 (**b**) eCRF$_{ALT}$ vs. VO$_{2peak}$, *n* = 55 (**c**) eCRF$_{PGA}$ vs. VO$_{2peak}$, *n* = 24. Blue dotted line illustrates mean bias, red dotted lines illustrate 95% limits of agreement. eCRF$_{ALT:}$ Alternative rheumatoid arthritis-specific eCRF model. eCRF$_{GEN}$: Estimated cardiorespiratory fitness, generic model. eCRF$_{PGA}$: Rheumatoid arthritis-specific eCRF model, assessed in patients with rheumatoid arthritis. VO$_{2peak}$: Peak oxygen uptake.
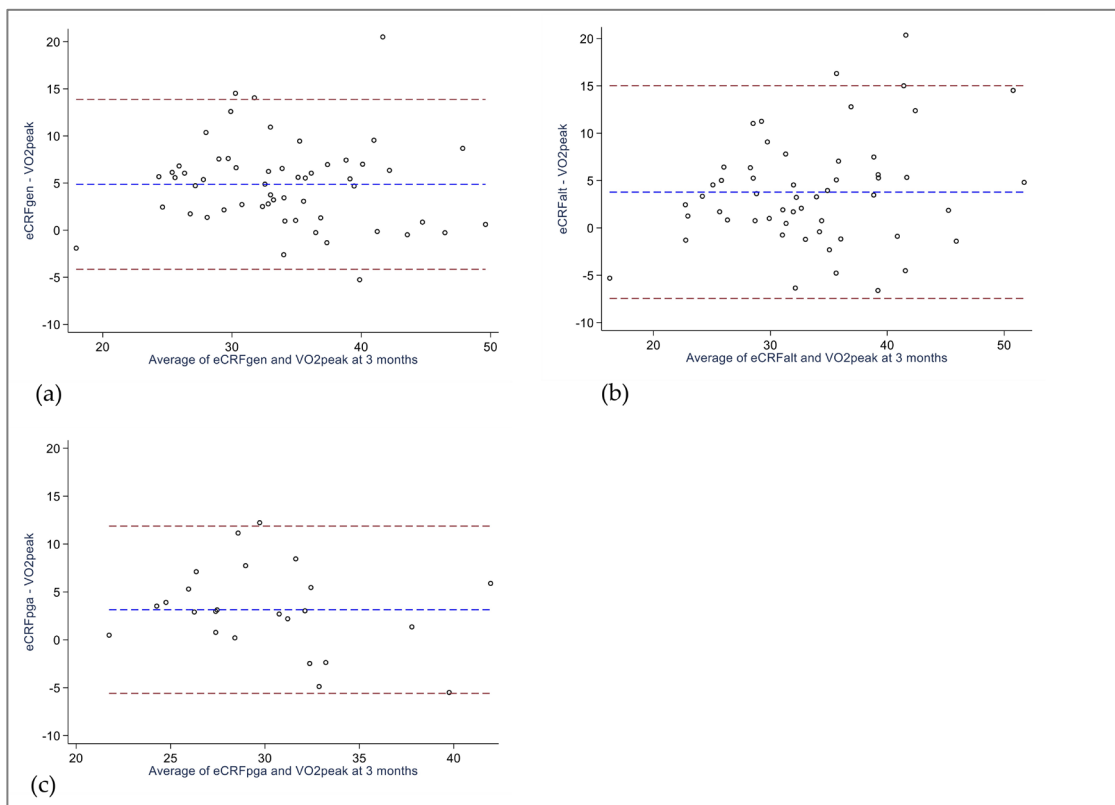
Figure 2. Bland–Altman plots of eCRF versus VO$_{2peak}$ at 3 months (**a**) eCRF$_{GEN}$ vs. VO$_{2peak}$, $n = 55$ (**b**) eCRF$_{ALT}$ vs. VO$_{2peak}$, $n = 54$ (**c**) eCRF$_{PGA}$ vs. VO$_{2peak}$, $n = 24$. Blue dotted line illustrates mean bias, red dotted lines illustrate 95% limits of agreement. eCRF$_{ALT:}$ Alternative rheumatoid arthritis-specific eCRF model. eCRF$_{GEN}$: Estimated cardiorespiratory fitness, generic model. eCRF$_{PGA}$: Rheumatoid arthritis-specific eCRF model, assessed in patients with rheumatoid arthritis. VO$_{2peak}$: Peak oxygen uptake.

Figure 3. Bland–Altman plots of change in eCRF versus change in VO_{2peak} from baseline to 3 months (**a**) $eCRF_{GEN}$ vs. VO_{2peak}, $n = 55$ (**b**) $eCRF_{ALT}$ vs. VO_{2peak}, $n = 54$ (**c**) $eCRF_{ALT}$ vs. VO_{2peak}, $n = 24$. Blue dotted line illustrates mean bias, red dotted lines illustrate 95% limits of agreement. $eCRF_{ALT}$: Alternative Rheumatoid Arthritis-specific eCRF model. $eCRF_{GEN}$: Estimated cardiorespiratory fitness, generic model. $eCRF_{PGA}$: Rheumatoid arthritis-specific eCRF model, assessed in patients with rheumatoid arthritis. VO_{2peak}: Peak oxygen uptake.

4. Discussion

Growing recognition of the value of CRF as a robust indicator of overall health has emphasized the need for valid and practical measures of CRF. The present study evaluated the validity and responsiveness of eCRF models in comparison to the criterion measure VO_{2peak} among patients with IJD. Our key findings revealed moderate correlations between eCRF models and VO_{2peak}, weak correlations between change scores, and large variability in agreement of individual measurements. Notably, all eCRF models demonstrated adequate performance in identifying larger improvements in VO_{2peak}.

In our data, we observed moderate agreement between eCRF and VO_{2peak}, indicating that while eCRF models may not precisely mirror VO_{2peak}, they can still capture a considerable correlation between these two measures of CRF among patients with IJD. These findings align with three other studies that have investigated the accuracy of several eCRF models in healthy adults and older individuals [5,22,42]. Collectively, this highlights a potential use for eCRF models to estimate VO_{2peak} at a group level. However, our Bland–Altman plots from baseline and 3 months into the study revealed wide limits of agreement that exceeded our pre-defined clinically acceptable difference. This observation parallels reports of an extensive range in the difference between eCRF and VO_{2peak} in other study samples [5,22,43]. Articles detailing the development of the eCRF models currently under investigation report a tendency to overestimate CRF, especially in individuals with VO_{2peak} below 30 mL/kg/min [21,23]. In line with this, our own study data with mean VO_{2peak} values of around 30–31 mL/kg/min demonstrated a consistent trend,

where both $eCRF_{GEN}$ and $eCRF_{ALT}$ overestimated VO_{2peak} at baseline and all eCRF models overestimated VO_{2peak} at 3 months. Collectively, this illustrates that eCRF models may not perform well in individual patients, and underlines the need for caution in applying eCRF to patients with presumably low fitness levels.

Assessing responsiveness lends insight into a measurement tool's capability to detect longitudinal change, and we believe the research design used herein allows us to draw inferences about the eCRF models' ability to capture change in VO_{2peak}. Our data demonstrated poor correlation between changes observed in the eCRF models and changes in VO_{2peak}. However, due to inherent restraints when examining change scores, lower correlation coefficients were anticipated. Beta coefficients derived from the regression analyses suggest that, depending on the specific eCRF model applied, a one-unit increase in eCRF corresponded to a change in VO_{2peak} ranging from 0.56 to 0.68 mL/kg/min. Along with the modest proportions of change in VO_{2peak} predicted by change in eCRF (R^2), this implies that eCRF models have weak ability to predict a true change in CRF. Although the mean differences between change captured by eCRF models and change in VO_{2peak} were, for the most part, small and statistically non-significant, the Bland–Altman limits of agreement extended beyond the range considered clinically acceptable. Few other studies have explored the longitudinal validity of eCRF models. Lannoy and Ross [44] examined the same $eCRF_{GEN}$ model used in our study within an RCT where participants were randomized to exercise at various intensity levels. Apart from participants randomized to high intensity exercise, no significant group differences were observed between changes in VO_{2peak} and eCRF at any timepoint. However, the authors of said study emphasized the presence of wide limits of agreement. Coupled with our results from the present study, this underscores that measures of change in eCRF and change in VO_{2peak} are not interchangeable at an individual level.

Epidemiological studies that examine a risk threshold of 3.5 mL/kg/min often rely on indirect measurement techniques such as eCRF and self-reported exercise. As discussed by Lannoy and Ross [44], within eCRF models, elements such as age and gender typically remain constant, while factors like BMI and resting heart rate might show small short-term variations. Accordingly, short-term changes in eCRF are largely reliant on shifts in self-reported physical activity behavior, suggesting that well-designed exercise interventions that improve CRF should be detectable using eCRF. However, several studies have demonstrated substantial variations in individual change in VO_{2peak} in response to a standardized exercise program, often attributed to differences in training dose, adherence to exercise, and genetic differences in exercise response [45]. eCRF models may struggle to capture individual response to exercise and will apply uniform change in CRF in response to change in the variables included in the model. In turn, physical activity may be prone to misclassifications, as the physical activity index in the current eCRF analyses relies heavily on self-reported intensity of exercise, and self-reported physical activity may be influenced by desirability bias [46,47]. Moreover, in the $eCRF_{PGA}$ model, disease activity (PGA) may correlate with changes in physical activity, considering that elevated disease activity associates with lower levels of physical activity [48]. In contrast to eCRF, CPET provides a precise measure of VO_{2peak}, with even a 1 mL/kg/min improvement linked to reduced risk of cardiovascular disease and mortality [3]. Although achieving the same level of accuracy from eCRF may be unrealistic, these models should be able to uncover larger changes in CRF. Our AUC estimates suggest that $eCRF_{GEN}$ and $eCRF_{ALT}$ models can correctly identify >80% of the patients with a $\geq$3.5 mL/kg/min improvement in VO_{2peak}. The $eCRF_{PGA}$ model performed even better, identifying 97% of the patients. Notably, the performance of the $eCRF_{PGA}$ model was tested in a small sample of patients, and caution is advised when interpreting these results. Furthermore, as only two patients in our cohort demonstrated $\geq$3.5 mL/kg/min deterioration in VO_{2peak}, the performance of eCRF models in capturing larger deteriorations among IJD patients remains unknown, and requires further investigation.

4.1. Clinical Implications and Future Research Avenues

Although eCRF has limitations in accuracy, the low cost and user-friendliness make it an appealing option. For instance, eCRF sidesteps the need for a maximal exercise test by using self-reported, easily accessible data, whereas VO_{2peak} assessed using CPET may exclude individuals with contraindications to maximal exercise tests or physical limitations [1], thereby resulting in the absence of CRF measures for these patients. Moreover, CRF obtained using eCRF is inversely related to cardiovascular disease and all-cause mortality [5,21,49], which adds valuable prognostic information. It may be important to distinguish between using eCRF in research and clinical contexts. For research that focuses on group trends, eCRF may be sufficient. In clinical care, eCRF may hold value as an initial screening tool to identify individual patients with low CRF that warrant more comprehensive tests such as CPET. However, clinicians need to be aware of the tendency for eCRF to overestimate CRF, especially in individuals with poor fitness levels [6,22,42], as this can have implications for risk interpretation and clinical assessment. Using eCRF in clinical practice as a surrogate for VO_{2peak} may therefore require a careful weighing of convenience versus limitations in accuracy. Given that CRF serves as a valuable measure of overall health, the potential benefits of eCRF may outweigh the inaccuracies, as having a rough CRF measure is better than none at all.

A path for future research could be to investigate whether adding eCRF models to cardiovascular disease risk assessment can improve the ability to predict cardiovascular disease and mortality in patients with IJD. Additionally, indirect tests that use variables such as workload and heart rate from maximal or submaximal exercise tests to estimate VO_{2peak} are considered second-best to CPET [1]. Assessing the validity of indirect exercise tests in modern-day IJD populations can offer insights into their potential as clinically feasible estimates of VO_{2peak}.

4.2. Strengths and Limitations

The primary strengths of the present study lie in our application of COSMIN recommendations to assess criterion validity and responsiveness, and the use of CPET to measure CRF. However, several limitations need to be addressed. Subtle biological variations and lack of consensus regarding the use of end criteria [29,30] illustrate that while CPET is recognized as the gold standard to measure CRF, the interpretation of test results is not an exact science. A treadmill was used in all CPETs in our sample, and our results may not extend to comparing eCRF models with VO_{2peak} from CPETs with a cycle ergometer. Our study cohort included participants from an RCT, and potential selection bias may limit the applicability of study results to the general IJD population. Caution is also advised in extrapolating results from the present study to different demographic groups and other eCRF models, as our cohort included individuals from Norway and the eCRF models under scrutiny were derived from Norwegian study samples. Lastly, the $eCRF_{PGA}$ model was evaluated in a small subset of patients, and results should be viewed as exploratory.

5. Conclusions

In summary, eCRF models demonstrated moderate criterion validity, suggesting a potential group-level applicability in patients with IJD. However, caution is advised in adopting eCRF for individual patients, due to the wide limits of agreement and a tendency to overestimate true CRF. eCRF models are less suited to measure change at the individual level, although these models have adequate ability to detect larger improvements in VO_{2peak} among patients with IJD. Balancing convenience with limitations in accuracy of individual measurements is crucial before applying eCRF models in research and clinical settings.

Table S1: Additional clinical characteristics from baseline sessions; Table S2: Cardiopulmonary Exercise Test characteristics from 3-month study visits.

References

1. Liguori, G.; American College of Sports Medicine. *ACSM's Guidelines for Exercise Testing and Prescription*; Lippincott Williams & Wilkins: Philadelphia, PA, USA, 2020.
2. Han, M.; Qie, R.; Shi, X.; Yang, Y.; Lu, J.; Hu, F.; Zhang, M.; Zhang, Z.; Hu, D.; Zhao, Y. Cardiorespiratory fitness and mortality from all causes, cardiovascular disease and cancer: Dose-response meta-analysis of cohort studies. *Br. J. Sports Med.* **2022**, *56*, 733–739. [CrossRef] [PubMed]
3. Imboden, M.T.; Harber, M.P.; Whaley, M.H.; Finch, W.H.; Bishop, D.L.; Fleenor, B.S.; Kaminsky, L.A. The Association between the Change in Directly Measured Cardiorespiratory Fitness across Time and Mortality Risk. *Prog. Cardiovasc. Dis.* **2019**, *62*, 157–162. [CrossRef] [PubMed]
4. Kodama, S.; Saito, K.; Tanaka, S.; Maki, M.; Yachi, Y.; Asumi, M.; Sugawara, A.; Totsuka, K.; Shimano, H.; Ohashi, Y.; et al. Cardiorespiratory fitness as a quantitative predictor of all-cause mortality and cardiovascular events in healthy men and women: A meta-analysis. *JAMA* **2009**, *301*, 2024–2035. [CrossRef] [PubMed]
5. Schumacher, B.T.; Di, C.; Bellettiere, J.; Lamonte, M.J.; Simonsick, E.M.; Parada, H., Jr.; Hooker, S.P.; Lacroix, A.Z. Validation, Recalibration, and Predictive Accuracy of Published VO 2max Prediction Equations for Adults Ages 50–96 Yr. *Med. Sci. Sports Exerc.* **2023**, *55*, 322–332. [CrossRef] [PubMed]
6. Ross, R.; Blair, S.N.; Arena, R.; Church, T.S.; Despres, J.P.; Franklin, B.A.; Haskell, W.L.; Kaminsky, L.A.; Levine, B.D.; Lavie, C.J.; 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. [CrossRef] [PubMed]
7. Sieper, J.; Poddubnyy, D. Axial spondyloarthritis. *Lancet* **2017**, *390*, 73–84. [CrossRef] [PubMed]
8. Sparks, J.A. Rheumatoid Arthritis. *Ann. Intern. Med.* **2019**, *170*, Itc1–Itc16. [CrossRef]
9. Ritchlin, C.T.; Colbert, R.A.; Gladman, D.D. Psoriatic Arthritis. *N. Engl. J. Med.* **2017**, *376*, 957–970. [CrossRef]
10. Conrad, N.; McInnes, I.B.; McMurray, J.J.V.; Sattar, N. Patients with a range of rheumatic diseases are at increased risk of cardiovascular disorders towards a re-evaluation of the European League against Rheumatism (EULAR)'s recommendations for cardiovascular risk management? *Ann. Rheum. Dis.* **2022**, *82*, 457–459. [CrossRef]

11. Agca, R.; Heslinga, S.C.; Rollefstad, S.; Heslinga, M.; McInnes, I.B.; Peters, M.J.; Kvien, T.K.; Dougados, M.; Radner, H.; Atzeni, F.; et al. EULAR recommendations for cardiovascular disease risk management in patients with rheumatoid arthritis and other forms of inflammatory joint disorders: 2015/2016 update. *Ann. Rheum. Dis.* **2017**, *76*, 17–28. [CrossRef]

12. Visseren, F.L.J.; Mach, F.; Smulders, Y.M.; Carballo, D.; Koskinas, K.C.; Bäck, M.; Benetos, A.; Biffi, A.; Boavida, J.M.; Capodanno, D.; et al. 2021 ESC Guidelines on cardiovascular disease prevention in clinical practice. *Eur. Heart J.* **2021**, *42*, 3227–3337. [CrossRef] [PubMed]

13. Metsios, G.S.; Koutedakis, Y.; Veldhuijzen van Zanten, J.J.; Stavropoulos-Kalinoglou, A.; Vitalis, P.; Duda, J.L.; Ntoumanis, N.; Rouse, P.C.; Kitas, G.D. Cardiorespiratory fitness levels and their association with cardiovascular profile in patients with rheumatoid arthritis: A cross-sectional study. *Rheumatology* **2015**, *54*, 2215–2220. [CrossRef] [PubMed]

14. Munsterman, T.; Takken, T.; Wittink, H. Are persons with rheumatoid arthritis deconditioned? A review of physical activity and aerobic capacity. *BMC Musculoskelet. Disord.* **2012**, *13*, 202. [CrossRef] [PubMed]

15. O'Dwyer, T.; O'Shea, F.; Wilson, F. Decreased physical activity and cardiorespiratory fitness in adults with ankylosing spondylitis: A cross-sectional controlled study. *Rheumatol. Int.* **2015**, *35*, 1863–1872. [CrossRef] [PubMed]

16. Halvorsen, S.; Vollestad, N.K.; Fongen, C.; Provan, S.A.; Semb, A.G.; Hagen, K.B.; Dagfinrud, H. Physical fitness in patients with ankylosing spondylitis: Comparison with population controls. *Phys. Ther.* **2012**, *92*, 298–309. [CrossRef] [PubMed]

17. Nordén, K.R.; Semb, A.G.; Dagfinrud, H.; Hisdal, J.; Ødegård, S.; Sexton, J.; Fongen, C.; Skandsen, J.; Blanck, T.; Metsios, G.S.; et al. Associations between cardiovascular risk factors, disease activity and cardiorespiratory fitness in patients with inflammatory joint disease: A cross-sectional analysis. *BMC Sports Sci. Med. Rehabil.* **2023**, *15*, 63. [CrossRef]

18. Liff, M.H.; Hoff, M.; Wisloff, U.; Videm, V. Reduced cardiorespiratory fitness is a mediator of excess all-cause mortality in rheumatoid arthritis: The Trøndelag Health Study. *RMD Open* **2021**, *7*, e001545. [CrossRef]

19. Videm, V.; Liff, M.H.; Hoff, M. Relative importance of inflammation and cardiorespiratory fitness for all-cause mortality risk in persons with rheumatoid arthritis: The population-based Trøndelag Health Study. *RMD Open* **2023**, *9*, e003194. [CrossRef]

20. Nes, B.M.; Janszky, I.; Vatten, L.J.; Nilsen, T.I.; Aspenes, S.T.; Wisloff, U. Estimating V.O 2peak from a nonexercise prediction model: The HUNT Study, Norway. *Med. Sci. Sports Exerc.* **2011**, *43*, 2024–2030. [CrossRef]

21. Nes, B.M.; Vatten, L.J.; Nauman, J.; Janszky, I.; Wisloff, U. A simple nonexercise model of cardiorespiratory fitness predicts long-term mortality. *Med. Sci. Sports Exerc.* **2014**, *46*, 1159–1165. [CrossRef]

22. Peterman, J.E.; Whaley, M.H.; Harber, M.P.; Fleenor, B.S.; Imboden, M.T.; Myers, J.; Arena, R.; Kaminsky, L.A. Comparison of non-exercise cardiorespiratory fitness prediction equations in apparently healthy adults. *Eur. J. Prev. Cardiol.* **2021**, *28*, 142–148. [CrossRef] [PubMed]

23. Liff, M.H.; Hoff, M.; Fremo, T.; Wisløff, U.; Videm, V. An Estimation Model for Cardiorespiratory Fitness in Adults with Rheumatoid Arthritis. *Med. Sci. Sports Exerc.* **2020**, *52*, 1248–1255. [CrossRef] [PubMed]

24. Mokkink, L.; Terwee, C.; de Vet, H. Key concepts in clinical epidemiology: Responsiveness, the longitudinal aspect of validity. *J. Clin. Epidemiol.* **2021**, *140*, 159–162. [CrossRef] [PubMed]

25. Nordén, K.R.; Dagfinrud, H.; Semb, A.G.; Hisdal, J.; Viktil, K.K.; Sexton, J.; Fongen, C.; Skandsen, J.; Blanck, T.; Metsios, G.S.; et al. Effect of high-intensity exercise on cardiorespiratory fitness, cardiovascular disease risk and disease activity in patients with inflammatory joint disease: Protocol for the ExeHeart randomised controlled trial. *BMJ Open* **2022**, *12*, e058634. [CrossRef]

26. Mokkink, L.B.; Prinsen, C.; Patrick, D.L.; Alonso, J.; Bouter, L.M.; De Vet, H.; Terwee, C.B. COSMIN Study Design Checklist for Patient-Reported Outcome Measurement Instruments. Available online: http://www.comsin.nl (accessed on 18 August 2023).

27. Kottner, J.; Audigé, L.; Brorson, S.; Donner, A.; Gajewski, B.J.; Hróbjartsson, A.; Roberts, C.; Shoukri, M.; Streiner, D.L. Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed. *J. Clin. Epidemiol.* **2011**, *64*, 96–106. [CrossRef] [PubMed]

28. Gagnier, J.J.; Lai, J.; Mokkink, L.B.; Terwee, C.B. COSMIN reporting guideline for studies on measurement properties of patient-reported outcome measures. *Qual. Life Res. Int. J. Qual. Life Asp. Treat. Care Rehabil.* **2021**, *30*, 2197–2218. [CrossRef] [PubMed]

29. American Thoracic, S.; American College of Chest, P. ATS/ACCP Statement on cardiopulmonary exercise testing. *Am. J. Respir. Crit. Care Med.* **2003**, *167*, 211–277. [CrossRef]

30. Pritchard, A.; Burns, P.; Correia, J.; Jamieson, P.; Moxon, P.; Purvis, J.; Thomas, M.; Tighe, H.; Sylvester, K.P. ARTP statement on cardiopulmonary exercise testing 2021. *BMJ Open Respir. Res.* **2021**, *8*, e001121. [CrossRef]

31. Balke, B.; Ware, R.W. An experimental study of physical fitness of Air Force personnel. *U. S. Armed Forces Med. J.* **1959**, *10*, 675–688.

32. Edvardsen, E.; Hem, E.; Anderssen, S.A. End criteria for reaching maximal oxygen uptake must be strict and adjusted to sex and age: A cross-sectional study. *PLoS ONE* **2014**, *9*, e85276. [CrossRef]

33. Kurtze, N.; Rangul, V.; Hustvedt, B.E.; Flanders, W.D. Reliability and validity of self-reported physical activity in the Nord-Trondelag Health Study: HUNT 1. *Scand. J. Public Health* **2008**, *36*, 52–61. [CrossRef]

34. Liljequist, D.; Elfving, B.; Skavberg Roaldsen, K. Intraclass correlation—A discussion and demonstration of basic features. *PLoS ONE* **2019**, *14*, e0219854. [CrossRef]

35. Husted, J.A.; Cook, R.J.; Farewell, V.T.; Gladman, D.D. Methods for assessing responsiveness: A critical review and recommendations. *J. Clin. Epidemiol.* **2000**, *53*, 459–468. [CrossRef] [PubMed]

36. Bland, J.M.; Altman, D.G. Correlation in restricted ranges of data. *BMJ* **2011**, *342*, d556. [CrossRef] [PubMed]

37. Bland, J.M.; Altman, D.G. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* **1986**, *1*, 307–310. [PubMed]

38. Giavarina, D. Understanding Bland Altman analysis. *Biochem. Med.* **2015**, *25*, 141–151. [CrossRef] [PubMed]

39. Laukkanen, J.A.; Isiozor, N.M.; Kunutsor, S.K. Objectively Assessed Cardiorespiratory Fitness and All-Cause Mortality Risk: An Updated Meta-analysis of 37 Cohort Studies Involving 2,258,029 Participants. *Mayo Clin. Proc.* **2022**, *97*, 1054–1073. [CrossRef] [PubMed]

40. Letnes, J.M.; Dalen, H.; Vesterbekkmo, E.K.; Wisloff, U.; Nes, B.M. Peak oxygen uptake and incident coronary heart disease in a healthy population: The HUNT Fitness Study. *Eur. Heart J.* **2019**, *40*, 1633–1639. [CrossRef]

41. Nahm, F.S. Receiver operating characteristic curve: Overview and practical use for clinicians. *Korean J. Anesth.* **2022**, *75*, 25–36. [CrossRef]

42. McAuley, E.; Szabo, A.N.; Mailey, E.L.; Erickson, K.I.; Voss, M.; White, S.M.; Wójcicki, T.R.; Gothe, N.; Olson, E.A.; Mullen, S.P.; et al. Non-Exercise Estimated Cardiorespiratory Fitness: Associations with Brain Structure, Cognition, and Memory Complaints in Older Adults. *Ment. Health Phys. Act.* **2011**, *4*, 5–11. [CrossRef]

43. Mailey, E.L.; White, S.M.; Wójcicki, T.R.; Szabo, A.N.; Kramer, A.F.; McAuley, E. Construct validation of a non-exercise measure of cardiorespiratory fitness in older adults. *BMC Public Health* **2010**, *10*, 59. [CrossRef] [PubMed]

44. Lannoy, L.; Ross, R. Nonexercise Equations for Determining Change in Cardiorespiratory Fitness. *Med. Sci. Sports Exerc.* **2020**, *52*, 1525–1531. [CrossRef] [PubMed]

45. Meyler, S.; Bottoms, L.; Muniz-Pumares, D. Biological and methodological factors affecting VO2max response variability to endurance training and the influence of exercise intensity prescription. *Exp. Physiol.* **2021**, *106*, 1410–1424. [CrossRef] [PubMed]

46. Houle, S.A.; Sui, X.; Blair, S.N.; Ross, R. Association Between Change in Nonexercise Estimated Cardiorespiratory Fitness and Mortality in Men. *Mayo Clin. Proc. Innov. Qual. Outcomes* **2022**, *6*, 106–113. [CrossRef] [PubMed]

47. Helmerhorst, H.J.; Brage, S.; Warren, J.; Besson, H.; Ekelund, U. A systematic review of reliability and objective criterion-related validity of physical activity questionnaires. *Int. J. Behav. Nutr. Phys. Act.* **2012**, *9*, 103. [CrossRef] [PubMed]

48. Bremander, A.; Malm, K.; Andersson, M.L. Physical activity in established rheumatoid arthritis and variables associated with maintenance of physical activity over a seven-year period—A longitudinal observational study. *BMC Rheumatol.* **2020**, *4*, 53. [CrossRef] [PubMed]

49. Lee, J.; Song, R.J.; Musa Yola, I.; Shrout, T.A.; Mitchell, G.F.; Vasan, R.S.; Xanthakis, V. Association of Estimated Cardiorespiratory Fitness in Midlife with Cardiometabolic Outcomes and Mortality. *JAMA Netw. Open* **2021**, *4*, e2131284. [CrossRef]