Received: 29 June 2023 | Accepted: 10 January 2024

DOI: 10.1111/ene.16223

european journal of neurology
the official journal of the european academy of neurology

ORIGINAL ARTICLE

Establishing how much improvement in lung function and distance walked is clinically important for adult patients with Pompe disease

Aglina Lika[1,2,3] | Eleni-Rosalina Andrinopoulou[2,3] | Nadine A. M. E. van der Beek[4] | Dimitris Rizopoulos[2,3] | Ans T. van der Ploeg[1] | Michelle E. Kruijshaar[1]

[1]Department of Pediatrics, Centre for Lysosomal and Metabolic Diseases, Erasmus MC University Medical Centre, Rotterdam, The Netherlands

[2]Department of Biostatistics, Erasmus MC University Medical Centre, Rotterdam, The Netherlands

[3]Department of Epidemiology, Erasmus MC University Medical Centre, Rotterdam, The Netherlands

[4]Department of Neurology, Centre for Lysosomal and Metabolic Diseases, Erasmus MC University Medical Centre, Rotterdam, The Netherlands

Correspondence
Michelle E. Kruijshaar, Department of Pediatrics, Centre for Lysosomal and Metabolic Diseases, Erasmus MC University Medical Centre, Dr Molewaterplein 40 3015 GD Rotterdam, The Netherlands.
Email: m.kruijshaar@erasmusmc.nl

Funding information
Sanofi Genzyme; ZonMw, Grant/Award Number: 05-09-2007, 152001005 and 80-83600-98-13007; Prinses Beatrix Spierfonds, Grant/Award Number: OP07-08, W.OR13-21 and W.OR15-10; TKI - Health Holland, Grant/Award Number: LSHM16008; SSWO - Sophia Children's Hospital Foundation, Grant/Award Number: 687

Abstract

Background and purpose: Pompe disease is a rare, inheritable, progressive metabolic myopathy. This study aimed to estimate the minimal clinically important difference (MCID) for an improvement in forced vital capacity in the upright seated position (FVC_{up}) and the 6-min walk test (6MWT) after a year of treatment with enzyme replacement therapy.

Methods: Data were obtained from two prospective follow-up studies. Between-group and within-group MCIDs were estimated using anchor-based methods. Additionally, a distribution-based method was used to generate supportive evidence. As anchors, self-reported change in health and in physical functioning, shortness of breath and a categorization of the Short-Form 36 Physical Component Summary score were used. Anchor appropriateness was assessed using Spearman correlations (absolute values ≥0.29) and a sufficient number of observations in each category.

Results: In all, 102 patients had at least one FVC_{up} or 6MWT measurement during enzyme replacement therapy. Based on the anchors assessed as appropriate, the between-group MCID for an improvement in FVC_{up} ranged from 2.47% to 4.83% points. For the 6MWT, it ranged from 0.35% to 7.47% points which is equivalent to a distance of 2.18–46.61 m and 1.97–42.13 m for, respectively, a man and a woman of age 50, height 1.75 m and weight 80 kg. The results of the distribution-based method were within these ranges when applied to change in the outcome values.

Conclusion: The MCIDs for FVC_{up} and 6MWT derived in this study can be used to interpret differences between and within groups of patients with Pompe disease in clinical trials and cohort studies.

KEYWORDS
anchor-based method, clinically meaningful threshold, forced vital capacity, minimal clinically important difference, minimal important change, Pompe disease, 6-minute walk test

INTRODUCTION

Pompe disease is a rare, inheritable and progressive metabolic myopathy. It is caused by partial or total deficiency of the lysosomal enzyme acid alpha-glucosidase, resulting in a build-up of lysosomal glycogen and subsequent cellular damage in virtually all body tissues, particularly in muscles [1]. Adult patients present with progressive muscle weakness, limitations in motor function and respiratory difficulties. Enzyme replacement therapy (ERT) with recombinant human alpha-glucosidase (alglucosidase alfa, Myozyme) has been available for treatment since 2006 and has been shown to improve or stabilize these outcomes [2-9], followed in the long term by a secondary decline [4,10,11,12]. New treatments are being investigated to improve the situation of these patients further with some improved forms of ERT recently having been approved [13-15].

The primary outcomes measured in clinical trials for treating adult patients with Pompe disease are usually forced vital capacity in the upright seated position (FVC_{up}) and the 6-min walk test (6MWT) [7,13,14]. However, the results of these trials can be limited since there is no threshold to determine how much a treatment group should improve on these end-points to be clinically important. A statistically significant difference does not also imply clinical importance for the patient. To guide clinical decision-making and avoid relying solely on statistical significance, the so-called minimal clinically important difference (MCID) has been proposed. The MCID is the smallest change in an outcome that a patient perceives as important, either beneficial or harmful, and that would lead the patient or clinician to consider a change in management [16]. At present, no published MCID values are available for Pompe disease, hampering the interpretation of trials.

The MCID can be determined by several methods, including two common ones: anchor-based methods and distribution-based methods [17-19]. Anchor-based methods link the outcomes of interest to an external criterion, either a factor that has clinical relevance (e.g., ventilator use) or patient-reported ratings of (changes in) health: the so-called anchor [17-19]. Distribution-based methods compare the change in the outcome of interest to some measure of statistical variability, such as the standard deviation (SD), and are typically used as supportive evidence since this method provides indirect information about the MCID [17-19]. The present study aimed to estimate the MCID for an improvement in FVC_{up} and 6MWT for interpreting changes between and within groups of patients with Pompe disease during their first year of treatment with ERT. Anchor-based methods were used and our findings were supported using distribution-based methods. For this study, data from two cohort studies were linked, one providing the clinical outcomes of interest and the other the anchors, resulting in a dataset of 102 adult Dutch patients who received ERT.

METHODS

Data

Data from two prospective follow-up studies were combined [10,20,21]. Both studies were conducted at the Centre for Lysosomal and Metabolic Diseases, Erasmus MC University Medical Centre, Rotterdam, the national referral centre for Pompe disease in the Netherlands.

The databases were locked at the end of December 2018 and included only adult patients with Pompe disease residing in the Netherlands who were followed in both longitudinal studies, received ERT and had at least one outcome measurement during ERT. In addition, data from up to 6 months before the start of ERT were included to ensure that there were sufficient 'baseline' data points.

Outcomes of interest

Forced vital capacity in the upright seated position (FVC_{up}) and the 6MWT were assessed every 3–12 months before and after the start of ERT since January 2005 as part of an ongoing clinical follow-up study [21,22]. FVC_{up} was measured using spirometry and the results were expressed as the percentage of the predicted normal values based on the subject's age, sex, race and height [23,24]. The 6MWT was used as a test of functional endurance in which the distance walked in 6 min was recorded [25]. The values were presented as a percentage of the predicted normal values to account for the effects of age, height, weight and sex [25,26].

Anchors and anchor groups

As anchors, patient-reported outcome measures were used, collected annually through an ongoing international questionnaire: the International Pompe Association (IPA)/Erasmus MC Pompe survey, and also partly alongside the clinical follow-up [20]. Questionnaires include the Medical Outcome Study 36-item Short-Form Health Survey (SF-36, version 1 before 2009 and version 2 after 2009) [27,28], and a Pompe-specific questionnaire specifically designed to assess symptoms and problems of the disease [20]. From these, the following items were included as potential anchors: GH_change, PF_change, SB_diff and PCS_cat. These are described in Table 1. The Pompe-specific questionnaire also includes questions on ventilation and wheelchair use. However, there was too little variation reported in the survey so these questions could not be used as anchors.

Using each anchor, patients were categorized into three groups ('better', 'same' and 'worse'), as presented in Table 1. For this, the categories 'much better' and 'a bit better' (and 'much'/'a bit worse')

TABLE 1 Anchors used and anchor groups.

Anchor	Description and information	Answer options	Anchor group
GH_change	Self-reported change in health. This is item 2 from SF-36: 'how has your health in general changed compared to 1 year ago?' This item is frequently used as anchor	FU1: much better	Better
		FU1: a bit better	
		FU1: the same	Same
		FU1: a bit worse	Worse
		FU1: much worse	
PF_change	Self-reported change in physical functioning. This is from an item in the Pompe-specific questionnaire: 'how has your physical functioning changed compared to 1 year ago'	FU1: much better	Better
		FU1: a bit better	
		FU1: the same	Same
		FU1: a bit worse	Worse
		FU1: much worse	
SB_diff • SB_diff_ir • SB_diff_sup • SB_diff_he • SB_diff_me	Observed difference in the answer to shortness of breath question (yes/no) from BL to FU1: • in rest • in supine position • in heavy exercise • in mild exercise *Note*: This was defined as anchor only for FVC_{up} (%)	BL: Yes → FU1: No	Better
		BL: Yes → FU1: Yes	Same
		BL: No → FU1: No	Same
		BL: No → FU1: Yes	Worse
PCS_cat	Categorization of the PCS score[a] of the SF-36 questionnaire based on the change from BL to FU1 *Note*: This is not a steadfast anchor since the PCS does not have a known MCID for the Pompe disease	>5	Better
		≤5 and ≥−5	Same
		<−5	Worse

Abbreviations: BL, baseline, i.e. the time-point closest to start of ERT in the time interval [−0.5, 0.49]; ERT, enzyme replacement therapy; FU1, 1-year follow-up, i.e. the time-point closest to 1 year in the time interval [0.5, 1.7]; FVC_{up}, forced vital capacity in the upright position; GH_change, self-reported change in health (question 2 of the SF-36 questionnaire); PCS_cat, a categorization of the change in Physical Component Summary (PCS) score; PF_change, self-reported change in physical functioning (question in the Pompe-specific questionnaire); SB_diff_ir, SB_diff_sup, SB_diff_he and SB_diff_me, change in shortness of breath in rest, in supine position, in heavy exercise and in mild exercise, respectively, from baseline to follow-up; SF-36, Medical Outcome Study 36-item Short-Form Health Survey [28].

[a]PCS norm scores were calculated using the Dutch 1998 norms, ensuring comparability of the results for both versions of the SF-36. Norm-based scores range from 0 to 100, with higher values indicating better quality of life.

of GH_change and PF_change were combined a priori because of the small number of patients. Combining two categories into one 'better' group could result in overestimating the 'minimal' clinically important difference for an improvement. Patients who deteriorated were not included in the calculation of the MCID for an improvement.

Statistical analysis

Time intervals

Given the observational nature of the data, measurements were not taken exactly at baseline (BL) and 1-year follow-up (FU1). To solve this issue, time points were selected using time intervals. In addition, for some anchors, modelling was used (see below and the Appendix, Part A). Time points were selected as follows:

- BL: the measurement in the time interval [−0.5, 0.49] that is closest to the start of ERT ($t=0$);
- FU1: the measurement in the time interval [0.5, 1.7] that is closest to $t=1$.

Time is expressed in years since the start of ERT. A slightly longer time is allowed on the right side to capture more patients.

The first year of treatment was chosen as the time frame for this study, as this pertains to the period in which clinical trials are usually performed. Also, it is the time frame in which the most improvement is seen [4-9].

Anchor-based MCID using time intervals

The between-group MCID was calculated from the data observed in the above time intervals using the mean change method [29,30]. Thus, it was calculated as the mean difference in the clinical outcome of interest (6MWT/FVC_{up}) from BL to FU1 in the anchor group 'better' minus the mean difference in the outcome of interest in the anchor group 'same':

$$MCID_{between} = \text{mean}\left[\left(outcome_{FU1} - outcome_{BL}\right)_{better}\right] - \text{mean}\left[\left(outcome_{FU1} - outcome_{BL}\right)_{same}\right]$$

The within-group MCID was the mean difference (from BL to FU1) of the outcome in the group 'better', as proposed by Jaeschke et al. [16]. These values are available in the results but are not emphasized further in the paper.

$$\text{MCID}_{\text{within}} = \text{mean}\left[\left(\text{outcome}_{\text{FU1}} - \text{outcome}_{\text{BL}}\right)_{\text{better}}\right]$$

To assess the appropriateness of the anchors, the Spearman correlation (r) was calculated between the patients' change (from BL to FU1) in the outcomes and the anchor groups ('better' and 'same'). Correlations ≥0.3 are recommended [31]. Nevertheless, given the rare nature of Pompe disease, the sample size limits the correlations and absolute correlations ≥0.29 were assumed to be sufficient. In addition, a threshold was set for the total number of patients available for the analysis (i.e., anchor groups 'better' plus 'same') of ≥35 patients and ≥eight patients per anchor group. The literature suggests to have at least 50 or 100 patients in total in these groups [32] but this was not possible for this rare disease.

Anchor-based MCID using a statistical model

Between-group MCIDs for GH_change and PF_change were also estimated using modelling, to remove the effect of variation in the timing of outcome and anchor measurements. For the remaining anchors modelling was too complex as these are based on the difference between two time points. Modelling was initially proposed by Angst et al. [19] to adjust for covariates that may not be equally distributed between anchor groups.

Briefly, a logistic linear mixed-effects model was fitted to estimate the values of the anchors at $t=1$ considering sex, disease duration at the start of ERT, time (nonlinear) and ERT. Next, a linear mixed-effects model was fitted for the FVC_{up}/6MWT to estimate its value at $t=0$ and $t=1$ for the categories 'better' and 'same' considering time (nonlinear), the imputed anchor value and their interaction. Estimates from 200 repetitions of the above process were pooled to provide the MCID estimates. Further explanation is provided in the Appendix, Part A.

Distribution-based MCID

Distribution-based methods do not consider the patient's perspective [33] and were used as supportive evidence. The MCID was calculated as half the standard deviation of the observed values in the BL interval: $\text{MCID} = 0.5 \times \text{SD}_{\text{BL}}$. This follows the method proposed by Norman et al. [34], who attributed their finding to the fact that 0.5 × SD represents the human mental discriminative capacity limit, a limit that would appear in most patient-reported outcome measures. This method was also applied to the change in outcome values from BL to FU1, as the purpose of the MCID is to provide insight into how much change is relevant.

Triangulation

Previous literature [35-37] suggests presenting a range of estimates for triangulation. The between-groups MCIDs estimated from all methods are therefore presented in a graph.

Analyses were performed with the R statistical software (version 4.2.0).

Standard protocol approvals, registrations and patient consents

Both studies were approved by the ethics committee of the Erasmus MC University Medical Centre and have been performed in accordance with the 1964 Declaration of Helsinki and its later amendments. Written informed consent was obtained from all participants prior to their inclusion. Consent for publication is not applicable.

RESULTS

Study population

There were 140 Dutch adult patients in the IPA/Erasmus MC Pompe survey until December 2018, 146 in the clinical follow-up study and 130 patients in both studies. After excluding patients who did not receive ERT, had no measurements during ERT and/or no FVC_{up} and 6MWT measurements, 102 patients were eligible for this study. The anchors GH_change, PF_change, SB_diff and PCS_cat were available for 88, 73, 66 and 83 patients, respectively (Appendix, Part B, Figure A1). The change in FVC_{up} and 6MWT from BL to FU1 could be calculated for 98 and 53 patients, respectively.

Table 2 shows the demographic and clinical characteristics of the study population and the time at which the measurements took place. Of all 102 Pompe patients, 54.9% were women, the median age at the start of ERT was 50 years and the median age at the start of symptoms 32.5 years, 30.4% were wheelchair dependent and 24.5% were ventilator dependent.

The median times at which the BL and FU1 measurements were taken were close to 0 and 1, respectively, both for the anchor and clinical outcomes. The clinical outcomes and the anchors were not measured at the same time. The median time difference was close to zero for all outcome–anchor combinations (Table 2). For some combinations the full range included 9 months but, for all the combinations, the interquartile range showed that 50% of the paired measurements were taken no more than 4 months apart. Further information is provided in the Appendix, Part C.

Anchor-based approach based on time intervals

Table 3 shows the results of the anchor-based approach using time intervals. For each anchor the number of patients, the mean

TABLE 2 Characteristics of the patients of interest ($n = 102$).

Demographic and clinical characteristics	Patients of interest ($n = 102$)
Women: number (%)	56 (54.9)
Age at start of symptoms in years: median (full range)	32.50 (2–62)
Age at start of ERT: median (full range)	50 (14–76)
Disease duration at start of ERT in years: median (full range)	13.76 (0.85–50.28)
Wheelchair dependent at start of ERT: number (%)	31 (30.4)
Full: number (%)	8 (7.8)
Partial: number (%)	23 (22.6)
Respiratory support at start of ERT: number (%)	25 (24.5)
Invasive: number (%)	2 (2.0)
Non-invasive: number (%)	23 (22.6)
Most frequent allele 1: number (%)	
c-32-13T>G(IVS1-13T>G)	97 (95.1)
Most frequent allele 2: number (%)	
c.525del	46 (45.1)

	Time of measurement in years: median (range)	
Measurement time since start ERT (*n*)	**BL**	**FU1**
FVC_{up} (%)	−0.04 (−0.41 to 0.31)	1.04 (0.50 to 1.54)
6MWT (%)	−0.03 (−0.29 to 0.38)	1.02 (0.50 to 1.35)
GH_change	–	1.04 (0.52 to 1.7)
PF_change	–	1.05 (0.50 to 1.66)
SB_diff	−0.11 (−0.50 to 0.49)	1.04 (0.50 to 1.66)
PCS_cat	−0.02 (−0.35 to 0.47)	1.04 (0.52 to 1.7)
Difference in time of outcome and anchor (*n*)		
FVC_{up} (%) – GH_change	–	0 (−0.66 to 0.52)
FVC_{up} (%) – PF_change	–	−0.04 (−0.85 to 0.61)
FVC_{up} (%) – SB_diff	0 (−0.67 to 0.54)	0.02 (−0.61 to 0.85)
FVC_{up} (%) – PCS_cat	0 (−0.54 to 0.52)	0 (−0.66 to 0.52)
6MWT (%) – GH_change	–	0.01 (−0.85 to 0.52)
6MWT (%) – PF_change	–	−0.16 (−0.47 to 0.52)
6MWT (%) – PCS_cat	0 (−0.37 to 0.63)	0.01 (−0.85 to 0.52)

Note: Both FVCup and 6MWT are expressed as a percentage of their predicted normal values.

Abbreviations: 6MWT, 6-min walk test; BL, baseline, i.e. time-point closest to start of ERT in the time interval [−0.5, 0.49]; ERT, enzyme replacement therapy; FU1, 1-year follow-up, i.e. time-point closest to 1 year in the time interval [0.5, 1.7]; FVC_{up}, forced vital capacity in the upright position; GH_change, self-reported change in health (question 2 of the SF-36 questionnaire); *n*, number of patients; PCS_cat, a categorization of the change in Physical Component Summary (PCS) score; PF_change, self-reported change in physical functioning (question in the Pompe-specific questionnaire); SB_diff, the change in shortness of breath in the first year of treatment with ERT.

change in the outcomes and the Spearman correlation (*r*) are presented. These are provided for all answer options and for the anchor groups that were used to calculate the MCID, that is, 'better' and 'same'.

The results presented in bold had anchors that met the appropriateness criteria ($|r| \geq 0.29$; $n \geq 35$ in both anchor groups and $n \geq 8$ per anchor group). For FVC_{up}, the anchors shortness of breath in the supine position (SB_diff_sup) and PCS_cat were 'appropriate' for the analysis, resulting in an MCID for between-group differences of 2.47% points and 4.83% points, respectively. The within-group MCID ranged from 1.26% points to 3.74% points.

For the 6MWT only the anchor GH_change was 'appropriate'. The MCID for an improvement in 6MWT between groups based on GH_change is 7.47% points which is equivalent to a distance of 46.61 m and 42.13 m for a man and a woman of age 50, height 1.75 m and weight 80 kg, respectively. For a man and a woman of age 70, height 1.70 m and weight 90 kg the distances are equivalent to 34.97 m and 31 m, respectively. The within-group estimate was 11.53% points.

TABLE 3 Anchor-based MCIDs using time intervals: number of patients, mean change and Spearman correlation for FVC_{up} and 6MWT by anchor.

	FVC_{up} (%)			6MWT (%)		
Anchor	Patients	Mean change (SE)	Spearman correlation[a]	Patients	Mean change (SE)	Spearman correlation[a]
GH_change	MCID = 1.56			MCID = **7.47**		
Better	32	1.27 (1.4)	−0.10	12	**11.53 (2.5)**	**−0.42**
Much better	10	5.49 (2.73)	−0.27	4	16.93 (6.49)	−0.37
A bit better	22	−0.65 (1.57)	0.01	8	8.83 (1.63)	−0.34
The same	40	−0.29 (0.86)	1	25	4.06 (1.63)	1
Worse	13	−3.92 (2.07)	–	6	5.26 (5.17)	–
A bit worse	12	−3.65 (2.23)	–	5	5.43 (6.33)	–
Much worse	1	−7.20 (NA)	–	1	4.46 (NA)	–
PF_change	MCID = 2.05			MCID = 3.29		
Better	38	0.69 (1.3)	−0.10	16	7.16 (2.7)	−0.23
Much better	8	6.18 (3)	−0.37	4	13.06 (9.19)	−0.3
A bit better	30	−0.78 (1.29)	−0.01	12	5.19 (2.06)	−0.22
The same	22	−1.36 (1.4)	1	7	3.87 (5.1)	1
Worse	11	−0.01 (1.77)	–	7	5.57 (3.28)	–
A bit worse	10	0.53 (1.86)	–	7	5.57 (3.28)	–
Much worse	1	−5.44 (NA)	–	0	0	–
SB_diff_he	MCID = −0.22			–		
Better	10	−1.35 (1.89)	0.06	–	–	–
The same	47	−1.13 (1.05)	1	–	–	–
Worse	6	3.15 (2.82)	–	–	–	–
SB_diff_me	MCID = 3.68			–		
Better	11	2.12 (1.89)	−0.17	–	–	–
The same	45	−1.56 (1.08)	1	–	–	–
Worse	7	−0.10 (2.18)	–	–	–	–
SB_diff_ir	MCID = 2.28			–		
Better	8	1.37 (1.75)	−0.10	–	–	–
The same	54	−0.91 (0.98)	1	–	–	–
Worse	1	−9.83 (NA)	–	–	–	–
SB_diff_sup	MCID = **2.47**			–		
Better	9	**1.26 (2.36)**	**−0.29**	–	–	–
The same	49	−1.21 (1)	1	–	–	–
Worse	5	0.04 (3.07)	–	–	–	–
PCS_cat	MCID = **4.83**			MCID = 6.39		
Better	17	**3.74 (1.7)**	**−0.32**	8	10.82 (4.16)	−0.25
The same	37	−1.09 (1.23)	1	20	4.43 (1.71)	1
Worse	26	−2.18 (1.24)	–	12	8.34 (2.83)	–

Note: MCID is the between-group minimal clinically important difference calculated as the mean change in the outcomes of the anchor group 'better' minus mean change in 'same'. The within-group MCIDs are the mean change values shown in the 'better' groups. Patients who replied 'much better' and 'a bit better' were included in the group 'better' and if they replied 'a bit worse' and 'much worse' in the group 'worse'.

Both outcomes are expressed as a percentage of their predicted normal values.

Abbreviations: 6MWT, 6-min walk test; FVC_{up}, forced vital capacity in the upright position; GH_change, self-reported change in health (question 2 of the SF-36 questionnaire); Mean change, mean difference in the value at baseline and 1-year follow-up; PCS_cat, a categorization of the Physical Component Summary (PCS) score; PF_change, self-reported change in physical functioning (question in the Pompe-specific questionnaire); SB_diff_he, SB_diff_me, SB_diff_ir and SB_diff_sup, change in shortness of breath in heavy exercise, in mild exercise, in rest and in supine position, respectively, from baseline to follow-up; SE, standard error.

[a]Correlation of the mean change in the outcome with the answer options (of the respective row and 'same'). In bold are the MCIDs with correlations with absolute value ≥0.29 and appropriate numbers of patients. The cells with a negative sign (–) are the ones that do not pertain to an improvement.

TABLE 4 Anchor-based MCIDs for FVC_{up} and 6MWT based on modelling of two anchors.

Outcome	FVC_{up} (%)		6MWT (%)	
Anchor	**GH_change**	**PF_change**	**GH_change**	**PF_change**
MCID	0.33	0.41	0.35[a]	0.58
SE pooled	0.88	0.94	1.48	2.11
Lower limit	−1.4	−1.43	−2.55	−3.57
Upper limit	2.07	2.24	3.25	4.72

Note: Both FVCup and 6MWT were expressed as a percentage of their predicted normal values.

Abbreviations: 6MWT, 6-min walk test; FVC_{up}, forced vital capacity in upright position; GH_change, self-reported change in health (question 2 of the SF-36 questionnaire); MCID, between-group minimal clinically important difference; PF_change, self-reported change in physical functioning (question in the Pompe-specific questionnaire); SE pooled, pooled standard error.

[a]This is the only estimate where the anchor is considered appropriate based on the time interval method.

TABLE 5 Distribution-based MCIDs using baseline outcome data and the change therein.

	Baseline (BL)			Change values (FU1 − BL)		
Outcome	**Patients (*n*)[a]**	**Mean of baseline values**	**MCID = 0.5 × $SD_{baseline}$**	**Patients (*n*)[b]**	**Mean of change values**	**MCID = 0.5 × $SD_{of\ changes}$**
FVC_{up} (%)	100	64.97	10.91	98	−0.27	3.65
6MWT (%)	55	59.85	10.50	53	5.81	4.53

Note: The change values emerged from subtracting the clinical outcome values observed in the baseline interval from those in the follow-up interval.Both outcomes are expressed as a percentage of their predicted normal values.

Abbreviations: 6MWT, 6-min walk test; BL, baseline, i.e. the time-point closest to start of ERT in the interval [−0.5, −0.49]; ERT, enzyme replacement therapy; FU1, 1-year follow-up, i.e. the time-point closest to 1 year in the time interval [0.5, 1.7]; FVC_{up}, forced vital capacity in upright position; MCID, minimal clinically important difference; *n*, number; SD, standard deviation.

[a]Patients with values in the baseline interval [−0.5, 0.49].

[b]Patients with values in both the baseline BL and the FU1 interval [0.5, 1.7].

Anchor-based approach based on modelling

The between-group MCIDs from this approach (Table 4) are considerably smaller than the results of the time interval method. The confidence intervals contain zero indicating that these results are not statistically significant. Based on the correlations obtained in the time interval method, only the MCID of 0.35% points estimated for the 6MWT using GH_change would be considered appropriate.

Distribution-based method

Table 5 shows the results of the distribution-based method using BL and the change (FU1 − BL) values. Based on the BL values, the MCIDs for FVC_{up} and the 6MWT are 10.91% points and 10.50% points, respectively, whilst using the change values they are 3.65% points and 4.53% points. The BL values result in higher MCIDs than the change values.

Triangulating the between-group results

Figure 1 presents all estimated MCIDs for FVC_{up} (left plot) and the 6MWT (right plot). For the anchor-based approaches, the between-group MCIDs are presented. For the interval estimates, the size of the dots corresponds to the correlation (absolute value) multiplied by the total number of patients ('better' and 'same') divided by 3. MCIDs from the modelling and distribution-based approach (diamond shapes and asterisk) have one size only since a correlation for these approaches cannot be calculated.

Using the anchor-based approaches, the MCIDs for FVC_{up} ranged from −0.22% points to 4.83% points. Based on the two anchors that met the appropriateness criteria (SB_diff_sup and PCS_cat; $|r| \geq 0.29$, $n \geq 8$ in each group and $n \geq 35$ in total for MCID), the between-group MCID for FVC_{up} ranges from 2.47% points to 4.83% points.

For the 6MWT, the anchor-based estimates ranged from 0.35% points to 7.47% points. Only one anchor (GH_change) met the appropriateness criteria. Therefore, the MCID for the 6MWT ranges from 0.35% points (modelling) to 7.47% points (interval).

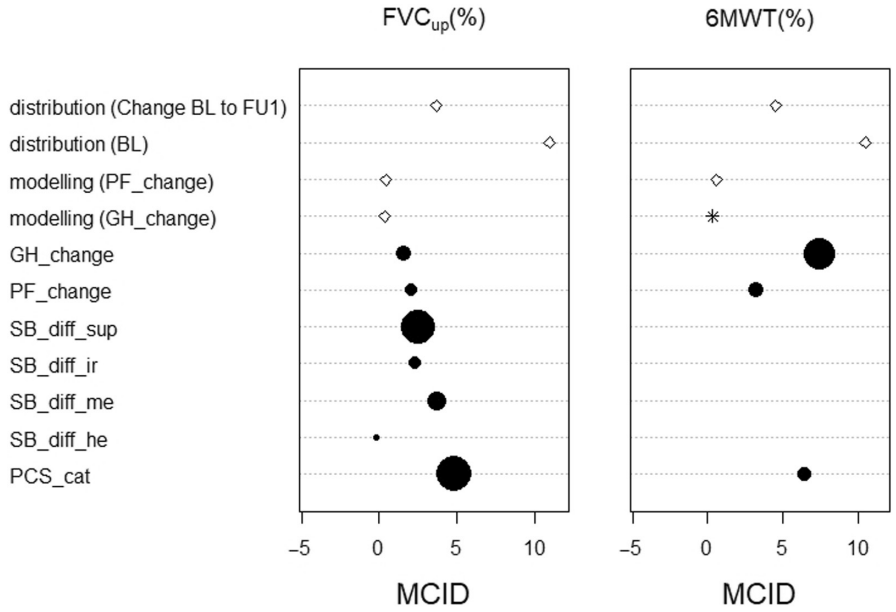

FIGURE 1 Estimated between-group minimal clinically important differences (MCIDs) for lung function (FVC$_{up}$, %) and distance walked (6MWT, %). The MCIDs were calculated based on the distribution-based method (0.5 × SD), modelling and the mean change method. Both FVCup and 6MWT are expressed as a percentage of their predicted normal values. FVC$_{up}$, forced vital capacity in upright position; 6MWT, 6-min walk test; distribution (change BL to FU1), distribution-based MCID calculated based on the change in values observed in the baseline (BL) interval [−0.5, 0.49] to those observed in the 1-year follow-up (FU1) interval [0.5, 1.7]; distribution (BL), distribution-based MCID calculated using values observed in the baseline interval; modelling (PF/GH_change), modelling MCID using specific anchor. The remaining MCIDs are calculated using the anchor-based mean change method using the anchors GH_change, self-reported change in health (question 2 of the SF-36 questionnaire); PF_change, self-reported change in physical functioning (question in the Pompe-specific questionnaire); SB_diff_sup, SB_diff_ir, SB_diff_me and SB_diff_he, change in shortness of breath from baseline to follow-up in supine position, in rest, in mild exercise and in heavy exercise, respectively; and PCS_cat, a categorization of the change in Physical Component Summary (PCS) score of the SF-36. The size of the MCIDs with bullets was calculated by multiplying the absolute correlation between the anchor and the clinical outcome with the total number of patients in the groups 'better' and 'same' divided by 3 (correlation × total number of patients/3). For the MCIDs with diamond symbols, the correlation could not be calculated. *The only estimate from the anchor-based modelling approach which is considered appropriate since the correlation between anchor and outcome was >0.29 in the anchor-based time interval approach.

The distribution-based estimates were provided as supportive evidence. The MCIDs based on the change values fall within the range estimated using the appropriate anchors. Nevertheless, the MCIDs based on the variation in the BL values are larger than any of the anchor estimates.

DISCUSSION

To our knowledge this is the first paper estimating MCIDs for FVC$_{up}$ and the 6MWT in adult patients with Pompe disease who were treated with ERT. Using the mean change method and modelling, it was estimated that the between-group MCID (after 1 year of treatment) ranged from 2.47% points to 4.83% points for FVC$_{up}$. This means that if one treatment group improves 2.5% points more than another this may already be clinically important and a 5% point difference is. The within-group MCID for FVC$_{up}$ (1.26% points to 3.74% points) was slightly lower.

For the 6MWT, the estimated between-group MCID ranged from 0.35% points to 7.47% points. This corresponds to a range of 2.2–46.6 m for a 50-year-old man of 1.75 m and 80 kg and 2.0–42.1 m for a woman of the same age, height and weight. To calculate the MCID in metres for patients with other sex, age, height and weight, the formulae of Enright and Sherrill [26] can be used. The within-group MCID for the 6MWT was higher (11.53% points), because patients remaining the 'same' on the anchor showed an improvement on the 6MWT.

A range of MCID values was obtained, since various anchors and approaches were applied. Anchor appropriateness was judged by the Spearman correlation and the number of patients. As a result, only one or two anchors were selected as being appropriate for each outcome. There is one other unpublished study estimating the MCID of FVC$_{up}$ and the 6MWT in Pompe disease, based on data from the COMET study and using different anchors (World Symposium 2023, Berger et al. [38] poster LB-11). This study reported a between-group MCID of 2.0% points (1.0–3.0) for FVC$_{up}$ and 33 m (17.0–50.0) for the 6MWT. These estimates overlap with our results, although their estimated range is a bit lower for FVC$_{up}$ and higher for the 6MWT.

Compared to estimates of the MCID for other diseases, our 6MWT results seem comparable. The review by Schrover et al. [39] shows a range of estimates for different diseases, ranging from 11 to 54 m. For Duchenne muscular dystrophy the MCID for the 6MWT using distribution-based methods was estimated by Henricson et al. [40] as 26.4 m and by McDonald et al. [41] as 28.5 m and 31.7 m.

The results from our interval and modelling approaches differed considerably, with MCIDs from modelling nearing zero. The modelling approach estimates anchor and outcome values at exactly 1 year of treatment by taking into consideration other confounders. This may introduce further uncertainty to the estimates. Therefore, it is suggested that the modelling results, which provided the lower boundary for the 6MWT, are interpreted with caution. Nevertheless, for a progressive disease like Pompe disease, it is not entirely unthinkable that no difference, that is, a stabilization, could also be perceived as a meaningful improvement.

The time interval approach also introduces uncertainty due to the fact that the paired anchor–outcome observations were not taken at the same time. Given that the median time difference for all outcome–anchor combinations was close to zero, it was felt that this uncertainty was limited. It was decided a priori to combine the outcome options 'much better' and 'a little better' of GH_change and PF_change into one 'better' group to increase the number of observations. This means that the upper boundary for the 6MWT, which is based on GH_change, may be inflated somewhat. The upper boundary for the FVC_{up} was based on the anchor PCS_cat, which used an increase ≥5 to differentiate patients who felt better. The true clinically meaningful threshold for the PCS in Pompe disease, however, is not known. Fu et al. [42] estimated that the MCID for the PCS in patients with stroke ranged from 1.8 to 3 points, whilst Copay et al. [43] estimated that the MCID for the PCS in lumbar spine surgery patients is 4.9 points.

Further limitations to the anchor-based methods in general include that they do not account for the presence of concomitant diseases and their treatments. For example, treatment for depression may improve GH_change but not FVC. Also, when starting treatment, optimism about the treatment effect may result in patients reporting feeling better than they actually do, which may underestimate the MCID. Last, these estimates apply to the average patient in our cohort and are not appropriate to interpret changes in patients with very poor or good FVC and 6MWT.

The distribution-based method, when applied to the change over time, suggests a threshold within the range of the MCIDs mentioned above. Nevertheless, when the traditional method of 0.5 × SD of the baseline variation between patients was used [34], a much higher estimate was obtained. Since our cohort includes a broad spectrum of patients with Pompe disease, ranging from very mildly to severely affected patients, this was to be expected. Clinical trials usually focus on a much narrower range of disease severity, and in that case the 0.5 × SD of the baseline variation will be much smaller. Therefore we believe that the 0.5 × SD of variation in the change values (changes from baseline to year 1) is a better measure to support the question of what change is clinically important.

Clinical application

The between-group MCIDs presented above can be used to interpret differences observed in both future and past trials, whilst the within-group MCIDs can be used to assess changes over time in one treatment group and in cohort studies. The initial placebo-controlled trial of alglucosidase alfa [7] estimated the treated group to have improved 3.4% points in FVC_{up} more than the placebo group. This falls within our estimated between-group MCID range and suggests that this improvement may be clinically meaningful for patients. The improvement on the 6MWT of 28.1 m may also be clinically important. Two more recent trials on safety and efficacy of avalglucosidase alfa and cipaglucosidase alfa plus miglustat reported patients treated with avalglucosidase alfa to have increased 2.89% points in FVC_{up} and 5.02% points in the 6MWT (2.43% points and 4.71% points more respectively than those treated with alglucosidase alfa) [13] and those treated with cipaglucosidase alfa plus miglustat to have a mean change of −0.9% points in FVC_{up} and 20.8 m in the 6MWT (3% points and 13.6 m more than those treated with alglucosidade alfa and placebo) [14]. All these increases, except for FVC_{up} in the cipaglucosidase trial, are within the range of our estimated MCIDs and suggest possible clinical importance.

CONCLUSION

It is concluded that for adult patients with Pompe disease the MCID of a beneficial effect on FVC_{up} in a trial ranges from 2.47% points to 4.83% points. For the 6MWT it ranges from 0.35% points to 7.47% points which is equivalent to a distance of 2.18–46.61 m and 1.97–42.13 m for a man and a woman of age 50, height 1.75 m and weight 80 kg, respectively. The thresholds presented in this study can be used to interpret group-level results from Pompe disease trials and cohort studies. Also, they can be used in sample size calculations for adequate powering of a study on Pompe disease.

AUTHOR CONTRIBUTIONS

Aglina Lika: Writing – original draft; formal analysis; investigation; methodology. **Eleni-Rosalina Andrinopoulou:** Formal analysis; writing – review and editing; supervision. **Nadine A. M. E. van der Beek:** Conceptualization; writing – review and editing; supervision; data curation. **Dimitris Rizopoulos:** Formal analysis; writing – review and editing; supervision. **Ans T. van der Ploeg:** Conceptualization; funding acquisition; data curation; supervision; writing – review and editing. **Michelle E. Kruijshaar:** Conceptualization; supervision; funding acquisition; data curation; investigation; writing – review and editing; methodology.

ACKNOWLEDGEMENTS

All patients are thanked for participating in both follow-up studies. The authors are grateful to all clinical staff who have helped to collect the physical outcomes over time.

FUNDING INFORMATION

This study was supported by Sanofi. Sanofi did not have any influence on the content or design of the study and the drafting and content of the manuscript nor did Sanofi cover the costs of the enzyme

therapy. This study was also supported by ZonMW The Netherlands Organization for Health Research and Development (projects 152001005, 80-83600-98-13007 and 05-09-2007); Prinses Beatrix Spierfonds (projects OP07-08, W.OR13-21 and W.OR15-10); TKI – Health Holland (project LSHM16008); SSWO-Sophia Children's Hospital Foundation (project 687). Several of the authors of this publication are members of the European Reference Networks for Hereditary Metabolic Disorders (Metab-ERN) and/or for Rare Neuromuscular Diseases (EURO-NMD) and/or of the Netherlands Neuromuscular Centre (NL-NMD).

CONFLICT OF INTEREST STATEMENT

A.T. van der Ploeg received funding for research, clinical trials and as an advisor from various industries working on ERT or next-generation therapies in the field of Pompe disease, other lysosomal storage diseases, and neuromuscular disorders under agreements with Erasmus MC University Medical Centre and the relevant industry. N.A.M.E. van der Beek received funding for research and as advisor from various industries working on ERT or next generation therapies in the field of Pompe disease under agreements with Erasmus MC University Medical Centre and the relevant industry. The other co-authors (A. Lika, D. Rizopoulos, E.R. Andrinopoulou and M.E. Kruijshaar) declare that they have no conflict of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are not publicly available to protect subject privacy.

ORCID

Aglina Lika https://orcid.org/0009-0004-7474-3568
Nadine A. M. E. van der Beek https://orcid.org/0000-0001-9161-3301
Ans T. van der Ploeg https://orcid.org/0000-0002-3359-1324
Michelle E. Kruijshaar https://orcid.org/0009-0009-8700-0591

REFERENCES

1. Reuser AJJ, Hirschhorn R, Kroos MA. Pompe disease: glycogen storage disease type II, acid α-glucosidase (acid maltase) deficiency. In: Valle DL, Antonarakis S, Ballabio A, Beaudet AL, Mitchell GA, eds. *The Online Metabolic and Molecular Bases of Inherited Disease*. McGraw-Hill Education; 2019.
2. van der Ploeg AT, Kruijshaar ME, Toscano A, et al. European consensus for starting and stopping enzyme replacement therapy in adult patients with Pompe disease: a 10-year experience. *Eur J Neurol*. 2017;24(6):768-e31.
3. Güngör D, Kruijshaar ME, Plug I, et al. Impact of enzyme replacement therapy on survival in adults with Pompe disease: results from a prospective international observational study. *Orphanet J Rare Dis*. 2013;8:49.
4. Anderson LJ, Henley W, Wyatt KM, et al. Effectiveness of enzyme replacement therapy in adults with late-onset Pompe disease: results from the NCS-LSD cohort study. *J Inherit Metab Dis*. 2014;37(6):945-952.
5. Angelini C, Semplicini C, Ravaglia S, et al. Observational clinical study in juvenile-adult glycogenosis type 2 patients undergoing enzyme replacement therapy for up to 4 years. *J Neurol*. 2012;259(5):952-958.
6. Bembi B, Pisa FE, Confalonieri M, et al. Long-term observational, non-randomized study of enzyme replacement therapy in late-onset glycogenosis type II. *J Inherit Metab Dis*. 2010;33(6):727-735.
7. van der Ploeg AT, Clemens PR, Corzo D, et al. A randomized study of alglucosidase alfa in late-onset Pompe's disease. *N Engl J Med*. 2010;362(15):1396-1406.
8. Strothotte S, Strigl-Pill N, Grunert B, et al. Enzyme replacement therapy with alglucosidase alfa in 44 patients with late-onset glycogen storage disease type 2: 12-month results of an observational clinical trial. *J Neurol*. 2010;257(1):91-97.
9. Regnery C, Kornblum C, Hanisch F, et al. 36 months observational clinical study of 38 adult Pompe disease patients under alglucosidase alfa enzyme replacement therapy. *J Inherit Metab Dis*. 2012;35(5):837-845.
10. Kuperus E, Kruijshaar ME, Wens SCA, et al. Long-term benefit of enzyme replacement therapy in Pompe disease: a 5-year prospective study. *Neurology*. 2017;89(23):2365-2373.
11. Harlaar L, Hogrel JY, Perniconi B, et al. Large variation in effects during 10 years of enzyme therapy in adults with Pompe disease. *Neurology*. 2019;93(19):e1756-e1767.
12. Güngör D, Kruijshaar ME, Plug I, et al. Quality of life and participation in daily life of adults with Pompe disease receiving enzyme replacement therapy: 10 years of international follow-up. *J Inherit Metab Dis*. 2016;39(2):253-260.
13. Diaz-Manera J, Kishnani PS, Kushlaf H, et al. Safety and efficacy of avalglucosidase alfa versus alglucosidase alfa in patients with late-onset Pompe disease (COMET): a phase 3, randomised, multicentre trial. *Lancet Neurol*. 2021;20(12):1012-1026.
14. Schoser B, Roberts M, Byrne BJ, et al. Safety and efficacy of cipaglucosidase alfa plus miglustat versus alglucosidase alfa plus placebo in late-onset Pompe disease (PROPEL): an international, randomised, double-blind, parallel-group, phase 3 trial. *Lancet Neurol*. 2021;20(12):1027-1037.
15. Stevens D, Milani-Nejad S, Mozaffar T. Pompe disease: a clinical, diagnostic, and therapeutic overview. *Curr Treat Options Neurol*. 2022;24(11):573-588.
16. Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. *Control Clin Trials*. 1989;10(4):407-415.
17. Crosby RD, Kolotkin RL, Williams GR. Defining clinically meaningful change in health-related quality of life. *J Clin Epidemiol*. 2003;56(5):395-407.
18. Revicki DA, Cella D, Hays RD, Sloan JA, Lenderking WR, Aaronson NK. Responsiveness and minimal important differences for patient reported outcomes. *Health Qual Life Outcomes*. 2006;4:70.
19. Angst F, Aeschlimann A, Angst J. The minimal clinically important difference raised the significance of outcome effects above the statistical level, with methodological implications for future studies. *J Clin Epidemiol*. 2017;82:128-136.
20. van der Meijden JC, Güngör D, Kruijshaar ME, Muir AD, Broekgaarden HA, van der Ploeg AT. Ten years of the international Pompe survey: patient reported outcomes as a reliable tool for studying treated and untreated children and adults with non-classic Pompe disease. *J Inherit Metab Dis*. 2015;38(3):495-503.
21. de Vries JM, van der Beek NA, Hop WC, et al. Effect of enzyme therapy and prognostic factors in 69 adults with Pompe disease: an open-label single-center study. *Orphanet J Rare Dis*. 2012;7:73.
22. van der Beek NA, de Vries JM, Hagemans ML, et al. Clinical features and predictors for disease natural progression in adults with Pompe disease: a nationwide prospective observational study. *Orphanet J Rare Dis*. 2012;7:88.
23. Quanjer PH, Stanojevic S, Cole TJ, et al. Mutli-ethnic reference values for spirometry for the 3-95 year age range: the global lung function 2012 equations. *Eur Respir J*. 2012;40(6):1324-1343.

24. American Thoracic Society/European Respiratory Society. ATS/ERS statement on respiratory muscle testing. *Am J Respir Crit Care Med*. 2002;166(4):518-624.
25. American Thoracic Society Committee on Proficiency Standards for Clinical Pulmonary Function Laboratories. ATS statement: guidelines for the six-minute walk test. *Am J Respir Crit Care Med*. 2002;166(1):111-117.
26. Enright PL, Sherrill DL. Reference equations for the six-minute walk in healthy adults. *Am J Respir Crit Care Med*. 1998;158(5 Pt 1):1384-1387.
27. Aaronson NK, Muller M, Cohen PD, et al. Translation, validation, and norming of the Dutch language version of the SF-36 health survey in community and chronic disease populations. *J Clin Epidemiol*. 1998;51(11):1055-1068.
28. Ware JE Jr, Sherbourne CD. The MOS 36-item Short-Form Health Survey (SF-36). I. Conceptual framework and item selection. *Med Care*. 1992;30(6):473-483.
29. Redelmeier DA, Lorig K. Assessing the clinical importance of symptomatic improvements. An illustration in rheumatology. *Arch Intern Med*. 1993;153(11):1337-1342.
30. Angst F. The minimal clinically important difference assigns significance to outcome effects. *J Rheumatol*. 2016;43:258-259.
31. Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *J Clin Epidemiol*. 2008;61(2):102-109.
32. Tsujimoto Y, Fujii T, Tsutsumi Y, et al. Minimal important changes in standard deviation units are highly variable and no universally applicable value can be determined. *J Clin Epidemiol*. 2022;145:92-100.
33. de Vet HC, Ostelo RW, Terwee CB, et al. Minimally important change determined by a visual method integrating an anchor-based and a distribution-based approach. *Qual Life Res*. 2007;16(1):131-142.
34. Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. *Med Care*. 2003;41(5):582-592.
35. Ren D, Wu T, Wan C, et al. Exploration of the methods of establishing the minimum clinical important difference based on anchor and its application in the quality of life measurement scale QLICP-ES (V2.0) for esophageal cancer. *Health Qual Life Outcomes*. 2021;19:173.
36. Turner D, Schünemann HJ, Griffith LE, et al. The minimal detectable change cannot reliably replace the minimal important difference. *J Clin Epidemiol*. 2010;63(1):28-36.
37. Wright AA, Cook CE, Baxter GD, Dockerty JD, Abbott JH. A comparison of 3 methodological approaches to defining major clinically important improvement of 4 performance measures in patients with hip osteoarthritis. *J Orthop Sports Phys Ther*. 2011;41(5):319-327.
38. Berger KI, Ivanescu C, Msihid J, et al. CO168 defining Clinically Meaningful Thresholds (CMT) for Forced Vital Capacity (FVC) and Six-Minute Walk Test (6MWT) in patients with Late-Onset Pompe Disease (LOPD). *Value Health*. 2023;26(6):S46.
39. Schrover R, Evans K, Giugliani R, Noble I, Bhattacharya K. Minimal clinically important difference for the 6-min walk test: literature review and application to Morquio A syndrome. *Orphanet J Rare Dis*. 2017;12(1):78.
40. Henricson E, Abresch R, Han JJ, et al. The 6-minute walk test and person-reported outcomes in boys with Duchenne muscular dystrophy and typically developing controls: longitudinal comparisons and clinically-meaningful changes over one year. *PLoS Curr*. 2013;5.
41. McDonald CM, Henricson EK, Abresch RT, et al. The 6-minute walk test and other clinical endpoints in Duchenne muscular dystrophy: reliability, concurrent validity, and minimal clinically important differences from a multicenter study. *Muscle Nerve*. 2013;48(3):357-368.
42. Fu V, Weatherall M, McNaughton H. Estimating the minimal clinically important difference for the physical component summary of the Short Form 36 for patients with stroke. *J Int Med Res*. 2021;49(12):3000605211067902.
43. Copay AG, Glassman SD, Subach BR, Berven S, Schuler TC, Carreon LY. Minimum clinically important difference in lumbar spine surgery patients: a choice of methods using the Oswestry disability index, medical outcomes study questionnaire Short Form 36, and pain scales. *Spine J*. 2008;8(6):968-974.
44. Mislevy RJ. Review of statistical analysis with missing data. *J Educ Stat*. 1991;16:150-155. doi:10.2307/1165119

How to cite this article: Lika A, Andrinopoulou E-R, van der Beek NAME, Rizopoulos D, van der Ploeg AT, Kruijshaar ME. Establishing how much improvement in lung function and distance walked is clinically important for adult patients with Pompe disease. *Eur J Neurol*. 2024;00:e16223. doi:10.1111/ene.16223

APPENDIX

Establishing minimal clinically important difference for lung function and distance walked in adult Pompe disease patients

A.1 | PART A: MODELLING APPROACH

A.1.1. | Why estimate anchor-based MCID using a statistical model?

As an alternative to using values observed in the time interval around the start of ERT ($t=0$/baseline) and 1-year follow-up ($t=1$/FU1), the between-group MCIDs were estimated for GH_change and PF_change using a modelling approach. The rationale for this was that it was wanted to remove the effect of variation in the time at which the anchor and the outcomes were assessed; with the time interval approach the anchor may be assessed earlier or later than the clinical outcome. The modelling approach allows this difference to be bypassed as well as adjustments to be made for other variables that may not be equally distributed between anchor groups, such as sex and disease duration. Modelling was only applied for GH_change and PF_change, as these are the self-reported changes in health and physical functioning. The remaining anchors require the calculation of a change between two observation points, which makes the modelling too complicated.

A.1.2. | Data used to fit the models

For the modelling approach, the measurements of the anchors (GH_change and PF_change) and the outcomes that were collected up to 1.7 years of treatment with ERT were used. Our interest was in the MCID in the first year of treatment, but included some follow-up time before $t=0$ and after $t=1$ to give the model sufficient data around these time points.

For the anchors, the measurements of 'same' and 'better' were used, and the measurements of 'worse' were excluded, since it was wanted to define the MCID for an improvement.

A.1.3. | Modelling

To estimate the values of the anchors (i.e., 'better' and 'same') of each patient at exactly the first year of treatment, a logistic mixed-effects model was fitted to the anchor measurements ('same' and 'better') correcting for sex, disease duration at the start of ERT, the nonlinear evolution of the anchor (natural splines of time with two degrees of freedom) and if the measurement was taken before or during ERT. Random intercept and random effects of the natural splines of time were also included in the model. Once this model was fitted to the existing data, the values for each patient were estimated at exactly $t = 1$, using the coefficients of the logistic model.

Next, these estimated values were added to the dataset with the FVC_{up}/6MWT measurements and a linear mixed-effects model was fitted for FVC_{up}/6MWT, assuming a nonlinear evolution of FVC_{up}/6MWT over time (natural splines of time with two degrees of freedom), the anchor value at the first year of treatment and their interaction. Random intercept and random effects of the natural splines of time were also included in the model.

Based on the model mentioned above, the expected value of the FVC_{up}/6MWT at exactly year 0 and 1 could be estimated for the groups 'better' and 'same', and the between-group MCID for FVC_{up}/6MWT was calculated as $MCID_{between} = (FVC_{up}/6MWT_{year=1} - FVC_{up}/6MWT_{year=0})_{better} - (FVC_{up}/6MWT_{year=1} - FVC_{up}/6MWT_{year=0})_{same}$.

The above process was repeated 200 times to account for the additional uncertainty in standard errors of the estimated values. The pooled estimates and the standard error (SE) of the MCIDs were obtained based on Rubin's rule [44].

Whilst correlations cannot be estimated for this method, the correlations obtained in the anchor-based time interval approach were used to give an indication of the appropriateness of these two anchors for the modelling approach.

A.2 | PART B: FLOWCHART

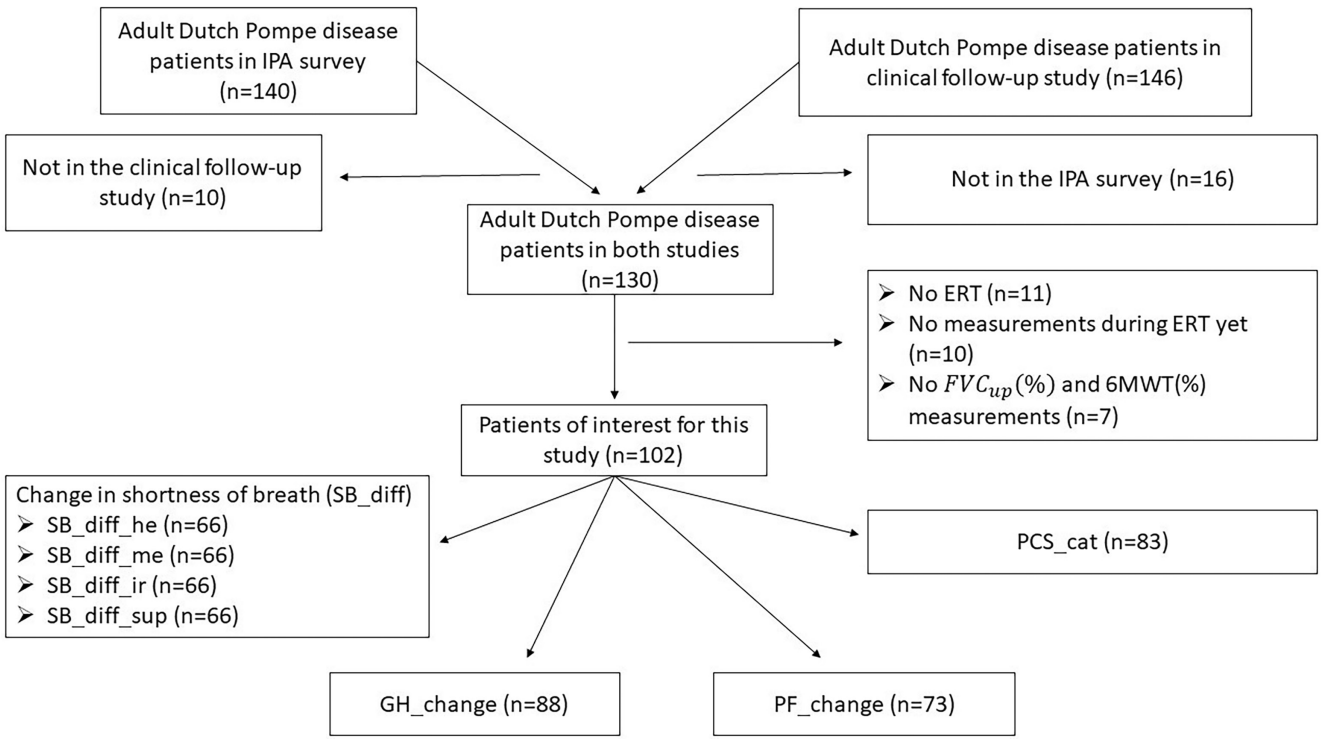

FIGURE A1 Flowchart of the study population. GH_change, patients with data on self-reported change in health (question 2 of the 36-item Short-Form Health Survey version 1 and 2) in the 1 year of follow-up interval (FU1 [0.5, 1.7]); *n*, number of patients; PCS_cat, patients with data on categorized variable of the Physical Component Summary measure in the baseline interval (BL [−0.5, 0.49]) and FU1; PF_change, patients with data on self-reported change in physical function (question of the IPA survey) in FU1; SB_diff, difference in the answer to the shortness of breath question in rest (SB_diff_ir), in supine position (SB_diff_sup), during heavy exercise (SB_diff_he) and during mild exercise (SB_diff_me) between BL and FU1.

A.3 | PART C: NUMBER OF PATIENTS AND TIME DIFFERENCE FOR OUTCOME–ANCHOR COMBINATIONS

The clinical outcome measures and the anchors used in this study were not measured at the same time. The median time difference between the anchors and the outcomes is shown in Table 2 of the paper and showed this was close to zero for all outcome–anchor combinations.

Whilst in some combinations the full range included 9 months, the interquartile range showed that at BL 50% of the outcome and anchor measurements were taken less than 1 month apart (except for the combination FVC_{up} and SB_diff, where 50% of the measurements were <3 months apart; interquartile range of the time difference [−0.17, 0.24]). At FU1, 50% of the outcome and anchor measurements were taken less than 4 months apart (<4 months, 6MWT with GH_change and PF_change; <3 months, 6MWT with PCS_cat; <3 months but >2 months, FVC_{up} with PF_change and SB_diff; <2 months, FVC_{up} with PCS_cat; <1.5 months, FVC_{up} with GH_change).

Table A1 provides the number of patients available per anchor/outcome measure in the BL and FU1 intervals. The lower part of the table presents the number of patients available for each anchor–outcome combination in both time intervals.

TABLE A1 Number of patients available for each anchor, outcome and anchor–outcome combination at the baseline and follow-up time interval.

	Number of patients with data in	
	Baseline interval	**Follow-up interval**
Measurement time since start ERT		
FVC_{up} (%)	100	99
6MWT (%)	55	59
GH_change	88	88
PF_change	73	73
SB_diff	6	6
PCS_cat	83	83
Difference in time of outcome and anchor		
FVC_{up} (%) – GH_change ($n=85$)	85	85
FVC_{up} (%) – PF_change ($n=71$)	71	71
FVC_{up} (%) – SB_diff ($n=63$)	63	63
FVC_{up} (%) – PCS_cat ($n=80$)	80	80
6MWT (%) – GH_change ($n=43$)	43	43
6MWT(%) – PF_change ($n=30$)	30	30
6MWT (%) – PCS_cat ($n=40$)	40	40

Abbreviations: 6MWT (%), 6-min walk test expressed as percentage of predicted normal values; baseline interval, values in the time interval [−0.5, 0.49]; ERT, enzyme replacement therapy; follow-up interval, values in the time interval [0.5, 1.7]; FVC_{up} (%), forced vital capacity in the upright position expressed as percentage of predicted normal values; GH_change, self-reported change in health (question 2 of the SF-36 questionnaire); PCS_cat, a categorization of the Physical Component Summary (PCS) score of the SF-36; PF_change, self-reported change in physical functioning (question in the Pompe-specific questionnaire); SB_diff, the change in shortness of breath in the first year of treatment with ERT.