scientific reports

Check for updates

OPEN

Impact of adapted physical activity and diet counselling on health-related quality of life in women undergoing adjuvant breast cancer therapy

Célia Touraine[1,2,7]✉, William Jacot[4], Sophie Gourgou[1,2], Marion Carayol[5], Pierre Senesse[6], Grégory Ninot[3,6] & Caroline Mollevi[1,2,8]

In the monocentric APAD1 trial, 143 women with non-metastatic breast cancer were randomised to undergo either an adapted physical activity and diet counselling (APAD) program or usual care. Health-related quality of life (HRQoL) was prospectively evaluated using the EORTC QLQ-C30 questionnaire at baseline, during treatment (adjuvant chemotherapy and radiotherapy) and during follow-up. Our objective was two-fold: to analyse the impact of APAD on HRQoL using three approaches; to illustrate the advantages and disadvantages of each approach, and derive methodological recommendations. Analytical approaches utilised were: statistical testing to compare the mean HRQoL scores between baseline and end of study in both groups and the mean HRQoL scores between the two groups at the different assessment times; linear mixed models that modelled the longitudinal score data in both groups and tested whether the score trajectories were different between the groups; a survival analysis comparing the time to deterioration of HRQoL between the groups using a minimal clinically important difference. This study shows a substantial clinical benefit of the APAD intervention on HRQoL, especially for global health status/HRQoL, functioning scales and the fatigue symptom scale. Furthermore, this study highlights the advantages and disadvantages of three standard approaches used to analyse HRQoL data.

Trial registration: The APAD1 study was registered with ClinicalTrials.gov (number NCT01495650, date 20/12/2011).

Keywords Health-related quality of life, Longitudinal data, Linear mixed models, Time to deterioration, Clinical trial, Breast cancer

Abbreviations

HRQoL	Health-related quality of life
APAD	Adapted physical activity and diet counselling
UC	Usual care
QLQ-C30	EORTC core quality of life questionnaire
LMM	Linear mixed model
MCID	Minimal clinically important difference

[1]Biometrics Unit, Cancer Institute of Montpellier, University of Montpellier, 208 avenue des apothicaires, Parc Euromédecine, 34298 Montpellier Cedex 5, France. [2]French National Platform Quality of Life and Cancer, Montpellier, France. [3]Desbrest Institute of Epidemiology and Public Health, University of Montpellier, INSERM, INRIA, Montpellier, France. [4]Cancer Institute of Montpellier, INSERM U1194, University of Montpellier, Montpellier, France. [5]Youth – Physical and Sports Activity – Health Laboratory (JAP2S), University of Toulon, Toulon, France. [6]Department of Supportive and Palliative Care, Cancer Institute of Montpellier, University of Montpellier, Montpellier, France. [7]Desbrest Institute of Epidemiology and Public Health, University of Montpellier, INSERM, ICM, Montpellier, France. [8]Desbrest Institute of Epidemiology and Public Health, University of Montpellier, INSERM, CHU Montpellier, Montpellier, France. ✉email: celia.touraine@gmail.com

TTD	Time to deterioration
HR	Hazard ratio
CI	Confidence interval

Breast cancer survival rates have increased in the last few decades due to earlier diagnosis and more effective treatments. Nevertheless, early breast cancer adjuvant therapy is associated with severe and clinically impactful side effects, most commonly fatigue[1,2]. Cancer-related fatigue has been shown to affect patients' health-related quality of life (HRQoL)[3], even more so than nausea, depression and pain[4]. Non-pharmacological interventions such as physical activity can reduce fatigue[5,6] and improve HRQoL in patients with breast cancer[7]. The primary objective of the randomised, controlled Adapted Physical Activity and Diet counselling (APAD1) trial[8] was to assess the impact of a diet-exercise intervention on cancer-related fatigue in women with early breast cancer receiving adjuvant treatment (chemotherapy and radiotherapy). One of the secondary objectives was to assess the impact of the diet-exercise intervention on HRQoL. To achieve this, HRQoL was prospectively assessed in patients receiving usual care (UC group) and patients receiving usual care plus the diet-exercise intervention (APAD group) using the EORTC core quality of life questionnaire (QLQ-C30). The impact of the diet-exercise intervention on the global health status/HRQoL scale and on the functional scales of the QLQ-C30 have been previously investigated[9]. The latter used a linear mixed model (LMM) approach, treating time as a categorical variable. In the current study, we performed a more in-depth analysis using three different and complementary approaches: statistical testing, time to deterioration and LMMs treating time as continuous. Further, our analysis included all of the domains covered by the QLQ-C30 including the symptom scales. Our aim was two-fold: (1) to complement previously published clinical results on the impact of APAD on HRQoL; (2) to illustrate the advantages and disadvantages of approaches commonly used to analyse HRQoL data, and derive methodological recommendations.

Methods

The APAD1 trial

Study design and patients

APAD1 was a monocentric randomised controlled interventional trial in women aged 18–75 years old with histologically proven and newly (less than 6 months) diagnosed non-metastatic breast cancer who had undergone curative surgery and for whom adjuvant treatment was planned. Adjuvant therapy consisted of 6 cycles of chemotherapy—either 6 cycles of FEC100 (Fluorouracil + Epirubicin 100 mg/m^2 + Cyclophosphamide) or 3 cycles of FEC100 and 3 cycles of docetaxel—with one cycle given every 3 weeks, followed by 6 weeks of radiotherapy. Exclusion criteria included metastatic cancer or HER2 positive or any other primary tumour, contra-indications to physical activity, pregnancy or breast feeding and inability to attend intervention sessions or assessments. A total of 143 patients were randomly assigned to the UC control group ($N=71$) or to the APAD intervention group ($N=72$).

APAD intervention

Contrary to the UC group, the APAD group underwent an additional diet-exercise intervention throughout the treatment period from chemotherapy initiation until end of radiotherapy (approximately 6 months). The adapted physical activity consisted of 3 sessions per week of aerobic or muscle-strengthening exercises. The diet counselling consisted of nine dietitian consultations that included nutritional status assessment and nutrition advice. Further details of the APAD diet-exercise intervention are available in the published APAD1 protocol[8].

HRQoL assessment

HRQoL was assessed using the EORTC QLQ-C30 questionnaire[10], version 3.0. The QLQ-C30 is a 30-item self-administered cancer specific questionnaire composed of five functional scales (physical, role, cognitive, emotional and social), nine symptom scales (fatigue, nausea and vomiting, pain, dyspnoea, insomnia, appetite loss, constipation, diarrhoea, financial difficulties) and a global health status/HRQoL scale. Questionnaires were collected at baseline (T0), end of chemotherapy (T1, ~4 months), end of radiotherapy (T2, ~6 months), end of 6-months' follow-up (T3, ~12 months) and 1-year follow-up (T4, ~18 months).

For each of the 15 scales we calculated a standardised score ranging from 0 to 100 from the item responses, as recommended by the EORTC[11]. Higher scores for functional scales and the global health status/HRQoL scale correspond to a higher level of HRQoL, whereas a higher score for a symptom scale corresponds to a lower level of HRQoL.

Statistical analysis

We described QLQ-C30 completion—defined at each assessment time by the ratio of the number of fully or partially completed forms to the number of expected forms—and amount of missing score data per scale. The different methods for analysing the HRQoL score data were then applied separately for each scale (summarised in the sections below). Analyses were conducted on all available data without any missing data imputation. The alpha level for all tests was 0.05. The LMM analysis was performed using R software Version 4.0 and the other analyses using STATA software Version 13.0 and STATA commands qlqc30[12] and qlqc30_TTD[13].

Statistical testing

We used Student tests to compare the mean scores between the two groups at the different assessment times. We used paired Student tests to compare the mean scores between baseline and end of study (T4, ~18 months) in each group.

Linear mixed models analysis

LMMs were used to analyse the longitudinal HRQoL score data while taking into account the correlation of the measurements from the same patient. Specifically, we used random intercept and slope models which specified the score of patient i at time t as follows:

$$Y_i(t) = \beta_0 + \beta_1 t + \beta_2 (group \times t) + b_{0i} + b_{1i} t + \epsilon_i(t) \quad (1)$$

where t was treated as a continuous variable, the random intercept b_{0i} and random slope b_{1i} corresponded to the individual deviations from the fixed intercept β_0 (mean score at baseline) and fixed slope β_1 (mean change by unit of time in the UC group), respectively; β_2 was the fixed group-by-time interaction effect and $group$ was set to 1 if patient i belonged to the APAD group, or 0 otherwise. The error terms were assumed to be mutually independent with $\epsilon_i(t) \sim N(0, \sigma^2)$ and independent of the vector of random effects $\begin{pmatrix} b_{0i} \\ b_{1i} \end{pmatrix}$. The random effects were assumed to be correlated with $\begin{pmatrix} b_{0i} \\ b_{1i} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2_{b_0} & \sigma_{b_0 b_1} \\ \sigma_{b_0 b_1} & \sigma^2_{b_1} \end{pmatrix}\right)$.

Note that $\beta_1 + \beta_2$ corresponded to the slope characterising the mean change by unit of time in the APAD group, so β_2 characterised the difference of the mean score trajectory in the APAD group in comparison to the UC group. To test the significance of the time effect in the APAD group, $\beta_1 + \beta_2$, we also performed LMMs where the group factor coding was reversed.

In general, randomisation ensures that the score at baseline is not different between the groups and testing the difference in baseline score between groups should not be done in randomised clinical trials[14]. However, in such a modelling context, we believe that if the mean score is found to be significantly different between the groups, an additional fixed effect for the group could be considered in the above model in order to properly fit the data and obtain slope parameters that reflect the true evolution of the score in each group.

Time-to-deterioration analysis

We used a survival model to analyse the time to deterioration (TTD) of HRQoL[15–18]. We considered deterioration to be a score decrease for the functional scales and global health status/HRQoL scale (respectively, a score increase for the symptom scales) by more than X points compared to the baseline score, where X represents the minimal clinically important difference (MCID). As recommended by Osoba et al.[19], we considered three definitions: a 5-point MCID (little change), a 10-point MCID (moderate change), and a 20-point MCID (large change). As recommended by Bascoul-Mollevi et al.[20], we did not analyse the six symptom scales dyspnoea, appetite loss, insomnia, constipation, diarrhoea and financial difficulties (single item-based scales). Time to deterioration was defined as the time interval between the baseline assessment and the first assessment at which a deterioration was observed. For patients with no observed deterioration, time to deterioration was right-censored at the last assessment with available score data. No imputation of deterioration status was made in the case of intermittent missing data prior to time to deterioration or right-censoring time. The survival functions were estimated in each group using the Kaplan-Meier estimator and compared using the log-rank test. Hazard ratios (HRs) with 95% confidence intervals (CIs) were estimated using a Cox model with the UC group considered as the reference.

Results

QLQ-C30 completion

Supplementary Table S1 describes QLQ-C30 completion over time and available score data by patient. Out of the 143 expected forms, there were only one missing form at baseline; thereafter, completion decreased slightly over time, reaching 80% (UC group) and 93% (APAD group) at the last assessment. As a result of partially completed questionnaires, the number of patients with an available score for a given scale might slightly differ from the number of completed forms. Although missing score data increased over time and was higher in the UC group than the APAD group, it remained low in all scales throughout the study.

Supplementary Table S1 also shows that the median visit times for HRQoL assessment (3.9, 6.7, 12.6 and 19 months) were close to the theoretical ones (4, 6, 12 and 18 months). However, for a few patients, the assessment times were greatly delayed.

Statistical testing

In the statistical testing analysis, the patients were evaluable for a given scale and a given test of comparison if they had available scores at the considered timepoints.

HRQoL change between baseline and end of study

A complete description of the baseline QLQ-C30 scores can be found in Supplementary Table S2. Table 1 reports the comparison of the 15 HRQoL scores between baseline (T0) and end of study (T4). In the UC group, we detected a significant deterioration in global health status/HRQoL (− 7.75 points, $p = 0.028$), physical functioning (− 6.47 points, $p = 0.013$), cognitive functioning (− 8.77 points, $p = 0.004$), and dyspnoea (+ 9.94 points, $p = 0.018$). In contrast, in the APAD group, we detected a significant improvement in global health status/HRQoL (+ 5.6 points, $p = 0.029$), a tendency for improvement in physical functioning (+ 2.99 points), a non-significant deterioration in cognitive functioning (− 3.73 points), and a non-significant deterioration in dyspnoea (+ 4.97 points). In addition, there were significant improvements in role functioning (+ 9.45 points, $p < 0.001$) and financial difficulties (− 5.47 points, $p = 0.033$) in the APAD group, while in the UC group there was, respectively, no change and a tendency for deterioration (+ 3.58 points). Notably, social functioning showed tendencies for

	UC group ($n=71$)										APAD group ($n=72$)									
	T0			T4							T0			T4						
	N	Mean	(SD)	N	Mean	(SD)	Estimated mean difference	(SD)		*p*	N	Mean	(SD)	N	Mean	(SD)	Estimated mean difference	(SD)		*p*
Global health status/HRQoL	57	71.20	(17.11)	57	63.45	(23.13)	-7.75	3.81	↘	**0.028**	67	66.29	(18.37)	67	71.89	(18.28)	5.6	3.17	↗	**0.029**
Functional scales																				
Physical functioning	57	89.04	(11.33)	57	82.57	(19.34)	-6.47	2.97	↘	**0.013**	67	85.87	(13.85)	67	88.86	(12.47)	2.99	2.28	↗	0.135
Role functioning	57	85.38	(22.06)	57	85.38	(20.43)	0.00	3.98	→	1.000	67	81.59	(19.05)	67	91.04	(15.71)	9.45	3.02	↗	**0.000**
Emotional functioning	57	69.01	(20.40)	57	73.39	(22.79)	4.38	4.05	↗	0.168	67	67.33	(17.63)	67	77.57	(19.45)	10.24	3.21	↗	0.117
Cognitive functioning	57	87.13	(17.54)	57	78.36	(22.49)	-8.77	3.78	↘	**0.004**	67	87.06	(16.37)	67	83.33	(16.67)	-3.73	2.85	↘	0.075
Social functioning	56	85.71	(21.18)	56	78.57	(27.10)	-7.14	4.60	↘	0.088	67	84.08	(22.92)	67	89.30	(16.08)	5.22	3.42	↗	0.090
Symptom scales/items																				
Fatigue	57	28.85	(23.27)	57	33.53	(26.10)	4.68	4,63	↗	0.285	67	27.36	(17.23)	67	27.11	(18.82)	-0.25	3.12	↘	0.926
Nausea and vomiting	57	4.39	(13.19)	57	4.09	(10.58)	-0.3	2.24	↘	0.892	67	4.73	(14.45)	67	1.99	(7.41)	-2.74	1.98	↘	0.132
Pain	57	25.15	(25.42)	57	31.87	(33.52)	6.72	5.57	↗	0.186	67	27.86	(25.36)	67	24.38	(26.32)	-3.48	4.47	↘	0.407
Dyspnoea	57	11.11	(21.21)	57	21.05	(27.91)	9.94	4.64	↗	**0.018**	67	8.96	(17.00)	67	13.93	(19.38)	4.97	3.15	↗	0.068
Insomnia	57	35.09	(33.58)	57	37.43	(37.31)	2.34	6.65	↗	0.666	65	38.97	(33.63)	65	32.82	(27.32)	-6.15	5.37	↘	0.122
Appetite loss	57	7.60	(15.45)	57	7.02	(13.71)	-0.58	2.74	↘	0.799	66	8.59	(17.84)	66	7.58	(18.30)	-1.01	3.15	↘	0.621
Constipation	57	12.87	(28.00)	57	18.71	(30.22)	5.84	5.46	↗	0.192	67	12.44	(22.35)	67	14.93	(26.13)	2.49	4.20	↗	0.357
Diarrhoea	57	4.68	(14.69)	57	4.68	(13.27)	0.00	2.62	→	1.000	67	3.48	(13.15)	67	4.98	(15.63)	1.50	2.50	↗	0.552
Financial difficulties	56	10.71	(23.01)	56	14.29	(26.10)	3.58	4.65	↗	0.224	67	14.43	(22.64)	67	8.96	(22.16)	-5.47	3.87	↘	**0.033**

Table 1. Statistical testing: QLQ-C30 scores comparison between baseline (T0) and end of study (T4) in both groups. p-values and arrows indicating the direction (increase or decrease of the mean score) are bolded where there is a significant difference; N = number of patients without missing values; SD = standard deviation.

improvement in the APAD group (+5.22 points, $p = 0.090$) and deterioration in the UC group (− 7.14 points, $p = 0.088$).

Between-group HRQoL comparison during treatment and follow-up

Table 2 reports the comparison of the 15 HRQoL scores between the UC and APAD groups at the end of chemotherapy (T1) and at the end of study (T4). The results obtained at the end of radiotherapy (T2) and at 6-months' follow-up (T3) are reported in Supplementary Table S3. At the end of chemotherapy, the mean scores were significantly higher in the APAD group compared to the UC group for all functional scales. The differences exceeded 10 points for role functioning (+13.10 points) and emotional functioning (+10.05 points). In addition, the mean score in the fatigue symptom scale was significantly lower in the APAD group compared to the UC group (− 10.40 points, $p = 0.023$). At the end of the study, the between-group differences were still significant for physical functioning (+8.44 points in the APAD group, $p = 0.031$) and social functioning (+10.73 points in the APAD group, $p = 0.007$). There was also a significant difference in global health status/HRQoL (+6.29 points in the APAD group, $p = 0.025$).

Linear mixed model analysis

All patients ($N = 143$) were evaluable for the LMM analysis since they had at least one score available out of the five assessment times in all scales (see Supplementary Table S1).

Table 3 shows the results (estimates, 95% CI and Wald test p-values) concerning the time effect β_1 (slope in the UC group) and the group-by-time interaction effect β_2 for each of the 15 scales; the results concerning the time effect $\beta_1 + \beta_2$ (slope in the APAD group) are also given. The slope estimates $\widehat{\beta_1}$ and $\widehat{\beta_1 + \beta_2}$ represent a mean time effect by month over the entire period from T0 to T4 in the UC and APAD groups, respectively. We related these results to the comparison of the observed mean scores between T0 and T4 performed in the statistical testing analysis (results in Table 1). The LMM analysis agreed with the statistical testing analysis in finding a deterioration in the UC group for global health status/HRQoL ($\widehat{\beta_1} = -0.090$), physical functioning ($\widehat{\beta_1} = -0.122$), cognitive functioning ($\widehat{\beta_1} = -0.235$) and dysponea ($\widehat{\beta_1} = 0.261$). However, the deterioration was significant only for cognitive functioning ($p = 0.048$). In addition, the LMM analysis found a significant improvement in emotional functioning in the UC group, with a mean increase of 0.282 points per month ($p = 0.035$). For the APAD group, the LMM analysis also agreed with the statistical testing analysis, finding significant improvements in global health status/HRQoL ($\widehat{\beta_1 + \beta_2} = 0.432$, $p < 0.001$), role functioning ($\widehat{\beta_1 + \beta_2} = 0.521$, $p < 0.001$) and financial difficulties ($\widehat{\beta_1 + \beta_2} = 0.366$, $p = 0.005$). Notably, significant improvements were also found in physical functioning ($\widehat{\beta_1 + \beta_2} = 0.196$, $p = 0.039$), emotional functioning ($\widehat{\beta_1 + \beta_2} = 0.474$, $p < 0.001$), social functioning ($\widehat{\beta_1 + \beta_2} = 0.599$, $p < 0.001$), fatigue ($\widehat{\beta_1 + \beta_2} = -0.420$, $p = 0.004$), nausea and vomiting ($\widehat{\beta_1 + \beta_2} = -0.167$, $p = 0.048$), and appetite loss ($\widehat{\beta_1 + \beta_2} = -0.261$, $p = 0.037$).

The estimate of the group-by-time interaction effect $\widehat{\beta_2}$ can also be related to the between-group comparison of the observed mean scores in the statistical testing analysis. Instead of performing multiple comparisons at the different time points (results for T1 and T4 in Table 2), the LMMs provide a single value that summarises the between-group difference over the entire period from T0 to T4. A substantial result from the LMM analysis is that HRQoL over time was better in the APAD group compared with the UC group in all dimensions ($\widehat{\beta_2} > 0$ for the global health status/HRQoL scale and for the functional scales, and $\widehat{\beta_2} < 0$ for the symptom scales). This difference was significant for five scales: global health status/HRQoL ($\widehat{\beta_2} = 0.522$, $p = 0.001$), physical functioning ($\widehat{\beta_2} = 0.317$, $p = 0.013$), role functioning ($\widehat{\beta_2} = 0.315$, $p = 0.032$), social functioning ($\widehat{\beta_2} = 0.533$, $p = 0.006$) and financial difficulties ($\widehat{\beta_2} = -0.356$, $p<0.05$).

We noted that the LMM parameter estimates govern the predicted score trajectories in the two groups. As a consequence, the evolution of the HRQoL scores in the two groups and the between-group differences can also be appreciated by depicting the predicted mean trajectories (see Fig. 1 for the ten scales where a significant time effect and/or a significant group-by-time interaction effect were found).

Time-to-deterioration analysis

For the TTD analysis, 142 patients were evaluable, that is, had available scores at baseline. Two out of these 142 patients had no available later score information and were right censored just after baseline. The 10-point and 20-point MCID analyses results are reported in Table 4. The 5-point MCID analysis results were similar to that of the 10-point MCID analysis, with the same significant scales (data not shown).

When defining deterioration as a score decrease of at least 10 points from baseline, we found the relative risk of deterioration in the APAD group was approximately half that in the UC group for global health status/HRQoL (HR = 0.48, $p = 0.001$), physical functioning (HR = 0.54, $p = 0.011$) and role functioning (HR = 0.47, $p = 0.002$). When defining deterioration as a score decrease of at least 20 points from baseline, the group effect was still significant for global health status/HRQoL ($p < 0.001$) and role functioning ($p = 0.006$) with the relative risk of deterioration approximately half in the APAD group compared to the UC group (HR = 0.36 and HR = 0.40, respectively). Figure 2 shows Kaplan-Meier survival curves plotting the estimated probability of not having deterioration over time in both groups (for the 10-point and 20-point MCID definitions) for global health status/HRQoL, physical functioning, and role functioning. At 18 months (planned time for end of study), the probabilities of not having deterioration by at least 10 points were 0.60 (95% CI 0.47, 0.70) in the APAD group versus 0.30 (95% CI 0.20, 0.42) in the UC group for global health status/HRQoL, 0.64 (95% CI 0.51, 0.74) in the APAD group versus 0.45 (95% CI 0.33, 0.56) in the UC group for physical functioning, and 0.73 (95% CI 0.59, 0.83) in the APAD group versus 0.37 (95% CI 0.25, 0.50) in the UC group for role functioning. Figure 2 also shows that the main time point at which a difference is observed between the two groups is the first assessment after baseline (~4 months), which corresponds to the end of chemotherapy. This suggests that the APAD

	T1: End of chemotherapy									T4: End of study								
	UC GROUP ($n=71$)			APAD GROUP ($n=72$)				Estimated mean difference	(SD)	UC GROUP ($n=71$)			APAD GROUP ($n=72$)				Estimated mean difference	(SD)
	N	Mean	(SD)	N	Mean	(SD)	p			N	Mean	(SD)	N	Mean	(SD)	p		
Global health status/RQoL	64	55.86	(21.14)	71	61.85	(19.85)	0.092	5.99	3.54	57	82.57	(19.34)	67	88.86	(12.47)	**0.025**	6.29	2.98
Functional scales																		
Physical functioning	64	75.94	(20.94)	71	83.38	(15.11)	**0.018**	7.44	3.17	57	63.45	(23.13)	67	71.89	(18.28)	**0.031**	8.44	3.79
Role functioning	64	71.88	(28.15)	71	84.98	(16.94)	**0.001**	13.10	4.05	57	85.38	(20.43)	67	91.04	(15.71)	0.084	5.66	3.32
Emotional functioning	64	67.06	(23.54)	71	77.11	(20.54)	**0.009**	10.05	3.82	57	73.39	(22.79)	67	77.57	(19.45)	0.273	4.18	3.84
Cognitive functioning	64	69.27	(30.17)	71	79.11	(21.77)	**0.030**	9.84	4.57	57	78.36	(22.49)	67	83.33	(16.67)	0.161	4.97	3.61
Social functioning	64	62.24	(25.41)	71	70.89	(23.69)	**0.043**	8.65	4.24	56	78.57	(27.10)	67	89.30	(16.08)	**0.007**	10.73	4.12
Symptom scales/items																		
Fatigue	64	51.56	(28.51)	71	41.16	(24.01)	**0.023**	-10.40	4.56	57	33.53	(26.10)	67	27.11	(18.81)	0.116	-6.42	4.15
Nausea and vomiting	64	7.03	(12.17)	71	5.63	(15.16)	0.559	-1.40	2.36	57	4.09	(10.58)	67	1.99	(7.41)	0.197	-2.10	1.67
Pain	64	31.77	(29.65)	71	32.16	(27.36)	0.937	0.39	4.93	57	31.87	(33.52)	67	24.38	(26.32)	0.166	-7.49	5.48
Dyspnoea	64	29.17	(33.33)	71	24.88	(24.38)	0.392	-4.29	5.07	57	21.05	(27.91)	67	13.93	(19.37)	0.098	-7.12	4.39
Insomnia	64	40.63	(34.36)	71	40.38	(31.33)	0.965	-0.25	5.68	57	37.43	(37.30)	66	32.83	(27.11)	0.432	-4.60	5.96
Appetite loss	64	20.31	(29.47)	71	15.71	(25.20)	0.332	-4.60	4.74	57	7.02	(13.71)	66	7.58	(18.30)	0.850	0.56	2.89
Constipation	64	22.92	(33.53)	71	28.17	(31.70)	0.351	5.25	5.63	57	18.71	(30.22)	67	14.93	(26.13)	0.456	-3.78	5.12
Diarrhoea	64	15.62	(25.87)	71	10.33	(22.25)	0.203	-5.29	4.17	57	4.68	(13.27)	67	4.98	(15.63)	0.910	0.30	2.60
Financial difficulties	64	17.71	(28.46)	71	16.90	(25.74)	0.863	-0.81	4.69	56	14.29	(26.10)	67	8.96	(22.16)	0.223	-5.33	4.42

Table 2. Statistical testing: QLQ-C30 scores comparison between groups at end of chemotherapy (T1) and end of study (T4). p-values are bolded where there is a significant difference; N = number of patients without missing values; SD = standard deviation.

	Time effect (UC group)			Time effect (APAD group)			Group-by-time interaction effect		
	$\widehat{\beta_1}$	95% CI	*p*	$\widehat{(\beta_1 + \beta_2)}$	95% CI	*p*	$\widehat{\beta_2}$	95% CI	*p*
Global health status/HRQoL	− 0.090	[− 0.339; 0.158]	0.478	0.432	[0.196; 0.668]	<**0.001**	0.522	[0.208; 0.835]	**0.001**
Functional scales									
Physical functioning	− 0.122	[− 0.317; 0.074]	0.223	0.196	[0.010; 0.381]	**0.039**	0.317	[0.067; 0.568]	**0.013**
Role functioning	0.205	[− 0.036; 0.445]	0.096	0.521	[0.287; 0.755]	<**0.001**	0.315	[0.028; 0.602]	**0.032**
Emotional functioning	0.282	[0.021; 0.543]	**0.035**	0.474	[0.226; 0.722]	<**0.001**	0.192	[− 0.137; 0.521]	0.253
Cognitive functioning	− 0.235	[− 0.466; − 0.003]	**0.048**	− 0.045	[− 0.264; 0.173]	0.685	0.189	[− 0.113; 0.491]	0.220
Social functioning	0.066	[− 0.241; 0.373]	0.675	0.599	[0.310; 0.888]	<**0.001**	0.533	[0.156; 0.911]	**0.006**
Symptom scales/items									
Fatigue	− 0.073	[− 0.375; 0.228]	0.634	− 0.420	[− 0.706; − 0.134]	**0.004**	− 0.345	[− 0.720; 0.029]	0.072
Nausea and vomiting	− 0.113	[− 0.284; 0.059]	0.198	− 0.167	[− 0.333; − 0.002]	**0.048**	− 0.054	[− 0.245; 0.136]	0.576
Pain	0.187	[− 0.205; 0.579]	0.350	− 0.188	[− 0.561; 0.184]	0.321	− 0.376	[− 0.872; 0.120]	0.138
Dyspnoea	0.261	[− 0.043; 0.565]	0.093	− 0.048	[− 0.338; 0.242]	0.745	− 0.308	[− 0.702; 0.087]	0.127
Insomnia	0.062	[− 0.322; 0.447]	0.751	− 0.321	[− 0.687; 0.045]	0.086	− 0.383	[− 0.875; 0.109]	0.128
Appetite loss	− 0.184	[− 0.439; 0.071]	0.158	− 0.261	[− 0.505; − 0.016]	**0.037**	− 0.077	[− 0.389; 0.235]	0.630
Constipation	0.169	[− 0.161; 0.499]	0.316	0.003	[− 0.307; 0.313]	0.986	− 0.166	[− 0.604; 0.272]	0.458
Diarrhoea	− 0.075	[− 0.298; 0.147]	0.507	− 0.152	[− 0.366; 0.063]	0.165	− 0.076	[− 0.335; 0.182]	0.561
Financial difficulties	− 0.010	[− 0.283; 0.263]	0.942	− 0.366	[− 0.621; − 0.111]	**0.005**	− 0.356	[− 0.711; − 0.001]	<**0.05**

Table 3. Linear mixed models analysis: results on the 15 scales of the QLQ-C30 questionnaire. p-values are bolded where there is a significant coefficient; CI = confidence interval.

intervention would have prevented a substantial number of women from experiencing chemotherapy-related deterioration in global health status/HRQoL, physical functioning and role functioning.

Summary

The results of our analysis were consistent between the three different approaches, with all revealing a beneficial impact of the diet-exercise intervention. Table 5 summarises the interpretation of the different results and their statistical significance by scale. We found a beneficial effect of the APAD intervention in the following scales (by the number of significant results): global health status/HRQoL, physical functioning and role functioning ($n = 6$), social functioning ($n = 5$), financial difficulties ($n = 4$), fatigue and emotional functioning ($n = 3$), cognitive functioning, nausea/vomiting and appetite loss ($n = 2$), and dyspnoea ($n = 1$).

Methodological comments

The three approaches considered to analyse the HRQoL data from the APAD1 study serve different purposes in addressing the research questions. Therefore, each of them could be preferred over the others depending on the purpose, but also on parameters such as data features, study design and methodological considerations. In this section, we would like to highlight some of their pitfalls in order to help choose the most appropriate one for a given situation or to avoid misinterpreting their results.

LMM approach and statistical testing analysis

Readability and multiplicity testing

The pairwise comparisons performed in the statistical testing analysis allow us to compare the observed mean scores at specific time points. However, the objective of the APAD1 trial was to evaluate the impact of the APAD intervention over the whole study period, specifically: the immediate impact during the intervention itself (2 planned visits for HRQoL assessment during treatment period), and the long-term impact after the intervention (2 planned visits for HRQoL assessment during follow-up). In this case, to avoid multiple testing and improve readability, a modelling approach for the longitudinal HRQoL data is preferable to pairwise comparisons. LMMs allow modelling of the HRQoL score trajectories in the two groups, and provide simplified and easy-to-interpret information. In the applied random coefficient models, the between-group difference was summarised over the whole period by the single estimate $\widehat{\beta}_2$, and the HRQoL evolution over time was governed by the single slope estimates $\widehat{\beta}_1$ (UC group) and $\widehat{\beta_1 + \beta_2}$ (APAD group). Further, the LMM estimates allow us to plot the predicted mean score trajectories, giving good insight at-a-glance into the HRQoL score evolution in each group.

Measurement error

An important advantage of the LMM analysis over the statistical testing analysis stems from its modelling approach. In a statistical testing analysis, the tests are performed directly on the *observed* scores. By contrast, the LMM analysis uses the whole population variability to provide predicted HRQoL score values for each patient, that is, estimates of the "true" HRQoL score values. Indeed, the observed score value of each patient is decomposed into the sum of a true unobserved score value and a measurement error. Thus, the predicted HRQoL score values correspond to the observed values after eliminating the measurement error.

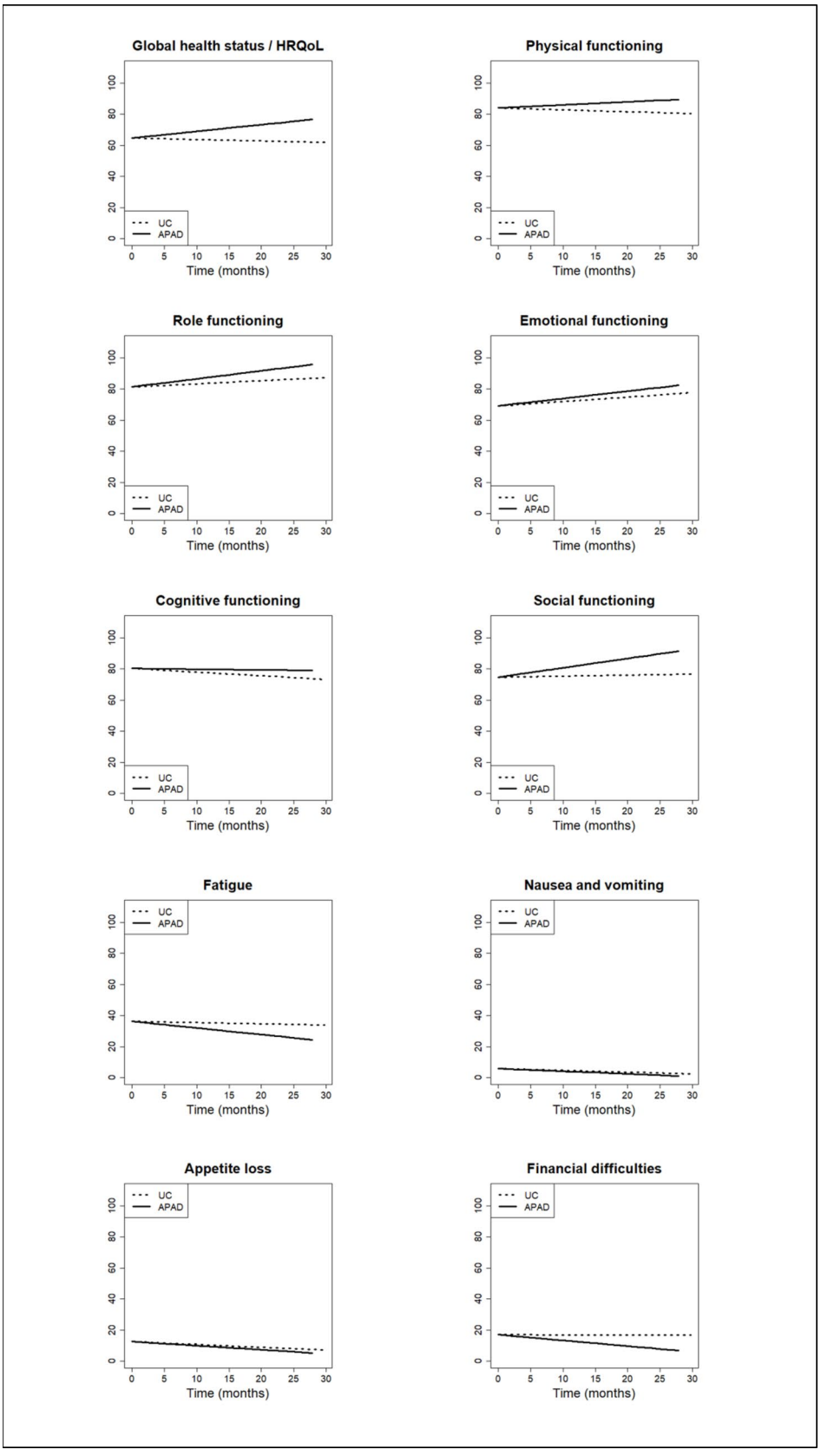

Fig. 1. Predicted mean score trajectories from the LMMs for the ten scales with significant time and/or group-by-time interaction effect. The predicted values at each time t are given by Eq. (1) where the random (thus, individual) part has been removed and the fixed effect parameters replaced by their estimates, that is, given by: $\widehat{Y}_{group=1}(t) = \widehat{\beta}_0 + \left(\widehat{\beta_1 + \beta_2}\right) t$ in the APAD group and $\widehat{Y}_{group=0}(t) = \widehat{\beta}_0 + \widehat{\beta}_1 t$ in the UC group.

	10-point MCID					20-point MCID					Minimal theoretical difference between two score values
	Number of events		Estimates			Number of events		Estimates			
	UC group (n = 71)	APAD group (n = 72)	HR	[95% CI]	*p*	UC group (n = 71)	APAD group (n = 72)	HR	[95% CI]	*p*	
Global Health Status/ HRQoL	48	32	0.48	[0.30–0.75]	**0.001**	36	17	0.36	[0.20–0.64]	< **0.001**	8.33
Functional scales											
Physical functioning	41	29	0.54	[0.34–0.88]	**0.011**	26	18	0.57	[0.31–1.04]	0.062	6.67
Role functioning	42	27	0.47	[0.29–0.77]	**0.002**	24	12	0.40	[0.2–0.79]	**0.006**	16.67
Emotional functioning	31	25	0.63	[0.37–1.07]	0.087	21	19	0.77	[0.41–1.43]	0.400	
Cognitive functioning	50	47	0.82	[0.55–1.22]	0.320	30	24	0.67	[0.39–1.15]	0.140	16.67
Social functioning	51	55	1.08	[0.74–1.58]	0.698	41	37	0.79	[0.50–1.23]	0.291	16.67
Symptom scales											
Fatigue	56	51	0.76	[0.52–1.11]	0.145	46	37	0.69	[0.45–1.07]	0.096	11.11
Nausea and vomiting	23	24	0.96	[0.54–1.70]	0.887	9	7	0.70	[0.26–1.88]	0.479	16.67
Pain	44	46	1.05	[0.69–1.58]	0.830	34	32	0.89	[0.55–1.44]	0.634	16.67

Table 4. TTD analysis: results of the analyses using a 10-point or 20-point MCID to define the deterioration event. p-values are bolded where there is a significant HR; HR = hazard ratio, CI = confidence interval.

TTD approach

Clinical interpretability

In the TTD approach, we are not modelling HRQoL score over time but time until HRQoL deterioration. This involves a classical survival analysis which has the advantage of producing outputs—namely, HRs and survival curves—that clinicians are familiar with. Another advantage of this approach is that the between-group comparison is performed using a deterioration definition in accordance with a score change that is clinically meaningful, although a potential limitation arises from controversy in the literature around the appropriate MCID to choose. However, the TTD approach has more serious issues than deciding whether a MCID of 5, 10 or 20 is the most appropriate for a given scale.

Measurement error

The definition of the deterioration event is derived from the observed scores; as a consequence, the HRQoL values are assumed to be observed without measurement error.

Cautiousness when interpreting the results

It is important to keep in mind that the number of possible score values is limited (4, 7, 10, 13, 16 for the scales of the QLQ-C30), so that the interval between two score values does not correspond with the usual MCID values. When applying the TTD approach to a particular scale, the smaller the number of possible values, the more misleading the interpretation of the results will be. For this reason, we chose not to apply this approach to the single item-based scales, where the score takes only values of 0, 33.33, 66.67, or 100. Given that a change of at least 33.33 points is needed to observe a deterioration in these scales, a TTD analysis using a MCID of 5, 10, or 20 would provide exactly the same results. Thus, in order to correctly interpret the results of a TTD analysis, it is important to consider the minimal difference that can be effectively observed for a given scale, in addition to the theoretical difference used as MCID.

We must also consider that the deterioration corresponds to a *minimal* change of score; consequently, two patients who deteriorate beyond the MCID point would be treated equally regardless of the size of the change. Similarly, a patient who improves is treated the same as a patient who deteriorates less than the MCID point (no event in both cases).

Observation in discrete time

Another issue is that the TTD approach uses a standard survival model, while the HRQoL was assessed only at study visits; therefore, the deterioration status was known only at these specific time points. Using a TTD approach implicitly assumes that the deterioration can occur only at the visit times and assumes that the deterioration status is unchanged between two visits for HRQoL assessment. This simplification is a source of bias if the timings of planned visits are different for each group, though this is generally not the case in clinical trials. Indeed, in such a case, more events of deterioration or shorter times to deterioration would be observed in one group compared to the other, simply because HRQoL would be assessed more frequently or at shifted time points. This simplification is also a source of bias where there is intermittent missing data (frequent in patient-reported outcomes), since patients without a score at a certain visit will be considered as still not deteriorated (if deterioration has not previously been observed). When the interval with missing data is large, consider right censoring the time to deterioration rather than assuming non-deterioration during a long period, even if a deterioration has been observed afterwards. Note that missing data will be particularly problematic if their amount varies according to the group, which can happen, for example, when follow-up care is better, or when toxicities resulting in missingness are more frequent, in one group.

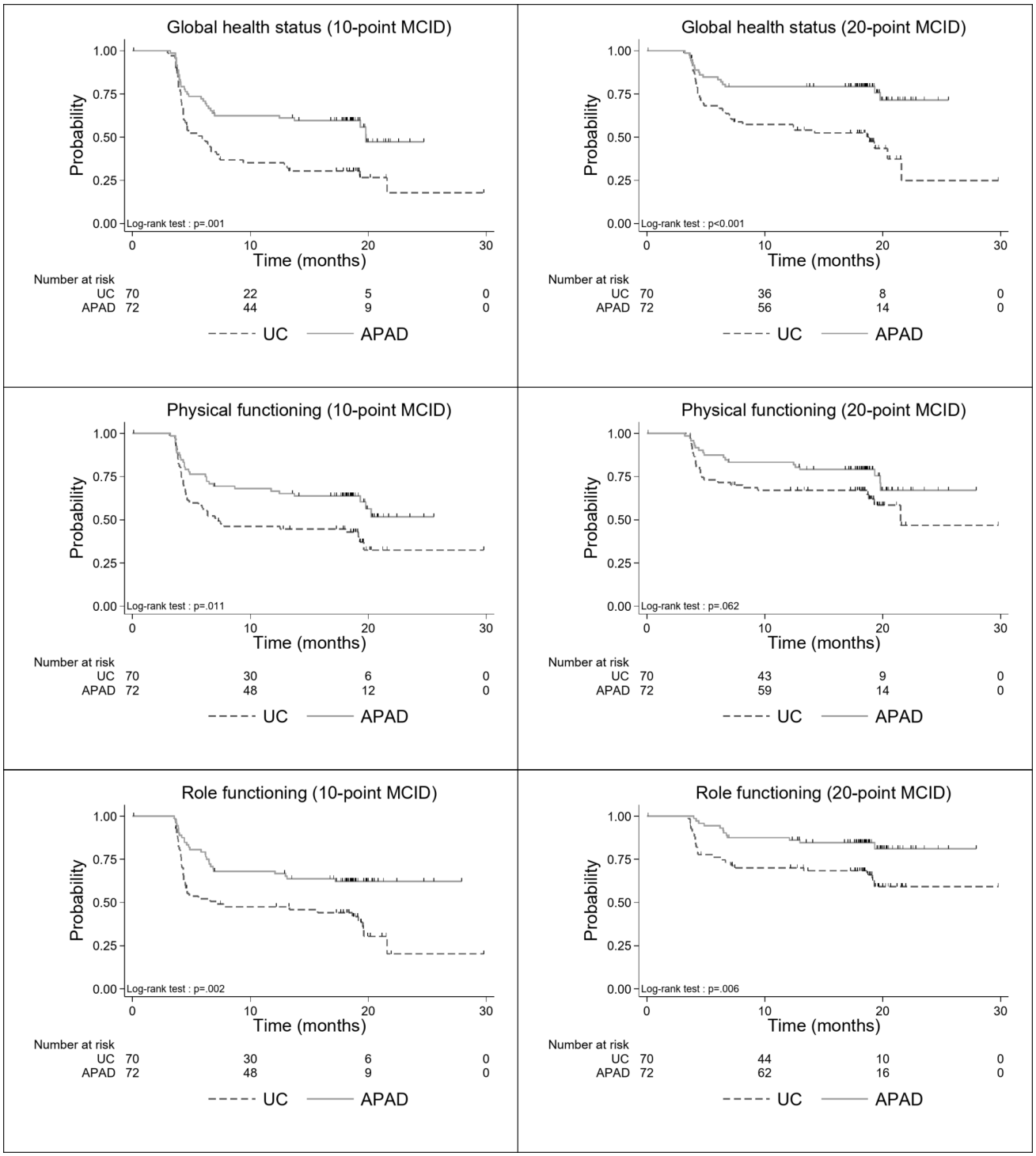

Fig. 2. TTD analysis in the UC and APAD groups for the scales where the survival estimates were significantly different for at least one definition of MCID (from top to bottom: global health status/HRQoL, physical functioning, role functioning): Kaplan-Meier curves of the 10-point MCID (left-hand side) and the 20-point MCID (right-hand side). Vertical lines on the curves represent right-censored times.

Discussion

The analyses performed in this article confirm the initial results already published, but also provide additional results, all of which support the beneficial effect of the diet-exercise intervention on HRQoL. Although a standard EORTC questionnaire was used with high completion rates, completion rates were slightly lower in the UC group. Furthermore, the degree of generalisability is limited by the fact that the study was monocentric (all patients were recruited at the Cancer Institute of Montpellier). The APAD programme is no longer applied

	Number of significant results	Comparison of the observed mean score between T0 and T4 (Table 1)		Evolution of the predicted mean score trajectory over the whole period (Table 3)		Increase of the observed mean score in the APAD group compared with the UC group (Table 2)		Difference of the predicted mean score trajectories of the two groups in favour of the APAD group (Table 3)	Decrease of the risk of deterioration in the APAD group compared with the UC group (Table 5)	
		Improvement in the APAD group	Deterioration in the UC group	Improvement in the APAD group	Deterioration in the UC group	At T1	At T4		MCID = 10 points	MCID = 20 points
QL	7	×	×	×			×	×	×	×
PF	6		×	×		×	×	×	×	
RF	6	×		×		×		×	×	×
SF	5			×	×	×	×	×		
FI	4	×		×	×			×	NA	NA
FA	3			×	×	×				
EF	3			×	×	×				
CF	2		×			×				
NV	2			×	×					
AP	2			×	×				NA	NA
DY	1		×						NA	NA
PA	0									
SL	0								NA	NA
CO	0								NA	NA
DI	0								NA	NA

Table 5. Summary of the results obtained from the three approaches (statistical testing, LMMs, and TTD analyses). QL = Global health status/HRQoL, PF = Physical functioning, RF = Role functioning, SF = Social functioning, FI = Financial difficulties, FA = Fatigue, EF = Emotional functioning, CF = Cognitive functioning, NV = Nausea and vomiting, AP = Appetite loss, DY = Dyspnoea, PA = Pain, SL = Insomnia, CO = Constipation, DI = Diarrhoea.

as such at the Cancer Institute of Montpellier, but physical activity and diet counselling are included in the supportive care programme for breast cancer patients during the recovery period. The latter includes 5 collective workshops, 4 group activities and individual follow-up sessions. This article suggests that it might be interesting to consider offering supportive care earlier, during the treatment phase. In this case, particular attention should be paid to balancing the intervention with the burden of adjuvant therapy for each patient. It should be noted that all French comprehensive cancer centres have a supportive care unit, whose services are fully covered by the health system and are never compulsory.

Using data from the APAD1 study, this article also illustrates the strengths and limitations of three main approaches used to analyse longitudinal HRQoL score data, namely: statistical testing, LMMs, and TTD analysis. These approaches may be complementary to capture the complexity of HRQoL and its measurement, or more or less appropriate depending on their intrinsic characteristics, the research question, data features and study design. For example, in a study with only one post-baseline HRQoL assessment, a statistical testing approach would be entirely appropriate. In order to choose the best approach for a given situation and to apply it in the most appropriate way, the reader can refer to the recommendations of the SISAQOL consortium[21]. Note that for the sake of simplicity, we have not adjusted for covariates, including baseline score, in the work presented in this article, although this is part of the SISAQOL recommendations. However, this was done in the previous publication on the APAD data with similar results.

In the LMM analysis, we used random coefficient models which provided easy-to-interpret parameters but involved the assumption of score trajectories that are linear over time. They may be insufficient to capture more complex forms of trajectories. For example, in scales closely related to the chemotherapy treatment, such as fatigue, we initially observed a deterioration between baseline and end of chemotherapy, followed by an improvement. This explains why the statistical testing analysis highlighted more between-group differences at the end of the chemotherapy than at the end of the study. By smoothing the actual trajectory, the LMM was able to provide a unique value that summarised how much this symptom decreased by some unit of time on average over the entire period, but missed changes in magnitude or direction (rapid increase then slow decrease). In such cases, LMMs that allow for flexible trajectories, such as models based on polynomials or splines, could be considered, as described in Winter et al.[22]. The estimated coefficients would have no direct interpretation, so it would be needless to report them. The outputs of interest would be the predicted trajectories and the result of the (likelihood-ratio) test to know whether the predicted curves are significantly different between the two groups. The previous LMM analysis performed by Carayol et al.[9] had avoided this linear assumption by treating time as a categorical variable. Such modelling involves many estimates to report, so suffers from similar readability issues as the statistical testing. Principally, the results cannot be interpreted and graphically depicted in terms of trajectories. Another class of models for longitudinal data, often referred to as covariance pattern models, could also be used. Rather than including random effects in the model, these models assume a covariance pattern defined within the residual matrix to account for the correlation of the measures repeated on the same patient. However, they also consider time as a categorical variable, are quite expensive regarding the number of parameters to estimate when there are 3 or 4 measurement times, and imply choosing a covariance structure adapted to the longitudinal design features.

Conclusion

In this article, we performed an in-depth investigation of the effect of the APAD intervention on all the dimensions of the QLQ-C30 questionnaire using three different approaches that all revealed a beneficial impact of the diet-exercise intervention on HRQoL. From our analysis, we also highlighted some methodological issues of the different approaches for analysis, and derived recommendations on their use. We argue that, with more than two time points of HRQoL assessment/interest, an approach for longitudinal data such as an LMM analysis is preferable/complementary to statistical testing for describing the evolution of HRQoL over time and comparing this between treatment groups. Indeed, an LMM analysis avoids multiple testing, takes into account the measurement error and, if time is treated as continuous, provides nice graphical outputs of the predicted mean trajectories. We revealed some limitations of the TTD approach. Despite being an appealing approach for clinicians by providing the familiar outputs of survival analysis, we argued that the readability of the results is only apparent. Indeed, the results are interpreted in the terms of a MCID, which in fact cannot be observed due to the ordinal nature of the HRQoL scores. We thus recommend using the TTD approach only with scales for which the score takes a reasonable number of values, and reporting the minimal difference that can be effectively observed by scale. Finally, TTD results can hide biases because deterioration is observed in discrete time at the study visits, this issue being worsened with missingness. We thus suggest using the TTD approach only when there are a reasonable number of timepoints and available HRQoL data, similarly distributed in the two treatment groups.

Data availability

The APAD1 trial data are not publicly available due to confidentiality requirements. The datasets used during the current study are however available from the corresponding author on reasonable request.

Received: 26 October 2024; Accepted: 21 February 2025
Published online: 10 March 2025

References

1. Bardwell, W. A. & Ancoli-Israel, S. Breast Cancer and fatigue. *Sleep. Med. Clin.* **3**, 61–71. https://doi.org/10.1016/j.jsmc.2007.10.011 (2008).

2. Hagen, K. B. et al. Fatigue, anxiety and depression overrule the role of oncological treatment in predicting self-reported health complaints in women with breast cancer compared to healthy controls. *Breast* **28**, 100–106. https://doi.org/10.1016/j.breast.2016.05.005 (2016).
3. Gupta, D., Lis, C. G. & Grutsch, J. F. The relationship between cancer-related fatigue and patient satisfaction with quality of life in cancer. *J. Pain Symptom Manag.* **34**, 40–47. https://doi.org/10.1016/j.jpainsymman.2006.10.012 (2007).
4. Stasi, R., Abriani, L., Beccaglia, P., Terzoli, E. & Amadori, S. Cancer-related fatigue: evolving concepts in evaluation and treatment. *Cancer* **98**, 1786–1801. https://doi.org/10.1002/cncr.11742 (2003).
5. Branstrom, R., Petersson, L. M., Saboonchi, F., Wennman-Larsen, A. & Alexanderson, K. Physical activity following a breast cancer diagnosis: implications for self-rated health and cancer-related symptoms. *Eur. J. Oncol. Nurs.* **19**, 680–685. https://doi.org/10.1016/j.ejon.2015.04.008 (2015).
6. van Vulpen, J. K., Peeters, P. H., Velthuis, M. J., van der Wall, E. & May, A. M. Effects of physical exercise during adjuvant breast cancer treatment on physical and psychosocial dimensions of cancer-related fatigue: A meta-analysis. *Maturitas* **85**, 104–111. https://doi.org/10.1016/j.maturitas.2015.12.007 (2016).
7. Canario, A. C. et al. Physical activity, fatigue and quality of life in breast cancer patients. *Rev. Assoc. Med. Bras. (1992).* **62**, 38–44. https://doi.org/10.1590/1806-9282.62.01.38 (2016).
8. Carayol, M. et al. Adapted physical activity and diet (APAD) during adjuvant breast cancer therapy: design and implementation of a prospective randomized controlled trial. *Contemp. Clin. Trials.* **36**, 531–543. https://doi.org/10.1016/j.cct.2013.09.016 (2013).
9. Carayol, M. et al. Short- and long-term impact of adapted physical activity and diet counseling during adjuvant breast cancer therapy: the APAD1 randomized controlled trial. *BMC Cancer*. **19**, 737. https://doi.org/10.1186/s12885-019-5896-6 (2019).
10. Aaronson, N. K. et al. The European organization for research and treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. *J. Natl. Cancer Inst.* **85**, 365–376. https://doi.org/10.1093/jnci/85.5.365 (1993).
11. Fayers, P. et al. EORTC QLQ-C30 Scoring Manual. (2001).
12. Bascoul-Mollevi, C., Castan, F., Azria, D. & Gourgou, S. EORTC QLQ-C30 descriptive analysis with the qlqc30 command. *Stata J.* **15**, 1060–1074 (2015).
13. Bascoul-Mollevi, C. et al. Longitudinal health-related quality of life analysis in oncology with time to event approaches, the STATA command qlqc30_TTD. *Comput. Methods Progr. Biomed.* **158**, 153–159. https://doi.org/10.1016/j.cmpb.2018.02.010 (2018).
14. de Boer, M. R., Waterlander, W. E., Kuijper, L. D., Steenhuis, I. H. & Twisk, J. W. Testing for baseline differences in randomized controlled trials: an unhealthy research behavior that is hard to eradicate. *Int. J. Behav. Nutr. Phys. Act.* **12**, 4. https://doi.org/10.1186/s12966-015-0162-z (2015).
15. Anota, A. et al. Time to health-related quality of life score deterioration as a modality of longitudinal analysis for health-related quality of life studies in oncology: do we need RECIST for quality of life to achieve standardization? *Qual. Life Res. Int. J. Qual. Life Aspects Treat. Care Rehabil.*. **24**, 5–18. https://doi.org/10.1007/s11136-013-0583-6 (2015).
16. Bonnetain, F. et al. Time until definitive quality of life score deterioration as a means of longitudinal analysis for treatment trials in patients with metastatic pancreatic adenocarcinoma. *Eur. J. Cancer (Oxford, England.* **46**, 2753–2762 (1990). https://doi.org/10.1016/j.ejca.2010.07.023 (2010).
17. Hamidou, Z. et al. Time to deterioration in quality of life score as a modality of longitudinal analysis in patients with breast cancer. *Oncologist* **16**, 1458–1468. https://doi.org/10.1634/theoncologist.2011-0085 (2011).
18. Jaeschke, R., Singer, J. & Guyatt, G. H. Measurement of health status. Ascertaining the minimal clinically important difference. *Control Clin. Trials.* **10**, 407–415. https://doi.org/10.1016/0197-2456(89)90005-6 (1989).
19. Osoba, D., Rodrigues, G., Myles, J., Zee, B. & Pater, J. Interpreting the significance of changes in health-related quality-of-life scores. *J. Clin. Oncol. J. Am. Soc. Clin. Oncol.* **16**, 139–144. https://doi.org/10.1200/JCO.1998.16.1.139 (1998).
20. Bascoul-Mollevi, C. et al. Longitudinal analysis of health-related quality of life in cancer clinical trials: methods and interpretation of results. *Qual. Life Res. Int. J. Qual. Life Aspects Treat. Care Rehabil.*. **30**, 91–103. https://doi.org/10.1007/s11136-020-02605-3 (2021).
21. Coens, C. et al. International standards for the analysis of quality-of-life and patient-reported outcome endpoints in cancer randomised controlled trials: recommendations of the SISAQOL consortium. *Lancet Oncol.* **21**, e83–e96. https://doi.org/10.1016/S1470-2045(19)30790-9 (2020).
22. Winter, A. et al. Flexible modeling of longitudinal health-related quality of life data accounting for informative dropout in a cancer clinical trial. *Qual. Life Res. Int. J. Qual. Life Aspects Treat. Care Rehabil.*. **32**, 669–679. https://doi.org/10.1007/s11136-022-03252-6 (2023).

Acknowledgements

Editorial assistance for English language improvement was provided by Angloscribe, Calvisson, France.

Author contributions

C.T. and C.M. performed statistical analysis and interpretation of the results. C.M. supervised the statistical analysis. C.T. performed the methodological review and wrote the Manuscript. C.M. critically commented the manuscript. All authors read and approval the final manuscript.

Funding

This work was supported by the "SIRIC Montpellier Cancer" (Grants INCa_Inserm_DGOS_12553 and INCa-DGOS-INSERM-ITMO Cancer_18004).

Declarations

Ethics approval and consent to participate

The APAD1 study received approval of all French institutional review boards. All participants provided written informed.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-91569-w.

Correspondence and requests for materials should be addressed to C.T.