Table 1: Baseline characteristics by treatment arm in the ITT population

Characteristic	Randomised groups	
	tCBT (n=500)	Usual care (n=496)
	Median (IQR)	*Median (IQR)*
Age (years)	58.8 (47.7 – 68.7)	59.5 (47.9 – 68.9)
	N (%)	*N (%)*
Gender		
Male	209 (41.8)	204 (41.0)
Female	291 (58.2)	292 (58.9)
Employment status		
Working (full or part-time)	277 (55.4)	244 (49.2)
Unable to work because of health	18 (3.6)	30 (6.0)
Retired	168 (33.6)	177 (35.7)
Other	37 (7.4)	45 (9.1)
CWP risk profile:		
Illness behaviour score > 4		
No	1 (0.2)	2 (0.5)
Yes	498 (99.6)	494 (99.5)
*Not known**	1 (0.2)	0 (0.0)
Somatic Symptoms Scale score > 2		
No	472 (94.4)	462 (93.1)
Yes	28 (5.6)	34 (6.9)
Sleep problems score > 4		
No	1 (0.2)	2 (0.4)
Yes	499 (99.8)	493 (99.4)
*Not known**	0 (0.0)	1 (0.2)
CWP risk profile factors present (N)		
2	474 (94.8)	466 (94.0)
3	26 (5.2)	30 (6.0)
	*Median (IQR) [n**]*	*Median (IQR) [n]*
Psychological distress (GHQ)	1 (0 – 4) [499]	1 (0 – 4) [494]
Quality of Life (EQ-5D-5L utility score)	0.74 (0.65 – 0.80) [499]	0.74 (0.64 – 0.80) [496]
ICECAP-A	0.91 (0.81 – 0.95) [495]	0.90 (0.79 – 0.95) [491]
Fibromyalgia Research Criteria		
WPI	3 (1 - 4) [499]	2 (1 – 4) [492]
SSS	4 (3 – 6) [497]	4 (3 – 5) [494]

EQ-5D-5L: Euroqol questionnaire – five dimensions – five levels; FRC: fibromyalgia research criteria; GHQ: general health questionnaire; ICECAP-A: ICEpop CAPability measure for Adults; IQR: interquartile range; SD: standard deviation; SSS: symptom severity scale; tCBT: telephone cognitive behavioural therapy; WPI: widespread pain index.

* Where individuals completed half or fewer items, the score was classified as not known, but individuals could still be eligible for recruitment based on their responses to other items answered.
** The number of persons for whom a scale score could be calculated

Table 2: Outcomes by treatment arm at 12 months (Intention to treat analysis)

Characteristic	Randomised groups	
	tCBT (n=500)	Usual care (n=496)
Primary Outcome	*N (%)*	*N (%)*
Chronic widespread pain		
No	315 (82.0)	364 (82.5)
Yes	69 (18.0)	77 (17.5)
Secondary Outcome		
Global impression of change		
Very much better	24 (6.5)	15 (3.5)
Much better	88 (23.7)	59 (13.8)
A little better	90 (24.3)	84 (19.6)
No change	83 (22.4)	126 (29.4)
A little worse	65 (17.5)	119 (27.7)
Much worse	18 (4.9)	23 (5.4)
Very much worse	3 (0.8)	3 (0.7)
Pain Reported		
No	79 (20.6)	68 (15.4)
Yes	305 (79.4)	373 (84.6)
CWP risk profile		
Somatic symptoms score		
0	210 (56.5)	228 (52.8)
1	103 (27.7)	123 (28.5)
2-5	59 (15.9)	81 (18.8)
Illness behaviour score *Mean (SD)[n*]*	8.21 (4.04) [371]	8.96 (4.19) [431]
Sleep problems score *Mean (SD)[n]*	8.20 (4.89) [373]	9.20 (5.16) [432]
Psychological Distress (GHQ score)		
0	201 (54.5)	202 (46.8)
1	59 (16.0)	54 (12.5)
2-5	68 (18.4)	113 (26.2)
6-12	41 (11.1)	63 (14.6)
	Mean (SD) [n]	*Mean (SD) [n]*
Chalder Fatigue Score	12.6 (4.5) [370]	13.6 (4.4) [433]
	Median (IQR) [n]	*Median (IQR) [n]*
Quality of Life (EQ-5D utility score)	0.74 (0.66 – 0.84) [371]	0.74 (0.65 – 0.82) [435]
ICECAP-A	0.91 (0.82 – 0.97) [368]	0.89 (0.78 – 0.95) [429]
Fibromyalgia Research Criteria		
WPI	2 (1 - 4) [366]	2 (1 – 4) [427]
SSS	3 (2 – 5) [369]	4 (2 – 5) [431]

EQ-5D-5L: Euroqol questionnaire – five dimensions – five levels; FRC: fibromyalgia research criteria; GHQ: general health questionnaire; ICECAP-A: ICEpop CAPability measure for Adults; IQR: interquartile range; SD: standard deviation; SSS: symptom severity scale; tCBT: telephone cognitive behavioural therapy; WPI: widespread pain index. * The number of persons for whom a scale score could be calculated

Table 3: Summary of the primary and secondary outcomes across follow-up points[1]

Outcome		Time point (months)	Analysis method (effect size)	Adjusted* effect size (95% CI)	p-value	Unadjusted effect size (95% CI)	p-value
Primary Outcome							
CWP		3	Logistic regression (OR)	1.08 (0.74, 1.58)	0.691	1.07 (0.75, 1.53)	0.716
CWP (per protocol)				1.15 (0.75, 1.75)	0.519	1.18 (0.77, 1.66)	0.522
CWP (with multiple imputation)				1.06 (0.74, 1.54)	0.749	1.05 (0.72, 1.53)	0.816
CWP**		**12**		**1.05 (0.75, 1.48)**	**0.771**	**1.04 (0.72, 1.48)**	**0.849**
CWP (per protocol)				1.11 (0.81, 1.50)	0.519	1.09 (0.74, 1.60)	0.673
CWP (with multiple imputation)				1.04 (0.75, 1.45)	0.982	1.03 (0.74, 1.42)	0.964
CWP		24		0.85 (0.68, 1.07)	0.163	0.84 (0.61, 1.18)	0.317
CWP (per protocol)				0.85 (0.64, 1.12)	0.241	0.84 (0.58, 1.20)	0.330
CWP (with multiple imputation)				0.85 (0.66, 1.09)	0.220	0.84 (0.65, 1.09)	0.196
CWP		3,12,24	GEE (OR)	1.00 (0.96, 1.04)	0.923	1.00 (0.96, 1.04)	0.835
Secondary Outcomes							
Global impression of change[2]		3	Ordinal Logistic Regression (OR)	0.42 (0.32, 0.55)	<0.001	0.43 (0.34, 0.56)	<0.001
		12		0.51 (0.39, 0.67)	<0.001	0.53 (0.41, 0.68)	<0.001
		24		0.55 (0.43, 0.70)	<0.001	0.58 (0.45, 0.73)	<0.001
CWP risk profile	Somatic symptom score	3	Ordinal logistic regression (OR)	0.79 (0.60, 1.03)	0.084	0.83 (0.64, 1.08)	0.173
		12		0.86 (0.71, 1.04)	0.112	0.85 (0.65, 1.11)	0.237
		24		0.81 (0.59, 1.12)	0.206	0.90 (0.67, 1.21)	0.498
	Illness behaviour score	3	Linear regression (Mean difference)	-0.17 (-0.58, 0.24)	0.385	-0.25 (-0.79, 0.29)	0.360
		12		-0.81 (-1.54, -0.09)	0.030	-0.74 (-1.32, -0.17)	0.011
		24		-1.25 (-2.15, -0.35)	0.010	-1.20 (-1.83, -0.58)	<0.001
	Sleep problems score	3		-0.62 (-1.26, 0.02)	0.057	-0.62 (-1.31, 0.08)	0.081

[1] Analyses shaded in grey favour tCBT over usual care at pre-specified significance level for secondary outcomes (p<0.01). Except for EQ-5D-5L, mean differences less than zero and odds ratios less than one favour the treatment group.

[2] OR of one point increase in global impression of change score (worsening of health)

		Time	Method	Estimate (95% CI)	p	Estimate (95% CI)	p
		12	Linear regression	-0.95 (-1.48, -0.42)	0.002	-1.00 (-1.70, -0.30)	0.005
		24	(Mean difference)	-0.51 (-1.25, 0.23)	0.161	-0.52 (-1.39, 0.16)	0.117
Psychological distress (GHQ)		3	Ordinal logistic	0.55 (0.43, 0.69)	<0.001	0.58 (0.45, 0.76)	<0.001
		12	regression (OR)	0.65 (0.50, 0.86)	0.002	0.70 (0.54, 0.90)	0.007
		24		0.76 (0.60, 0.96)	0.024	0.74 (0.56, 0.98)	0.037
Chalder fatigue score		3	Linear Regression	-1.36 (-2.10, -0.64)	0.001	-1.40 (-1.97, -0.82)	<0.001
		12	(Mean difference)	-1.02 (-1.63, -0.42)	0.003	-1.03 (-1.64, -0.42)	0.001
		24		-0.93 (-1.62, -0.23)	0.012	-0.93 (-1.58, -0.27)	0.006
Quality of Life (EQ-5D-5L utility score)		3	Linear Regression	0.009 (-0.009, 0.028)	0.304	0.021 (-0.004, 0.046)	0.101
		12	(Mean difference)	0.024 (0.009, 0.040)	0.004	0.037 (0.010, 0.064)	0.007
		24		0.030 (0.009, 0.050)	0.008	0.040 (0.011, 0.069)	0.007
ICECAP-A tariff		3	Ordinal logistic	1.14 (0.89, 1.48)	0.304	1.17 (0.86, 1.59)	0.323
		12	regression (OR)	1.39 (0.94, 2.04)	0.096	1.39 (1.01, 1.91)	0.042
		24		0.88 (0.67, 1.15)	0.338	0.99 (0.70, 1.41)	0.966
Fibromyalgia research criteria	Widespread Pain Index	3	Poisson	0.98 (0.90, 1.07)	0.698	1.01 (0.93, 1.10)	0.771
		12	Regression (IRR)	0.88 (0.80, 0.98)	0.018	0.92 (0.84, 0.99)	0.036
		24		0.88 (0.78, 0.98)	0.022	0.92 (0.84, 1.00)	0.058
	Symptom Severity Scale	3	Linear Regression	-0.28 (-0.52, -0.04)	0.026	-0.25 (-0.57, 0.65)	0.118
		12	(Mean difference)	-0.52 (-0.75, -0.28)	<0.001	-0.59 (-0.91, -0.27)	<0.001
		24		-0.29 (-0.55, -0.02)	0.040	-0.28 (-0.61, 0.05)	0.100

CI: confidence interval; CWP: chronic widespread pain; EQ5D-5D-5L: Euroqol questionnaire – five dimensions – five levels; GEE: generalised estimating equations with unstructured correlation structure; GHQ: general health questionnaire; ICECAP-A: ICEpop CAPability measure for Adults; IRR: incident rate ratio; OR: odds ratio; *Adjusted analyses control for the number of risk factors (two or three), age, gender, baseline score (if applicable) and centre (random effect). Analyses are intention-to-treat unless otherwise stated. ** Primary outcome.

Table 4: Adjusted[3] mean incremental costs, incremental QALYs, and incremental cost-effectiveness ratio over 24 months between tCBT vs usual care

Analysis	Incremental mean costs, £ (95% CI)[4]	Incremental mean QALYs (95% CI)	ICER (£/QALY)
Imputed dataset (NHS perspective)	42.30 (-451.19 to 597.90)	0.023 (0.007 to 0.039)	1,828
SA: Complete cases (NHS perspective)	229.86 (-228.74 to 734.09)	0.024 (-0.005 to 0.053)	9,608
SA: Imputed dataset (NHS + private care perspective)	90.12 (-475.79 to 772.98)	0.022 (0.007 to 0.039)	4,022
SA: Imputed dataset using actual trial expenses (NHS perspective)[5]	101.07 (-373.14 to 641.98)	0.023 (0.007 to 0.039)	4,367
SA: Imputed dataset using the cost of a complete tCBT course (NHS perspective)[6]	353.59 (-80.46 to 1,238.07)	0.023 (0.007 to 0.039)	15,280
SA: Imputed dataset using ICECAP (NHS perspective)[7]	-127.90 (-603.19 to 545.33)	0.013 (0.003 to 0.023)[8]	NA

QALYs, quality-adjusted life-years; ICER, incremental cost-effectiveness ratio; NA, not applicable; SA, sensitivity analysis.

---

[3] Adjusted for baseline differences (age, gender, number of risk factors present, employment status, centre, baseline EQ-5D health utility score and baseline cost).

[4] Bootstrapped non-parametric 95% confidence interval (2.5th/97.5th centile)/Generalised linear model with γ distribution and log-link function to estimate incremental costs and generalised linear model with Poisson distribution and power 0.5 link function to estimate incremental QALYs/years of full capacity. Discounted at 3.5% per year

[5] Included the actual trial expenses per tCBT participant, £301 regardless of the length of each tCBT session and treatment completion.

[6] Included the cost of a complete tCBT course per participant, £443 using the average duration of each tCBT session.

[7] Adjusted for baseline differences (age, gender, number of risk factors present, employment status, centre, baseline ICECAP value and baseline cost).

[8] Incremental years of full capability.