Appendix H Statistical Appendix

Confidential

The GlaxoSmithKline group of companies

Paroxetine

BRL29060

Statistical Appendix

Protocol BRL29060/676

[redacted] M.Sc*, [redacted] *, [redacted] M.Sc*

*Biomedical Data Sciences

Document Number: BRL-029060/RSD-101T26/1

Issue Date: 10th May 2002

Copyright 2001 the GlaxoSmithKline group of companies. All rights reserved.
Unauthorized copying or use of this information is prohibited.

Table of Contents

List of Tables

List of Figures

1 Introduction

This appendix describes the supplementary statistical analyses performed for study BRL29060/676.

2 Demographic and Baseline Characteristics

A total of 322 patients were randomised, 165 (51.2%) to paroxetine and 157 (48.8%) to placebo; the age breakdown was 92 (28.6%) children and 230 (71.4%) adolescents. Fewer patients were actually randomised than was planned due to a lower than expected attrition rate between randomisation and first post dose assessment. The required number of evaluable patients was still reached.

Of these, 319 patients (163 on paroxetine, 156 on placebo) were included in the intention-to-treat (ITT) population, which was defined as all patients who were randomised into the study, who received at least one dose of double blind medication and who had either a post-baseline assessment or an AE. The first dose of randomised medication was administered on 30th November 1999 and the last dose of randomised medication (including taper) was administered on 19th October 2001.

Three randomised patients were not included in the ITT population, two of whom did not have a post-baseline assessment or an AE (one on paroxetine and one on placebo). The third patient (a paroxetine patient) was randomised in error at the screening visit. This patient has no baseline assessment and the post baseline assessment is after the last dose of medication.

Table 1 shows the characteristics of the patients in the study by treatment group. There was no marked imbalance between the treatment groups in any of the patient characteristics apart from gender, where there were more females than males in the paroxetine group and more males than females in the placebo group.

Table 1: Patient Characteristics

	Paroxetine	Placebo
ITT population	163	156
Age Group (Children:Adolescents) *	(46:117)	(45:111)
Age : Mean (SD)	13 (2.81)	13.3 (2.73)
Gender (Male:Female)	(71:92)	(89:67)
Race n (%)		
White	139 (85.3)	131 (84.
Black	4 (2.5)	6 (3
Oriental	2 (1.2)	2 (1.3)
Other	18 (11.0)	17 (10
Country n (%)		
USA	93 (57.1)	87 (55.8)
Canada	15 (9.2)	14 (9.0)
South Africa	50 (30.7)	50 (32.1)
Belgium	5 (3.1)	5 (3.2)
Baseline CGI Severity of Illness Score: Median (Minimum, Maximum)	5 (3, 7)	5 (3, 7)
Baseline LSAS-CA Score: Mean (SD)	77.6 (28.72)	77.7 (27.05)
Baseline D-GSADS-A Score: Mean (SD)	84.4 (25.42)	81.9 (26.25)
Baseline SPAI-C Score: Mean (SD)	28.1 (11.71)	29.5 (11.06)
Baseline SPAI Difference Score: Mean (SD)	98.7 (31.56)	90.9 (32.23)
Baseline GAF Score: Mean (SD)	53.0 (6.85)	53.5 (7.51)
Baseline CDRS-R Score: Mean (SD)	29.5 (10.43)	30.8 (11.90)

* Children are patients aged 11 or less at their last birthday; and Adolescents are patients aged 12 or greater at their last birthday

3 Covariates and Covariate Groupings

The following covariates were pre-specified and adjusted for in the models:

- Age as a categorical variable (based on the FDA definitions:

  Children are patients aged 11 or less at their last birthday; and Adolescents are patients aged 12 or greater at their last birthday)

  (Note that this covariate is not applicable for the D-GSADS-A, SPAI-C or SPAI analyses).

- Gender

- Baseline efficacy scores (i.e., for the CGI Global Improvement analysis baseline CGI Severity of Illness Score was included, for the LSAS-CA analysis baseline LSAS-CA was included, for the D-GSADS-A analysis baseline D-GSADS-A was included, for the SPAI-C analysis baseline SPAI-C was included, for the SPAI analysis baseline SPAI was included, for the GAF analysis baseline GAF Score was included, for the CDRS-R analysis baseline CDRS-R was included).

- Centre

If any centre recruited less than 8 patients, centres were grouped together to allow the centre effect and the centre-by-treatment interaction to be assessed. The preferred method of grouping was within country. Outlined below is the algorithm used to group the centres.

If there were only a few small centres, within a country, then these centres were to be combined as follows:

For each country, the centres with less than 8 patients were to be ranked, the centre with the smallest number of patients being ranked one. Centres with equal numbers of patients were to be ranked depending on centre number, the lowest centre number being allocated the lowest rank. The centre with the lowest rank was to then be combined with the centre with the highest rank. The centre with the second lowest rank was to be combined with the centre with the second highest rank, and so on to form centre groups within a country. This procedure was to be repeated for centre groups with less than 8 patients until there were no centre groups with less than 8 patients.

However, if there were a large number of small centres, within a country, then all centres were to be combined for that country.

If there were less than 8 patients in total within a country then the following country groupings were to be used where necessary:

USA & Canada

South Africa & Belgium

Note: The process of combining centres was performed on the Intention-to-Treat population and repeated on the Per-Protocol population, before the study was unblinded. Additionally, the algorithm was repeated on the ITT population excluding centre 001. Therefore, it was possible that centre groupings may differ between the Intention-to-Treat analyses (including centre 001), the Per-Protocol analyses and the ITT analyses excluding centre 001.

Following this algorithm resulted in 26 centre groupings to be used in the ITT analyses and the ITT analyses excluding centre 001. For the PP population, when the Belgian centres were combined they still had less than eight patients. Therefore the Belgian centres were grouped with the South African centres to form one level of the country grouping, the other two levels were terms for USA and Canada individually.

However, when the analyses were carried out for the ITT population on the week 16 LOCF dataset, the model including the term for centre grouping did not converge. This was due to insufficient observations in some of the centres, for example all patients within a centre in a specific treatment group being responders or non-responders. For this reason it was decided to group the centres into countries. Upon fitting this model to the data, the treatment-by-country interaction did not converge, due to all patients on paroxetine in Belgium being responders. To allow this interaction to be assessed, it was decided to group Belgium and South Africa together, as this was the pre-specified country grouping given above. The country grouping term thus had three levels, as for the PP population, i.e. 'USA', 'Canada' and 'South Africa and Belgium'. As the centre grouping term did not converge for the primary ITT LOCF analysis, the country grouping term was then also used in the analyses excluding centre 001. Additionally, the country grouping term was carried forward to the analyses of all the secondary endpoints (excluding CGI Severity of Illness, which doesn't make any adjustment for covariates).

It was pre-specified that baseline CGI severity of illness should be included as a categorical covariate in the CGI responder analyses. However, due to an inadequate number of patients in two of the categories the classifications were collapsed into three groupings to facilitate model convergence. These were:

- 'Mildly ill' or 'Moderately ill',
- 'Markedly ill',
- 'Severely ill' or 'Among the most extremely ill patients'.

There were no patients with Baseline CGI severity of illness classified as 'Normal, not at all ill' or 'Borderline mentally ill'.

4 Definitions and Procedures for Analysis and Reporting

4.1 Datasets

The Last Observation Carried Forward (LOCF) dataset contains all data for the week 16 visit, plus the last on-treatment assessment prior to that visit for patients who were not assessed at that visit (this includes early withdrawals).

The PP population comprised 73.4% (234/319) of the ITT population, hence the PP analyses for the primary variable are included below (both Observed Cases (OC) and LOCF are considered).

The 70% LOCF endpoint is the latest time point where at least 70% of patients in each treatment group remain in the study. This occurred at week 12, and hence the 70% LOCF dataset was created.

A potential issue was discovered at one of the centres (centre 001), whereby the blind was broken for all randomised patients upon their completion of the double blind phase of the study. This situation was investigated and the principal investigator confirmed that the unblinding was carried out at the request of the patients parents. It was also confirmed that the physicians who made the efficacy and safety assessments remained blinded to treatment. As a precaution it was decided, prior to breaking the treatment blind, that all patients from this centre would be excluded from the per protocol population. Additionally, a supplementary analysis of the primary efficacy variable was to be performed for the ITT population with centre 001 excluded, in order to assess the overall impact of this centre. However, it should be noted that the ITT population including patients from centre 001 is considered primary.

4.2 Changes in Statistical Methodology

Please note that there are some subtle differences between the analyses specified in the Reporting and Analysis Plan (RAP) and those specified in the protocol. The changes to the RAP were made following FDA comments (letter received June 26th, 2000) on the statistical section of a previously completed paediatric protocol (SBBRL29060/701 - A randomised, multicenter, 8-week, double-blind, placebo-controlled flexible dose study to evaluate the efficacy and safety of paroxetine in children and adolescents with major depressive disorder). Since protocol 701 was a paediatric study of similar design, the FDA comments were

applied to this study too. Specifically, where additional analyses are conducted in the situation where assumptions do not hold for the analyses of continuous variables, they are for assessing robustness of conclusions and do not replace primary inferences. Further, reference to checking for linearity of covariates has been removed. Additionally, it has been clarified that investigation of interactions is limited to the primary variable at the primary timepoint of interest for the primary dataset, and is to assess robustness of conclusions from the primary analysis.

4.3 Treatment of Missing Values

Missing data was handled as detailed in the RAP and is summarised below.

For the individual items of each scale, the raw data values were listed. Any subtotals or total scores that were listed and/or used in the analyses were adjusted to include the relevant imputations outlined below.

4.3.1 CGI Global Improvement Item - Proportion of Responders

Responders were defined as patients who had a global improvement score of 1 or 2 at the particular endpoint.

The proportion of responders based on the global improvement item is defined as:

$$\frac{Number\ of\ patients\ with\ a\ response\ of\ 1\ or\ 2\ at\ the\ visit}{Number\ of\ patients\ with\ a\ CGI\ assessment\ at\ that\ visit\ *}$$

*i.e. the sum of responders and non-responders.

Patients with a 0 value (i.e. not assessed) at the time point of analysis were considered as missing.

4.3.2 CGI Severity of Illness Item

Patients with a 0 value (i.e. not assessed) at baseline or the time point of analysis were considered as missing.

4.3.3 LSAS-CA Total Score

The LSAS-CA provides six subscale scores: total fear, fear of social interaction, fear of performance, total avoidance, avoidance of social interaction, and

avoidance of performance. An overall total score was obtained by summing the total fear and total avoidance scores.

If there were more than 8 missing items out of the 48 Fear/Anxiety or Avoidance items for a patient at a particular timepoint, then the validity of the test was questionable and the overall total score was not calculated[1]. In such a case, the patient's data was excluded from the analysis and the summary tables for this variable at that timepoint.

If at least 40 of the 48 items making up the total were present at a particular timepoint, the missing value(s) was allowed for by calculating the total score as:

$$\frac{(Sum\ of\ the\ scores\ for\ items\ present)\ x\ 48}{Number\ of\ items\ answered}$$

If this resulted in a fractional value it was rounded to the nearest whole number.

4.3.4 D-GSADS-A Total Score

The D-GSADS-A provides four subscale scores: Fear and anxiety score (items A 1-18, anxiety column), Avoidance score (items A 1-18, avoidance column), Affective distress score (items C 1-6), and Somatic distress score (items C 7-11). An overall total score was obtained by summing the four subscale scores (**Note:** items B (3 fear/avoidance seminal items) were NOT included in the calculation of the total score).

Only 1 item was allowed to be missing out of the 18 making up the Fear and anxiety score at a particular timepoint, and only 1 item was allowed to be missing out of the 18 making up the Avoidance score[2]. In each of these cases, the missing value was allowed for by calculating the particular subscale score as:

$$\frac{(Sum\ of\ the\ scores\ for\ items\ present)\ x\ 18}{17}$$

If either of these calculations resulted in a fractional value the relevant subscale score was rounded to the nearest whole number.

---

1 As recommended by the Author

2 As recommended by the Author

No missing items were allowed out of the 6 making up the Affective distress score at a particular timepoint and no missing items were allowed out of the 5 making up the Somatic distress score.

If more than 1 question was missing in either the Fear and anxiety score or the Avoidance score, or if any of the questions were missing in either the Affective distress score or the Somatic distress score, for a patient at a particular timepoint, then that patient's data was excluded from the analysis and the summary tables for this variable at that timepoint.

4.3.5 SPAI-C Total Score

The SPAI-C provides seven separate scores: A1, A2, A3, B1, B2, C and D. The question numbers that make up each of these separate scores are as follows:

A1 = 15 – 20 inclusive

A2 = 9 – 14 inclusive

A3 = 24

B1 = 1 – 8 inclusive

B2 = 22, 23

C = 21

D = 25, 26

Note that question numbers 1 – 8 inclusive and numbers 22 and 23 each have 1 part, question numbers 9 – 20 inclusive and number 24 each have 3 parts, question 21 has 4 parts, and question numbers 25 and 26 each have 5 parts.

Simple arithmetic calculations were then performed as per the instructions below:

- A1, A2, and A3 are summed to get subtotal 1
- B1 is added to B2 to obtain subtotal 2
- C is subtotal 3
- The sum of D is subtotal 4.

If there were 3 or more missing items in a subtotal (subtotal1, subtotal2, subtotal3 or subtotal4), for a patient at a particular timepoint, then the validity of the test was questionable and that subtotal and the total score was not calculated[3]. The patient's data was excluded from the analysis and the summary tables for this variable at that timepoint.

If there were only 1 or 2 omitted items in a subtotal for a patient at a timepoint then a corrected value for that subtotal was obtained using the calculations documented in the appendix of the SPAI-C manual. After obtaining corrected subtotals in the case of missing items, proceed as follows:

- Subtotal 1 is divided by 3
- Subtotal 2 is not divided by anything
- Subtotal 3 is divided by 4
- Subtotal 4 is divided by 5

The results of these calculations were rounded to the nearest whole number (e.g. 3.4 would have been rounded to 3, 10.5 would have been rounded to 11, 17.8 would have been rounded to 18, etc.). The sum of these 4 subtotals is equal to the Total score. The subtotals and total score were obtained programmatically.

4.3.6 SPAI Difference Score

The SPAI provides six separate scores: SP_1, SP_2, SP_3, SP_4, SP_5 and Ag. The question numbers that make up each of these separate scores are as follows:

SP_1 = 1 – 8 inclusive and 27, 28, 29

SP_2 = 20 – 25 inclusive and 30

SP_3 = 9 – 19 inclusive

SP_4 = 31

SP_5 = 26 and 32

Ag = 33 – 45 inclusive

[3] As recommended by the Author

Note that question number 31 has 3 parts, numbers 9 – 25 inclusive and number 30 each have 4 parts, and question numbers 26 and 32 each have 5 parts. All other questions have 1 part.

If there were 4 or more missing items in a subtotal (SP_1, SP_2, SP_3, SP_4, SP_5, or Ag), for a patient at a particular timepoint, then the validity of the test was questionable and that subtotal and the total score were not calculated[4]. The patient's data was excluded from the analysis and the summary tables for this variable at that timepoint.

If there were only 1, 2 or 3 omitted items in a subtotal for a patient at a timepoint then a corrected value for that subtotal was obtained using the calculations documented in the appendix of the SPAI manual.

After obtaining corrected subtotals in the case of missing items, simple arithmetic calculations were then performed to obtain the total social phobia score, total agoraphobia score and the difference score, as per the instructions below:

- SP_2 is added to SP_3, and the resultant sum is divided by 4 to get Z_1
- SP_4 is divided by 3 to get Z_2
- SP_5 is divided by 5 to get Z_3
- The results of these calculations (Z_1 to Z_3) were rounded to the nearest whole number (e.g. 3.4 would have been rounded to 3, 10.5 would have been rounded to 11, 17.8 would have been rounded to 18, etc)
- The sum of SP_1, Z_1, Z_2 and Z_3 is equal to the Total Social Phobia Score (SP)
- Ag gives the Total Agoraphobia Score
- The Difference Score was obtained by subtracting the Total Agoraphobia Score from the Total Social Phobia Score (SP-Ag)

The subtotals and difference score were obtained programmatically.

4 As recommended by the Author

4.3.7 Global Assessment Functioning (GAF)

Patients with a 0 value (i.e. inadequate information) at baseline or the time point of analysis were considered as missing.

4.3.8 CDRS-R Total Score

The CDRS-R total score is the sum of the responses to the 17 questions as recorded in the eCRF. The highest possible score is 113 which represents the most severe measure of depression, and the lowest is 17 for a patient not suffering from depression. If a minimum of 15 questions making up the score were present at a particular timepoint, the missing value(s) was allowed for by calculating the total score as:

$$\text{Observed Total Score} * \left(1 + \frac{\text{Sum of Denominator(s) of the missing value(s)}}{\text{Sum of Denominators of the non - missing values}}\right)$$

Note: Denominator refers to the maximum possible value for a question (either 5 or 7).

As the Total Score was imputed when there were 1 or 2 missing values only, the above formula was simplified for the five following possible scenarios:

One missing question:

1. Missing answer for a 5 item question: i.e. **Observed total score * (1 + 5/108)**

2. Missing answer for a 7 item question: i.e. **Observed total score * (1 + 7/106)**

Two missing questions:

1. Missing answer for a 5 item question and for a 7 item question: i.e. **Observed total score * (1 + 12/101)**

2. Missing answer for two 5 item questions: i.e. **Observed total score * (1 + 10/103)**

3. Missing answer for two 7 item questions: i.e. **Observed total score * (1 + 14/99)**

If the calculation resulted in a fractional value it was rounded to the nearest whole number.

If less than 15 questions were answered for a patient at a particular timepoint then that patient's data was excluded from the analysis and the summary tables for the variable at that timepoint[5].

4.4 Summary of Decisions Regarding the Age Specific Scales

The Social Phobia Anxiety Inventory (SPAI), Social Phobia Anxiety Inventory for Children (SPAI-C), and Dalhousie Generalised Social Anxiety Disorder Scale for Adolescents (D-GSADS-A) are only appropriate for certain subsets of patients within the study (dependant on age). i.e.:

SPAI-C:

"The SPAI-C allows practitioners to assess children aged 8-14 years with a third grade reading level. If the client is over the age of 14, the use of the adult SPAI is recommended."

The protocol specified that the scale was intended to be used in this study for children aged 8-13, but, for the analysis, an exception was made in the protocol requirements such that patients aged 14 or 15 years who inadvertently completed this scale would be included in the analyses of this endpoint. Patient's aged 16 years and above who inadvertently complete the SPAI-C were excluded from the analyses of this endpoint.

SPAI:

"developed originally for adults, but the authors have used the instrument with 12-18 year olds".

"the authors have found the SPAI to be effective for clients aged 14 and above".

The protocol specified that the scale was intended to be used in this study for children aged 14 and over, but, for the analysis, an exception was made in the protocol requirements such that patients aged 13 years who inadvertently completed this scale would be included in the analyses of this endpoint. Patients aged 12 years and below who inadvertently complete the SPAI were excluded from the analyses of this endpoint.

---

[5] As agreed with the Author

D-GSADS-A:

The protocol specified that the scale was intended to be used in this study for the subgroup of patients aged 11-17 years.

However these three scales have sometimes been used to assess patients who are not within the specified age range. Prior to breaking the study blind, decisions were made about how to treat these patients (these decisions are documented in the RAP) and a description of how the data was analysed, tabulated and listed is given below.

NOTE: Age is as calculated at the Screening visit (screening visit date - date of birth).

Listings

All data was listed.

Summary Tables and Analysis

SPAI / SPAI-C:

The following rules were applied, in the order specified, to determine whether the patients were considered in the SPAI or SPAI-C tables. Note that no patient occurred in summaries for both scales.

1) SPAI-C: Data from patients aged 16 or over were excluded from the analyses of this endpoint and are not tabulated

SPAI: Data from patients aged less than 13 were excluded from the analyses of this endpoint and are not tabulated

i.e. as per the RAP, some leeway was allowed in that patients aged 14 or 15 could complete the SPAI-C, similarly patients aged 13 could complete the SPAI, throughout the study.

2) Patients who are aged 13 to 15 (inclusive) who have completed both SPAI and SPAI-C scales are only considered for the scale that was assessed at baseline

3) Patients who are aged 13 to 15 (inclusive) who have completed both SPAI and SPAI-C scales but both or neither scales were assessed at baseline are considered as:

patients aged 14 / 15 -> SPAI

patients aged 13 -> SPAI-C

NOTE: The data included in the summary statistics tables and analysis tables both follow these rules and will generally therefore include the same patients at each visit (given the patient completed the scale). However, the analysis tables also required the patient to have had an assessment at both baseline and the specific visit being considered in order to calculate change from baseline and additionally required complete information on covariates, hence the analysis tables may sometimes have less data included than the corresponding visit in the summary table.

D-GSADS-A:

Only patients aged 11 or greater were considered.

5 Patient Withdrawals

Table 2 shows patient withdrawals during the treatment phase for ITT patients.

Table 2: Patient Withdrawals During the Treatment Phase

Intention-to-Treat Population

	Paroxetine (N=163)	**Placebo (N=156)**
Adverse Experience	10 (6.1%)	3 (1.9%)
Lack of Efficacy	6 (3.7%)	22 (14.1%)
Protocol Deviation (including non-compliance)	11 (6.7%)	11 (7.1%)
Lost to follow-up	4 (2.5%)	10 (6.4%)
Other	9 (5.5%)	7 (4.5%)
Total Withdrawn	40 (24.5%)	53 (34.0%)

In total 29.2% (93/319) of patients in the ITT population withdrew during the treatment phase. Total withdrawals were slightly higher in the placebo group compared to the paroxetine group (53/156, 34.0% compared to 40/163, 24.5%). The number of patients withdrawing due to an AE was slightly higher on paroxetine than placebo, whilst the number withdrawn due to Lack of Efficacy was slightly higher on placebo than paroxetine.

The total number of patients withdrawing in each age group was similar, with 34.1% (31/91) of children and 27.2% (62/228) of adolescents withdrawing. The proportion of children withdrawing on paroxetine (15/46, 32.6%) was similar to the proportion on placebo (16/45, 35.6%), however the proportion of adolescents withdrawing on placebo (37/111, 33.3%) was higher than on paroxetine (25/117, 21.4%).

6 Results

6.1 Primary Efficacy Variable – CGI Global Improvement Item – Proportion of Responders

Primary inference is based on the week 16 LOCF dataset for the ITT population.

An analysis at week 16 OC was additionally carried out to assess the robustness of the results. Similarly, for the primary variable only, analysis was also carried out for the Per-Protocol population. The primary variable was also analysed for the ITT population excluding centre 001 to assess the overall impact of this centre. See Section 4.1 for a description of the issue regarding this centre.

6.1.1 Overall Assessment of Treatment Effect

Since all main effects were to be included in the final model, regardless of statistical significance, the final model for the primary analysis contained the following terms (as defined in Section 3):

- Country Grouping, Baseline CGI Severity of Illness Score, Age Group, Gender.

The results from the logistic regression analysis using the SAS procedure GENMOD are provided in Table 3 (ITT population), Table 6 (PP population) and Table 7 (ITT population excluding centre 001).

6.1.2 Intention-to-Treat Population

The week 16 LOCF ITT dataset for the proportion of responders based on the CGI Global Improvement Item contained 161 patients treated with paroxetine and 154 patients treated with placebo. There were four patients in the ITT population that were not included in this primary analysis. Two of these patients (both placebo patients) had no CGI Global Improvement assessments, one patient (paroxetine patient) had no CGI Global Improvement assessments that were on treatment (they did have a post treatment CGI Global Improvement assessment) and the other patient (paroxetine patient) had no CGI Severity of Illness Baseline Score.

Table 3 summarises the treatment comparisons in the ITT population.

Table 3: Summary of Analysis for CGI Global Improvement Item – Proportion of Responders, Intention-to-Treat Population – Adjusted for Country Grouping, CGI Severity of Illness Baseline Score, Age Group and Gender

	Paroxetine			Placebo			Treatment Comparisons*			
	n	%	N	n	%	N	Odds Ratio	95% CI Lower Limit	95% CI Upper Limit	p-value
Week 16 OC	106	85.5	124	51	51.5	99	6.56	3.29	13.05	<0.001
Week 16 LOCF Endpoint	125	77.6	161	59	38.3	154	7.02	4.07	12.11	<0.001
70% LOCF Endpoint	118	73.3	161	57	37.0	154	5.37	3.21	8.98	<0.001

Source: Table 14.1.2b
* The odds ratios represent the odds of improving with paroxetine relative to that with placebo
Note: Responders are patients who have a score of 1 or 2
Note: Percentage of responders is unadjusted, whilst the odds ratios is adjusted for the terms in the model.
See Section 3 for a description of the covariate groupings

The proportion of patients treated with paroxetine that were CGI Global Improvement Responders at week 16 LOCF endpoint was 125/161 (77.6%) and the proportion of placebo treated patients was 59/154 (38.3%). The odds of being a CGI Global Improvement responder on paroxetine compared to placebo at week 16 LOCF for the Intention-to-Treat population is 7.02 (95% CI: [4.07, 12.11], $p<0.001$); showing a statistically significant benefit of paroxetine over placebo. The study was powered to detect a clinically meaningful difference of 20 percentage points between paroxetine and placebo and this difference was exceeded, with there being a difference of nearly 40 percentage points observed.

Therefore, there is statistically significant evidence that patients treated with paroxetine have a greater response in the CGI Global Improvement Item at week 16 LOCF endpoint than patients treated with placebo.

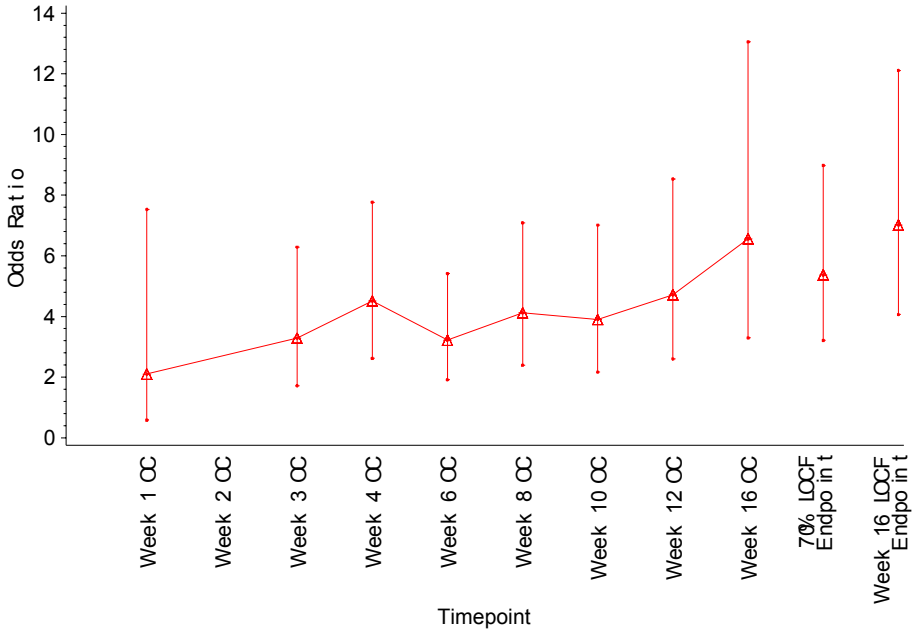
The 70% LOCF endpoint is the last visit where at least 70% of patients in each treatment group remain in the study, this occurred at week 12. Analysis of this endpoint supports the conclusions drawn from the week 16 LOCF endpoint, in that there is statistically significant evidence that patients treated with paroxetine have a greater response in the CGI Global Improvement Item than patients treated with placebo.

The Observed Cases (OC) dataset at week 16 also supports the conclusions drawn from the week 16 LOCF endpoint, in that there is statistically significant evidence

that patients treated with paroxetine have a greater response in the CGI Global Improvement Item than patients treated with placebo.

Figure 1 shows the proportion of CGI responders analysis results at each visit based on the ITT population. Note that results for the analysis at week 2 are not shown as there were convergence problems fitting the specified model.

Figure 1: Proportion of CGI Responders Analysis Results at each Visit - Odds Ratio and 95% Confidence Interval



If there was no difference between the treatments the odds ratio would be one, indicating that the odds of responding on paroxetine are the same as the odds of responding on placebo. It can be seen from Figure 1 that the confidence interval is consistently above this value, indicating that the odds of being a CGI Global Improvement responder on paroxetine are increasingly better than the odds on placebo, over the course of the study.

The covariate significance table for this primary model is shown in Table 4.

Table 4: Summary of Analysis for CGI Global Improvement - Proportion of Responders – Covariate Significance, Week 16 LOCF, Intention-to-Treat Population

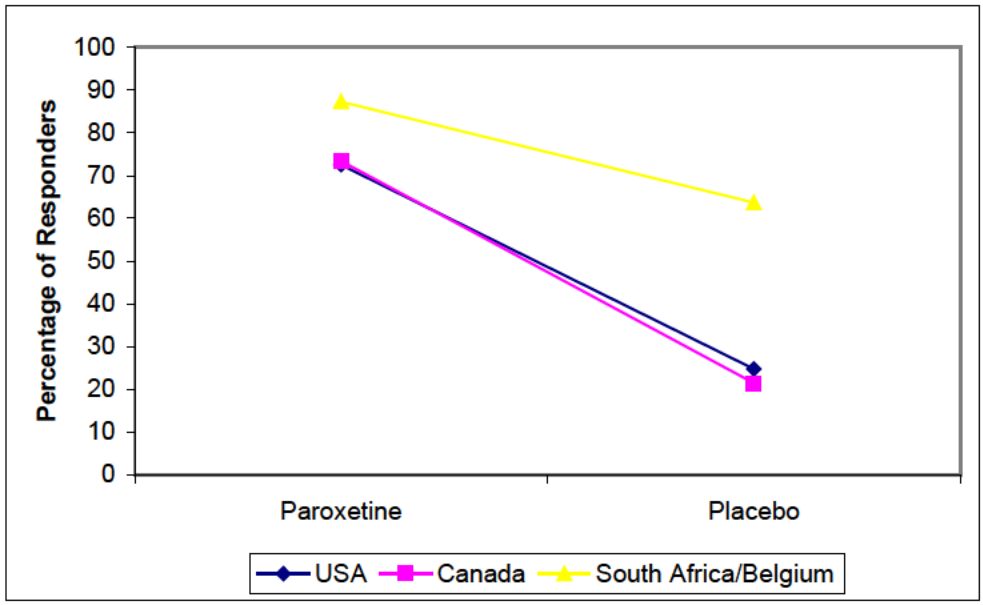
Terms in model	Change in Deviance*	Change in Degrees of Freedom**	P-value***
Country Grouping	25.53	2	<0.001
Baseline Score	2.76	2	0.251
Age Group	3.29	1	0.070
Gender	0.25	1	0.619

Source: Table 14.1.2.1
* Increase in deviance from removing the term from the full model
** Increase in degrees of freedom from removing the term from the full model
*** By comparison to the chi-squared distribution
See Section 3 for a description of the covariate groupings

Table 4 shows that there is a statistically significant difference in response between patients from the different countries (South Africa and Belgium, USA or Canada). These differences are independent of which treatment the patient received. There is no evidence of any variation in response due to varying baseline scores, age group and gender.

Patients in the South Africa / Belgium country group have a greater proportion of responders than those in either USA or Canada, irrespective of treatment group. The odds of being a CGI Global Improvement responder in South Africa / Belgium compared to the USA is 4.38 (i.e. a patient in South Africa / Belgium is over four times more likely to respond than a patient in the USA) and the odds of being a CGI Global Improvement responder in Canada compared to the USA is 1.17. These are both at week 16 LOCF for the ITT population and adjusted for the other covariates in the model. Figure 2 shows the percentage of responders for CGI Global Improvement at week 16 LOCF for patients in each country group split by treatment group.

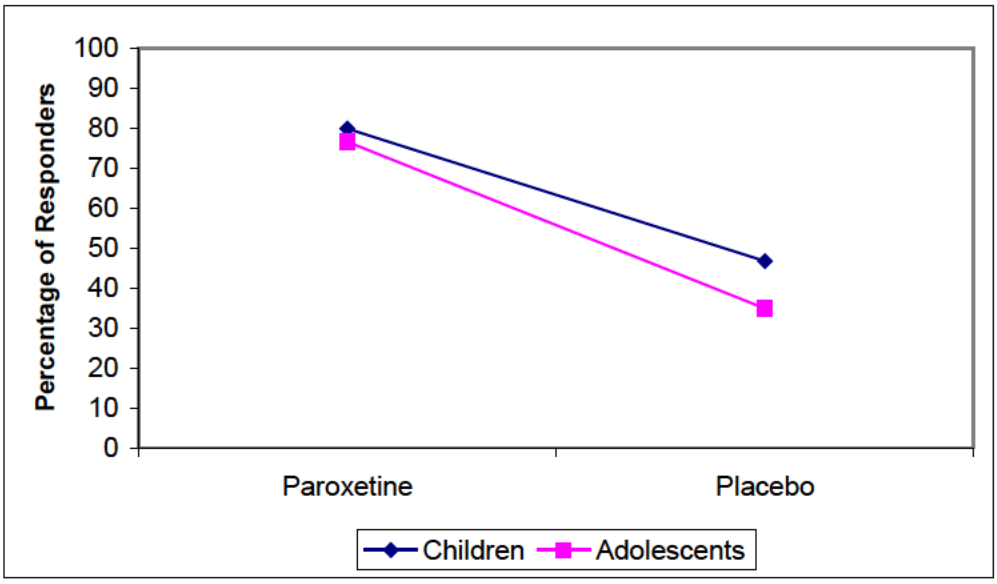
Figure 2: Percentage of Responders for CGI Global Improvement at Week 16 LOCF for Patients in each Country Grouping



It can be seen from Figure 2 that patients in 'South Africa / Belgium' have a greater proportion of responders than patients in the other two countries, irrespective of which treatment they received. The lines drawn onto this plot are reasonably parallel, which displays that there is no interaction between country group and treatment.

Even though the age group term in the model isn't statistically significant, there is some evidence to suggest that there is a difference in response between patients in the different age groups. The odds of being a CGI Global Improvement responder at week 16 LOCF for the ITT population for an adolescent compared to a child is 0.58 (adjusted estimate). Figure 3 shows the percentage of responders for CGI Global Improvement at week 16 LOCF for children and adolescents split by treatment group.

Figure 3: Percentage of Responders for CGI Global Improvement at Week 16 LOCF by Age Group



It can be seen from Figure 3 that children have a greater proportion of responders than adolescents, irrespective of which treatment they received. The lines drawn onto this plot are virtually parallel, indicating that there is no interaction between age group and treatment.

6.1.3 Effect of Missing Values on the Analysis

There were less patients in the placebo treatment group than in the paroxetine group at week 16 OC. For the CGI Global Improvement Item endpoint, 23.0% (37/161) of week 16 LOCF paroxetine patients did not have an assessment at week 16 compared to 35.7% (55/154) of placebo patients. The conclusions from both the OC and LOCF analyses were consistent and therefore the effect of withdrawals on the analysis has been minimal.

6.1.4 Interactions

The consistency of the treatment effect across the covariates was assessed by considering treatment by covariate interactions one at a time with all the main effects in the model. Table 5 shows the significance of each of the treatment by covariate interactions for the primary endpoint.

Table 5: Significance of Treatment by Covariate Interactions for the Analysis of the CGI Global Improvement Item – Proportion of Responders, Week 16 LOCF, Intention-to-Treat Population

Terms in model	Change in Deviance*	Change in Degrees of Freedom**	P-value***
Treatment * Country Grouping	1.88	2	0.392
Treatment * Baseline Score	1.06	2	0.590
Treatment * Age Group	0.30	1	0.585
Treatment * Gender	0.23	1	0.629

* Increase in deviance from removing the term from the full model
** Increase in degrees of freedom from removing the term from the full model
*** By comparison to the chi-squared distribution
See Section 3 for a description of the covariate groupings

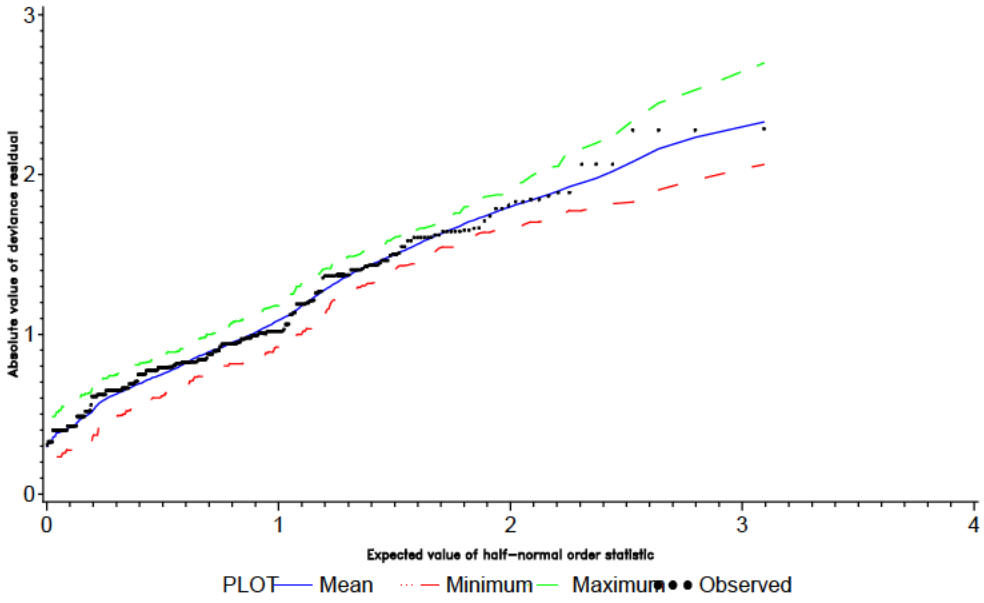
It can be seen from Table 5 that for the primary endpoint there was no evidence of any statistically significant treatment by covariate interactions at the 10% significance level. This indicates that the treatment effect is consistent across country grouping, baseline CGI Severity of Illness Score, age group and gender.

6.1.5 Model Diagnostics

Diagnostic plots were examined and gave no reason to suspect that the underlying assumptions of the primary model were invalid.

Figure 4 shows a half normal probability plot with a simulated envelope. The simulated envelope can be taken as a guide as to what pattern can be expected for residuals from a logistic regression.

Figure 4: Half Normal Probability Plot - ITT Population Week 16 LOCF



It can be seen from Figure 4 that the residuals are contained within the simulated envelope, indicating that the distributional assumptions are reasonable.

6.1.6 Per-Protocol Population

The week 16 LOCF PP dataset for the proportion of CGI Global Improvement responders contained 124 patients treated with paroxetine and 110 patients treated with placebo.

Table 6 summarises the treatment comparisons in the PP population.

Table 6: Summary of Analysis for CGI Global Improvement - Proportion of Responders, Per-Protocol Population – Adjusted for Country Grouping, Baseline CGI Severity of Illness Score, Age Group and Gender

	Paroxetine			Placebo			Treatment Comparisons*			
								95% CI		
	n	%	N	n	%	N	Odds Ratio	Lower Limit	Upper Limit	p-value
Week 16 OC	91	86.7	105	42	49.4	85	7.80	3.60	16.90	<0.001
Week 16 LOCF Endpoint	101	81.5	124	43	39.1	110	8.41	4.36	16.21	<0.001
70% LOCF Endpoint	94	75.8	124	42	38.2	110	5.53	3.03	10.07	<0.001

Source: Table 14.1.2c
* The odds ratios represent the odds of improving with paroxetine relative to that with placebo
Note: Responders are patients who have a score of 1 or 2
Note: Percentage of Responders is unadjusted, whilst the odds ratio is adjusted for the terms in the model.
See Section 3 for a description of the covariate groupings

The proportion of patients treated with paroxetine that were CGI Global Improvement Responders at week 16 LOCF endpoint was 101/124 (81.5%) and the proportion of placebo treated patients was 43/110 (39.1%). The odds of being a CGI Global Improvement responder on paroxetine compared to placebo at week 16 LOCF for the Per-Protocol population is 8.41 (95% CI: [4.36, 16.21], $p<0.001$); showing a statistically significant benefit of paroxetine over placebo.

Therefore, there is statistically significant evidence from the PP analysis that patients treated with paroxetine have a greater response in the CGI Global Improvement Item at week 16 LOCF endpoint than patients treated with placebo, which is consistent with the ITT analysis.

The analysis at the 70% LOCF endpoint (week 12) supports the conclusions drawn from the analysis of the primary dataset.

The Observed Cases (OC) dataset at week 16 endpoint also showed similar results to the week 16 LOCF endpoint, supporting the conclusions drawn.

6.1.7 Intention-to-Treat Population Excluding Centre 001

The week 16 LOCF ITT dataset, excluding centre 001, for the proportion of CGI Global Improvement responders contained 159 patients treated with paroxetine and 152 patients treated with placebo.

Table 7 summarises the treatment comparisons in the ITT population excluding centre 001.

Table 7: Summary of Analysis for CGI Global Improvement – Proportion of Responders, Intention-to-Treat Population Excluding Centre 001 – Adjusted for Country Grouping, Baseline CGI Severity of Illness Score, Age Group and Gender

	Paroxetine			Placebo			Treatment Comparisons*			
	n	%	N	n	%	N	Odds Ratio	95% CI Lower Limit	95% CI Upper Limit	p-value
Week 16 OC	105	86.1	122	51	52.6	97	6.67	3.31	13.47	<0.001
Week 16 LOCF Endpoint	124	78.0	159	59	38.8	152	7.07	4.07	12.29	<0.001
70% LOCF Endpoint	117	73.6	159	56	36.8	152	5.60	3.32	9.44	<0.001

Source: Table 14.1.2bZ
* The odds ratios represent the odds of improving with paroxetine relative to that with placebo
Note: Responders are patients who have a score of 1 or 2
Note: Percentage of Responders is unadjusted, whilst the odds ratio is adjusted for the terms in the model.
See Section 3 for a description of the covariate groupings

The proportion of patients treated with paroxetine that were CGI Global Improvement Responders at week 16 LOCF endpoint was 124/159 (78.0%) and the proportion of placebo treated patients was 59/152 (38.8%). The odds of being a CGI Global Improvement responder on paroxetine compared to placebo at week 16 LOCF for the Intention-to-Treat population excluding Centre 001 is 7.07 (95% CI: [4.07, 12.29], $p<0.001$); showing a statistically significant benefit of paroxetine over placebo.

Therefore, there is statistically significant evidence from the ITT (excluding centre 001) analysis that patients treated with paroxetine have a greater response in the CGI Global Improvement Item at week 16 LOCF endpoint than patients treated with placebo, which is consistent with the primary ITT analysis.

The analysis at the 70% LOCF endpoint (week 12) supports the conclusions drawn from the analysis of the primary dataset.

The Observed Cases (OC) dataset at week 16 endpoint also showed similar results to the week 16 LOCF endpoint, supporting the conclusions drawn.

Thus the inclusion of centre 001 does not affect the overall study conclusions.

6.2 Secondary Efficacy Variables

6.2.1 CGI Severity of Illness Score

The results from the non-parametric Wilcoxon rank sum test using the SAS procedure NPAR1WAY are provided in Table 8.

This procedure does not allow adjustment for covariates, however the analysis is presented separately for each age group.

Table 8: Summary of Analysis for Change from Baseline for CGI Severity of Illness Score Intention-to-Treat Population

	Paroxetine					Placebo					Treatment Comparisons	
	Mean	Median	Min	Max	N	Mean	Median	Min	Max	N	Median Difference	p-value*
Children												
Baseline	4.5	4.0	3	6	45	4.5	4.0	3	6	45		
Change from baseline to:												
Week 16 OC	-2.0	-2.0	-5	0	32	-1.3	-1.0	-4	1	29	-1.0	0.044
Week 16 LOCF Endpoint	-1.9	-2.0	-5	0	45	-0.9	-1.0	-4	1	45	-1.0	<0.001
70% LOCF Endpoint	-1.6	-1.0	-4	0	45	-0.8	-1.0	-4	0	45	-1.0	<0.001
Adolescents												
Baseline	4.7	5.0	3	7	117	4.7	5.0	3	7	110		
Change from baseline to:												
Week 16 OC	-2.3	-3.0	-5	1	92	-1.3	-1.0	-5	0	71	-1.0	<0.001
Week 16 LOCF Endpoint	-2.0	-2.0	-5	1	116	-1.0	-1.0	-5	0	109	-1.0	<0.001
70% LOCF Endpoint	-1.8	-2.0	-5	1	116	-0.9	0.0	-4	0	109	-1.0	<0.001

Source: Table 14.2.3
* P-value from Wilcoxon Rank Sum Test

For children, the median difference between paroxetine and placebo at week 16 LOCF for the ITT population is -1 ($p<0.001$), indicating that there is evidence of a statistically significant benefit of paroxetine over placebo.

Similarly for adolescents, the median difference between paroxetine and placebo at week 16 LOCF for the ITT population is -1 ($p<0.001$) indicating that there is evidence of a statistically significant benefit of paroxetine over placebo.

Similar results were observed for the week 16 OC analyses and the 70% LOCF analyses.

6.2.2 LSAS-CA Total Score

The results from the analysis of covariance modelling using the SAS procedure GLM are provided in Table 9.

Table 9: Summary of Analysis for Change from Baseline in LSAS-CA Total Score, Intention-to-Treat Population – Adjusted for Country Grouping, Baseline LSAS-CA Score, Age Group and Gender

	Paroxetine			Placebo			Treatment Comparisons			
								95% CI		
	LSM+	s.e.+	N	LSM+	s.e.+	N	Diff*	Lower Limit	Upper Limit	p-value
Baseline	77.6	28.72	161	77.7	27.05	155				
Change From Baseline to:										
Week 16 OC	-49.0	2.64	124	-25.7	2.75	101	-23.31	-29.59	-17.03	<0.001
Week 16 LOCF Endpoint	-48.0	2.64	159	-24.3	2.67	150	-23.75	-29.77	-17.74	<0.001
70% LOCF Endpoint	-44.3	2.59	159	-22.2	2.63	150	-22.08	-27.99	-16.16	<0.001

Source: Table 14.3.2
* Diff: Difference in adjusted least square means are shown (Paroxetine minus Placebo)
+ LSM = Least Square Mean: Note that for Baseline, raw means not Least Square means and Standard Deviations not Standard Errors are presented
See Section 3 for a description of the covariate groupings

Patients treated with paroxetine showed an adjusted mean change from baseline to week 16 LOCF in LSAS-CA total score of –48.0 points (s.e. 2.64) and placebo treated patients showed an adjusted mean change of –24.3 points (s.e. 2.67). The adjusted mean difference between paroxetine and placebo at week 16 LOCF for the ITT population is 23.75 points in favour of paroxetine (95% CI: [-29.77, -17.74], p<0.001) providing evidence of a statistically significant benefit of paroxetine over placebo.

Similar results were observed for the week 16 OC analysis and the 70% LOCF analysis.

Diagnostic plots were examined and gave no reason to suspect that the underlying assumptions of the model were invalid.

6.2.3 D-GSADS-A Total Score

The results from the analysis of covariance modelling using the SAS procedure GLM are provided in Table 10. See Section 4.4 for a summary of which patients were included in these analyses.

Table 10: Summary of Analysis for Change from Baseline in D-GSADS-A Total Score, Intention-to-Treat Population – Adjusted for Country Grouping, Baseline D-GSADS-A Score and Gender

	Paroxetine			Placebo			Treatment Comparisons			
								95% CI		
	LSM+	s.e.+	N	LSM+	s.e.+	N	Diff*	Lower Limit	Upper Limit	p-value
Baseline	84.4	25.42	126	81.9	26.25	125				
Change From Baseline to:										
Week 16 OC	-46.8	2.76	97	-23.5	2.90	84	-23.26	-30.51	-16.01	<0.001
Week 16 LOCF Endpoint	-42.9	2.66	124	-21.1	2.71	120	-21.86	-28.56	-15.16	<0.001
70% LOCF Endpoint	-40.1	2.52	124	-18.5	2.57	120	-21.53	-27.88	-15.17	<0.001

Source: Table 14.4.2

* Diff: Difference in adjusted least square means are shown (Paroxetine minus Placebo)

+ LSM = Least Square Mean: Note that for Baseline, raw means not Least Square means and Standard Deviations not Standard Errors are presented

See Section 3 for a description of the covariate groupings

Patients treated with paroxetine showed an adjusted mean change from baseline to week 16 LOCF in D-GSADS-A total score of –42.9 points (s.e. 2.66) and placebo treated patients showed an adjusted mean change of –21.1 points (s.e. 2.71). The adjusted mean difference between paroxetine and placebo at week 16 LOCF for the ITT population is 21.86 points in favour of paroxetine (95% CI: [-28.56, -15.16], p<0.001) providing evidence of a statistically significant benefit of paroxetine over placebo.

Similar results were observed for the week 16 OC analysis and the 70% LOCF analysis.

Diagnostic plots were examined and gave no reason to suspect that the underlying assumptions of the model were invalid.

6.2.4 SPAI-C Total Score

The results from the analysis of covariance modelling using the SAS procedure GLM are provided in Table 11. See Section 4.4 for a summary of which patients were included in these analyses.

Table 11: Summary of Analysis for Change from Baseline in SPAI-C Total Score, Intention-to-Treat Population – Adjusted for Country Grouping, Baseline SPAI-C Score and Gender

	Paroxetine			Placebo			Treatment Comparisons			
								95% CI		
	LSM+	s.e.+	N	LSM+	s.e.+	N	Diff*	Lower Limit	Upper Limit	p-value
Baseline	28.1	11.71	71	29.5	11.06	66				
Change From Baseline to:										
Week 16 OC	-18.1	1.64	51	-8.7	1.75	41	-9.36	-13.55	-5.17	<0.001
Week 16 LOCF Endpoint	-17.6	1.59	69	-8.1	1.62	66	-9.44	-13.19	-5.69	<0.001
70% LOCF Endpoint	-16.8	1.57	69	-8.1	1.60	66	-8.75	-12.44	-5.05	<0.001

Source: Table 14.5.2
* Diff: Difference in adjusted least square means are shown (Paroxetine minus Placebo)
+ LSM = Least Square Mean: Note that for Baseline, raw means not Least Square means and Standard Deviations not Standard Errors are presented
See Section 3 for a description of the covariate groupings

Patients treated with paroxetine showed an adjusted mean change from baseline to week 16 LOCF in SPAI-C total score of –17.6 points (s.e. 1.59) and placebo treated patients showed an adjusted mean change of –8.1 points (s.e. 1.62). The adjusted mean difference between paroxetine and placebo at week 16 LOCF for the ITT population is 9.44 points in favour of paroxetine (95% CI: [-13.19, -5.69], p<0.001) providing evidence of a statistically significant benefit of paroxetine over placebo.

Similar results were observed for the week 16 OC analysis and the 70% LOCF analysis.

Diagnostic plots were examined and gave no reason to suspect that the underlying assumptions of the model were invalid.

6.2.5 SPAI Difference Score

The results from the analysis of covariance modelling using the SAS procedure GLM are provided in Table 12. See Section 4.4 for a summary of which patients were included in these analyses.

Table 12: Summary of Analysis for Change from Baseline in SPAI Difference Score, Intention-to-Treat Population – Adjusted for Country Grouping, Baseline SPAI Difference Score and Gender

	Paroxetine			Placebo			Treatment Comparisons			
								95% CI		
	LSM+	s.e.+	N	LSM+	s.e.+	N	Diff*	Lower Limit	Upper Limit	p-value
Baseline	98.7	31.56	81	90.9	32.23	84				
Change From Baseline to:										
Week 16 OC	-56.0	4.98	61	-24.6	4.95	54	-31.37	-43.62	-19.12	<0.001
Week 16 LOCF Endpoint	-51.9	4.53	77	-19.1	4.40	81	-32.80	-43.57	-22.03	<0.001
70% LOCF Endpoint	-47.2	4.19	77	-16.9	4.07	81	-30.35	-40.30	-20.40	<0.001

Source: Table 14.6.2

* Diff: Difference in adjusted least square means are shown (Paroxetine minus Placebo)

+ LSM = Least Square Mean: Note that for Baseline, raw means not Least Square means and Standard Deviations not Standard Errors are presented

See Section 3 for a description of the covariate groupings

Patients treated with paroxetine showed an adjusted mean change from baseline to week 16 LOCF in SPAI difference score of –51.9 points (s.e. 4.53) and placebo treated patients showed an adjusted mean change of –19.1 points (s.e. 4.40). The adjusted mean difference between paroxetine and placebo at week 16 LOCF for the ITT population is 32.80 points in favour of paroxetine (95% CI: [-43.57, -22.03], p<0.001) providing evidence of a statistically significant benefit of paroxetine over placebo.

Similar results were observed for the week 16 OC analysis and the 70% LOCF analysis.

Diagnostic plots were examined and gave no reason to suspect that the underlying assumptions of the model were invalid.

6.2.6 GAF Score

The results from the analysis of covariance modelling using the SAS procedure GLM are provided in Table 13.

Table 13: Summary of Analysis for Change from Baseline in GAF Score, Intention-to-Treat Population – Adjusted for Country Grouping, Baseline GAF Score, Age Group and Gender

	Paroxetine			Placebo			Treatment Comparisons			
								95% CI		
	LSM+	s.e.+	N	LSM+	s.e.+	N	Diff*	Lower Limit	Upper Limit	p-value
Baseline	53.0	6.85	162	53.5	7.51	155				
Change From Baseline to:										
Week 16 OC	19.5	1.24	124	10.4	1.30	101	9.17	6.21	12.13	<0.001
Week 16 LOCF Endpoint	17.1	1.14	159	8.4	1.15	151	8.74	6.15	11.34	<0.001
70% LOCF Endpoint	15.0	1.01	159	7.4	1.02	151	7.58	5.28	9.88	<0.001

Source: Table 14.7.2

* Diff: Difference in adjusted least square means are shown (Paroxetine minus Placebo)

+ LSM = Least Square Mean: Note that for Baseline, raw means not Least Square means and Standard Deviations not Standard Errors are presented

See Section 3 for a description of the covariate groupings

Patients treated with paroxetine showed an adjusted mean change from baseline to week 16 LOCF in GAF score of 17.1 points (s.e. 1.14) and placebo treated patients showed an adjusted mean change of 8.4 points (s.e. 1.15). The adjusted mean difference between paroxetine and placebo at week 16 LOCF for the ITT population is 8.74 points in favour of paroxetine (95% CI: [6.15, 11.34], $p<0.001$) providing evidence of a statistically significant benefit of paroxetine over placebo.

Similar results were observed for the week 16 OC analysis and the 70% LOCF analysis.

The diagnostic plots produced for this analysis were Normal probability plots with a simulated envelope. These were examined and gave some reason to suspect that the underlying assumptions of the model may be invalid. In order to assess the robustness of the conclusions drawn from the parametric analysis, a non-parametric analysis was performed using the Wilcoxon Rank Sum test, taking no account of any covariates. This additional analysis supports the conclusions

drawn above, specifically that there is a statistically significant benefit of paroxetine over placebo.

6.3 Other Efficacy Variable

6.3.1 CDRS-R Total Score

The results from the analysis of covariance modelling using the SAS procedure GLM are provided in Table 14.

Table 14: Summary of Analysis for Change from Baseline in CDRS-R Total Score, Intention-to-Treat Population – Adjusted for Country Grouping, Baseline CDRS-R Score, Age Group and Gender

	Paroxetine			Placebo			Treatment Comparisons			
								95% CI		
	LSM+	s.e.+	N	LSM+	s.e.+	N	Diff*	Lower Limit	Upper Limit	p-value
Baseline	29.5	10.43	162	30.8	11.90	155				
Change From Baseline to:										
Week 16 OC	-6.2	0.79	124	-2.1	0.83	100	-4.03	-5.91	-2.15	<0.001
Week 16 LOCF Endpoint	-4.8	0.97	145	-1.1	1.00	126	-3.61	-5.88	-1.34	0.002

Source: Table 14.8.2

* Diff: Difference in adjusted least square means are shown (Paroxetine minus Placebo)

+ LSM = Least Square Mean: Note that for Baseline, raw means not Least Square means and Standard Deviations not Standard Errors are presented

See Section 3 for a description of the covariate groupings

Patients treated with paroxetine showed an adjusted mean change from baseline to week 16 LOCF in CDRS-R total score of –4.8 points (s.e. 0.97) and placebo treated patients showed an adjusted mean change of –1.1 points (s.e. 1.00). The adjusted mean difference between paroxetine and placebo at week 16 LOCF for the ITT population is 3.61 points in favour of paroxetine (95% CI: [-5.88, -1.34], p=0.002) providing evidence of a statistically significant benefit of paroxetine over placebo.

Similar results were observed for the week 16 OC analysis. The 70% LOCF analysis was not carried out as CDRS-R scores were only collected at baseline and week 16 / early withdrawal.

Diagnostic plots were examined and gave some reason to suspect that the underlying assumptions of the model may be invalid. An additional non-

parametric analysis was performed using the Wilcoxon Rank Sum test, taking no account of any covariates. This additional analysis supports the conclusions drawn above.

6.4 Safety

6.4.1 Adverse Experiences

One hundred and forty four patients out of one hundred and sixty three patients (88.3%) in the ITT population randomised to paroxetine reported gender non-specific emergent adverse experiences during the treatment phase, compared with 125/156 patients (80.1%) receiving placebo.

The most common gender non-specific adverse experiences for patients on paroxetine were headache, infection, respiratory disorder, abdominal pain, asthenia, insomnia, somnolence, rhinitis and nausea whilst the most common adverse experiences for patients on placebo were headache, infection, rhinitis and respiratory disorder.

One patient out of seventy one male patients on paroxetine (1.4%) reported a male specific AE, whilst 6/92 female patients on paroxetine (6.5%) and 4/67 female patients on placebo (6.0%) reported female specific AE's.

Seventy six out of ninety one (83.5%) children reported gender-non-specific emergent adverse experiences during the treatment phase (41/46 (89.1%) on paroxetine, 35/45 (77.8%) on placebo), whilst 193/228 (84.6%) adolescents reported gender non-specific emergent adverse experiences during the treatment phase (103/117 (88.0%) on paroxetine, 90/111 (81.1%) on placebo).

6.4.2 Serious Adverse Experiences

Three out of one hundred and sixty five patients (1.8%) randomised to paroxetine reported gender non-specific serious adverse experiences during the treatment, taper or follow-up phase compared with 1/157 placebo patients (0.6%). No male or female specific serious adverse experiences were reported for either treatment group.

7 Conclusions

The results for the primary endpoint (CGI Global Improvement Item – Proportion of Responders) have provided evidence that paroxetine is more efficacious than placebo in treating children and adolescents with Social Anxiety Disorder/Social Phobia.

The results for all the secondary endpoints have also provided evidence that paroxetine is more efficacious than placebo in treating children and adolescents with Social Anxiety Disorder/Social Phobia, and support the primary endpoint results.

For the primary endpoint there was no evidence of any statistically significant treatment by covariate interactions, indicating that the treatment effect is consistent across country grouping, baseline CGI Severity of Illness Score, age group and gender.

In total, just over 29% of patients withdrew during the treatment phase, with slightly less on paroxetine than placebo. Although the total number of withdrawals in each age group was similar and the proportion of children withdrawing on paroxetine was similar to that on placebo, the proportion of adolescents withdrawing on paroxetine was lower than on placebo.

The most common gender non-specific adverse experiences for patients on paroxetine were headache, infection, respiratory disorder, abdominal pain, asthenia, insomnia, somnolence, rhinitis and nausea whilst the most common adverse experiences for patients on placebo were headache, infection, rhinitis and respiratory disorder. The only gender specific adverse experience reported in more than 5% of patients was dysmenorrhea. The percentage of emergent adverse experiences reported during the treatment phase was similar in the two age groups. Three (1.8%) patients on paroxetine reported a gender non-specific serious adverse event (SAE) compared to one (0.6%) on placebo. There were no gender specific SAE's reported.

8 References

1. COOK, R.D. & WEISBURG, S.. 1982. Residuals and Influence in Regression. Chapman & Hall.

2. HASTIE, T.J. & TIMBSHIRANI, R.J.. 1990. Generalized Additive Models. Chapman & Hall.

Attachment 1 Summary of Patients with Comorbidity

The tables below give a summary of the numbers of patients that have at least one comorbid condition, based on the clinicians overall diagnosis from the ADIS C/P.

Table 15 shows the number of children and adolescents in each treatment group who had Social Anxiety Disorder/Social Phobia for the ITT population.

Table 15: Number of Patients with a Diagnosis of Social Phobia / Social Anxiety Disorder

Age Group	Paroxetine	Placebo	Total
Children	46	45	91
Adolescents	117	111	228
Total	163	156	319

As expected, all patients have a diagnosis of Social Anxiety Disorder / Social Phobia.

Table 16 shows the number of children and adolescents in each treatment group who had Social Anxiety Disorder / Social Phobia and any other psychiatric condition for the ITT population.

Table 16: Number of Patients with a Diagnosis of Social Anxiety Disorder / Social Phobia and any other Psychiatric Condition

Age Group	Paroxetine	Placebo	Total
Children	30	30	60
Adolescents	79	64	143
Total	109	94	203

The table shows that in total 203 patients had a psychiatric condition in addition to Social Anxiety Disorder / Social Phobia . Of these 203 patients there were slightly more on paroxetine than on placebo (109 on paroxetine and 94 on placebo). There are an equal number of children with an additional psychiatric condition across the treatment groups (30 in each treatment group) and there are slightly more adolescents in the paroxetine group than in the placebo group (79 on paroxetine compared to 64 on placebo).

Table 17 shows the number of children and adolescents in each treatment group with Social Anxiety Disorder / Social Phobia and any other psychiatric condition, excluding School Refusal Behaviour and Interpersonal Relationship, in the ITT population.

Table 17: Number of Patients with a Diagnosis of Social Anxiety Disorder / Social Phobia and any other Psychiatric Condition, Excluding School Refusal Behaviour and Interpersonal Relationships

Age Group	Paroxetine	Placebo	Total
Children	29	26	55
Adolescents	63	50	113
Total	92	76	168

The table shows that in total 168 patients had a psychiatric condition (excluding School Refusal Behaviour and Interpersonal Relationships) in addition to Social Anxiety Disorder / Social Phobia . Of these 168 patients there are slightly more on paroxetine than on placebo (92 on paroxetine and 76 on placebo). There is a similar number of children with an additional psychiatric condition (excluding School Refusal Behaviour and Interpersonal Relationships) across the treatment groups (29 on paroxetine, 26 on placebo) and there are slightly more adolescents in the paroxetine group than in the placebo group (63 on paroxetine compared to 50 on placebo).