1. Introduction

This document describes the statistical analysis and reporting to be undertaken for paroxetine adult suicidality data. The data include trials submitted, or planned to be submitted, as part of the adult suicidality review for the Food and Drug Administration (FDA) in September and December 2005. The data also include trials which, because of their duration, fall outside the scope of the FDA review.

Adult, randomized, parallel group, placebo controlled trials in which the total number of patients treated with paroxetine and placebo was at least 30, comprised the full set of trials included in the analysis (Appendix 1). The subset of trials which were less than 17 weeks in duration formed the set of trials submitted to the FDA.

2. Objectives

2.1 Primary objective

The primary objective is to compare the incidence of definitive suicidal behaviour and ideation on paroxetine and placebo. The null hypothesis is that there is no difference between the two treatment groups. The alternative hypothesis is that a difference between the two treatment groups does exist.

2.2 Secondary objectives

The secondary objectives are:

- To compare the incidence of other measures of suicidal behaviour and ideation on paroxetine and placebo.
- To compare the incidence of suicidal behaviour and ideation on paroxetine and placebo in disease and demographic subgroups.
- To compare the efficacy of paroxetine and placebo.

3. Indications

3.1 Datasets to be analysed

The following datasets will be analysed, according to the indication being investigated:

1. All Indications
2. All Depression (i.e. 3-7 below)
3. Major Depressive Disorder (MDD)
4. Intermittent Brief Depression (IBD)
5. Dysthymia*
6. Bipolar Disorder*
7. Depression with Chronic Back Pain*
8. All Non-Depression (i.e. 9-16 below)
9. Panic Disorder
10. Obsessive Compulsive Disorder (OCD)
11. Social Anxiety Disorder (SAD)
12. Generalized Anxiety Disorder (GAD)

13. Post-traumatic Stress Disorder (PTSD)
14. Pre-menstrual Dysphoric Disorder (PMDD)
15. Detoxification in Alcoholics (EtOH)*
16. Fibromyalgia*

*indicates that the indication contains only one trial

The trials included in each indication are specified in Appendix 1.

3.2 Long term extension trials

Where data are included from long term extension trials, in which patients continue on medication to which they were randomized in an acute trial, the data from the long term phase will be displayed within the results of the original trial in which patients were randomized. The trials affected are specified in Appendix 1.

4. Definitions

4.1 Columbia University Suicidality Classifications of Adverse Events

As part of the FDA's adult suicidality data review, potential cases of suicidality were identified from searches of adverse event terms, a review of all deaths and serious adverse events (SAEs) and from a review of comments fields on Case Report Forms (CRFs). Cases were only included in the list of potential events if they occurred during the double-blind phase of treatment or within one day following the cessation of randomized treatment.

For all potential events a detailed narrative was prepared by Drug Safety Alliance, Inc (DSA) and these narratives were forwarded to Dr Kelly Posner at the Columbia University Medical Centre, who randomly assigned them to a group of independent suicide experts for review and classification. Each narrative was reviewed by three expert raters and assigned a code according to the following classifications specified by the FDA:

1. Completed suicide
2. Suicide attempt
3. Preparatory acts toward imminent suicidal behaviour
4. Suicidal ideation
5. Self-injurious behaviour, intent unknown
6. Not enough information (fatal)
7. Self-injurious behaviour, no suicidal intent
8. Other: accident; psychiatric, medical
9. Not enough information (non-fatal).

Categories 1-4 above will be referred to collectively as *Definitive Suicidal Behaviour and Ideation*. Categories 1-3 above will be referred to collectively as *Definitive Suicidal Behaviour*.

4.2 Rating Scale Emergent Ideation and Behaviour (Depression only)

Data from the suicidality questions on the Hamilton Depression Rating Scale (HAMD, item 3) and Montgomery Asberg Depression Rating Scale (MADRS, item 10) will also be used to assess the risk of suicidality in trials of depression. In other indications the HAMD and MADRS were generally administered only at baseline, and for this reason they will not be used to assess emergent suicidality in those indications.

Item 3 of the HAMD measures suicidal thoughts and behaviour on the following scale:

0 = Absent
1 = Feels life is not worth living
2 = Wishes he were dead or any thoughts of possible death to self
3 = Suicidal ideas or gesture
4 = Attempts at suicide.

Item 10 of the MADRS measures suicidal thoughts and behaviour on the following scale:

0 = Enjoys life or takes it as it comes
1
2 = Weary of life. Only fleeting suicidal thoughts
3
4 = Probably better off dead. Suicidal thoughts are common, and suicide is considered as a possible solution, but without specific plans or intention.
5
6 = Explicit plans for suicide when there is an opportunity. Active preparation for suicide.

Emergent suicidal behaviour and ideation on HAMD (item 3) or MADRS (item 10) will be defined as any case where a patient's pre-treatment baseline score was 0 or 1 and where this score increased to a score of ≥ 3 while on double-blind treatment (including one day after the cessation of the treatment).

Emergent suicidal behaviour on HAMD (item 3) or MADRS (item 10) will be defined as any case where a patient's pre-treatment baseline score was 0 or 1 and the patient had a post-baseline score of 4 (HAMD) or 6 (MADRS) while on double-blind treatment (including one day after the cessation of the treatment).

In any trial where both the HAMD and MADRS were used, data will be assessed independently on each scale, and a patient will be considered to have satisfied the definition of emergent ideation or behaviour if the criteria have been met for one or both of the scales.

4.3 Intent to Treat Population

All analyses will be based on the Intent-to-Treat population, which is defined as all patients who were randomized and received at least one dose of trial medication. This is the same definition that was used in populating datasets for submission to the FDA.

Additionally, analyses of change from baseline require that at least one post-baseline measurement was taken for the parameter of interest.

5. Assessment of Risk

5.1 Endpoints

The following endpoints will be analysed:

1. Definitive Suicidal Behaviour and Ideation
2. Rating Scale Emergent Suicidal Behaviour and Ideation
3. Composite Suicidal Behaviour and Ideation (i.e. 1 or 2 above)
4. Definitive Suicidal Behaviour
5. Rating Scale Emergent Suicidal Behaviour
6. Composite Suicidal Behaviour (i.e. 4 or 5 above)

Definitive Suicidal Behaviour and Ideation (endpoint 1) is the primary endpoint. Endpoints 2, 3, 5, and 6 pertain only to the Depression indications.

Endpoints 1 and 4 above are obtainable directly from datasets supplied to FDA as part of the adult suicidality submissions (short term trials less than 17 weeks duration only). Endpoints 2, 3, 5, and 6 require additional data not included as part of the FDA datasets.

5.2 Statistical methods

For each endpoint the incidence of events will be compared between treatment groups (paroxetine and placebo). The analysis will be adjusted for trial using the exact method of StatXact® (PROC STRATIFY, StatXact for SAS®). Together with the incidence of the event in each treatment group, an estimate of the common odds ratio will be presented, together with 95% “mid-p” confidence interval (CI). For the overall estimates of treatment effect, but not for each individual trial, a p-value will be presented. For the adjusted analysis the exact p-value will be calculated by the method of summing all probabilities less than or equal to the observed.

To assess the robustness of the method of adjusting for trial, an overall estimate of the odds ratio and its 95% CI will also be obtained by adjusting for trial using the Mantel Haenszel method (with 0.5 continuity correction).

The number needed to harm (NNH) will also be presented. The NNH is equal to the reciprocal of the probability difference and is interpreted as the number of patients who need to be treated to incur one additional adverse outcome over a fixed time period. Larger values of NNH correspond to a lower risk of an adverse outcome on paroxetine relative to placebo. If the value of the NNH is negative this indicates that

the adverse outcome is less likely on paroxetine than on placebo and it is referred to as the number needed to treat (NNT) i.e. the number of patients who need to be treated to prevent one additional adverse outcome across a fixed time period. For adjusted analyses the NNH and NNT will be calculated using a conversion from the adjusted odds ratio (Appendix 2).

The results will be presented in a table similar to Table 1.

Table 1: Incidence of Definitive Suicidal Behaviour and Ideation by Indication, Treatment Group, and Trial
Indication = Major Depressive Disorder

Trial	Paroxetine	Placebo	OR (95% CI)	P-value	NNT/NNH
Overall (exact, adjusted)	x / xxx (x.xx%)	y / yyy (y.yy%)	z.z (z.z, z.z)	0.xxx	z.z
Overall (Mantel Haenszel)	x / xxx (x.xx%)	y / yyy (y.yy%)	z.z (z.z, z.z)	0.xxx	z.z
Trial 1	x / xxx (x.xx%)	y / yyy (y.yy%)	z.z (z.z, z.z)		z.z
Trial 2	x / xxx (x.xx%)	y / yyy (y.yy%)	z.z (z.z, z.z)		z.z
Trial 3	x / xxx (x.xx%)	y / yyy (y.yy%)	z.z (z.z, z.z)		z.z

The results will also be presented in a Forest plot, displaying the results for each trial and for the overall results. An additional Forest plot may also be produced in which small trials are combined into a single group; this is dependent on the results and the visual clarity of the initial Forest plot.

Heterogeneity of results across trials will be assessed using Zelen's exact test, or with the Breslow-Day test if Zelen's test cannot be computed.

For datasets 1, 2, and 8 (i.e. All Indications Combined, All Depression Combined, and All Non-Depression Combined) the analysis will be adjusted for trial, but the tables and figures will be presented by indication, not by trial.

5.3 Risk factors

In addition to the overall analysis, results will be presented according to the following risk factors:

1. Baseline suicidal ideation, defined as the presence of one or more of:
   HAMD item 3 score ≥ 3
   MADRS item 10 score ≥ 3
   Symptom Checklist 90 (SCL-90) Q15 score ≥ 1
   Beck Depression Inventory (BDI) Q7 score ≥ 2
   BDI Q9 score ≥ 1
   One or more previous suicide attempts (trials 057 and 106).
2. Age (continuous covariate)
3. Age group (18-24, 25-64, ≥ 65)
4. Gender

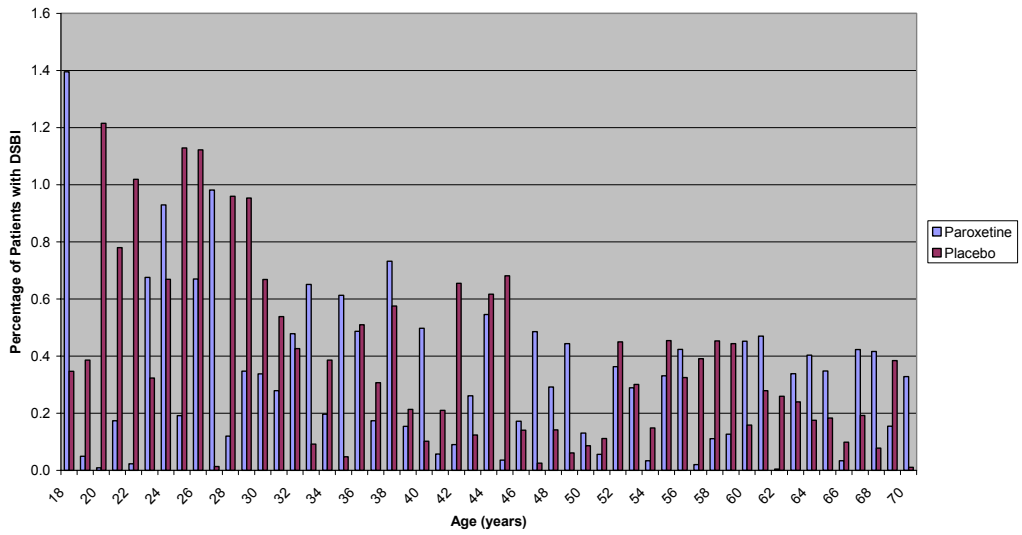
For categorical covariates (i.e. 1, 3 and 4 above) the data will be analysed and presented in the same way as the exact analysis described in section 5.2. A table similar to Table 2 will be produced to present the results according to each risk factor. A Forest plot will be presented showing the results for each risk factor.

Table 2: Incidence of Definitive Suicidal Behaviour and Ideation by Indication, Treatment Group, and Risk Factors Indication = Major Depressive Disorder					
Trial	**Paroxetine**	**Placebo**	**OR (95% CI)**	**P-value**	**NNT/NNH**
Overall (exact, adjusted)	x / xxx (x.xx%)	y / yyy (y.yy%)	z.z (z.z, z.z)	0.xxx	z.z
Baseline suicidal ideation					
Absent	x / xxx (x.xx%)	y / yyy (y.yy%)	z.z (z.z, z.z)		z.z
Present	x / xxx (x.xx%)	y / yyy (y.yy%)	z.z (z.z, z.z)		z.z
Age Group					
18-24	x / xxx (x.xx%)	y / yyy (y.yy%)	z.z (z.z, z.z)		z.z
25-64	x / xxx (x.xx%)	y / yyy (y.yy%)	z.z (z.z, z.z)		z.z
≥65	x / xxx (x.xx%)	y / yyy (y.yy%)	z.z (z.z, z.z)		z.z
Gender					
Male	x / xxx (x.xx%)	y / yyy (y.yy%)	z.z (z.z, z.z)		z.z
Female	x / xxx (x.xx%)	y / yyy (y.yy%)	z.z (z.z, z.z)		z.z

For the analysis of age as a continuous covariate a generalized linear model will be fitted modelling the log-odds of an event as a linear function of age. In the event that the linear model does not provide a good fit, other models (e.g. quadratic) will be assessed. The results will be presented graphically showing the predicted odds ratio as a function of age. The incidence of an event for each individual age will be presented by treatment group in a histogram, similar to that in Figure 1 (Note: the data in Figure 1 are illustrative only, they are not data collected from paroxetine clinical trials). The number and percentage of patients with Definitive Suicidal Behaviour and Ideation will be tabulated by age.

In the initial stage, risk factor and subgroup analyses will be conducted only for the primary endpoint of Definitive Suicidal Behaviour and Ideation. Risk factor and subgroup analyses may be conducted subsequently for other endpoints if there are notable differences between the conclusions derived from the different endpoints.

Figure 1: Percentage of Patients with Definitive Suicidal Behaviour and Ideation by Treatment Group and Age



Note: in the figure that is produced the data will be clustered for each treatment group, not for each age group in the way shown in this illustration.

6. Assessment of Efficacy

6.1 Statistical methods

Continuous efficacy measures (i.e. "change from baseline") will be analysed using a general linear model, which will contain terms for baseline score, trial, and treatment group. If the model fails to converge then the smallest trials will be combined with each other until convergence is achieved. Categorical efficacy measures (i.e. "responder" analyses) will be analysed using the exact method described in section 5.2.

For efficacy measures the last time point eligible for inclusion as an endpoint measure will be the first day following the cessation of treatment.

For responder analyses the number needed to treat (NNT) will also be presented, and will be calculated using a conversion from the adjusted odds ratio (Appendix 2). The NNT is interpreted as the number of patients who need to be treated to achieve one additional positive outcome across a fixed time period. Smaller values of NNT correspond to a larger benefit of paroxetine relative to placebo. If the value of the NNT is negative this indicates that a positive outcome is more likely on placebo than on paroxetine and it is referred to as the number needed to harm (NNH) i.e. the number of patients who need to be treated to produce one additional adverse outcome across a fixed time period.

Trials with both an acute phase and an extension phase (see section 2.2) will be considered as a single trial, with the baseline measurement defined according to the point of randomization into the trial. For Last Observation Carried Forward (LOCF) analyses, the endpoint will always be measured for the acute phase of the trial.

6.2 Efficacy measures

6.2.1 All Depression Combined

Efficacy will be assessed using the following measures:

1. Change from baseline on the HAMD total score to LOCF endpoint.
2. Change from baseline on the MADRS total score to LOCF endpoint.
3. Change from baseline on the HAMD item 3 to LOCF endpoint.
4. Change from baseline on the MADRS item 10 to LOCF endpoint.
5. Declining suicidal ideation, i.e. a baseline score of ≥ 3 on HAMD item 3 or MADRS item 10, reducing to a score of 0 or 1 at endpoint.
6. Responder analysis, i.e. reduction of 50% or more between baseline and endpoint on HAMD total score (or MADRS total score for trials where HAMD not recorded).

6.2.2 Major Depressive Disorder

Efficacy will be assessed using the following measures:

1. Change from baseline on the HAMD total score to LOCF endpoint.
2. Change from baseline on the MADRS total score to LOCF endpoint.
3. Change from baseline on the HAMD item 3 to LOCF endpoint.
4. Change from baseline on the MADRS item 10 to LOCF endpoint.
5. Declining suicidal ideation, i.e. a baseline score of ≥ 3 on HAMD item 3 or MADRS item 10, reducing to a score of 0 or 1 at endpoint.
6. Responder analysis, i.e. reduction of 50% or more between baseline and endpoint on HAMD total score (or MADRS total score for trials where HAMD not recorded).

6.2.3 Intermittent Brief Depression

Efficacy will be assessed using the following measures:

1. Change from baseline on the MADRS total score to LOCF endpoint.
2. Change from baseline on the MADRS item 10 to LOCF endpoint.
3. Declining suicidal ideation, i.e. a baseline score of ≥ 3 on MADRS item 10, reducing to a score of 0 or 1 at endpoint.
4. Responder analysis, i.e. reduction of 50% or more between baseline and endpoint on MADRS total score.

6.2.4 Dysthymia

Efficacy will be assessed using the following measures:

1. Change from baseline on the HAMD total score to LOCF endpoint.
2. Change from baseline on the HAMD item 3 to LOCF endpoint.
3. Declining suicidal ideation, i.e. a baseline score of ≥ 3 on HAMD item 3, reducing to a score of 0 or 1 at endpoint.
4. Responder analysis, i.e. reduction of 50% or more between baseline and endpoint on HAMD total score.

6.2.5 Bipolar Disorder

Efficacy will be assessed using the following measures:

1. Change from baseline on the HAMD total score to LOCF endpoint.
2. Change from baseline on the HAMD item 3 to LOCF endpoint.
3. Declining suicidal ideation, i.e. a baseline score of ≥ 3 on HAMD item 3, reducing to a score of 0 or 1 at endpoint.
4. Responder analysis, i.e. reduction of 50% or more between baseline and endpoint on HAMD total score.

6.2.6 Depression with Chronic Back Pain

Efficacy will be assessed using the following measures:

1. Change from baseline on the MADRS total score to LOCF endpoint.
2. Change from baseline on the MADRS item 10 to LOCF endpoint.
3. Declining suicidal ideation, i.e. a baseline score of ≥ 3 on MADRS item 10, reducing to a score of 0 or 1 at endpoint.
4. Responder analysis, i.e. reduction of 50% or more between baseline and endpoint on MADRS total score.

6.2.7 All Non-Depression Combined

Efficacy will be assessed using the following measure:

1. Responder analysis, i.e. Clinical Global Impression (CGI) Improvement Score of 1 (Very Much Improved) or 2 (Much Improved) at LOCF endpoint.

6.2.8 Panic Disorder

Efficacy will be assessed using the following measure:

1. Responder analysis, i.e. Clinical Global Impression (CGI) Improvement Score of 1 (Very Much Improved) or 2 (Much Improved) at LOCF endpoint.

The number of panic attacks is an endpoint common to all panic disorder trials, and being submitted as part of the FDA submission, but the time period in which the number of attacks is measured varies between trials. Consequently, this measure is not being used in this pooled analysis of data from panic disorder trials.

6.2.9 Obsessive Compulsive Disorder

Efficacy will be assessed using the following measures:

1. Change from baseline on the Yale Brown Obsessive Compulsive Scale (YBOCS) total score to LOCF endpoint.
2. Responder analysis, i.e. Clinical Global Impression (CGI) Improvement Score of 1 (Very Much Improved) or 2 (Much Improved) at LOCF endpoint.

6.2.10 Social Anxiety Disorder

Efficacy will be assessed using the following measures:

1. Change from baseline on the Liebowitz Social Anxiety Scale (LSAS) total score to LOCF endpoint.
2. Responder analysis, i.e. Clinical Global Impression (CGI) Improvement Score of 1 (Very Much Improved) or 2 (Much Improved) at LOCF endpoint.

6.2.11 Generalized Anxiety Disorder

Efficacy will be assessed using the following measures:

1. Change from baseline on the Hamilton Anxiety Scale (HAMA) total score to LOCF endpoint.
2. Responder analysis, i.e. Clinical Global Impression (CGI) Improvement Score of 1 (Very Much Improved) or 2 (Much Improved) at LOCF endpoint.

6.2.12 Post-traumatic Stress Disorder

Efficacy will be assessed using the following measures:

1. Change from baseline on the Clinician-Administered PTSD Scale (CAPS-2) total score to LOCF endpoint.
2. Responder analysis, i.e. Clinical Global Impression (CGI) Improvement Score of 1 (Very Much Improved) or 2 (Much Improved) at LOCF endpoint.

6.2.13 Pre-menstrual dysphoric disorder

Efficacy will be assessed using the following measures:

1. Change from baseline in Mean Luteal Phase Visual Analogue Scale (VAS) Mood Score to LOCF endpoint (trials 677, 688, 689 only).
2. Responder analysis, i.e. Clinical Global Impression (CGI) Improvement Score of 1 (Very Much Improved) or 2 (Much Improved) at LOCF endpoint.

6.2.14 Detoxification in Alcoholics

Efficacy will be assessed using the following measures:

1. Change from baseline on Total Daily Alcohol Consumption to LOCF endpoint.
2. Responder analysis, i.e. Clinical Global Impression (CGI) Improvement Score of 1 (Very Much Improved) or 2 (Much Improved) at LOCF endpoint.

6.2.15 Fibromyalgia

Efficacy will be assessed using the following measures:

1. Change from baseline on the VAS to LOCF endpoint.
2. Responder analysis, i.e. Clinical Global Impression (CGI) Improvement Score of 1 (Very Much Improved) or 2 (Much Improved) at LOCF endpoint.

7. Comparison of Adverse Event Reporting by Time

To assess whether there has been a change over time in adverse event reporting in general, or specifically for adverse events relating to suicidality, the incidence of Definitive Suicidal Behaviour and Ideation, and the incidence of any adverse event, will be plotted by trial and treatment group in a histogram. Trials will be ordered by start date. A corresponding summary table will also be produced.

8. List of Tables and Figures

The tables and figures to be produced are listed below.

Number	Title	Comments
Table 1.01	Demographic Characteristics by Indication, Treatment Group and Trial	
Table 2.01	Number and Percent of Patients with Definitive Suicidal Behaviour and Ideation by Indication, Treatment Group and Trial	Ordered by Start Date
Table 2.02	Number and Percent of Patients with Rating Scale Emergent Suicidal Behaviour and Ideation by Indication, Treatment Group and Trial	Ordered by Start Date. Depression indications only
Table 2.03	Number and Percent of Patients with Composite Suicidal Behaviour and Ideation by Indication, Treatment Group and Trial	Ordered by Start Date. Depression indications only
Table 2.04	Number and Percent of Patients with Definitive Suicidal Behaviour by Indication, Treatment Group and Trial	Ordered by Start Date
Table 2.05	Number and Percent of Patients with Rating Scale Emergent Suicidal Behaviour by Indication, Treatment Group and Trial	Ordered by Start Date. Depression indications only
Table 2.06	Number and Percent of Patients with Composite Suicidal Behaviour by Indication, Treatment Group and Trial	Ordered by Start Date. Depression indications only
Table 2.07	Number and Percent of Patients with Any Adverse Event by Indication, Treatment Group and Trial	Ordered by Start Date
Table 2.08	Number and Percent of Patients with Definitive Suicidal Behaviour and Ideation by Indication, Treatment Group and Risk Factors	
Table 3.01	Change from Baseline on HAMD Total Score to LOCF endpoint by Indication, Treatment Group and Risk Factors	Depression indications only
Table 3.02	Change from Baseline on MADRS Total Score to LOCF endpoint by Indication, Treatment Group and Risk Factors	Depression indications only
Table 3.03	Change from Baseline on HAMD Item 3 to LOCF endpoint by Indication, Treatment Group and Risk Factors	Depression indications only
Table 3.04	Change from Baseline on MADRS Item 10 to LOCF endpoint by Indication, Treatment Group and Risk Factors	Depression indications only
Table 3.05	Number and Percent of Patients with Declining Suicidal Ideation by Indication, Treatment Group and Risk Factors	Depression indications only

Table 3.06	Number and Percent of Patients with >= 50% Reduction in HAMD or MADRS Baseline Score by Indication, Treatment Group and Risk Factors	Depression indications only
Table 3.07	Number and Percent of CGI Responders (Very Much Improved or Much Improved) at LOCF endpoint by Indication, Treatment Group and Risk Factors	Non-Depression Indications only
Table 3.08	Change from Baseline on YBOCS Total Score to LOCF endpoint by Treatment Group and Risk Factors	OCD only
Table 3.09	Change from Baseline on LSAS Total Score to LOCF endpoint by Treatment Group and Risk Factors	SAD only
Table 3.10	Change from Baseline on HAMA Total Score to LOCF endpoint by Treatment Group and Risk Factors	GAD only
Table 3.11	Change from Baseline on CAPS-2 Total Score to LOCF endpoint by Treatment Group and Risk Factors	PTSD only
Table 3.12	Change from Baseline on Mean Luteal Phase VAS Mood Score to LOCF endpoint by Treatment Group and Risk Factors	PMDD only
Table 3.13	Change from Baseline on Total Daily Alcohol Consumption to LOCF endpoint by Treatment Group and Risk Factors	EtOH only
Table 3.14	Change from Baseline on VAS Score to LOCF endpoint by Treatment Group and Risk Factors	Fibromyalgia only
Figure 2.01	Definitive Suicidal Behaviour and Ideation by Indication and Trial	Forest plot
Figure 2.02	Rating Scale Emergent Suicidal Behaviour and Ideation by Indication and Trial	Forest plot. Depression indications only
Figure 2.03	Composite Suicidal Behaviour and Ideation by Indication and Trial	Forest plot. Depression indications only
Figure 2.04	Definitive Suicidal Behaviour by Indication and Trial	Forest plot
Figure 2.05	Rating Scale Emergent Suicidal Behaviour by Indication and Trial	Forest plot. Depression indications only
Figure 2.06	Composite Suicidal Behaviour by Indication and Trial	Forest plot. Depression indications only
Figure 2.07	Definitive Suicidal Behaviour and Ideation by Indication and Baseline Suicidal Ideation	Forest plot
Figure 2.08	Definitive Suicidal Behaviour and Ideation by Indication and Age	Model

Figure 2.09	Definitive Suicidal Behaviour and Ideation by Indication and Age Group	Forest plot
Figure 2.10	Definitive Suicidal Behaviour and Ideation by Indication and Gender	Forest plot
Figure 2.11	Percent of Patients with Definitive Suicidal Behaviour and Ideation by Indication, Treatment Group and Trial	Histogram
Figure 2.12	Percent of Patients with Any Adverse Event by Indication, Treatment Group and Trial	Histogram, Ordered by Start Date
Figure 2.13	Percent of Patients with Definitive Suicidal Behaviour and Ideation by Indication, Treatment Group and Age	Histogram

Appendix 1: List of trials

Trial	Indication	Trial	Indication
276 (MDUK09 Edwards)	MDD	118	OCD
279 (MDUK12 Trimble)	MDD	136	OCD
274 (MDUK06 Naylor)	MDD	241 (LTX of 136)	OCD
001	MDD	414	OCD
002	MDD	660	OCD
009	MDD	108	Panic
003	MDD	120	Panic
115	MDD	187	Panic
128	MDD	222 (LTX of 120)	Panic
251	MDD	223	Panic
448	MDD	228 (LTX of 187)	Panic
449	MDD	494	Panic
487	MDD	495	Panic
625	MDD	497	Panic
785	MDD	384	Panic
810	MDD	410	Panic
NKD20006*	MDD	400	PMDD
874	MDD	427 (LT)	PMDD
442	MDD	658	PMDD
057 (LT)	IBD	677	PMDD
106 (LT)	IBD	688	PMDD
327	Dysthymia	689	PMDD
352	Bipolar	711 (LTX of 677, 688, 689)	PMDD
298	Dep with back pain	627	PTSD
433	Fibromyalgia	648	PTSD
201	EtOH	651	PTSD
637	GAD	382	SAD
641	GAD	454	SAD
642	GAD	502	SAD
791	GAD	790	SAD
116	OCD	661	SAD

* paroxetine was the active comparator.

LTX = Long Term Extension, not included in FDA review

LT = Long Term, not included in FDA review

Appendix 2: Calculation of Number Needed to Treat (NNT) and Number Needed to Harm (NNH)

NNT and NNH for Unadjusted Analyses

The NNT/NNH is calculated as the reciprocal of the probability difference for an event, i.e.

$$NNT = 1/(Pt - Pc)$$

where Pt is the probability of a beneficial outcome in the test group (paroxetine)
and Pc is the probability of a beneficial outcome in the control group (placebo).

If the NNT is negative then it is referred to as the NNH. If the outcome is a measure of harm (e.g. an adverse event) then the NNT and NNH are reversed.

NNT and NNH for Adjusted Analyses

Where an analysis is adjusted for trial or other covariates, the NNT/NNH can be calculated by converting the adjusted odds ratio:

$$NNT = \frac{Pc * (AOR - 1) + 1}{Pc * (AOR - 1) * (1 - Pc)}$$

where Pc is the probability of a beneficial outcome in the control group (placebo) and AOR is the adjusted odds ratio for a beneficial outcome (expressed as the odds of the outcome on paroxetine relative to placebo).

If the NNT is negative then it is referred to as the NNH. If the outcome is a measure of harm (e.g. an adverse event) then the NNT and NNH are reversed.