Australian Government
Department of Health
Therapeutic Goods Administration

AusPAR Attachment 2

Extract from the Clinical Evaluation Report for Nalmefene (as hydrochloride dihydrate)

Proprietary Product Name: Selincro

Sponsor: Lundbeck Australia Pty Ltd

First round CER report: 10 February 2014

Second round CER report: 22 September 2014

TGA Health Safety Regulation

About the Therapeutic Goods Administration (TGA)

- The Therapeutic Goods Administration (TGA) is part of the Australian Government Department of Health, and is responsible for regulating medicines and medical devices.
- The TGA administers the *Therapeutic Goods Act 1989* (the Act), applying a risk management approach designed to ensure therapeutic goods supplied in Australia meet acceptable standards of quality, safety and efficacy (performance), when necessary.
- The work of the TGA is based on applying scientific and clinical expertise to decision-making, to ensure that the benefits to consumers outweigh any risks associated with the use of medicines and medical devices.
- The TGA relies on the public, healthcare professionals and industry to report problems with medicines or medical devices. TGA investigates reports received by it to determine any necessary regulatory action.
- To report a problem with a medicine or medical device, please see the information on the TGA website <https://www.tga.gov.au>.

About the Extract from the Clinical Evaluation Report

- This document provides a more detailed evaluation of the clinical findings, extracted from the Clinical Evaluation Report (CER) prepared by the TGA. This extract does not include sections from the CER regarding product documentation or post market activities.
- The words [Information redacted], where they appear in this document, indicate that confidential information has been deleted.
- For the most recent Product Information (PI), please refer to the TGA website <https://www.tga.gov.au/product-information-pi>.

Contents

List of abbreviations

Abbreviation	Meaning
ADS	Alcohol Dependence Scale
ALAT	alanine aminotransferase
AME	absorption, metabolism, and excretion
ANCOVA	analysis of covariance
APTS	all-patients-treated set
ASAT	aspartate aminotransferase
AUC	area under the plasma concentration versus time curve
BAC	blood alcohol concentration
BMI	body mass index
BOCF	baseline observation carried forward
BRENDA	Biopsychosocial evaluation Report to the patient on assessment Empathic understanding of the patient's situation Needs collaboratively identified by the patient and treatment provider Direct advice to the patient on how to meet those needs Assess reaction of the patient to advice and adjust as necessary for best care
CER	Clinical Evaluation Report
CER_1	Clinical Evaluation Report – First round
CER_2	Clinical Evaluation Report – Second round
CGI-I	Clinical Global Impression – Global Improvement
CGI-S	Clinical Global Impression – Severity of Illness
CI	confidence interval
CL	clearance
CL/F	oral clearance

Abbreviation	Meaning
C_{max}	maximum observed concentration
CYP	cytochrome P450 isoenzyme
DRL	drinking risk level
DSM-IV-TR™ Revision	Diagnostic and Statistical Manual of Mental Disorders, 4th Edition, Text
ECG	electrocardiogram
EMA	European Medicines Agency
ESRD	end-stage renal disease
FAS	full-analysis set
FDA	United States Food and Drug Administration
GGT	gamma-glutamyltransferase
HDAB	(at-least) High DRL at Baseline
HDAR	(at-least) High DRL at Randomisation
HDD	heavy drinking day
ICH	International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use
IMP	investigational medicinal product
ITBM	intended to be marketed
ka	absorption rate constant
kel	elimination rate constant
LBM	lean body mass
LDAR	Low DRL at Randomisation
LFT	liver function test
LOCF	last observation carried forward
LREG	logistic regression
MAA	Marketing Authorisation Application

Abbreviation	Meaning
MDAR	(at-least) Medium DRL at Randomisation
MedDRA	Medical Dictionary for Regulatory Activities
MI	multiple imputation
MMRM	mixed model repeated measures
NMF	nalmefene
NNTB	number-needed-to-treat-to-benefit
OC	observed cases
PBO	placebo
PCS	potentially clinically significant
PCtab	Patheon Inc., Canada – clear film-coated tablet
PET	positron emission tomography
PMI	placebo mean imputation
PMM	pattern mixture model
PYE	patient years of exposure
QTc	heart rate corrected QT interval
QTcI	heart rate corrected QT interval based on an individual correction method
RLDRL below	response defined as a downward shift from baseline in DRL to low DRL or
RSDRL	response defined as a downward shift from baseline of at least two DRL levels or to low DRL
SAE	serious adverse event
S-D	single dose
SD	standard deviation
SMQ	standardised MedDRA query
t½	apparent elimination half-life
TAC	total alcohol consumption

Abbreviation	Meaning
TEAE	treatment-emergent adverse event
TLFB	timeline followback
t_{max}	time to maximum observed concentration
UGT	uridine diphosphate glucuronosyltransferase
UItab	University of Iowa – clear film-coated tablet
UKATT	United Kingdom Alcohol Treatment Trial
Vz/F	volume of distribution
WHO	World Health Organization

1. Background

1.1. Submission type

This is a Category 1 (Type A) submission to register a new chemical entity.

1.2. Drug class and therapeutic indication

The indication in the Sponsor's application form and the original proposed Product Information (PI) sheet reads as follows:

> *"SELINCRO is indicated for the reduction of alcohol consumption in adult patients with alcohol dependence who have a high Drinking Risk Level (DRL), without physical withdrawal syndrome and who do not require immediate detoxification.*
>
> *SELINCRO should be prescribed in conjunction with psychosocial support focused on treatment adherence and reducing alcohol consumption."*

The proposed indication does not include a comment on whether subjects should have failed non-pharmacological measures to treat their alcohol dependence prior to being prescribed nalmefene. Elsewhere in the proposed PI, however, the following comment appears during description of the pivotal efficacy studies:

> *In Studies 1 (12014A; n = 579) and 2 (12023A; n = 655), 18%, and 33%, of the total population, respectively, considerably reduced their alcohol consumption in the period between screening and randomisation. Of the patients with a high or very high DRL at baseline, 35% experienced improvement due to non-pharmacological effects in the period between the initial visit (screening) and randomisation. At randomisation, these patients consumed such a small amount of alcohol that there was little room for further improvement (floor effect).* ***Therefore, the patients who maintained a high or very high DRL at randomisation [that is, show persisting high DRL despite the non pharmacological effects associated with the enrolment process] were defined post hoc as the target population.*** *(Emphasis and [explanatory addition] added.)*

The quoted paragraph indicates that the Sponsor now defines the target population according to very specific post hoc criteria – continued high or very high DRL despite the non-pharmacological, psychosocial effects at work prior to randomisation in the pivotal studies – and that this was the subgroup in which the evidence for therapeutic efficacy was more favourable. Accordingly, the proposed indication should be altered to match. In the Sponsor's primary submission, the Sponsor was attempting to use a very restrictive definition of the target group in their post hoc selection of favourable efficacy results, while nonetheless pursuing a broader definition of the target group for their marketing authorisation – and neither of these definitions corresponded with the prospectively defined target group of the pivotal studies, which was broader still.

The situation is further complicated by the Sponsor's reference to 'high or very high DRL at randomisation' in the bolded section above, which makes sense in a trial setting but maps imprecisely to the clinical setting – such that the Sponsor and the evaluator disagree on how the target group should be identified in the clinical setting.

It is of additional concern that the Sponsor's original Briefing Document used a very broad definition of the target group:

> *"Selincro is indicated for the reduction of alcohol consumption, in conjunction with psychosocial support, in adult patients with alcohol dependence"*

This broad definition is at odds with the Sponsor's subsequent focus on one particular post hoc subgroup analysis, and with their insistence that the prospective efficacy results in the total pivotal population should be omitted from the PI and de-emphasised in this evaluation report.

The approved European Summary of Product Characteristics expresses the indication as follows:

> *Selincro is indicated for the reduction of alcohol consumption in adult patients with alcohol dependence who have a high drinking risk level (DRL) [see section 5.1], without physical withdrawal symptoms and who do not require immediate detoxification. Selincro should only be prescribed in conjunction with continuous psychosocial support focused on treatment adherence and reducing alcohol consumption.* ***Selincro should be initiated only in patients who continue to have a high DRL two weeks after initial assessment.*** *(Emphasis added.)*

The European wording for the indication is more appropriate than that proposed for the Australian PI in the Sponsor's original submission, but it does not clearly indicate the purpose of the two-week wait.

Also, the indication should specify what 'high DRL' means, using terms familiar to most prescribers.

In the first-round Clinical Evaluation Report (CER_1), the following indication was suggested:

> *SELINCRO is indicated for the reduction of alcohol consumption in adult patients with alcohol dependence who continue to have a high Drinking Risk Level (alcohol > 60 g/day for men and > 40 g/day for women) despite psychosocial interventions including counselling and documentation of alcohol intake during a pre-treatment baseline period of at least 2 weeks.*
>
> *SELINCRO should only be prescribed in conjunction with psychosocial support focused on treatment adherence and reducing alcohol consumption. It should not be used in patients who have physical withdrawal syndrome or require immediate detoxification.*

In their Section 31 response, the Sponsor has explicitly rejected this wording, along with the evaluator's interpretation of how the Sponsor's post hoc definition of the target group ('at least high DRL at Randomisation and Baseline') might be applied to clinical practice, proposing yet another indication:

> *SELINCRO is indicated for the reduction of alcohol consumption in adult patients with alcohol dependence who continue to have a high Drinking Risk Level (alcohol > 60 g/day for men and > 40 g/day for women) for at least 2 weeks after a comprehensive clinical assessment, including documentation of alcohol intake.*
>
> *SELINCRO should only be prescribed in conjunction with psychosocial support focused on treatment adherence and reducing alcohol consumption. It should not be used in patients who have physical withdrawal syndrome and require immediate detoxification.*

In the evaluator's opinion, this version is still inadequate.

1.3. Dosage forms and strengths

The submission proposes registration of the following dosage form:

- SELINCRO nalmefene 18mg (as hydrochrloride dehydrate)

In different contexts, this dose is variously referred to as nalmefene 20mg (referring to the quantity of nalmefene hydrochloride) or as 18mg (referring to the quantity of nalmefene base). In all of the major efficacy studies and most of the clinical pharmacology studies, the 20mg designation was used, and this report will follow that convention.

1.4. Dosage and administration

The proposed dose is nalmefene 18mg daily [20mg daily] as needed, taken before the perceived period of risk.

The proposed PI provides the following dosing instructions:

> *SELINCRO is to be taken as-needed: on each day the patient perceives a risk of drinking alcohol; one tablet should be taken, preferably 1 - 2 hours prior to the anticipated time of drinking. If the patient has started drinking alcohol without taking SELINCRO, the patient should take one tablet as soon as possible. The maximum dose of SELINCRO is one tablet per day.*

2. Clinical rationale

2.1. Evolution of clinical role for nalmefene

Nalmefene binds to opioid receptors, including μ, δ, and κ receptors. In vitro studies have shown that it is a selective opioid receptor antagonist at the μ and δ receptors and a partial agonist at κ receptors. The Sponsor proposes that antagonist activity at μ opioid receptors (mu-receptors) is the most important activity with respect to the current proposed indication.

Nalmefene was initially developed as an opioid antagonist, and parenteral forms have been approved for use in the treatment of opioid overdose in a number of countries including the United States (1995) and Canada (1996). It has since been discontinued in the United States, for business reasons, and similarly it appears not to have made it to the Canadian market. Parenteral forms are still available for treatment of opioid overdose in Mexico and China.

Nalmefene has also been tested for efficacy in a large number of indications, ranging from cystitis to rheumatoid arthritis, with an eventual focus on addiction disorders, including gambling and smoking. Finally, it was assessed for efficacy in alcohol dependence, the currently proposed indication. The drug has passed through a number of sponsors, each of whom has had different ideas about its potential therapeutic role. The current Sponsor, Lundbeck, acquired the rights for nalmefene in 2006, and has performed three efficacy/safety studies in the setting of alcohol dependence. The previous Sponsor, Biotie, assessed nalmefene in alcohol use disorders, a more loosely defined category.

The drug was not primarily developed for treatment of alcohol dependence, and the clinical rationale for its use in alcohol dependence is somewhat post hoc. Indeed, the Sponsor's conception of how the drug should be used is still evolving. After the pivotal studies described in this report, but prior to submission, the current Sponsor significantly modified the therapeutic target; the target group described in the proposed PI, for instance, is different to that identified prospectively in the pivotal efficacy studies.

The precise mechanism of action (MOA) of nalmefene in the treatment of alcohol dependence is still unclear. The Sponsor points out that alcohol consumption results in mesolimbic dopamine release that is facilitated by the release of β-endorphins, and this provides positive reinforcement. The Sponsor also suggests that "After repeated exposure to high doses of alcohol, neuroadaptations occur in several neurotransmitter neuropeptide systems, including the opioid receptor system, which leads to negative reinforcement that drives continued alcohol intake." The relationship between positive and negative enforcement is not made clear in the Sponsor's Clinical Overview, and details of the proposed neuroadapatations involved in negative reinforcement were not provided in the original clinical submission. The Clinical Overview suggests that antagonism of mu-receptors plays a role, but a detailed pharmacological rationale was not provided. The Sponsor's submission also mentions that preclinical in vivo studies have shown that nalmefene reduces alcohol consumption, but assessment of those preclinical studies

is beyond the scope of this Clinical Evaluation Report. The Sponsor proposes that the therapeutic effect is possibly mediated by modulating cortico-mesolimbic functions, and concludes: "The most likely mechanism of action of nalmefene is to reduce the reinforcing effects of alcohol and thereby help the patient to reduce drinking."

Overall, understanding of the MOA of nalmefene appeared vague in the Sponsor's original submission, and this did not provide a strong basis for expecting substantial efficacy in humans. On the other hand, naltrexone, another opioid antagonist, has been approved for treatment of alcohol dependence in some countries, so there is some indirect evidence that opioid antagonism may be useful in alcohol dependence.

In response to the first-round Clinical Evaluation Report (CER1), the Sponsor has provided additional information about the proposed MOA of nalmefene.

2.2. Clinical need for new treatments for alcohol dependence

The Sponsor mounts a clear case that new treatments for alcohol dependence are needed. Alcoholism is obviously a major clinical and social problem, with extensive repercussions including liver disease, heart disease, neurological toxicity, an increased risk of injury, and social disruption. The evidence showing alcoholism to be harmful does not need to be reviewed here, but the Sponsor's Clinical Overview provides a good summary (pages 10 to 16).

Alcohol dependence is rightly considered a disease, because of the underlying neurotransmitter changes that mediate addiction, but alcohol intake also reflects conscious choices on the part of the drinker. This means that education and other psychosocial interventions play an important role in curtailing excessive drinking. Alcohol dependence is difficult to study in standard placebo-controlled studies, because the endpoint, drinking behaviour, is under partial voluntary control.

The main existing treatments for alcohol dependence consist of psychosocial support programs, with pharmacological approaches playing a secondary role. The main psychosocial measures include Cognitive Behavioural Therapy, Motivational Enhancement Therapy, and 12-Step Facilitation. These can be combined with acute detoxification approaches, which usually require inpatient monitoring and management of withdrawal effects. Research has shown that all of the major psychosocial interventions have similar efficacy and that they share many overlapping design features. The most important aspect of all psychosocial interventions is the quality of the therapeutic relationship between the treatment provider and the patient.

In the submitted Lundbeck studies, blinded nalmefene treatment was combined with a psychosocial approach known as BRENDA, standing for:

- Biopsychosocial evaluation
- Report to the patient on assessment
- Empathic understanding of the patient's situation
- Needs collaboratively identified by the patient and treatment provider
- Direct advice to the patient on how to meet those needs
- Assess reaction of the patient to advice and adjust as necessary for best care

Existing *pharmacological* treatments for alcohol dependence include disulfiram, acamprosate, and naltrexone, all of which are currently approved in the EU for this indication. The same drugs are also indicated for the maintenance of abstinence as part of a counselling programme.

The role of these drugs has been summarised by the European Medicines Agency (EMA) in their *"Guideline on the development of medicinal products for the treatment of alcohol dependence"*, as follows:[1]

> *Disulfiram is classified as an aversive treatment modality and primarily applied as a test of motivation or compliance with therapy. It interferes with alcohol metabolism, causing accumulation of toxic acetaldehyde. If alcohol is consumed simultaneously (although it is strongly recommended to avoid this), Disulfiram causes severe headache, nausea, and with higher amount of alcohol also more dangerous toxic effects.*
>
> *Acamprosate, a GABA agonist and functional glutamate antagonist, is used as an anti-craving substance in several EU countries for preventing relapses in abstinent alcohol users. It has shown higher abstinence rates and longer periods of abstinence respectively, compared to placebo in several but not all trials.*
>
> *Naltrexone, a non-selective opiate antagonist, binds with receptors for endogenous opioids and appears to modify some of the reinforcing effects of alcohol and to prevent the reinstatement of extinguished alcohol-seeking behaviour induced by alcohol-associated cues. Naltrexone treated alcohol-dependent patients have been reported to drink less frequently and smaller quantities. However, no benefit in continued abstinence vs. placebo has been shown.*

In general, it is expected that use of these drugs is combined with psychosocial treatments.

2.3. Alcohol reduction as a therapeutic goal

In previous studies of treatments aiming to achieve abstinence, 40 to 75% of the patients who follow an abstinence treatment plan relapse within the first 12 months.[2] In one 5-year follow-up study,[3] 44% of the patients had remained abstinent, 38% had relapsed to heavy drinking, and 7% were drinking at non-heavy levels.

Given that success rates in achieving abstinence are poor with available interventions, and not every patient who suffers from alcohol dependence has abstinence as their goal, the Sponsor proposes that a treatment aimed at reducing alcohol intake in problem drinkers could serve a useful role. This is reasonable, but raises the question of how much reduction is a worthwhile therapeutic goal.

The World Health Organisation's *"International Guide for Monitoring Alcohol Consumption and Related Harm"* defines several brackets of risk, as shown in the table below; the narrowest bracket ranges from 40-60 g/day in men and from 20 - 40g/day in women. As a rough guide, then, a shift in alcohol consumption of ≥ 20g/day could be considered potentially clinically significant.

[1] European Medicines Agency, Committee for Medicinal Products for Human Use (CHMP), "Guideline on the development of medicinal products for the treatment of alcohol dependence (EMA/CHMP/EWP/20097/2008)", 18 February 2010.
[2] Sadock BJ, Sadock VA. Kaplan and Sadock's Comprehensive textbook of psychiatry. Philadelphia: Lippincott Williams & Wilkins, 2004; Miller WR, Walters ST, Bennett ME. How effective is alcoholism treatment in the United States? J Stud Alcohol 2001; 62: 211-220.
[3] Gual A, Lligoña A, Colom J. Five-year outcome in alcohol dependence. A naturalistic study of 850 patients in Catalonia. Alcohol. 1999; 34: 183-192.

Table 1. Drinking Risk Levels (DRLs) of Alcohol Consumption

DRL	Total Alcohol Consumption (g/day)	
	Men	Women
Low	>1 and ≤40	>1 and ≤20
Medium	>40 and ≤60	>20 and ≤40
High	>60 and ≤100	>40 and ≤60
Very high	>100	>60

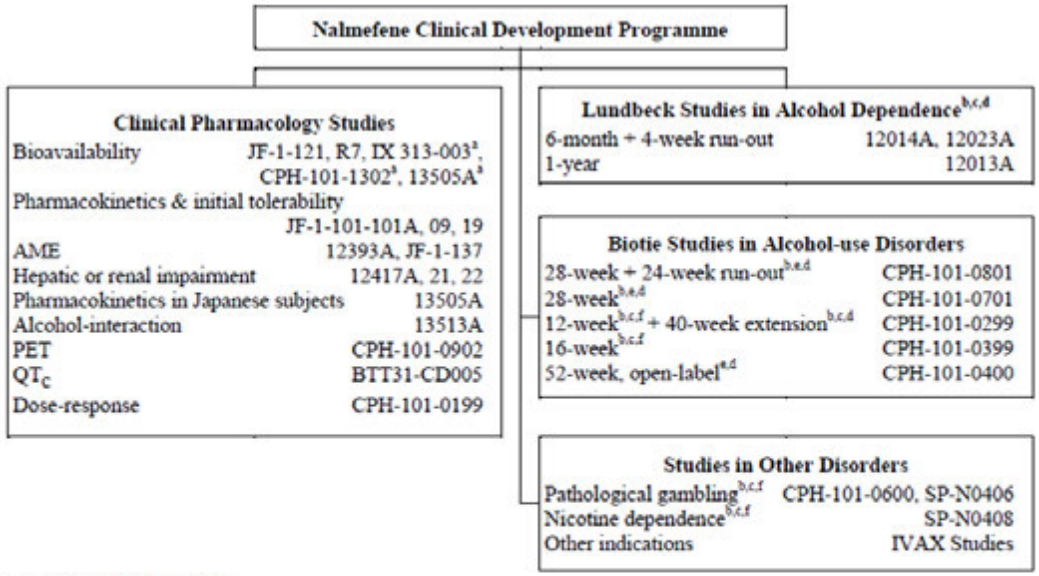
International Guide for Monitoring Alcohol Consumption and Related Harm[3]

3. Contents of the clinical dossier

3.1. Scope of the clinical dossier

The panel below shows the relationship between the submitted studies. The submitted dossier contained 17 clinical pharmacology studies (13505A is listed twice below), 3 major studies in Alcohol Dependence performed by the current Sponsor, Lundbeck, two of which can be considered pivotal, and 5 studies in Alcohol-Use Disorders by a previous Sponsor, Biotie. A number of older efficacy studies for other indications are also listed in the panel, but are only relevant to the assessment of safety.

Figure 1. Overview of Nalmefene Clinical Development Program



a Including food effect
b Randomised, double-blind, placebo-controlled
c Fixed-dose
d As-needed dosing
e Flexible-dose
f Daily dosing

The submission contained the following clinical information:

- 17 clinical pharmacology studies, all of which included some pharmacokinetic data; 3 provided pharmacodynamic data (an alcohol interaction study, a PET study and a dose-response study), and 1 was primarily a safety study (the QTc study).
- 1 population pharmacokinetic analysis.
- 2 pivotal Lundbeck efficacy studies.

- 1 supportive Lundbeck safety/efficacy study.
- 5 Biotie studies in Alcohol-use Disorders.
- Literature References.
- A tabulated Summary of all clinical studies.

3.2. Paediatric data

The submission did not included paediatric data.

3.3. Good clinical practice

The Lundbeck studies, including the pivotal efficacy studies, included statements of compliance with Good Clinical Practice (GCP). Several older studies by previous sponsors failed to declare compliance with GCP, and a couple of very early PK studies were clearly non-compliant with GCP – for instance, they lacked any discussion of safety.

Overall, compliance with GCP was declared in the studies that mattered. Regrettably, compliance with GCP in the conduct of these studies does not necessarily mean that appropriate practices were followed in the reporting of the results.

The ICH Guidelines for GCP includes the following statement about the need for prospectively identified endpoints:

> *"6.4 Trial Design*
>
> *The scientific integrity of the trial and the credibility of the data from the trial depend substantially on the trial design. A description of the trial design, should include:*
>
> > *6.4.1 A specific statement of the primary endpoints and the secondary endpoints, if any, to be measured during the trial."*

The study protocols contained clear statements of the primary endpoints and the proposed method of analysis, but the Sponsor's submission worked against the intent of the ICH guidelines by shifting the focus from prospectively stated endpoints to post hoc analyses. This was most marked in the proposed PI, where the prospective endpoints for the pivotal studies were not even mentioned, and instead p-values for a post hoc subgroup were provided.

4. Pharmacokinetics

4.1. Studies providing pharmacokinetic data

Tables 2 and 3 show an overview of studies.

Table 2. Overview of Clinical Pharmacology Studies

Study Year[a]	Type of Study Study Design	Treatment Duration	Tablet	Doses[b](mg)	Number of Subjects Treated
Clinical Pharmacokinetic and Initial Tolerability Studies					
JF-1-101-101A 1984	*Single-ascending-dose* Randomised, double-blind, parallel-group, placebo-controlled, single-dose	1 single dose (20 subjects) 2 single doses (4 subjects)	NA	NMF: 2, 6, 12mg i.v. 24mg i.v. or i.m. PBO: i.v.	24
09 1989	*Single-ascending-dose* Open-label, parallel-group, single-dose	3 single doses	NA	NMF: 0.5, 1, 2mg, i.v.	18
19 1992	*Single-ascending-dose/elderly* Open-label, parallel-group, single-dose	3 single doses	NA	NMF: 0.5, 1, 2mg, i.v.	36
Absorption, Metabolism, and Excretion (AME) Studies					
12393A 2010	*AME* Open-label, single-dose	1 single dose	NA	^{14}C-NMF: 20mg, oral solution	6
JF-1-137 Before May-1986	*AME* Open-label, single-dose	2 single doses	NA	^{14}C-NMF: 5mg, i.v. 200mg, oral solution	4
Effect of Intrinsic Factors					
13505A 2011	*Pharmacokinetics in Japanese subjects/food effect* Randomised, double-blind, placebo-controlled, single-dose and multiple-ascending-dose	Part A: 2 single doses Part B: 5 days	PCtab	Part A: NMF: 20mg PBO Part B: NMF: 20 or 40mg PBO	Part A: 13 Part B: 81
Effects of Diseases					
12417A 2010	*Hepatic impairment* Open-label, three-group (healthy or mild or moderate hepatic impairment), single-dose	1 single dose	PCtab	NMF: 20mg	24
21 1993	*Hepatic impairment* Open-label, four-group (healthy or mild or moderate or severe hepatic impairment), single-dose	1 single dose	NA	NMF: 2.0mg, i.v.	24
22 1993	*Renal impairment* Open-label, single-dose, historical control (healthy subjects) from Study 21	2 single doses	NA	NMF: 1.0mg, i.v.	9

NMF = nalmefene hydrochloride; PBO = placebo; AME = absorption, metabolism, and excretion; PET = positron emission tomography; i.v. = intravenous; p.o. = oral; i.m. = intramuscular

Study Year[a]	Type of Study Study Design	Treatment Duration	Tablet	Doses[b](mg)	Number of Subjects Treated
Alcohol-interaction Study					
13513A 2011	*Drug-drug interaction* Randomised, double-blind, four-period crossover (NMF+ethanol; NMF+PBO; PBO+ethanol; PBO+PBO), placebo-controlled, single-dose	4 single doses	PCtab	NMF: 20mg Ethanol: 0.6g/kg PBO	46
Pharmacodynamic Study					
CPH-101-0902 2002	*PET/occupancy* Open-label, two-period	1 single dose + 7 days	PCtab	NMF: 20mg	12
QTc Study					
BTT31-CD005 2008	*QT_c* Randomised, double-blind, parallel-group, placebo-controlled, multiple-dose, moxifloxacin-controlled	7 days	PCtab	NMF: 20, 40, 80mg PBO Moxifloxacin: 400mg for 1 day (Day 7)	270
Other					
CPH-101-0199 1999	*Dose response and efficacy* Randomised, parallel-group, placebo-controlled	2 single doses (same drug and dose)	Ultab, PCtab	NMF: 20, 60mg Naltrexone: 50mg PBO	32

NMF = nalmefene hydrochloride; PBO = placebo; AME = absorption, metabolism, and excretion; PET = positron emission tomography; i.v. = intravenous; p.o. = oral; i.m. = intramuscular

a Year of last subject/patient last visit

b Oral tablets, unless otherwise indicated. Doses are based on nalmefene hydrochloride; 20mg nalmefene hydrochloride corresponds to 18.1mg nalmefene base.

Table 3. Submitted pharmacokinetic studies.

PK topic	Subtopic	Study ID	*
PK in healthy adults	General PK - Single dose oral	JF-1-121	*
	IV, oral	R7	
	IV, IM	JF-101-101A	
	IV	09	*
	Multi-dose	13505A	*
		CPH-101-0902	
		BTT31CD005	
	Pooled Analysis of Multiple PK Studies	14019A	*
	Bioequivalence† - Single dose. Tablet vs Solution	JF-1-121	*
		IX 313-003	
	Bioavailability IV vs Oral	R7	*
	Mass Balance	12393A	*
		JF-1-137	*
	Food effect	IX-313-003	*
		CPH-101-1302	*
		13505A	
PK in special populations	Target population § - Single dose	-	
	Multi-dose	-	
	Hepatic impairment	12417A	*
		21	*
	Renal impairment	22	*
	Neonates/infants/children/adolescents	-	
	Elderly	19	*

PK topic	Subtopic	Study ID	*
Genetic/gender related PK	Males versus females	13505A	*
	Race	13505A	*
PK interactions	Ethanol	13513A	*
Population PK analyses	Healthy subjects	12735A	*

* Indicates the primary aim of the study.
† Bioequivalence of different formulations.
§ Subjects who would be eligible to receive the drug if approved for the proposed indication.

None of the pharmacokinetic studies had such severe deficiencies that they were excluded from consideration, but the bioavailability study, R7, was inadequate in the context of the current submission, because it was small, with only four subjects receiving matching IV and oral doses, and the doses administered were well below the proposed dose. Also, it produced estimates of bioavailability that were markedly inconsistent with a pooled analysis of other studies. The Sponsor has since argued that the pooled analysis is likely to have been more reliable.

4.2. Summary of pharmacokinetics

The information included in the proposed PI and in the summary below is primarily derived from an integrated PK analysis (Study 14019A) which was in turn derived from source data in the original conventional pharmacokinetic studies, pooled and reanalysed by the current Sponsor.

4.2.1. Pharmacokinetics in healthy subjects

The main PK parameters of nalmefene are show in the table below, which represents the integrated PK data from all oral dose studies in fasting, healthy volunteers, recalculated from the source data by the current Sponsor (Study 14019A). The values cited for tmax, t½ and AUC are broadly consistent with the results of the individual studies from which these estimates were derived. The proposed PI uses these recalculated values, which is appropriate.

Table 4. Mean Dose-normalised Pharmacokinetic Parameters Derived from the Integrated Pharmacokinetic Analysis of Nalmefene Data — Study 14019A

	$dAUC_{0\text{-}tau,M\text{-}D}$ (ng·h/mL)	$dAUC_{0\text{-}inf,S\text{-}D}$ (ng·h/mL)	$dC_{max,S\text{-}D}$ (ng/mL)	$dC_{max,M\text{-}D}$ (ng/mL)	t_{max} (h)	$t_{½}$ (h)	CL/F (L/h)	V_Z/F (L)
Nalmefene								
N	135	56	57	135	192	222	191	176
Mean	154	131	16.5	22.3	1.5	12.5	169	3200
SD	48.9	39.6	6.63	9.10	0.5 - 6.0	5.41	286	3770
Nalmefene 3-O-conjugates								
N	12	41	41	12	53	60	NA	NA
Mean	915	1010	165	186	1.0	11.7	NA	NA
SD	232	507	97.2	62.0	0.5 - 3.0	2.33	NA	NA
Nornalmefene								
N	135	14	41	135	176	81	NA	NA
Mean	33.3	29.7	2.36	3.24	2.0	31.9	NA	NA
SD	13.1	10.0	0.862	1.26	0.5 - 4.25	18.1	NA	NA
Nornalmefene 3-O-conjugates								
N	12	32	40	12	52	52	NA	NA
Mean	195	221	14.9	17.5	3.0	16.3	NA	NA
SD	50.4	60.3	4.53	4.17	1.5 - 6.0	4.91	NA	NA

Median and range (min-max) are presented for t_{max}.
Estimation of $t_{½}$ is based on oral single and multiple dosing and i.v. dosing of nalmefene.
S-D = single dose; M-D = multiple dose; SD = standard deviation; NA = not applicable

a Based on data from biopharmaceutic and clinical pharmacology studies of subjects in fasting state, with normal renal and hepatic function, who were <67 years of age. All estimated pharmacokinetic parameters, except for $t_{½}$, are based on data after single and multiple oral dosing of nalmefene. The data have been dose-normalised to 20mg nalmefene.

4.2.1.1. Absorption

As shown in the above table, nalmefene is rapidly absorbed after a single oral administration of 18 mg (20mg of nalmefene hydrochloride), reaching a peak concentration (Cmax) of 16.5 ng/ml after approximately 1.5 hours, producing an exposure (AUCinf) of 131 ng·h/mL.

Based on population pharmacokinetic modelling, the absorption of nalmefene from tablets can be described by a first-order rate constant of approximately 0.75/h.

Two mass-balance studies with oral radio-labelled 14C-nalmefene have been conducted in human subjects, but the first (JF-1-1137) was technically inadequate because of poor sensitivity in the radio profiling of samples, so absorption data are primarily derived from the second study (12393A).

Most of a radiolabelled dose is absorbed from the gut, with only a small proportion of radioactivity (~20%) recovered from faeces. Following a single oral dose of 20mg 14C-nalmefene, mean total recovery of drug-related material was 91% at 240 hours post-dose, with 71% of the total radioactivity recovered from urine, on average, compared to 20% (CV 2.6%) in faeces.

4.2.1.1.1. Sites and mechanisms of absorption

The precise sites and mechanisms of absorption were not discussed by the Sponsor.

4.2.1.2. Bioavailability

4.2.1.2.1. Absolute bioavailability

The adequacy of data for estimation of absolute bioavailability was discussed by Lundbeck and the TGA prior to submission. During initial pre-submission planning, the TGA noted that Lundbeck had not clearly indicated whether an absolute bioavailability study had been performed.

Lundbeck responded as follows:

> *Lundbeck will provide study R7 (21,266- R7), a dedicated absolute oral bioavailability study, and its associated PK report. This study, which is entitled "Phase I Rising Dose Tolerance and Oral Bioavailability Study of Nalmefene in Healthy Volunteers", is referred to in Module 2.7.1 and the Module 5 Table of Contents that were included with the PPF submission dated 22 July 2013. The absolute oral bioavailability of nalmefene has been estimated in this study where subjects were dosed single intravenously and orally in a crossover design. Furthermore, Study 12735A provides an additional estimate of oral absolute bioavailability, which was determined from PK modelling that included the results from R7 and eight other studies. Of these eight studies four included intravenous dosing and contributed to the population model and the measure of absolute bioavailability. A value of the absolute oral bioavailability was calculated as 41%, which is reported in the clinical overview, clinical summaries and the draft Product Information.*

In the draft PI, the absolute oral bioavailability of nalmefene is described as 41%. This value is derived from pooling Study R7 with eight additional studies, rather than directly derived from a conventional bioavailability study.

Study R7, performed in 1983 by Key Pharmaceuticals, was not sufficiently robust to provide accurate data on bioavailability. It involved two small groups of 4 subjects each. Group 1 (n = 4) received 0.5 and 2.0 mg IV nalmefene, and 2.0, 8.0 and 32.0 mg of oral nalmefene and placebo. Group 2 (n = 4) received 1.0 and 2.0 mg IV nalmefene, and 4.0, 16.0 and 64.0 mg of oral nalmefene and placebo. Thus, only four subjects (in Group 1) received the same dose of IV and oral nalmefene, and the maximum dose received IV was one tenth of that proposed for clinical use. Estimates of bioavailability were performed 23 years after the original study, in a reanalysis by Biotie, and did not produce stable estimates in the range suggested by the PI:

> *When the dose-normalized values for area under the concentration by time curve until infinity after po and iv dosing were compared subject-wise, the bioavailability was 60-72% in three subjects and seemed 120% in one. [R7 Pharmacokinetic Report]*

That is, bioavailability seemed to range from 60-120%, and was not similar to 41% in any subject.

Because of the low intravenous dose used in Study R7, plasma nalmefene levels rapidly fell below the limit of quantification, so Biotie conceded that the terminal elimination phase was probably not observed, adding further doubt to dose-normalised estimates of exposure. Finally, this study was performed with earlier formulations of nalmefene, not the one proposed for release, and no relevant studies of relative bioavailability have been performed with the ITBM formulation.

Thus, while 41% is the Sponsor's current best estimate for bioavailability, the evidence for this value is indirect and the estimate should be considered uncertain.

4.2.1.2.2. Bioavailability relative to an oral solution or micronised suspension

The bioavailabilities of oral solution and oral tablet are similar. In the population PK analysis based on 9 Phase I studies, the absorption rate constant for nalmefene as an oral solution (1.40/h) was substantially higher than that estimated for an oral tablet (0.751/h).

4.2.1.2.3. Bioequivalence of clinical trial and market formulations

The formulation used in the pivotal Lundbeck studies is ostensibly the same as the formulation proposed for the market.

4.2.1.2.4. Bioequivalence of different dosage forms and strengths

The Sponsor is only proposing one dosage form and strength. Earlier PK studies were performed with different doses and formulations, and can only be considered approximately equivalent to the formulation proposed for clinical use.

4.2.1.2.5. Bioequivalence to relevant registered products

Not applicable.

4.2.1.2.6. Influence of food

Different PK studies yielded slightly different estimates of the possibility of a significant food effect. In one study (Study CPH-101-1302), co-administration with high-fat food increased the total exposure (AUC) to nalmefene by 30% and increased the peak concentration (Cmax) by ~50%, but delayed the time to peak concentration (tmax) by ~30 min. In a study of Japanese men (Study 13505A), no substantial food effect was noted. In the population PK analysis, food intake significantly increased oral bioavailability, by ~30%.

Given the acceptable tolerability of higher doses of nalmefene, this food effect is unlikely to be clinically significant.

4.2.1.2.7. Dose proportionality

Dose proportionality was assessed in Study BTT31-CD005, in which 20mg and 80mg nalmefene were administered orally, and in Study 13505A, in which 20mg and 40mg nalmefene were administered orally for up to 5 days.

In Study BTT31-CD005, increasing the dose 4-fold from 20mg to 80mg increased Cmax ~4.4 fold, and increased AUC0-tau ~4.3 fold. In Study 13505A, body-weight adjusted oral clearance was similar for 20 mg and 40mg in Japanese men, Japanese women, and Caucasian women. A minor increase of 30% in body weight-adjusted clearance was observed between the 20mg and 40mg dose group in Caucasian men. Thus, overall, oral doses at and above the recommended dose range show approximate dose proportionality.

Intravenously administered nalmefene at doses between 0.5 and 24mg also showed approximately linear dose proportionality (Studies JF-1-101-101A, 09, and 19).

4.2.1.2.8. Bioavailability during multiple-dosing

As discussed above, no direct, absolute bioavailability study of nalmefene was submitted, apart from an old study, R7, which produced unreliable results. The PK of multiple doses of nalmefene were compared to single dose PK in Study 13505A and there were no important differences.

4.2.1.2.9. Effect of administration timing

Apart from the food effect, no studies assessed the effects of timing of administration. There are no a priori reasons for suspecting a substantial variation in PK at different administration times.

4.2.1.3. Distribution

4.2.1.3.1. Volume of distribution

Distribution was estimated in the pooled analysis of the Phase I studies. The estimated volume of distribution (Vz/F) was approximately 3200 L (SD 3770 L), indicating that nalmefene undergoes extensive and highly variable extravascular distribution.

4.2.1.3.2. Plasma protein binding

The average protein-bound fraction of nalmefene in plasma was estimated to be ~29%, based on administration of radiolabelled nalmefene (Study BTT31-AD036). This fraction remained consistent over a range of concentrations up to 400ng/mL.

4.2.1.3.3. Erythrocyte distribution

Based on administration of radiolabelled nalmefene, and subsequent comparison of the radioactivity of whole blood versus plasma, some nalmefene-related compounds appear to be excluded from red blood cells. The mean exposure to total radioactivity in plasma (based on AUC0-inf) was approximately 1.5 times that observed in whole blood (with individual ratios ranging from 1.2 to 1.9; see Study 19).

4.2.1.3.4. Tissue distribution

In a positron emission tomography (PET) study involving single and repeated daily dosing with 20mg nalmefene (Study CPH-101-0902), PET signals were consistent with 94-100% receptor occupancy within 3 hours after dosing, suggesting that nalmefene readily crosses the blood-brain barrier.

Based on the volume of distribution (~3200 L), nalmefene is extensively distributed to extravascular tissues, but no clinical study assessed the specific sites of distribution.

4.2.1.4. Metabolism

The primary mechanism of clearance of nalmefene is by glucuronide conjugation, followed by renal excretion of metabolites along with some unaltered nalmefene.

The oral clearance of nalmefene (CL/F) was estimated to be ~169 L/h. The terminal half-life was estimated to be 12.5 hours.

4.2.1.4.1. Metabolites identified in humans

Metabolism of nalmefene is extensive, and includes a number of common biotransformations, such as hydroxylation, N-dealkylation, glucuronic acid conjugation, and sulphation. The tables below list the metabolites recovered from faces (Table 5) and plasma (Table 6).

In plasma, nalmefene only accounted for ~4.5% of the total exposure (AUCinf) to nalmefene-related compounds (Study 12393A). In total, 9 metabolites were quantified in human plasma, and the predominant metabolite in plasma was nalmefene 3-O-glucuronide (Study 13081).

Table 5. Nalmefene and Metabolites in Pooled Urine and Faeces from Healthy Men (n = 6) following a Single Oral Dose of 20mg 14C-Nalmefene — Study 13081

Metabolite	Metabolite ID	% of Administered Dose		
		Urine[a]	Faeces[b]	Urine+Faeces
Total of drug-related material (%[c])		70.16	18.83	88.99
Nornalmefene 3-O-glucuronide	M1	3.20	-	3.20
Nornalmefene 3-O-sulphate	M2	1.77	-	1.77
Nornalmefene	M3	2.17	5.37	7.54
Glucuronide of hydroxy nalmefene	M4	-	0.36	0.36
Glucuronide of hydroxy nalmefene	M5	1.67	0.66	2.34
Nalmefene 3-O-glucuronide	M6	54.38	0.11	54.49
Unidentified	M7	-	1.48	1.48
Nalmefene 3-O-sulphate	M8	-	0.40	0.40
Glucuronide of nalmefene	M9	-	0.13	0.13
Nalmefene		2.83	2.62	5.46
Others		-	0.47	0.47
Total		66.03	11.62	77.64

a Collection interval: 0-96 hours
b Collection interval: 0-144 hours
c The total concentration of radiolabelled drug-related material was determined in Study 13235.

Table 6. Nalmefene and Metabolites in Pooled Plasma from Healthy Men (n = 6) following a Single Oral Dose of 20mg 14C-Nalmefene — Study 13081

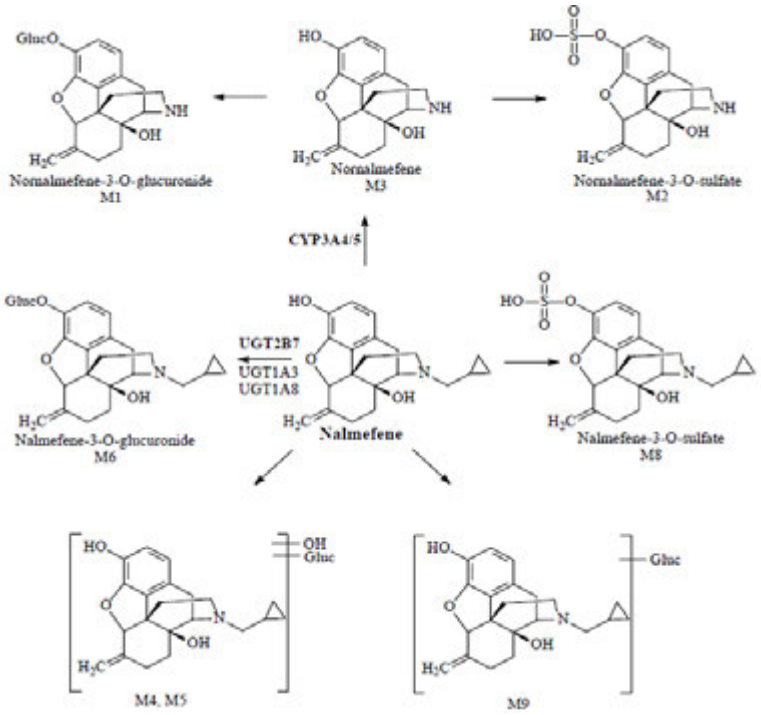
Metabolite	Metabolite ID	Concentration (ng equivalent/mL)								
		1 h	2 h	3 h	4 h	6 h	12 h	24 h	48 h	72 h
Total of drug-related material[a]		199	196	152	110	72	61	29	13	5
Nornalmefene 3-O-glucuronide	M1	2	4	4	4	2	2	-	-	-
Nornalmefene 3-O-sulphate	M2	4	7	8	7	6	4	2	-	-
Nornalmefene	M3	3	3	2	1	1	1	-	-	-
Glucuronide of hydroxy nalmefene	M4	2	3	2	2	1	-	-	-	-
Glucuronide of hydroxy nalmefene	M5	8	8	7	4	3	3	1	-	-
Nalmefene 3-O-glucuronide	M6	161	157	115	82	51	44	18	8	4
Nalmefene 3-O-sulphate	M8	1	1	1	-	-	-	-	-	-
Glucuronide of nalmefene	M9	1	1	-	-	-	-	-	-	-
Unidentified	M10	2	1	1	-	-	-	-	-	-
Nalmefene		13	9	7	5	4	4	2	-	-
Total		196	193	148	105	69	57	23	8	4

a The total concentration of radiolabelled drug-related material was determined in Study 13235.

4.2.1.4.2. Sites of metabolism and mechanisms / enzyme systems involved

As illustrated in the figure below, the main enzyme systems involved in the metabolism of nalmefene are CYP3A4/5 and UGT2B7, with a minor contribution from UGT1A3 and UGT1A8.

Figure 2. Biotransformation of Nalmefene



The major metabolite is nalmefene 3-O-glucuronide, produced by the UGT2B7 enzyme (with contributions by UGT1A3 and UGT1A8). A small proportion of nalmefene is converted to nalmefene 3-O-sulphate by sulphation, and another small proportion is converted to nornalmefene by CYP3A4/5. Nornalmefene undergoes similar conversions, producing nornalmefene 3-O-glucuronide and nornalmefene 3 - O-sulphate.

Two unknown metabolites have also been identified: M7 found in faeces and M10 found in plasma. These were found at low levels, as shown in the tables above.

4.2.1.4.2.1. Active metabolites

The only significant active metabolite of nalmefene is nalmefene 3 - O-sulphate. It has a potency at opioid receptors that is similar to nalmefene, but it is present in relatively low concentrations (<10% of nalmefene concentrations) and therefore makes only a small contribution to the overall pharmacological effect.

4.2.1.4.2.2. Other metabolites

The other metabolites identified (figure and tables above) do not appear to contribute significantly to the overall pharmacological effect on opioid receptors.

4.2.1.4.3. Interconversion between enantiomers

Interconversion between enantiomers does not appear to occur to an appreciable extent. The chiral centre of nalmefene derives its configuration from its natural source, and this appears to be stable.

4.2.1.4.4. Pharmacokinetics of metabolites

The PK of the major nalmefene metabolites are summarised in the table below.

Table 7. Pharmacokinetic Parameters of Nalmefene and Metabolites in Plasma following a Single Oral Dose of 20mg Nalmefene - Study 12417A

Parameter	Healthy Subjects (N=8)	Mild Hepatic Impairment (N=8)	Moderate Hepatic Impairment (N=8)
Nalmefene			
AUC_{0-t} (ng·h/mL)	94.7 (29.5)	141 (17.5)	293 (40.5)
AUC_{0-inf} (ng·h/mL)	97.4 (28.5)	145 (16.7)	297 (39.9)
C_{max} (ng/mL)	15.1 (34.5)	16.4 (27.3)	25.3 (38.3)
t_{max} (h)	1.0 (1.0, 1.0)	1.0 (1.0, 2.0)	1.0 (1.0, 3.0)
$t_{½}$ (h)	11.1 (13.7)	12.3 (22.3)	11.5 (30.8)
CL/F (L/h)	220 (28.7)	142 (20.0)	89.0 (78.5)
V_z/F (L)	3465 (25.7)	2477 (24.0)	1258 (37.5)
Nalmefene 3-O-conjugates			
AUC_{0-t} (ng·h/mL)	1179 (19.7)	1656 (46.8)	1183 (30.2)
AUC_{0-inf} (ng·h/mL)	1199 (19.5)	1682 (46.4)	1200 (30.5)
C_{max} (ng/mL)	195 (19.0)	209 (30.8)	141 (64.7)
t_{max} (h)	1.0 (1.0, 2.0)	1.0 (1.0, 2.0)	2.0 (1.0, 4.0)
$t_{½}$ (h)	10.1 (15.9)	10.8 (20.4)	11.2 (30.7)
MR	12.8 (20.0)	11.6 (38.9)	5.63 (104)
Nornalmefene			
AUC_{0-t} (ng·h/mL)	14.6 (34.9)	17.1 (39.5)	19.4 (39.2)
AUC_{0-inf} (ng·h/mL)	20.9 (27.4)[a]	23.4 (18.1)[b]	23.4 (37.9)
C_{max} (ng/mL)	1.60 (24.9)	1.32 (44.2)	1.57 (36.8)
t_{max} (h)	1.0 (1.0, 1.0)	1.0 (1.0, 2.0)	1.0 (1.0, 3.0)
$t_{½}$ (h)	25.8 (35.1)[b]	22.6 (23.0)[a]	17.7 (32.7)
MR	0.221 (35.5)[b]	0.163 (20.1)[a]	0.088 (39.7)
Nornalmefene 3-O-conjugates			
AUC_{0-t} (ng·h/mL)	156 (25.4)	182 (50.1)	83.7 (54.9)
AUC_{0-inf} (ng·h/mL)	179 (18.6)	198 (50.2)	96.7 (50.2)
C_{max} (ng/mL)	13.9 (24.1)	12.0 (36.5)	5.92 (74.3)
t_{max} (h)	3.0 (2.0, 3.0)	3.0 (2.0, 6.0)	3.5 (1.0, 6.0)
$t_{½}$ (h)	13.5 (33.0)	13.5 (44.7)	12.2 (22.8)
MR	1.92 (23.9)	1.36 (43.8)	0.50 (122)

Arithmetic mean (CV%) data are presented for AUC_{0-t}, AUC_{0-inf}, C_{max}, $t_{½}$, CL/F, V_z/F, and MR.
Median (min, max) data are presented for t_{max}.
N = number of subjects; MR = metabolic ratio
a N = 6
b N = 7

4.2.1.4.5. Consequences of genetic polymorphism

The Sponsor did not address the possibility of genetic polymorphism affecting the PK of nalmefene.

4.2.1.5. Excretion

4.2.1.5.1. Routes and mechanisms of excretion

Nalmefene metabolites are predominantly excreted by the renal route, along with a small amount of unchanged parent compound (< 3% of administered dose), but some non-renal elimination also occurs. About half (54%) of the total administered dose is excreted in urine as nalmefene 3 - O-glucuronide, whereas other metabolites are present in the urine in much lower amounts (< 3% each).

As noted below, faecal elimination accounts for about 19% of total excretion.

4.2.1.5.2. Mass balance studies

In total, compounds collected from faeces account for ~19% of an administered radioactive dose, compared to urine which accounts for 70% (leaving 11% unrecovered, Study 12393A).

Nornalmefene was the predominant metabolite recovered from faeces, accounting for about 5% of an administered dose. Conjugated compounds, in general, were much more likely to be eliminated renally.

In an earlier study (JF-1-137), recovery from faeces was slightly greater: following oral dosing with 14C-nalmefene, overall recovery of drug-related material was 74%, with 49% recovered from urine and 25% from faeces over a period of 7 days. In the same study, 36 to 69% of the total dose was excreted in urine and 17 to 18% in faeces after intravenous dosing. Thus, overall, non-renal clearance of metabolites plays a relatively minor role.

4.2.1.5.3. Renal clearance

Renal clearance accounts for about 70% of an administered dose of nalmefene, with the main compound in urine being nalmefene 3-O-glucuronide, which accounts for 54% of an administered dose. Clearance of nalmefene from the plasma does not depend closely on renal function, because the primary mechanism of plasma clearance of the parent compound is hepatic metabolism, but studies in the setting of severe renal impairment show evidence of delayed clearance.

4.2.1.6. Inter- and intra-individual variability of pharmacokinetics

The inter- and intra-individual variability of the main PK parameters was estimated in the integrated PK analysis (Study 14019A), and is shown below expressed as the percent coefficient of variation (CV%). The variability was up to ~45% for the inter-subject variability (first value in each pair in the table), and up to ~31% for the intra-subject variability, depending on the PK parameter under consideration. This is an acceptable degree of variability, unlikely to have significant clinical consequences given the broad therapeutic ratio of nalmefene.

Table 8. Inter- and Intra-subject Variability (CV%) of Pharmacokinetic Parameters of Nalmefene, Nornalmefene and their Conjugates following Oral Administration of Nalmefene - Study 14019A

	Inter-/Intra-subject Variability			
Parameter	**Nalmefene**	**Nalmefene Conjugates**	**Nornalmefene[a]**	**Nornalmefene Conjugates**
$AUC_{0\text{-}inf}$ (ng·h/mL)	36.9/9.9	27.6/12.4	28.5/NA	39.7/8.8
C_{max} (ng/mL)	44.2/30.7	40.3/22.7	41.8/25.8	40.6/12.7
$t_{½}$ (h)	29.5/15.2	21.6/11.5	45.3/36.8	29.1/23.4
CL/F (L/h)	36.9/9.8	NA	NA	NA

NA = not applicable

a The results for $AUC_{0\text{-}inf}$ and $t_{½}$ are somewhat uncertain since the LLOQs were high compared to the observed concentrations.

4.2.2. Pharmacokinetics in the target population

Pharmacokinetics in the target population were not directly assessed. A population-PK study was submitted, but it was based on results in healthy volunteers.

Subjects with alcohol dependence are at increased risk of hepatic impairment, which increases exposure to nalmefene as discussed below. It is unclear if chronic alcohol use modifies the activity of the hepatic enzyme systems involved in the metabolism of nalmefene, and this issue should be addressed by the Sponsor. Acute exposure to alcohol does not appear to modify exposure to nalmefene (but, importantly, nalmefene appears to increase exposure to alcohol by ~9%, with broad confidence limits for this PK effect that include the possibility of no interaction.

4.2.3. Pharmacokinetics in other special populations

The population-PK analysis assessed the effects of gender, age and ethnicity on the PK of nalmefene, finding no important effects. For central volume of distribution (V2), age emerged as a significant covariate, with V2 decreasing with age. Over the range of 18 to 80 years assessed, the estimated V2 decreased from 287 to 156 L. This is unlikely to be clinically important, given the acceptable tolerability of nalmefene at doses higher than that proposed for clinical use.

The main covariate affecting clearance (CL) of nalmefene was lean body mass (LBM), which showed a statistically significant effect. CL increased with LBM and, over the range of LBM assessed (37.3 to 74.2 kg), estimated CL ranged from 46.7 to 71.8 L/h. This is unlikely to be clinically important.

4.2.3.1. Pharmacokinetics in subjects with hepatic impairment

Two studies have been performed assessing the PK of nalmefene in subjects with hepatic impairment: in an early study (Study 21, n = 12 with impairment, n = 12 controls), intravenously administered nalmefene (2.0mg) was administered to patients with mild, moderate, or severe hepatic impairment and to normal subjects; in a later study (12417A), oral nalmefene (20mg) was administered to patients with mild or moderate hepatic impairment, in comparison to normal subjects.

In the early, low-dose IV study, patients with mild hepatic impairment had an AUCinf that was similar to healthy subjects (34.3ng·h/mL), whereas patients with moderate or severe impairment had AUCinf that was 45% (49.8 ng·h/mL) or 51% (51.9ng·h/mL) higher, respectively. The elimination half-life was similar in all groups, but the total body clearance was 22% to 33% lower in patients with hepatic impairment, compared to healthy subjects.

The PI refers to data from the later study, which used the proposed 20mg oral dose. In patients with mild hepatic impairment, exposure (AUC) increased 1.5 times and oral clearance decreased by approximately 35%. In patients with moderate hepatic impairment, exposure increased 2.9 times for AUC and 1.7 times for Cmax, while oral clearance decreased by ~60%. This represents a more substantial effect than revealed with the earlier, low-dose study. No consistent changes were seen in tmax or elimination half-life for any of the groups, which is consistent with the earlier study.

Table 9. Statistical Analysis of the Effect of Hepatic Impairment on the Pharmacokinetics of Nalmefene - Study 12417A

	Mean Value (N=8)	(N=8)	Ratio	95% CI
Hepatic Function	**Mild Impairment**	**Normal**		
AUC_{0-t} (ng·h/mL)	141 (17.5)	94.7 (29.5)	1.52	(1.13, 2.05)
AUC_{0-inf} (ng·h/mL)	145 (16.7)	97.4 (28.5)	1.53	(1.15, 2.03)
C_{max} (ng/mL)	16.4 (27.3)	15.1 (34.5)	1.12	(0.75, 1.68)
$t_{½}$(h)	12.3 (22.3)	11.1 (13.7)	1.3	(-1.5, 4.1)
Hepatic Function	**Moderate Impairment**	**Normal**		
AUC_{0-t} (ng·h/mL)	293 (40.5)	94.7 (29.5)	2.88	(1.72, 4.84)
AUC_{0-inf} (ng·h/mL)	297 (39.9)	97.4 (28.5)	2.87	(1.73, 4.76)
C_{max} (ng/mL)	25.3 (38.3)	15.1 (34.5)	1.72	(1.13, 2.64)
$t_{½}$(h)	11.5 (30.8)	11.1 (13.7)	0.60	(-2.6, 3.9)
Hepatic Function	**Moderate Impairment**	**Mild Impairment**		
AUC_{0-t} (ng·h/mL)	293 (40.5)	141 (17.5)	1.89	(1.15, 3.10)
AUC_{0-inf} (ng·h/mL)	297 (39.9)	145 (16.7)	1.88	(1.16, 3.05)
C_{max} (ng/mL)	25.3 (38.3)	16.4 (27.3)	1.54	(1.02, 2.31)
$t_{½}$(h)	11.5 (30.8)	12.3 (22.3)	-0.7	(-4.1, 2.8)

Arithmetic mean (CV%) data are presented for AUC_{0-t}, AUC_{0-inf}, C_{max}, and $t_{½}$.

Because no study has assessed the proposed oral dose in subjects with severe hepatic impairment, the PI carries a warning saying "Pharmacokinetic data after oral administration of nalmefene to patients with severe hepatic impairment are not available." This is appropriate, given that the early study (which did include subjects with severe impairment, but used IV administration) assessed such a very low dose and produced under-estimates of the effect of moderate impairment, relative to the effect seen in the later study.

4.2.3.2. Pharmacokinetics in subjects with renal impairment

No study of the PK of oral nalmefene has been conducted in patients with renal impairment. The only submitted study dealing with the issue was an old study (Study 22, n = 8) that lacked a proper control group, and instead used the control group (n = 12) from the hepatic impairment study (Study 21). The dose administered in the renal patients and the controls was different: 1mg IV for the renal patients and 2mg IV for the controls, so the data required dose-normalisation, adding to the uncertainty of the results.

After dose-normalisation, administration of 1 mg nalmefene IV in patients with severe renal impairment (ESRD) resulted in a 1.6-fold larger exposure (dose-adjusted AUCinf), than in healthy subjects. Patients with renal impairment also had a lower Cmax (by a factor of ~ 2.1 to 4.6). The elimination half-life (~26 hours) was longer than that seen in healthy subjects (~10 hours).

Nalmefene conjugates showed an increase in Cmax and prolongation of t½, as shown in the table below.

Table 10. Dose-normalised Pharmacokinetic Parameters of Nalmefene and Conjugated Nalmefene in Patients with ESRD and in Healthy Subjects following a Single Intravenous Dose of 1.0 or 2.0mg Nalmefene - Studies 22 and 21

Parameter	Patients with ESRD (N=8)[b]		Healthy Subjects (N=12)[c]
	1 Day after Haemodialysis	4 Hours prior to Haemodialysis	
Nalmefene			
$AUC_{0\text{-}inf}$ (ng·h/mL)	28.4	27.5	17.2
C_{max} (ng/mL)	6.9	15.2	31.5
$t_{½}$ (h)	26.1	25.7	10.2
CL (L/h)	46.1[d]	44.0[d]	62.4[e]
Nalmefene Conjugates			
$AUC_{0\text{-}inf}$ (ng·h/mL)	384	235	44.7
C_{max} (ng/mL)	7.6	6.5	4.2
$t_{½}$ (h)	27.3	21.6	13.9

Arithmetic mean data are presented for $AUC_{0\text{-}inf}$, C_{max}, $t_{½}$, and CL.
N = number of subjects; ESRD = end-stage renal disease

a The data have been dose-normalised to 1.0mg nalmefene.
b Data from Study 22; 1.0mg nalmefene as an intravenous bolus (administered over 15 seconds)
c Data from Study 21; 2.0mg nalmefene as an intravenous bolus (administered over 15 seconds)
d Data from Study 22, transformed from L/h·kg to L/h (Study 14019A)
e Data from Study 21, transformed from L/h·kg to L/h (Study 14019A)

4.2.3.3. Pharmacokinetics according to age

No study has specifically assessed the PK of nalmefene with oral dosing in patients ≥ 65 years of age. One study with low-dose IV administration (Study 19) showed exposure that was similar to another study of IV administration in younger adults (Study 09). This suggests that there are no major PK changes in the healthy elderly population, but a direct comparison of younger and older subjects within the same study would be more appropriate. The Sponsor's initial version of the proposed PI falsely implied that such a study had been done, but the Sponsor has accepted a suggestion to remove this implication from the latest PI. The effect of age was also explored in the population-PK analysis, showing that volume of distribution declines with age, but this is unlikely to be clinically important.

4.2.3.4. Pharmacokinetics related to genetic factors

The Sponsor did not submit any studies investigating potential genetic differences in pharmacokinetics, apart from those relating to race and gender. In a study of Japanese and Caucasian men and women, no important differences were found in weight-adjusted PK (Study 13505A).

4.2.4. Pharmacokinetic interactions

4.2.4.1. Pharmacokinetic interactions demonstrated in human studies

Only one drug interaction study was performed, Study 13513A (n = 46), which was a randomised, double-blind, placebo-controlled, four-period crossover, single-dose study evaluating the potential PK and PD interactions between nalmefene and ethanol in healthy volunteers.

Each subject received four single-dose treatments:

- placebo tablet and placebo solution
- placebo tablet and ethanol solution 0.6g/kg
- nalmefene 20mg and placebo solution
- nalmefene 20mg and ethanol solution 0.6g/kg

The results are listed below, comparing the nalmefene-ethanol combination (NE) with the nalmefene-placebo combination (NP). The substitution of ethanol for placebo made little difference to the PK of nalmefene, and the 90%CIs for the ratio of NE parameters vs NP parameters were within the conventional bioequivalence interval (0.8 to 1.25), and were also broadly consistent with the inter-subject variability demonstrated in other studies. The greatest effect on nalmefene PK was observed with nalmefene Cmax, which increased by 22.5% when combined with ethanol.

Table 11. Pharmacokinetic Parameters of Nalmefene following a Single Oral Dose of 20mg Nalmefene Administered to Subjects Exposed to Ethanol (0.6g/kg) or Placebo Solution - Study 13513A

Parameter	Nalmefene plus Ethanol (0.6g/kg) or Placebo Solution	Plasma Nalmefene	Nalmefene plus Ethanol *versus* Nalmefene plus Placebo Solution Ratio (90% CI)
AUC_{0-t} (ng·h/mL)	NE	134 (24.3)[a]	1.037 (0.999, 1.076)
	NP	129 (26.1)[b]	
AUC_{0-inf} (ng·h/mL)	NE	141 (22.5)[c]	1.047 (1.012, 1.084)
	NP	135 (23.8)[d]	
C_{max} (ng/mL)	NE	18.4 (24.8)[a]	1.225 (1.096, 1.371)
	NP	15.2 (32.6)[b]	
t_{max} (h)	NE	1.50 (0.50, 2.50)[a]	
	NP	1.50 (0.50, 5.00)[b]	
$t_{½}$ (h)	NE	11.9 (18.2)[c]	
	NP	11.8 (18.6)[d]	
CL/F (L/h)	NE	150 (27.4)[c]	
	NP	157 (27.1)[d]	
V_z/F (L)	NE	2534 (23.0)[c]	
	NP	2647 (26.7)[d]	

Arithmetic mean (CV%) data are presented for AUC_{0-t}, AUC_{0-inf}, C_{max}, $t_{½}$, CL/F, and V_z/F.
Median (min, max) data are presented for t_{max}.
NE = nalmefene plus ethanol; NP = nalmefene plus placebo solution; N = number of subjects

a N=32
b N=41
c N=30
d N=39

Given the wide therapeutic window for nalmefene, the combination with alcohol does not pose any particular issues on the basis of nalmefene PK.

The study also assessed the PK of ethanol when combined with nalmefene (NE) versus ethanol by itself (PE). This analysis is more important than a traditional drug interaction study, because reducing ethanol exposure is the sole therapeutic purpose of nalmefene treatment, and even small increases in ethanol exposure could offset the therapeutic effect.

Table 12. Pharmacokinetic Parameters of Ethanol following a Single Oral Dose of 0.6g/kg Ethanol Administered to Subjects Exposed to 20mg Nalmefene or Placebo - Study 13513A

Parameter	Ethanol plus Nalmefene (20mg) or Placebo Tablet	Ethanol		Nalmefene plus Ethanol *versus* Placebo Tablet plus Ethanol (all subjects) Ratio (90% CI)
		Men	Women	
AUC_{0-t} (mmol·h/L)	NE	42.0 (25.6)[a]	52.1 (30.9)[a]	1.086 (0.977, 1.208)
	PE	38.4 (35.4)[b]	46.8 (37.2)[c]	
C_{max} (mmol/L)	NE	14.7 (19.5)[a]	18.1 (31.5)[a]	1.003 (0.913, 1.101)
	PE	15.2 (<2.2, 19.1)[b]	18.5 (24.8)[c]	
t_{max} (h)	NE	2.25 (1.50, 3.00)[a]	2.50 (1.50, 4.00)[a]	
	PE	2.00 (1.50, 3.00)[d]	2.00 (1.50, 3.00)[c]	

Arithmetic mean (CV%) data are presented for AUC_{0-t} and C_{max} (NE for men and women and PE for women).
Median (min, max) data are presented for t_{max} and C_{max} (PE for men).
NE = nalmefene plus ethanol; PE = placebo tablet plus ethanol; N = number of subjects; CI = confidence interval

a N=16 c N=21
b N=18 d N=17

As shown in the table above, combining nalmefene with ethanol appears to increase the AUC for ethanol by ~9%, with a 90%CI potentially consistent with an increase of up to 21% (ratio 1.086, 90%CI 0.977 to 1.208), but also potentially consistent with no change in alcohol exposure. The 95%CI would be expected to be broader still.

On any single drinking occasion, an increase in alcohol exposure of 9% would not be of particular concern, although it could shift a drinker over the legal limit for driving. As a chronic effect, though, this potentially represents a serious hidden efficacy cost for nalmefene, which was not mentioned once in the Sponsor's efficacy analysis. This issue is discussed further in the Efficacy section, and was also discussed in the Sponsor's Section 31 Response. The matter was also referred to an independent statistician, who agreed that the observed PK effect increases the uncertainty surrounding estimates of the efficacy of nalmefene.[4]

4.2.4.2. Clinical implications of in vitro findings

According to the Sponsor's summary of the preclinical in vitro studies, drug-interactions involving nalmefene and its metabolites are unlikely when nalmefene is coadministered with drugs metabolised by the most common CYP450 and UGT enzymes or membrane transporters.

On the other hand, coadministration of nalmefene with drugs that are potent inhibitors of the UGT2B7 enzyme system (diclofenac, fluconazole, medroxyprogesterone acetate, meclofenamic acid) could significantly increase exposure to nalmefene. Given the broad therapeutic index of nalmefene and good tolerability at higher doses, this interaction is not likely to present a clinically significant problem with occasional use of UGT2B7 inhibitors, but caution should be exercised with long-term concurrent use of these drugs.

Coadministration with a UGT inducer (dexamethasone, phenobarbital, rifampicin, omeprazole) would be expected to lower nalmefene exposure, which could lead to inadequate efficacy, but at least this would not pose safety concerns.

[4] The Sponsor has objected strongly to identification of this potental PK interaction as a potential hidden efficacy cost, calling it an 'assumption'. It is important to note that the first-round CER does not assume that the 9% increase in exposure would necessarily be replicated in future studies. The mean PK interaction over a larger population of drinkers could be favourable, neutral, or unfavourable, but the point is that the magnitude of this interaction is not currently known with adequate precision, and *the current mean estimate is unfavourable.* Thus it is the Sponsor who has made an assumption, working on the basis that there is a 0% interaction with no surrounding uncertainty bounds. The onus of prooof here lies clearly with the Sponsor.

4.2.5. Pooled PK analysis

The Sponsor performed a population PK analysis. The Sponsor also reanalysed the pooled data set derived from Lundbeck PK studies and earlier PK studies by other Sponsors. The proposed PI generally includes the values derived from this pooled analysis.

Table 13. Overview of Reports of Pooled Pharmacokinetic Data

Report Year[a]	Type of Report	NMF Doses (mg)	Number of Subjects Included
12735A 2009	Population PK modelling of nalmefene in healthy subjects	i.v.: 0.5-24mg; p.o.: 2-80mg (tablets and/or solution)	243 (i.v.: 86; p.o. tablets: 157; p.o. solution: 10; fasted state: 243, fed state: 16)
14019A 2011	Pooling and cross-study analysis of PK parameters	i.m.: 24mg; i.v.: 0.5-64mg; p.o.: 2-200mg (tablets and/or solution)	293 (single doses: i.m.: 4; i.v.: 106; p.o. tablets: 58; p.o. solution: 16; multiple doses: p.o. tablets: 135)

PK = pharmacokinetics; p.o. = oral; i.v. = intravenous; i.m = intramuscular

a Year of final report

4.3. Evaluator's conclusions based on the CER round 1

The PK of nalmefene has been reasonably well characterised in healthy volunteers, but three issues remain slightly unresolved.

Firstly, the PK of nalmefene in the target population has not been studied directly, though it might be expected from the hepatic-impairment studies that subjects with alcoholic liver disease would have increased exposure.

Secondly, there has been no adequate study of absolute bioavailability, which has instead been inferred from the pooled analysis of multiple Phase I studies. Study R7 was ostensibly submitted as a bioavailability study, but it produced results inconsistent with the pooled results.

Thirdly, and most importantly, there is preliminary evidence of a clinically relevant PK interaction between nalmefene and alcohol that appears to increase exposure to alcohol by ~9%. The 90%CI for the estimate of the interaction fell within the conventional bioequivalence range (0.8 to 1.25) but, when considered in the context of a drug administered with the sole intent of reducing alcohol exposure, the observed increase in exposure of 9% could be enough to compromise the proposed clinical benefit. Furthermore, it remains possible that the apparent efficacy observed with nalmefene partly consists of drinkers titrating their drinking to compensate for the increased exposure associated with nalmefene.

Currently, the uncertainty bounds for this estimate are broad, and the 90%CI includes the possibility of no effect, as well as the possibility that exposure to alcohol is increased by ~21% when drinking occurs after nalmefene administration; the 95%CI would be expected to be broader still. This could be more than enough to negate and even reverse the proposed benefits of nalmefene. In their original submission, the Sponsor did not comment on this issue or even appear to notice that an increase in exposure to alcohol has efficacy implications.[5]

[5] In their response to CER_1, the Sponsor not only tried to defend this omission, but explicitly insisted that the efficacy implications of this PK interaction should be ignored because the estimates fall within the conventional bioequivalence range. This reveals a curious double standard, because the Sponsor attempts to promote minor *decreases* in alcohol exposure (mediated by PD means) as a major clinical benefit, while simultaneously dismissing comparable *increases* in exposure (mediated by PK means) as insignificant.

5. Pharmacodynamics

5.1. Studies providing pharmacodynamic data

Three PD studies were submitted: a PET study assessing nalmefene occupancy of opioid receptors, a thorough QT study, and a "dose-response" study attempting to assess the influence of nalmefene on short-term drinking behaviour in volunteers.

Individual summaries of these three pharmacodynamic studies are presented. The PET study and QT study were acceptable and provided useful information. The dose-response study was performed by an earlier Sponsor and appeared to be of marginal value. It was a small study that assessed drinking behaviour in volunteers who were invited to participate in two drinking sessions, but given no particular incentive to curtail drinking. As shown by the modest results obtained in the pivotal studies, any effects of nalmefene on drinking behaviour are subtle, and unlikely to be revealed in a short term study of this nature. In fact, this study produced completely inconsistent results across the two drinking sessions, and so it should be rejected.

5.2. Summary of pharmacodynamics

5.2.1. Mechanism of action

No clinical studies clarified the MOA of nalmefene.

5.2.2. Pharmacodynamic effects

5.2.2.1. Primary pharmacodynamic effects

The only useful PD study assessing primary PD effects was the PET study that assessed opioid receptor occupancy. This study is described; it showed that occupancy of mu-opioid receptors was high (94 to 100%) 3 hours after dosing, and that the decline of mu-opioid receptor occupancy was delayed relative to the elimination of nalmefene and its metabolites from the blood stream. The occupancy was still high (83 to 100%) at 26 hours post-dose, and occupancy was still over 50% at 50 hrs post-dose, in all brain areas studied. Peak occupancy was similar after single and repeated dosing, indicating that nalmefene does not accumulate significantly in the brain across doses.

As noted above, the Sponsor also submitted a "dose-response" study, which attempted to investigate the effect of nalmefene on short-term drinking behaviour. This study was small and produced inconsistent results, so it adds little to the overall understanding of the efficacy of nalmefene in treating alcohol disorders.

5.2.2.2. Secondary pharmacodynamic effects

The pharmacodynamic effects of nalmefene on cardiac repolarisation were explored in the QT study, which is described in detail. No clinically significant effect on QT interval was detected with nalmefene.

5.2.3. Time course of pharmacodynamic effects

No useful PD data studied the time course of the pharmacodynamic effects, apart from the PET study described above. The rejected dose-response study had two different drinking sessions, but the results in the two sessions were so different that no conclusions can be drawn.

5.2.4. Relationship between drug concentration and PD effects

No specific data are available on the relation between nalmefene concentrations and the desire to consume alcohol or the ability of alcohol dependent subjects to refrain from drinking.

5.2.5. Genetic-, gender- and age-related differences in pharmacodynamic response

No Phase I PD studies assessed the pharmacodynamic response to nalmefene in sufficient detail to compare results based on age, gender or genetic differences. Subgroup analyses in the major efficacy studies addressed this to some degree.

5.2.6. Pharmacodynamic interactions

As an opioid antagonist, nalmefene would be expected to counteract both the positive and negative effects of opioids. Commencement of nalmefene could induce a narcotic withdrawal syndrome in subjects already taking significant doses of opioids, and should therefore be avoided in this context. Concurrent use of nalmefene could reduce the efficacy of opioids prescribed for pain, diarrhoea or other indications, creating difficulties in adjusting doses. The draft PI carries appropriate warnings in this regard.

5.3. Evaluator's conclusions based on the CER round 1

The mechanism of action of nalmefene was not clearly described in the initial submission, but the Sponsor has since provided additional information on the proposed MOA in their Section 31 Response.

The submitted PET study confirmed that nalmefene binds to opioid receptors, and the QT study confirmed that it has acceptable safety in terms of cardiac repolarisation.

An old dose-response study of nalmefene did not produce reliable results.

6. Dosage selection for the pivotal studies

The current Sponsor, Lundbeck, did not perform a dose-response study, and no dose-response study has been performed for the proposed indication of Alcohol Dependence in high-risk drinkers – all three pivotal studies used the same 20mg dose.

In a previous study by an earlier Sponsor, Biotie, which was performed for the indication of Alcohol-Use Disorder, nalmefene doses of 10mg and 40mg were assessed, and an apparent treatment effect was achieved for the 40mg dose but not the 10mg dose (Study CPH-101-0399). Tolerability at 40mg was not ideal. The overall incidence of AEs was the same at both doses (92%, compared to 88% with placebo), but AEs were more common at the higher dose for the categories of "nervous system disorders", "general disorders and administration site conditions" and "gastrointestinal disorders". Conversely, "psychiatric disorders" were more common at the lower dose, as shown in the table below.

Table 14. Adverse events by system organ class

System Organ Class	*Placebo*	*NMF 10*	*NMF 40*
Any adverse events reported	44 (88%)	46 (92%)	46 (92%)
Nervous system disorders	20 (40%)	26 (52%)	32 (64%)
General disorders and administration site conditions	18 (36%)	23 (46%)	27 (54%)
Gastrointestinal disorders	14 (28%)	25 (50%)	26 (52%)
Psychiatric disorders	6 (12%)	14 (28%)	9 (18%)
Skin & subcutaneous tissue disorders	9 (18%)	8 (16%)	7 (14%)
Musculoskeletal, connective tissue and bone disorders	8 (16%)	5 (10%)	6 (12%)
Ear and labyrinth disorders	1 (2%)	3 (6%)	5 (10%)
Investigations	2 (4%)	4 (8%)	4 (8%)
Immune system disorders	-	-	3 (6%)
Metabolism and nutrition disorders	2 (4%)	6 (12%)	3 (6%)
Cardiac disorders	1 (2%)	4 (8%)	2 (4%)
Eye disorders	3 (6%)	1 (2%)	2 (4%)
Infections and infestations	5 (10%)	8 (16%)	2 (4%)
Injury and poisoning	3 (6%)	2 (4%)	2 (4%)
Renal and urinary disorders	2 (4%)	-	2 (4%)
Respiratory, thoracic and mediastinal disorders	4 (8%)	2 (4%)	2 (4%)
Reproductive system and breast disorders	1 (2%)	-	1 (2%)
Blood and lymphatic system disorders	2 (4%)	-	-
Surgical and medical procedures	1 (2%)	2 (4%)	-
Vascular disorders	1 (2%)	1 (2%)	-

Subsequent studies by Biotie targeted 20mg, but allowed subjects to alter the dose as needed to 10mg or 40mg – increasing the dose if efficacy was perceived to be poor, and reducing the dose if they encountered side effects. Despite the freedom to modify the dose, most subjects continued the target 20mg dose, indicating that this dose was reasonably tolerated. Efficacy of the 20mg dose in these Biotie studies was not clearly confirmed, with 2 of 3 placebo-controlled Biotie studies of this dose producing a negative outcome.

No Biotie study directly compared the efficacy of 20mg to 40mg, or compared 20mg to 10mg, so the comparative efficacy of doses in this range remains unclear.

Despite the short-comings in the earlier Biotie study program, Lundbeck performed three Phase III studies that all used the same 20mg dose, on the basis that one Biotie study had reached a positive outcome at this dose. Thus, there is still no adequate comparative efficacy data assessing other possible doses, but this is of relatively minor concern given the other problems with the efficacy data.

7. Clinical efficacy

Efficacy data for the proposed indication comes from 8 studies, including 3 performed by the current Sponsor, Lundbeck, and 5 performed by an earlier Sponsor, Biotie. These studies are listed below. In addition to having different Sponsors, the two groups of studies had different entry requirements and they are therefore distinguished in the Summary of Clinical Efficacy by using slightly different terms for the target condition: “Alcohol Dependence” for the Lundbeck studies, and “Alcohol-use Disorders” for the Biotie studies.

Table 15. Overview of Clinical Studies in Alcohol Dependence (Lundbeck Sponsored)

Study	Study Design	Number of Patients in FAS	
		PBO	NMF
12014A	24-week (+ 4-week run-out), randomised, double-blind, placebo-controlled, fixed-dose (20mg), as-needed dosing	289	290
12023A	24-week (+ 4-week run-out), randomised, double-blind, placebo-controlled, fixed-dose (20mg), as-needed dosing	326	329
12013A	52-week, randomised, double-blind, placebo-controlled, fixed-dose (20mg), as-needed dosing	137	415
Total		**752**	**1034**

FAS = full-analysis set; PBO = placebo; NMF = nalmefene

Table 16. Overview of Clinical Studies in Alcohol-Use Disorders (Biotie Sponsored)

Study	Study Design	Number of Patients in ITT Population	
		PBO	NMF
Controlled studies			
CPH-101-0801	28-week (+ 24-week run-out), randomised, double-blind, placebo-controlled, flexible-dose (10, 20 [target dose], or 40mg), as-needed dosing	159	236
CPH-101-0701	28-week, randomised, double-blind, placebo-controlled, flexible-dose (10, 20 [target dose], or 40mg), as-needed dosing	82	85
CPH-101-0299	12-week daily dosing + 40-week as-needed dosing extension, randomised, double-blind, placebo-controlled, fixed-dose (5, 20, or 40mg)	58	5mg: 61 20mg: 59 40mg: 59
CPH-101-0399	16-week, randomised, double-blind, placebo-controlled, fixed-dose (10 or 40mg), daily dosing	50	10mg: 50 40mg: 50
Uncontrolled study			
CPH-101-0400	52-week, open-label, flexible-dose (10, 20, or 40mg), as-needed dosing		60
Total		**349**	**660**

ITT = intent-to-treat

Two of the Lundbeck studies (12014A and 12023A) can be considered pivotal, and were submitted by the Sponsor as confirmatory studies. These two studies shared an identical design, and they are described together in this report. (Study 12023A is considered the more reliable of the two pivotal studies, because it was less potentially susceptible to withdrawal bias.)

The third Lundbeck study (12013A) has an overall design that could have led to its being considered pivotal, if the protocol had been specified prospectively, but this study was originally conceived as a safety and tolerability study, with efficacy considerations finalised once the study was underway. This study should therefore be considered a major supportive study.

Of the 5 Biotie studies, only one (CPH -101-0801) was sufficiently large that it could be considered a major supportive study. Three other placebo-controlled Biotie studies (CPH -101-0701, CPH -101-0299 and CPH -101-0399) should only be considered minor supportive studies and one of these (CPH-101-0701) was associated with such a high withdrawal rate (71%) that the results are meaningless. The uncontrolled Biotie study (CPH-101-0400) should be rejected.

In place of the cumbersome names used to designate each study in the submission, this report will use the following abbreviations.

Table 17. Study Abbreviations

Abbreviation	Original Designation
Lundbeck14	12014A
Lundbeck23	12023A
Lundbeck13	12013A
Biotie801	CPH-101-0801
Biotie701	CPH-101-0701
Biotie299	CPH-101-0299
Biotie399	CPH-101-0399
Biotie400	CPH-101-0400

7.1. Pivotal studies

In evaluating these studies, it became clear that post hoc analyses played a very prominent part in the Sponsor's presentation of the results. In particular, the Sponsor emphasised results in one particular subgroup, which was not mentioned at all in the study protocol, but which was given increasing prominence in the Summary of Clinical Efficacy and the Clinical Overview, culminating in the proposed PI which only discussed the results in this subgroup, omitting the primary efficacy results completely.

The post hoc change in target group was justified in the PI as follows (emphasis added):

> *In Studies 1 (12014A; n = 579) and 2 (12023A; n = 655), 18%, and 33%, of the total population, respectively, considerably reduced their alcohol consumption in the period between screening and randomisation. Of the patients with a high or very high DRL at baseline, 35% experienced improvement due to non-pharmacological effects in the period between the initial visit (screening) and randomisation. At randomisation, these patients consumed such a small amount of alcohol that there was little room for further improvement (floor effect). Therefore,* ***the patients who maintained a high or very high DRL at randomisation were defined post hoc as the target population.*** *In this post hoc population, the treatment effect was larger than that in the total population. Drinking Risk Level (DRL) was defined according to WHO criteria.*

Despite the Sponsor's preferred emphasis on one particular post hoc analysis, this evaluation will concentrate on the prospectively specified endpoints.[6] The post hoc results are of interest, and appear to identify a subgroup worthy of further study, but it is statistically invalid to identify this group on the basis of the results and then apply standard statistical tests to those same results as if they had been obtained prospectively. (An independent statistician has also declared the Sponsor's post hoc approach to be invalid.)

[6] The term 'prospecitively *specified* endpoints' obviously includes the full specification of those endpoints, including the population in which those endpoints are to be assessed. Nonetheelss, the Sponsor flagged this statement as 'incorrect' because their preferred *post hoc* analysis used the same efficacy variables as that proposed prospectively, albeit in a different population. This is hair-splitting, and misses the point.

An additional issue is that this subgroup was identified in a particular trial context, being separated from the total, prospective cohort on the basis of their behavioural response to entry in the trial itself, and the Sponsor wishes to apply a specific interpretation of how that trial context should be mapped to the clinical setting – in particular, the Sponsor does not want to restrict nalmefene to subjects who continue to exhibit high DRL after introduction of psychosocial measures, There is no empirical basis allowing a precise mapping of the post hoc subgroup to a clinically identifiable target population, and the evaluator and Sponsor disagree on how this mapping should be approached. As in other areas of disagreement, the onus of proof lies with the Sponsor.

The results in the Sponsor's favoured post hoc subgroup – the only results they chose to disclose their proposed PI – are discussed.

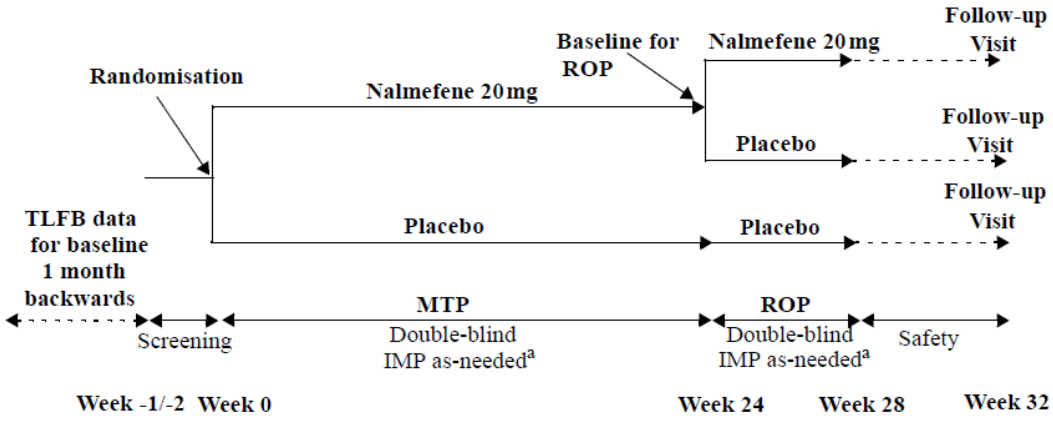
7.1.1. Study design, objectives, locations and dates

These two studies (Lundbeck14, n = 579; Lundbeck23, n = 655) shared an identical design[7] and are therefore considered together in this report. Both studies were double-blind, placebo-controlled parallel-group studies that assessed the efficacy of as-needed dosing with nalmefene 20mg versus placebo in the treatment of alcohol dependence. Treatment duration was 24 weeks, followed by a 4-week run-out period (ROP) in which subjects on active treatment were randomly and blindly assigned to continue active treatment or switch to placebo, in order to assess the effects of withdrawal of treatment.

In all subjects, formal psychosocial supportive measures (the BRENDA program, described below) were supplied at Randomisation and at each visit throughout the study (Weeks 0, 1, 2, 4, 8, 12, 16, 20, 24, 28). (The TLFB data strongly suggest that a powerful psychosocial effect was also produced by the process of enrolling patients and performing initial monitoring of alcohol intake – mean alcohol intake declined substantially prior to Randomisation.)

The primary therapeutic objective in both studies was to reduce alcohol intake. Subjects were advised to take investigational medicinal product (IMP), which was either nalmefene 20mg or placebo, on days when they felt they were at risk of consuming alcohol. Subjects were asked to take their tablet 1 to 2 hours prior to the anticipated time of drinking. If the patient started drinking alcohol without taking a study tablet, they were supposed to take it as soon as possible. Tablets could be taken up to once daily.

Figure 4. Study Design - Lundbeck14 and Lundbeck23



TLFB = timeline followback; MTP = Main Treatment Period; ROP = Run-out Period
a BRENDA was provided at Weeks 0, 1, 2, 4, 8, 12, 16, 20, 24, and 28.

[7] The Sponsor objected to the term 'nearly identical' in the first-round CER, claiming that the study designs were actually identical, but the studies used different imputation methods for their key secondary endpoint.

Lundbeck14 was conducted at 39 sites in 4 countries, as follows: 4 in Austria, 11 in Finland, 16 in Germany, and 8 in Sweden. The first patient-visit was on 18th December 2008, and the last on 14th October 2010.

Lundbeck23 was conducted in 57 sites, as follows: 7 in Belgium, 3 in the Czech Republic, 16 in France, 10 in Italy, 7 in Poland, 4 in Portugal, and 10 in Spain. The first patient-visit was on 16th March 2009, and the last on 22nd March 2011.

7.1.2. Inclusion and exclusion criteria

The main entry criterion was a primary diagnosis of Alcohol Dependence according to the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV-TR™). Additional criteria were that the subjects:

- were ≥ 18 years of age
- had a blood alcohol concentration (BAC) < 0.02% at the Screening Visit
- had documented ≥ 6 heavy drinking days (HDDs) in the 4 weeks preceding the Screening Visit
- had an average alcohol consumption at *medium risk level or above* (> 40g/day for men; > 20g/day for women) in the 4 weeks preceding the Screening Visit
- if female, were willing to take measures to avoid pregnancy

The main exclusion criteria were:

- > 14 consecutive non-drinking days in the 4 weeks prior to Screening
- significant medical or psychiatric issues apart from Alcohol Dependence
- dependence on additional drugs
- aspartate aminotransferase (ASAT) or alanine aminotransferase (ALAT) values > 3 times the upper limit of the reference range that were considered clinically significant

The Sponsor also listed the following more detailed exclusion criteria, (quoted verbatim):[8]

1. The patient has had < 6 HDDs in the 4 weeks preceding the Screening Visit.
2. The patient has had an average alcohol consumption below medium risk levels.
3. The patient has had > 14 consecutive abstinent days in the 4 weeks preceding the Screening Visit.
4. The patient has a Revised Clinical Institute Withdrawal Assessment for Alcohol (CIWA-Ar) score ≥ 10.
5. The patient has:
    a. a DSM-IV Axis I disorder other than alcohol dependence or nicotine dependence as evaluated using the Mini-International Neuropsychiatric Interview (MINI)
    b. an antisocial personality disorder as evaluated using the MINI
    c. other disorders for which the treatment takes priority over treatment of the drinking problem or are likely to interfere with study treatment or impair treatment compliance
    d. (Use of cannabis is not a reason for exclusion unless it fulfils the criteria for cannabis dependence.)
6. The patient is at risk of suicide as evaluated using the suicidality module of the MINI.

[8] The criteria are those listed for Lundbeck14; the criteria for Lundbeck23 were essentially the same.

7. The patient has a history of delirium tremens or alcohol withdrawal seizures.
8. The patient has a significant cognitive disorder likely to interfere with the patient's understanding of the study and its procedures.
9. The patient has reported, or urine drug screen has revealed, current use of substances of abuse other than alcohol, cannabis, nicotine, or benzodiazepines.
10. The patient has seizure disorder, mental retardation, or encephalopathy.
11. The patient has a clinically significant unstable illness, for example, hepatic or renal insufficiency, or a cardiovascular, pulmonary, gastrointestinal, endocrine, neurological, infectious, neoplastic, or metabolic disturbance.
12. The patient has clinically significant abnormal vital signs.
13. The patient has S-aspartate aminotransferase (ASAT) and/or S-ALAT values >3 times of upper limit of the reference range, or one or more laboratory values outside the reference range, based on the blood and urine samples taken at the Screening Visit, that are, in the investigator's opinion clinically significant.
14. The patient has a clinically significant abnormal ECG.
15. The patient has a history of severe drug allergy or hypersensitivity.
16. The patient has reported current or recent (within 3 months prior to the Screening Visit) treatment with disulfiram, acamprosate, topiramate, naltrexone, or carbimide, or with any opioid antagonists (carbimide was added with SA01).
17. The patient has reported current or recent (within 1 week prior to the Screening Visit) treatment with opioid agonists or partial agonists.
18. The patient has reported current or recent (within 8 weeks prior to the Screening Visit) treatment with antipsychotics or antidepressants.
19. The patient has taken/takes disallowed recent or concomitant medication or it is anticipated that the patient will require treatment with at least one of the disallowed concomitant medications during the study.
20. The patient has a disease or takes medication that could, in the investigator's opinion, interfere with the assessments of safety, tolerability, or efficacy.
21. The patient has been treated with any investigational medicinal product (IMP) within 30 days or 5 half-lives (whichever is longer) prior to the Screening Visit.
22. The patient is currently participating or has recently (within 4 weeks prior to the Screening Visit) participated in a treatment or support programme for alcohol-use disorders.
23. The patient is pregnant or breast-feeding.
24. The patient is, in the investigator's opinion, unlikely to comply with the protocol or is unsuitable for any reason.
25. The patient is a member of the site personnel or their immediate families.
26. The patient is under forced treatment.
27. The patient has previously participated in clinical studies with nalmefene.

Overall, these criteria were reasonable, and were aimed at recruiting subjects from the target population in whom a fair assessment could be made of the efficacy of nalmefene, without confounding treatments or conditions, and without exposing patients to unacceptable safety risks.

Note that the entry criteria reveal the Sponsor's initial concept of their target population, and this is at odds with the target population subsequently claimed in their proposed PI.

Study treatments:

All subjects were randomised to nalmefene 20mg or placebo, with an additional randomisation of nalmefene subjects to withdrawal or continuation of active treatment in the run-out period.

Subjects could take up to one tablet of their assigned treatment on any one day. They were asked to take the treatment on an as-needed basis, 1-2 hours prior to the anticipated time of drinking.

Concomitant medications were allowed, with some restrictions as listed in the table below.

Table18. Restrictions on Recent and Concomitant Medication

Drug Class	Disallowed Prior to the Screening Visit	Disallowed During the Study for		
		Chronic Use	Episodic Use	Comments or Exceptions
Agents used for treatment of alcohol-use disorders	Within 3 months	X	X	Including oral and injectable naltrexone, acamprosate, topiramate, disulfiram, and carbimide[a]
Any IMP	Within 30 days or 5 half-lives (whichever is longer)	X	X	
Analgesics	Within 1 week for opioid agonists and partial agonists	X	X	Non-opioids were allowed.
Anorexics		X	X	
Antianginal agents		X	X	
Antibiotics				Nitrofurantoin, metronidazole, and tinidazole were not allowed due to their interaction risks with alcohol.
Anticoagulants		X	X	
Anticonvulsants[b]	X	X	X	Barbiturates were allowed in patients not taking oral contraceptives.
Antidepressants	Within 8 weeks[b]	X	X	
Antidiarrhoeal agents				Opioids were not allowed.
Antifungal agents: Systemic Topical		X	X	Antifungal agents for topical use were allowed.
Antineoplastics	X	X	X	
Antipsychotics	Within 8 weeks[b]	X	X	
Anxiolytics		X		This applied to benzodiazepines. See also antidepressants.
Cough/cold agents				Opioids were not allowed.
Insulin	X	X	X	
Muscle relaxants		X		
Psychotropic agents not otherwise specified (including herbal agents)		X	X	
Sedatives/hypnotics		X		
Steroids: Systemic Topical Inhalant		X		Systemic corticosteroids were not allowed. Topical and inhalant use were allowed. Oral contraceptives and hormone replacement therapy were allowed.

a Carbamide was added with SA01.
b Added with SA01

7.1.3. Efficacy variables and endpoints

The Sponsor listed a large number of efficacy variables, as reproduced below.

Drinking measures derived from the patient diary (timeline followback, TLFB):

- number of heavy-drinking days (HDDs, a day with alcohol consumption ≥ 60g for men and ≥ 40g for women)
- Total Alcohol Consumption (TAC, defined as mean daily alcohol consumption in g/day over 28 days)
- Response defined as a shift in drinking risk level (RSDRL, defined as a downward shift from baseline in DRL; for patients at very high risk at Baseline, a shift to medium risk or below, and for patients at high or medium risk at Baseline, a shift to low risk or below)
- TAC response (defined as a ≥ 30%, ≥ 50%, or ≥ 70% reduction in TAC from Baseline; response based on a ≥ 70% reduction was added as a post hoc analysis)
- Response defined as the achievement of a low drinking risk level (R*L*DRL, defined as a downward shift in DRL to low risk or below)
- Number of non-drinking days (NDDs)

Alcohol dependence symptoms and clinical status:

- Clinical Global Impression – Global Improvement (CGI-I) score
- CGI-I response, defined as a CGI-I score ≤ 2 (added as a *post hoc* analysis)
- Clinical Global Impression – Severity of Illness (CGI-S) score
- Alcohol Dependence Scale (ADS) total score
- Drinker Inventory of Consequences (DrInC-2R) score

Liver function and other clinical safety laboratory tests:

- gamma-glutamyl transferase (GGT)
- alanine aminotransferase (ALAT)
- mean corpuscular volume (MCV)
- percent carbohydrate-deficient transferrin (%CDT)

Pharmacoeconomic outcomes:

- 36-item Short-form Health Survey (SF-36) subscale scores
- EuroQol (EQ-5D) utility index and visual analogue scale (VAS) scores
- Resource Use Measurement Questionnaire – Alcohol Dependence (RUMQ-ADP)
- Brief Measure of Readiness to Change Questionnaire (BMRCQ) subscale scores: importance, confidence, and readiness.

The two co-primary efficacy variables for both studies were the two direct measures of alcohol consumption:

- Number of Heavy Drinking Days (HDDs) per 28-day period
- Total Alcohol Consumption (TAC) in g/day, averaged over 28 days

These co-primary variables measure different aspects of alcohol consumption, and are both important. Some forms of alcohol-related harm, such as the risk of injury, may be related to days of excessive consumption. Others are related to the overall alcohol consumption.

Estimating alcohol intake may be difficult, but this is a difficulty inherent to all studies of Alcohol Dependence, and the Sponsor took reasonable measures to get information that was as accurate as possible. Directly relying on the patient to record each drink in a diary is unreliable, because subjects may drink at times when they are away from their diary; they may be self-

conscious about making an entry while drinking socially; and their reliability may become further compromised during a drinking session, in which individual drinks may be difficult to track. To get around this, the Sponsor used a timeline followback (TLFB) method devised by Sobell et al[9] in which patients provide retrospective estimates of daily drinking, using memory aids, such as a calendar. In essence, the patients are required to recall what they were doing each day, and note whether they drank and then estimate how much. If a patient missed a visit, the TLFB was extended to cover the missing days. In using this method, patients were asked to report their alcohol intake in terms of "standard units", defined in standard drink conversion cards distributed to the patients. Months with < 7 days of data were discarded from the analysis; these are likely to represent months in which drinking was excessive, leading to an inherent bias in the TLFB data towards more favourable results – but without a particular bias in favour of active therapy.

The major secondary efficacy variables were the response rates, which attempted to capture the proportion of patients with a clinically worthwhile shift in drinking-related health risk, and the clinicians' impressions:

- Response-Shift in DRL (RSDRL)
- Response-Low in DRL (RLDRL)
- Clinical Global Impression – Global Improvement (CGI-I) score
- Clinical Global Impression – Severity of Illness (CGI-S) score

All of these were considered secondary endpoints, but the Sponsor identified the RSDRL at Month 6 as the key secondary efficacy endpoint.

The laboratory measures, based on liver function tests, were used as indirect, surrogate indicators of alcohol consumption; they provide some objective support for the diary measures.

In addition to the large number of efficacy variables, the Sponsor also defined a number of efficacy analyses, based on different statistical approaches to the same efficacy variables. The most important analysis was the primary efficacy analysis, defined by the Sponsor as follows:

> *The changes from Baseline I in monthly number of HDDs and monthly TAC were analysed using a mixed model repeated measures (MMRM) model, using observed cases (OC), with the Baseline I score as a covariate, and site, sex, time in months (Months 1 to 6), and treatment as fixed effects, with the estimated treatment difference at Month 6 tested at the 5% level of significance. [Lundbeck14 Synopsis]*

The primary efficacy analysis for both studies was identical.

Although this represents the major confirmatory analysis on which the studies must be judged, the Sponsor also assessed the co-primary endpoints in a number of other ways, to check for the robustness of the results. In particular, the Sponsor assessed a number of different methods for imputing missing data from withdrawing patients. These analyses are discussed further under 'Statistical Methods' below.

The key secondary efficacy analysis was an assessment of RSDRL at 6 months using a logistic regression (LREG) model. As with the primary efficacy analysis, this result was subjected to a number of additional sensitivity analyses using different methods of imputation.

Despite the very large number of efficacy variables and endpoints listed in the protocol, the two endpoints that were eventually presented in the PI (HDDs and TAC in subjects with high or very high DRL at Randomisation) did not feature at all in the prospective protocol.

[9] Sobell LC, Sobell MB. Timeline Follow-back: a technique for assessing self-reported ethanol consumption. In: Litten RZ, Allen JP, editors. Measuring alcohol consumption: psychosocial and biological methods. Totowa, NJ, US: Humana Press; 1992. p 41-72.

7.1.4. Randomisation and blinding methods

Subjects were randomised to placebo or nalmefene in a 1:1 ratio, according to a central randomisation list that was computer generated by Lundbeck, using random numbers assigned to each patient at screening. Block randomisation was used to ensure that balanced numbers of patients entered each treatment group within each site. Randomisation for the run-out period was performed with the original randomisation.

Blinding was approached by storing the randomisation codes centrally, and making treatment assignment unknown to all patients and clinicians. Placebo and nalmefene tablets appeared identical.

In the Biotie studies, a bittering agent was used in the placebo tablets, but there was no evidence in Lundbeck's original submission that a bittering agent was used in the Lundbeck studies. (A search of the Lundbeck14 study report for the keyword "bitter" produces a single adverse event of wine tasting bitter, but no mention of a bittering agent to assist blinding. A similar search of the Lundbeck23 report reveals two adverse events of bitter taste. All reports of bitterness occurred in nalmefene recipients.) In their Section 31 Response, the Sponsor has since clarified that a bittering agent was used, but it is of concern that this issue was not given more prominence in their original submission.

Similarly, the capacity for un-blinding via telltale side effects was not discussed by the Sponsor, and no attempt was made to test for possible un-blinding by asking subjects to guess their assigned treatment allocation. This is unfortunate, because several side effects may have provided an unblinding signal, as discussed in Sectio.

When this issue was raised in the first-round CER, the Sponsor rejected the possibility of unblinding, annotating CER1 with the following comment:

> *There was a substantial decrease in alcohol consumption in both treatment groups with absence of high drop out in the placebo group and with large adherence to treatment in the placebo group. In such context it is difficult to see the relevance of the Assessor's point concerning potential for unblinding due to telltale adverse events.*

In this context, it should be recalled that:

- subjects sought treatment and entered the study because they wanted to cut their drinking;
- many subjects cut their drinking before they received any treatment;
- regardless of treatment assignment, subjects continued to receive psychosocial treatment to help them cut drinking;
- subjects agreed to enter the study in the full knowledge that they might receive placebo.

It is thus fanciful to propose that many subjects would abandon their initial desire to curtail their drinking and reject the benefits of on-going psychosocial interventions merely because they suspected they were receiving placebo. Subjects began with an intent to cut their drinking, and many cut their drinking before receiving treatment, so no conclusions about blinding can be drawn from the fact that many placebo recipients continued to follow through with this initial intent.

Even if the Sponsor's point were valid, though, and somehow the continued minor improvements in the placebo group post-Randomisation were proof that no members of the placebo group were unblinded, this would still tell us nothing about unblinding in the active group. For individual subjects, a lack of side effects is essentially a lack of evidence, and not a reliable indicator of receiving placebo – subjects without side effects could simply assume that the drug had few side effects. The presence of side effects, however, is potentially a strong indicator of active treatment, and this might be expected to produce an additional psychological incentive to curtail drinking.

In a study where the primary efficacy variable is under voluntary control, the Sponsor's failure to address unblinding is an important omission. Unblinded subjects on active treatment might have been more embarrassed about their drinking, if they perceived every drinking occasion as a failure of the active agent to control their addiction, leading to:

- genuinely lower intake; or
- a greater reluctance to report drinking.

Thus, unblinding could have had complex effects on the efficacy variables completely unrelated to the actual proposed mechanism of action of the drug. These issues were not discussed by the Sponsor in their original submission, and have been dismissed by the Sponsor in their Section 31 Response, but they remain a major methodological concern.

7.1.5. Analysis populations

The Sponsor defined the following analysis sets:

- all-patients-randomised set (APRS) – all randomised patients
- all-patients-treated set (APTS) – patients in the APRS, excluding those with no recorded study drug intake and all study drug returned
- full-analysis set (FAS) – all patients in the APTS who had at least one valid post-baseline assessment in the main treatment period of both co-primary efficacy variables (HDD and TAC) and had an average alcohol consumption at medium risk or above according to World Health Organization criteria (> 40g/day for men and > 20g/day for women) in the 4 weeks preceding the Screening Visit.

The results in the proposed PI are not based on any of these prospectively defined populations.

7.1.6. Sample size

Sample size estimations were based on the intended primary efficacy analysis of both co-primary endpoints, using a Mixed Model Repeated Measures (MMRM) approach, assuming a correlation of 0.7 between the co-primary efficacy variables. The estimation also assumed a standard deviation for the change from baseline in HDDs of 7 days and a change from baseline in TAC of 36.5g/day, and a withdrawal rate of up to 35% by Month 6.

Based on these parameters, a minimum of 600 enrolled patients (300 per group) was estimated to provide a power of at least 90% for detecting a difference of 3 HDDs and 12g/day in the TAC at a traditional significance level of 5%.

Recruitment in Lundbeck14 achieved this target, with ~300 subjects in each group (placebo n = 298, nalmefene n = 306) at randomisation. In Lundbeck23, the target was extended to 350 patients per group because a blinded interim review of the data indicated higher than anticipated standard deviations and lower than anticipated correlations for the co-primary endpoints. The revised recruitment target for Lundbeck23 was exceeded (placebo n = 360, nalmefene n = 358).

Of the two co-primary endpoints across each pivotal study (4 endpoints in total) only 3 achieved statistical significance. This partly reflects that the observed treatment effect was smaller than anticipated, rather than necessarily indicating that the studies were underpowered.

7.1.7. Statistical methods

The statistical software used in both pivotal studies was SAS®, Version 9.2. In addition, SADs Version 4.0 was used for generating the analysis data. All major (prospective) efficacy analyses were conducted using the Full Analysis Set (FAS), but the results reported in the PI are based on a subgroup within the FAS.

There were very few differences between analysis methods in the two studies, and these are mentioned below where relevant. In general, the major analyses were identical, but Lundbeck23 used additional post hoc analyses with different imputation methods. One important difference between the two studies was analysis of the key secondary endpoint (RSDRL), where the two studies used different imputation methods, and this led to different results.

For the co-primary efficacy endpoints in both studies, the changes from Baseline[10] in monthly number of HDDs and monthly TAC were analysed using a mixed model repeated measures (MMRM) approach, using observed cases (OC), with the Baseline score as a covariate, and site, sex, time in months and treatment as fixed effects. The Baseline score-by-time interaction and the treatment-by-time interaction were also included in the model.

The estimated treatment difference at Month 6 was tested at the 5% level of significance ($p < 0.05$). No statistical correction for the use of multiple endpoints was employed, so both co-primary endpoints had to show a significant treatment effect at Month 6 for the study to be considered positive. According to the Sponsor:

> *The null hypothesis of no difference in treatment effect was tested against the alternative hypothesis that there was a difference in treatment effect in the MMRM analyses. The null hypothesis was to be rejected for both co-primary endpoints at the 5% significance level to consider nalmefene 20mg as-needed use to be efficacious.*

Sensitivity analyses were also performed using: an MMRM in which monthly observations were disregarded if there were < 14 days of data (instead of < 7 days as in the primary analysis); and analyses of covariance (ANCOVA) by month, using OC, last observation carried forward (LOCF), baseline observation carried forward (BOCF), or placebo mean imputation (PMI, based on the mean reduction observed at Month 1 in the placebo group, adjusted for sex). In both studies, another imputation method was used, known as multiple imputation (MI). This technique was described by the Sponsor as follows:

> *Another sensitivity analysis for each of the co-primary endpoints was performed using a multiple imputation (MI) method, which assumed that patients who withdraw differ from completers and that their future outcomes (conditional on the past) are the same as those in the placebo group (with the same past).[11] Multiple imputation was based on the pattern-mixture model using 50 simulations. The 50 complete data sets were analysed using the MMRM model as for the primary analysis of co-primary endpoints. The different estimated treatment effects and standard errors across the data sets were combined to produce a unique point estimate and standard error, taking into account the uncertainty of the imputation process.*

Subgroup analyses were performed for the patients with a high or very high DRL at Baseline (HDAB), using an MMRM approach similar to that used for the co-primary efficacy analyses, as well as with a post hoc ANCOVA using LOCF. This important subgroup is considered separately below.

An important omission in the Sponsor's submission was a corresponding subgroup analysis in patients with medium DRL at Baseline (MDAB). To a limited extent, the Sponsor addressed this omission in the Summary of Clinical Efficacy, where results in medium DRL subjects were included in a couple of summary tables. The Sponsor also provided more data on the medium DRL subgroup in their Section 31 Response.

[10] The sponsor defined two baselines, Baseline I was the baseline for the Main Treatment Period (MTP), and Baseline II was the baseline for the Run-Out Period (ROP). This report will simply refer to the baseline of the main treatment period as "the Baseline". Baseline variables derived from the TLFB data (alcohol intake estimates) were based on the month (28 days) preceding the Screening Visit, and for other efficacy assessments, the Baseline value was defined as the value at the Screening Visit.

[11] Little R, Yau L. Intent-to-treat analysis for longitudinal studies with drop-outs. Biometrics 1996; 52: 1324-1333.

It should be noted that most of the Sponsor's discussion of the pivotal studies was outside the scope of the Statistical Analysis Plan. For instance, *post hoc* analyses were performed to compare the effect of nalmefene versus placebo in the patients who, at Randomisation, still fulfilled the Screening requirements regarding alcohol consumption. This was done with an MMRM model including a consumption-by-time-by-treatment interaction, as well as with ANCOVA using LOCF. Post hoc analyses were also performed on subjects who exhibited at least high-risk DRL at both Baseline and Randomisation (High DRL at Randomisation, HDAR).

The key secondary efficacy endpoint was RSDRL at Month 6, which was analysed with a logistic regression (LREG) model, using country, sex, Baseline DRL, and treatment as fixed effects, and with missing values imputed as non-response (NR imputation) for Lundbeck14 and as the MMRM-predicted TAC for Lundbeck23. This difference in imputation method led to substantially different outcomes in the two studies for this endpoint, as discussed below.

In view of the multiple endpoints, the null hypothesis (hypothesis of no treatment effect) had to be rejected for both co-primary endpoints at the 5% level in order to proceed with formal testing of the key secondary RSDRL endpoint. This means that the RSDRL endpoint should not have been subjected to formal statistical analysis in Lundbeck23, and results for this endpoint should be considered descriptive.

Sensitivity analyses were performed on the RSDRL endpoint with different imputation methods, including OC or LOCF, and by an approach where monthly observations were disregarded if there were < 14 days of data. A post hoc analysis was also performed in which imputation was based on individual patient-predicted values of TAC at each month derived from the MMRM model used in the primary analysis of TAC. In Lundbeck23, a non-response imputation method was also assessed, as well as a sustained response imputation. In this imputation method, a sustained response was defined as a response at the current month and previous month, and LOCF was applied to sustained responses but not to transient or unconfirmed responses.

A number of secondary efficacy analyses were performed, including an analysis of responder rates by other definitions (based on a ≥ 30%, ≥ 50%, or ≥ 70% reduction in monthly TAC, or by RLDRL); these were analysed using an LREG model.

Changes from Baseline in monthly number of NDDs, DrInC-2R scores, CGI-S scores, %CDT, SF-36 subscale scores, EQ-5D utility index and VAS scores, and BMRCQ scores in the MTP were analysed using an MMRM model similar to the one used for the co-primary efficacy analyses. The CGI-I score was analysed using a similar MMRM model, but with the CGI-S Baseline I score as a covariate.

The log-transformed GGT, ALAT and MCV values were analysed using an MMRM model, using observed cases, with the log-transformed Baseline value as a covariate, and site, sex, time in weeks, and treatment as fixed effects. Log-transformed Baseline-score-by-time interactions and treatment-by-time interactions were also included in the model.

For the minor endpoint of ADS total score, the change from Baseline to each month was analysed using an ANCOVA, with missing data imputed by BOCF.

7.1.8. Participant flow

Participant flow is summarised in the panels below. The proportion of patients that withdrew was very high and unequal across treatment groups in Lundbeck14 (the only pivotal study with a positive overall result), raising the possibility that withdrawal bias contributed to the positive findings. This seems particularly likely in view of the fact that adverse events were the major reason for withdrawals in the nalmefene group, as shown in the third panel below.

The proportion of withdrawals in Lundbeck23 was lower, and similar across treatment groups, but still higher than desirable for a pivotal study.[12]

In trying to interpret the tables below, it should be noted that the first two tables below refer to all randomised patients, whereas the third table expresses withdrawals as a proportion of subjects in the Full Analysis Set. The FAS population is considered most relevant, because this is the group used for the efficacy analysis.

Table 19. Patient Disposition – Lundbeck14

	PBO	NMF	Total
Patients randomised	298	306	604
Patients treated	296 (100%)	302 (100%)	598 (100%)
Patients completed	200 (68%)	138 (46%)	338 (56%)
Patients withdrawn	96 (32%)	164 (54%)	260 (44%)
FAS	289	290	579

Table 20. Patient Disposition – Lundbeck23

	PBO	NMF	Total
Patients randomised	360	358	718
Patients treated	337 (100%)	341 (100%)	678 (100%)
Patients completed	205 (61%)	194 (57%)	399 (59%)
Patients withdrawn	132 (39%)	147 (43%)	279 (41%)
FAS	326	329	655

Table 21. Withdrawals by Primary Reason – Lundbeck14 and Lundbeck23

	12014A		12023A	
	PBO n (%)	NMF n (%)	PBO n (%)	NMF n (%)
FAS	289 (100)	290 (100)	326 (100)	329 (100)
Patients completed[a]	213 (74)	152 (52)	229 (70)	212 (64)
Patients withdrawn	76 (26)	138 (48)	97 (30)	117 (36)
Primary Reason				
Adverse events	17 (6)	57 (20)	4 (1)	12 (4)
Lack of efficacy	19 (7)	17 (6)	13 (4)	7 (2)
Non-compliance		7 (2)	3 (1)	7 (2)
Protocol violation	4 (1)	10 (3)	25 (8)	19 (6)
Withdrawal of consent	25 (9)	31 (11)	32 (10)	43 (13)
Lost to follow-up	7 (2)	12 (4)	11 (3)	12 (4)
Other	4 (1)	4 (1)	9 (3)	17 (5)

a Patients with TLFB data at Month 6

7.1.9. Major protocol violations/deviations

The overall number of protocol violations was not originally summarised by the Sponsor in a convenient format, but instead individual violations were included in multi-page tables. Major deviations included the recruitment of subjects who did not satisfy eligibility criteria, use of

[12] The evaluator points out that the sponsor objected to this sentence in the first-round Clinical Evaluation Report (CER1) with the comment: "The sponsor does not agree with this statement." This leaves the sponsor with the onus of demonstrating that it is desirable to have a withdrawal rate >41% in a pivotal study.

disallowed concomitant medication, visits outside designated windows, and a number of other deviations in data collection.

In response to a request for further information, the Sponsor provided the tables below. It still remains unclear how many of these deviations could be considered "Major deviations" but it appears that the overall incidence of deviations was within the expected bounds for studies of this nature.

Table 22. Protocol deviations in Lundbeck14

Deviation Category	Incidence (%)
Deviation in informed consent procedure	4.6
Violation of inclusion criterion (other than informed consent procedure)	4.1
Violation of exclusion criterion	14
Use of disallowed concomitant medication	12
Procedural compliance deviations: Visit 2 (Randomisation Visit) outside of visit window (that is, ≥21 days after Visit 1)	2.5
Procedural compliance deviations: Visits 3 to 12 outside of visit window (that is, ≥7 days for Visits 3 and 4, or ≥15 days for Visits 5 to 12)	38
Procedural compliance deviations: Other deviations	73

Table 23. Protocol deviations in Lundbeck23

Deviation Category	Incidence (%)
Deviation in informed consent procedure	4.7
Violation of inclusion criterion (other than informed consent procedure)	5.6
Violation of exclusion criterion	30
Use of disallowed concomitant medication	7.8
Procedural compliance deviations: Visit 2 (Randomisation Visit) outside of visit window (that is, ≥21 days after Visit 1)	4.7
Procedural compliance deviations: Visits 3 to 12 outside of visit window (that is, ≥7 days for Visits 3 and 4, or ≥15 days for Visits 5 to 12)	39
Procedural compliance deviations: Other deviations	58

7.1.10. Baseline data

Baseline characteristics in the pooled population across both pivotal studies are summarised below, and baseline data for the two individual studies are shown in the subsequent tables.

Table 24. Baseline Efficacy Summary– Lundbeck14 and Lundbeck23 Pooled

	PBO	NMF	Total
Number of Patients	658	664	1322
DRL, n (%)			
Unknown		1 (0.2)	1 (<0.1)
Low	8 (1.2)	6 (0.9)	14 (1.1)
Medium	142 (21.6)	136 (20.5)	278 (21.0)
High	253 (38.4)	243 (36.6)	496 (37.5)
Very high	255 (38.8)	278 (41.9)	533 (40.3)
Number of HDDs (days/month)			
n	658	663	1321
Mean ± SD	18.9 ± 7.0	19.6 ± 7.1	19.3 ± 7.1
TAC (g/day)			
n	658	663	1321
Mean ± SD	86.6 ± 45.3	88.8 ± 44.8	87.7 ± 45.1
CGI-S			
n	655	664	1319
Mean ± SD	4.0 ± 1.5	4.0 ± 1.5	4.0 ± 1.5
DrInc-2R Total Score			
n	650	660	1310
Mean ± SD	41.0 ± 21.7	41.7± 22.3	41.3 ± 22.0
ADS Total Score			
n	657	663	1320
Mean ± SD	13.5 ± 5.7	13.8 ± 5.8	13.6 ± 5.8

Table 25. Baseline Efficacy Summary– Lundbeck14

	PBO	NMF	Total
Number of Patients	298	306	604
DRL, n (%)			
Unknown		1 (0)	1 (0)
Low	2 (1)	1 (0)	3 (<1)
Medium	60 (20)	68 (22)	128 (21)
High	119 (40)	114 (37)	233 (39)
Very High	117 (39)	122 (40)	239 (40)
Number of HDDs (days/month)			
n	298	305	603
Mean ± SD	19.5 ± 7.0	19.5 ± 7.3	19.5 ± 7.1
TAC (g/day)			
n	298	305	603
Mean ± SD	84.1 ± 41.5	84.8 ± 42.1	84.5 ± 41.8
Maximum Number of Consecutive NDDs[a] (days/month), n (%)			
0	94 (32)	111 (36)	205 (34)
1-3	140 (47)	137 (45)	277 (46)
4-7	51 (17)	47 (15)	98 (16)
8-14	13 (4)	9 (3)	22 (4)
>14		1 (<1)	1 (<1)
CGI-S Score			
n	298	306	604
Mean ± SD	4.0 ± 1.5	4.0 ± 1.5	4.0 ± 1.5
DrInc-2R Total Score			
n	297	304	601
Mean ± SD	35.0 ± 18.1	35.8 ± 18.7	35.4 ± 18.4
ADS Total Score			
n	298	306	604
Mean ± SD	12.2 ± 4.9	12.9 ± 5.8	12.5 ± 5.4

a NDD = non-drinking day

Table 26. Baseline Efficacy Summary– Lundbeck23

	PBO	NMF	Total
Number of Patients	360	358	718
DRL, n (%)			
Low	6 (2)	5 (1)	11 (1.5)
Medium	82 (23)	68 (19)	150 (21)
High	134 (37)	129 (36)	263 (37)
Very High	138 (38)	156 (44)	294 (41)
Number of HDDs (days/month)			
n	360	358	718
Mean ± SD	18.4 ± 7.0	19.7 ± 7.0	19.4 ± 7.0
TAC (g/day)			
n	360	358	718
Mean ± SD	88.8 ± 48.2	92.2 ± 46.9	90.5 ± 47.5
Maximum Number of Consecutive NDDs (days/month), n (%)			
0	121 (34)	134 (37)	225 (36)
1-3	145 (40)	138 (38)	283 (39)
4-7	81 (22)	67 (19)	148 (21)
8-14	13 (4)	19 (5)	32 (4)
CGI-S Score			
n	357	358	715
Mean ± SD	4.0 ± 1.4	4.0 ± 1.4	4.0 ± 1.4
DrInc-2R Total Score			
n	353	356	709
Mean ± SD	46.0 ± 23.1	46.7 ± 23.9	46.4 ± 23.5
ADS Total Score			
n	359	357	716
Mean ± SD	14.6 ± 6.2	14.5 ± 5.7	14.5 ± 5.9

Overall, there were no important baseline differences between the nalmefene and placebo groups at Baseline, and the matching between groups was acceptable.

7.1.11. As-needed usage

Investigational medicinal product (IMP) was only taken on days when subjects perceived themselves to be at risk of drinking (or excessive drinking). Overall, nalmefene subjects took IMP on 48% of days in Lundbeck14, and on 57% of days in Lundbeck23. Placebo recipients took IMP on 64% and 65% of days in the two studies, respectively. As shown in the tables, subjects sometimes drank without taking any IMP (this happened on 11-22% of days), reflecting non-compliance with study instructions on those days. Overall, as a proportion of drinking days, nalmefene recipients were non-compliant 38% of the time in Lundbeck14 (22/58), and 26% of the time in Lundbeck 23 (13/50).

Incomplete compliance could have compromised the demonstration of a treatment effect, but compliance would be expected to be even worse if nalmefene were used in a non-study setting, potentially leading to a weaker treatment effect in clinical practice than the already-weak effect shown in these studies.

Table 27. Mean IMP versus Drinking – Lundbeck14

	IMP Intake (% of days)	No IMP Intake (% of days)
NMF		
Drinking	36%	22%
No drinking	12%	30%
PBO		
Drinking	49%	13%
No drinking	15%	23%

Table 28. Mean IMP versus Drinking – Lundbeck23

	IMP Intake (% of days)	No IMP Intake (% of days)
NMF		
Drinking	37%	13%
No drinking	20%	30%
PBO		
Drinking	40%	11%
No drinking	25%	24%

7.1.12. Results for the co-primary efficacy endpoints

The pivotal results in this submission are summarised in the table below, which shows the two co-primary endpoints for each of the two pivotal studies. Each of these results was analysed by two different techniques, which gives eight p-values, but the primary, prospectively identified confirmatory analysis was the MMRM (with the ANCOVA being merely supportive).

Table 29. Difference to Placebo in HDDs and TAC (FAS) – Total Population – Lundbeck14 and Lundbeck23

Efficacy Variable Analysis	12014A					12023A				
	n		Mean	95% CI	p-value	n		Mean	95% CI	p-value
	PBO	NMF				PBO	NMF			
Number of HDDs (days/month)										
MMRM	213	152	-2.3	[-3.8; -0.8]	0.002	229	212	-1.7	[-3.1; -0.4]	0.012
ANCOVA, LOCF	289	290	-1.7	[-3.0; -0.4]	0.010	326	329	-1.8	[-3.0; -0.6]	0.004
TAC (g/day)										
MMRM	213	152	-11.0	[-16.8; -5.1]	<0.001	229	212	-5.0	[-10.6; 0.7]	0.088
ANCOVA, LOCF	289	290	-8.8	[-14.3; -3.3]	0.002	326	329	-5.9	[-11.1; -0.7]	0.026

CI = confidence interval

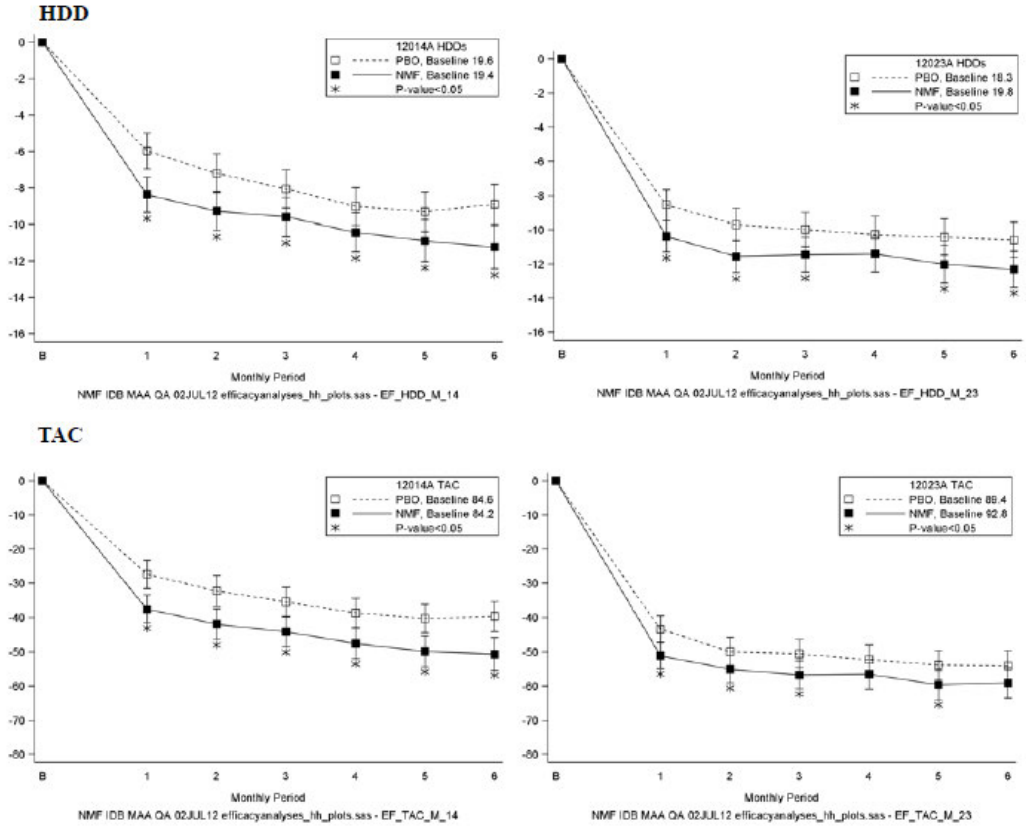
The MMRM results should be considered the main results for assessment of the treatment effect, but all the results should be interpreted with caution given the high withdrawal rates. The MMRM imputation method restricted the analysis to observed cases, but there are doubts whether observed cases are truly representative of the entire population. A patient who withdraws is inherently different from one who persists in a study, even if they have had similar alcohol intake up until that point. The subject's motivation to continue documenting alcohol intake and to take a treatment intended to curtail intake is likely to be related to the subject's intrinsic motivation to control intake, so it is almost inevitable that the withdrawing cohort were less motivated to control their drinking than those who remained in the study. For a heavy drinker, turning up at each visit to report continued heavy drinking could be embarrassing and could encourage withdrawal from the study. It is likely that the presence of drug-related side effects provided additional disincentives to continue, so that the enrichment of the nalmefene group with motivated subjects was greater than that seen in the placebo group.

The ANCOVA results were not always congruent with the MMRM results. Of note, the ANCOVA results are shown with LOCF imputation, which is an optimistic imputation method because it assumes that any reductions in alcohol intake achieved while in the study are continued despite withdrawal from the study. That is, patients whose motivation was so poor that they discontinued from the study had their alcohol intake locked in from a timepoint when they were still cooperative – the data for those patients was artificially protected from showing any relapse in alcoholism. In Lundbeck14, more nalmefene recipients than placebo recipients had their data handled in this way because of the higher withdrawal rates in recipients of active treatment.

Three of the four co-primary end-points can be considered statistically significant by MMRM: reduction in HDD was significant in both pivotal studies, but reduction in TAC was only significant in one of the pivotal studies (Lundbeck14). According to the Sponsor's prospective analysis plan, this means that Lundbeck23 should be considered negative overall, because it did not achieve significance for both of its primary endpoints.

Evolution of the two co-primary efficacy variables over time in both of the pivotal studies is shown in the figure below, reproduced from the sponsor's Clinical Overview. Importantly, the provided figure omits the pre-treatment assessment of the efficacy variables at Randomisation, which occurred between baseline ('B') and the subsequent post-treatment assessment at Month 1. This omission disguises an important feature of the results, discussed later, which was that many subjects showed a major improvement prior to commencing pharmacological treatment.

Figure 5. Changes from Baseline in HDDs (days/month) and TAC (g/day) (FAS, MMRM) – Total Population – Lundbeck14 (left) and Lundbeck23 (right)



More context is provided in the tables below for Lundbeck14 (top table) and Lundbeck23 (bottom table), including the mean Baseline value of the efficacy variables, the mean changes from Baseline, and the mean group differences (and 95%CI) in the changes from Baseline. This format reveals that the between-group differences for both HDD and TAC, despite being

statistically significant, were small as a proportion of the changes from Baseline, and smaller still as a proportion of total Baseline values.

For instance, for number of HDDs per month, the mean reduction achieved with active treatment in Lundbeck14, relative to placebo, was 2.3 days from an initial 19.4 days; the 95%CI included the possibility that the treatment effect was less than one HDD per month. In the second study, Lundbeck23, the difference relative to placebo in HDDs per month was 1.7 days from a baseline value of 19.8 days, and the 95%CI was consistent with a mean treatment effect of less than half a day per month. Note that the HDD reduction does not necessarily represent days in which heavy drinking was replaced by days of abstinence, but could represent days of moderate rather than heavy drinking.

For TAC, the mean difference achieved in Lundbeck14, relative to placebo, was 11g/day from a baseline intake of 84 g/day. In Lundbeck23, the difference was 5g/day from a baseline of 93 g/day, and the 95%CI was consistent with the treatment effect being less than one gram of alcohol per day. A standard drink contains 10g of alcohol, so the treatment effect in Lundbeck23 was about half a drink and potentially less than a tenth of a drink.

Of the two pivotal studies, Lundbeck23 is more likely to be reflective of results in real life, because withdrawal in the treatment groups was reasonably similar. In this study, the estimated size of the treatment effect was small, and the overall result was statistically negative.

Thus, despite being statistically significant for 3 of 4 pivotal co-primary endpoints, the clinical effect demonstrated for nalmefene treatment in the two pivotal studies appears to be modest, particularly in Lundbeck23, the more reliable study. These results are even less impressive if considered in the context of a potential PK effect, demonstrated in Study 13513A, that increases exposure to ethanol by ~9% when ethanol is taken with nalmefene. Furthermore, if the observed treatment effect has been even slightly inflated by unblinding, the true benefit in clinical practice could be marginal.

Table 30. Results for the Co-Primary Efficacy Variables at Month 6 (FAS) – Lundbeck14

Variable / Treatment Group	Baseline		Change from Baseline to Month 6		Difference to PBO		
	N	Mean ± SD	N	Mean ± SE	Mean ± SE	95% CI	p-value
Number of HDDs (days/month)							
MMRM							
PBO	289	19.6 ± 6.9	213	-8.9 ± 0.6			
NMF	290	19.4 ± 7.3	152	-11.2 ± 0.6	-2.3 ± 0.8	[-3.8; -0.8]	0.002
LOCF							
PBO	289	19.6 ± 6.9	289	-8.4 ± 0.6			
NMF	290	19.4 ± 7.3	290	-10.2 ± 0.6	-1.7 ± 0.7	[-3.0; -0.4]	0.010
TAC (g/day)							
MMRM							
PBO	289	85 ± 42	213	-39.7 ± 2.2			
NMF	290	84 ± 42	152	-50.7 ± 2.4	-11.0 ± 3.0	[-16.8; -5.1]	<0.001
LOCF							
PBO	289	85 ± 42	289	-37.7 ± 2.3			
NMF	290	84 ± 42	290	-46.5 ± 2.3	-8.8 ± 2.8	[-14.3; -3.3]	0.002

Baseline values are based on FAS, OC.

Table 31. Results for the Co-Primary Efficacy Variables at Month 6 (FAS) – Lundbeck23

Variable / Treatment Group	Baseline		Change from Baseline to Month 6		Difference to PBO		
	N	Mean ± SD	N	Mean ± SE	Mean ± SE	95% CI	p-value
Number of HDDs (days/month)							
MMRM							
PBO	326	18.3 ± 7.0	229	-10.6 ± 0.5			
NMF	329	19.8 ± 6.8	212	-12.3 ± 0.5	-1.7 ± 0.7	[-3.1; -0.4]	0.012
LOCF							
PBO	326	18.3 ± 7.0	326	-10.0 ± 0.5			
NMF	329	19.8 ± 6.8	329	-11.8 ± 0.5	-1.8 ± 0.6	[-3.0; -0.6]	0.004
TAC (g/day)							
MMRM							
PBO	326	89 ± 48	229	-54.1 ± 2.2			
NMF	329	93 ± 46	212	-59.0 ± 2.3	-5.0 ± 2.9	[-10.6; 0.7]	0.088
LOCF							
PBO	326	89 ± 48	326	-51.7 ± 2.2			
NMF	329	93 ± 46	329	-57.6 ± 2.2	-5.9 ± 2.6	[-11.1; -0.7]	0.026

Baseline values are based on FAS, OC.

The Sponsor performed a number of sensitivity analyses, with different methods of imputation. With optimistic imputation methods, such as LOCF, all four co-primary endpoints became positive. With pessimistic imputation methods, such as BOCF, all four became negative (and for Lundbeck14, BOCF imputation actually produced a trend in favour of placebo for both endpoints). The results of these analyses are displayed graphically below.

Table 32. Sensitivity Analyses – Changes from Baseline to Month 6 in HDDS (days/month) – Total Population – Lundbeck14 (left) and Lundbeck23 (right)

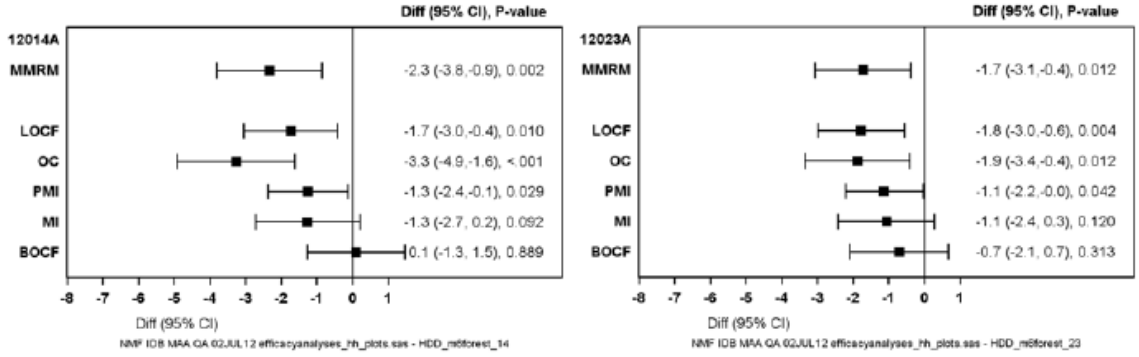
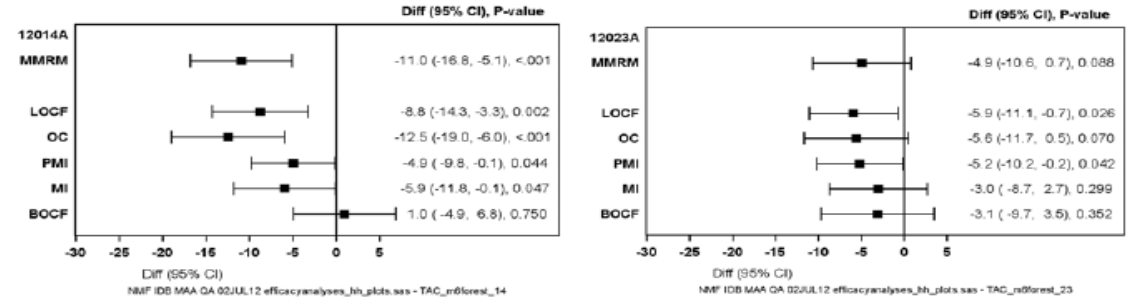


Table 33. Sensitivity Analyses – Changes from Baseline to Month 6 in TAC (g/day) – Total Population – Lundbeck14 (left) and Lundbeck23 (right)



This imputation analysis suggests that the results are not robust, and that the conclusion drawn about each study depends on the assumptions made about withdrawing subjects.

7.1.13. Results in low-risk and high-risk drinkers

Three different approaches were taken to perform subgroup analyses of high-risk drinkers, based on three different definitions of high risk:

- at-least high DRL at Baseline (HDAB, the prospectively identified high-risk group)
- at-least medium DRL at Baseline and Randomisation (MDAR, post hoc)
- at-least high DRL at Baseline and Randomisation (HDAR, post hoc)

The third of these methods is considered in a separate section in view of the fact that the Sponsor decided, post hoc, to consider the HDAR subgroup as 'the target group' for nalmefene.

The prospectively identified high-risk subgroup consisted of subjects with high or very high DRL at Baseline (HDAB). Results in this subgroup were positive for both efficacy variables in Lundbeck14 (upper panel, below) but negative for TAC in Lundbeck23 (lower panel, below). When the two studies were combined, results for both efficacy variables were significantly positive in this high-risk subgroup. Conversely, results in subjects with medium DRL at Baseline were not significant.

Table 34. Lundbeck14 Results in Patients with High or Very High DRL at Baseline for Co-Primary Efficacy Analyses (MMRM; FAS; OC; Main Treatment Period)

Efficacy Variable Treatment Group	Baseline I		Adjusted Change from Baseline I to Month 6		Difference to PBO		
	N	Mean ± SD	N	Mean ± SE	Mean ± SE	95% CI	p-value
Number of HDDs							
PBO	230	21.6 ± 6.2	165	-9.9 ± 0.7			
NMF	222	21.8 ± 6.3	111	-12.5 ± 0.8	-2.6 ± 0.9	[-4.4; -0.7]	0.006
TAC							
PBO	230	95 ± 40	165	-46.6 ± 2.8			
NMF	222	96 ± 41	111	-58.8 ± 3.0	-12.2 ± 3.6	[-19.3; -5.2]	<0.001

Baseline I values were based on FAS, OC; changes from Baseline I and differences to placebo were based on MMRM; FAS, OC values.

Table 35. Lundbeck23 Results in Patients with High or Very High DRL at Baseline for Co-Primary Efficacy Analyses (MMRM; FAS; OC; Main Treatment Period)

Efficacy Variable Treatment Group	Baseline I		Adjusted Change from Baseline I to Month 6		Difference to PBO		
	N	Mean ± SD	N	Mean ± SE	Mean ± SE	95% CI	p-value
Number of HDDs							
PBO	247	20.6 ± 6.3	175	-11.5 ± 0.6			
NMF	265	21.6 ± 5.9	176	-13.6 ± 0.7	-2.1 ± 0.8	[-3.7; -0.5]	0.010
TAC							
PBO	247	103 ± 48	175	-61.1 ± 2.8			
NMF	265	103 ± 45	176	-67.8 ± 2.8	-6.7 ± 3.6	[-13.6; 0.3]	0.062

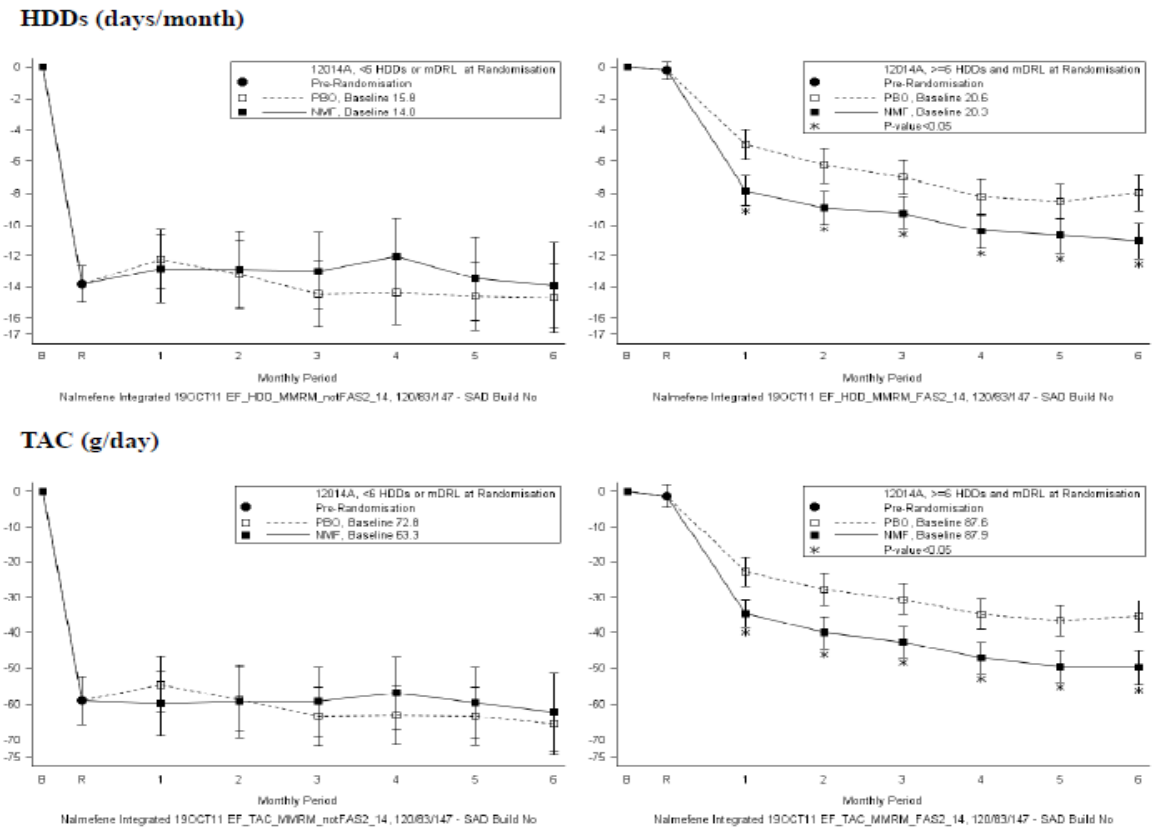
Baseline I values were based on FAS, OC; changes from Baseline I and differences to placebo were based on MMRM; FAS, OC values.

In another approach (employed post hoc but given greater emphasis in the Sponsor's SCE and Clinical Overview than the prospectively declared method), the Sponsor subdivided the population according to the subjects' alcohol consumption at Randomisation and Baseline, rather than just at Baseline. Although all subjects were required to have an alcohol intake consistent with at least medium DRL and ≥ 6 HDDs/month at Baseline, some had improved by the time of Randomisation to the extent that they would not have qualified for the study if they

had shown the same low intake at Screening. According to the protocol, all of these subjects were still eligible for inclusion in the FAS, because eligibility was defined by consumption at Screening, not by consumption at Randomisation. In retrospect, a better design would have involved re-screening the subjects for eligibility immediately prior to Randomisation. Subjects who lowered their intake to low DRL levels (or < 6 HDDs/month) by Randomisation can be referred to as the Low DRL At Randomisation (LDAR) subgroup, and those who continued to drink at medium risk levels or higher (and had ≥ 6 HDDs/month) can be referred to as the (at-least)Medium DRL At Randomisation (MDAR) subgroup.

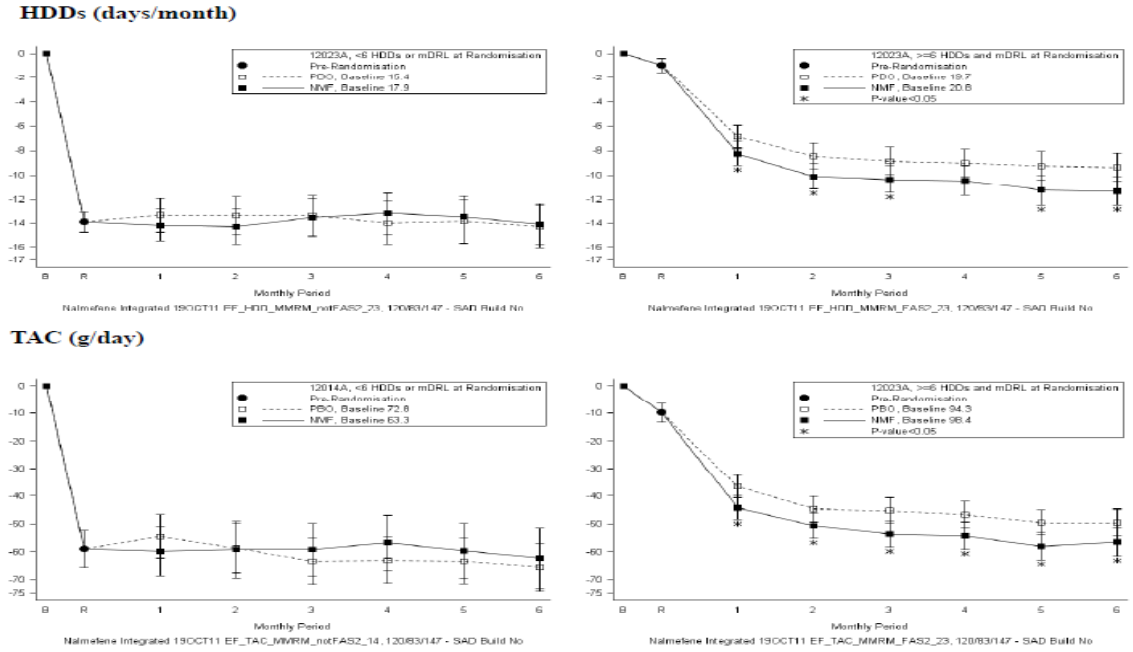
The figures below show the evolution of the primary efficacy variables in the LDAR group (left graphs in each panel) and those who continued to show at least-medium DRL and ≥ 6 HDDs/month at Randomisation (MDAR group, right graphs).

Figure 6. Changes from Baseline in HDDs and TAC in Patients Categorised According to Alcohol Consumption at Randomisation – Lundbeck14 (FAS, MMRM)



(Graphs on the left refer to subjects with < 6HDDs/month ***or*** < medium DRL at Randomisation ; graphs on the right refer to subjects with ≥ 6 HDDs/month ***and*** ≥ medium DRL at Randomisation).

Figure 7. Changes from Baseline in HDDs and TAC in Patients Categorised According to Alcohol Consumption at Randomisation – Lundbeck23 (FAS, MMRM)



The main feature seen in all of the graphs on the left is a major reduction in alcohol intake between Baseline and Randomisation, followed by minimal changes over the next six months. These graphs suggest that the subjects in the LDAR subgroup represent a distinct population who respond to non-pharmacological measures and may not need or benefit from pharmacological intervention. In fact, in both studies, those randomised to nalmefene from this LDAR subgroup showed a slightly higher intake at 6 months (relative to baseline) than those randomised to placebo, with no clear separation between the two treatment groups.

The LDAR group represent 18% and 33% of the enrolled patients in the Lundbeck14 and Lundbeck23 Studies, respectively, and the higher proportion seen in the Lundbeck23 study may account in part for the weaker results in this study. Despite the fact that they did not appear to benefit from treatment, these patients took study drug on 39% of days in the nalmefene group and on 48% of days in the placebo group.

Statistical analysis of the number of HDDs by MMRM, as shown in the table below, was consistent with a lack of a treatment effect in this LDAR subgroup (identified as the 'No' group in the table). Differences from placebo were either slightly positive, indicating higher drinking levels, or in the case of the LOCF analysis of Lundbeck23, slightly negative, but with 95%CIs including no difference.

The subsequent table shows a similar analysis for TAC, with the 'No' (LDAR) subgroup showing no significant treatment effect and even slightly adverse trends relative to placebo.

By contrast, both tables show a significant treatment effect in the 'Yes' group, consisting of the 82% of subjects with (at least) Medium DRL At Randomisation (MDAR). At Randomisation, these patients (by definition) still had ≥6 HDDs/month and at least a medium DRL; their alcohol consumption at Randomisation corresponded to a mean of 20.0 HDDs/month in both the nalmefene and placebo groups and a mean TAC of 84g/day and 88g/day in the two groups, respectively.

Unlike the primary efficacy analysis performed on the total population of each study, which failed to show significance for the TAC endpoint in Lundbeck23, this post hoc subgroup analysis achieved significance for both HDD and TAC in both Lundbeck14 and Lundbeck23.

Table 36. Adjusted Change from Baseline in HDDs (days/month) at Month 6, in Patients Categorised by Alcohol Consumption at Randomisation (FAS) – Lundbeck14 and Lundbeck23

>=6 HDDs and mDRL#		12014A PBO	12014A NMF	12023A PBO	12023A NMF
No	Baseline				
	N	58	44	105	113
	Mean	15.75	14.02	15.39	17.88
	SD	6.31	6.05	6.54	6.60
Yes	Baseline				
	N	231	246	221	216
	Mean	20.57	20.34	19.71	20.84
	SD	6.76	7.10	6.83	6.61
MMRM, No	Adjusted Change				
	N	47	25	66	69
	Mean	-14.69	-13.92	-14.26	-14.09
	95% CI	(-16.88;-12.50)	(-16.66;-11.18)	(-16.03;-12.49)	(-15.78;-12.40)
	Difference to PBO				
	Mean		0.77		0.17
	95% CI		(-2.53;4.08)		(-2.13;2.47)
	P-value		0.6464		0.8861
MMRM, Yes	Adjusted Change				
	N	166	127	163	143
	Mean	-7.97	-11.04	-9.33	-11.34
	95% CI	(-9.10;-6.84)	(-12.23;-9.84)	(-10.46;-8.20)	(-12.54;-10.14)
	Difference to PBO				
	Mean		-3.07		-2.01
	95% CI		(-4.62;-1.51)		(-3.57;-0.45)
	P-value		0.0001		0.0117
LOCF, No	Adjusted Change				
	N	58	44	105	113
	Mean	-14.16	-12.75	-13.61	-14.00
	95% CI	(-16.39;-11.94)	(-15.32;-10.18)	(-15.33;-11.90)	(-15.69;-12.31)
	Difference to PBO				
	Mean		1.41		-0.39
	95% CI		(-1.73;4.55)		(-2.48;1.71)
	P-value		0.3769		0.7168
LOCF, Yes	Adjusted Change				
	N	231	246	221	216
	Mean	-7.18	-9.90	-8.50	-10.68
	95% CI	(-8.33;-6.04)	(-11.01;-8.78)	(-9.65;-7.36)	(-11.88;-9.49)
	Difference to PBO				
	Mean		-2.71		-2.18
	95% CI		(-4.13;-1.30)		(-3.63;-0.72)
	P-value		0.0002		0.0034

Alcohol consumption at randomisation is based on the TLFB data collected between the Screening Visit and Randomisation, extrapolated to 4 weeks.
Patients were categorised (no,yes) as having at least 6 HDDs and at least medium DRL based on their alcohol consumption at Randomisation
Nalmefene IDB HDD_IntFAS2_1423 03NOV2011:12:04:29 1001/120/83/147 - SAD Build Numbers

Table 37. Adjusted Change from Baseline in TAC (g/day) at Month 6, in Patients Categorised by Alcohol Consumption at Randomisation (FAS) – Lundbeck14 and Lundbeck23

>=6 HDDs and mDRL#		12014A PBO	12014A NMF	12023A PBO	12023A NMF
No	Baseline				
	N	58	44	105	113
	Mean	72.8	63.3	79.3	82.1
	SD	42.1	28.1	49.3	41.4
Yes	Baseline				
	N	231	246	221	216
	Mean	87.6	87.9	94.3	98.4
	SD	41.1	43.1	46.9	47.6
MMRM, No	Adjusted Change				
	N	47	25	66	69
	Mean	-65.6	-62.4	-66.7	-65.3
	95% CI	(-74.2;-57.0)	(-73.2;-51.5)	(-74.2;-59.1)	(-72.7;-57.9)
	Difference to PBO				
	Mean		3.2		1.4
	95% CI		(-9.7;16.2)		(-8.3;11.1)
	P-value		0.6228		0.7837
MMRM, Yes	Adjusted Change				
	N	166	127	163	143
	Mean	-35.3	-49.8	-49.5	-56.5
	95% CI	(-39.8;-30.7)	(-54.6;-45.0)	(-54.3;-44.6)	(-61.6;-51.3)
	Difference to PBO				
	Mean		-14.5		-7.0
	95% CI		(-20.6;-8.5)		(-13.6;-0.4)
	P-value		<.0001		0.0371
LOCF, No	Adjusted Change				
	N	58	44	105	113
	Mean	-63.7	-56.5	-65.2	-65.6
	95% CI	(-72.9;-54.4)	(-67.1;-45.9)	(-72.5;-57.8)	(-72.9;-58.2)
	Difference to PBO				
	Mean		7.2		-0.4
	95% CI		(-5.8;20.2)		(-9.4;8.6)
	P-value		0.2795		0.9323
LOCF, Yes	Adjusted Change				
	N	231	246	221	216
	Mean	-32.1	-45.6	-46.2	-54.3
	95% CI	(-36.9;-27.3)	(-50.2;-40.9)	(-51.1;-41.2)	(-59.4;-49.3)
	Difference to PBO				
	Mean		-13.5		-8.1
	95% CI		(-19.3;-7.6)		(-14.4;-1.9)
	P-value		<.0001		0.0110

Alcohol consumption at randomisation is based on the TLFB data collected between the Screening Visit and Randomisation, extrapolated to 4 weeks.

This post hoc analysis is somewhat reassuring, suggesting that properly targeted patients may show consistent benefit for across the major efficacy variables, but it also suggests that a small proportion of subjects who initially appear to be eligible (~1 in 5 to 1 in 3) will show a reasonable response to non-pharmacological measures, and that medication does not increase the benefit in such subjects. This should be reflected in the Product Information, so that subjects are not treated unnecessarily. Given that this represents a post hoc revision of the target group, the utility of nalmefene in higher-risk (MDAR) subjects identified in this manner should also be confirmed prospectively.

7.1.14. Results in high-risk drinkers with at-least-high DRL at Baseline and Randomisation

A third approach to analysing high-risk subjects was to consider those with high or very high DRL at both Baseline and Randomisation (at-least High DRL At Randomisation, HDAR); this represents an even higher risk subgroup within the MDAR subgroup. These subjects represent an important target for therapy, because they had a high risk of alcohol-related harm prior to enrolment and did not respond to the non-pharmacological measures introduced prior to Randomisation. They are also the group that the Sponsor now considers to be the "target" group for nalmefene.[13]

Across both pivotal studies, the total number of subjects in this high-risk subgroup was 667, which is ~50% of the total study population (n = 1322). On average, these subjects had ~23 HDDs per month at Baseline, well in excess of the ≥ 6 HDDs required for eligibility.

Table 38. Baseline Efficacy Summary – Patients with a High or Very High DRL at Baseline and Randomisation – Lundbeck14 and Lundbeck23, Pooled

	PBO	NMF	Total
Number of Patients	332	335	667
DRL, n (%)			
High	139 (41.9)	124 (37.0)	263 (39.4)
Very high	193 (58.1)	211 (63.0)	404 (60.6)
Number of HDDs (days/month)			
n	332	335	667
Mean ± SD	22.4 ± 6.0	22.9 ± 5.9	22.6 ± 5.9
TAC (g/day)			
n	332	335	667
Mean ± SD	103.3 ± 44.5	107.7 ± 45.5	105.5± 45.0
CGI-S			
n	331	335	666
Mean ± SD	4.3 ± 1.4	4.3 ± 1.4	4.3 ± 1.4
DrInc-2R Total Score			
n	329	333	662
Mean ± SD	42.2 ± 22.2	41.1 ± 22.3	41.6 ± 22.2
ADS Total Score			
n	331	334	665
Mean ± SD	13.3 ± 5.7	14.0 ± 6.0	13.7 ± 5.8

Patient withdrawals in this high-risk group were quite high (47% in Lundbeck14 and 36% in Lundbeck23), with an excess of withdrawals in the nalmefene group compared to the placebo group in Lundbeck14, but identical withdrawal rates in Lundbeck23, as shown in the table below. Amongst these high-risk subjects, the main reason for nalmefene withdrawals in Lundbeck14 was adverse events, but this was not a major issue in Lundbeck23 (as shown in the second table below). Reasons for this difference are not clear, but might be cultural, given that the studies were conducted in different regions but had identical designs. As in the primary efficacy analysis, the post hoc subgroup results of Lundbeck14 should be considered less reliable than Lundbeck23.

[13] The evaluator points out that the sponsor criticised the first-round Clinical Evaluation Report for not highlighting this particular subgroup analysis and giving it the prominence they felt it deserved. Despite the sponsor's preference for discussing the results achieved in this subgroup, it remains one of many possible subgroups that could have been identified post hoc, and the HDAR results do not constitute an intrinsic component of the pivotal studies as they were conceived and conducted. This group has been given somewhat more prominence in the second-round CER, but the sponsor's post hoc treatment of these results is considered statistically invalid.

Table 39. Patient Disposition – Patients with a High or Very High DRL at Baseline and Randomisation – Lundbeck14 and Lundbeck23

	12014A			12023A		
	PBO	NMF	Total	PBO	NMF	Total
Patients randomised	170	180	350	162	155	317
Patients treated	169 (100%)	179 (100%)	348 (100%)	158 (100%)	152 (100%)	310 (100%)
Patients completed	107 (63%)	77 (43%)	184 (53%)	101 (64%)	97 (64%)	198 (64%)
Patients withdrawn	62 (37%)	102 (57%)	164 (47%)	57 (36%)	55 (36%)	112 (36%)
FAS	167	171	338	155	148	303

Table 40. Withdrawals by Primary Reason – Patients with a High or Very High DRL at Baseline and Randomisation – Lundbeck14 and Lundbeck23

	12014A		12023A	
	PBO n (%)	NMF n (%)	PBO n (%)	NMF n (%)
FAS	167 (100)	171 (100)	155 (100)	148 (100)
Patients completed[a]	114 (68)	85 (50)	111 (72)	103 (70)
Patients withdrawn	53 (32)	86 (50)	44 (28)	45 (30)
Primary Reason				
Adverse events	11 (7)	41 (24)	3 (2)	4 (3)
Lack of efficacy	15 (9)	12 (7)	8 (5)	4 (3)
Non-compliance		5 (3)	1 (1)	1 (1)
Protocol violation	2 (1)	5 (3)	8 (5)	10 (7)
Withdrawal of consent	17 (10)	16 (9)	15 (10)	15 (10)
Lost to follow-up	5 (3)	5 (3)	4 (3)	3 (2)
Other	3 (2)	2 (1)	5 (3)	8 (5)

a Patients with TLFB data at Month 6

The figure below shows the results in this subgroup for both co-primary endpoints, with Lundbeck14 results on the left and Lundbeck23 results on the right. Although the error bars overlap in most of the curves, the treatment effect is apparent after the first month of treatment, and remains consistent throughout the 6 month studies. The subsequent tables show the quantified results for each primary endpoint and each study; all endpoints were significantly positive in this post hoc subgroup.

Figure 8. Post Hoc Analysis: Patients with a High or Very High DRL at Baseline and Randomisation – Changes from Baseline in HDDs (days/month) and TAC (g/day) – Lundbeck14 and Lundbeck23 (FAS, MMRM)

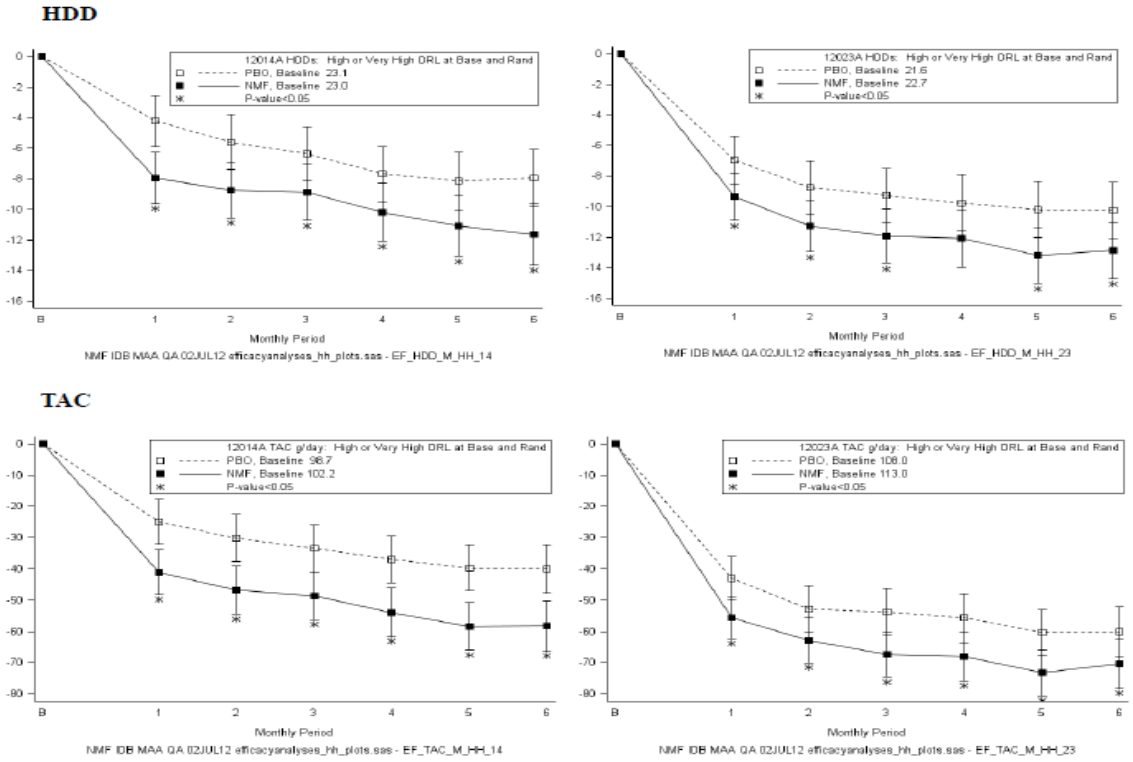


Table 41. Post Hoc Analysis: Patients with a High or Very High DRL at Baseline and Randomisation – Changes from Baseline in HDDs (days/month) – Lundbeck14 and Lundbeck23 (FAS)

Variable / Treatment Group	Baseline		Change from Baseline to Month 6		Difference to PBO		
	N	Mean ± SD	N	Mean ± SE	Mean ± SE	95% CI	p-value
Study 12014A							
MMRM							
PBO	167	23.1 ± 5.4	114	-8.0 ± 1.0			
NMF	171	23.0 ± 5.9	85	-11.6 ± 1.0	-3.7 ± 1.1	[-5.9; -1.5]	0.001
LOCF							
PBO	167	23.1 ± 5.4	167	-7.6 ± 1.0			
NMF	171	23.0 ± 5.9	171	-10.6 ± 1.0	-3.0± 1.0	[-4.8; -1.1]	0.002
Study 12023A							
MMRM							
PBO	155	21.6 ± 6.4	111	-10.2 ± 0.9			
NMF	148	22.7 ± 6.0	103	-12.9 ± 0.9	-2.7 ± 1.2	[-5.0 -0.3]	0.025
LOCF							
PBO	155	21.6 ± 6.4	155	-9.5 ± 0.9			
NMF	148	22.7 ± 6.0	148	-12.2 ± 0.9	-2.7 ± 1.1	[-4.8 -0.6]	0.013

Baseline values are based on FAS, OC.

Table 42. Post Hoc Analysis: Patients with a High or Very High DRL at Baseline and Randomisation – Changes from Baseline in TAC (g/day) –Lundbeck14 and Lundbeck23 (FAS)

Variable Treatment Group	Baseline		Change from Baseline to Month 6		Difference to PBO		
	N	Mean ± SD	N	Mean ± SE	Mean ± SE	95% CI	p-value
Study 12014A							
MMRM							
PBO	167	99 ± 40	114	-40.0 ±3.9			
NMF	171	102 ± 43	85	-58.3± 4.1	-18.3 ± 4.4	[-26.9; -9.7]	<0.001
LOCF							
PBO	167	99 ± 40	167	-37.8 ± 4.1			
NMF	171	102 ± 43	171	-54.2 ± 4.1	-16.3 ± 4.0	[-24.2; -8.4]	<0.001
Study 12023A							
MMRM							
PBO	155	108 ± 47	111	-60.1 ± 4.0			
NMF	148	113 ± 48	103	-70.4 ± 4.0	-10.3 ± 5.0	[-20.2; -0.5]	0.040
LOCF							
PBO	155	108 ± 47	155	-57.7 ± 4.0			
NMF	148	113 ± 48	148	-68.6 ± 3.9	-10.9 ± 4.6	[-20.0; -1.8]	0.019

Baseline values are based on FAS, OC.

Even in this post hoc subgroup, though, the clinical value of the observed results remains modest. In terms of HDDs, nalmefene was associated with a reduction of 3.7 or 2.7 days/month relative to placebo, in Lundbeck14 and Lundbeck23, respectively, by the MMRM method. The reduction in TAC was 18.3 and 10.3 g/day in the two studies, respectively, with the 95%CI in Lundbeck23 extending to within 0.5g/d of no effect; this represents one twentieth of a standard drink. The results were more favourable in Lundbeck14, but this study had a much higher withdrawal rate and was potentially more susceptible to withdrawal bias. (The Sponsor dismissed the possibility of withdrawal bias, but provided no alternative explanation for the superior results in Lundbeck14).

The Sponsor performed a number of sensitivity analyses to determine the robustness of these results, and the results are shown in the figures below.

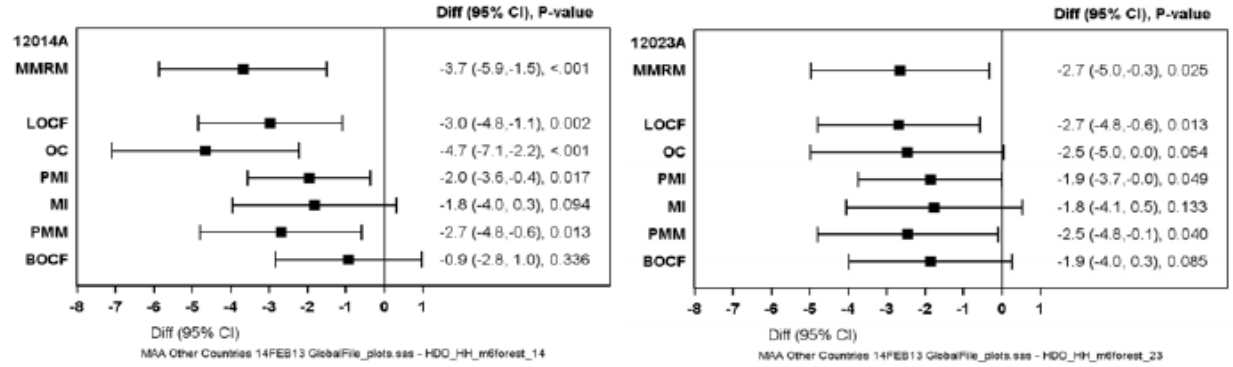
In Lundbeck14, significant improvements in HDDs were demonstrated by most analysis methods, but not for the Multiple Imputation (MI) method or Baseline Observation Carried Forward (BOCF) method. Given that the MI method attempts to minimise withdrawal bias by modelling the unequal withdrawal from two different treatment groups, this sensitivity analysis casts some doubt on the validity of the results, but the TAC results in this study were positive for every imputation method except BOCF. The failure of the BOCF method to give positive results is of less concern, because it is a very pessimistic imputation method, locking in the high Baseline drinking levels in all withdrawing patients, as if those subjects all reverted to their previous drinking habits. According to the Sponsor, there is empirical evidence in the alcohol-dependence literature that this does not occur, and that alcohol consumption remains reduced for months after an intervention, but it is doubtful that such empirical evidence was based on subjects who withdrew from study, so its applicability to questions of withdrawal bias is limited.

Results in Lundbeck23 were less favourable, with most imputation methods failing to produce a statistically significant result for HDDs and TAC even in this post hoc subgroup. The optimistic approach of imputing the last observation (LOCF) produced positive results, the pessimistic approach of imputing the Baseline value (BOCF) produced negative results, and most other methods produced results intermediate between those two. Multiple Imputation (MI) methods showed no significant treatment effect in this subgroup for either co-primary endpoint in Lundbeck23; although the mean treatment effect by the MI method was numerically similar to

that shown in the primary analysis, the 95%CI was broad and crossed the no-effect line. This may reflect the fact that the MI method requires modelling of the data with many assumptions, each of which adds uncertainty, and the study was not originally powered for this approach. Nonetheless, this sensitivity analysis highlights the fact that results in Lundbeck23 were weak, even with post hoc selection of a more responsive subgroup – and recall that that Lundbeck23 was actually a negative study according to the primary efficacy analysis.

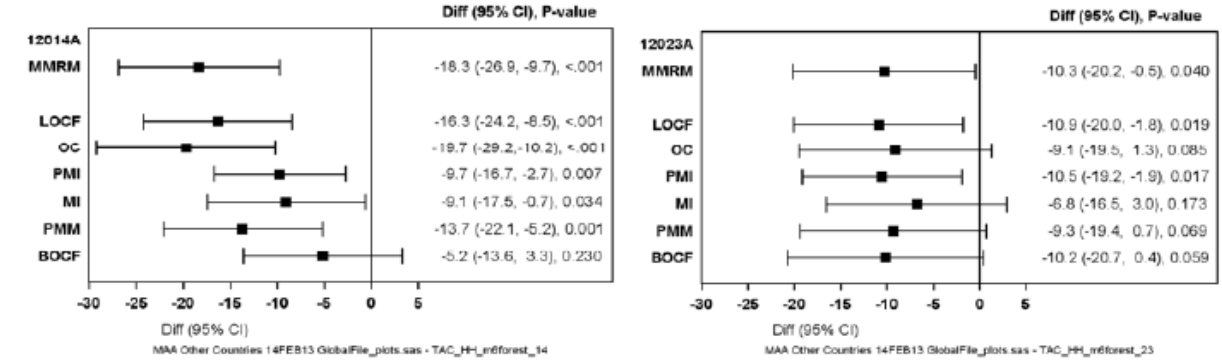
The discrepancy between the two pivotal studies casts doubt on the overall results, particularly in view of the fact that Lundbeck23, which was less likely to be affected by withdrawal bias, produced the weaker estimates of the treatment effect. (It should also be noted that the prospectively identified high-risk subgroup – subjects with high or very high DRL at Baseline – did not show a significant treatment effect for reduction in TAC in Lundbeck23.)

Figure 9. Patients with a High or Very High DRL at Baseline and Randomisation – Changes from Baseline in HDDs (days/month) – Lundbeck14 and Lundbeck23 (FAS)



MMRM is the co-primary efficacy analysis and is included for comparison.
Negative values indicate a greater reduction in the number of HDDs and TAC in the nalmefene group than in the placebo group.

Figure 10. Patients with a High or Very High DRL at Baseline and Randomisation – Changes from Baseline in TAC (g/day) –Lundbeck14 and Lundbeck23 (FAS)



MMRM is the co-primary efficacy analysis and is included for comparison.
Negative values indicate a greater reduction in the number of HDDs and TAC in the nalmefene group than in the placebo group.

Despite the fact that this high-risk subgroup was identified post hoc, and was only one of three high-risk subgroups considered, the proposed PI only presents the results in this subgroup; that is, the overall primary results of the pivotal studies are not mentioned in the PI, the results in the prospectively identified high-risk subgroup are not mentioned, and the results in those with qualifying consumption (at-least medium DRL) at Randomisation are not mentioned, but this group's results are presented as the main outcome of the pivotal studies. Note that this subgroup differs from the initial, prospectively identified target population in two ways: the DRL of interest has been increased to at-least high (instead of at-least medium), and the

principal time at which the DRL criterion was applied has been shifted to Randomisation (as well as still being applied at Baseline) instead of being applied just at Baseline.[14] Thus, it represents an approach, with no correction for the use of multiple analyses. Such an emphasis on one subgroup with favourable results is not appropriate.

7.1.15. Responder analyses

The key secondary analysis for both pivotal studies was based on the response rate (RSDRL), where a response was defined as a downward shift in DRL category of at least two categories or a shift to low DRL or below. This endpoint was analysed with logistic regression (LREG) in both studies, but with a different imputation method in each: Lundbeck14 used Non-Response (NR) imputation, where missing data was pessimistically counted as a non-response; Lundbeck23 used individualised MMRM-predicted TAC to impute missing data. This method accounted for the alcohol consumption up to the point of the missing data and imputed a value consistent with what other subjects with a similar drinking history had consumed. (This imputation method is optimistic, because it assumes that withdrawing subjects continue on the same trajectory as continuing subjects, ignoring the possibility that the continuing subjects are likely to be more motivated than withdrawing subjects.)

Results for the RSDRL are shown in the tables below (Lundbeck14, first table; Lundbeck23, second table).

At Month 6 in Lundbeck14, the response rate was 44.3% in the placebo group, compared to 36.9% in the nalmefene group; in other words, by the prospective, main analysis method for this endpoint, the response rate was significantly inferior with active treatment ($p = 0.039$).

By contrast, at Month 6 in Lundbeck23, the response rate was 63.2% in the placebo group and 67.2% in the nalmefene group, consistent with a marginally better (but not statistically significant) response with active treatment ($p = 0.1833$).

Thus, by their own prospectively identified analysis methods, Lundbeck14 was not only negative for its key secondary endpoint but significantly favoured placebo, and Lundbeck23 was negative for its key secondary endpoint.

[14] The evaluator notes that drinking levels generally fell between Baseline and Randomisation, so the important timepoint at which this high-DRL subgroup was defined was at Randomisation. In many places in CER1, this important timepoint was emphasised in the interests of brevity, and it was not repeatedly stated that the relevant subgroup also showed at-least high DRL at Baseline. Thus, the CER1 version of the sentence was "Note that this subgroup differs from the initial, prospectively identified target population in two ways: the DRL of interest has been increased to at-least high, and the time at which the DRL criterion was applied has been shifted to Randomisation instead of Baseline." The fact that high DRL was also present at Baseline in the HDAR subgroup had already been covered in the initial description of the subgroup, so it was not repeated. The sponsor flagged this sentence as an error, and also objected to the expression 'at-least high DRL at Randomisation', insisting that the high-DRL at Baseline should be mentioned again during every reference to this subgroup. This nitpick has essentially no bearing on the issues at hand. Indeed, the nitpickiness of the sponsor's objection is revealed by considering their own PI, where the same abbreviation is used that was flagged as erroneous in CER1: "Therefore, the patients who maintained a high or very high DRL at randomisation were defined post hoc as the target population. In this post hoc population, the treatment effect was larger than that in the total population."

Table 43. Main Treatment Period, Lundbeck14: Adjusted Odds Ratio for Response, Shift in DRL (RSDRL, FAS, Non-Response Imputation, LREG)

Treatment Group	Month	N	Responders n	Responders (%)	Adjusted Odds Ratio with 95% CI Odds Ratio	Lower	Upper	LR Test p-value
PBO	1	289	98	(33.9)				
	2	289	124	(42.9)				
	3	289	132	(45.7)				
	4	289	131	(45.3)				
	5	289	132	(45.7)				
	6	289	128	(44.3)				
NMF	1	290	144	(49.7)	1.94	1.38	2.73	0.0001
	2	290	151	(52.1)	1.43	1.02	2.01	0.0394
	3	290	133	(45.9)	0.97	0.69	1.36	0.8519
	4	290	124	(42.8)	0.86	0.61	1.21	0.3896
	5	290	115	(39.7)	0.75	0.53	1.05	0.0907
	6	290	107	(36.9)	0.70	0.50	0.98	0.0390

Shift is defined as Baseline I Very High Risk DRL to Medium or Below or from Baseline I High or Medium Risk DRL to Low or Below

12014A FINAL ET_DRL44 02AUG2011:15:46:46 1001/120 - TGML/SAD Build Numbers

Table 44. Main Treatment Period, Lundbeck23: Adjusted Odds Ratio for Response, Shift in DRL (RSDRL, FAS, MMRM-Imputation, LREG)

Treatment Group	Month	N	Responders n	Responders (%)	Adjusted Odds Ratio with 95% CI Odds Ratio	Lower	Upper	LR Test p-value
PBO	1	326	171	(52.5)				
	2	326	194	(59.5)				
	3	326	197	(60.4)				
	4	326	201	(61.7)				
	5	326	201	(61.7)				
	6	326	206	(63.2)				
NMF	1	329	208	(63.2)	1.93	1.34	2.81	0.0004
	2	329	211	(64.1)	1.37	0.96	1.97	0.0854
	3	329	218	(66.3)	1.43	1.00	2.04	0.0495
	4	329	216	(65.7)	1.27	0.89	1.81	0.1835
	5	329	224	(68.1)	1.46	1.02	2.08	0.0385
	6	329	221	(67.2)	1.28	0.89	1.83	0.1833

Shift is defined as Baseline I Very High Risk DRL to Medium or Below or from Baseline I High or Medium Risk DRL to Low or Below
Missing values imputed by response based on individual patient-predicted TAC from the MMRM-model in primary analysis

12023A FINAL Report ET_DRL48 14SEP2011:11:34:28 1001/83 - TGML/SAD Build Numbers

The negative results for this key secondary endpoint across both pivotal studies represent a substantial failure of the study program to achieve a major, clinically relevant endpoint. In retrospect, this partly reflects the use of a pessimistic imputation method in Lundbeck14, where withdrawal rates were higher with active treatment (leading to a greater need for imputation and more imputed non-responses with active treatment). Poor results for this endpoint also reflect the weakness of the clinical treatment effect, though, because even optimistic imputation of values inferred from the MMRM model led to a non-significant result in Lundbeck23.

These unfavourable results were included in the individual study reports for each study, but they were de-emphasized in the Summary of Clinical Efficacy and the Clinical Overview.

The Clinical Overview provided the following explanation of the adverse results for this endpoint:

> *The results of the responder analyses using different imputation methods were consistently in favour of nalmefene, with the exception of the analysis that imputed missing values as non-response [see Panel 80 below]. However, for the majority of the patients who*

withdrew, their alcohol consumption was stable (or decreasing) up to the time of withdrawal. The data from the patients who withdrew and who had TLFB data after last dose of IMP indicate that the patients maintain their lower level of alcohol consumption after discontinuing IMP, and published data also indicate that patients are able to maintain a low stable alcohol consumption after withdrawing from a study. Thus, assuming non-response for all withdrawn patients is not supported by empirical evidence. Furthermore, the LOCF sustained response analysis confirmed the efficacy of nalmefene; this analysis is considered conservative and unlikely to be biased in favour of the nalmefene group as the patients who withdrew early (that is, prior to Month 2) were, by default, non-responders.

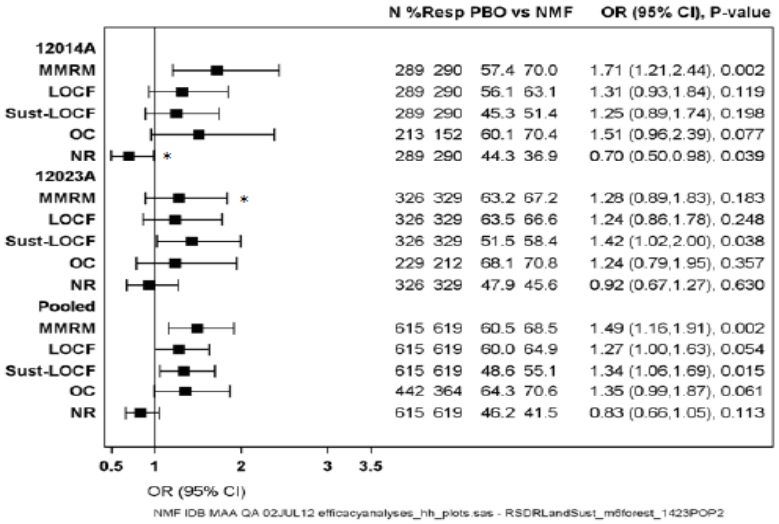
It is reasonable to suggest that the Non-Response imputation method was unnecessarily pessimistic. What is missing from this explanation, though, is an acknowledgement that the NR imputation method was the main prospective analysis method for this endpoint in a pivotal study, and was considered acceptable when the study was conceived.

The failure of both studies to achieve their key secondary endpoint was barely discussed by the Sponsor, just as the overall negative outcome of the primary endpoint for Lundbeck23 was barely mentioned; the Sponsor's provided summaries give roughly equal weight to the prospectively designated major statistical methods (which showed Lundbeck23 to be negative for its primary efficacy analysis and both studies to be negative for their key secondary analyses) and a variety of secondary or post hoc methods (which found positive results for the same variables). Thus, the Sponsor has presented the data in a more favourable light than is justified, using multiple statistical methods and then defending the ones that give favourable results. This shift in emphasis would not have been necessary if the prospective results had been robustly positive.

The figure below shows the results for RSDL in each study individually, by a variety of imputation methods, and then for the pooled data. Note that the main prospective imputation method was NR for Lundbeck14, and MMRM-predicted TAC for Lundbeck23. (These key results have been marked with an asterisk during preparation of this evaluation report, and were not marked in the original figure, which listed the main method for Lundbeck14 last.)

The pooled results resembled the results of the individual studies, with the results favouring nalmefene when MMRM-predicted TAC was used, but revealing a trend in favour of placebo when the pessimistic NR method was used. The attributable response rate in the pooled population, using MMRM, was 8% (60.5% with placebo and 68.5% with nalmefene), indicating that 12.5 patients would need to be treated to achieve one response.

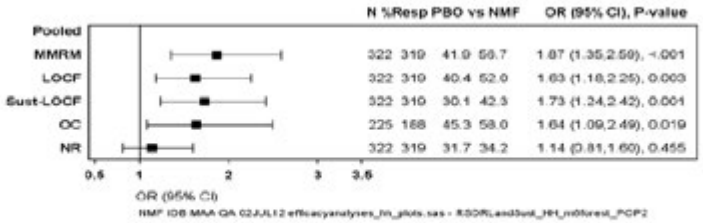
Figure 11. Key Secondary Analysis (RSDRL) – Proportion of Responders (%) and Odds Ratio for Response at Month 6 (FAS, LREG) – Lundbeck14, Lundbeck23, and Pooled Pivotal Studies



MMRM = missing values imputed using MMRM-predicted TAC
NR = missing values imputed as non-response

For high-risk drinkers (as defined by the Sponsor's preferred post hoc method, those with high or very high DRL at Baseline and Randomisation), the benefit in terms of response rate appeared to be greater than in the overall cohort, though there was still variation in the significance of the results depending on the imputation method. By the MMRM (OC) method, the response rate in this subgroup was 41.9% in the placebo group and 58.7% in the nalmefene group, an attributable difference of 16.8%, consistent with a number-needed-to-treat (NNT) of ~6 patients.

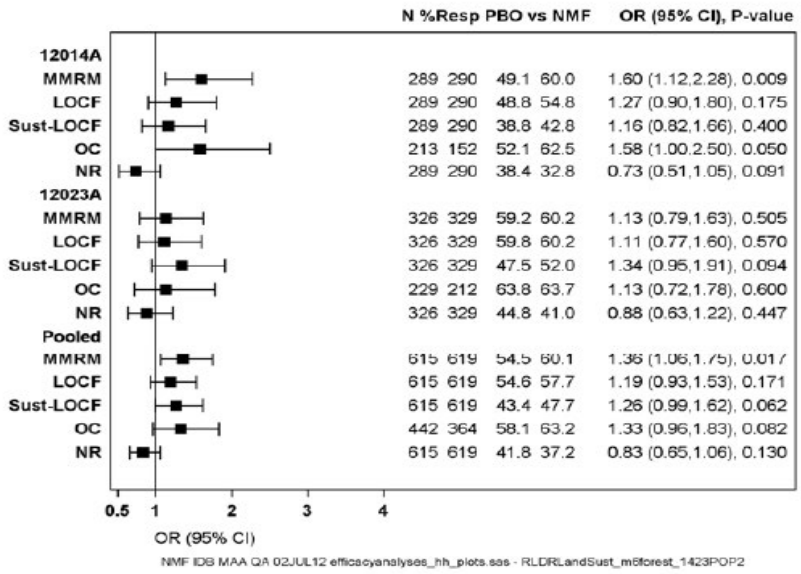
Figure 12. Post Hoc Analysis: Proportion (%) of Responders (RSDRL) and Odds Ratio for Response at Month 6 (FAS, LREG) – Patients with a High or Very High DRL at Baseline and Randomisation – Lundbeck14 and Lundbeck23, Pooled



MMRM = missing values imputed using MMRM-predicted TAC
NR = missing values imputed as non-response

Other definitions of response produced broadly similar results, and showed the same dependency on imputation method. For the response rate defined as a shift to low DRL (RLDRL), the individual studies were negative for most imputation methods, but the pooled analysis showed a positive treatment effect for the MMRM method. The pooled analysis of RLDL was negative for every other imputation method, highlighting the non-robustness of these results.

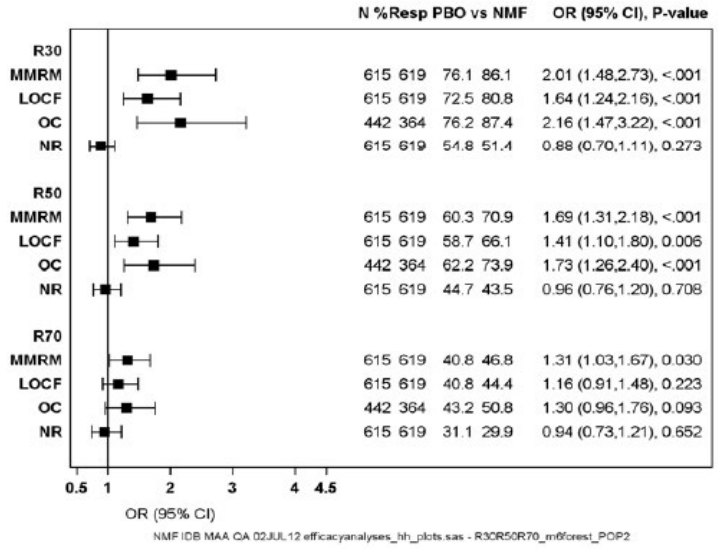
Figure 13. RLDRL – Proportion of Responders (%) and Odds Ratio for Response at Month 6 (FAS, LREG) – Lundbeck14, Lundbeck23, Individually and Pooled



MMRM = missing values imputed using MMRM-predicted TAC
NR = missing values imputed as non-response

When responses were defined according to percentage reductions in TAC (≥ 30%, ≥ 50%, ≥ 70%), the broader definitions showed more favourable treatment effects, as shown in the figure below. The MMRM imputation method produced positive results for all of these response definitions, but the attributable response rate was modest (~7 to 10%, depending on the definition of response).

Figure 14. Proportion of Responders (%) Based on ≥ 30%, ≥ 50% and ≥ 70% Reduction in TAC and OR for Response at Month 6 (FAS, LREG) – Lundbeck14 and Lundbeck23, Pooled



MMRM = missing values imputed using MMRM-predicted TAC
NR = missing values imputed as non-response

In their response to CER1, the Sponsor requested more emphasis on the results for this secondary efficacy variable in their chosen post hoc subgroup, the HDAR subjects.

The following table is derived from the proposed PI, but with a revised title – the original, PI version omitted any mention of the analysis being a post hoc analysis, and referred to Baseline as 'Screening'.

Table 45. Pooled *Post Hoc* Responder Analysis Results in Patients with a High or Very High DRL at Baseline and Randomisation

Response	Placebo	Nalmefene	Odds Ratio (95% CI)	p-value
MMRM[a]				
TAC R70[b]	25.8%	38.2%	1.88 (1.32; 2.70)	< 0.001
0-4 HDD[c]	20.5%	30.4%	1.91 (1.30; 2.83)	0.001
NR[d]				
TAC R70[b]	19.9%	25.4%	1.44 (0.97; 2.13)	0.067
0-4 HDD[c]	16.8%	22.3%	1.54 (1.02; 2.35)	0.040
a Analysis uses patient-predicted TAC or HDD values derived from the MMRM model in the primary analysis for patients who withdrew				
b ≥ 70% reduction from baseline in TAC at Month 6 (28-day period)				
c 0 to 4 HDDs/month at Month 6 (28-day period)				
d Analysis treats patients who withdrew as non-responder				

In this post hoc subgroup, the results look more favourable than in the overall population. This could reflect efficacy, the psychological impact of unblinding, the effects of withdrawal bias, the results of selecting a favourable dataset, or some combination of these. Unfortunately, the cited p-values are invalid because of the post hoc nature of the analysis and the Sponsor's failure to correct for multiple statistical analyses.

7.1.16. Results for other efficacy outcomes[15]

For the CGI-S and CGI-I, a significant treatment effect was observed in both studies at most time points, as shown in the tables and figures below. At the main time point of 6 Months, Lundbeck14 was clearly positive for both CGI-S and CGI-I (p < 0.001), but Lundbeck23 was only moderately positive for CGI-S (p = 0.029) and was negative for CGI-I (p = 0.111).

Table 46. Changes from Baseline to Week 24 in CGI-S (FAS, MMRM) – Lundbeck14

Treatment Group	Baseline		Change from Baseline to Week 24		Difference to PBO		
	N	Mean ± SD	N	Mean ± SD	Mean	95% CI	p-value
PBO	289	4.0 ± 1.5	210	-0.90 ± 0.08			
NMF	290	4.0 ± 1.5	152	-1.27 ± 0.08	-0.37 ± 0.10	[-0.57; -0.16]	<0.001
Baseline values are based on FAS, OC.							

[15] The evaluator notes that these are minor endpoints, and so they are only presented for the main study population, not for every subgroup.

Table 47. CGI-I Scores at Week 24 (FAS, MMRM) – Lundbeck14

Treatment Group	Score at Week 24		Difference to PBO		
	N	Mean ± SD	Mean	95% CI	p-value
PBO	210	2.65 ± 0.07			
NMF	152	2.30 ± 0.08	-0.34 ± 0.10	[-0.53; -0.15]	<0.001

Table 48. Changes from Baseline to Week 24 in CGI-S (FAS, MMRM) – Lundbeck23

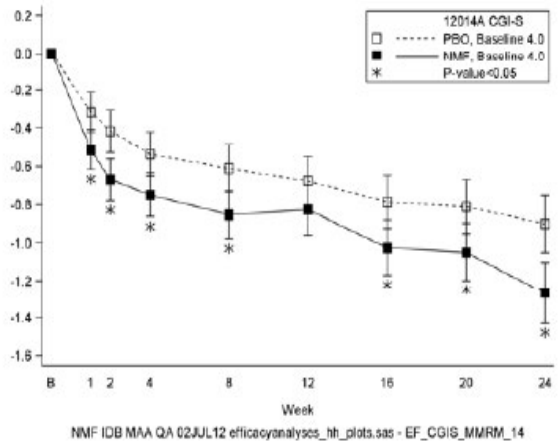
Treatment Group	Baseline		Change from Baseline to Week 24		Difference to PBO		
	N	Mean ± SD	N	Mean ± SD	Mean	95% CI	p-value
PBO	323	4.0 ± 1.4	225	-1.04 ± 0.08			
NMF	329	4.1 ± 1.4	203	-1.27 ± 0.08	-0.23 ± 0.11	[-0.44; -0.02]	0.029
Baseline values are based on FAS, OC.							

Table 49. CGI-I Scores at Week 24 (FAS, MMRM) – Lundbeck23

Treatment Group	Score at Week 24		Difference to PBO		
	N	Mean ± SD	Mean	95% CI	p-value
PBO	225	2.68 ± 0.08			
NMF	203	2.51 ± 0.08	-0.17 ± 0.11	[-0.38; -0.04]	0.111

Figure 15. Changes from Baseline in CGI-S Scores and CGI-I Scores – Lundbeck14

CGI-S



CGI-I

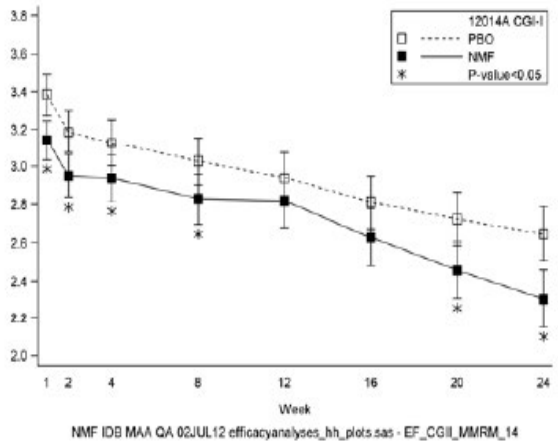
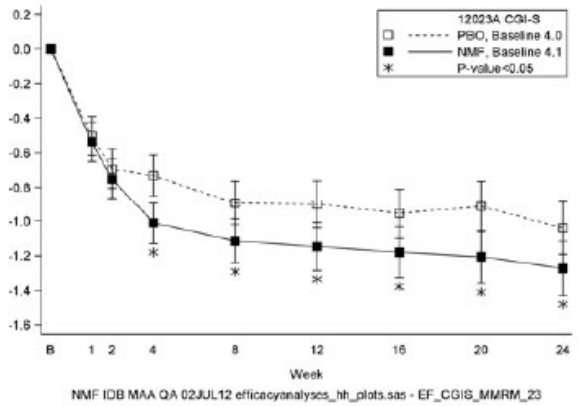
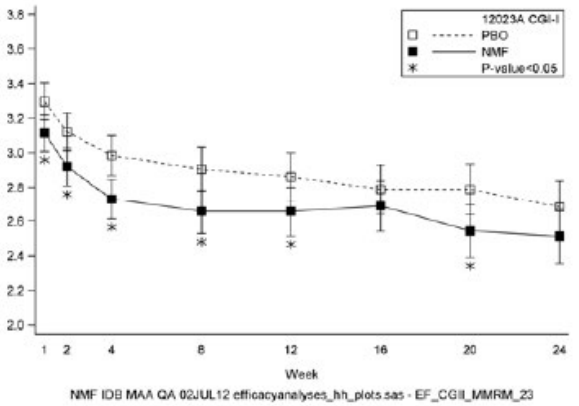


Figure 16. Changes from Baseline in CGI-S Scores and CGI-I Scores – Lundbeck23

CGI-S



CGI-I



For most other efficacy endpoints, the pattern observed in the primary analysis was repeated: results were clearly significant in Lundbeck14, but borderline in Lundbeck23. For efficacy as assessed by the surrogate endpoint of liver function tests, significant reductions in ALAT were observed in both studies, but the significance was marginal in Lundbeck23 (Lundbeck14 p = 0.011, Lundbeck23, p = 0.049). Significant reductions in GGT were only observed in Lundbeck14 (p=0.009), with Lundbeck23 showing no substantial trend (p = 0.529).

Table 50. GGT and ALAT at Week 24 (FAS, MMRM) – Lundbeck14

Variable Treatment Group	Geometric Mean at Baseline		Geometric Mean at Week 24		Ratio to PBO		
	N	Mean	N	Mean	Ratio	95% CI	p-value
GGT (IU/L)							
PBO	289	53.0	211	45.7			
NMF	290	51.0	158	40.3	0.88	[0.80; 0.97]	0.009
ALAT (IU/L)							
PBO	288	28.9	209	28.1			
NMF	290	29.1	158	25.4	0.90	[0.84; 0.98]	0.011

Baseline values are based on FAS, OC.

Table 51. GGT and ALAT at Week 24 (FAS, MMRM) – Lundbeck23

Variable Treatment Group	Geometric Mean at Baseline		Geometric Mean at Week 24		Ratio to PBO		
	N	Mean	N	Mean	Ratio	95% CI	p-value
GGT (IU/L)							
PBO	324	52.2	224	45.0			
NMF	329	52.6	207	43.4	0.96	[0.86; 1.08]	0.529
ALAT (IU/L)							
PBO	324	28.1	222	27.2			
NMF	329	28.8	205	25.0	0.92	[0.84; 1.00]	0.049

Baseline values are based on FAS, OC.

The magnitude of the changes was small for ALAT in all treatment groups, as shown above. The fall in GGT was greater, but the difference from placebo was minimal in Lundbeck23 (the attributable fall in GGT was 3.4 IU/L in Lundbeck14 and only ~2 IU/L in Lundbeck23).

The changes from baseline in Non-Drinking Days (NDDs) were disappointing in both studies, failing to achieve statistical significance. In Lundbeck14, the mean placebo-subtracted increase in NDDs was 1.2 days/month (p = 0.096) and in Lundbeck23, the mean increase was less than a day (0.6 days, p = 0.437). Thus, even temporary abstinence was not significantly promoted by active treatment.

Table 52. Changes from Baseline to Month 6 in NDDs (days/month) (FAS, MMRM) – Lundbeck14 and Lundbeck23

Study Treatment Group	Baseline		Change from Baseline to Month 6		Difference to PBO		
	N	Mean ± SD	N	Mean ± SE	Mean ± SE	95% CI	p-value
12014A							
PBO	289	5.5 ± 5.8	213	6.5 ± 0.5			
NMF	290	5.4 ± 6.1	152	7.7 ± 0.6	1.2 ± 0.7	[-0.2; 2.6]	0.096
12023A							
PBO	326	5.4 ± 5.9	229	8.6 ± 0.6			
NMF	329	5.0 ± 5.8	212	9.2 ± 0.6	0.6 ± 0.8	[-0.9; 2.1]	0.437

Baseline values are based on FAS, OC.

For a variety of psychosocial and quality-of-life (QOL) assessments, there was a trend in favour of active treatment in Lundbeck14, and for some measures (including many components of the SF-36 and one part of the EQ-5D), the between-group difference was statistically significant and in favour of active treatment in Lundbeck23.

Results for NDDs, ADS, DrInC-2R, SF-36 and EQ-5D in Lundbeck14 are shown in the three tables below. None of the p-values achieved significance.

Table 53. Summary of Minor Efficacy Variables (MMRM; FAS, OC) – Lundbeck14

Efficacy Variable Treatment Group	Baseline I		Adjusted Change from Baseline I to Month 6/Week 24		Difference to PBO		
	N	Mean ± SD	N	Mean ± SE	Mean ± SE	95% CI	p-value
Number of NDDs							
PBO	289	5.5 ± 5.8	213	6.5 ± 0.5			
NMF	290	5.4 ± 6.1	152	7.7 ± 0.6	1.2 ± 0.7	[-0.2; 2.6]	0.096
ADS[a]							
PBO	298	12.2 ± 4.9	289	-4.3 ± 0.3			
NMF	306	12.9 ± 5.8	290	-4.6 ± 0.3	-0.3 ± 0.4	[-1.1; 0.5]	0.432
DrInC-2R							
PBO	289	34.7 ± 18.2	213	-13.8 ± 1.0			
NMF	289	36.0 ± 18.7	158	-16.0 ± 1.1	-2.2 ± 1.3	[-4.7; 0.4]	0.100

Baseline I values were based on FAS, OC; changes from Baseline I and differences to placebo were based on MMRM; FAS, OC values.
Cross-reference: Tables 97 to 102

a Changes from Baseline I and differences to placebo were based on ANCOVA; FAS, BOCF values.

Table 54. Adjusted Mean Change from Baseline in SF-36 Subscale and Component Scores (MMRM; FAS, OC) – Lundbeck14

SF-36 Subscale Treatment Group	Baseline I		Adjusted Change from Baseline I to Week 24		Difference to PBO		
	N	Mean ± SD	N	Mean ± SE	Mean ± SE	95% CI	p-value
Physical functioning							
PBO	288	51.5 ± 6.7	211	0.8 ± 0.4			
NMF	286	51.2 ± 7.0	157	0.8 ± 0.4	0.0 ± 0.5	[-1.0; 1.0]	0.999
Role-physical							
PBO	289	47.9 ± 8.3	211	2.4 ± 0.5			
NMF	288	47.9 ± 8.5	158	3.0 ± 0.6	0.5 ± 0.7	[-0.9; 2.0]	0.448
Bodily pain							
PBO	289	50.8 ± 10.9	211	1.4 ± 0.7			
NMF	289	51.0 ± 10.8	159	2.3 ± 0.8	0.9 ± 0.9	[-1.0; 2.7]	0.346
General health							
PBO	289	46.8 ± 9.4	209	2.1 ± 0.5			
NMF	287	47.4 ± 9.6	157	2.2 ± 0.6	0.0 ± 0.7	[-1.3; 1.4]	0.937
Vitality							
PBO	287	49.0 ± 9.7	208	3.0 ± 0.6			
NMF	287	49.2 ± 10.0	159	2.0 ± 0.6	-0.9 ± 0.8	[-2.5; 0.6]	0.228
Social functioning							
PBO	286	47.2 ± 10.7	208	2.9 ± 0.6			
NMF	282	47.3 ± 9.7	157	3.2 ± 0.7	0.4 ± 0.8	[-1.2; 1.9]	0.664
Role-emotional							
PBO	289	45.3 ± 10.9	211	1.9 ± 0.7			
NMF	288	45.5 ± 10.3	157	3.6 ± 0.8	1.7 ± 0.9	[-0.1; 3.5]	0.066
Mental health							
PBO	288	45.5 ± 10.5	210	2.0 ± 0.6			
NMF	287	44.9 ± 10.6	158	2.4 ± 0.7	0.4 ± 0.8	[-1.3; 2.1]	0.625
Physical Component Summary							
PBO	289	51.3 ± 7.8	209	1.4 ± 0.4			
NMF	285	51.4 ± 8.5	152	1.5 ± 0.5	0.0 ± 0.6	[-1.1; 1.2]	0.941
Mental Component Summary							
PBO	289	44.7 ± 11.5	209	2.5 ± 0.7			
NMF	285	44.6 ± 11.2	152	3.2 ± 0.8	0.7 ± 0.9	[-1.1; 2.5]	0.454

Cross-reference: Tables 115 to 134

Table 55. Adjusted Mean Change in EQ-5D Subscale Scores (MMRM; FAS, OC) – Lundbeck14

EuroQol Variable Treatment Group	Baseline I		Adjusted Change from Baseline I to Week 24		Difference to PBO		
	N	Mean ± SD	N	Mean ± SE	Mean ± SE	95% CI	p-value
Utility index							
PBO	289	0.8 ± 0.2	210	0.0 ± 0.0			
NMF	289	0.8 ± 0.2	154	0.0 ± 0.0	0.0 ± 0.0	[0.0; 0.0]	0.581
Health state							
PBO	284	71.4 ± 17.7	208	3.4 ± 1.1			
NMF	283	71.4 ± 17.2	153	4.7 ± 1.2	1.3 ± 1.4	[-1.5; 4.1]	0.3675

Cross-reference: Tables 136 and 138

Similarly, results for NDDs, ADS, DrInC-2R, and SF-36 in Lundbeck23 are shown in the three tables below. Significant results were obtained in the SF-36 and one part (Health State) of the EQ-5D.

Table 56. Summary of Minor Efficacy Variables (MMRM; FAS, OC) – Lundbeck23

Efficacy Variable Treatment Group	Baseline I		Adjusted Change from Baseline I to Month 6/Week 24		Difference to PBO		
	N	Mean ± SD	N	Mean ± SE	Mean ± SE	95% CI	p-value
Number of NDDs							
PBO	326	5.4 ± 5.9	229	8.6 ± 0.6			
NMF	329	5.0 ± 5.8	212	9.2 ± 0.6	0.6 ± 0.8	[-0.9; 2.1]	0.437
ADS[a]							
PBO	359	14.6 ± 6.2	325	-5.1 ± 0.4			
NMF	357	14.5 ± 5.7	328	-5.9 ± 0.4	-0.7 ± 0.4	[-1.6; 0.2]	0.105
DrInC-2R							
PBO	324	45.8 ± 22.9	230	-19.0 ± 1.5			
NMF	328	47.2 ± 23.7	209	-20.4 ± 1.5	-1.5 ± 1.8	[-5.0; 2.1]	0.422

Baseline I values were based on FAS, OC; changes from Baseline I and differences to placebo were based on MMRM; FAS, OC values.
Cross-reference: Tables 99 to 104

a Changes from Baseline I and differences to placebo were based on ANCOVA; FAS, BOCF values.

Table 57. Adjusted Mean Change from Baseline in SF-36 Subscale and Component Scores (MMRM; FAS, OC) – Lundbeck23

SF36 Subscale Treatment Group	Baseline I		Adjusted Change from Baseline I to Week 24		Difference to PBO		
	N	Mean ± SD	N	Mean ± SE	Mean ± SE	95% CI	p-value
Physical functioning							
PBO	317	51.5 ± 7.2	224	0.1 ± 0.4			
NMF	319	51.4 ± 7.3	203	1.6 ± 0.4	1.5 ± 0.5	[0.5; 2.5]	0.004
Role-physical							
PBO	317	45.2 ± 9.7	225	2.5 ± 0.5			
NMF	329	44.9 ± 9.9	204	4.6 ± 0.5	2.1 ± 0.7	[0.8; 3.4]	0.002
Bodily pain							
PBO	323	49.0 ± 10.6	229	0.7 ± 0.6			
NMF	328	49.6 ± 10.5	207	3.5 ± 0.6	2.7 ± 0.8	[1.2; 4.3]	<0.001
General health							
PBO	321	43.4 ± 9.2	227	1.8 ± 0.6			
NMF	326	43.9 ± 9.0	204	3.8 ± 0.6	2.0 ± 0.7	[0.6; 3.4]	0.006
Vitality							
PBO	324	46.2 ± 9.7	230	2.7 ± 0.6			
NMF	328	46.6 ± 9.8	206	5.6 ± 0.7	2.9 ± 0.8	[1.4; 4.5]	<0.001
Social functioning							
PBO	315	40.6 ± 10.9	224	3.4 ± 0.6			
NMF	319	41.4 ± 11.0	204	6.3 ± 0.7	2.9 ± 0.8	[1.4; 4.5]	<0.001
Role-emotional							
PBO	320	40.6 ± 11.3	226	3.1 ± 0.7			
NMF	328	40.4 ± 12.0	206	5.4 ± 0.7	2.4 ± 0.9	[0.7; 4.0]	0.007
Mental health							
PBO	324	38.5 ± 11.0	230	2.9 ± 0.7			
NMF	325	38.6 ± 11.0	205	6.2 ± 0.7	3.3 ± 0.9	[1.6; 5.0]	<0.001
Physical Component Summary							
PBO	315	51.2 ± 7.9	218	0.3 ± 0.4			
NMF	326	51.4 ± 7.4	201	2.2 ± 0.4	1.9 ± 0.5	[0.9; 2.9]	<0.001
Mental Component Summary							
PBO	315	37.4 ± 11.6	218	3.6 ± 0.7			
NMF	326	37.5 ± 11.8	201	6.8 ± 0.8	3.2 ± 0.9	[1.4; 5.0]	<0.001

Cross-reference: Tables 117 to 136

Table 58. Adjusted Mean Change in EQ-5D Subscale Scores (MMRM; FAS, OC) – Lundbeck23

EuroQol Variable Treatment Group	Baseline I		Adjusted Change from Baseline I to Week 24		Difference to PBO		
	N	Mean ± SD	N	Mean ± SE	Mean ± SE	95% CI	p-value
Utility index							
PBO	323	0.8 ± 0.2	228	0.02 ± 0.01			
NMF	329	0.8 ± 0.2	207	0.05 ± 0.01	0.03 ± 0.02	[-0.01; 0.06]	0.100
Health state							
PBO	319	66.9 ± 17.4	221	3.51 ± 1.05			
NMF	323	68.1 ± 17.3	200	6.67 ± 1.11	3.17 ± 1.37	[0.48; 5.85]	0.021

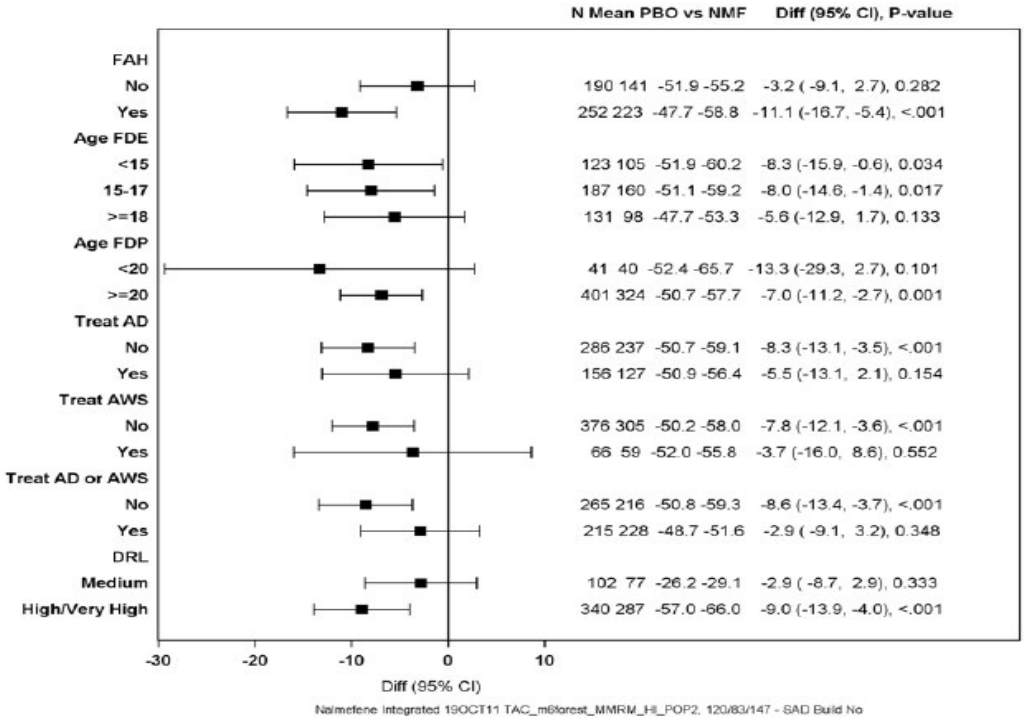
Cross-reference: Tables 137 to 140

7.1.17. Subgroup analyses

Subgroup analyses related to various definitions of high-risk alcohol dependence have already been considered below.

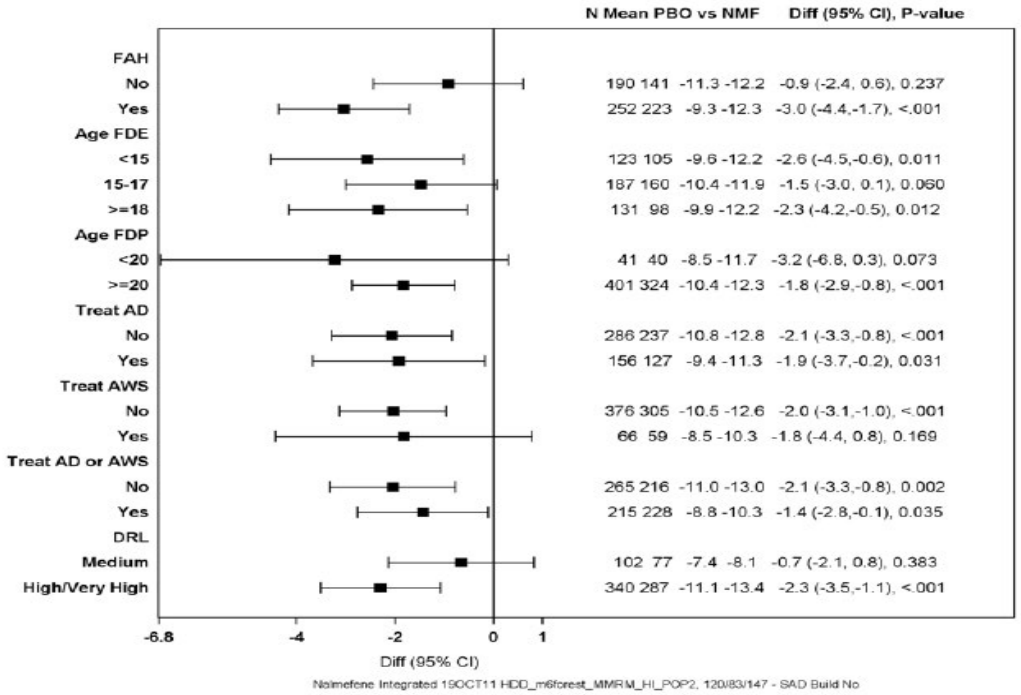
TAC results for subgroups defined by age, previous treatment, and baseline DRL are shown below for the pooled population of both pivotal studies. A similar analysis of HDDs is shown in the subsequent table. Overall, the results for both efficacy variables were consistent across different subgroups, but many of the comparisons were underpowered. Results in medium-DRL subjects appear weak.

Figure 17. Efficacy at Month 6 by Disease Variable (FAS, MMRM) – TAC (g/day) – Lundbeck14 and Lundbeck23, Pooled



FAH = family alcohol history; Age FDE = age at first drinking experience; Age FDP = age at first drinking problem; Treat AD = treatment for alcohol dependence; Treat AWS = treatment for alcohol withdrawal symptoms; DRL = drinking risk level

Figure 18. Efficacy at Month 6 by Disease Variable (FAS, MMRM) – HDDs (days/month) – Lundbeck14 and Lundbeck23, Pooled



FAH = family alcohol history; Age FDE = age at first drinking experience; Age FDP = age at first drinking problem; Treat AD = treatment for alcohol dependence; Treat AWS = treatment for alcohol withdrawal symptoms; DRL = drinking risk level

7.1.18. Results in the run-out period

Results in the Run-Out Period showed no significant benefit in continuing active treatment, compared to its randomised withdrawal. In Lundbeck14 and Lundbeck23, the number of HDDs and the TAC were both slightly lower in subjects that continued nalmefene compared to those that switched to placebo. In Lundbeck14, the between-group difference was about one HDD per month and 3.2 g alcohol per day; in Lundbeck23, the difference was 0.1 HDDs per month, and 1.6 g per day. This is of marginal clinical significance.

A failure of the groups to show a significant difference in the Run-Out Period could reflect the weakness of the therapeutic effect of nalmefene as well as a lack of statistical power in this phase of the study, but it could also reflect persistence of benefit. Given the marginal results achieved in the main treatment period, and the small between-group differenes in the ROP, the first explanation appears more likely. The fact that subjects continued to exhibit lowered alcohol intake after randomised cessation of nalmefene supports the notion that psychological factors played a large role in their initial reduction in drinking.

This analysis at least suggests that there was no rebound increase in alcohol consumption on ceasing nalmefene.

Table 59. Run-Out Period: Changes from ROP Baseline in HDDs and TAC (FAS, OC, ANCOVA) – Lundbeck14

Variable / Treatment Group	ROP Baseline		Change from ROP Baseline to Month 7		Difference to NMF-PBO		
	N	Mean ± SD	N	Mean ± SE	Mean ± SE	95% CI	p-value
Number of HDDs (days/month)							
NMF-PBO	72	6.5 ± 7.6	72	0.4 ± 0.5			
NMF-NMF	69	7.7 ± 7.2	67	-0.6 ± 0.5	-1.0 ± 0.6	[-2.2; 0.2]	0.107
TAC (g/day)							
NMF-PBO	72	29 ± 27	72	0.9 ± 1.8			
NMF-NMF	69	30 ± 24	67	-2.3 ± 1.8	-3.2 ± 2.0	[-7.2; 0.8]	0.112
ROP baseline values are based on FAS, OC.							

Table 60. Run-Out Period: Changes from ROP Baseline in HDDs and TAC (FAS, OC, ANCOVA) – Lundbeck23

Variable / Treatment Group	ROP Baseline		Change from ROP Baseline to Month 7		Difference to NMF-PBO		
	N	Mean ± SD	N	Mean ± SE	Mean ± SE	95% CI	p-value
Number of HDDs (days/month)							
NMF-PBO	98	6.2 ± 8.9	96	1.0 ± 0.7			
NMF-NMF	100	6.7 ± 8.8	97	0.9 ± 0.7	-0.1 ± 0.8	[-1.7; 1.6]	0.917
TAC (g/day)							
NMF-PBO	98	29 ± 37	96	5.3 ± 3.5			
NMF-NMF	100	31 ± 36	97	3.7 ± 3.4	-1.6 ± 4.1	[-9.7; 6.6]	0.706
ROP baseline values are based on FAS, OC.							

7.2. Major supportive efficacy studies

7.2.1. Major supportive lundbeck study (Lundbeck13, 12013A)

7.2.1.1. Study design, objectives, locations and dates

This study was a Phase III, randomised, double-blind, placebo-controlled, parallel-group study of nalmefene 20mg as-needed in the treatment of alcohol dependence. Its initial primary objective was the study of the safety and tolerability of nalmefene over 52 weeks, and efficacy was added as an additional primary objective later. The two co-primary efficacy endpoints (HDDs and TAC) were not specified until the 4th protocol amendment on 31st July, 2009, approximately four months after study commencement on 24th March 2009. These endpoints were selected to match the two pivotal efficacy studies, Lundbeck14 and Lundbeck23, which have been discussed above, and the time-point for the major efficacy assessment was set at 24 weeks, also matching those other Lundbeck studies. The total study duration was 52 weeks, with the weeks beyond the primary efficacy time point primarily satisfying the safety monitoring objective, although efficacy data continued to be collected for a total of 52 weeks.

An additional objective was added in the 2nd protocol amendment: "to study how genotype may affect treatment response to nalmefene". The results in relation to this objective were not reported in the main study summary.

The study was performed at 60 sites – 5 in Czech Republic, 5 in Estonia, 2 in Hungary, 4 in Latvia, 2 in Lithuania, 15 in Poland, 8 in Russia, 4 in Slovakia, 10 in Ukraine, and 5 in the United Kingdom. The first patient visit was on 24th March 2009, and the last patient visit was on 3rd November 2010.

7.2.1.2. Inclusion and exclusion criteria

The main inclusion criteria closely resembled those in the two pivotal studies. Patients were eligible if they were outpatients with a primary diagnosis of alcohol dependence according to DSM-IV-TR™, and also satisfied the following major criteria:

- they were ≥ 18 years of age
- they had a blood alcohol concentration (BAC) < 0.02% at the Screening Visit
- they had had ≥ 6 HDDs in the 4 weeks preceding the Screening Visit
- they did not have serum aspartate aminotransferase (ASAT) and/or serum alanine aminotransferase (ALAT) values > 3 times upper limit of the reference range, that were considered clinically significant.

Unlike the pivotal studies, the subjects were still eligible for enrolment if they had a low DRL at Baseline (the pivotal studies required at least medium risk); for the efficacy analysis, such subjects were excluded.

Minor inclusion and exclusion criteria were essentially the same as those used in the pivotal studies.

7.2.1.3. Study treatments

Subjects received nalmefene 20mg on an as-needed basis, up to once per day, or matching placebo.

7.2.1.4. Efficacy variables and outcomes

The efficacy variables were the same as described previously for the pivotal studies, as listed below, except that CGI-I response was added post hoc as an additional response measure.

Drinking measures derived from the timeline followback (TLFB):

- number of HDDs (an HDD was defined as a day with alcohol consumption ≥ 60g for men and ≥ 40g for women)
- TAC, defined as mean daily alcohol consumption in g/day over a month (28 days)
- RSDRL response
- TAC response (defined as a ≥ 30%, ≥ 50%, or ≥ 70% reduction in TAC from baseline)
- RLDRL response (defined as a downward shift in DRL to low risk or below)
- number of non-drinking days (NDDs)

Alcohol dependence symptoms and clinical status:

- Clinical Global Impression – Global Improvement (CGI-I) score
- CGI-I response, defined as a CGI-I score ≤ 2 (added as a post hoc analysis)
- Clinical Global Impression – Severity of Illness (CGI-S) score
- Drinker Inventory of Consequences (DrInC-2R) score

Liver function and other clinical safety laboratory tests:

- GGT
- ALAT
- mean corpuscular volume (MCV)
- percent carbohydrate-deficient transferrin (%CDT)

Pharmacoeconomic outcomes:

- 36-item Short-form Health Survey (SF-36) subscale scores
- EuroQol (EQ-5D) utility index and visual analogue scale (VAS) scores
- Resource Use Measurement Questionnaire – Alcohol Dependence (RUMQ-ADP)
- Brief Measure of Readiness to Change Questionnaire (BMRCQ) subscale scores: importance, confidence, and readiness

These assessments were performed as described for the pivotal studies.

The co-primary efficacy endpoints were change from baseline to Month 6 in HDDs and TAC, analysed by MMRM.

The key secondary endpoint was the response rate (RSRDL), analysed with logistic regression (LREG).

Additional secondary endpoints included:

- responder rates based on a ≥ 30%, ≥ 50%, or ≥ 70% reduction from baseline in monthly TAC
- changes from baseline in monthly NDDs, DrInC-2R scores, CGI-S scores, %CDT, SF-36 subscale scores, EQ-5D utility index and VAS scores, and BMRCQ scores
- log-transformed GGT, ALAT, and MCV values
- CGI-I score
- response-rate based on a CGI-I score ≤ 2.

In general, these endpoints resembled those in the pivotal studies. The primary efficacy analysis based on the two co-primary endpoints of HDDs and TAC, as well as the secondary efficacy analysis based on the RSDRL, were identical to those used in Lundbeck23 and very similar to those used in Lundbeck14. (The two pivotal studies had differed in their imputation methods for the RSDRL analysis, and this study employed the method used in Lundbeck23, which yielded positive results, rather than the one used in Lundbeck14, which yielded negative results).

7.2.1.5. Randomisation and blinding methods

Subjects were randomised unequally to active treatment or placebo in a 3:1 ratio, using a centralised randomisation program generated by Lundbeck. The program used block randomisation in blocks of 4 to balance patient assignments at each site.

Blinding was approached by using placebo tablets that appeared identical to the nalmefene tablets, keeping the randomisation codes hidden at a central location.

As in the pivotal studies, it is possible that some subjects became unblinded because of telltale side effects.

7.2.1.6. Analysis populations

As in the pivotal studies, the Sponsor described three main populations for analysis:

- all-patients-randomised set (APRS) – all randomised patients
- all-patients-treated set (APTS) –patients in the APRS excluding those with no study-drug intake
- full-analysis set (FAS) – all patients in the APTS who had at least one valid post-baseline assessment of both co-primary efficacy variables and were at medium risk or above according to WHO criteria (> 40g/day for men, > 20g/day for women) at Baseline

Unlike the pivotal studies, this study allowed recruitment of subjects with low DRL at baseline, and 16% of subjects were in this category. These subjects were not part of the FAS, according to the definitions above. This is appropriate, because such subjects would not be expected to show a substantial response to treatment and do not represent the target population for nalmefene. They were primarily recruited to facilitate collection of safety data.

7.2.1.7. Sample size

Sample size estimations were based on anticipated results in the MMRM model, assuming a standard deviation for the change from baseline in number of HDDs of 7 days and the change from baseline in TAC of 36.5g/day. A target of 668 patients was planned for enrolment, based on a 20% withdrawal rate by Month 6 and assuming a correlation of 0.7 between the co-primary efficacy variables. Using a standard significance level of 5% ($p < 0.05$), it was estimated that a total of 668 patients, randomised in a 3:1 ratio, would provide power of ≥ 90% for detecting a treatment-related reduction of 3 HDDs/month and a TAC reduction of 12 g/day, relative to placebo.

7.2.1.8. Statistical methods

Statistical methods in this study closely resembled those already described for the pivotal studies. All efficacy analyses were based on the FAS and were tested at the 5% level of significance. The principal statistical software used was the same as in the pivotal studies, SAS, Version 9.2.

The co-primary efficacy analyses were based on the changes from baseline in monthly number of HDDs and monthly TAC, analysed by MMRM, using observed cases (OC), and with the baseline score as a covariate, and site, sex, time in months (Months 1 to 13), and treatment as fixed effects. The baseline score-by-time interaction and the treatment-by-time interaction were also included in the model.

The key secondary efficacy analysis was based on the RSDRL at Month 6, using a logistic regression (LREG) model, with country, sex, baseline DRL, and treatment as fixed effects. Missing values were imputed using individual-patient predicted values of TAC derived from the MMRM model, as had been done for Lundbeck23, rather than a No-Response imputation method, as had been done for Lundbeck14.

Secondary efficacy analyses included:

- an assessment of responder rates based on a ≥ 30%, ≥ 50%, or ≥ 70% reduction from baseline in monthly TAC, analysed using an LREG model as for the RSDRL analysis.
- changes from baseline in monthly number of NDDs, DrInC-2R scores, CGI-S scores, %CDT, SF-36 subscale scores, EQ-5D utility index and VAS scores, and BMRCQ scores, analysed using an MMRM model similar to the one used for the co-primary efficacy analyses.

These statistical approaches were reasonable overall, but – as revealed by a number of sensitivity analyses – some of the results were quite sensitive to the statistical method chosen.

No correction was made for using multiple endpoints; instead, the null hypothesis (of no significant treatment effect) had to be rejected for both co-primary endpoints at the 5% level in order to consider the drug efficacious. Also, formal testing of the key secondary endpoint was to be performed only if both co-primary endpoints were significantly positive. In the individual study report, the Sponsor was explicit on this issue:

> *The null hypothesis was to be rejected for both co-primary endpoints at the 5% significance level to consider nalmefene 20mg as-needed use to be efficacious."*
>
> *The null hypothesis was to be rejected for both co-primary endpoints at the 5% level of significance in order to proceed with formal testing of the key secondary endpoint.*

Despite making these pronouncements, the Sponsor appeared to ignore them in their presentation of the results. As it turned out, the null hypothesis was not rejected for either co-primary endpoint – that is, no significant treatment effect was shown for either primary endpoint – so no formal testing of the key secondary endpoint should have been performed. Nonetheless, 95%CIs and p-values were cited for the key secondary endpoint and for all subsequent endpoints in the hierarchy.

7.2.1.9. Participant flow

Patient disposition is summarised in the figure and table below. Withdrawals were relatively common (36.5%), but the withdrawal rate was acceptable for a study of this nature; compliance with the study procedures would be expected to be impaired by the condition under study. Withdrawals were more common in the active group (placebo 31.7% vs nalmefene 38.1%), with adverse events accounting for most of the excess (withdrawals due to AEs: placebo 2/166, 1.2% vs nalmefene 43/508, 8.5%). The other major reason for withdrawal, "Withdrawal of consent" occurred with a similar incidence in each treatment group (placebo 35/166, 21% vs nalmefene 94/509, 18.5%).

As in the pivotal studies, this unequal withdrawal rate poses problems of interpretation, and renders the results susceptible to the effects of different imputation methods.

Figure 19. Patient Disposition – Lundbeck13

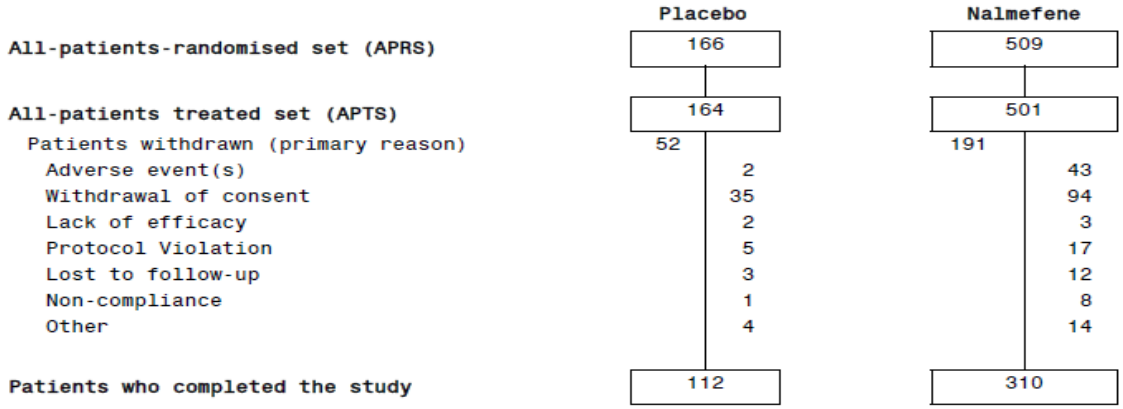


Table 61. Patient Disposition – Lundbeck13

	PBO		NMF		TOTAL	
	n	%	n	%	n	%
Patients Randomised	166		509		675	
Patients Treated	164		501		665	
Patients Completed	112	(68.3)	310	(61.9)	422	(63.5)
Patients Withdrawn	52	(31.7)	191	(38.1)	243	(36.5)
Efficacy Data Sets						
Full Analysis Set	137		415		552	

12013A FINAL Report ST_DI01 07SEP2011:13:07:05 1001/147 - TGML/SAD Build Numbers

7.2.1.10. Major protocol violations/deviations

Protocol deviations were not summarised in a convenient format in the Sponsor's original submission, but instead individual violations were included in several tables, some of which covered multiple pages. Deviations included the recruitment of ineligible subjects, use of disallowed concomitant medication, visits outside designated windows, and a number of other deviations in data collection.

The Sponsor provided the table below as part of their Section 31 Response.

Table 62. Protocol deviations in Lundbeck13

Deviation Category	Incidence (%)
Deviation in informed consent procedure	8.9
Violation of inclusion criterion (other than informed consent procedure)	0
Violation of exclusion criterion	7.0
Use of disallowed concomitant medication	15
Procedural compliance deviations: Visit 2 (Randomisation Visit) outside of visit window (that is, ≥21 days after Visit 1)	1.6
Procedural compliance deviations: Visits 3 to 12 outside of visit window (that is, ≥7 days for Visits 3 and 4, or ≥15 days for Visits 5 to 12)	23
Procedural compliance deviations: Other deviations	42

Overall, the number of protocol deviations was acceptable for a study of this nature.

7.2.1.11. Baseline data

Baseline characteristics including demographics and disease characteristics are summarised in the tables below. There were no important demographic differences. The two treatment groups were also reasonably well-matched in terms of baseline level of alcohol use, as reflected in mean number of HDDs, TAC, and DRL. Mean alcohol intake at baseline was ~75g/day (placebo 74.6 g/day, nalmefene 75.2 g/day). On average, clinicians rated the two groups as having a similar level of disease severity on the CGI-S scale (mean CGI-S: placebo 3.92, nalmefene 4.00).

Table 63. Patient Demographics – Lundbeck13

		PBO		NMF		TOTAL	
Number of Patients		166		509		675	
Age (years)	N	166		509		675	
	MEAN	44.27		44.26		44.26	
	STD	11.99		11.24		11.42	
	MIN	18.00		19.00		18.00	
	MAX	72.00		77.00		77.00	
	MEDIAN	44.00		44.00		44.00	
Age group (years) n (%)	<25	8	(4.8)	14	(2.8)	22	(3.3)
	>=25 and <35	30	(18.1)	91	(17.9)	121	(17.9)
	>=35 and <45	47	(28.3)	160	(31.4)	207	(30.7)
	>=45 and <55	44	(26.5)	153	(30.1)	197	(29.2)
	>=55 and <65	30	(18.1)	64	(12.6)	94	(13.9)
	>=65	7	(4.2)	27	(5.3)	34	(5.0)
Sex n (%)	F	39	(23.5)	116	(22.8)	155	(23.0)
	M	127	(76.5)	393	(77.2)	520	(77.0)
Race n (%)	ASIAN			1	(0.2)	1	(0.1)
	BLACK			1	(0.2)	1	(0.1)
	CAUCASIAN	165	(99.4)	506	(99.4)	671	(99.4)
	OTHER	1	(0.6)	1	(0.2)	2	(0.3)

12013A FINAL Report Results ST_DE01 19SEP2011:09:47:41 1001/147 - TGML/SAD Build Numbers

Table 64. Baseline Weight, Height, BMI – Lundbeck13

		PBO	NMF	TOTAL
Number of Patients		166	509	675
Weight (kg)	N	165	506	671
	MEAN	79.38	80.07	79.90
	STD	16.41	15.41	15.65
	MIN	46.00	45.00	45.00
	MAX	160.00	138.70	160.00
	MEDIAN	77.30	78.50	78.00
Height (cm)	N	165	504	669
	MEAN	174.86	175.06	175.01
	STD	8.39	8.44	8.42
	MIN	158.00	150.00	150.00
	MAX	202.00	200.00	202.00
	MEDIAN	176.00	175.50	176.00
BMI (kg/m2)	N	165	504	669
	MEAN	25.84	26.04	25.99
	STD	4.34	4.18	4.22
	MIN	16.30	15.60	15.60
	MAX	43.40	43.80	43.80
	MEDIAN	25.40	25.55	25.50
BMI Categories n (%)	<18.5 kg/m2 (underweight)	3 (1.8)	6 (1.2)	9 (1.3)
	>=18.5 and <25 kg/m2 (normal)	76 (45.8)	217 (42.6)	293 (43.4)
	>=25 and <30 kg/m2 (overweight)	63 (38.0)	192 (37.7)	255 (37.8)
	>=30 and <40 kg/m2 (obese)	22 (13.3)	88 (17.3)	110 (16.3)
	>=40 kg/m2 (severely obese)	1 (0.6)	1 (0.2)	2 (0.3)
	Unknown	1 (0.6)	5 (1.0)	6 (0.9)

12013A FINAL Report Results ST_DE03 19SEP2011:09:48:07 1001/147 - TGML/SAD Build Numbers

Table 65. Socio-Demographics – Lundbeck13

		PBO	NMF	TOTAL
Number of Patients		166	509	675
Living Arrangement n (%)	Alone	26 (15.7)	73 (14.3)	99 (14.7)
	Single with children	9 (5.4)	19 (3.7)	28 (4.1)
	With spouse or partner	58 (34.9)	156 (30.6)	214 (31.7)
	With spouse or partner and children	55 (33.1)	209 (41.1)	264 (39.1)
	Other	18 (10.8)	52 (10.2)	70 (10.4)
Education Level n (%)	Not known	3 (1.8)	13 (2.6)	16 (2.4)
	Other general/practical education	21 (12.7)	56 (11.0)	77 (11.4)
	Primary education or less/Primary S	13 (7.8)	50 (9.8)	63 (9.3)
	Secondary education/High School	78 (47.0)	225 (44.2)	303 (44.9)
	Tertiary/further education/Universi	51 (30.7)	165 (32.4)	216 (32.0)

12013A FINAL Report Results ST_DE02 19SEP2011:09:48:00 1001/147 - TGML/SAD Build Numbers

Table 66. Baseline Efficacy Summary (FAS) – Lundbeck13

		PBO		NMF		TOTAL	
Number of Patients		137		415		552	
Drinking Risk n (%)	Medium	49	(35.8)	163	(39.3)	212	(38.4)
	High	57	(41.6)	143	(34.5)	200	(36.2)
	Very high	31	(22.6)	109	(26.3)	140	(25.4)
Number of HDDs per Month (28 days)	N	137		415		552	
	MEAN	14.67		15.17		15.05	
	STD	6.07		6.12		6.11	
	MIN	6.00		6.00		6.00	
	MAX	28.00		28.00		28.00	
	MEDIAN	13.00		13.00		13.00	
Total Alcohol Consumption (g/day)	N	137		415		552	
	MEAN	74.60		75.20		75.05	
	STD	40.97		39.41		39.77	
	MIN	23.00		21.00		21.00	
	MAX	283.00		447.00		447.00	
	MEDIAN	66.00		65.00		65.00	
CGI-S	N	135		409		544	
	MEAN	3.92		4.00		3.98	
	STD	1.06		1.12		1.10	
	MIN	1.00		1.00		1.00	
	MAX	6.00		6.00		6.00	
	MEDIAN	4.00		4.00		4.00	
GGT (IU/L)	N	137		414		551	
	MEAN	79.29		71.03		73.08	
	STD	125.80		114.52		117.36	
	MIN	5.00		4.00		4.00	
	MAX	928.00		1415.0		1415.0	
	MEDIAN	40.00		37.00		38.00	
ALAT (IU/L)	N	137		415		552	
	MEAN	32.04		34.06		33.56	
	STD	19.93		22.64		22.00	
	MIN	5.00		4.00		4.00	
	MAX	100.00		144.00		144.00	
	MEDIAN	26.00		28.00		27.00	
MCV (fL)	N	134		412		546	
	MEAN	96.25		96.63		96.54	
	STD	5.17		5.68		5.56	
	MIN	81.00		73.00		73.00	
	MAX	115.00		118.00		118.00	
	MEDIAN	96.00		96.00		96.00	
%CDT (%)	N	134		413		547	
	MEAN	2.50		2.57		2.55	
	STD	1.40		1.55		1.52	
	MIN	1.05		0.88		0.88	
	MAX	8.91		12.90		12.90	
	MEDIAN	2.02		2.03		2.03	
DrInC-2R Total	N	137		415		552	

12013A FINAL Report Results ET_BL01_FAS 19SEP2011:09:48:53 1001/147 - TGML/SAD Build Numbers

7.2.1.12. Results for the co-primary efficacy endpoints

This study was negative for both of its co-primary endpoints. The number of HDDs at Month 6 was reduced, relative to Baseline, in both treatment groups, but the reduction in the two groups was similar, 8.9 HDDs/month in the placebo group, compared to 9.8 HDDs/month in the nalmefene group, a mean difference of -0.9 days (with a considerable degree of uncertainty on either side of this estimate; 95%CI -2.1 to +0.4 days).

The average reduction in alcohol intake per day was substantial in both groups, indicating benefit from the non-pharmacological aspects of being involved in the study. The placebo group reduced their TAC by 45.6 g/day from a mean baseline of 75 g/day; the nalmefene group reduced their TAC by 49.0 g/day from the same mean baseline TAC. The mean difference between the two groups was minor (-3.5 g/day), with a 95%CI that included no difference (95%CI -9.2 to +2.2).

These results represent a rather weak trend in favour of active treatment, which would be of doubtful clinical significance even if the differences were confirmed in a higher-powered study. If withdrawal bias, un-blinding from telltale side effects or any other methodological issue has inflated these observed differences, then even this weak trend could be spurious.

Table 67. Results for Co-Primary Efficacy Analysis (FAS, OC, MMRM) – Lundbeck13

Efficacy Variable / Treatment Group	Baseline		Adjusted Change from Baseline to Month 6		Difference to PBO		
	N	Mean ± SD	N	Mean ± SE	Mean ± SE	95% CI	p-value
Number of HDDs							
PBO	137	14.7 ± 6.1	110	-8.9 ± 0.6			
NMF	415	15.2 ± 6.1	320	-9.8 ± 0.4	-0.9 ± 0.6	[-2.1; 0.4]	0.160
TAC							
PBO	137	75 ± 41	110	-45.6 ± 2.6			
NMF	415	75 ± 39	320	-49.0 ± 1.6	-3.5 ± 2.9	[-9.2; 2.2]	0.232

Baseline values were based on FAS, OC; changes from baseline and differences to placebo were based on MMRM; FAS, OC values.

The Sponsor's Summary of Clinical Efficacy (SCE) reported results for this study that were technically correct, but in a way that partially disguised the fact that this was a negative study. Firstly, the Sponsor introduced the Overview of this study (Section 2.1.3.1 of the SCE) with this comment:

> *The overall objectives of the study were to evaluate the long-term safety and tolerability of as-needed dosing of 20mg nalmefene in patients with alcohol dependence during a 52-week treatment period. A secondary objective was to explore the efficacy of nalmefene versus placebo during the 52-week treatment period; a protocol amendment specified efficacy comparisons at Month 6.*

This comment could easily be read as indicating that the highest ranking efficacy objective was assessment of efficacy over 52 weeks, and that the efficacy comparisons at Month 6 were a minor afterthought. In fact, the study synopsis indicates that the 24-week assessment was primary for efficacy:

> *Primary objectives:*
>
> *– safety:*
>
> *• to evaluate the long-term safety and tolerability of as-needed use of 20mg nalmefene versus placebo over a period of 52 weeks in patients with alcohol dependence*
>
> *– efficacy (added with SA04):*
>
> *• to evaluate the effect of as-needed use of 20mg nalmefene on alcohol consumption using the monthly number of heavy drinking days (HDDs) and the monthly total alcohol consumption (TAC) in patients with alcohol dependence* ***during a treatment period of 24 weeks*** *(co-primary efficacy endpoints)*
>
> *[Study 1213A Study Report Body, p3/1234, emphasis added].*

Secondly, in presenting the actual results for the co-primary endpoints, the Sponsor's SCE fails to distinguish between the primary 24-week time point and the later 52-week time point, giving the results at each time point equal emphasis:

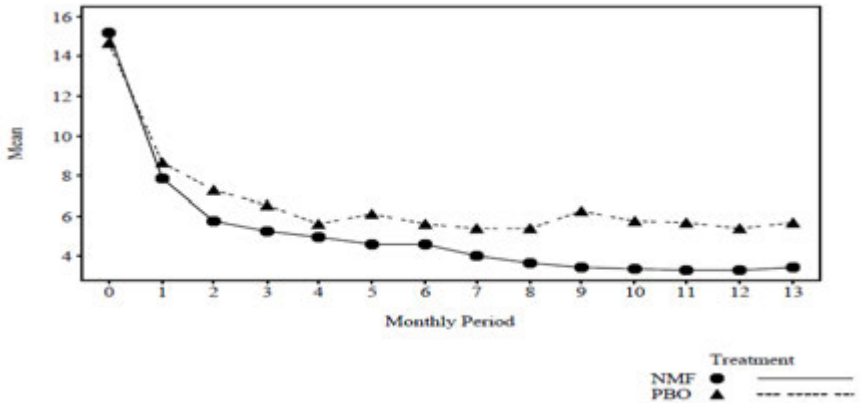
> *Nalmefene was numerically better than placebo in reducing the number of HDDs and TAC at Month 6. The effect of nalmefene was evident already at Month 1 and maintained throughout the treatment period. The difference to placebo was in favour ($p < 0.05$) of nalmefene at the majority of the timepoints, including Month 13 (Panels 42, 43, and 44)*

In this part of the SCE, which is the Sponsor's first paragraph under the heading "Co-primary Efficacy Analyses', the p-value for the actual, pre-specified co-primary efficacy analysis is not even mentioned; instead, the p-value for other time points is given prominence. This is misleading. Nowhere in the SCE is it mentioned that this study was negative according to its

pre-specified Statistical Analysis Plan. The results at Month 13 represent a secondary endpoint, and as such are discussed separately in this report.

Even the individual study report has been written in a way that de-emphasizes the fact that this study was negative, with the HDD results produced without error bars, as shown below. (The SCE figures used error bars, but highlighted time points in which the treatment groups were significantly different).

Figure 20. Monthly HDDs (FAS, OC) – Lundbeck13



In the previously describe pivotal studies (Lundbeck14 and Lundbeck23), a secondary analysis was performed using ANCOVA with LOCF, and this approach achieved significance where the MMRM had not. For this supportive study, however, not even the LOCF approach was significantly in favour of active treatment, as shown below.

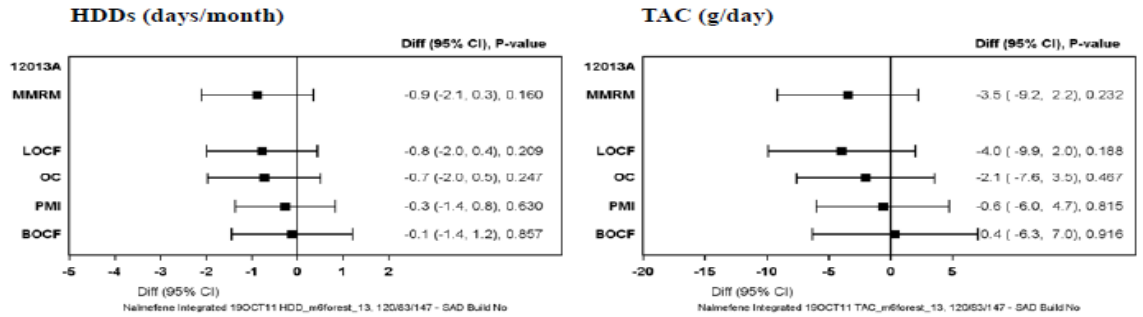
Table 68. Results for the Co-Primary Efficacy Variables at Month 6 (FAS) – Lundbeck13

Variable Treatment Group	Baseline		Change from Baseline to Month 6		Difference to PBO		
	N	Mean ± SD	N	Mean ± SE	Mean ± SE	95% CI	p-value
Number of HDDs (days/month)							
MMRM							
PBO	137	14.7 ± 6.1	110	-8.9 ± 0.6			
NMF	415	15.2 ± 6.1	320	-9.8 ± 0.4	-0.9 ± 0.6	[-2.1; 0.4]	0.160
LOCF							
PBO	137	14.7 ± 6.1	137	-9.0 ± 0.6			
NMF	415	15.2 ± 6.1	415	-9.7 ± 0.4	-0.8 ± 0.6	[-2.0; 0.4]	0.209
TAC (g/day)							
MMRM							
PBO	137	75 ± 41	110	-45.6 ± 2.6			
NMF	415	75 ± 39	320	-49.0 ± 1.6	-3.5 ± 2.9	[-9.2; 2.2]	0.232
LOCF							
PBO	137	75 ± 41	137	-45.5 ± 2.8			
NMF	415	75 ± 39	415	-49.5 ± 1.8	-4.0 ± 3.2	[-9.9; 2.0]	0.188

Baseline values are based on FAS, OC.

The Sponsor performed a number of sensitivity analyses to explore these results further. Usually, the main point of a sensitivity analysis is to show that a positive result is robust enough that the outcome does not depend critically on the statistical methodology or imputation methods. In this case, however, the sensitivity analysis merely confirmed that multiple different imputation methods consistently produced a negative result at the main efficacy time point, Month 6.

Figure 21. Sensitivity Analyses – Changes from Baseline in HDDs and TAC at Month 6 – Lundbeck13



MMRM is the co-primary efficacy analysis and is included for comparison.
Negative values indicate a greater reduction in the number of HDDs and TAC in the nalmefene group than in the placebo group.

7.2.1.13. Results in low-risk and high-risk drinkers

The protocol prospectively specified a subgroup analysis of higher-risk drinkers, identified as those with high or very high DRL at Baseline. (Note that this is different to the high-risk subgroup identified post hoc and featuring prominently in the SCE, which was based on subjects with high or very high DRL at Baseline and Randomisation; this post hoc high-risk group is discussed separately below.)

Efficacy results in this prospectively identified subgroup at the main analysis time point are shown below. No significant treatment effect was identified.

Table 69. Summary of Results for Patients with a High or Very High DRL at Baseline for the Co-Primary Efficacy Analyses at Month 6 (MMRM; FAS, OC)

Efficacy Variable Treatment Group	Baseline		Adjusted Change from Baseline to Month 6		Difference to PBO		
	N	Mean ± SD	N	Mean ± SE	Mean ± SE	95% CI	p-value
Number of HDDs							
PBO	88	16.7 ± 6.4	68	-10.5 ± 0.9			
NMF	252	18.0 ± 6.0	188	-11.6 ± 0.6	-1.1 ± 1.0	[-3.0; 0.8]	0.253
TAC							
PBO	88	91 ± 42	68	-57.7 ± 4.1			
NMF	252	94 ± 40	188	-63.3 ± 2.9	-5.6 ± 4.6	[-14.6; 3.4]	0.219

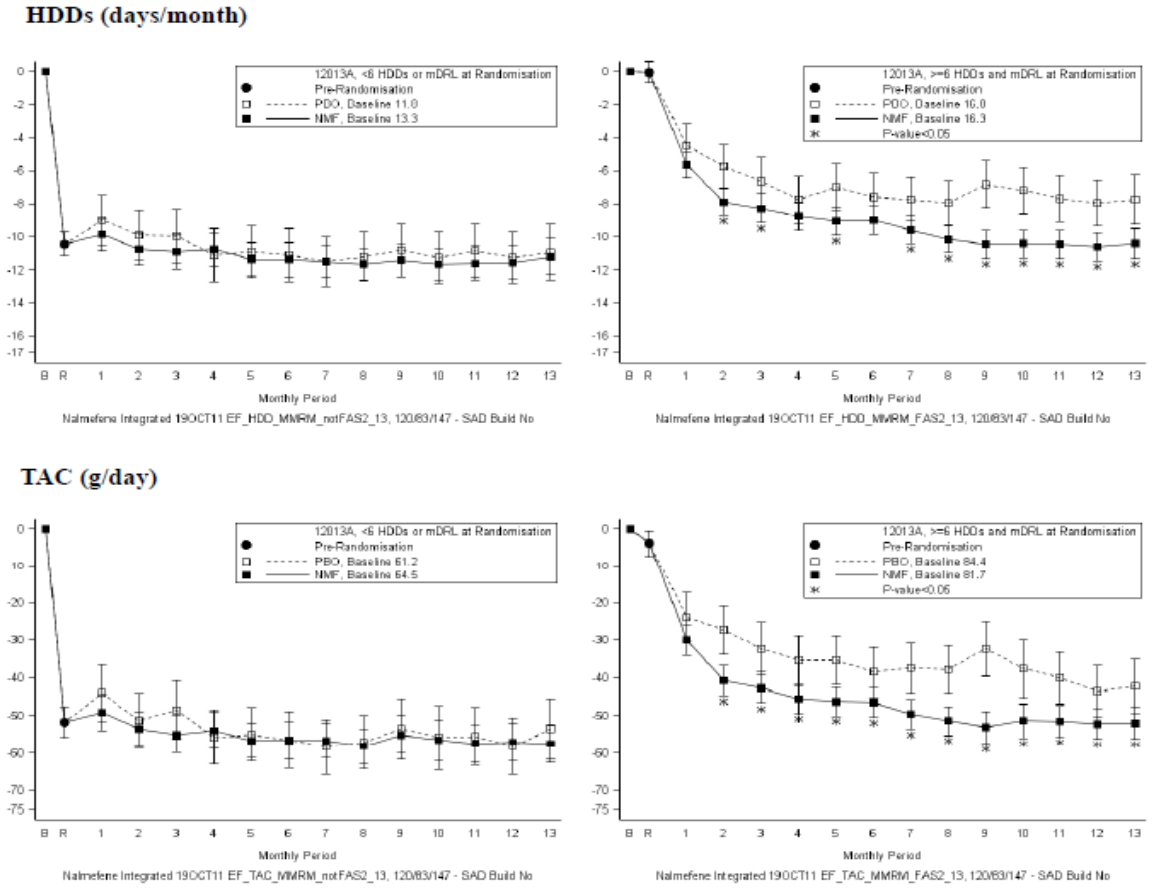
Baseline values were based on FAS, OC; changes from baseline and differences to placebo were based on MMRM; FAS, OC values.

As in the pivotal study, the Sponsor realised that several subjects (58/137 placebo recipients [42%] and 157/415 nalmefene recipients [38%]) reduced their alcohol intake substantially prior to randomisation, such that their drinking levels were no longer consistent with medium DRL. In these subjects, treatment had little effect because alcohol consumption was already low prior to treatment. The Sponsor therefore performed a post hoc subgroup analysis based on consumption levels at randomisation, which might be expected to identify a group in whom a positive treatment effect was observed.

Results in this LDAR subgroup (Low DRL At Randomisation) are shown in the left graphs below, with the results in the MDAR subgroup (at-least Medium DRL At Randomisation) on the right. As in the pivotal studies, LDAR subjects continued to have low intake throughout the study, regardless of assigned treatment, and there was no apparent treatment effect.

For the MDAR subjects, a statistically significant treatment effect was observed at multiple time points; this included Month 6 for TAC, but not for HDDs.

Figure 22. Changes from Baseline in HDDs and TAC – Patients Categorised According to Alcohol Consumption at Randomisation (FAS, MMRM) – Lundbeck13



This data is somewhat reassuring, because it suggests that a significant treatment effect can be achieved with nalmefene when the drug is correctly targeted, but it is important to note that the subgroup was identified post hoc, and even then it did not achieve significance for both co-primary efficacy variables at the designated time point.

Together with similar observations in the pivotal studies, this post hoc subgroup analysis suggest that the failure of Lundbeck13 to achieve a positive result for either of its co-primary efficacy analyses partly reflects dilution of the study population with subjects who had already responded to non-pharmacological factors prior to randomisation. This suggests that a better study design might have been to randomise only the subjects who were still drinking at qualifying levels at Randomisation.

The magnitude of the benefit in this subgroup, as shown below, was modest, amounting to 2.7 HDDs per month (p = 0.002) and 9.8 g/day for TAC (p = 0.0163). This could be regarded as clinically worthwhile if the same results could be obtained prospectively.

Table 70. Adjusted Change from Baseline in HDDs at Month 13, Patients Categorised According to Alcohol Consumption at Randomisation (FAS) – Lundbeck13

>=6 HDDs and mDRL#		12013A PBO	NMF
No	Baseline		
	N	58	157
	Mean	11.79	13.28
	SD	4.65	5.30
Yes	Baseline		
	N	79	258
	Mean	16.78	16.32
	SD	6.14	6.31
MMRM, No	Adjusted Change		
	N	43	110
	Mean	-10.94	-11.20
	95% CI	(-12.63;-9.24)	(-12.29;-10.12)
	Difference to PBO		
	Mean		-0.26
	95% CI		(-2.19;1.66)
	P-value		0.7879
MMRM, Yes	Adjusted Change		
	N	54	148
	Mean	-7.73	-10.41
	95% CI	(-9.25;-6.21)	(-11.33;-9.49)
	Difference to PBO		
	Mean		-2.68
	95% CI		(-4.37;-0.99)
	P-value		0.0020
LOCF, No	Adjusted Change		
	N	58	157
	Mean	-10.14	-11.20
	95% CI	(-11.94;-8.35)	(-12.36;-10.04)
	Difference to PBO		
	Mean		-1.06
	95% CI		(-3.06;0.95)
	P-value		0.3012
LOCF, Yes	Adjusted Change		
	N	79	258
	Mean	-7.39	-9.28
	95% CI	(-8.97;-5.81)	(-10.25;-8.31)
	Difference to PBO		
	Mean		-1.90
	95% CI		(-3.55;-0.24)
	P-value		0.0250

Alcohol consumption at randomisation is based on the TLFB data collected between the Screening Visit and Randomisation, extrapolated to 4 weeks.
Patients were categorised (no,yes) as having at least 6 HDDs and at least medium DRL based on their alcohol consumption at Randomisation

Table 71. Adjusted Change from Baseline in TAC (g/day) at Month 13, Patients Categorised According to Alcohol Consumption at Randomisation (FAS) – Lundbeck13

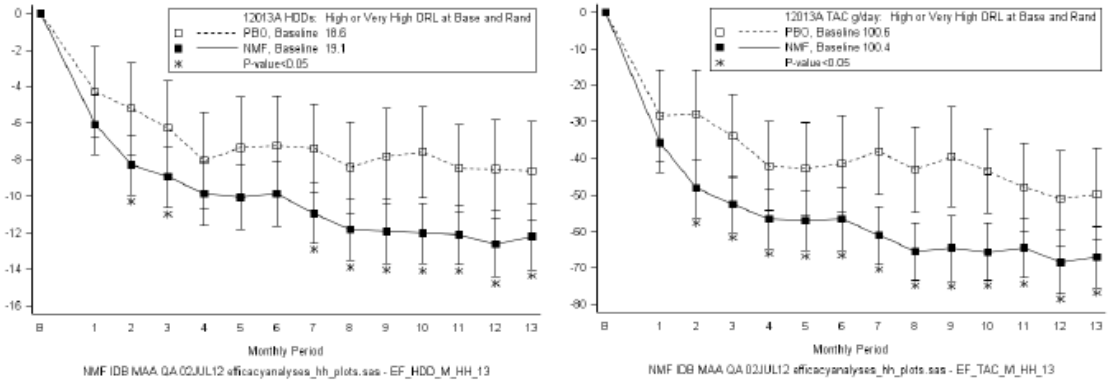
>=6 HDDs and mDRL#		12013A PBO	NMF
No	Baseline		
	N	58	157
	Mean	61.2	64.5
	SD	30.6	31.4
Yes	Baseline		
	N	79	258
	Mean	84.4	81.7
	SD	44.8	42.3
MMRM, No	Adjusted Change		
	N	43	110
	Mean	-53.4	-57.5
	95% CI	(-61.4;-45.4)	(-62.6;-52.3)
	Difference to PBO		
	Mean		-4.1
	95% CI		(-13.2;5.0)
	P-value		0.3741
MMRM, Yes	Adjusted Change		
	N	54	148
	Mean	-42.1	-51.8
	95% CI	(-49.2;-35.0)	(-56.2;-47.5)
	Difference to PBO		
	Mean		-9.8
	95% CI		(-17.7;-1.8)
	P-value		0.0163
LOCF, No	Adjusted Change		
	N	58	157
	Mean	-52.4	-56.9
	95% CI	(-61.5;-43.2)	(-62.9;-50.9)
	Difference to PBO		
	Mean		-4.5
	95% CI		(-14.7;5.7)
	P-value		0.3851
LOCF, Yes	Adjusted Change		
	N	79	258
	Mean	-37.1	-47.9
	95% CI	(-45.1;-29.1)	(-52.8;-42.9)
	Difference to PBO		
	Mean		-10.8
	95% CI		(-19.2;-2.3)
	P-value		0.0125

Alcohol consumption at randomisation is based on the TLFB data collected between the Screening Visit and Randomisation, extrapolated to 4 weeks.
Patients were categorised (no,yes) as having at least 6 HDDs and at least medium DRL based on their alcohol consumption at Randomisation

Similar results were obtained in the high-risk subgroup of subjects with high or very high DRL at Baseline and Randomisation (HDAR), which represent a more severe subset of drinkers within the MDAR subgroup. According to the proposed PI, this subgroup represented just 27% of the original study population. In this high-risk subgroup, a significant treatment effect was demonstrated at most time points for TAC, including Month 6. For HDDs, the Month 6 results were not significant, but some earlier time points did show a significant treatment effect; a sustained significant effect did not appear until Month 7 (see the table and figure below).

Results towards the end of the treatment period are potentially more susceptible to withdrawal bias than earlier results, so this evidence is only partially supportive. As noted by the SPonsor in the proposed PI, "In this post hoc target population, more patients receiving nalmefene withdrew (45%) as compared to those receiving placebo (31%)." Furthermore, the fact that this subgroup was identified post hoc means that the p-values cited for the betwene-group differences are invalid. As discussed earlier, the SAP specified a different high-risk subgroup (high or very high DRL at Baseline), which showed no significant treatment effect for either of the co-primary efficacy variables at Month 6.

Figure 23. Change from Baseline in HDDs and TAC – Patients with a High or Very High DRL at Baseline and Randomisation (FAS, MMRM) – Lundbeck13



7.2.1.14. Responder analysis

According to the Sponsor's SAP, no formal analysis of secondary endpoints should have been performed, because the null hypothesis of no treatment effect was not rejected in the co-primary efficacy analysis. Despite this, the Sponsor performed formal statistical analysis of all secondary endpoints.

The key secondary efficacy analysis was based on the RSDRL at Month 6, and results are shown below. No significant treatment effect was identified, although there were favourable numerical trends at various time points. For the pre-specified key secondary efficacy analysis at Month 6, the response rate was 76% in the placebo group and 78% in the nalmefene group, a trivial difference that did not reach statistical significance (p = 0.816).

Despite the Sponsor's various attempts to impute sensible values for missing data, it remains unclear to what extent the fall in alcohol consumption and improvement in response rate throughout the study, and particularly in the last few months, reflects progressive withdrawal of less motivated patients and progressive enrichment of the remaining study cohort with more motivated patients.

By Month 13, the response rate in the placebo group was 75%, and in the nalmefene group it was 83%. This difference was nearly significant (p = 0.053), but these late results could be susceptible to withdrawal bias. The magnitude of the difference in response rate at the end of the study (~8% higher with nalmefene) is broadly similar to the 6.4% difference in withdrawal rate (placebo 31.7% vs nalmefene 38.1%).

Note that ¾ of placebo recipients showed a satisfactory response to treatment over the course of a year. This suggests that most subjects identified as having alcohol dependence can show a sustained response to non-pharmacological measures.

Table 72. Adjusted OR Response, Shift in DRL (RSDRL, FAS, MMRM-Imputation, LREG)

Treatment Group	Month	N	Responders n	Responders (%)	Adjusted Odds Ratio with 95% CI Odds Ratio	Lower	Upper	LR Test p-value
PBO	1	137	72	(52.6)				
	2	137	88	(64.2)				
	3	137	93	(67.9)				
	4	137	99	(72.3)				
	5	137	99	(72.3)				
	6	137	104	(75.9)				
	7	137	103	(75.2)				
	8	137	105	(76.6)				
	9	137	101	(73.7)				
	10	137	104	(75.9)				
	11	137	103	(75.2)				
	12	137	104	(75.9)				
	13	137	103	(75.2)				
NMF	1	415	253	(61.0)	1.44	0.95	2.19	0.0887
	2	415	296	(71.3)	1.39	0.89	2.17	0.1504
	3	415	307	(74.0)	1.36	0.85	2.15	0.2021
	4	415	313	(75.4)	1.18	0.73	1.87	0.4995
	5	415	319	(76.9)	1.24	0.77	1.98	0.3760
	6	415	324	(78.1)	1.06	0.64	1.74	0.8162
	7	415	334	(80.5)	1.36	0.82	2.25	0.2335
	8	415	342	(82.4)	1.42	0.84	2.36	0.1915
	9	415	339	(81.7)	1.63	0.99	2.65	0.0548
	10	415	337	(81.2)	1.35	0.80	2.24	0.2633
	11	415	344	(82.9)	1.65	0.98	2.75	0.0583
	12	415	345	(83.1)	1.58	0.94	2.64	0.0843
	13	415	345	(83.1)	1.67	0.99	2.78	0.0534

Shift is defined as Baseline Very High Risk DRL to Medium or Below or from Baseline High or Medium Risk DRL to Low or Below
Missing values imputed by response based on individual patient-predicted TAC from the MMRM-model in primary analysis

12013A FINAL Report Results ET_DRL48 06SEP2011:14:28:19 1001/147 - TGML/SAD Build Numbers

For other definitions of response, the results were generally disappointing, with only the ≥ 30% TAC reduction showing a significant treatment effect (p = 0.048). Note that this marginal p-value has not been adjusted for multiple comparisons, and is not formally valid given that secondary endpoints were not supposed to be subjected to any formal analysis in the event of a failed primary efficacy analysis.

Table 73. Proportion (%) of Responders and OR for Response at Month 6 (FAS, LREG)

Response Criterion	PBO N	PBO %	NMF N	NMF %	OR for Response	95% CI	p-value
≥30% reduction in TAC	137	81	415	88	1.78	[1.01; 3.12]	0.048
≥50% reduction in TAC	137	72	415	75	1.13	[0.70; 1.80]	0.604
≥70% reduction in TAC	137	54	415	59	1.21	[0.79; 1.85]	0.368
Shift in DRL to *low* or below	137	73	415	73	0.96	[0.58; 1.56]	0.860
Subgroup Responder Analyses[a]							
≥30% reduction in TAC	88	78	252	88	2.15	[1.05; 4.39]	0.038
≥50% reduction in TAC	88	72	252	75	1.05	[0.57; 1.90]	0.861
≥70% reduction in TAC	88	53	252	60	1.24	[0.73; 2.10]	0.430
Shift in DRL to *low* or below	88	66	252	65	0.89	[0.49; 1.57]	0.685

Cross-reference: Tables 93, 95, 97, and 99, and 100 to 103

a Patients with a *high* or *very high* DRL at baseline

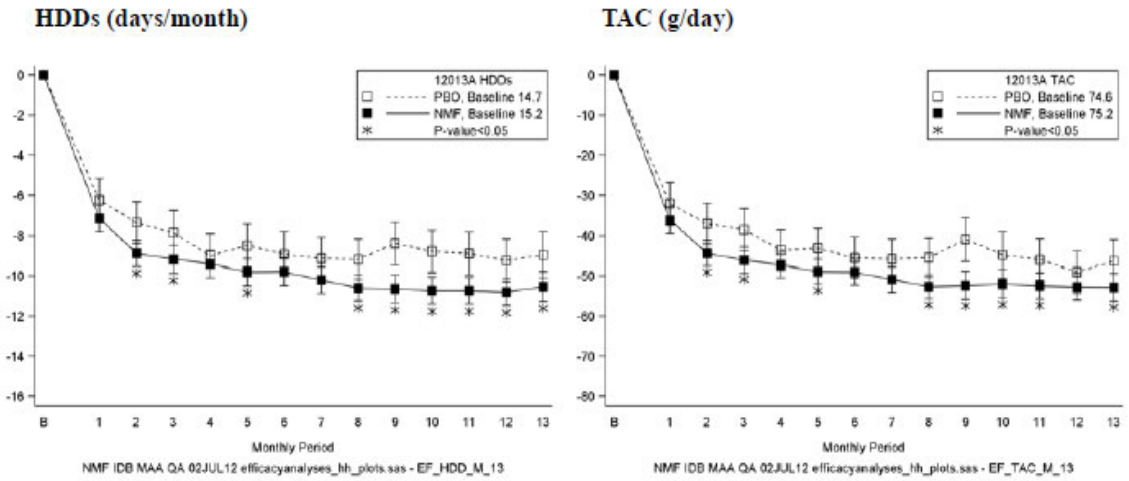
7.2.1.15. Results for other efficacy endpoints

Results at Month 13 for HDDs and TAC were not specified as important endpoints in the prospective SAP, and should therefore be considered minor secondary endpoints (despite this,

the Month 13 results were inappropriately reported by the Sponsor under the misleading heading of 'Co-primary efficacy analysis' in the Summary of Clinical Efficacy). According to the SAP, no minor endpoint should have been subjected to formal analysis after the failure of the primary efficacy analysis, so the Month 13 results are not statistically valid. The Sponsor presented analysis of these endpoints, which are discussed here for completion, but the p-values should be rejected according to the SAP.

At the Month 13 time point, the two treatment groups were statistically separated, with less alcohol consumption in the nalmefene group as indicated by both HDDs and TAC. This provides some evidence of efficacy for nalmefene, but these results must be interpreted with caution given the high withdrawal rate and the Sponsor's deviation from the prospective SAP. Also, it should be noted that several other time points did not show a significant difference for TAC, including Month 12 (see the figure below).

Figure 24. Changes from Baseline in HDDs and TAC (FAS, MMRM) – Lundbeck13



The magnitude of the benefit at Month 13 was small (1.6 HDDs/month and 6.5 g/day), which adds to the concerns about the validity of using Month 13 results as evidence of efficacy. In particular, the fairly small numerical differences between the treatment groups, combined with the large and unequal withdrawal rate in the two groups (placebo 31.7% vs nalmefene 38.1%), raises the possibility that the separation of the curves is at least partly due to withdrawal bias. This could happen, for instance, if poorly motivated subjects who would have recorded higher TAC if they had stayed in the study instead simply withdrew from the study.

Other secondary endpoints generally favoured active treatment, but most statistical comparisons were either narrowly positive (p-values close to 0.05) or showed no significant treatment effect. For CGI-S, as shown below, there was a slight benefit for active treatment (0.2 points) but the 95%CI reached zero difference (p = 0.046). For CGI-I, the results were not significant. Note that none of these p-values was adjusted for multiple comparisons, and none was formally valid, given that the primary efficacy analysis failed to reject the null hypothesis.

Table 74. Results for CGI-S and CGI-I at Month 6 (MMRM; FAS, OC) – Lundbeck13

Efficacy Variable Treatment Group	Baseline		Adjusted Change from Baseline to Week 24		Difference to PBO		
	N	Mean ± SD	N	Mean ± SE	Mean ± SE	95% CI	p-value
CGI-S							
PBO	135	3.9 ± 1.1	104	-0.8 ± 0.1			
NMF	409	4.0 ± 1.1	306	-0.9 ± 0.0	-0.2± 0.1	[-0.4; 0.0]	0.046
CGI-S[a]							
PBO	130	4.0 ± 0.9	100	-0.8 ± 0.1			
NMF	389	4.2 ± 0.9	288	-1.0 ± 0.0	-0.2 ± 0.1	[-0.4; 0.0]	0.047
CGI-I							
PBO			104	2.7 ± 0.1[b]			
NMF			306	2.5 ± 0.1[b]	-0.1 ± 0.1	[-0.4; 0.1]	0.217

Baseline values were based on FAS, OC; changes from baseline and differences to placebo were based on MMRM; FAS, OC values.
Cross-reference: Tables 83, 85, 86, 87, and 90

a Excluding patients with CGI-S=1 at baseline
b Adjusted CGI-I score at Week 24

For number of Non-Drinking Days (NDDs) and for the psychosocial impact measure DrInC-2R, no benefit was observed with active treatment.

Table 75. Summary of Minor Efficacy Variables (MMRM; FAS, OC) – Lundbeck13

Efficacy Variable Treatment Group	Baseline		Adjusted Change from Baseline to Month 6/Week 24		Difference to PBO		
	N	Mean ± SD	N	Mean ± SE	Mean ± SE	95% CI	p-value
Number of NDDs							
PBO	137	9.5 ± 6.2	110	9.9 ± 0.6			
NMF	415	8.9 ± 5.9	320	9.8 ± 0.4	-0.1 ± 0.7	[-1.5; 1.4]	0.930
DrInC-2R							
PBO	137	47.8 ± 25.9	108	-19.5 ± 2.1			
NMF	415	49.0 ± 24.0	318	-18.8 ± 1.3	0.8 ± 2.3	[-3.8; 5.3]	0.746

Baseline values were based on FAS, OC; changes from baseline and differences to placebo were based on MMRM; FAS, OC values.

For the objective surrogate measures based on liver function tests, no significant treatment effect was observed, and there were no substantial trends in favour of nalmefene: ALAT levels were very similar in the two groups at Week 24. (The Sponsor identified some other time points where the groups showed significant differences, as shown in the subsequent tables, but these provide only very weak support for the claim of efficacy, as they represent p-values at time points that have not been adjusted for multiple endpoints, much less multiple time points, and which are part of a process that was not formally valid in terms of the prospectively specified hierarchical testing procedure.)

Table 76. Summary of GGT, ALAT and MCV (MMRM; FAS, OC) – Lundbeck13

Efficacy Variable Treatment Group	Baseline		Adjusted Geometric Mean at Week 24		Ratio to PBO		
	N	Mean ± SD	N	Mean	Ratio	95% CI	p-value
GGT (IU/L)							
PBO	133	78.4 ± 126.2	108	34.5			
NMF	408	70.6 ± 114.6	319	32.2	0.93	[0.83; 1.05]	0.273
ALAT (IU/L)							
PBO	134	32.0 ± 20.1	108	25.8			
NMF	409	34.1 ± 22.8	318	25.6	0.99	[0.90; 1.10]	0.916
MCV ($\times 10^{-15}$L)							
PBO	126	96.4 ± 5.2	103	95.6			
NMF	396	96.6 ± 5.7	296	95.5	1.00	[0.99; 1.01]	0.726

Table 77. Adjusted GGT Values – Back Transformed from Log Scales (FAS, OC, MMRM)

Treatment Group	Week	N	Geometric Mean	Ratio to PBO	95% CI Lower	95% CI Upper	p-value
PBO	12	122	33.30				
	24	108	34.50				
	36	103	37.53				
	52	98	41.25				
NMF	12	361	34.47	1.04	0.91	1.17	0.5813
	24	319	32.24	0.93	0.83	1.05	0.2730
	36	283	31.31	0.83	0.73	0.95	0.0074
	52	259	31.99	0.78	0.67	0.90	0.0011

12013A FINAL Report Results ST_LA06 06SEP2011:18:39:50 1001/147 - TGML/SAD Build Numbers

Table 78. Adjusted ALAT Values – Back Transformed from Log Scales (FAS, OC, MMRM)

Treatment Group	Week	N	Geometric Mean	Ratio to PBO	95% CI Lower	95% CI Upper	p-value
PBO	12	122	24.98				
	24	108	25.78				
	36	103	25.95				
	52	97	27.81				
NMF	12	357	26.41	1.06	0.95	1.17	0.2855
	24	318	25.63	0.99	0.90	1.10	0.9158
	36	281	24.15	0.93	0.84	1.04	0.1906
	52	259	24.55	0.88	0.79	0.99	0.0366

12013A FINAL Report Results ST_LA07 06SEP2011:18:39:58 1001/147 - TGML/SAD Build Numbers

Table 79. Adjusted Changes from Baseline in %CDT Scores (FAS, OC, MMRM)

Treatment Group	Week	N	Mean	SE	Difference to PBO Mean	Difference to PBO SE	95% CI Lower	95% CI Upper	p-value
PBO	12	120	0.13	0.10					
	24	106	0.13	0.11					
	36	100	0.17	0.10					
	52	94	0.41	0.14					
NMF	12	355	-0.06	0.07	-0.19	0.11	-0.42	0.03	0.0903
	24	313	-0.14	0.07	-0.27	0.12	-0.51	-0.04	0.0224
	36	282	-0.10	0.07	-0.27	0.11	-0.49	-0.05	0.0148
	52	255	0.15	0.09	-0.25	0.16	-0.56	0.05	0.1054

12013A FINAL Report Results ST_LA09 06SEP2011:18:40:13 1001/147 - TGML/SAD Build Numbers

7.2.2. Major supportive biotie study (Biotie801, CPH-101-801)

7.2.2.1. Study design, objectives, locations and dates

This double-blind, placebo-controlled, parallel-group study (n = 403) was performed by the previous Sponsor, Biotie, to assess the efficacy and safety of as-needed nalmefene when used to reduce alcohol intake in subjects with heavy alcohol intake who reported difficulty in controlling their use of alcohol. It differed from the Lundbeck studies in several ways, including

the use of a flexible dosing regimen. The target dose was 20mg, as in the subsequent Lundbeck studies, but doses of 10 or 40mg were allowed from Week 3 onwards.

The main treatment period lasted 28 weeks, and this was followed by a randomised Run-Out Period (ROP) lasting for 24 weeks. Only nalmefene subjects who completed the first phase, were willing to enter the second phase and had a "positive treatment response" (according to the investigator) were eligible for the ROP; these subjects (n = 57) were re-randomised to nalmefene or placebo and followed for an additional 24 weeks.

The study was conducted in a number of centres in Finland, from 17th December, 2001 to 29th October, 2004.

7.2.2.2. Inclusion and exclusion criteria

Eligible subjects were aged 18 or higher, were drinking heavily and having difficulties in controlling drinking.

Subjects were also required to have no serious medical or psychiatric problems. Subjects who needed immediate abstinence or inpatient detoxification were not eligible.

Unlike the Lundbeck studies, subjects were not required to have a formal diagnosis of alcohol dependence. At screening, however, 93% of the patients did meet the DSM-IV criteria for alcohol dependence, so the population was broadly similar to that studied in the pivotal Lundbeck studies.

7.2.2.3. Study treatments

Subjects were advised to take nalmefene 20mg (or matching placebo) on as-need basis when drinking was imminent. From Week 3 onwards, subjects could adjust the dose to 10mg or 40mg based on efficacy and tolerability.

In practice, the mean dose was close to the recommended dose of 20mg. The mean dose of study-drug taken during the 28-week treatment period was 19.3mg in the nalmefene group, compared to 25.1mg in the placebo group. This average does not include non-dosing days.

In the randomised withdrawal period (ROP), subjects on active treatment who had shown a favourable response were re-randomised to either continue active treatment or switch to placebo in a blinded fashion.

7.2.2.4. Efficacy variables and outcomes

Efficacy assessments relied substantially on the Timeline Follow Followback (TLFB) method, as described for the pivotal Lundbeck studies, plus a few questionnaires and laboratory tests.

The primary efficacy variable was the number of HDDs per month, where HDDs were defined as days on which a male subject consumed ≥ 5 standard drinks, or a female consumed ≥ 4 standard drinks. A standard drink was estimated to contain 12g of alcohol, although this ranged across 10-14g based on brand and type of beverage consumed.

Additional efficacy variables included:

- Hazardous Drinking Days (HzDDs) per month. A HzDD was defined as a day on which a male subject consumed ≥ 3 or a female consumed ≥ 2 standard drinks, so alcohol consumption on a HzDD could be up to 2 standard drinks less than on HDDs.
- Cumulative number of non-hazardous drinking days.
- Monthly and cumulative number of abstinence days.
- Cumulative number of non-heavy drinking days.
- Monthly number of very heavy drinking days (VHDDs). A VHDD was defined as a day on which male subject consumed ≥ 10 or a female subject consumed ≥ 8 standard drinks.

- Drinks per drinking day (= intensity of drinking), obtained by dividing the total number of drinks by the number of days when any drinking occurred in each month.
- Mean weekly consumption.
- Monthly ratio of heavy drinking days to drinking days.
- Blood alcohol concentration.
- Erythrocyte mean corpuscular volume.
- Alanine aminotransferase, gammaglutamyl transferase, and carbohydrate deficient transferrin.
- Alcohol dependence scale (ADS).
- Drinker inventory of consequences (DrInC, an earlier version of the DrInC R2).
- Subject's and Investigator's Clinical Global Impression (CGI) of efficacy on a 7-point scale, based on the question: "How would you describe the effect of the treatment on your [or 'the subject's'] drinking problem? Compare your [or 'the'] present situation to the situation before the start of the treatment."

The primary efficacy endpoint was the number of HDDs during the 7 one-month periods spanning the main treatment period, as analysed by Poisson analysis or RM-ANOVA (with the method selected according to whether the group means and medians for HDDs were < 5HDDs, in which case a Poisson approach would be used, or ≥ 5 HDDs, in which case RM-ANOVA would be used). It was intended that the Poisson analysis would be based on the overall difference in mean number of HDDs, whereas the RM-ANOVA would be based on the significance of the treatment-by-time interaction.

No key secondary endpoint was identified; all secondary efficacy variables appeared to carry equal weight in the prospective analysis plan.

7.2.2.5. Randomisation and blinding methods

Subjects were randomised unequally using a software package (SAS Release 8.01), initially aiming at higher proportion of nalmefene subjects (240 subjects) than placebo subjects (160 subjects), using random permuted blocks of size 10, balanced by study site.

Blinding was attempted by using identically appearing nalmefene and placebo tablets with medications distributed by subject number only. A bittering agent was used in the placebo tablets, to assist with maintenance of blinding if the subject bit through the coating.

7.2.2.6. Analysis populations

All statistical analyses were based on the intent-to-treat (ITT) population, which included all randomised subjects, except that all data from one centre with unacceptable protocol deviations were excluded.

7.2.2.7. Sample size

Sample size calculations were based on data from an earlier Biotie study (CPH-101-0399). In that study, the data were approximately normally distributed after 16 weeks treatment and the largest standard deviation for HDDs in the three treatment groups was 7.9.

The Sponsor assumed that a between-group difference in the mean HDD values of ≥ 3 would be clinically significant. Demonstrating this difference with a statistical power of 80% and a significance level of 5% (two-sided) would require a group size of 110. Allowing for dropouts, a group size of approximately 160 was required in the main treatment period, and this was set as the target size for the placebo group. To maintain reasonable power during the Run-Out Period (ROP), in which treatment-responsive subjects underwent randomised withdrawal of active

treatment, an initial target of 240 subjects in the nalmefene group was planned, with the expectation that 200 of these subjects would show a treatment response and therefore be eligible and available to be re-randomised at week 28.

The recruitment targets were met (placebo n = 161, nalmefene n = 242), and the study was positive for its primary endpoint, so the study can be considered adequately powered for the main treatment period.

In practice, only 57 subjects entered the ROP, so the second phase of the study was underpowered.

7.2.2.8. Statistical methods

All statistical analyses were performed in the intent-to-treat (ITT) population, and p-values < 0.05 were considered significant. No correction was performed for the use of multiple endpoints. The main treatment period and ROP were analysed separately.

The primary efficacy variable was the number of HDDs during the 7 one-month periods following the screening visit.

The Sponsor proposed two alternative approaches to the primary analysis, with the choice of method depending on the distribution of HDDs. If the group means and medians during the study months 1-7 were mainly below 5 HDDs/month, the HDDs were to be modelled with Poisson distribution. The second approach, the one actually used for the analysis of the primary efficacy variable, was analysis of variance for repeated measures (RM-ANOVA). The RM-ANOVA model for the number of HDDs was to include the effects of treatment group, time and the treatment-by-time interaction, as well as the centre type, gender and the optional factors of family alcohol history and alcohol dependence, as covariates. The baseline number of HDD was to be included as a dependent variable. The main interest was to be in the significance of the treatment group by time interaction, describing the differences between the treatment groups in the monthly changes of the mean HDDs.

A similar approach was taken for HDDs during the randomised withdrawal period, but this was considered a supportive analysis.

Secondary efficacy variables derived directly from the TLFB (NDDs, HzDDs and VHDDs) were to be analysed in the same way as HDDs. For cumulative counts of non-heavy drinking days, abstinence days and non-hazardous drinking days, ANOVA was to be used.

Monthly total alcohol consumption and DrInC scores were to be analysed in the same way as the primary efficacy variable.

Differences between the treatment groups in the distribution of CGI scores were to be analysed using ordinal logistic regression, for the initial 28-week period only.

7.2.2.9. Participant flow

Patient disposition is summarised in the table below. Overall, 63% of patients completed the study and 37% withdrew, but withdrawal was different in the two groups (placebo 32%, nalmefene 40%), which raises the possibility of withdrawal bias.

Table 80. Patient Disposition – Biotie801

	PBO	NMF	Total
Patients randomised	161 (100%)	242 (100%)	403 (100%)
Patients completed	110 (68%)	145 (60%)	255 (63%)
Patients withdrawn	51 (32%)	97 (40%)	148 (37%)
ITT[a]	159	236	395
Patients with treatment response at Week 28	*48*	*86*	*134*
Patients re-randomised to 24-week ROP		*57 (NMF-NMF 30; NMF-PBO 27)*	
Patients completed 52 weeks		*NMF-NMF 26 NMF-PBO 23*	

a The statistical inference was based on the analyses performed for the ITT, which included all randomised patients except those for Site 33 (Helsinki 2); Site 33 was excluded due to *Good Clinical Practice* issues.

7.2.2.10. Major protocol violations/deviations

The Biotie study report provides narrative accounts of many individual protocol deviations, without summarising the overall incidence of protocol deviations. The deviations mentioned include incomplete recording of adverse events, failures to obtain proper consent, inaccurate dosing, disallowed concomitant medication, and missing study visits. Overall, the number of violations appeared typical of a study of this nature and is unlikely to have had a major impact on the findings.

7.2.2.11. Baseline data

The treatment groups appeared to be adequately matched at baseline for key demographic features, as shown below.

Table 81. Summary of Gender, Age and BMI – Biotie801

	Placebo N=161	Nalmefene N=242	Total N=403
Gender	29 (18%) Females 132 (82%) Males	46 (19%) Females 196 (81%) Males	75 (19%) Females 328 (81%) Males
Age, years	48.8 (8.4)	49.5 (9.1)	49.2 (8.8)
BMI, kg/m^2	27.8 (4.7)	28 (4.8)	27.9 (4.8)

Mean (with SD) is given for age and body mass index (BMI). Source: sections 14.1.1.1-2, 14.1.1.6.

The history of alcoholism in the two treatment groups also appeared to be matched, but this information was not presented in a convenient tabular format. As shown in the tables below, the baseline count of HDDs was similar in the two groups.

7.2.2.12. Results for the primary efficacy endpoint

The primary efficacy analysis across all 7 months of the main treatment period showed a significant reduction in HDDs in the nalmefene group by RM-ANOVA (p = 0.0065, treatment-by-time interaction).

The mean number of HDDs in each group is shown in the figure and table below. At Month 7, the mean reduction from baseline was 5.6 HDDs in the placebo group, compared to 6.4 HDDs[16] in the nalmefene group, a difference of 0.8 days. This is of only moderate clinical utility; Biotie's sample size estimation procedure had assumed that a difference of 3 HDDs would be clinically significant, and the observed difference is only 27% of this. Allowing for the possibility that the treatment effect has been inflated by a withdrawal bias, the clinical benefit could be even less.

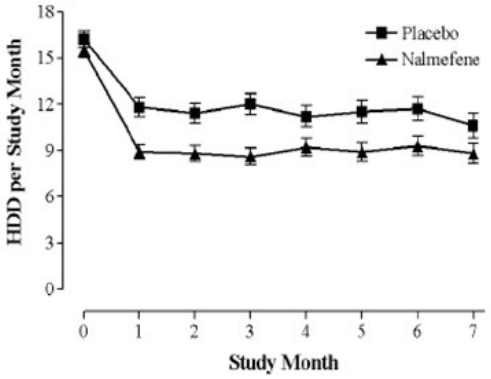
[16] A typographical error in CER1 changed the reported nalmefene reduction; the actual between-group difference of 0.8 HDDs was less than that reported in CER1, so this study was actually weaker than initially suspected.

Table 82. Change in HDDs by Study Month – Biotie801

Study month	Placebo							Nalmefene						
	HDD count			Percent			N	HDD count			Percent			N
	Median	Mean	SD	Median	Mean	SD		Median	Mean	SD	Median	Mean	SD	
1	-3.0	-4.5	7.0	-22.2	-24.2	41.8	143	-6.0	-6.6	7.0	-43.8	-40.4	41.1	215
2	-4.0	-4.8	6.7	-28.6	-26.5	44.0	133	-5.0	-6.4	7.2	-46.2	-40.6	42.9	193
3	-3.0	-4.3	7.0	-26.9	-22.0	46.9	128	-6.0	-6.7	7.3	-50.0	-42.7	43.0	174
4	-3.0	-4.9	7.1	-27.3	-26.5	46.0	123	-6.0	-6.0	7.4	-42.1	-38.7	46.5	159
5	-4.0	-4.8	7.5	-29.6	-24.7	48.3	118	-6.0	-6.4	7.7	-46.2	-40.2	46.7	147
6	-3.0	-4.2	7.4	-20.0	-21.9	50.3	114	-6.0	-6.1	7.7	-46.2	-39.1	45.1	147
7	-5.0	-5.6	7.3	-33.3	-32.8	44.2	102	-6.0	-6.4	6.8	-43.8	-42.8	43.2	129

The change is presented *vs.* baseline. N, subject count. Source: section 14.2.1.1

Figure 25. HDDs (days/month) (OC) – Biotie801



Treatment Group	N	HDDs (days/month)	
Month		Mean	SD
Placebo			
Baseline	159	16.2	6.9
1	143	11.8	7.5
7	102	10.6	8.3
Nalmefene			
Baseline	236	15.5	6.9
1	217	8.9	6.8
7	131	8.8	7.3

A sensitivity analysis, based on different methods of imputing missing data, showed that pessimistic imputation (BOCF) showed a significant treatment effect for the first 3 months by the Kruskal-Wallis test, though significance was not achieved at later dates. Optimistic imputation methods (LOCF) produced significant results at all time points (see the table below).

Table 83. Statistical Scenarios for HDDs – Biotie801

The effect of imputation models on p values obtained with Kruskal-Wallis test at baseline and at monthly assessments. "Original" refers to not replacing missing values, "Best case" to carrying the last observation forward, and "Worst case" to replacing the missing values with baseline HDD count.

Study month	Original			Best case			Worst case		
	Placebo N	Nalmefene N	Kruskal-Wallis p value	Placebo N	Nalmefene N	Kruskal-Wallis p value	Placebo N	Nalmefene N	Kruskal-Wallis p value
0	159	236	0.3086	159	236	0.3086	159	236	0.3086
1	143	217	0.0001	159	238	0.0002	159	238	0.0002
2	133	195	0.0008	159	238	0.0007	159	238	0.0103
3	128	176	<0.0001	159	238	<0.0001	159	238	0.0053
4	123	161	0.0156	159	238	0.0021	159	238	0.2346
5	118	149	0.0015	159	238	0.0004	159	238	0.1224
6	114	149	0.0120	159	238	0.0008	159	238	0.1039
7	102	131	0.1143	159	238	0.0127	159	238	0.5743

Month 0 refers to baseline. Source: section 14.2.1.1.

7.2.2.13. Results for other efficacy outcomes

Changes in alcohol consumption recorded in the TLFB were generally consistent with the primary analysis of HDDs, but not all endpoints achieved significance.

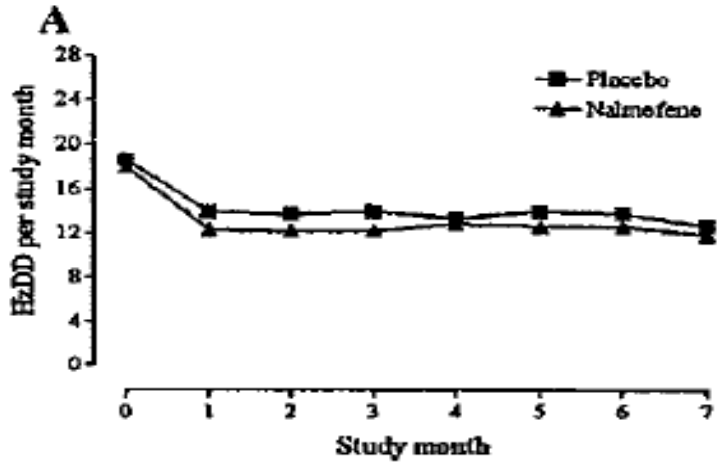
For Hazardous Drinking Days (HzDDs), there was a reduction in both groups, as shown in the table below, but there was no statistically significant difference in the decrease between the groups (p = 0.3034, repeated measures ANOVA). Not surprisingly, the complementary efficacy variable, Non-hazardous drinking days did not show a significant difference either.

Table 84. Hazardous Drinking Days – Biotie801

Study month	Placebo								Nalmefene							
	Count				Change				Count				Change			
	Median	Mean	SD	N	Median	Mean	SD	N	Median	Mean	SD	N	Median	Mean	SD	N
0	19	18.6	6.8	159	-	-	-	-	17	18.0	6.8	236	-	-	-	-
1	13	13.9	8.2	143	-4	-4.6	7.2	143	11	12.3	7.7	217	-4	-5.8	6.8	215
2	12	13.7	7.8	133	-3	-4.6	6.6	133	11	12.2	8.2	195	-5	-5.7	6.9	193
3	13	13.9	7.9	128	-3.5	-4.4	7.0	128	11	12.2	8.3	176	-5	-5.7	6.9	174
4	13	13.3	7.9	123	-3	-4.8	6.7	123	12	12.8	8.2	161	-4	-5.1	7.1	159
5	12.5	13.9	8.3	118	-3	-4.4	6.9	118	11	12.6	8.3	149	-5	-5.5	7.3	147
6	12	13.8	8.1	114	-2	-4.0	6.8	114	11	12.6	8.2	149	-5	-5.5	7.2	147
7	11	12.7	8.2	102	-4	-5.5	6.9	102	11	11.9	8.2	131	-5	-6.2	6.9	129

The change is presented *vs.* baseline (month 0). N, subject count. Source: section 14.2.1.2.

Figure 26. Hazardous Drinking Days (HzDDs) – Biotie801



For Non-Drinking Days (NDDs), also known as Abstinence Days, there was a significant increase in both groups over time. The treatment-by-visit interaction appeared significant (p = 0.0499) but visual inspection of the monthly values does not suggest any clear treatment effect. The increase in mean NDDs from baseline to Month 7 was numerically greater in the placebo group (increase of 5.2 NDDs, from 8.5 to 13.7) than in the nalmefene group (increase of 5 NDDs, from 8.8 to 13.8), but the nalmefene group showed greater increases at intermediate time points.

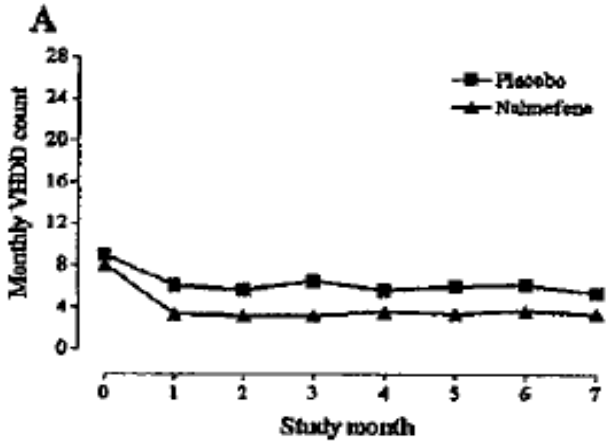
Table 85. Number of Abstinence Days by Study Month – Biotie801

Study month	Placebo				Nalmefene			
	Median	Mean	SD	N	Median	Mean	SD	N
0	8	8.5	6.6	159	8.5	8.8	6.6	236
1	12	12.2	8.4	143	16	13.7	8.2	217
2	13	12.6	8.1	133	16	14.0	8.3	195
3	13	12.8	8.0	128	16	14.1	8.7	176
4	14	13.3	8.0	123	14	13.2	8.5	161
5	13	12.4	8.4	118	15	13.4	8.8	149
6	13	12.6	8.2	114	15	13.2	8.9	149
7	14	13.7	8.2	102	14	13.8	8.9	131

Month 0 refers to baseline. N, subject count. Source: section 14.2.1.3

The median monthly number of VHDDs showed a significant treatment effect; it was reduced from 8 at baseline to 4-5 in the placebo group and from 7 to 1 - 2 in the nalmefene group, with the between-group difference being statistically significant (p < 0.0001 by linear mixed model).

Figure 27. Very Heavy Drinking Days (VHDDs) – Biotie801



The median intensity of drinking (standard drinks per drinking day) was reduced from 8 .9 to 7.4 drinks per drinking day in the placebo group and from 8 .7 to 6.0 drinks per drinking day in the nalmefene group. This was significant by the treatment-by-time interaction (p = 0.0134).

For the CGI analysis, the results were also significantly in favour of nalmefene. The distributions of CGI scores at Week 28 (or at early termination) were significantly better for nalmefene recipients both for subject and investigator-assessed scores (p < 0.001; ordinal logistic regression, OC).

In the nalmefene group, 43% of the patients reported 'much' or 'very much' improvement, compared to 28% of the patients in the placebo group. The corresponding percentages for the investigator-assessed scores were slightly better, but with a similar group difference (nalmefene 48% and placebo 33%).

No significant or important differences emerged between the groups in a number of pharmaco-economic variables.

For the surrogate endpoints based on liver function tests, mean GGT and ALAT values decreased in both groups, but significantly more in the nalmefene group than in the placebo group (GGT, p = 0.002; ALAT, p = 0.009; RM-ANOVA, treatment-by-visit interaction effect).

Table 86. GGT and ALAT (OC) – Biotie801

Treatment Group	N	GGT (U/L)		ALAT (U/L)	
Week		Mean	SD	Mean	SD
Placebo					
Baseline	159	107.4	(172.8)	46.9	(34.2)
12	130	109.5	(173.2)	49.1	(40.0)
28	115	115.4	(189.8)	52.0	(45.4)
Nalmefene					
Baseline	238	102.7	(141.3)	43.8	(29.6)
12	179	81.2	(112.6)	41.0	(47.4)
28	148	77.2	(89.1)	37.5	(24.5)

Table 87. MCV, CDT, ALAT and GGT – Biotie801

Study month		MCV, fl			CDT, %			ALAT, U l^{-1}			GGT, U l^{-1}		
		0	3	7	0	3	7	0	3	7	0	3	7
Placebo	Median	96	96	96	3.0	3.0	2.6	35	36	40	64	59	66
	Mean	96.4	96.4	96.3	3.4	3.6	3.3	46.9	49.1	52.0	107.4	109.5	115.4
	SD	4.9	4.8	4.9	1.6	1.9	2.0	34.2	40.0	45.4	172.8	173.2	189.8
	N	159	129	114	159	130	115	159	130	115	159	130	115
Nalmefene	Median	96	96	96	2.8	3.0	2.5	34.5	29	28	57	42	45
	Mean	96.8	96.4	96.1	3.3	3.4	3.1	43.8	41.0	37.5	102.7	81.2	77.2
	SD	4.1	4.0	3.8	1.5	1.9	1.8	29.6	47.4	24.5	141.3	112.6	89.1
	N	238	179	147	237	179	148	238	179	148	238	179	148

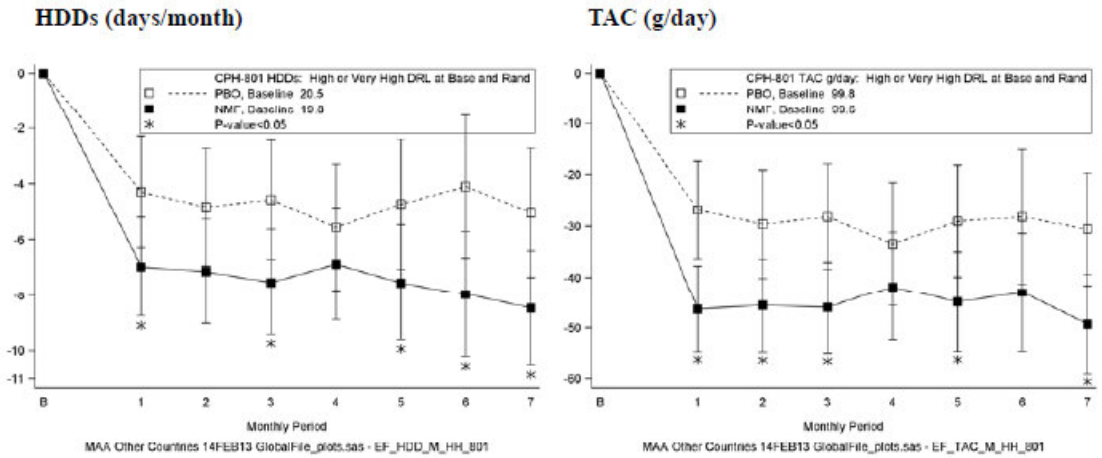
MCV, mean corpuscular volume; CDT, carbohydrate-deficient transferrin; ALAT, alanine aminotransferase; GGT, gamma-glutamyl transferase. Month 0 refers to baseline. Source: sections 14.2.3.1-4.

Blood alcohol was measured at all treatment visits, but did not show a significant difference between groups: 90% of subjects had zero blood alcohol at any one visit, and 95% had levels < 0.05%.

7.2.2.14. Results in high-risk drinkers

Lundbeck, the current Sponsor, performed a post hoc subgroup analysis of the Biotie data, based on the results in patients with high or very high DRL at Baseline and Randomisation. This analysis was included in the SCE, but was not in the original Biotie study report. In this subgroup, significant reductions in both HDDs and TAC were demonstrated at Month 7 by MMRM, as well as at several other time points, but the Month 6 results failed to achieve significance for TAC despite post hoc selection of a favourable subgroup.

Figure 28. Changes from Baseline in HDDs and TAC (FAS, MMRM) – Patients with a High or Very High DRL at Baseline and Randomisation – Biotie801



7.2.2.15. Run-out period

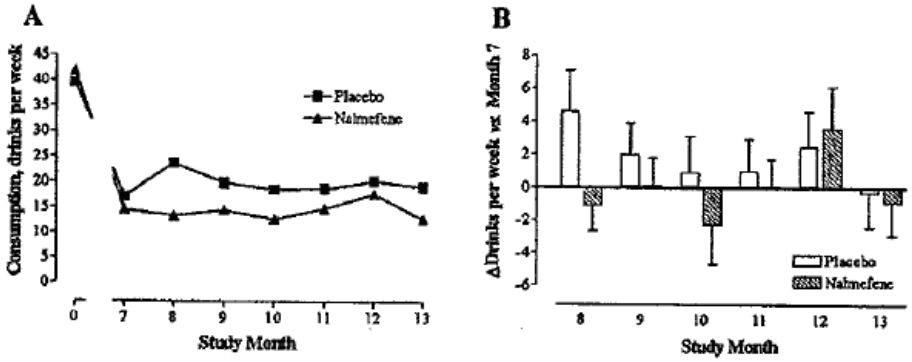
Efficacy analysis in the ROP was underpowered (n = 57). Eighty-six of the 145 patients in nalmefene group completing the main treatment period were considered to have had a positive treatment response, but only 57 of these were willing to continue treatment. As shown in the figure and table below, there was a trend suggesting a higher mean intake in those randomised to withdrawal of nalmefene (switching to placebo), but this was not significant. After showing a favourable response to treatment, subjects randomised to continue or discontinue nalmefene both showed a persistent low level of HDDs compared to their pre-treatment baseline.

Although these results are inconclusive because of low patient numbers, the lack of a major rebound effect could be considered a favourable feature of nalmefene treatment, but it does

raise the possibility that the main benefits achieved in the study population were not pharmacologically mediated.

Figure 29. Alcohol Consumption During Randomised Withdrawal – Biotie801

A. Mean alcohol consumption by study month. B. Change since re-randomization. The values are standard drinks per week. Error bars are standard error of the mean. Source: section 14.6.1.5.



7.3. Minor supportive efficacy studies

7.3.1. Supportive study biotie701 (CPH-101-0701)

7.3.1.1. Study design, objectives, locations and dates

This double-blind, placebo-controlled, parallel-group Biotie study assessed the efficacy of as-needed nalmefene when used to reduce alcohol consumption in the setting of heavy drinking. The target dose was 20mg, but doses of 10mg or 40mg were allowed from Week 3 onwards.

7.3.1.2. Inclusion and exclusion criteria

The study was aimed at subjects with moderate or high alcohol intake who wanted to reduce their alcohol intake.

The main inclusion criteria were as follows:

- a desire to reduce or gain better control of alcohol consumption
- difficulty in controlling drinking
- a positive family history of alcohol problems
- at least 18 heavy drinking days and no more than 14 consecutive abstinence days during the 12 weeks preceding the first screening visit (i.e. ≥ 6 HDDs/month)
- age 18 years or older.

The main exclusion criteria were:

- a severe hepatic or renal disorder
- dementia
- a seizure disorder
- encephalopathy
- any other disorder taking priority over treating the drinking problem or likely to interfere with study treatment or compliance
- a medical need for immediate total abstinence
- dependence on drugs other than alcohol, or current use of illicit drugs

- treatment with disulfiram, naltrexone, acamprosate or opioid agonists
- pregnancy
- nursing

Only 77% of the patients met the DSM-IV criteria for alcohol dependence, so this study assessed a broader population than in the pivotal Lundbeck studies. The minimum number of HDDs at Screening was similar to the Lundbeck studies (≥ 6 HDDs/month).

7.3.1.3. Study treatments

Subjects were advised to take nalmefene 20mg (or matching placebo) on as-need basis when drinking was imminent. From Week 3 onwards, subjects could adjust the dose to 10mg or 40mg based on efficacy and tolerability.

The mean dose of study-drug taken during the 28-week treatment period was 22.9mg in the nalmefene group and 26.4mg in the placebo group

7.3.1.4. Efficacy variables and outcomes

Efficacy variables were as described for the larger Biotie Study (Biotie801), in Section 7.2.2.4.

The primary efficacy endpoint was the number of HDDs, as assessed by a linear mixed model, with the main statistical test for significance based on the treatment-by-time interaction in RM-ANOVA, over the course of the main treatment period.

No secondary endpoint was identified as the key secondary endpoint.

7.3.1.5. Randomisation and blinding methods

Randomisation was performed using a centralised software package, using random permuted blocks, of six, balanced by study site.

Blinding was maintained by using placebo tablets that were identical in appearance to nalmefene tablets. As in other Biotie studies, a bittering agent was also used in the placebo tablets.

7.3.1.6. Analysis populations

The primary analysis population was the intent-to-treat population consisting of all randomised subjects. All treated subjects contributed safety data.

7.3.1.7. Sample size

To assist with sample size estimations, data from a previous from a previous study (CPH-101-0399) was used, but only subjects with a positive family history were included in the estimation. A simple t-test approach was used, despite the fact that the data were subsequently analysed by RM-ANOVA.

The earlier data were approximately normally distributed, and the largest standard deviation seen in the three treatment groups was 9 HDDs. At 16 weeks, the observed mean difference between placebo and 40 mg nalmefene was 5 HDDs. To demonstrate a similar difference with a statistical power of 80% and two-sided significance level of 5%, the Sponsor estimated that 52 subjects per group would be required. Allowing for a drop out rate of 30%, a final sample size of 75 per group was selected.

Recruitment targets were exceeded, but the dropout rate was much higher than expected (71%), so the study was ultimately underpowered. The size of the treatment effect was also much smaller (~1.4 days) than the 5 HDDs anticipated in the sample size estimation, so the failure to achieve statistical significance is not solely because of low patient numbers.

7.3.1.8. Statistical methods

The statistical methods used in this study were as described for the larger Biotie study (Biotie801). As in that study, the primary analysis technique depended upon the number of HDDs, with a Poisson analysis intended if HDDs were mostly < 5, and an RM_ANOVA intended for higher HDD counts. The primary analysis eventually proceeded with an RM-ANOVA.

A significance level of 5% was used ($p < 0.05$) without any apparent correction for the use of multiple endpoints.

7.3.1.9. Participant flow

Patient disposition is summarised in the table below. The withdrawal rate was very high (71% overall in the main treatment period), which basically invalidates the study. Of the 49 subjects who completed the main treatment period, all entered the Safety Follow-up, but another 8 of these withdrew, so the final withdrawal rate was 75% (126/167). About a third of the nalmefene recipients withdrew because of AEs, compared to only a low proportion of placebo recipients, raising substantial concerns about withdrawal bias and unblinding.

Table 88. Patient Disposition – Biotie701

	PBO	NMF	Total
Patients randomised	82	85	167
Weeks 0 to 28			
Patients completed	28 (34%)	21 (25%)	49 (29%)
Patients withdrawn	54 (66%)	64 (75%)	118 (71%)
Safety Follow-up; Weeks 29 to 40			
Patients completed	24 (86%)	17 (81%)	41 (84%)
Patients withdrawn	4 (14%)	4 (19%)	8 (16%)

Table 89. Reasons for Premature Discontinuation – Biotie701

Reason	Total	Adverse event	Lack of efficacy or clear clinical deterioration	Protocol violation	Consent withdrawal	Lost to follow-up	Other
Total	118 (70.7)	23 (19.5)	6 (5.1)	2 (1.7)	44 (37.3)	41 (34.7)	2 (1.7)
Placebo	54 (65.9)	3 (5.6)	3 (5.6)	2 (3.7)	26 (48.1)	20 (37.0)	0 (0)
Nalmefene	64 (75.3)	20 (31.3)	3 (4.7)	0 (0)	18 (28.1)	21 (32.8)	2 (3.1)

Number of subjects (with percentage in parentheses) is given. Source: section 14.4.3.3.

7.3.1.10. Major protocol violations/deviations

Major protocol violations leading to withdrawal from the study were uncommon, affecting 2 subjects in the placebo group and none in the nalmefene group. A variety of minor protocol violations were described by the Sponsor, including administration of placebo for two weeks to a single nalmefene recipient and several missed visits (48 events). There were 9 cases of forbidden medication being used.

These deviations are within the expected limits for a study of this nature and are unlikely to have contributed substantially to the findings of the study.

7.3.1.11. Baseline data

The two treatment groups were reasonably well-matched in terms of age and gender distribution, social situation, and baseline drinking habits, as shown in the tables below. Mean HDD counts at baseline were very similar (placebo 21.0, nalmefene 21.1 HDDs/month).

Table 90. Age and Gender – Biotie701

Age	Placebo			Nalmefene			Total		
	Male	Female	All	Male	Female	All	Male	Female	All
Mean	45.8	43.2	44.8	46.0	45.5	45.8	45.9	44.4	45.3
SD	11.8	7.4	10.4	9.0	8.2	8.6	10.4	7.8	9.5
Min	24	29	24	26	30	26	24	29	24
Max	76	59	76	68	63	68	76	63	76

Table 91. Social Characteristics – Biotie701

		Placebo N (%)	Nalmefene N (%)	Total N (%)
Educational status	Secondary school or less	29 (35.4)	34 (40.0)	63 (37.7)
	Upper secondary school	15 (18.3)	17 (20.0)	32 (19.2)
	College	28 (34.1)	23 (27.1)	51 (30.5)
	University graduate	10 (12.2)	11 (12.9)	21 (12.6)
Current employment status	Student	2 (2.4)	3 (3.5)	5 (3.0)
	Full time employed	43 (52.4)	29 (34.1)	72 (43.1)
	Part time employed	6 (7.3)	11 (12.9)	17 (10.2)
	Unemployed	18 (22.0)	32 (37.6)	50 (29.9)
	Retired	8 (9.8)	5 (5.9)	13 (7.8)
	Other	5 (6.1)	5 (5.9)	10 (6.0)
Current marital status	Never married	16 (19.5)	20 (23.5)	36 (21.6)
	Married	39 (47.6)	35 (41.2)	74 (44.3)
	Separated or divorced	25 (30.5)	28 (32.9)	53 (31.7)
	Widowed	2 (2.4)	2 (2.4)	4 (2.4)
Living arrangement	Alone	17 (20.7)	19 (22.4)	36 (21.6)
	With spouse or partner	26 (31.7)	24 (28.2)	50 (29.9)
	With spouse or partner and children	25 (30.5)	28 (32.9)	53 (31.7)
	Single with children	7 (8.5)	9 (10.6)	16 (9.6)
	Other	7 (8.5)	5 (5.9)	12 (7.2)

Subject count (%) is shown. Source: sections 14.1.1.7-14.1.1.10.

Table 92. Baseline Drinking – Biotie701

	Placebo N=81			Nalmefene N=84			Total N=165		
	Mean	SD	Median	Mean	SD	Median	Mean	SD	Median
HDD count	21.0	7.3	24	21.1	7.4	24.5	21.0	7.4	24
VHDD count	11.2	9.5	10	11.3	9.6	9.5	11.2	9.6	10
Abstinence day count	4.9	6.4	1	5.5	6.6	0.5	5.2	6.5	1
Drinks per week	86.6	44.5	76.7	81.2	41.9	71.1	83.8	43.1	73.6
Drinks per drinking day	15.3	7.5	12.8	14.6	6.0	14.4	14.9	6.8	13.2
Ratio of HDD to drinking days	90.6	16.7	100.0	94.3	15.1	100.0	92.5	15.9	100.0

Source: section 14.2.1.1.

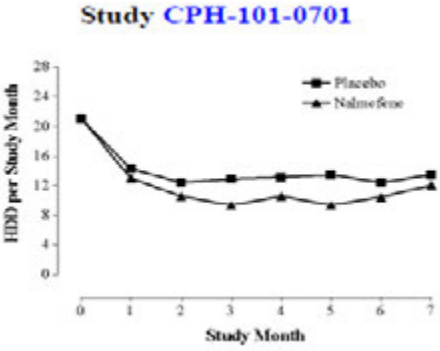
7.3.1.12. Results for the primary efficacy outcome

The difference between the treatment groups in the reduction of HDDs was numerically in favour of nalmefene but this did not achieve statistical significance (p = 0.088; RM-ANOVA, treatment-by-visit interaction). The mean reduction in HDDs in the placebo group was 7.6 HDDs, compared to 9 HDDs in the nalmefene group, a difference of 1.4 HDDs; no 95%CIs were provided.

The difference between groups became statistically significant upon adding interaction terms to the model (gender-by-time, centre-type-by-time, group-by-centre-type and group-by-gender) in one of several pre-specified supportive analyses. No correction was made for the use of multiple statistical models, so this supportive analysis adds little weight. A significant difference between groups was also observed at the isolated time point of Month 3, which again carries little weight as this time point was not part of the primary analysis.

In view of the very high withdrawal rate (71% in the main treatment period), the data has been obtained from a small non-random subset of the original study population, and no firm conclusions can be drawn from this study.

Figure 30. HDDs (days/month) (OC) – Biotie701



Treatment Group	N	HDDs (days/month)	
Month		Mean	SD
Placebo			
Baseline	81	21.0	7.3
1	67	14.3	9.2
7	31	13.4	9.2
Nalmefene			
Baseline	84	21.1	7.4
1	60	13.0	8.9
7	21	12.1	10.2

Table 93. Heavy Drinking Days by Study Month

Study Month	Placebo				Nalmefene			
	Median	Mean	SD	N	Median	Mean	SD	N
Baseline	24	21.0	7.3	81	24.5	21.1	7.4	84
1	15	14.3	9.2	67	10	13.0	8.9	60
2	11	12.4	9.4	57	8	10.6	8.9	46
3	12.5	12.9	9.3	52	7.5	9.4	9.5	38
4	13.5	13.1	9.1	44	8	10.6	9.6	36
5	13	13.4	8.7	40	8	9.4	9.1	33
6	13	12.4	9.4	37	8	10.5	10.3	30
7	13	13.4	9.2	31	9	12.1	10.2	21

7.3.1.13. Results for other efficacy outcomes

For a range of alcohol intake measures derived from the TLFB, minor differences were noted between groups that failed to reach statistical significance.

For the secondary endpoints of patient and investigator CGI, there was a significant treatment effect. According to patient assessments, 44% of the patients in the nalmefene group and 32% (misquoted as 22% in the SCE) of the patients in the placebo group reported “much” or “very much” improvement. The overall distributions of CGI scores were significantly different across the two treatment groups (p=0.0038; ordinal logistic regression). For investigator assessments, “much” or “very much” improved scores were reported in 42% and 18% of the nalmefene and placebo groups, respectively (p = 0.016; ordinal logistic regression).

For the psychosocial impact measures, DrInC and ADS, no significant treatment effect was noted.

(Note that the study report claims, in its Efficacy Conclusions section, “marginally significant” results were obtained. “A statistically marginally significant difference favoring nalmefene was noted also in the DrInC and ADS scores during the treatment period .” Biotie appears to have used this term to describe strong trends that failed to achieve significance, but not even strong trends were observed for DrInC. The between-group difference was less than one point from a mean score of 27.9; p = 0.52. Biotie could have been referring to isolated components or secondary analyses of the DRInC. For ADS, by the end of the treatment period, the mean scores had declined from 18.5 to 13.4 and 12.2 in the placebo and nalmefene groups, respectively, and this difference approached significance; p = 0.0833.)

For the laboratory based endpoints, no significant treatment effect was demonstrated. There was a trend suggesting lower mean ALAT in the placebo group for the 28-week treatment period, which approached significance (p = 0.0505; RM-ANOVA, treatment-by-visit interaction). The differences in favour of placebo were most marked at Week 28, but were not persistent. At week 40, the ALAT values in the two groups were similar (placebo 38.3 IU/L, nalmefene 40.3 IU/L). There were no important differences between the treatment groups in mean GGT.

Figure 31. GGT and ALAT (OC) – Biotie701

Treatment Group / Week	N	GGT (U/L) Mean	GGT (U/L) (SD)	ALAT (U/L) Mean	ALAT (U/L) (SD)
Placebo					
Screening	82	128.5	(165.5)	45.1	(41.4)
12	48	114.1	(142.5)	37.8	(37.1)
28	49	93.6	(113.0)	32.1	(18.6)
40	24	121.4	(136.7)	38.3	(26.7)
Nalmefene					
Screening	85	119.3	(172.5)	41.0	(52.9)
12	35	82.1	(100.0)	35.5	(34.2)
28	48	105.5	(254.8)	39.5	(40.1)
40	17	74.1	(71.1)	40.3	(29.8)

7.3.1.14. Subgroup analyses

Formal subgroup analyses were not performed. They would have been inappropriate given that the main analysis was underpowered.

The effects of gender, family history of alcoholism and centre were explored within the RM-ANOVA model. Because of low patient numbers and a failure of the primary endpoint to achieve significance, these exploratory analyses did not add any useful insights.

7.3.2. Supportive study biotie299 (CPH-101-0299)

7.3.2.1. Study design, objectives, locations and dates

This 52-week study was a placebo-controlled, double-blind, parallel-group dose-ranging study of the efficacy of nalmefene in the treatment of heavy alcohol drinkers with impaired control of their alcohol intake.

The primary objective was to determine the efficacy and dose-response of nalmefene in reducing heavy alcohol drinking and promoting abstinence in alcohol dependent subjects. Secondary objectives were to assess the safety of nalmefene, and to explore the viability of as-needed use of nalmefene in preventing relapse to heavy drinking.

In contrast to the pivotal studies, this study tested fixed dosing as well as as-needed dosing. It also differed from the pivotal studies in that it explored the efficacy of multiple doses, ranging from 5mg to 40mg nalmefene. Treatment began with a fixed daily dosing phase of 12 weeks, followed by a 40-week extension period during which study drug was taken on an as-needed basis.

The study was conducted from 25th February 2000 to 5th October 2001, in 13 centres in the United States.

7.3.2.2. Inclusion and exclusion criteria

The main eligibility criteria were that subjects had to be aged 21 or higher, have alcohol dependence according to DSM-IV, and report drinking heavily and having difficulties in controlling their drinking.

In particular, subjects had to have alcohol dependence including at least one of the following DSM-IV criteria:

- alcohol is often taken in larger amounts or over a longer period than intended
- persistent desire to cut down or control drinking

They had to have had at least 8 heavy drinking days within the last 6 weeks prior to screening, with an HDD defined as ≥ 5 standard drinks for male, or ≥ 4 standard drinks for female.

Subjects had to have a period of abstinence of 3 consecutive days immediately prior to randomisation. This was a somewhat unusual requirement, because it selected for patients able to control their alcohol intake, at least briefly, without the need of medication. Such subjects might be expected to do quite well with placebo treatment, making it difficult to demonstrate efficacy; they might also stop drinking in response to unblinding.

Subjects required a CIWA-Ar (Clinical Institute of Withdrawal Assessment for Alcohol –revised) score less than 8 at Randomisation, indicating that alcohol cessation would be safe.

Subjects were not eligible if they needed immediate abstinence or inpatient detoxification, or had serious medical or psychiatric problems (dementia, seizure disorder, mental retardation, encephalopathy or any other serious medical comorbidity). Subjects could not be pregnant, nursing or at risk of pregnancy. Subjects were also ineligible if they were repeatedly positive on a drug screen test (positive at screening and in a re-test ≥ 3 days later).

7.3.2.3. Study treatments

Patients were randomly assigned to one of four treatment groups:

- placebo
- nalmefene 5mg
- nalmefene 20mg
- nalmefene 40mg

Patients took their assigned treatment daily for 12 weeks, then on as as-needed basis for 40 weeks, up to once per day, whenever they felt alcohol consumption was imminent.

7.3.2.4. Efficacy variables and outcomes

The primary efficacy variable was HDDs, as in the other Biotie studies.

The other main efficacy variables resembled those from previously described studies:

- Drinking data were recorded with the TLFB method, as described in the pivotal studies, and the data was expressed as HDDs, VHDDs, NDDs, monthly ratio of heavy drinking days to drinking days, mean weekly consumption, maximum intensity of drinking, drinks per drinking day.
- Blood alcohol concentration was assessed with a breathalyser.
- Blood was collected for laboratory analysis of mean corpuscular volume, alanine aminotransferase, gamma-glutamyl transferase, carbohydrate deficient transferrin
- Subjects completed self-administered questionnaires on the severity of alcohol dependence and on adverse consequences of alcohol abuse (obsessive compulsive drinking scale OCDS, alcohol dependence scale ADS, drinker inventory of consequences DrInC.)
- The subject's drinking pattern was assessed by his/her spouse/significant other.

The two study phases, the 12 week fixed daily dosing phase and the 40 week as-needed phase, were analysed separately, and neither was formally designated as more important than the other, but subjects were required to complete 12 weeks to enter the per-protocol population and the exploration of as-needed dosing was a lower ranking objective than the assessment of efficacy. This and other indicators imply that the primary analysis phase was the first 12 weeks, so efficacy in the first 12 weeks, as reflected in HDDs, should be considered the primary endpoint.

7.3.2.5. Randomisation and blinding methods

Randomisation was performed using a centralised software package, and blinding was attempted, as in other Biotie studies, by using identically appearing and similarly tasting placebo and nalmefene tablets.

7.3.2.6. Analysis populations

The Sponsor identified three analysis populations.

The intent-to-treat (ITT) population consisted of all treated subjects for whom there was at least one post-baseline efficacy assessment.

The per protocol (PP) population consisted of subjects who:

- Completed 12 weeks of study medication use
- Did not miss more than 5 doses of medication in the month
- Did not use disulfiram or naltrexone
- Satisfied all inclusion and exclusion criteria

The safety population consisted of all randomised subjects.

7.3.2.7. Sample size

The Sponsor performed a sample-size estimation based on a two-sided t-test, assuming that 60% of the patients in the placebo group would have one or more HDDs per week. A sample size of n = 240 (n = 60 per group) would result in a power of 0.80 (with a Type I error of 0.05) for detecting a between-group difference of 20 percentage points in the proportions of heavy versus non-heavy drinking.

7.3.2.8. Statistical methods

The 12 week daily dosing phase and the 40 week extension phase were analysed separately.

The primary assessment of efficacy for the 12 week daily dosing phase was based on the change from screening in the number of HDDs per month. The primary efficacy variable, HDDs, was analysed by Poisson regression analysis for repeated measures, but repeated-measures ANOVA was also performed.

For the 40 week extension phase, the number of HDDs was calculated over 5 two-month periods, and the primary efficacy assessment was based on the changes from the third study month (weeks 8 - 12) of the number of heavy drinking days over the 5 two-month periods.

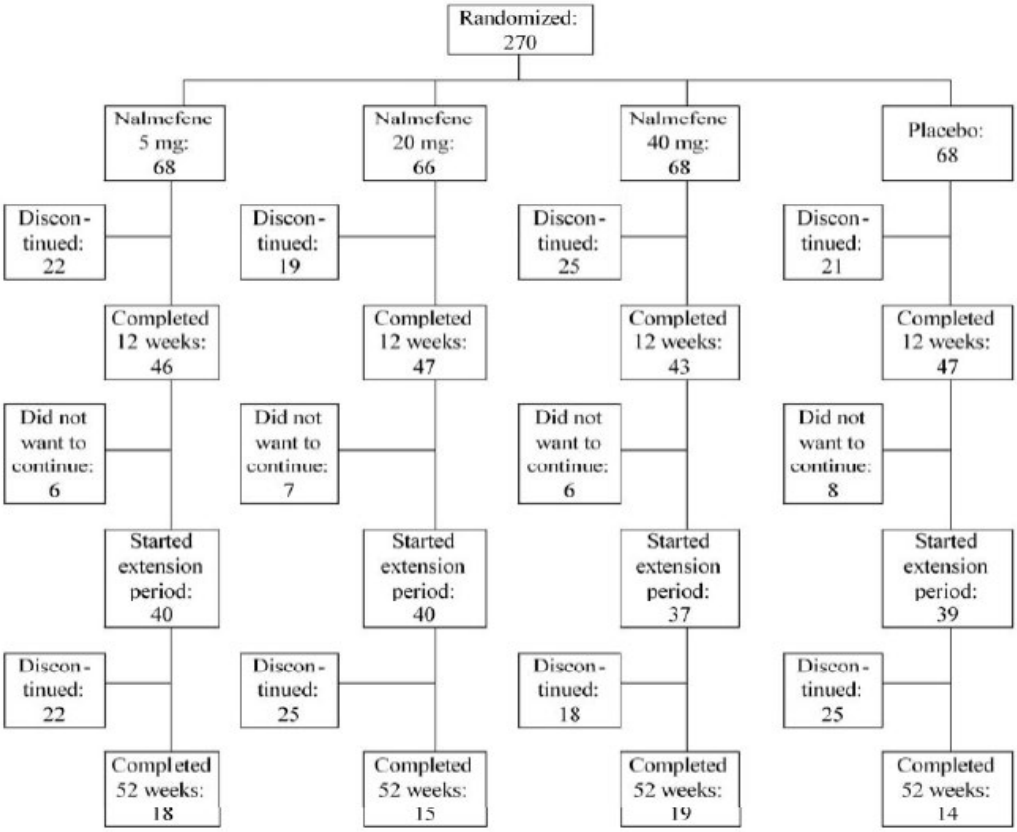
The primary efficacy analysis population was the intent-to-treat population.

A standard significance level of p = 0.05 was used, without any explicit correction for analysis of multiple endpoints.

7.3.2.9. Participant flow

Patient disposition is summarised in the figure below. Discontinuations were common, even in the first 12 weeks (19-25 subjects from 68 subjects in each group). Unlike many other nalmefene studies, the withdrawals were reasonably balanced across the treatment groups, making it less likely that the study was substantially compromised by withdrawal bias.

Figure 32. Disposition of Study Subjects – Biotie299.



7.3.2.10. Major protocol violations/deviations

The most common protocol deviations are summarised in the table below. These were not formally separated into major and minor violations. Protocol violations were common, but within expected bounds for a study of this nature.

Table 94. Most Common Protocol Deviations

Protocol deviation	Number of subjects
Visit not within time frame	206
Study medication blister cards not returned by subject	154
One or more measurements/assessments not done	150
Missing visit(s)	57
Disallowed concomitant medication	54
Vital signs not measured from same arm during the whole study	45
Study medication blister cards dispensed or used in wrong order	10

7.3.2.11. Baseline data

The four treatment groups were reasonably well-matched at baseline in terms of their basic demographics, family history of alcoholism and, most importantly, their actual alcohol habits prior to intervention, as shown in the tables below.

The mean monthly HDD count in the four groups was 20.8, 20.5, 19.6 and 19.6 in the placebo, 5mg, 20mg and 40mg groups, respectively.

Overall, the results of this study do not appear to have been unduly influenced by any significant mismatch at baseline.

Table 95. Summary of age, gender and BMI – Biotie299

	Placebo N=68	5 mg N=68	20 mg N=66	40 mg N=68	Total
Age, years	45.1 (11.1)	45.4 (11.3)	46.5 (10.9)	43.6 (9.1)	45.1 (10.6)
Gender	15 (22.1%) F 53 (77.9%) M	23 (33.8%) F 45 (66.2%) M	18 (27.3%) F 48 (72.7%) M	20 (29.4%) F 48 (70.6%) M	76 (28.1%) F 194 (71.9%) M
BMI, kg/m^2	28.8 (6.4)	27.9 (4.5)	27.1 (4.9)	27.1 (4.7)	27.7 (5.2)

Mean (with SD) is given for age and body mass index (BMI). F=female, M=male. Source: sections 14.1.1-2, 14.1.6.

Table 96. Summary of age, gender and BMI – Biotie299

	Placebo N=68	5 mg N=68	20 mg N=66	40 mg N=68	Total N=270
Father	28 (41.2)	33 (48.5)	32 (48.5)	38 (55.9)	131 (48.5)
Mother	11 (16.2)	12 (17.6)	17 (25.8)	16 (23.5)	56 (20.7)
Brother	24 (35.3)	17 (25.0)	23 (34.8)	23 (33.8)	87 (32.2)
Sister	11 (16.2)	8 (11.8)	11 (16.7)	8 (11.8)	38 (14.1)
Grandparents	22 (32.4)	23 (33.8)	23 (34.8)	25 (36.8)	93 (34.4)
Uncle	30 (44.1)	20 (29.4)	33 (50.0)	28 (41.2)	111 (41.1)
Aunt	8 (11.8)	9 (13.2)	12 (18.2)	19 (27.9)	48 (17.8)

Percentage given in the parentheses is calculated from all randomized subjects, whether family history was positive or negative. Source: sections 14.1.23-29.

Table 97. Baseline drinking – Biotie299

	Placebo N=68			5 mg N=68			20 mg N=66			40 mg N=68		
	Mean	SD	Median	Mean	SD	Median	Mean	SD	Median	Mean	SD	Median
HDD count	20.8	6.5	22.0	20.5	6.9	22.5	19.6	7.8	23.0	19.6	7.9	20.0
VHDD count	11.4	10.4	8.0	11.7	9.8	9.0	9.5	9.7	7.0	10.4	10.1	9.0
Abstinence day count	4.1	5.1	2.0	6.0	6.5	3.0	5.3	6.0	4.0	6.1	6.9	3.0
Drinks per week	63.9	41.6	52.0	61.6	37.8	54.4	55.5	34.2	48.5	54.3	31.5	44.7

N, number of subjects. Source: sections 14.1.43.1-4.

7.3.2.12. Results for the primary efficacy outcome

At baseline, the mean number of heavy drinking days was ~19-21 HDDs per month across the four treatment groups. During the 12 weeks of fixed daily dosing, the mean number of HDDs decreased by ~13-15 HDDs to approximately 5-8 HDDs per month. There were no important differences across the four treatment groups, and no statistical evidence of a treatment effect (p = 0.4564, main effect of treatment, Poisson regression). Over the first two months of treatment, the placebo group showed the greatest mean reduction in HDDs, but in the third month the active treatment groups showed a slightly greater mean effect; at all stages there was substantial overlap between all four treatment groups.

Figure 33. Number of HDDs by Month – Biotie299

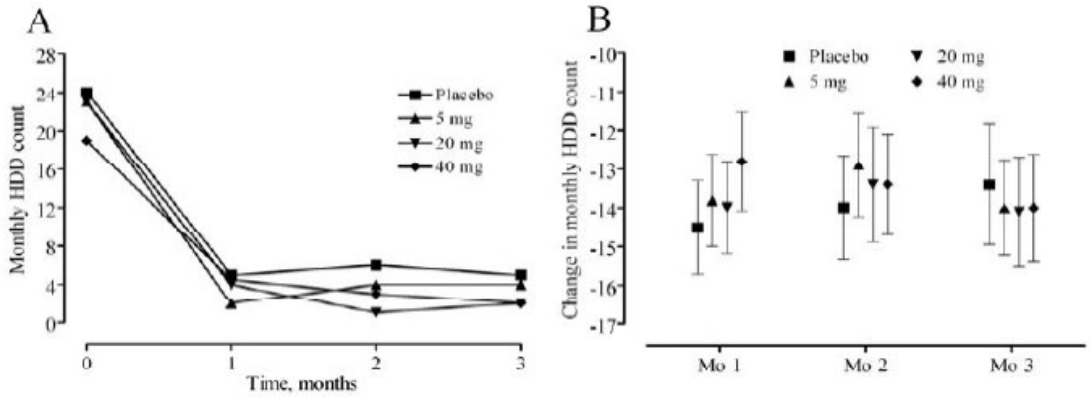


Table 98. Number of HDDs by Month – Biotie299

Group		Observed				Change from baseline		
		Screening	Mon 1	Mon 2	Mon 3	Mon 1	Mon 2	Mon 3
0	N	58	58	51	48	58	51	48
	Mean	21.2	6.7	7.4	8.1	-14.5	-14.0	-13.4
	SD	6.7	7.0	7.9	9.1	9.1	9.6	10.8
	Q1	16.0	0.0	0.0	0.0	-22.0	-22.0	-22.0
	Median	24.0	5.0	6.0	5.0	-16.5	-16.0	-15.5
	Q3	27.0	9.0	12.0	14.0	-5.0	-6.0	-5.5
5	N	61	61	56	53	61	56	53
	Mean	20.6	6.7	7.8	6.9	-13.8	-12.9	-14.0
	SD	7.1	8.6	9.0	8.4	9.3	10.0	9.0
	Q1	14.0	0.0	0.0	0.0	-23.0	-24.0	-20.0
	Median	23.0	2.0	4.0	4.0	-12.0	-11.5	-13.0
	Q3	27.0	11.0	14.5	11.0	-7.0	-4.0	-7.0
20	N	59	59	55	51	59	55	51
	Mean	19.5	5.6	5.9	5.0	-14.0	-13.4	-14.1
	SD	7.7	6.7	8.4	7.6	9.2	11.0	10.1
	Q1	12.0	0.0	0.0	0.0	-23.0	-23.0	-24.0
	Median	23.0	4.0	1.0	2.0	-14.0	-13.0	-14.0
	Q3	27.0	9.0	9.0	5.0	-7.0	-6.0	-7.0
40	N	59	58	54	47	58	54	47
	Mean	19.1	6.1	5.6	5.0	-12.8	-13.4	-14.0
	SD	8.1	7.1	6.7	7.0	9.7	9.5	9.5
	Q1	11.0	0.0	0.0	0.0	-22.0	-21.0	-22.0
	Median	19.0	4.5	3.0	2.0	-11.5	-13.0	-15.0
	Q3	27.0	10.0	8.0	8.0	-5.0	-5.0	-7.0

Source: Section 14.2.1

7.3.2.13. Results for other efficacy outcomes

As seen with the primary efficacy variable, no significant treatment effect was observed for various other measures of alcohol intake over the first 12 weeks of treatment, including monthly number of very heavy drinking days, number of abstinence days, proportion of heavy drinking days of all drinking days, mean weekly alcohol consumption, maximum intensity of drinking, and average intensity of drinking. Graphs for some of these variables are shown in the figures below; these indicate substantial overlap between groups with no consistent trend in favour of active treatment. Tables showing quantitative values for these variables did not add any useful insights; these can be obtained from the original Study Report.

Figure 34. Very Heavy Drinking Days During Initial Period – Biotie299

A. Median monthly counts. B. Mean change from baseline. Error bars are standard error of the mean. Source: section 14.2.3.

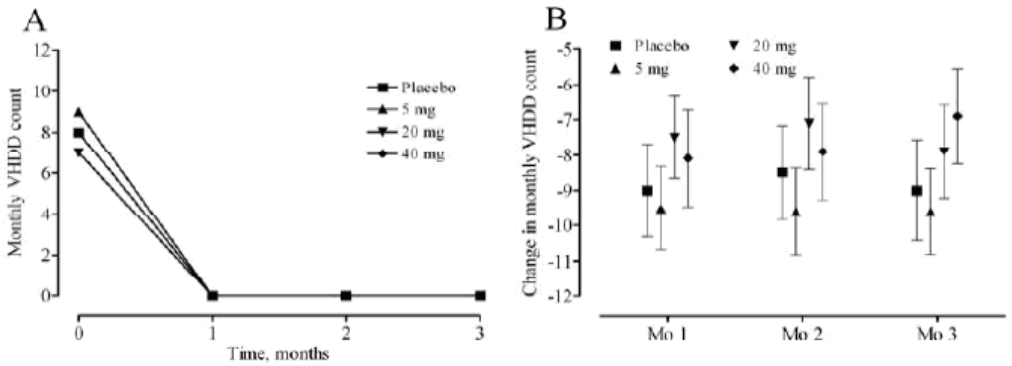


Table 99. Abstinence Days During Initial Period – Biotie299

A. Monthly median. B. Mean change from baseline with standard error. Source: section 14.2.2.

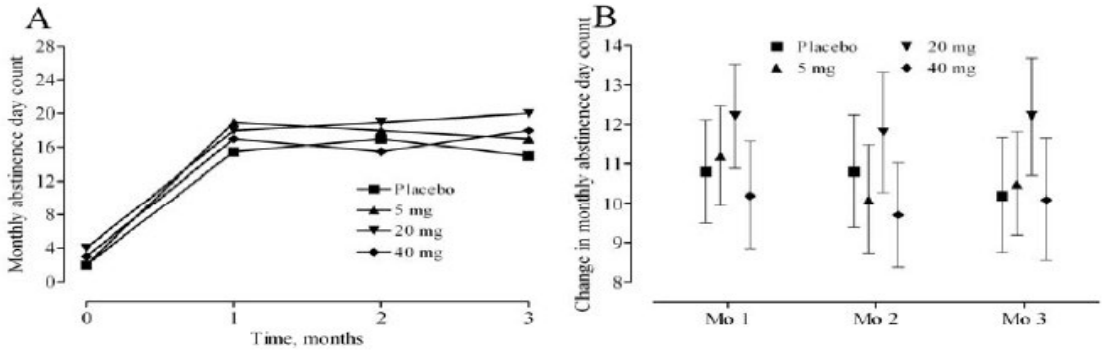
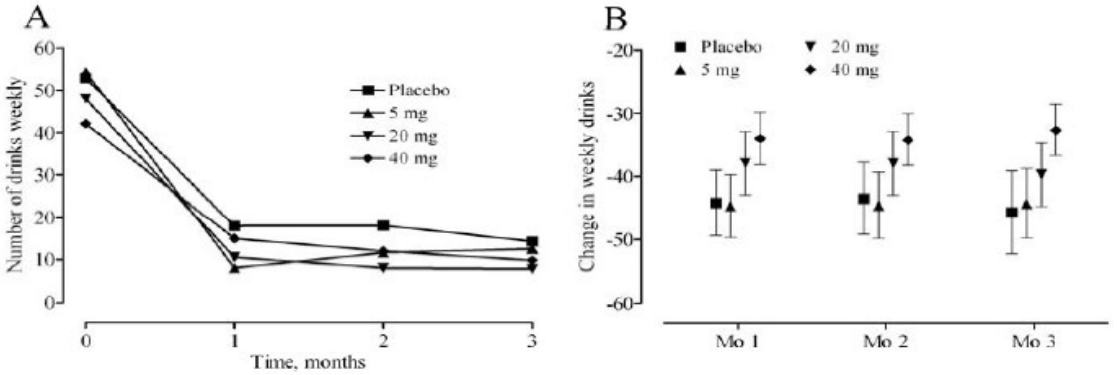


Figure 35. TAC During Initial Period – Biotie299

A. Median number of drinks weekly by group. B. Mean change from baseline in number of drinks weekly. Error bars are standard error of the mean. Source: section 14.2.4.



The surrogate efficacy measures derived from laboratory tests showed some evidence of reduced alcohol intake in all treatment groups, with reductions in mean serum gamma-glutamyl transferase and carbohydrate-deficient transferrin over 12 weeks, but there was no discernable between-group difference.

Table 100. Serum CDT, Initial Period – Biotie299

	Placebo						5 mg						20 mg						40 mg					
	N	Mean	SD	Q1	Median	Q3	N	Mean	SD	Q1	Median	Q3	N	Mean	SD	Q1	Median	Q3	N	Mean	SD	Q1	Median	Q3
Screening	67	3.1	1.2	2.2	2.8	3.9	68	3.7	2.4	2.2	2.9	4.1	64	3.6	1.9	2.3	3.0	3.8	67	3.7	2.4	2.3	2.9	4.1
Week 6	48	2.6	0.9	2.1	2.4	3.0	50	2.7	1.2	2.0	2.4	2.9	46	2.8	1.1	2.2	2.4	3.2	51	3.2	2.4	2.1	2.4	3.7
Week 12	47	2.7	0.9	2.0	2.4	3.2	47	3.2	2.4	2.0	2.4	3.1	47	3.2	1.8	2.2	2.6	3.2	45	3.2	2.0	2.2	2.4	3.5

Values are % of total transferrin. The normal values are <2.6 %. Source: section 14.2.9

Table 101. Serum GGT, Initial Period – Biotie299

	Placebo						5 mg						20 mg						40 mg					
	N	Mean	SD	Q1	Median	Q3	N	Mean	SD	Q1	Median	Q3	N	Mean	SD	Q1	Median	Q3	N	Mean	SD	Q1	Median	Q3
Screen	68	73.6	72.7	25	53	85	68	73.5	80.9	26	47.5	87	64	75.7	82.0	29	44	77	68	86.0	135	27	44	85
Week 6	50	64.0	61.1	21.0	40	82	51	46.9	42.2	22	34	62	47	55.2	72.5	21	29	54	52	62.3	56.8	24	39	83
Week 12	50	58.4	64	20.0	37	69	50	56.7	84.8	21	29	63	47	44.4	46.3	22	31	45	49	46.6	43.4	23	33	52

Values are U l^{-1}.Reference range: Male: 0-55 (13-64 y), Female: 0-45; 65+y (0-75). Source: Section 14.2.9

The psychosocial cost of drinking was also assessed by a number of measures (alcohol dependence scale, the drinker inventory of consequences, obsessive compulsive drinking scale, and collateral estimate from spouse or partner). These assessments showed some favourable changes during the 12 weeks of treatment, but the changes were similar across all treatment groups (data not shown).

7.3.2.14. Subgroup analysis

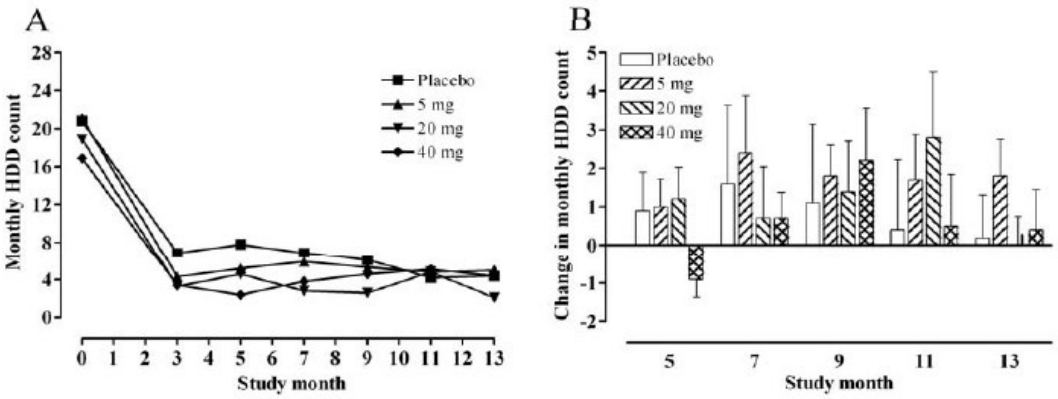
No subgroup analyses were performed.

7.3.2.15. Results in the as-needed dosing phase

Results in the flexible dosing phase were similarly disappointing. For the main efficacy variable of HDDs, the mean count remained stable over the subsequent 40 weeks, close to the level achieved at the end of the initial 12-week period. There were no consistent trends in favour of active treatment or significant differences between groups (p = 0.9579, main effect of treatment; Poisson regression).

Table 102. Heavy Drinking During Extension Period – Biotie299

A. Monthly mean HDD count in the extension population. B. Changes in the monthly HDD count since month 3. Error bars are standard error of the mean. Source: sections 15.2.2.1 and 15.2.3.1.



For various other measures of alcohol intake, as shown in the figures below, there was a similar lack of any important differences between groups. None of the between-group comparisons approached statistical significance.

Table 103. Very Heavy Drinking During Extension Period – Biotie299

A. Monthly mean VHDD count in the extension population. B. Change in the monthly VHDD count with standard error since end of daily dosing. Source: sections 15.2.2.3 and 15.2.3.3.

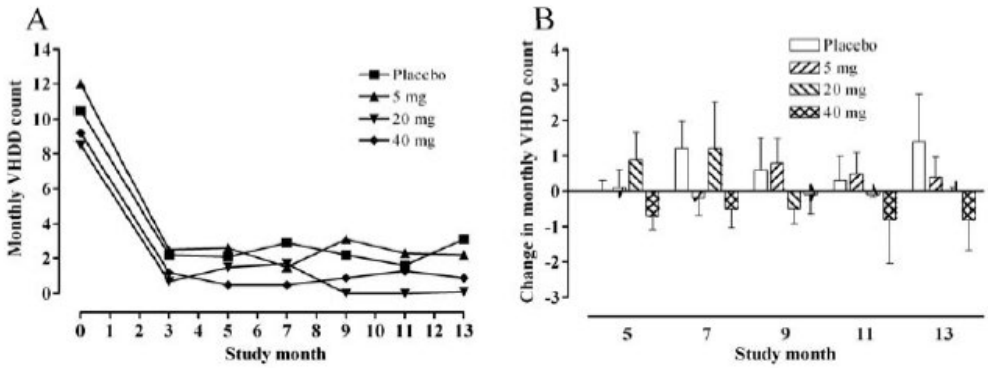


Table 104. Abstinence Days During Extension Period – Biotie299

A. Monthly mean abstinence day count in the extension population. B. Change in the monthly abstinence day count since end of daily dosing. Error bars are standard error of the mean. Source: sections 15.2.2.2 and 15.2.3.2.

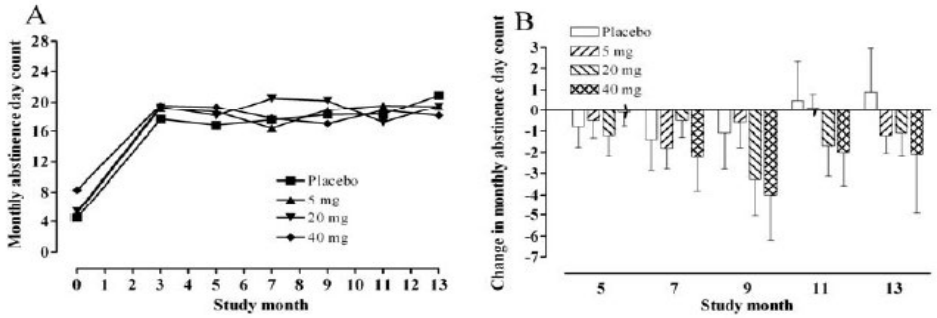
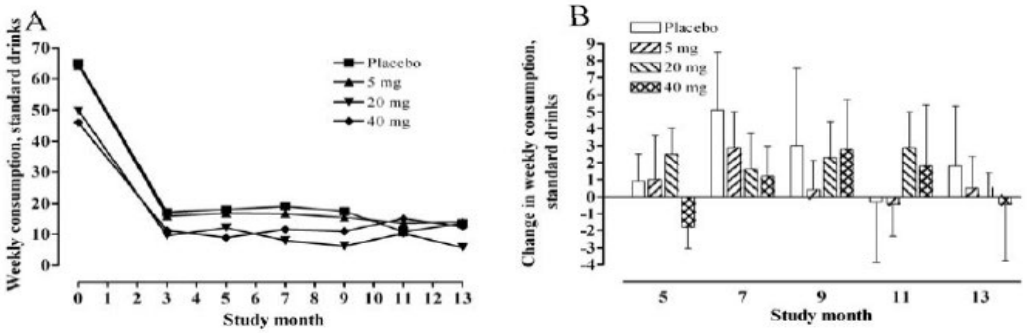


Figure 36. TAC During Extension Period – Biotie299

A. Weekly mean alcohol consumption in the extension population. B. Change in the consumption since month 3. Error bars are standard error of the mean. Source: sections 15.2.2.4 and 15.2.3.4.



There was also no statistically or clinically significant difference between the groups in the psychosocial measures during the flexible dosing phase (data not shown).

7.3.3. Supportive study biotie399 (CPH-101-0399)

7.3.3.1. Study design, objectives, locations and dates

This 16-week study was a randomised, placebo-controlled, double-blind, parallel-group study that assessed the efficacy of nalmefene at two different doses (10mg or 40mg) in the treatment of subjects with heavy alcohol consumption and impaired control of their intake. This study did not assess the proposed 20mg dose of nalmefene.

The primary objective was to assess the efficacy and dose-response of nalmefene in reducing heavy alcohol drinking, with safety assessment as a secondary objective.

The study was conducted in 6 centres in Finland, from 3rd January 2000 to 28th July 2000.

7.3.3.2. Inclusion and exclusion criteria

The inclusion criteria were broadly similar to other submitted studies. The study sought patients over 18 years with heavy alcohol use and impaired control, without significant comorbidities.

In particular, subjects were eligible if they had:

- a desire to cut down or control drinking
- difficulty in controlling drinking (taking alcohol in larger amounts or for a longer than was intended)

- ≥ 8 heavy drinking days within the last month prior screening (HDD defined as ≥ 5 standard drinks for males, ≥ 4 for female)

The main exclusion criteria were:

- significant psychiatric or somatic co-morbidity
- abuse or dependence on other substances

7.3.3.3. Study treatments

This study differed from the pivotal studies in that it did not assess the proposed 20mg dose, it did not employ as-needed dosing, and it did not include formal psychosocial supportive measures.

The target doses were 0mg (placebo), 10mg or 40mg per day. Subjects achieved this by taking two tablets once daily (placebo, 5mg or 20mg tablets).

During the first week of the study, to promote tolerance and to reduce the risk of un-blinding, study drug was introduced at a lower dose: all nalmefene subjects took nalmefene 5mg, and all placebo subjects took a single placebo tablet.

7.3.3.4. Efficacy variables and outcomes

The primary efficacy variable was the number of heavy drinking days (HDDs) per month, as in all of the other Biotie studies in alcohol-use disorders.

- Secondary efficacy variables included:
- Number of very heavy drinking days (VHDDs)
- Number of abstinence days (NDDs)
- Ratio of heavy drinking days to drinking days
- Total alcohol consumption (TAC)
- Maximum and average intensity drinking
- Obsessive Compulsive Drinking Scale (OCDS)
- Alcohol Dependence Scale (ADS)
- Drinkers Inventory of Consequences-scale (DrInC)
- Clinical laboratory markers (ALAT, GGT, CDT, MCV)
- Drinking pattern assessed the subject's spouse or significant other (Spouses were asked to respond to the question, 'How often has your significant other been drinking?' using the scale 'Not at all', 'Drank on just a few days', 'Drank on several days/month', 'Drank at least once a week', 'Drank every day')
- Discontinuation of the study

The primary efficacy endpoint was defined as the overall difference in monthly HDDs for the 16-week treatment phase, analysed by RM-ANOVA.

7.3.3.5. Randomisation and blinding methods

Subjects were randomised equally to the three dose groups, using a centralised computer-based approach. Subjects were randomised in blocks of six, and treatment was allocated to centres in blocks to ensure balance at each centre across treatment groups.

Blinding was ostensibly maintained by storing the randomisation codes centrally and by using tablets that appeared identical in all three treatment groups.

7.3.3.6. Analysis populations

All subjects were in the intent-to-treat (ITT) population. No other efficacy subset was defined.

7.3.3.7. Sample size

The Sponsor used an approximation method for sample size estimations, because power computations are not readily available for non-linear mixed-effects regression models. The Sponsor treated the weekly results from the planned 16-week assessment as a cluster of partially correlated results, with an intra-subject correlation of r = 0.5. They also assumed that subjects would exhibit heavy drinking on about 70% of days. Assuming a two-sided comparison with a standard significance level of 0.05, it was estimated that 40 patents would need to complete assessment in each group to provide 80% power to detect a difference of 35% in the proportions of HDDs vs non-HDDs per month. Allowing for a dropout rate of 20%, 50 subjects in group (150 in total) would need to be recruited.

This study was positive for its primary endpoint, suggesting that these assumptions were broadly appropriate.

7.3.3.8. Statistical methods

Statistical analysis was done with SAS software (v8.1). The primary statistical test, used on the primary efficacy variable and most secondary variables, was RM-ANOVA, looking at the significance of the treatment-group-by-time interaction, using a standard threshold of $p < 0.05$.

Treatment group, gender and centre were used as independent variables. In nonparametric tests only treatment group was used as an explanatory variable. Categorical baseline variables were analysed with generalised logit functions, whereas continuous variables were analysed with analysis of variance. Normality assumptions were checked visually, with Shapiro-Wilk tests were used as supportive methods.

Corrections for the use of multiple different endpoints were not employed.

No secondary endpoint can be considered a key endpoint, and there was no formal ranking of secondary endpoints, so a hierarchical testing procedure cannot be applied, even in retrospect. Given that the secondary endpoints contained a mix of positive and negative results, and that the positive results were generally of only moderate significance ($0.01 < p < 0.05$); these secondary endpoints should basically be seen as weakly supportive.

7.3.3.9. Participant flow

Patient disposition is summarised in the table below. A total of 150 subjects were enrolled (50 per group) and all entered the ITT analysis, but some discontinued medication for reasons shown below. The total number of withdrawals was similar in each group (9, 11 and 10 in the placebo, 10mg and 40mg groups, respectively).

Table 105. Reasons for Discontinuation of Study Medication

Reason	*Placebo*	*NMF 10*	*NMF 40*	*All*
N	**9**	**11**	**10**	**30**
Adverse event	3	6	6	15
Lack of efficacy/clinical deterioration	3	1	1	5
Lost to follow up	1	2	1	4
Consent withdrawal	1	2	2	5
Other	1	-	-	1

7.3.3.10. Major protocol violations/deviations

A clear distinction between major and minor protocol deviations was not specified. Major protocol deviations were not common, but 12 subjects were enrolled despite not fully meeting the eligibility criteria. For 8 of these subjects, the TLFB showed they did not have ≥ 8 HDDs within the previous month (two in the placebo group, three in nalmefene 10mg group and three in nalmefene 40mg group). For 8 female subjects, enrolment was inappropriate because they had childbearing potential and were not using reliable contraception.

Also, 34 subjects missed tablets at some stage, 10 subjects took their total daily dose in two divided doses, and 12 took a reduced dose. Five subjects received opioids during the study, 3 in the setting of a serious adverse event. Occasional assessments were mistimed or missing.

Overall, these protocol deviations are within the expected limits for a study of this nature and they are unlikely to have had a major impact on the results.

7.3.3.11. Baseline data

Baseline data is summarised in the tables below, including basic demographics, social situation, drinking history and baseline drinking pattern. There were minor differences between groups, but no major mismatches that are likely to have affected the efficacy analysis. The mean number of HDDs in each group at baseline was 14.3, 15.2 and 14.9 in the placebo, 10mg and 40mg groups, respectively.

Table 106. Age (years) by Group and Gender – Biotie399

	Placebo	*NMF 10*	*NMF 40*
Male			
n	39	39	44
Mean	49.1	49.0	50.4
SD	8.8	8.0	9.0
Median	49.0	47.0	50.5
Q1 to Q3	43.0 to 55.0	43.0 to 57.0	42.0 to 57.0
Min – Max	29 – 66	34 – 63	32 – 67
Female			
n	11	11	6
Mean	44.3	50.8	49.0
SD	10.5	7.5	5.0
Median	43.0	51.0	50.0
Q1 to Q3	36.0 to 49.0	44.0 to 58.0	46.0 to 53.0
Min – Max	29 – 64	41 – 63	41 – 54

Table 107. Social History by Group – Biotie399

	Placebo	*NMF 10*	*NMF 40*
n	50	50	50
Education			
Secondary school or less	7(14%)	9 (18%)	16 (32%)
Vocational school	16 (32%)	15 (30%)	18 (36%)
Upper secondary school graduate	12 (24%)	8 (16%)	3 (6%)
College	12 (24%)	11 (22%)	12 (24%)
University graduate	3 (6%)	7 (14%)	1 (2%)
Current Employment Status			
Student	-	1 (2%)	-
Employed, Full time	29 (58%)	23 (46%)	24 (48%)
Employed, Part time	1 (2%)	2 (4%)	2 (4%)
Unemployed	10 (20%)	9 (18%)	10 (20%)
Retired	10 (20%)	13 (26%)	14 (28%)
Other	-	2 (4%)	-
Marital Status			
Never married	7 (14%)	6 (12%)	9 (18%)
Married	23 (46%)	29 (58%)	29 (58%)
Separated or divorced	18 (36%)	14 (28%)	10 (20%)
Widow (er)	2 (4%)	1 (2%)	2 (4%)
Living Arrangement			
Alone	9 (18%)	14 (28%)	16 (32%)
With spouse or partner	17 (34%)	18 (36%)	16 (32%)
With spouse or partner and children	18 (36%)	17 (34%)	17 (34%)
Single with children	4 (8%)	-	1 (2%)
Other	2 (4%)	1 (2%)	-

Table 108. Alcohol Drinking History by Group – Biotie399

	Placebo	*NMF 10*	*NMF 40*
n	50	50	50
History of Alcohol Problems in Family			
Father	16 (32%)	22 (44%)	18 (36%)
Mother	1 (2%)	2 (4%)	3 (6%)
Grandparents	7 (14%)	5 (10%)	4 (8%)
Siblings	13 (26%)	6 (12%)	9 (18%)
Uncles	13 (26%)	11 (22%)	14 (28%)
Aunts	-	1 (2%)	-
Total	32 (64%)	32 (64%)	31 (62%)
Age at First Drinking Experience			
Mean	16.5	17.2	17.0
SD	4.8	4.8	3.6
Median	16.0	16.0	16.5
Q1 to Q3	15.0 to 17.0	15.0 to 18.0	15.0 to 18.0
Min – Max	10.0 – 45.0	8.0 – 30.0	11.0 – 29.0
Age at the Onset of Problem Drinking			
Mean	36.8	37.3	36.0
SD	10.9	10.6	11.4
Median	38.0	35.0	35.0
Q1 to Q3	30.0 to 44.0	30.0 to 45.0	27.0 to 42.0
Min – Max	16.0 – 66.0	12.0 - 61.0	18 – 60.0

Table 109. Previous Alcohol Treatment – Biotie399

	Placebo	*NMF 10*	*NMF 40*
Any Previous Treatment for Alcohol Problems	18 (36%)	27 (54%)	22 (44%)
Previous Detoxifications	6 (12%)	9 (18%)	10 (20%)
Previous Attendance at AA/Self-help groups	18 (36%)	27 (54%)	22 (44%)

Table 110. Baseline Drinking – Biotie399

	Placebo	*NMF 10*	*NMF 40*
N	50	50	50
Number of HDDs			
Mean	14.3	15.2	14.9
SD	6.5	7.3	6.7
Median	12.2	13.0	12.5
Q1 to Q3	9.3 to 18.3	9.3 to 22.3	10.0 to 19.3
Min – Max	5.3 – 28	3.0 – 28	6.7 – 28
Number of VHDDs			
Mean	6.7	8.2	7.5
SD	5.2	6.9	6.6
Median	6.0	6.3	7.2
Q1 to Q3	3.0 to 9.0	3.7 to 10.3	2.3 to 10.3
Min – Max	0.0 – 24.7	0.0 – 27.7	0.0 – 27.7
Number of Abstinence days			
Mean	10.0	10.1	9.7
SD	6.6	7.2	7.2
Median	11.2	10.0	9.3
Q1 to Q3	3.3 to 16.0	3.3 to 16.3	2.3 to 16.3
Min – Max	0.0 – 22.0	0.0 – 24.0	0.0 – 21.3
Total Consumption (drinks/week)			
Mean	38.6	42.0	38.9
SD	18.4	23.6	18.3
Median	36.4	35.8	34.7
Q1 to Q3	25.1 to 50.2	23.9 to 60.8	25.3 to 48.8
Min – Max	14.7 – 92.7	6.5 – 104.2	15.0 - 89.9

7.3.3.12. Results for the primary efficacy outcome

At the end of four months, the reduction in HDDs was greatest within the nalmefene 40mg group, intermediate in the placebo group, and least in the nalmefene 10mg group. The reduction in the 40mg group appeared in the first month and persisted across the study, whereas the placebo group and the 10mg group showed smaller reductions and changed relative ranking in HDDs through the study.

Statistical analysis of all three groups revealed a significant treatment group time interaction ($p = 0.015$), which primarily reflects the difference between the nalmefene 40mg group and the placebo group. Pairwise comparisons revealed a significant treatment group time interaction between placebo and NMF 40 groups ($p = 0.01$), but a non-significant treatment group time interaction between placebo and 10 mg ($p = 0.063$) and between 10 mg and 40 mg ($p = 0.14$).

The spread of HDDs by the end of the study period was small, and examination of individual months shows that the significant difference between the nalmefene 40mg and placebo groups primarily reflects the first two months in the study, when reductions in the nalmefene 40mg group occurred rapidly and reductions in the placebo group had not yet occurred. The treatment group time interactions were significant for Month 0 versus Month 1 and for Month 0 versus Month 2 ($p = 0.0096$ and $p = 0.018$, respectively), but not for Month 0 versus Month 3 or for Month 0 versus Month 4 ($p = 0.65$ and $p = 0.24$, respectively).

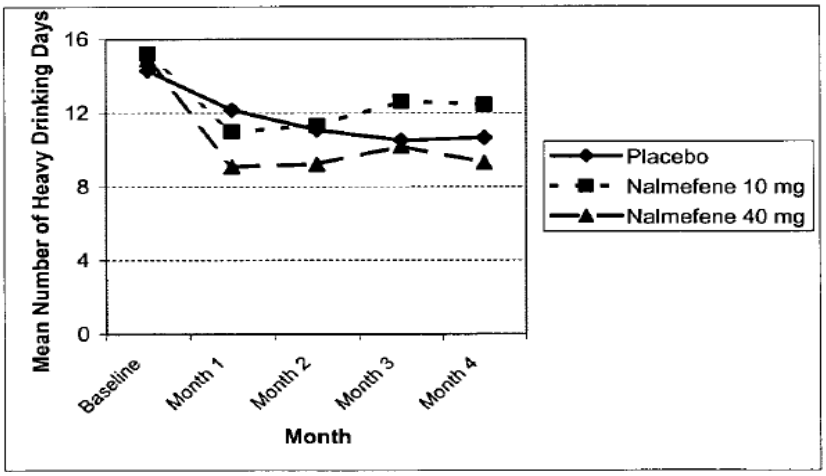
In the context of the Lundbeck studies, which tended to show a treatment difference emerging later during treatment (only reaching significance by Month 7 in the one-year Lundbeck Lundbeck13), this apparent early superiority of nalmefene is inconsistent and of uncertain importance. One difference between this study and the later Lundbeck studies is that the Lundbeck studies included psychosocial measures (BRENDA), which may have helped placebo recipients to achieve an early reduction in alcohol intake. The difference in the apparent timing of the therapeutic effect might also arise from the higher dose used in this study, (which would imply that the dose in the later Lundbeck studies was inadequate). Cultural differences might

also have played a role, as all centres in this study were from Finland, compared to the pivotal studies, which were international.

Table 111. Number of HDDs by Group and Month – Biotie399

	Baseline	Month 1	Month 2	Month 3	Month 4
Placebo					
N	50	48	45	45	45
Mean	14.3	12.1	11.1	10.5	10.6
SD	6.5	7.7	7.1	7.3	8.5
Median	12.2	10.5	10.0	10.0	9.0
Q1 to Q3	9.3 to 18.3	7.5 to 16.5	6.0 to 16.0	6.0 to 15.0	4.0 to 17.0
Min – Max	5.3 – 28	0.0 – 28	0.0 – 28	0.0 – 28	0.0 – 28
NMF 10					
N	50	50	48	46	45
Mean	15.2	11.0	11.3	12.6	12.4
SD	7.3	8.3	7.7	7.8	7.7
Median	13.0	9.5	9.0	12.0	12.0
Q1 to Q3	9.3 to 22.3	4.0 to 16.0	5.0 to 17.5	6.0 to 17.0	6.0 to 17.0
Min – Max	3.0 – 28	0.0 – 28	0.0 – 28	0.0 – 28	0.0 – 28
NMF 40					
N	50	48	48	47	46
Mean	14.9	9.1	9.2	10.2	9.3
SD	6.7	7.4	8.2	7.5	8.5
Median	12.5	7.5	8.0	9.0	8.5
Q1 to Q3	10.0 to 19.3	4.0 to 13.0	2.0 to 13.0	5.0 to 14.0	1.0 to 14.0
Min – Max	6.7 – 28	0.0 – 28	0.0 – 28	0.0 – 28	0.0 – 28

Figure 37. Mean Number of HDDs by Group – Biotie399



In assessing the clinical relevance of the treatment effect, only the long-term effects are important: a treatment that produced a benefit over placebo just for the first two months would not be of much clinical utility because most complications of alcoholism take years to become manifest. The magnitude of the benefit observed at the end of Month 4 was small: subjects taking nalmefene 40mg/d had a mean of 9.3 HDDs per month, compared to 10.6 HDDs in the placebo group, a difference of 1.3 days.

7.3.3.13. Results for other efficacy outcomes

The number of VHDDS is shown in the table below. There were improvements in both groups, but there was no significant treatment group time interaction (p = 0.18).

Table 112. Number of Very Heavy Drinking Days by Group and Month – Biotie399

	Baseline	Month 1	Month 2	Month 3	Month 4
Placebo					
N	50	48	45	45	45
Mean	6.7	5.3	4.8	4.7	4.3
SD	5.2	5.0	5.0	5.3	5.3
Median	6.0	4.0	4.0	4.0	3.0
Q1 to Q3	3.0 to 9.0	2.0 to 8.0	1.0 to 7.0	0.0 to 7.0	0.0 to 7.0
Min – Max	0 – 24.7	0 – 28	0 – 24	0 – 28	0 – 27
NMF 10					
N	50	50	48	46	45
Mean	8.2	5.9	5.3	5.6	4.9
SD	6.9	7.1	5.4	5.2	5.0
Median	6.3	3.0	3.0	4.0	3.0
Q1 to Q3	3.7 to 10.3	1.0 to 10.0	1.5 to 8.0	1.0 to 9.0	2.0 to 8.0
Min – Max	0 – 27.7	0 – 25	0 – 22	0 – 22	0 – 22
NMF 40					
N	50	48	48	47	46
Mean	7.5	3.8	3.9	4.5	3.5
SD	6.6	4.7	5.8	5.6	5.8
Median	7.2	2.0	1.0	3.0	1.0
Q1 to Q3	2.3 to 10.3	0.0 to 6.5	0.0 to 5.0	0.0 to 6.0	0.0 to 6.0
Min – Max	0 – 27.7	0 – 24	0 – 28	0 – 28	0 – 28

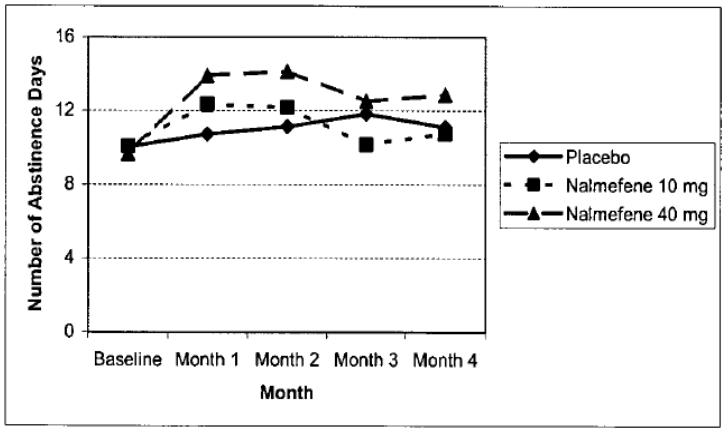
The number of NNDs (days of abstinence) increased in all treatment groups, and there was a significant difference across the groups (group time interaction, p = 0.016). Pairwise comparisons between groups showed significant differences between placebo and 10 mg (treatment group time: p = 0.023), between placebo and 40 mg (treatment group time: p=0.0089) but not between 10 mg and 40 mg (treatment group time: p = 0.43). This provides moderate support for the overall efficacy analysis, but there was no correction for the use of multiple endpoints and several other secondary endpoints were negative.

The mean number of days of abstinence at Month 4 was only slightly better in the 40mg group than the placebo group, with 1.8 extra days per month spent abstinent in the 40mg group compared to the placebo group (placebo 11.1 NDDs, nalmefene 10mg 10.8 NDDs, nalmefene 40mg 12.9 NDDs).

Table 113. Number of Abstinence Days by Group and Month – Biotie399

	Baseline	Month 1	Month 2	Month 3	Month 4
Placebo					
n	50	48	45	45	45
Mean	10.0	10.7	11.1	11.8	11.1
SD	6.6	7.9	7.7	7.9	8.8
Median	11.2	10.0	11.0	12.0	11.0
Q1 to Q3	3.3 to 16.0	5.0 to 16.0	4.0 to 17.0	5.0 to 17.0	2.0 to 16.0
Min – Max	0 – 22.0	0 – 28	0 – 28	0 – 28	0 – 28
NMF 10					
n	50	50	48	46	45
Mean	10.1	12.4	12.2	10.2	10.8
SD	7.2	8.0	7.5	7.5	7.4
Median	10.0	13.0	12.5	10.0	10.0
Q1 to Q3	3.3 to 16.3	6.0 to 18.0	6.0 to 19.0	4.0 to 17.0	5.0 to 17.0
Min – Max	0 – 24.0	0 – 27	0 – 27	0 – 28	0 – 28
NMF 40					
n	50	48	48	47	46
Mean	9.7	13.9	14.1	12.5	12.9
SD	7.2	8.0	8.7	8.2	8.9
Median	9.3	14.0	13.5	12.0	14.0
Q1 to Q3	2.3 to 16.3	9.0 to 20.0	8.0 to 21.5	6.0 to 19.0	3.0 to 20.0
Min – Max	0 – 21.3	0 – 28	0 – 28	0 – 28	0 – 28

Figure 38. Mean Number of Abstinence Days by Month and Group – Biotie399



The mean weekly alcohol consumption is shown in the table below. Although there was a decrease in all groups, the difference between groups was not significant (group time interaction, p = 0.14). There was a trend in favour of nalmefene 40mg. In the 40mg group, the mean reduction by Month 4 was 14.3 g/week, compared to 11.7 g/week in the 10mg group and 9.4 in the placebo group g/week, from baselines of ~39-42 g/week.

Table 114. Mean Weekly Alcohol Consumption and Changes from Baseline by Month

	Baseline	Month 1	Month 2	Month 3	Month 4
Placebo					
N	50	48	45	45	45
Mean (SD)	38.6 (18.4)	32.4 (21.2)	30.6 (20.9)	29.7 (20.5)	28.6 (22.2)
Change from baseline		-6.2 (17.3)	-7.4 (18.1)	-8.4 (20.0)	-9.4 (25.7)
% change from baseline		-12.9 (42.9)	-14.4 (44.2)	-16.1 (52.3)	-14.9 (62.4)
NMF 10					
N	50	50	48	46	45
Mean (SD)	42 (23.6)	30.9 (24.4)	31.0 (22.2)	32. (19.6)	29.8 (18.3)
Change from baseline		-11.1 (19.9)	1.0 (5.8) -10.8 (20.1)	-9.1 (18.9)	-11.7 (17.6)
% change from baseline		-23.2 (46.1)	-21.4 (42.4)	-13.7 (50.1)	-20.1 (50.0)
NMF 40					
N	50	48	48	47	46
Mean (SD)	38.9 (18.3)	24.6 (17.3)	24.1 (20.6)	27.5 (19.0)	24.2 (20.1)
Change from baseline		-13.6 (18.4)	-14.1 (17.8)	-10.7 (17.8)	-14.3 (19.6)
% change from baseline		-31.8 (43.8)	-37.6 (46.4)	-26.2 (48.0)	-35.3 (55.7)

Laboratory monitoring did not reveal significant differences between groups, as shown in the tables below. For changes in MCV, the Kruskal-Wallis test approached significance (p = 0.052), but the observed changes were clinically trivial. For ALAT and GGT, there were no important differences between groups, and no strong trends emerged.

Table 115. MCV at Screening, Week 6 and Week 16 – Biotie399

	Screening	Week 6	Change from screening to Week 6	Week 16	Change from screening to Week 16
Placebo					
N	50	45	45	44	44
Median	96.0	96.0	0	96.5	0
Q1 to Q3	93 to 99	93 to 98	-1 to 0	93 to 99	-1 to 1
Min,Max	78, 105	75, 105	-7, 3	76, 108	-9, 4
NMF 10					
N	50	47	47	42	42
Median	97.0	97.0	0	97.5	0
Q1 to Q3	93 to 100	92 to 100	-1 to 1	91 to 100	-1, 1
Min,Max	86, 108	85, 107	-3, 5	86, 107	-4, 5
NMF 40					
N	50	47	47	45	45
Median	96	96	-1	95	-1
Q1 to Q3	93 to 98	92 to 98	-2 to 0	93 to 97	-2 to 0
Min,Max	86, 103	85, 103	-3, 2	85, 102	-6, 2
Kruskal-Wallis test p	0.55	0.21	0.19	0.27	0.052

Table 116. GGT (U/L) at Screening, Week 6 and Week 16 – Biotie399

	Screening	Week 6	Change from screening to Week 6	Week 16	Change from screening to Week 16
Placebo					
N	50	45	45	44	44
Median	59.0	42.0	-6	53.0	-3
Q1 to Q3	31 to 108	28 to 100	-17 to 2	30 to 121.5	-14 to 9
Min,Max	7, 433	8, 443	-195, 206	7, 307	-176, 152
NMF 10	50	47	47	42	42
Median	65.5	51.0	-5	58.0	-6.5
Q1 to Q3	40 to 115	37 to 107	-18 to 5	35 to 122	-34 to 3
Min-Max	14, 372	12, 677	-240, 330	10, 412	-242, 92
NMF 40	50	47	47	45	45
Median	81.5	45.0	-14	55.0	-13
Q1 to Q3	42 to 139	29 to 102	-48 to -4	30 to 103	-24 to –1
Min-Max	18, 908	13, 519	-524, 65	15, 622	-291, 117
Kruskal-Wallis test p	0.14	0.48	0.035	0.90	0.14

Table 117. ALT (U/L) at Screening, Week 6 and Week 16 – Biotie399

	Screening	Week 6	Change from screening to Week 6	Week 16	Change from screening to Week 16
Placebo					
N	50	45	45	44	44
Median	30.0	26.0	1	27.5	0
Q1 to Q3	20 to 50	18 to 42	-6 to 7	18.5 to 55	-7 to 7.5
Min,Max	9, 175	11, 194	-86, 94	5, 245	-85, 145
NMF 10					
N	50	47	47	42	42
Median	32.0	29.0	-2	29.5	-4.5
Q1 to Q3	23 to 66	25 to 58	-9 to 5	22 to 60	-12 to 4
Min,Max	8, 174	7, 208	-104, 108	6, 159	-130, 44
NMF 40					
N	50	47	47	45	45
Median	39.5	30.0	-4	34.0	-6
Q1 to Q3	26 to 55	22 to 44	-19 to 5	24 to 48	-14 to 6
Min,Max	15, 162	9, 790	-103, 750	12, 250	-45, 88
Kruskal-Wallis test p	0.098	0.35	0.17	0.65	0.34

Table 118. CDT (U/L) at Screening, Week 6 and Week 16 – Biotie399

	Screening	Week 6	Change from screening to Week 6	Week 16	Change from screening to Week 16
Placebo					
N	50	45	45	44	44
Median	17.0	19.0	2	20.0	2.5
Q1 to Q3	12 to 22	15 to 25	-2 to 6	15 to 24	0 to 6.0
Min,Max	8, 55	9, 47	-29, 23	10, 56	-15, 20
NMF 10					
N	50	46	46	42	42
Median	16.5	16.5	1	19.0	2.5
Q1 to Q3	12 to 27	13 to 23	-4 to 4	15 to 31	0 to 6
Min,Max	6, 76	6, 63	-33, 17	10, 83	-45, 54
NMF 40					
N	50	47	47	45	45
Median	16.5	17.0	1	18.0	3
Q1 to Q3	12 to 24	13 to 20	-2 to 3	16 to 25	1 to 6
Min,Max	7, 37	8, 48	-12, 17	9, 49	-16 ,15
Kruskal-Wallis test					
p	0.96	0.62	0.38	0.93	0.95

Psychosocial assessments also failed to show a treatment effect. Mean ADS scores decreased in all the treatment groups, consistent with reduced alcohol dependence, and this was significant (effect of time, $p < 0.0001$), but the between-group differences for ADS were not statistically significant (treatment group time interaction: $p = 0.77$).

Similarly, DrInC-Recent total scores were lower at week 16 than at baseline in all treatment groups (effect of visit, $p < 0 .0001$), but with no significant differences between the treatment groups (group visit, $p = 0 .59$).

Table 119. Total DrInC-Recent Scores at Week 16 and Change from Inclusion

	Week 16 Recent	Change from inclusion to Week 16
Placebo (n)	45	45
Mean	29.0	-6.3
SD	22.3	16
NMF 10 (n)	45	45
Mean	27.2	-6.1
SD	18.4	17.1
NMF 40 (n)	46	45
Mean	22.9	-8.9
SD	17.9	16.0
p-values (rm-ANOVA)	Group: p=0.33 Visit: p<0.0001 Group*visit: p=0.59	

The OCDS scores showed a similar pattern of significant improvement in all groups, but no significant difference between groups (group visit, $p = 0.22$).

Table 120. Total OCDS Scores at Screening, Week 6 and Week 16

	Screening	Week 6	Week 16	Change from screening to Week 6	Change from screening to Week 16
Placebo (n)	50	46	45	46	45
Mean	17.1	13.4	10.7	-3.7	-6.2
SD	5.8	7.0	5.9	6.6	5.5
NMF 10 (n)	50	48	45	48	45
Mean	18.5	13.3	13.0	-4.7	-4.6
SD	6.2	5.0	5.1	5.8	6.3
NMF 40 (n)	50	47	47	47	47
Mean	16.9	11.1	10.4	-5.7	-6.3
SD	5.8	5.5	6.2	6.4	6.7

7.3.3.14. Subgroup analysis

No subgroup analyses were performed.

7.3.4. Rejected study biotie400 (CPH-101-0400)

7.3.4.1. Study design, objectives, locations and dates

This was a small, open-label, uncontrolled feasibility study to assess the treatment of alcohol excess with as-needed nalmefene for up to 52 weeks. The target dose was 20mg, but subjects were permitted to adjust the dose to 10mg or 40mg based on tolerability and perceived efficacy.

7.3.4.2. Inclusion and exclusion criteria

Subjects were eligible for inclusion if they were aged 18 or older, drinking heavily and reported difficulties in controlling their drinking. Subjects were excluded if they needed inpatient detoxification, or had serious medical or psychiatric problems.

7.3.4.3. Study treatments

The starting dose of nalmefene hydrochloride for each subject was 20 mg, used as needed, when drinking was imminent, preferably at least an hour before alcohol intake.

Medication was taken on 89% of drinking days, showing reasonably good compliance with these instructions.

Subjects were allowed to increase the dose to 40 mg at 2 weeks or later if the treatment seemed ineffective (if the score of the CGI ≥ 3, consistent with ratings from “minimally improved” to “very much worsened”).

Subjects were also allowed to reduce the dose from 20 mg to 10 mg or from 40 mg to 20 mg, either temporarily or for the rest of the study, if they had apparent side effects.

7.3.4.4. Efficacy variables and outcomes

The main efficacy variables were similar to many of the other Biotie studies, and included:

- Number of HDDs
- Number of hazardous and non-hazardous drinking days
- Number of abstinence days,
- Number of non-heavy drinking days
- Number of very heavy drinking days
- Drinks per drinking day,

- Mean weekly consumption,
- Monthly ratio of heavy drinking days to drinking days,
- Blood alcohol concentration,
- Mean corpuscular volume,
- ALAT and GGT
- carbohydrate deficient transferrin,
- Obsessive compulsive drinking scale,
- Alcohol dependence scale,
- Drinker inventory of consequences
- Subject's clinical global impression.

The primary efficacy variable was the number of HDDs, but there was no comparative group so there was no formal statistical comparison suitable to be used as a primary efficacy endpoint.

The Sponsor performed a statistical assessment of the change in HDDs over time, but this is not a suitable efficacy endpoint and should be considered merely descriptive. HDDs would be expected to decrease in any interventional study of alcoholism, and indeed similar assessments of the placebo groups of the other submitted studies showed significant decreases compared to baseline.

7.3.4.5. Randomisation and blinding methods

All subjects received open-label treatment.

7.3.4.6. Analysis populations

All subjects entered the ITT population and were analysed.

7.3.4.7. Sample size

No formal sample size estimations were performed, which was appropriate given the uncontrolled nature of the study.

7.3.4.8. Statistical methods

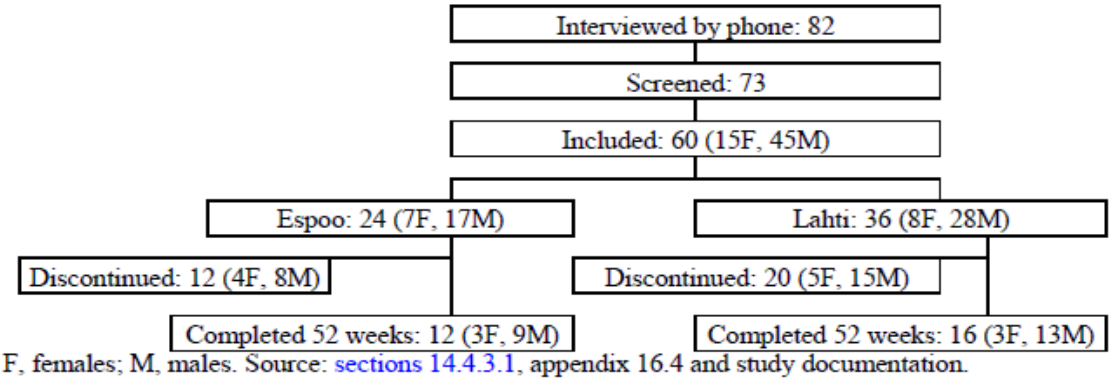
All counts based on number of drinking days, including HDDs, were analysed with generalised

linear mixed-effects models, using a Poisson or binomial distribution assumption. Efficacy laboratory parameters were analysed with RM-ANOVA. Scores for the DrInC and ADS were analysed with repeated-measures Poisson regression, whereas the OCDS was analysed with a linear mixed-effects model. CGI was primarily used as a descriptive measure, but the effects of gender and centre were tested with logistic regression.

7.3.4.9. Participant flow

Patient disposition is summarised in the figure below, by centre. The initial population included 15 females and 45 males (total n = 60). Only 28 subjects (47%) completed the 52 weeks in the study. This low completion rate adds to other methodological concerns. Reasons for discontinuation are shown in the table.

Figure 39. Disposition of Subjects – Biotie400



F, females; M, males. Source: sections 14.4.3.1, appendix 16.4 and study documentation.

Table 121. Reasons for Discontinuation – Biotie400

Reason	Espoo N (%)	Lahti N (%)	Total N (%)
Total	12 (50.0)	20 (55.6)	32 (53.3)
Consent withdrawn	5 (20.8)	10 (27.8)	15 (25.0)
Adverse event	3 (12.5)	7 (19.4)	10 (16.7)
Lost to follow-up	4 (16.7)	2 (5.6)	6 (10.0)
Other	-	1 (2.8)	1 (1.7)

Source: section 14.4.3.3

7.3.4.10. Major protocol violations/deviations

The study report did not clearly differentiate minor and major protocol deviations. Some of the more important by-subject deviations are shown in the table below. Among the more serious systemic deviations, the protocol did not specify definitions of a standard drink.

Table 122. Summary of Some Deviations by Centre – Biotie400

Deviation	Espoo	Lahti	Total
Electronic CRF created, although subject did not sign the consent	1	0	1
Investigator signed subject's informed consent at inclusion visit	5	0	5
Informed consent: study nurse dated it instead of the subject; wrong signature month or subject's address missing	3	0	3
Subject did not sign the updated subject information	-[a]	3	-
Female subject used only condom for contraception	1	1	2
More than 14 days between screening and inclusion visits (range 15 – 19 days)	11	6	17
Use of illicit drugs found in the urine drug screen	1	0	1
Subject was not sober at inclusion visit	1	0	1
More than 14 consecutive abstinence days prior screening	1	0	1
Incorrect subject initials used	2	0	2
Subject did not return all study medication blisters	10	12	22
Subject used expired study medication	0	1	1
Week 28 electrocardiogram was taken 2 to 4 weeks late	11	0	11
Results of screening carbohydrate deficient transferrin not entered to the database from the laboratory printout	3	0	3
Urine drug screen results not entered to the database from the laboratory printout	2	0	2
Screening mean corpuscular volume value probably a laboratory error	1	1	2
Appendix 16.4 includes TLFB data with "signed" status instead of "Finalized".	0	1	1
The eCRFs and SAS datasets include wrong data in week 16 obsessive-compulsive drinking scale and visit dataset because of an entry error. This data has been deleted from the appendix 16.4.	0	1	1
Missing visits	1	0	1

Subject count is shown. [a]By-subject information not available in the study documentation. Source: appendix 16.2.2 and the study documentation.

7.3.4.11. Baseline data

All subjects were Caucasian. The age range was 32-75 (mean 49.9) years. The subjects' body mass index was 18-38 (mean 27.5) kg/m2. Nine subjects were unemployed and 36 were full-time employed. The age at the onset of problem drinking ranged from 16 to 60 years (mean 35). Half of the subjects had been treated for alcohol problems previously.

In general, this means the population broadly resembled those treated in the other submitted studies.

Table 123. Summary of Gender, Age and BMI – Biotie400

	Espoo (N=24)	Lahti (N=36)	Total (N=60)
Females : Males	7 (29.2%) : 17 (70.8%)	8 (22.2%) : 28 (77.8%)	15 (25.0%) : 45 (75.0%)
Age (years)	49.5 (8.9; 34-69)	50.2 (8.8; 32-75)	49.9 (8.8; 32-75)
Body mass index (kg/m^2)	26.8 (4.7; 18-36)	28.0 (5.0; 20-38)	27.5 (4.9; 18-38)

Means (with standard deviation and range) are given. Source: 14.1.1.1, 14.3.6.1-3.

Table 124. Summary of Social Status – Biotie400

		Espoo N (%)	Lahti N (%)	Total N (%)
Marital status	Married	10 (41.7)	24 (66.7)	34 (56.7)
	Never Married	3 (12.5)	1 (2.8)	4 (6.7)
	Divorced	10 (41.7)	10 (27.8)	20 (33.3)
	Widower	1 (4.2)	1 (2.8)	2 (3.3)
Living status	Alone	7 (29.2)	5 (13.9)	12 (20.0)
	With spouse or partner	9 (37.5)	17 (47.2)	26 (43.3)
	With spouse or partner and children	5 (20.8)	11 (30.6)	16 (26.7)
	Single with children	2 (8.3)	3 (8.3)	5 (8.3)
	Other	1 (4.2)	-	1 (1.7)
Employment	Student	-	-	-
	Employed, full time	17 (70.8)	19 (52.8)	36 (60.0)
	Employed, part time	1 (4.2)	2 (5.6)	3 (5.0)
	Unemployed	1 (4.2)	8 (22.2)	9 (15.0)
	Retired	4 (16.7)	7 (19.4)	11 (18.3)
	Other	1 (4.2)	-	1 (1.7)
Education	Secondary school or less	6 (25.0)	8 (22.2)	14 (23.3)
	Vocational school	12 (50.0)	14 (38.9)	26 (43.3)
	Upper secondary school graduate	-	12 (33.3)	12 (20.0)
	College	1 (4.2)	-	1 (1.7)
	University graduate	5 (20.8)	2 (5.6)	7 (11.7)

Source: sections 14.1.1.4-8.

7.3.4.12. Results for the primary efficacy outcome

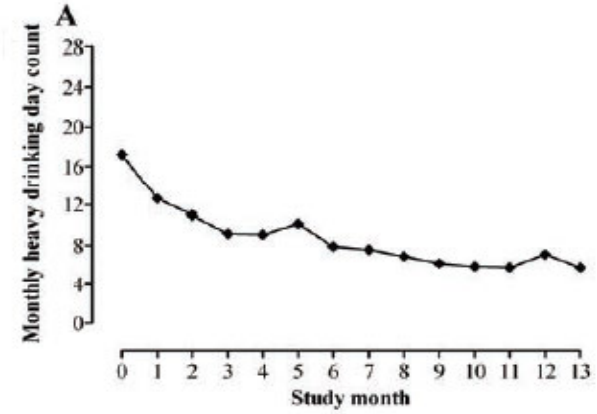
The median count of HDDs/month fell from 17 at baseline to 3-6 for the second half of the study. The decrease was largest during the first two months, a pattern observed in nearly all treatment groups of the submitted studies, including placebo groups. Whether the fall observed in this study represents a pharmacological effect is completely unknown, but a placebo effect is likely to account for most of the observed reduction, possibly supplemented by withdrawal bias.

Table 125. Heavy Drinking Day Count by Study Month – Biotie400

	Month	0	1	2	3	4	5	6	7	8	9	10	11	12	13
	N	59	51	49	44	43	37	35	31	30	30	30	29	28	24
Count	Median	17	12	10	6.5	5	11	6	6	5.5	6	4	4	5.5	3
	Mean	17.2	12.6	11.0	9.2	9.1	10.2	7.9	7.6	6.9	6.2	5.9	5.8	7.1	5.8
	SD	6.3	7.6	8.6	8.5	8.4	7.4	6.8	7.1	6.2	5.7	6.1	5.9	7.2	7.0
Change	Median		-4	-8	-8	-9	-9	-10	-10	-10	-11	-12	-11	-10	-10.5
	Mean		-4.9	-6.6	-8.4	-8.3	-7.8	-9.5	-9.6	-10.7	-11.3	-11.7	-11.8	-10.4	-11.5
	SD		7.2	7.8	8.1	8.1	7.0	7.1	6.8	6.3	5.9	6.2	6.3	7.3	7.0
Per cent change	Median		-22.2	-50.0	-58.6	-62.5	-45.5	-66.7	-68.8	-69.4	-68.9	-73.9	-75.0	-68.8	-79.4
	Mean		-25.5	-37.4	-47.9	-48.5	-43.6	-54.0	-56.9	-62.1	-65.6	-67.8	-68.2	-61.3	-68.4
	SD		41.8	45.1	42.7	43.5	39.4	38.2	37.2	30.1	27.1	29.1	26.1	36.3	33.9

Month 0 refers to baseline. N, subject count; SD, standard deviation. Source: section 14.2.1.1-3.

Figure 40. Heavy Drinking Day Count by Month – Biotie400



The Sponsor explored the influence of family alcohol history on HDDS, and found it had no significant effect.

7.3.4.13. Results for other efficacy outcomes

Changes in other counts derived from TLFB were generally consistent with the HDDs, and face the same problems of interpretation.

The serum ALAT and GGT and the CDT% showed significant improvement during the study, consistent with the observed reduction in alcohol intake. The MCV did not change.

The CGI and the psychometric scores (ADS, OCDS and DrInC) showed improvement during the study, consistent with the observed reduction in alcohol intake and subject to the same problems of interpretation.

7.4. Published studies of nalmefene in alcohol dependence

The Sponsor mentioned the following published studies of nalmefene in Alcohol Dependence. An assessment of these studies is beyond the scope of this evaluation. All of the studies listed are small and the duration was inadequate, so they have little potential for significantly modifying the overall weight of evidence assessing the efficacy of nalmefene. Also, these studies tested a range of doses, with relatively few patients receiving the proposed 20mg dose. The drug was not taken one hour prior to expected alcohol consumption, as proposed in the current submission.

Table 126. Published (Investigator Initiated) Nalmefene Studies in Alcohol Dependence

Publication	Study Design[a]	Number of Patients in ITT Population		
		PBO	Naltrexone	NMF
Publications based on original data				
Mason *et al.* 1999[25]	12-week, randomised, double-blind, placebo-controlled, fixed-dose (20 or 80mg/day), twice-daily dosing	35		20mg+ 80mg: 70
Mason *et al.* 1994[26]	12-week, randomised, double-blind, placebo-controlled, fixed-dose (10 or 40mg/day), twice-daily dosing	7		10mg: 7 40mg: 7
Drobes *et al.* 2003[27]	1-week, laboratory-based, single-blind, placebo-controlled, active-reference, fixed-dose (naltrexone 50mg/day; nalmefene 40mg/day)	77	67	57
Publications based on previously published data				
Mason 1996[28b]	12-week, dose-ranging, fixed-dose (10, 20, 40, or 80mg/day)			n/a
Drobes *et al.* 2004[29c]	1-week, laboratory-based, single-blind, placebo-controlled, active-reference, fixed-dose (naltrexone 50mg/day; nalmefene 40mg/day)	n/a	n/a	n/a
Total		**119**	**67**	**141**

a Doses are based on nalmefene hydrochloride; nalmefene hydrochloride 20mg corresponds to 18.1mg nalmefene base.

b This publication compared the results from 14 patients who received nalmefene 10mg/day or 40mg/day with the results from 16 patients who received nalmefene 20mg/day or 80mg/day. Although not explicitly stated, the patients included in this publication appear to be a subset of those in the two clinical studies reported by Mason *et al.* (1994 and 1999).[25,26]

c This publication evaluated the effects of alcohol on subjective craving, stimulation, and sedation in the study by Drobes *et al.* (2003).[27]

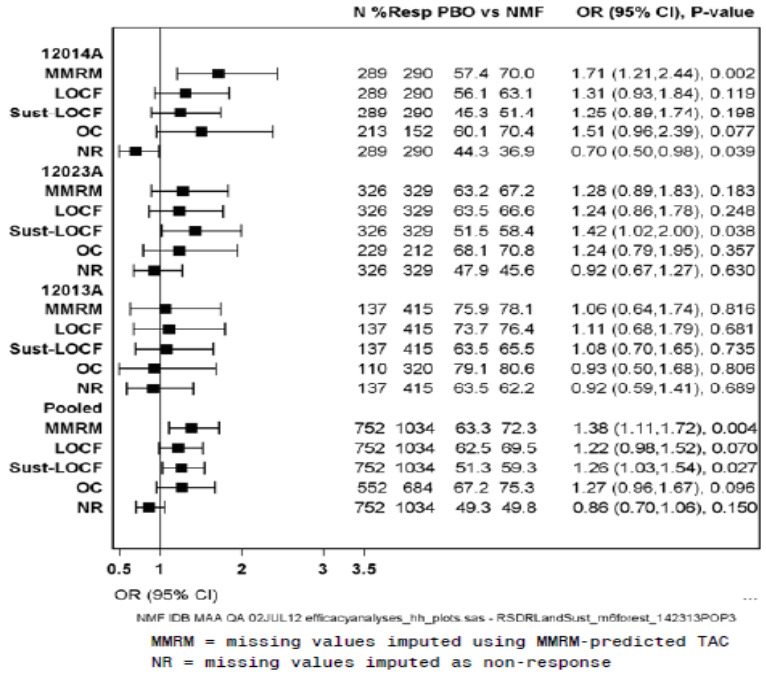
7.5. Analyses performed across studies

7.5.1. Pooled analysis across the 3 Lundbeck Studies

For several major endpoints, data from the two pivotal studies, Lundbeck14 and Lundbeck23, were pooled, which is appropriate as these studies shared almost identical designs. These pooled results have been presented in the discussion of the efficacy results of those studies.

The Sponsor also pooled the response rates (RSDRL) across the three Lundbeck studies, as shown below. This pooling was partially appropriate, as entry criteria and response definitions were similar across the three studies, but the original study plans varied in their proposed imputation methods. The response rates in all three studies, as well as the pooled data set, are shown below. Using an optimistic imputation method (MMRM), the pooled results were significantly in favour of active treatment (odds ratio 1.38). With pessimistic imputation methods (NR), there was a moderate trend in favour of placebo, which failed to reach significance. Given the high and unequal withdrawal rate in Lundbeck14 (already discussed), and the divergent results based on different imputation methods, this pooled analysis should be considered unreliable.

Figure 41. RSDRL at Month 6 – Proportion (%) of Responders and OR for Response (FAS, LREG) – Lundbeck14, Lundbeck23 and Lundbeck13, and Pooled Lundbeck Studies



7.5.2. Sponsor's estimations of clinical impact

The Sponsor's Clinical Overview included a section in which the adverse consequences of excess alcohol were modelled, based on previously published estimates of risk for 8 alcohol-related diseases (ischaemic heart disease, ischaemic stroke, traffic-related injuries, non-traffic-injuries, liver cirrhosis, pancreatitis, pneumonia, and haemorrhagic stroke). The modelling process was described in more detail in Module 5.3.5.3, Reports of Analyses of Data from More than One Study (clinical-relevance-report.pdf). A detailed evaluation of this complex model is beyond the scope of this evaluation, but the main features are discussed below.

To build the model, alcohol consumption was simulated for 200,00 patients but then expressed in terms of a notional cohort of 100,000 patients, using drinking patterns for each patient based on the characteristics of patients with a high or very high DRL at Baseline and Randomisation in the Lundbeck studies. These drinking patterns included days with no alcohol consumption and days with high alcohol consumption, with a distribution resembling that seen in the real patient cohort.

The next step was to convert those simulated drinking patterns into risk estimates. The risk of an event was determined according to the individual level and pattern of alcohol consumption and gender for all patients in the notional cohort. The risk equations used for this step were developed independently by a team of experts led by Dr Jürgen Rehm (Chair, Addiction Policy, and Professor, University of Toronto). The model was used to produce estimates for a range of drinking levels extending from 100 HDDs per year to > 220 HDDs per year, as shown in the table below. (A similar approach was taken for TAC).

Of note, the model has not been validated as a predictive tool, so the results of this analysis should be considered speculative.

Table 127. Number of Events per 100,000 Patient-years by Number of HDDs

Number of HDDs (days)	Ischaemic Heart Disease	Ischaemic Stroke	Traffic Injuries	Other Injuries	Cirrhosis	Pancreatitis	Pneumonia	Haemorrhagic Stroke
<100	1563	494	303	2434	330	141	1827	149
100-120	1707	538	398	3085	428	188	1957	171
120-140	1819	572	465	3602	505	235	2046	188
140-160	1928	606	531	4112	582	291	2120	203
160-180	2027	637	598	4553	686	391	2217	223
180-200	2118	666	668	4953	836	596	2340	248
200-220	2211	695	747	5361	1084	1136	2523	284
>220	2300	724	822	5742	1378	2251	2705	323

The Sponsor used the model to convert reductions in HDDs observed in the studies to theoretical event reductions for the 8 alcohol-related diseases in a cohort of 100,000 patients. This was done in two ways: using the absolute mean reduction in HDDs (which includes the effect of psychosocial interventions and the placebo effect), and the placebo-adjusted reduction in HDDs compared to placebo.

In introducing this analysis, the Sponsor suggested that nalmefene produced an absolute fall from 23 HDDs per 'month' (4-week cycle) to 11 HDDs per month. This matches the monthly drinking levels observed in the high-risk cohort that showed high or very high DRL at Baseline and Randomisation. The Sponsor equated these levels with the highest and second-lowest risk brackets in the risk model, as follows:

> *"To evaluate the clinical relevance of the absolute reduction from baseline to Month 6 in the number of HDDs in the two 6-month studies, the model estimated the number of events in a cohort of 100,000 patients with alcohol dependence who have >220 HDDs/year and in a cohort of 100,000 patients with alcohol dependence who have 100 to 120 HDDs/year. The difference between these cohorts corresponds to reducing drinking from 23 HDDs/month to 11 HDDs/month."* (Clinical Overview)

The assumptions behind this statement were not clear in the Clinical Overview, but the associated document, clinical-relevance-report.pdf, indicates that the estimates are based on a high-risk subgroup identified post hoc, using an optimistic MMRM/OC imputation method. Nonetheless, given that the claimed reductions in this cohort were from 23 HDDs/4-weeks to 11 HDDs/4-weeks, which is equivalent to a fall from 299 HDDs to 143 HDDs per 52-week year, the Sponsor's focus on the two brackets >220 and 100-120 appears unjustified. Given the incongruities in these figures, it was proposed in the first-round CER that the Sponsor might have been basing their calculations on their preferred post hoc subgroup, but they have since stated that the total population was considered. It was also suggested that the Sponsor may have they miscalculated the number of 4-week cycles per year, and the Sponsor has since agreed that they used an approximation of ten 4-week cycles per year.[17]

[17] The evaluator notes that the sponsor stated: "We fully acknowledge that what has been used is an approximation using a conservative approach (using a multiplier of 10 cycles of 4 weeks instead of 13). It is not a miscalculation but an approximation using a conservative approach." This seems a very odd and inaccurate approximation to insert into a complex modelling process, but a full exploration of this model is beyond the scope of the CER.

Comparing modelled risks in those two brackets, the Sponsor then estimated the absolute benefit of treating 100,000 patients for one year, using data from the table above. Ostensibly, such treatment would prevent 7,773 events:

- 593 events of ischaemic heart disease
- 186 events of ischaemic stroke
- 424 events of traffic-related injuries
- 2657 events of non-traffic-related injuries
- 950 events of liver cirrhosis
- 2063 events of pancreatitis
- 748 events of pneumonia
- 152 haemorrhagic stroke[18]

Many of the diseases listed take more than a year to develop, so the concept of benefit after one year is not clinically realistic; to achieve 100,000 patient-years of nalmefene treatment and have a benefit on the incidence of cirrhosis, a smaller number of subjects would need to receive nalmefene for periods much longer than a year, without discontinuation or relapse.

Re-expressing those figures as events per hundred patient-years, which is more clinically meaningful from an individual clinician's perspective, the Sponsor's analysis suggests that about one case of cirrhosis would be prevented and about 3 injuries, with less than one case prevented for each of the other diseases. Note that this estimate includes the placebo effect and continued psychosocial support, and assumes that subjects do not relapse into heavy drinking.

For the attributable reduction in risk of harm, compared to placebo, the Sponsor assumed that subjects would shift by one additional risk bracket beyond the placebo response, from the 120-140 HDDs/year bracket to the 100-120 HDDs/year. The placebo-subtracted reduction in HDDs/month in the two studies was 2.3 and 1.7 for Lundbeck14 and Lundbeck23, respectively, corresponding to 30 HDDs/year and 22 HDDs/year. A one-bracket shift is therefore a reasonable estimate (assuming no biases have inflated the pivotal efficacy results).

Comparing the risks in the above table in the two relevant brackets (100-120 HDDs and 120-140 HDDs), the Sponsor estimated that treating 100,000 patients for one year would prevent 960 events:

- 112 events of ischaemic heart disease
- 34 events of ischaemic stroke
- 67 events of traffic-related injuries
- 517 events of non-traffic-related injuries
- 77 events of liver cirrhosis
- 47 events of pancreatitis
- 89 events of pneumonia
- 17 events of haemorrhagic stroke

Again, this estimate ignores the fact that cirrhosis takes years to develop. The total harm reduction is equivalent to just under one (0.96) event prevented per hundred patient-years of

[18] The haemorrhagic strokes were left off this list in CER_1.

treatment, assuming these estimates have not been inflated by methodological issues such as withdrawal bias and unblinding.

More than half of the events prevented (517/960, 54%) represent non-traffic injuries, which could include a range of minor injuries. Removing these, treatment might be expected to prevent 0.44 serious events per 100 patient-years of treatment, equivalent to 0.0044 events per patient, or one event per 226 patients.

The Sponsor finished the section on clinical relevance with the following statement (emphasis added):

> *To further justify the clinical relevance, a new additional analysis was performed (Clinical relevance of nalmefene versus placebo [Clinical Relevance Report]). In this analysis, which combined the LOCF-estimated treatment effect versus placebo in the patients with a* ***high or very high DRL at baseline and randomisation*** *with the results from a comprehensive meta-analysis on reduced drinking and mortality in patients treated for alcoholism,* ***the difference between nalmefene and placebo is predicted to lead to an 8% (95% CI: 3%, 13%) reduction in mortality.*** *A sensitivity analysis based on a more conservative assumption predicted a 4% (95% CI: 1%, 8%) reduction in mortality.*

Note that the 8% appears to be a relative reduction, and only refers to mortality attributable to alcohol. The absolute reduction in annual mortality would be expected to be very small, but the Sponsor has not provided an estimate of actual number of lives saved per 100,000 high-risk patients treated. (Note that, in the PK context, the Sponsor considered a 9% increase in exposure to be so small as to be not worth considering, even though this increase could also be explored through a mortality model to produce a relative increased risk of death attributable to the PK effect.)

Table 128. Sponsor's Estimate of Change in Mortality (Nalmefene vs Placebo) Assuming OR = 0.41 for Reduced Drinking Including Abstinence

		Mortality Risks (Nalmefene *versus* Placebo)	
		Point Estimate	**95% CI**
Pooled 6–month Studies	Women	0.971	(0.887, 1.056)
	Men	0.916	(0.849, 0.983)
1-year Study	Women	0.789	(0.579, 1.000)
	Men	0.867	(0.728, 1.006)

One of the key assumptions underlying the 8% estimate was that nalmefene would be associated with an odds ratio of 0.41 for 'reduced drinking including abstinence' (as noted in the title of the Sponsor's table, above). In this respect, it is worth reviewing the odds of a reduced-drinking response as observed in the pivotal studies, using the Sponsor's prospective definitions of response and the prospective target group. Response rate is one of the key secondary efficacy endpoints recommended by the EMA, who proposed it as one way of gauging clinical relevance of a purported treatment for alcohol dependence. It should be noted that, in the two pivotal studies, responses were not more common with nalmefene treatment. In fact, in Lundbeck14, the Month 6 response rate was 44.3% in the placebo group, compared to 36.9% in the nalmefene group; that is, the response rate was significantly inferior with active treatment (p = 0.039). In Lundbeck23, the response rate was 63.2% in the placebo group and 67.2% in the nalmefene group, consistent with a marginally better (but not statistically significant) response

with active treatment (p = 0.1833).[19] The proposed odds ratio of 0.41 is therefore implausible, and it is not supported by any prospective evidence.

Given that cirrhosis and other complications of drinking take years to develop, the suggested 8% reduction in mortality inherently assumes long periods of reduced drinking, achieved with the combination of nalmefene and continued psychosocial support, but there is no evidence of sustained multi-year efficacy for nalmefene. In fact, it seems very unlikely that the intensive monitoring of alcohol intake that was performed in the pivotal studies would be continued for years on end by clinicians prescribing nalmefene. Also, these claims of mortality reduction are based on optimistic interpretations of the efficacy evidence which, as discussed below, was generally weak and disappointing. It is completely unclear how the mortality calculations would be modified by appropriate adjustment for unblinding, withdrawal bias and possible PK interactions.

7.6. Evaluator's conclusions based on the CER round 1

The submitted evidence only weakly supports the Sponsor's claim of efficacy. The major concerns arising on evaluation of the evidence are discussed in separate sections below. First, there appears to be a potential PK interaction between nalmefene and alcohol that undermines the proposed therapeutic effect. Second, there was a large potential for unblinding in the pivotal studies, which is of particular concern given that the endpoint was behaviourally mediated and under some voluntary control, and also because alcohol intake was estimated retrospectively rather than recorded directly. Third, there was a high withdrawal rate, so all endpoints depend critically on imputation methods, and there was a large potential for withdrawal bias. Fourth, most of the submitted studies were negative even without adjustments for the first three concerns. Fifth, the mean magnitude of the benefit was small, and of modest clinical value. Sixth, the proposed PI emphasizes results in a subgroup that was defined post hoc, which means that none of the cited p-values are legitimate. Seventh, the level of psychosocial support was carefully controlled in the pivotal studies; if reliance on medication led to any reduction in psychosocial measures, this would be expected to compromise or even reverse any benefit from nalmefene.

7.6.1. Potential compromise of efficacy by PK interactions

A PK interaction study (13513A) suggested that nalmefene increases exposure to ethanol by ~9%, although there is substantial uncertainty surrounding this estimate and the effect could increase exposure by 21%, or not at all. The ratio of the AUC for ethanol with nalmefene compared to ethanol without nalmefene was 1.086 (90%CI 0.977 to 1.208). This result was not statistically significant, but lack of a statistically significant PK effect is not the same as statistical proof of a lack of effect, and 8.6% is the best current estimate for the magnitude of the effect.

[19] The evaluator states that it should be recalled that the prospective imputation methods differed in the two studies, which partly accounts for the differences mentioned here. The differences in response rates were less marked when using the same imputation method. See the pivotal efficacy results for other imputation approaches.

Table 129. PK Parameters of Ethanol following a Single Oral Dose of 0.6g/kg Ethanol Administered to Subjects Exposed to Nalmefene 20mg or Placebo – Study 13513A

Parameter	Ethanol plus Nalmefene (20mg) or Placebo Tablet	Ethanol		Nalmefene plus Ethanol *versus* Placebo Tablet plus Ethanol (all subjects) Ratio (90% CI)
		Men	Women	
AUC_{0-t} (mmol·h/L)	NE	42.0 (25.6)[a]	52.1 (30.9)[a]	1.086 (0.977, 1.208)
	PE	38.4 (35.4)[b]	46.8 (37.2)[c]	
C_{max} (mmol/L)	NE	14.7 (19.5)[a]	18.1 (31.5)[a]	1.003 (0.913, 1.101)
	PE	15.2 (<2.2, 19.1)[b]	18.5 (24.8)[c]	
t_{max} (h)	NE	2.25 (1.50, 3.00)[a]	2.50 (1.50, 4.00)[a]	
	PE	2.00 (1.50, 3.00)[d]	2.00 (1.50, 3.00)[c]	

Arithmetic mean (CV%) data are presented for AUC_{0-t} and C_{max} (NE for men and women and PE for women).
Median (min, max) data are presented for t_{max} and C_{max} (PE for men).
NE = nalmefene plus ethanol; PE = placebo tablet plus ethanol; N = number of subjects; CI = confidence interval

a N=16 c N=21
b N=18 d N=17

The treatment effect in the pivotal efficacy studies was modest, but it appears even less impressive when the PK interaction between ethanol and nalmefene is used to estimate the effective TAC in nalmefene recipients. If the AUC for ethanol is increased by a factor of 1.086, every 1.0 gram of ethanol consumed by a subject taking nalmefene is potentially producing the same exposure that would normally be produced by 1.086 g of ethanol, and this adjustment needs to be applied to the pivotal efficacy results.

To put this in context, the estimated treatment effect for TAC was 5g/d in Lundbeck23, the pivotal study least compromised by unequal withdrawals, and the mean baseline TAC was 89 g/d. In both pivotal studies, the reduction from baseline in the active group was to a new level of ~34g/d.[20] Multiplying this alcohol consumption by the PK interaction effect (1.086) increases the effective TAC in the nalmefene group to ~37g/d (~3g higher than without the adjustment), negating more than half the 5g treatment benefit. If the upper estimate of the 90%CI were used for the PK interaction (1.208), the effective TAC in the nalmefene group would be 41 g/d, which is 7g/d more than a simple accounting of drinks would suggest, and enough to reverse the treatment effect (so that the nalmefene group could have had ~2g/d more alcohol exposure at Month 6 than if they had taken placebo).

In practice, the PK effect on TAC may actually have been less than this, because of poor compliance, and the PK interaction between nalmefene and ethanol has not been studied rigorously so the true effect is uncertain, but this is an important point that was completely ignored by the Sponsor.

In this respect, the fact that the interaction as measured by the AUC ratio fell within the standard bioequivalence limits of 0.8 to 1.25 is not particularly reassuring, The standard bioequivalence limits are appropriate for a drug with a broad therapeutic index, but they are not appropriate when modifying exposure to a toxic agent such as ethanol is the sole purpose of treatment. For instance, if the efficacy endpoints of the pivotal studies were similarly assessed according to bioequivalence criteria (unadjusted for the PK interaction), then ethanol consumption with and without nalmefene would be considered equivalent. In Lundbeck14, mean TAC was reduced by 39.7 g/day in the placebo group, compared to 50.7 g/day in the nalmefene group, a ratio for placebo:nalmefene TAC of 0.78 (or 1.28 for nalmefene:placebo). This ratio is just outside the standard bioequivalence limits. In Lundbeck23, the TAC was reduced by 54.1 g/day in the placebo group, compared to 59.0 in the nalmefene group, a ratio of 0.92 (or 1.09 for nalmefene:placebo), which is within the bioequivalence limit. It would not be

[20] The evaluator points out that in Lundbeck14, the reduction was from a baseline intake of 84 g/day. In Study 23, the reduction was from a baseline of 93 g/day. The reduction relative to baseline was 51g and 59g in the two studies, respectively, giving TAC at Month 6 of 84-51=33g/d in Lundbeck14 and 93-59=34g/d in Lundbeck23.

consistent to consider an exposure ratio of 1.086 unimportant in the PK context and then, in the same submission, propose that a TAC-reduction ratio of 1.09 represents a worthwhile clinical effect.

The Sponsor appears to believe that uncertainty on this issue works in their favour (because there is still a possibility of no PK interaction), but the residual uncertainty on this issue actually weakens their efficacy claims. The upper half of the 90%CI for the PK interaction ranges from an increase in alcohol exposure of ~9% (enough to seriously compromise the efficacy of nalmefene) to an increase of 21% (enough to reverse the efficacy of nalmefene and potentially lead to increased alcohol-related complictions in nalmefene users).The confidence interval for the estimated size of the treatment effect also exhibits a spread, and the lower half of that interval includes benefits small enough to be cancelled by the mean PK interaction. The Sponsor's claims currently rest on the hope that the treatment effect is in the upper half of the uncertainty range while the PK effect is in the lower half of the uncertainty range. Thus, far from demonstrating that the treatment effect is significant with the traditional 95% certainty, the Sponsor has not yet demonstrated that an overall beneficial effect (in their original target population) is even likely.

A more favourable balance between the PK estimates and efficacy estimates could be inferred if the Sponsor's preferred post hoc subgroup were accepted as the target group, but uncertainty would nonetheless weaken the Sponsor's claims. Correct statistical accounting of the residual uncertainties in the PK and efficacy domains would require that efficacy estimates in the post hoc efficacy analysis incorporated residual PK uncertainties; that is, errors in one domain would need to be propagated into the next. The Sponsor has not only failed to perform such an analysis, but does not even seem to recognise that the issue is important.

7.6.2. Potential for unblinding in the pivotal studies

The primary endpoints in both pivotal studies were based on a timeline followback (TLFB) method, in which subjects estimate their drinking behaviour by looking back over their social activities and reporting the number of drinks consumed. Unblinding could have affected the primary endpoints by affecting drinking behaviour or by affecting reporting of drinking.

Because alcohol intake in the pivotal studies was voluntary, each decision to drink depended on a range of subconscious and conscious psychological factors. The pivotal studies provided direct proof of the importance of psychological factors, because many subjects curtailed their drinking between Baseline and Randomisation, and the overall reduction in the placebo group was substantially greater than the additional reduction associated with active treatment. If active treatment caused unblinding via side effects, and this magnified the psychological effect of treatment even slightly, this could be enough to account for the minor additional benefit of active treatment over placebo.

Unblinding is a particularly important issue because of the stigma associated with alcoholism. Subjects with alcohol dependence are usually embarrassed and concerned by their drinking, and by their perceived lack of control over their drinking. It seems likely that subjects who knew they had just taken an active agent aimed at improving control would experience even higher levels of remorse or embarrassment when they drank excessively than subjects who thought they were taking placebo. That is, each drink beyond their preferred limits would not only prove to themselves that they had poor control, but also establish that their self control was so poor it could not be rescued pharmacologically. Thus, the primary endpoints could be highly susceptible to the effects of unblinding at the behavioural phase, producing a genuine reduction in drinking, but by a spurious mechanism unrelated to the actual pharmacology of the drug, much as one might expect a bitter placebo to produce different behaviour to a bland-tasting placebo.

The TLFB and all other methods relying on patient reports are also inherently prone to recall bias, because subjects who are embarrassed about their drinking may be motivated to reduce

their estimates. Subjects who believe they are receiving a new drug designed to curb excess alcohol could easily suffer even more embarrassment than usual if, despite the drug, they have nonetheless kept drinking heavily. Conversely, subjects who believe they are receiving placebo could easily feel justified in continuing higher levels of drinking because they may feel they have been denied pharmacological assistance. The primary endpoints would, therefore, be highly susceptible to unblinding in the reporting phase.

It is not known whether unblinding occurred in the pivotal studies, but there is ample indirect evidence suggesting that unblinding may have been substantial.

(In the first-round Clinical Evaluation Report, concern was raised about the fact that nalmefene has a bitter taste. In the Biotie studies, a bittering agent was used in the placebo tablets to minimise this problem, but no mention of a bittering agent was found in the Lundbeck study reports. The Sponsor has since explained that a bittering agent was used in the Lundbeck studies, which is reassuring, but it is of concern that the Sponsor did not realise that unblinding was an important issue and did think to mention this in their initial submission.)

Nalmefene produces a range of side effects including dizziness, nausea, headaches and sleep disturbance. In the alcohol-nalmefene interaction study, adverse events occurred in 81% of subjects exposed to nalmefene 20mg (with placebo to ethanol), compared to 15% of subjects exposed to nalmefene placebo (with placebo to ethanol). This suggests that most subjects experienced a tell-tale side effect from nalmefene, and many subjects may have realised they were receiving active treatment from the first day of treatment. Given that milder symptoms of sedation or light-headedness might not have been reported as AEs, an even higher proportion might have been able to deduce their assigned treatment.

Table 130. Summary of Adverse Events (All Causes, Safety Set)

	Pre-treatment (N=46)	20 mg Nalmefene + Ethanol (N=43)	20 mg Nalmefene + Placebo (to Ethanol) (N=43)	Placebo (to Nalmefene) + Ethanol (N=42)	Placebo (to Nalmefene) + Placebo (to Ethanol) (N=41)	Overall (N=46)
Subjects with adverse events	1 (2%)	43 (100%)	35 (81%)	42 (100%)	6 (15%)	46 (100%)
Subjects with SAEs	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Subjects with adverse events leading to withdrawal	0 (0%)	2 (5%)	2 (5%)	0 (%)	0 (0%)	4 (9%)
Total number of adverse events	1	224	152	80	7	464

Source: Table 51
N = number of subjects

In the pivotal studies, AEs eventually occurred in most placebo recipients, making it difficult to determine the true rate of tell-tale side effects in the nalmefene group from a raw count of subjects reporting AEs. The overall proportion of patients in the Alcohol Dependence Pool (ADP) with TEAEs was 62.7% in the placebo group, compared to 74.7% in the nalmefene group, an absolute excess of 12%. Of the 37.3% of nalmefene recipients who would not be expected to report a TEAE based on the placebo rate, about one third (12% vs 37.3%) reported a TEAE. If TEAEs within the first day of treatment are considered, the incidence of these early TEAES with placebo was 17.7%, compared to 40.8% in the nalmefene group, an absolute excess of 23.1%. Withdrawals from the pivotal studies were attributed to AEs in 6% of the placebo group and 20% of the nalmefene groups, an excess of 14%. Clearly, subjects withdrawing due to AEs are very likely to believe they are receiving active treatment, but the 14% excess does not include additional patients with milder side effects who decided to persist with treatment.

The unblinding signal from telltale side effects might be expected to be even stronger than these figures suggest, because some side effects are intrinsically more likely to be drug-related and the total count of AEs in both groups included a number of AEs unlikely to be interpreted as drug side effects (such as pharyngitis or common viral infections); subtracting these from the

total AE counts would increase the relative excess of nalmefene-related AEs. Also, the total AE count does not reflect the timing of AEs in relation to drug ingestion. Most subjects would be capable of deducing that they were on active treatment if they became dizzy one to two hours after the ingestion of every tablet, whereas a placebo recipient experiencing dizziness at random times would probably not make the same inference.

A variable dosing pattern was used in the pivotal studies, with tablets taken on some days but not others, and tablets were taken at variable times according to need. Compared to regular dosing, this variability would give patients an even greater chance to observe the difference between days on which they took the tablet and days when they did not, and to note the onset of telltale side effects in relation to the variable timing of the tablet. For instance, insomnia was a common side effect, reported in 13% of nalmefene recipients, and only 5% of placebo recipients. Nalmefene recipients troubled by this symptom could easily note that they slept better on days when they did not take nalmefene, whereas placebo recipients with insomnia are unlikely to have noted the same pattern. Similar logic applies to nausea, which affected 22.1% of nalmefene recipients, compared to only 5.9% of placebo recipients. Dizziness was also much more common in nalmefene recipients, affecting 18.2%, compared to 5.5% of placebo recipients. Nausea and dizziness attributable to nalmefene is likely to have had a somewhat predictable relation to the timing of the dose, and to be minimal on non-dosing days, allowing many subjects to realise they were on active treatment.

Despite these obvious methodological challenges, unblinding in the pivotal studies was not assessed, and this important issue was not discussed by the Sponsor. It would have been easy to ask subjects to guess what treatment they thought they were taking, and those answers could have been used to estimate how many subjects had been unblinded, but this important step was not performed.

In the absence of any assessment of unblinding, it is impossible to know how much the primary endpoints would need to be adjusted to account for unblinding. It seems plausible, however, that 20-50% of subjects receiving nalmefene could have become aware of their assigned treatment. If it is assumed that unblinded subjects consumed one standard drink less per day (10g/day) than they otherwise would have, the average reduction in TAC due to unblinding could be about 2-5g. In the pivotal studies, mean TAC was only reduced by 11g in Lundbeck14 and by 5g (half a drink) in Lundbeck23, so the unblinding effect could potentially account for a large part of the apparent treatment effect. If it were assumed that unblinded subjects reduced intake or reporting by 2 standard drinks per day, the apparent treatment effect would essentially be negated. (The apparent treatment effect in the Sponsor's preferred post hoc subgroup was greater – 18g in Lundbeck14 and 10g in Lundbeck23, but the issue remains important, and only bias equivalent to one standard drink would be sufficient to negate the effect in the more reliable of the two studies.) Of course, it is unknown to what extent unblinding occurred, and how much it might have modified drinking behaviour, because the Sponsor did not attempt to quantify this problem, but these back-of-the-enevelpe calculations suggest that the issue could be important.

By itself, this single issue may not be enough to invalidate the apparent efficacy of nalmefene in the pivotal studies, but it adds to concerns about withdrawal bias and the PK effect.

7.6.3. Potential for withdrawal bias in the pivotal studies

Withdrawal bias arises from the non-random withdrawal of subjects from a study, with resulting modification of the pooled results in the remaining subjects. The most common way it manifests is withdrawal of subjects who are doing badly, enriching the remaining cohort with subjects who are doing well. When this is coupled with a higher withdrawal rate in the active group, because of side effects, then a spurious treatment effect will appear because the active group is more enriched with good outcomes than the placebo group.

Withdrawal bias is a common problem in placebo-controlled studies, but it is of particular concern when the withdrawal rate is substantial, when there are a priori reasons to suspect that the withdrawing population is not representative of the overall cohort, and when pessimistic imputation methods produce inferior outcomes to optimistic imputation methods. All three of these concerning factors were present in the pivotal studies.

Firstly, considering the Full Analysis Set, the withdrawal rate in the nalmefene group was close to half (48%) in Lundbeck14, and roughly double that seen in the placebo group (26%). In Lundbeck23, the withdrawal rate in the nalmefene group was less marked (36%) and was closer to that seen in the placebo group (30%). It is of interest, then, that these otherwise similar studies showed considerable differences in the results for their primary endpoints, with much greater effects seen in Lundbeck14, which was positive for its combined co-primary endpoint, than in Lundbeck23, which was negative for its combined co-primary endpoint.

Table 131. Withdrawals by Primary Reason – Lundbeck14 and Lundbeck23

	12014A		12023A	
	PBO n (%)	NMF n (%)	PBO n (%)	NMF n (%)
FAS	289 (100)	290 (100)	326 (100)	329 (100)
Patients completed[a]	213 (74)	152 (52)	229 (70)	212 (64)
Patients withdrawn	76 (26)	138 (48)	97 (30)	117 (36)
Primary Reason				
Adverse events	17 (6)	57 (20)	4 (1)	12 (4)
Lack of efficacy	19 (7)	17 (6)	13 (4)	7 (2)
Non-compliance		7 (2)	3 (1)	7 (2)
Protocol violation	4 (1)	10 (3)	25 (8)	19 (6)
Withdrawal of consent	25 (9)	31 (11)	32 (10)	43 (13)
Lost to follow-up	7 (2)	12 (4)	11 (3)	12 (4)
Other	4 (1)	4 (1)	9 (3)	17 (5)

a Patients with TLFB data at Month 6

Secondly, there are reasons to suspect that withdrawal was non-random. In Study 12014A, the withdrawal rate in the nalmefene group was nearly double that in the placebo group, suggesting that many subjects were experiencing drug-related side effects. These side effects would have created a disincentive to continue in the study, but whether they actually led to withdrawal is likely to have depended on other factors, such as how onerous the subjects were finding the process of trying to limit their alcohol intake. It is plausible that the more onerous the subjects were finding the process, and the more tempted they were to resume unrestricted drinking, the more likely they were to find the side effects too burdensome to continue with. Also, if subjects had to put up with side effects and could simultaneously see that the urge to drink or their actual intake was not being helped by treatment, they would discouraged from staying in the study. This could easily lead to enhanced withdrawal of nalmefene subjects unable to continue low levels of drinking and enrichment of the persisting cohort with nalmefene recipients who were managing their alcohol intake more easily.

Thirdly, pessimistic imputation methods revealed that all of the major outcomes depended critically on what assumptions were made about withdrawing subjects. The figure below shows the co-primary endpoints for each pivotal study, reanalysed with a number of different imputation methods. The most optimistic method (LOCF) effectively paints each withdrawing subject as immune to relapse, locking in the low levels of alcohol intake they have achieved within the study. The most pessimistic method (BOCF) effectively assumes that withdrawing subjects immediately relapse and return to their baseline drinking habits. Results with the pessimistic BOCF imputation were negative across both endpoints in both studies, and in Lundbeck14, the trend for both endpoints was actually in favour of placebo.

The Sponsor proposed a number of mathematical approaches to the imputation process, but none of these can capture the true difference between a continuing patient and a withdrawing patient. Withdrawing patients were, by definition, not studied after their decision to withdraw, so their true drinking habits are unknown.

Figure 42. Sensitivity Analyses – Changes from Baseline to Month 6 in HDDS (days/month) – Lundbeck14 and Lundbeck23 (FAS, Total Population)

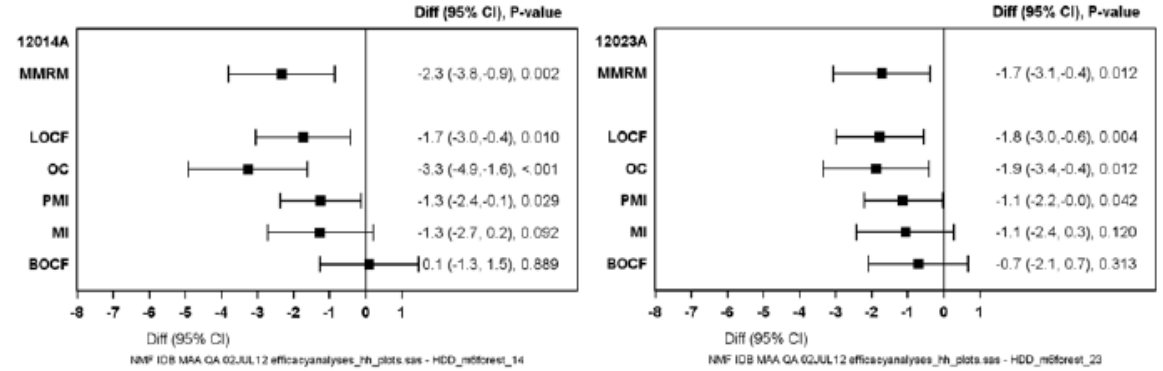
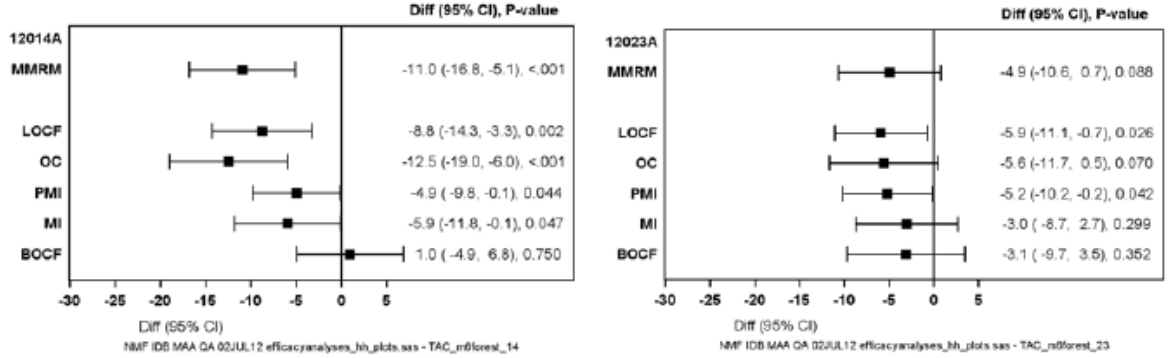


Figure 43. Sensitivity Analyses – Changes from Baseline to Month 6 in TAC (g/day) – Lundbeck14 and Lundbeck23 (FAS, Total Population)



It is impossible to define a foolproof imputation method that eliminates the risk of withdrawal bias, and it is also difficult to estimate the potential contribution of this issue to the apparent treatment effect. If the difference between the two pivotal studies is assumed to be largely attributable to their differing withdrawal rates, however, this gives a very rough idea of how serious the problem might be in relation to the treatment effect. The mean reduction in TAC in Lundbeck14, which suffered from highly imbalanced withdrawal, was 11g. The mean reduction in Lundbeck23, which had less unbalanced withdrawals, was only 5g, a difference in estimated treatment effect of 6g. If withdrawal bias accounted for even half of this difference, just 3g, it would be of major concern, especially in conjunction with the other issues raised. If unblinding spuriously reduced reported intake in the active group by 2-5g, and a PK interaction produced a hidden difference of 3g, and withdrawal bias produced an effect of ~3g in Lundbeck14, then the treatment effect of both studies could be negated.

7.6.4. Modest magnitude of clinical benefit

Even before making any adjustments to account for a possible PK interaction, unblinding and withdrawal bias, the magnitude of the clinical benefit observed in the pivotal studies was small. The two primary endpoints for the two pivotal studies are shown in the table below. Only the MMRM values are directly relevant, because this was the primary prospective analysis technique. Relative to placebo, the number of HDDs per month was reduced by 2.3 and 1.7 in

Lundbeck14 and Lundbeck23, respectively, and the daily alcohol intake was reduced by 11g and 5g, respectively. Lundbeck23 was slightly larger and much less affected by unequal withdrawal, so its results may be more reflective of the true benefit. Also, the reduction in alcohol intake in the third Lundbeck study (shown in the second table below) was 3.5g, so two of the three studies produce estimates for TAC reduction in the range 3.5 to 5g, and the odd study out was the one with the greatest potential for withdrawal bias.

Table 132. Difference to Placebo in HDDs and TAC at Month 6 – (Total Population, FAS) – Lundbeck14 and Lundbeck23

Efficacy Variable Analysis	12014A					12023A				
	n		Mean	95% CI	p-value	n		Mean	95% CI	p-value
	PBO	NMF				PBO	NMF			
Number of HDDs (days/month)										
MMRM	213	152	-2.3	[-3.8; -0.8]	0.002	229	212	-1.7	[-3.1; -0.4]	0.012
ANCOVA, LOCF	289	290	-1.7	[-3.0; -0.4]	0.010	326	329	-1.8	[-3.0; -0.6]	0.004
TAC (g/day)										
MMRM	213	152	-11.0	[-16.8; -5.1]	<0.001	229	212	-5.0	[-10.6; 0.7]	0.088
ANCOVA, LOCF	289	290	-8.8	[-14.3; -3.3]	0.002	326	329	-5.9	[-11.1; -0.7]	0.026

CI = confidence interval

Table 133. Difference to Placebo in HDDs and TAC at Month 6 – (Total Population) – Lundbeck13

Efficacy Variable Treatment Group	Baseline		Adjusted Change from Baseline to Month 6		Difference to PBO		
	N	Mean ± SD	N	Mean ± SE	Mean ± SE	95% CI	p-value
Number of HDDs							
PBO	137	14.7 ± 6.1	110	-8.9 ± 0.6			
NMF	415	15.2 ± 6.1	320	-9.8 ± 0.4	-0.9 ± 0.6	[-2.1; 0.4]	0.160
TAC							
PBO	137	75 ± 41	110	-45.6 ± 2.6			
NMF	415	75 ± 39	320	-49.0 ± 1.6	-3.5 ± 2.9	[-9.2; 2.2]	0.232

Baseline values were based on FAS, OC; changes from baseline and differences to placebo were based on MMRM; FAS, OC values.

The underlying treatment effect for nalmefene 20mg is therefore likely to be similar to the results obtained in Lundbeck23 and Lundbeck13: the number of heavy drinking days might be reduced by 1 or 2 days per month (usually changing these to moderate drinking days, not days of abstinence; the number of days of abstinence was only increased by 0.6 days in Lundbeck23, with a confidence interval that included zero). The daily alcohol intake might be reduced by 3.5 - 5g, or less than half a drink.

The Sponsor's initial ideas of what would constitute a clinically meaningful response are revealed by their prospective definition of response rate, which was specified as a key secondary endpoint. At Month 6 in Lundbeck14, the response rate was 44.3% in the placebo group, compared to 36.9% in the nalmefene group; in other words, the response rate was inferior with active treatment (p=0.039). In Lundbeck23, the response rate at Month 6 was 63.2% in the placebo group and 67.2% in the nalmefene group, consistent with a marginal (4%) better response rate with active treatment (p=0.1833). If the results of Lundbeck23 were reproduced in clinical practice, 25 patients would need to be treated with nalmefene to obtain one response – assuming no methodological problems require an adjustment of the results. If the results of Lundbeck14 were reproduced in clinical practice, patients would be more likely to have a worthwhile response to placebo than to nalmefene.

In the Sponsor's favoured post hoc subgroup, subjects with at least high DRL at Baseline and Randomisation, the apparent treatment effect was better than in the original target population, but still modest: TAC was reduced by ~18g in Lundbeck14 and by ~10g in Lundbeck23, less than one DRL category and much less than originally envisaged in the Sponsor's power calculations.

7.6.5. Negative outcomes in the majority of submitted studies

Of the 8 submitted efficacy studies, two were positive, four were negative, and two provided no evidence on the efficacy of the proposed dose (in one of these inconclusive studies the 20mg dose was not tested and results for other doses produced conflicting outcomes; in the other, there was no control group).

Of the three Lundbeck studies (shown in bold in the table below), one was positive and two were negative for their co-primary endpoint. Only one of two pivotal studies was positive, and as already noted, that positive result could have been subject to withdrawal bias and unblinding, and may also have been compromised by a PK interaction increasing exposure to ethanol in subjects taking nalmefene. The pivotal study that was least susceptible to withdrawal bias was negative.

Table 134. Overall Summary of Endpoints in Submitted Efficacy Studies

Efficacy Study	Target Dose	(Co) Primary Endpoint	Duration	Results for (Co) Primary Endpoint		Comment
12014A	**20mg**	**HDDs & TAC**	**6 Months**	**+**		**High and unequal withdrawals. Sponsor focussed on subjects with high DRL at randomisation**
12023A	**20mg**	**HDDs & TAC**	**6 Months**		**- (negative for 1 of 2 E.P.)**	**Sponsor focussed on subjects with high DRL at randomisation**
12013A	**20mg**	**HDDs & TAC**	**12 Months (Primary analysis at 6 months)**		**- (negative for 2 of 2 E.P.)**	**Sponsor focussed on 13 month results**
Biotie 801	20mg	HDDs	6 Months	+		*Adjustable dose (10-40mg), not fixed as currently proposed*
Biotie 701	20mg	HDDs	7 Months		-	*Adjustable dose (10-40mg), not fixed as currently proposed*
Biotie 299	20mg	HDDs	12 Months		-	*Regular daily dose for 3 months then as needed*
Biotie 399	10mg 40mg	HDDs	4 Months	+ for 40mg	- for 10mg	*Proposed dose not tested*
Biotie 400	20mg	HDDs	12 Months		Not applicable	*Rejected study, no control*

+ - significant
- - not significant
E.P. - endpoint

7.6.6. Inappropriate emphasis on post hoc analyses

Given all of the issues already outlined, including a modest positive result in one of 3 Lundbeck studies and negative results in the other two Lundbeck studies, as well as a number of methodological concerns even in the positive study, it is perhaps unsurprising that the Sponsor's Clinical Overview, Summary of Clinical Efficacy and draft Product Information put very little emphasis on the prospectively identified primary endpoints, which produced

disappointing outcomes, and instead discussed more favourable results in a subgroup identified post hoc.

The statistical principles underlying placebo-controlled studies have been well established over decades and a fundamental feature of any satisfactory placebo-controlled study is that is has clearly defined prospective endpoints and analysis methods. Otherwise, the data can be examined in several different ways and the final presentation of the data can be chosen to suit a particular conclusion. Traditionally, the results of placebo-controlled studies have been reported with p-values, which represent the probability that a between-group difference at least as great as the one observed could arise by chance in the absence of any true underlying group differences. A study is often accepted if its primary endpoint satisfies the threshold $p<0.05$. A natural consequence of this traditional threshold is that about one endpoint in twenty will satisfy the significance threshold because of random variation alone. If twenty unbiased studies are performed with drugs that are no more or less efficacious than placebo, then about one would be expected to produce a spuriously positive result. Similarly, if twenty independent endpoints are assessed in a single unbiased study of a drug that has no effect, about one of those twenty endpoints would be expected to satisfy the $p < 0.05$ criterion. If investigators are allowed to assess multiple statistical endpoints and then pick the winners post hoc, spuriously positive endpoints can be generated from many or most data sets.

Because of this, p-values should be adjusted for the number of endpoints assessed, and the statistical tests should be stated prospectively. Post hoc analyses should, in general, be treated as hypothesis-generating observations and should be reported descriptively; they cannot by definition be regarded as hypothesis-confirming.

These fundamental principles appear to have been known to the Sponsor when the pivotal studies were designed, but abandoned when the results came in. In particular, the Sponsor focussed much of the efficacy discussion on a particular high-risk group that exhibited high-risk drinking not only at baseline but also at randomisation. The PI presented results for this subgroup instead of reporting the results for the primary endpoint – in fact, the actual prospective efficacy results of the pivotal studies were not even mentioned in the PI.

High-risk drinkers certainly belong to a group of interest, but there are many potential ways of identifying high-risk drinkers and the Sponsor had a chance to choose the target population during the design phase of the pivotal studies. The entry criteria for their pivotal studies stipulated that subjects had to have at least moderate DRL at study entry, have ≥6 heavy drinking days (HDDs) in the 4 weeks preceding the Screening Visit, have an average alcohol consumption at medium risk level or above (>40g/day for men; >20g/day for women) in the 4 weeks preceding the Screening Visit, satisfy a DSM diagnosis of alcohol dependence, and not have more than 14 days of abstinence in the preceding 4 weeks, and so on. It was only post hoc that the Sponsor realised that the treatment effect in this target group was relatively modest compared to the non-pharmacological effect of simply signing up to a study and monitoring intake prior to randomisation, and it was only post hoc that the emphasis shifted to the subgroup who still had high or very high DRL at Randomisation (HDAR).

Although it may seem self-evident, in retrospect, that the real target of nalmefene treatment is the set of patients who cannot curtail their drinking after an initial counselling and monitoring process, this was not obvious when the studies were designed – or the studies would have been designed differently. In retrospect, the design of these studies was questionable, and subjects who reduced their drinking prior to randomisation were poor subjects for assessment. The fact that the Sponsor went ahead with this questionable study design suggests that the Sponsor was not yet ready to perform a final confirmatory study.

If the Sponsor had performed an appropriate Phase II study with this questionable methodology, the problem of subjects responding prior to randomisation could have been noticed at that point, and the specific hypothesis that nalmefene is effective in HDAR patients

could then have been tested prospectively in a new study. In fact, because the Sponsor has changed the target population in response to the submitted studies, the 3 Lundbeck studies could be considered Phase II studies that have generated a plausible hypothesis for further testing, rather than as Phase III studies confirming a mature hypothesis in a rigorous prospective fashion.

Note that many other high-risk subgroups could have been identified, including those still at medium or higher DRL at randomisation (given the entry criteria, this would seem a more natural target population to focus on, even after noticing the enrolment effect). The protocol even specified a high-risk subgroup prospectively: subjects with high or very high DRL at Baseline. Without adjustments for using multiple endpoints, results in this subgroup were positive for both efficacy variables in Lundbeck14 but negative for TAC in Lundbeck23 (and hence negative for Lundbeck23 overall, given the requirement for both halves of the co-primary endpoint to be positive). With appropriate adjustments for the use of multiple endpoints, results in this subgroup would be vastly inferior to those reported. The PI does not even mention this prospectively identified high-risk group, only the post hoc HDAR group.

The prospective protocols for the pivotal studies included more than twenty endpoints (precise counting is not possible given that the endpoints were not formally ranked). The two endpoints reported in the draft PI (HDDs and TAC in HDAR subjects) were not among those > 20 endpoints. Even if these two post hoc endpoints were ranked as the first two post hoc endpoints to be considered, this would mean they ranked 21st or lower in the overall list of endpoints.

One method for coping with multiple endpoints is to lower the threshold significance according to the number of endpoints (i.e. insisting on $p<0.025$ for two endpoints, and progressively lower values for higher numbers of endpoints). If the significance threshold were adjusted in this manner, much lower p-values would be needed to consider the HDAR results as significant. Another method is to use a closed hierarchical procedure in which p-values are not tested at all for any lower ranking endpoint once a higher endpoint has failed to achieve significance – this is the approach that the Sponsor announced that they would take for the key secondary endpoint. By either of these two methods, the post hoc results in the HDAR subgroup would not be considered significant even if these endpoints had been announced prospectively. Given that they were announced retrospectively, results in this subgroup should be considered observational in nature and p-values should not be cited at all.

Despite all of this, the draft PI presents unadjusted p-values for these HDAR endpoints without mentioning the actual p-values achieved for the prospective endpoints.

Similar considerations apply to the major Lundbeck supportive study (Lundbeck13), which was negative for its primary endpoint (HDDs at Month 6). Most of the discussion in the Clinical Overview focussed on the Month 13 results, instead, which represents post hoc selection of a favourable timepoint. Also, later timepoints were potentially more affected by withdrawal bias than the primary, Month 6 timepoint. Again, the draft PI fails to mention, even briefly, the actual primary endpoint in this study.

The Sponsor's failure to acknowledge and report the primary results of its own studies in the PI is unacceptable. At a minimum, the draft PI should be rewritten to reflect the actual, prospective primary results of all three Lundbeck studies.

7.6.7. Uncertain impact of nalmefene on psychosocial treatments

Busy clinicians who prescribe nalmefene may be tempted to rely on the pharmacological action of the drug and cut back on psychosocial treatments, which are necessarily time consuming to administer. The design of the pivotal studies was such that this effect could not occur: clinicians were notionally unaware of treatment assignment, so psychosocial treatment measures were applied equally to each treatment group. There is no guarantee that this artificial feature of the study environment would translate into clinical practice, where clinicians would know the subject was on active treatment, and potentially provide less psychosocial care as a result.

The placebo group's overall reduction in alcohol consumption (40g, 54g and 46g across Lundbeck14, Lundbeck23 and Lundbeck13, respectively) was much greater than the apparent treatment effect (11g, 5g and 3.5 g across the three studies, respectively). This indicates that psychosocial measures are much more important than the pharmacological effect of nalmefene.

7.6.8. Overall conclusions on efficacy

The efficacy of nalmefene in the treatment of alcohol dependence remains uncertain. The submitted studies suggest that nalmefene has some efficacy under carefully controlled trial conditions, but they fail to provide conclusive evidence of a clinically significant effect.

Of the three submitted Lundbeck studies, only one (Lundbeck14) was positive according to its primary endpoints; the other two were negative. The single positive study had a withdrawal rate of ~50% in the active group, and the withdrawal rate in the active group was about twice that of the placebo group, creating a real risk of withdrawal bias. In other words, the only positive study was severely compromised, and the less compromised studies were negative. In the most robust Lundbeck study (Lundbeck23) – robust because it had a similar withdrawal rate in the two treatment arms – the effect of nalmefene treatment was clinically modest, amounting to ~5g of alcohol per day (not statistically different from zero grams), and even this modest effect could be negated by plausible adjustments for unblinding and/or for PK interactions between nalmefene and alcohol. The treatment effect in the third study was even smaller (TAC 3.5g/day, 95%CI -9.2 to +.2 g/day, p = 0.232).

In the proposed PI, the Sponsor has emphasized efficacy results in a subgroup identified post hoc, which raises major concerns about the application of statistical tests designed for prospective hypothesis testing. Even in this subgroup, the benefit was modest, especially in the second study, which is likely to have been the more robust of the two pivotal studies. As shown in the table below (copied from the proposed PI), the treatment benefit in this post hoc subgroup amounted to ~18g in the first study and ~10g, or about one standard drink per day in the second study (Lundbeck23). (Although the table title in the proposed PI refers to these results as "Co-primary Endpoints", they cannot be considered primary because analysis of this group was not listed amongst the >20 endpoints mentioned in the prospective protocol.) Even in this post hoc subgroup, the benefit could be considerably less than 10g if appropriate adjustments were made for unblinding or PK interaction effects. Furthermore the only p-values cited in the draft PI are invalid, given that they refer to a subgroup that was identified post hoc, and no adjustment has been made for the use of multiple endpoints.

Table 135. Post Hoc Results Cited in Proposed PI – HDDs and TAC in Patients with a High or Very High DRL at Screening and Randomisation

Endpoint	Study 1 Difference to placebo			Study 2 Difference to Placebo		
	Mean	95% CI	p-value	Mean	95% CI	p-value
HDD (days/month)	-3.7	-5.9;-1.5	<0.001	-2.7	-5.0; -0.3	0.025
TAC (g/day)	-18.3	-26.9;-9.7	<0.001	-10.3	-20.2;-0.5	0.040

8. Clinical safety

8.1. Studies providing evaluable safety data

The following studies provided evaluable safety data:

- 3 efficacy studies performed by Lundbeck assessing nalmefene in the treatment of alcohol dependence (12014A, 12023A, 12013A);
- 5 studies performed by Biotie in alcohol-use disorders (Studies CPH-101-0801, CPH-101-0701, CPH-101-0299, CPH-101-0399, and CPH-101-0400);
- 2 studies in pathological gambling (one by Biotie and one by Somaxon Pharmaceuticals, Inc);
- 1 study in nicotine dependence conducted by Somaxon;
- 17 clinical pharmacology studies conducted by Lundbeck, Biotie, Key Pharmaceuticals, and IVAX Corporation;
- 47 studies in various other indications, such as pruritic conditions, rheumatoid arthritis, and
- interstitial cystitis, conducted by IVAX, Key Pharmaceuticals, or in the context of investigator-sponsored Investigational New Drug (IND) Applications (collectively designated as "the IVAX Studies" by the current Sponsor).

The primary data pool for safety assessment comes from the three Lundbeck studies in Alcohol Dependence. Biotie also made an integrated safety database using pooled the safety data from studies in Alcohol-use Disorders. The other studies, which assessed a range of doses, routes and indications, are not easily pooled.

8.1.1. Pivotal efficacy studies

In the pivotal Lundbeck efficacy studies and the major Biotie studies, the following safety data were collected:

- General adverse events (AEs) were assessed by interviewing subjects at each visit and also noting unscheduled attendances and hospital admissions or abnormal laboratory results, if considered clinically significant.
- AEs were graded by severity and by presumed causal relation to study drug, and coded according to the Medical Dictionary for Regulatory Activities (MedDRA).
- Treatment-emergent adverse events (TEAEs) were defined as AEs with an onset on or after the day of first study-drug intake or, if present since baseline, AEs that had increased in intensity during the study.
- Laboratory tests, including liver function monitoring, creatinine, urea and electrolytes, and haematology monitoring were performed at regular intervals.
- Vital signs were recorded at each visit.
- Subjects underwent regular ECG assessment.

8.1.2. Pivotal studies that assessed safety as a primary outcome

Lundbeck13 (12013A) was initially designed as a Safety study, but had efficacy endpoints added later; it has already been discussed in the Efficacy Section. Safety assessments in this study were similar to the two pivotal studies and were therefore combined in the Alcohol Dependence Pool. No other major studies assessed safety as a primary outcome.

8.1.3. Other studies

Efficacy studies in other indications and clinical pharmacology studies had variable safety monitoring. Nearly all studies recorded adverse events, apart from a couple of initial PK studies performed by earlier sponsors in the 1980s. Most studies incorporated laboratory monitoring, vital sign assessment and ECG monitoring.

Because of patchy and inconsistent approaches to monitoring in several early studies, the primary safety analysis is based on the three Lundbeck studies (Alcohol Dependence Pool), with supportive data from the Biotie efficacy studies (Alcohol-use Disorders Pool). Patient disposition in these two pools is summarised below (All Patients Randomised Set, APRS).

All patients in the active groups of the Lundbeck studies (Alcohol Dependence Pool) took nalmefene 20mg once daily as needed, whereas dosing in the Biotie studies (Alcohol-use Disorders Pool) varied from 5mg to 40mg.

Table 136. Patient Disposition (APRS) – Alcohol Dependence Pool

	PBO		NMF		TOTAL	
	n	%	n	%	n	%
Patients Randomised	824		1173		1997	
Patients Treated	797		1144		1941	
Patients Completed	527	(66.1)	653	(57.1)	1180	(60.8)
Patients Withdrawn	270	(33.9)	491	(42.9)	761	(39.2)
Efficacy Data Sets						
Full Analysis Set	752		1034		1786	

Nalmefene Integrated ST_DI01_POP3 26OCT2011:14:47:29 1001/120/83/147 - TGML/SAD Build Numbers

Table 137. Patient Disposition (APRS) – Alcohol-Use Disorders Pool

	PBO		NMF 5mg		NMF 10mg		NMF 20mg[a]		NMF 40mg	
	n	(%)	n	(%)	n	(%)	n	(%)	n	(%)
Patients randomised (APRS)	361		68		50		453		118	
Patients completed[b,c]	226	(62.6)	46	(67.6)	39	(78.0)	241	(53.2)	83	(70.3)
Patients withdrawn[b]	135	(37.4)	22	(32.4)	11	(22.0)	212	(46.8)	35	(29.7)
Reason										
Adverse event(s)	11	(3.0)	4	(5.9)	6	(12.0)	77	(17.0)	14	(11.9)
Lack of efficacy	24	(6.6)	5	(7.4)	1	(2.0)	7	(1.6)	3	(2.5)
Withdrawal of consent	45	(12.5)	1	(1.5)	2	(4.0)	68	(15.0)	4	(3.4)
Lost to follow-up	47	(13.0)	11	(16.2)	2	(4.0)	49	(10.8)	12	(10.2)
Other	6	(1.7)	1	(1.5)			7	(1.5)	2	(1.7)
Unknown	2	(0.6)					4	(0.9)		
Patients re-randomised in Study CPH-101-0801[d]	27						30			
Patients completed	23	(85.2)					26	(86.7)		
Patients withdrawn	4	(14.8)					4	(13.3)		
Reason										
Adverse event(s)							1	(3.3)		
Other	4	(14.8)					3	(10.0)		
Patients in optional extension in Study CPH-101-0299[e]	39		40				40		37	
Patients completed	14	(35.9)	18	(45.0)			15	(37.5)	19	(51.4)
Patients withdrawn	25	(64.1)	22	(55.0)			25	(62.5)	18	(48.6)
Reason										
Adverse event(s)	2	(5.1)	1	(2.5)			1	(2.5)	2	(5.4)
Other	23	(59.0)	21	(52.5)			24	(60.0)	16	(43.2)

Patients studied in the context of other addiction indications are summarised in the table below.

Table 138. Patient Disposition (APRS) – Studies in Other Addiction Indications

Study CPH-101-0600	PBO		NMF 25mg		NMF 50mg		NMF 100mg	
	n	(%)	n	(%)	n	(%)	n	(%)
Patients randomised	51		52[a]		52		52	
Patients completed	24	47.1	19	36.5	15	28.8	15	28.8
Patients withdrawn	27	52.9	32	61.5	37	71.2	37	71.2
Reason								
Adverse event(s)	3	5.9	12	23.1	17	32.7	21	40.4
Lack of efficacy	3	5.9	1	1.9			1	1.9
Withdrawal of consent	2	3.9	5	9.6	2	3.8	3	5.8
Lost to follow-up	13	25.5	11	21.2	12	23.1	10	19.2
Other	6	11.8	3	5.8	6	11.5	2	3.8

Study SP-N0406	PBO		NMF 20mg		NMF 40mg	
	n	(%)	n	(%)	n	(%)
Patients randomised	74		77		82	
Patients completed	46	62.2	44	57.1	36	43.9
Patients withdrawn	28	37.8	33	42.9	46	56.1
Reason						
Adverse event(s)	7	9.5	23	29.9	22	26.8
Lack of efficacy	2	2.7				
Withdrawal of consent	8	10.8	6	7.8	14	17.1
Lost to follow-up	7	9.5	3	3.9	5	6.1
Other	4	5.4	1	1.4	5	6.1

Study SP-N0408	PBO		NMF 40mg		NMF 80mg	
	n	(%)	n	(%)	n	(%)
Patients randomised	32		32		12	
Patients completed	22	68.8	21	65.6	8	66.7
Patients withdrawn	10	31.2	11	34.4	4	33.3
Reason						
Adverse event(s)	4	12.5	4	12.5	2	16.7
Lack of efficacy			2	6.2		
Withdrawal of consent	5	15.6	3	9.4		
Lost to follow-up					1	8.3
Other	1	3.1	2	6.2	1	8.3

a Information missing for 1 patient

8.2. Patient exposure

Overall, approximately 3,090 patients have received nalmefene at doses up to 100mg/day for up to 52 weeks. These patients include:

- 1,144 patients in 3 studies in patients with alcohol dependence;
- 689 patients in 5 studies in patients with alcohol-use disorders;
- 357 patients in 3 studies in other addiction indications;
- 486 subjects in 17 clinical pharmacology studies in healthy subjects (373 who received oral nalmefene in doses up to 80mg, 113 who received IV nalmefene at doses up to 24mg, and 4 who received IM nalmefene at a dose of 24mg);
- 901 patients in 47 studies in various other indications, such as pruritic conditions, rheumatoid
- arthritis, and interstitial cystitis.

Exposure in the primary safety pool (APTS, All Patients Treated Set) is summarised below, and amounts to 312 patient-years. Study-drug was not taken every day, but only when subjects thought that drinking was imminent. Actual consumption of study drug is summarised in the subsequent table.

Table 139. Overall Time in Study and Exposure (APTS) – Alcohol Dependence Pool

	Treatment Period				Run-out Period					
	PBO		NMF		PBO-PBO		NMF-PBO		NMF-NMF	
	n	PY	n	PY	n	PY	n	PY	n	PY
Time in study	797	391	1144	630	415	33	172	14	171	15
Exposure[a]	784[b]	226	1134	305	378	29	143	7	142	7

Time in study = total number of patient days (patient years [PY]) in the specified period
Exposure = total number of patient days (PY) with IMP intake in the specified period

a Includes only patients with TLFB records
b The APTS includes an additional 5 patients whose TLFB records indicate no IMP intake but who did not return all of their IMP

Table 140. IMP Intake (APTS) – Alcohol Dependence Pool

Treatment Group	Patients N	Summary Statistics	TLFB Days	IMP Days	IMP %
PBO	789	Mean	173	104	62.1
		Min	5	0	0.0
		Max	448	373	100
		Median	168	97	66.7
		p10	35	16	20.0
		p25	139	44	39.3
		p75	178	152	88.7
		p90	364	174	98.2
NMF	1134	Mean	192	98	50.8
		Min	1	1	0.3
		Max	468	387	100
		Median	169	69	48.3
		p10	28	5	9.5
		p25	86	23	21.4
		p75	358	150	81.5
		p90	367	232	96.4

TLFB Days with IMP and % Days with IMP are summary statistics based on individual patient.

In the Alcohol-use Disorders Pool, a total of 689 patients were exposed to nalmefene, amounting to 250 patient-years, including 212 patient-years in which subjects took 20mg or more of nalmefene.

Table 141. Exposure to Nalmefene by Mean daily Dose (APRS) – Alcohol-Use Disorders Pool

	Mean Daily Dose						
	>0 and ≤5mg	>5 and ≤10mg	>10 and ≤20mg	>20 and ≤30mg	>30mg	Any Dose	Percent[a]
Duration (Weeks)	Number of Patients						
	n	n	n	n	n	n	
>0 and ≤1	8	3	26		7	44	6.5
>1 and ≤2	9	5	12		3	29	4.3
>2 and ≤4	12	9	20	5	7	53	7.9
>4 and ≤12	44	26	39	3	19	131	19.4
>12 and ≤24	45	53	30	5	58	191	28.3
>24 and ≤48	38	41	46	13	14	152	22.6
>48 and ≤96	25	20	16	5	8	74	11.0
Any	181	157	189	31	116	674	100.0
Percent[a]	26.9	23.3	28.0	4.6	17.2	100.0	

a Percent of total population with TFLB records

Exposure to nalmefene in studies for other addiction indications is summarised below. This exposure amounts to ~53 patient-years (assuming patients took one dose per day).

Table 142. Exposure (Safety Population) – Studies in Other Addiction Indications

Study CPH-101-0600[a]	PBO	NMF 25mg	NMF 50mg	NMF 100mg
Number of patients	51	51	52	52
Number of days of exposure				
Mean (SD)	75.4 (41.2)	67.4 (47.8)	49.5 (47.9)	44.5 (46.9)
Median [min,max]	87 [1,116]	64 [1,152]	30 [1,120]	21.5 [1,127]
Study SP-N0406[b,c]	**PBO**	**NMF 20mg**	**NMF 40mg**	
Number of patients	73	77	80	
Number of doses				
Mean (SD)	62.8 (28.8)	55.2 (34.2)	51.5 (32.3)	
Median [min,max]	81 [2,94]	79 [1,89]	60 [1,96]	
Study SP-N0408[b,c]	**PBO**		**NMF 40mg**	**NMF 80mg**
Number of patients	32		32	12
Number of doses				
Mean (SD)	66.0 (22.7)		63.5 (30.0)	62.7 (32.1)
Median [min,max]	83 [12,85]		82 [2,86]	82.5 [3,86]

SD = standard deviation

a Safety population is based on the all patients randomised set.
b Exposure for the double-blind titration period and double-blind treatment period.
c Patients received 1 dose of IMP per day.

Exposure in the IVAX studies is somewhat unclear, because mean daily dose was often not recorded. The Summary of Clinical Safety summarises this exposure as follows:

> *Of the 901 patients, 5% received nalmefene 30mg/day to 60mg/day, 5% received 60mg/day to 90mg/day, and 9% received more than 90mg/day. Nine percent of the patients were treated for to 8 weeks (30 to 90mg/day), 4% were treated for 8 to 12 weeks (unknown dose), 2% were treated for 24 to 48 weeks (the majority received >90mg/day), and nearly 4% were treated for 48 weeks (> 90mg/day).*

Exposure in the clinical pharmacology program was limited to single doses or very short treatment periods of about one week. A total of 486 subjects were exposed, 373 who received oral nalmefene in doses up to 80mg, 113 who received IV nalmefene at doses up to 24mg, and 4 who received IM nalmefene at a dose of 24mg. Adverse events reported in this population resembled the profile in larger studies.

8.3. Adverse events

8.3.1. All adverse events (irrespective of relationship to study treatment)

8.3.1.1. Alcohol dependence pool

The overall proportion of patients in the Alcohol Dependence Pool (ADP) with TEAEs was 62.7% in the placebo group, compared to 74.7% in the nalmefene group, an absolute excess of 12%. That is, of the 37.3% of patients who would not be expected to report a TEAE based on the placebo rate, about one third (12% vs 37.3%) reported a TEAE.

A multi-page listing of all TEAEs in the ADP is included. The tables below show TEAEs occurring at higher frequencies: the first table below shows TEAEs that occurred with ≥ 3% incidence in system organ classes (SOCs) that showed a pooled TEAE incidence of ≥ 10%; the subsequent table shows individual TEAEs occurring at an incidence of ≥ 5%; the third table shows TEAEs where the excess in the nalmefene group was statistically significant.

Table 143. TEAEs with ≥ 3% Incidence in SOCs with a ≥ 10% Incidence of TEAEs (APTS) – Alcohol Dependence Pool

SOC Preferred Term	PBO		NMF	
	n	%	n	%
Patients treated	797		1144	
Nervous system disorders	155	19	432	38
Dizziness	44	6	208	18
Headache	66	8	141	12
Somnolence	23	3	59	5
Tremor	11	1	40	3
Disturbance in attention	4	1	30	3
Gastrointestinal disorders	151	19	390	34
Nausea	47	6	253	22
Vomiting	18	2	100	9
Diarrhoea	35	4	43	4
Dry mouth	12	2	34	3
Psychiatric disorders	128	16	342	30
Insomnia	43	5	153	13
Anxiety	27	3	42	4
Sleep disorder	5	1	38	3
Infections and infestations	174	22	220	19
Nasopharyngitis	73	9	107	9
General disorders and administration site conditions	70	9	218	19
Fatigue	37	5	95	8
Asthenia	5	1	33	3
Musculoskeletal and connective tissue disorders	70	9	134	12
Back pain	23	3	37	3
Injury, poisoning and procedural complications	96	12	85	7
Accidental overdose	34	4	25	2
Fall	23	3	17	1

Cross-reference: Table 96

Table 144. Frequent Adverse Events by Preferred Term (APTS) – Alcohol Dependence Pool

Preferred Term	PBO n	PBO (%)	NMF n	NMF (%)
Number of Patients	797		1144	
Patients with TEAEs	500	(62.7)	855	(74.7)
Nausea	47	(5.9)	253	(22.1)
Dizziness	44	(5.5)	208	(18.2)
Insomnia	43	(5.4)	153	(13.4)
Headache	66	(8.3)	141	(12.3)
Nasopharyngitis	73	(9.2)	107	(9.4)
Vomiting	18	(2.3)	100	(8.7)
Fatigue	37	(4.6)	95	(8.3)
Somnolence	23	(2.9)	59	(5.2)

Table 145. Duration and Time to First Onset of TEAES with Fishers Exact Test p-value < 0.05 and an Incidence of ≥ 1% - Alcohol Dependence Pool

	Incidence						Duration (median, days)		Time to First Onset (median, days)	
	PBO			NMF						
Patients Preferred Term	e	n	%	e	n	%	PBO	NMF	PBO	NMF
Patients Treated (APTS)		797			1144					
Nausea	59	47	5.9	470	253	22.1	3.0	3.0	8.0	0.0
Dizziness	63	44	5.5	359	208	18.2	2.5	3.0	1.0	0.0
Insomnia	57	43	5.4	236	153	13.4	8.0	5.0	47.0	4.0
Headache	101	66	8.3	245	141	12.3	3.0	2.0	11.0	7.0
Vomiting	22	18	2.3	139	100	8.7	2.0	1.0	51.0	9.0
Fatigue	44	37	4.6	219	95	8.3	20.0	7.0	17.0	1.0
Somnolence	36	23	2.9	184	59	5.2	10.0	2.0	0.0	1.0
Decreased Appetite	9	9	1.1	179	56	4.9	88.0	9.0	27.0	1.5
Hyperhidrosis	8	8	1.0	71	49	4.3	13.0	4.0	2.5	0.0
Tremor	15	11	1.4	70	40	3.5	3.0	3.0	63.0	23.5
Sleep Disorder	5	5	0.6	43	38	3.3	46.0	4.0	7.0	0.0
Dry Mouth	12	12	1.5	56	34	3.0	8.5	8.0	0.5	1.0
Asthenia	6	5	0.6	40	33	2.9	3.0	2.0	7.0	3.0
Tachycardia	14	8	1.0	41	33	2.9	57.0	3.0	35.0	41.0
Disturbance In Attention	4	4	0.5	78	30	2.6	28.0	3.5	13.0	0.0
Accidental Overdose	43	34	4.3	38	25	2.2	1.0	1.0	32.0	50.0
Malaise	5	2	0.3	105	23	2.0	1.0	2.0	56.5	1.0
Muscle Spasms	0	0	0.0	29	19	1.7		5.0		7.0
Paraesthesia	2	2	0.3	43	19	1.7		3.0	8.0	0.0
Fall	24	23	2.9	17	17	1.5	1.0	1.0	108.0	97.0
Feeling Abnormal	1	1	0.1	29	17	1.5	1.0	5.0	21.0	0.0
Weight Decreased	3	3	0.4	17	16	1.4	98.5	63.0	118.0	96.5
Hypoaesthesia	1	1	0.1	30	15	1.3	22.0	6.0	12.0	0.0
Confusional State	2	2	0.3	24	14	1.2	1.5	2.0	3.5	0.0
Palpitations	1	1	0.1	13	13	1.1	7.0	3.0	5.0	4.0
Restlessness	3	2	0.3	15	13	1.1	56.0	49.0	1.0	1.0
Gastroenteritis	13	13	1.6	7	6	0.5	3.0	6.0	47.0	99.5

e=number of events; n=number of patients; %=frequency relative to the APTS
Medians, among patients with the event in question, are estimated using the Kaplan-Meier method
Longest duration and the shortest time to first onset from first dose are used for each patient within a preferred term

The symptoms in nalmefene recipients that occurred with a clear excess incidence relative to placebo recipients were nausea, dizziness, insomnia, headache, vomiting, fatigue and somnolence. Nasopharyngitis was common in both placebo recipients and nalmefene recipients, reflecting the high incidence of this symptom in the general community. Nausea was particularly common in nalmefene recipients, affecting 22.1%, compared to only 5.9% of placebo recipients. Dizziness was also much more common in nalmefene recipients, affecting 18.2%, compared to 5.5% of placebo recipients. These two symptoms alone could account for a substantial proportion of patients guessing their assigned treatment, particularly if the nausea or dizziness occurred soon after taking a tablet, or if it was absent on days when the subject did not take nalmefene.

A precise temporal accounting of when AEs occurred in relation to dosing was not available, but the Sponsor did present data on the TEAEs that occurred within the first day after treatment with study drug. In the placebo group, nausea occurred in 1.5% of subjects on the first day, and dizziness occurred in 3.0%. In the nalmefene group, nausea occurred in 14.3% of subjects, and dizziness in 13.9%. Overall, nervous system disorders occurred in 7.2% of placebo recipients during the first day of treatment, compared to 23.9% of nalmefene recipients. This does not represent a major tolerability issue, especially for a drug taken on as as-needed basis, but it does represent a probable source of unblinding. Milder symptoms of nervous system disturbance or nausea may have gone unreported.

Table 146. TEAEs with an Incidence of ≥ 3% with an Onset Within 1 Day of First Dose by Preferred Term (APTS) – Alcohol Dependence Pool

	PBO		NMF	
Preferred Term	n	(%)	n	(%)
Number of Patients	797		1144	
Patients with TEAEs	141	(17.7)	467	(40.8)
Nausea	12	(1.5)	164	(14.3)
Dizziness	24	(3.0)	159	(13.9)
Insomnia	11	(1.4)	70	(6.1)
Headache	17	(2.1)	59	(5.2)
Fatigue	12	(1.5)	55	(4.8)
Vomiting	2	(0.3)	39	(3.4)

Most TEAEs were mild or moderate in intensity, but severe TEAEs were more common in nalmefene recipients (14%) than placebo recipients (9.0%). The types of TEAEs rated as severe reflected the overall spectrum of TEAEs and included dizziness, nausea, insomnia, vomiting and headache.

Table 147. Severe TEAEs with an Incidence of ≥ 1% in Either Treatment Group (APTS) – Alcohol Dependence Pool

	PBO		NMF	
Preferred Term	n	(%)	n	(%)
Number of Patients	797		1144	
Patients with Severe TEAEs	72	(9.0)	160	(14.0)
Dizziness	3	(0.4)	27	(2.4)
Nausea	4	(0.5)	26	(2.3)
Insomnia	2	(0.3)	24	(2.1)
Vomiting	0	(0.0)	17	(1.5)
Headache	3	(0.4)	12	(1.0)

8.3.1.2. Other studies

In general, other studies revealed a distribution of AEs similar to that seen in the Alcohol Dependence Pool (ADP). In the Alcohol-use Disorders Pool (AUDP), nausea, insomnia and dizziness were the most common AEs, and occurred more frequently in nalmefene recipients than in placebo recipients.

Table 148. TEAEs with an Incidence of ≥ 5% in Patients Treated with Nalmefene (Safety Population) – Alcohol-Use Disorders Pool

Preferred Term	PBO		NMF 5mg		NMF 10mg		NMF 20mg[a]		NMF 40mg		NMF Total	
	n	%	n	%	n	%	n	%	n	%	n	%
Patients treated	361		68		50		453		118		689	
Patients with TEAEs												
Nausea	39	10.8	10	14.7	11	22.0	142	31.3	26	22.0	189	27.4
Insomnia	48	13.3	15	22.1	11	22.0	128	28.3	35	29.7	189	27.4
Dizziness	32	8.9	20	29.4	16	32.0	114	25.2	30	25.4	180	26.1
Headache	74	20.5	17	25.0	10	20.0	96	21.2	32	27.1	155	22.5
Fatigue	40	11.1	11	16.2	5	10.0	89	19.6	15	12.7	120	17.4
Influenza-like illness	47	13.0	11	16.2	9	18.0	67	14.8	31	26.3	118	17.1
Vomiting NOS	17	4.7	4	5.9	4	8.0	61	13.5	8	6.8	77	11.2
Malaise	4	1.1			6	12.0	49	10.8	1	0.8	56	8.1
Diarrhoea NOS	35	9.7	6	8.8	4	8.0	27	6.0	18	15.3	55	8.0
Sweating increased	10	2.8	2	2.9	2	4.0	43	9.5	3	2.5	50	7.3
Alcoholic hangover	14	3.9	1	1.5	5	10.0	37	8.2	2	1.7	45	6.5
Dry mouth	7	1.9	5	7.4	3	6.0	27	6.0	9	7.6	44	6.4
Back pain	26	7.2	4	5.9	1	2.0	31	6.8	5	4.2	41	6.0
Appetite decreased	3	0.8	5	7.4	5	10.0	24	5.3	3	2.5	37	5.4

NOS = not otherwise specified

Cross-reference: Table 149

a Patients in the nalmefene groups in the flexible-dose studies CPH-101-0801, CPH-101-0701, and CPH-101-0400 are included in the nalmefene 20mg group, although their actual dose may differ.

Studies of nalmefene for other indications were generally listed by study, rather than pooled. The overall pattern of AEs in individual studies was similar to those seen in the major safety pools (ADP and AUDP), with nausea, dizziness, sleep disturbance and headache being the most common problems.

8.3.2. Treatment-related adverse events (adverse drug reactions)

8.3.2.1. Alcohol dependence pool

For all AEs, investigators were required to guess at the causal relationship between study drug and the AE. This process is inherently inaccurate, and may often reflect the investigators' preconceived ideas about the likely side effects of a treatment. Events considered related to treatment ("treatment-related TEAEs") were common in both the placebo and nalmefene groups. In the placebo group, the incidence of mild, moderate and severe treatment-related TEAEs was 17.2%, 14.6% and 4.3%, respectively, according to the multi-page table reproduced. In the nalmefene group, the incidences were 17.2%, 32.0% and 10.5%. The most common severe TEAES that were thought to be related to treatment were dizziness, nausea, insomnia and vomiting, all of which were more common in nalmefene recipients.[21]

[21] The evaluator points out that the sponsor's tabulation of severe related TEAEs was slightly inconsistent.

Table 149. Related, Severe TEAEs with an Incidence of ≥ 1% in Either Treatment Group (APTS) – Alcohol Dependence Pool

	PBO		NMF	
Preferred Term	n	(%)	n	(%)
Number of Patients	797		1144	
Patients with Related, severe TEAEs	33	(4.1)	120	(10.5)
Dizziness	3	(0.4)	27	(2.4)
Nausea	4	(0.5)	26	(2.3)
Insomnia	0	(0.0)	24	(2.1)
Vomiting	0	(0.0)	15	(1.3)

8.3.2.2. Other studies

In the minor studies, a similar approach was taken to identifying AEs that were thought to be related to treatment. The overall pattern of "related" AEs was similar to that seen in the major studies, but this data was not presented in a convenient summary table and the data is not shown in this report. In general, the treatment-related AEs in the minor studies raised no new safety concerns.

8.3.3. Serious adverse events

8.3.3.1. Alcohol dependence pool

Serious Adverse Events (SAEs) are summarised in the table below, which shows SAEs that occurred in more than 1 patient in either treatment group. The overall incidence of SAEs was similar in the two groups (placebo 4.4%, nalmefene 5.0%), and the only SAE substantially more likely to occur in the nalmefene group was alcohol withdrawal syndrome, which was ~7 times more common (placebo 0.1%, nalmefene 0.7%). This is potentially consistent with nalmefene exhibiting efficacy and encouraging alcohol cessation, but the observation is based on a low number of patients.

Compared to the TEAEs, the SAEs did not feature the typical side effects of dizziness, nausea, and headache, indicating that these AEs were usually not serious, and represent a tolerability issue rather than a major safety concern.

Table 150. SAEs in More than 1 Patient in Either Treatment Group by Preferred Term (APTS) – Alcohol Dependence Pool

	PBO		NMF	
Preferred Term	n	(%)	n	(%)
Number of Patients	797		1144	
Patients with Treatment-emergent SAEs	35	(4.4%)	57	(5.0%)
Alcohol withdrawal syndrome	1	(0.1%)	8	(0.7%)
Alcoholism	2	(0.3%)	4	(0.3%)
Fall	1	(0.1%)	3	(0.3%)
Alcohol poisoning	2	(0.3%)	2	(0.2%)
Atrial fibrillation	0	(0.0%)	2	(0.2%)
Depression	0	(0.0%)	2	(0.2%)
Disorientation	0	(0.0%)	2	(0.2%)
Non-cardiac chest pain	0	(0.0%)	2	(0.2%)
Alcohol abuse	2	(0.3%)	1	(0.1%)
Fibula fracture	2	(0.3%)	1	(0.1%)
Pneumonia	2	(0.3%)	1	(0.1%)
Completed suicide	2	(0.3%)	0	(0.0%)
Convulsion	2	(0.3%)	0	(0.0%)
Hypertension	2	(0.3%)	0	(0.0%)
Intentional overdose	2	(0.3%)	0	(0.0%)
Pneumothorax	2	(0.3%)	0	(0.0%)
Pyothorax	2	(0.3%)	0	(0.0%)
Rib fracture	2	(0.3%)	0	(0.0%)
Tibia fracture	2	(0.3%)	0	(0.0%)

Dictionary: MedDRA 13.0

Nalmefene Integrated ST_AE05_POP3_ge2Pat 28OCT2011:15:26:16 1001/120/83/147 - TGML/SAD Build Numbers

SAEs in the run-out period were relatively rare, and did not raise any particular concerns.

Table 151. Run-Out Period: SAEs by Preferred Term (APTS) – Alcohol Dependence Pool

	PBO-PBO		NMF-PBO		NMF-NMF	
Preferred Term	n	(%)	n	(%)	n	(%)
Number of Patients	415		172		171	
Patients with Treatment-emergent SAEs	2	(0.5)	1	(0.6)	0	(0.0)
Alcoholism	0	(0.0)	1	(0.6)	0	(0.0)
Intentional overdose	1	(0.2)	0	(0.0)	0	(0.0)
Multiple injuries	1	(0.2)	0	(0.0)	0	(0.0)
Road traffic accident	1	(0.2)	0	(0.0)	0	(0.0)
Suicidal behaviour	1	(0.2)	0	(0.0)	0	(0.0)

Dictionary: MedDRA 13.0

Nalmefene Integrated ST_AE25_POP2 21OCT2011:15:33:35 1001/120/83/147 - TGML/SAD Build Numbers

8.3.3.2. Other studies

SAEs in the Alcohol-use Disorder Pool (AUDP) are shown in the table below (for SAEs that occurred in more than one patient). No clear pattern of concerning SAEs emerged in these studies. The overall proportion of SAEs was 4% in both treatment groups (placebo and pooled nalmefene groups)

Table 152. SAEs in More than 1 Patient in Either Treatment Group by Preferred Term (APTS) – Alcohol-Use Disorders Pool

Preferred Term	PBO		NMF 5mg		NMF 10mg		NMF 20mg[a]		NMF 40mg		NMF Total	
	n	%	n	%	n	%	n	%	n	%	n	%
Patients treated	361		68		50		453		118		689	
Patients with SAEs												
Depression							3	0.7			3	0.4
Pancreatitis	1	0.3					3	0.7			3	0.4
Ankle fracture							2	0.4			2	0.3
Melaena							2	0.4			2	0.3
Alcohol intoxication acute			1	1.5			1	0.2			2	0.3
Appendicitis							1	0.2	1	0.8	2	0.3
Chest pain	1	0.3					1	0.2	1	0.8	2	0.3

a Patients in the nalmefene groups in the flexible-dose studies CPH-101-0801, CPH-101-0701, and CPH-101-0400 are included in the nalmefene 20mg group, although their actual dose may differ.

SAEs in the other addiction studies raised no specific safety concerns.

Table 153. SAEs – Studies in Other Addiction Indications

Study	Patient	SAE no.	Preferred Term	SAE Term	Treatment	Related?
CPH-101-0600		CPH-1033-02	Pain in jaw; chest pain; reaction to drug excipient; pain in limb	Reaction to dye used in stress test	NMF 100 mg/day	No
		CPH-1056-02	Kidney infection NOS	Kidney infection	Placebo	No
SP-N0406		-	Deep vein thrombosis	Right lower extremity deep venous thrombosis	NMF 20 mg/day	U
		-	Intentional selfinjury	Suicidal gesture	NMF 40 mg/day	No
		-	Appendicitis	Appendicitis	Placebo	No
		-	Suicidal ideation	Suicidal ideation	NMF 20 mg/day	No
		-	Chest pain	Chest pain	NMF 20 mg/day	P
SP-N0408		-	Orthostatic hypotension	Postural hypotensive event	NMF 80 mg/day	No
		-	Hepatic enzyme increased	Elevated liver enzymes	NMF 80 mg/day	Yes

NOS = not otherwise specified; U = unlikely; P = possible

SAEs in the IVAX studies were not clearly tabulated, but 1.6% of nalmefene recipients in those studies (14 patients) reported an SAE, which is a lower incidence than noted in the alcohol studies, possibly reflecting a shorter duration of monitoring in the IVAX studies.

8.3.4. Discontinuation due to adverse events

8.3.4.1. Alcohol dependence pool

TEAEs that led to withdrawal are summarised in the table below (for TEAEs with an incidence ≥ 0.5%) and in the multi-page listing. Overall, withdrawal rates due to AEs were more than twice as common in nalmefene recipients (13.0%) compared to placebo recipients (5.9%). The most common individual disorders leading to withdrawal in the nalmefene group were dizziness (3.1%), nausea (2.6%), fatigue (1.3%) and headache (1.1%), none of which led to any withdrawals in the placebo group.

Table 154. TEAEs with an Incidence of ≥ 0.5% in Either Treatment Group Leading to Withdrawal by Preferred Term (APTS) – Alcohol Dependence Pool

	PBO		NMF	
Preferred Term	n	(%)	n	(%)
Number of Patients	797		1144	
Patients with TEAEs Leading to Withdrawal	47	(5.9)	149	(13.0)
Dizziness	0	(0.0)	36	(3.1)
Nausea	0	(0.0)	30	(2.6)
Fatigue	0	(0.0)	15	(1.3)
Headache	0	(0.0)	13	(1.1)
Insomnia	2	(0.3)	10	(0.9)
Vomiting	1	(0.1)	9	(0.8)
Hyperhidrosis	0	(0.0)	8	(0.7)
Alcohol withdrawal syndrome	3	(0.4)	7	(0.6)
Disturbance in attention	0	(0.0)	7	(0.6)
Depression	7	(0.9)	5	(0.4)

{SS} Sex Specific
Dictionary: MedDRA 13.0
Nalmefene Integrated ST_AE03_pct_POP3 21OCT2011:15:38:16 1001/120/83/147 - TGML/SAD Build Numbers

8.3.4.2. Other studies

Discontinuations in other studies followed a similar pattern, as summarised in the tables below.

Table 155. TEAEs Leading to Withdrawal in ≥ 1% Patients Treated with Nalmefene (all doses) – Alcohol-Use Disorders Pool

	Placebo		NMF 5mg		NMF 10mg		NMF 20mg[a]		NMF 40mg		NMF Total	
Preferred Term	n	%	n	%	n	%	n	%	n	%	n	%
Patients treated	361		68		50		453		118		689	
Patients with TEAEs leading to withdrawal	13	3.6	5	7.4	6	12.0	78	17.2	16	13.6	105	15.2
Nausea	4	1.1	3	4.4	2	4.0	22[b]	4.9	5	4.2	32	4.6
Dizziness	2	0.6	2	2.9	2	4.0	13[b]	2.9	3	2.5	20	2.9
Insomnia	2	0.6	1	1.5	1	2.0	9[b]	2.0	1	0.8	12	1.7
Headache	1	0.3					7	1.5	3	2.5	10	1.5
Vomiting NOS	1	0.3					5	1.1			5	0.7

NOS = not otherwise specified
a Patients in the nalmefene groups in the flexible-dose studies CPH-101-0801, CPH-101-0701, and CPH-101-0400 are included in the nalmefene 20mg group, although their actual dose may differ.
b Details about timing of the event relative to first dose of IMP are not available for 1 or more patients.

Table 156. Adverse Events Leading to Withdrawal – Studies in Other Addiction Indications

Study CPH-101-0600	PBO		NMF 25mg		NMF 50mg		NMF 100mg		NMF Total	
(>2% in NMF Total Group)	n	%	n	%	n	%	n	%	n	%
Patients treated	51		52		52		52		156	
Patients with TEAEs leading to withdrawal	3	5.9	12	23.1	17	32.7	21	40.4	50	32.1
Nausea			6	11.5	9	17.3	6	11.5	21	13.5
Insomnia			6	11.5	8	15.4	4	7.7	18	11.5
Headache	1	2.0	1	1.9	5	9.6	4	7.7	10	6.4
Dizziness			1	1.9	2	3.8	5	9.6	8	5.1
Diarrhoea NOS			1	1.9	4	7.7	2	3.8	7	4.5
Vomiting NOS					3	5.8	3	5.8	6	3.8
Dry mouth			3	5.8	2	3.8	1	1.9	6	3.8
Fatigue			2	3.8	2	3.8	1	1.9	5	3.2
Feeling abnormal	1	2.0	1	1.9			3	5.8	4	2.6
Sweating increased			2	3.8	1	1.9	1	1.9	4	2.6
Study SP-N0406	**PBO**		**NMF 20mg**		**NMF 40mg**				**NMF Total**	
(>2% in NMF Total Group)	n	%	n	%	n	%			n	%
Patients treated	73		77		80				157	
Patients with TEAEs leading to withdrawal	7	9.6	23	29.9	22	27.5			45	28.7
Nausea	1	1.4	6	7.8	5	6.3			11	7.0
Insomnia			6	7.8	3	3.8			9	5.7
Vomiting	1	1.4	3	3.9	2	2.5			5	3.2
Alanine aminotransferase increased			2	2.6	2	2.5			4	2.5
Study SP-N0408	**PBO**				**NMF 40mg**		**NMF 80mg**		**NMF Total**	
(>1 Patient in NMF Total Group)	n	%			n	%	n	%	n	%
Patients treated	32				32		12		44	
Patients with TEAEs leading to withdrawal	4	12.5			4	12.5	2	16.7	6	13.6
Insomnia	1	3.1			1	3.1	2	16.7	3	6.8
Nausea					1	3.1	1	8.3	2	4.5

8.3.5. Deaths

A total of 8 deaths were reported within the Summary of Clinical Safety, including 4 in the ADP and 4 in other studies. Overall, 4 deaths occurred on nalmefene treatment, 3 on placebo, and one death in an early study occurred in a subject for whom treatment allocation was unknown.

Table 157. Deaths (Safety Populations) – All Studies

Study (Indication)	Treatment & Dose	Patient/Sex/ Age (years)	Adverse Event(s)	Time from First/ Last Dose to Adverse Event Onset (Days)	Causality
12014A (alcohol dependence)	PBO as-needed		Completed suicide	33/5[a]	NR
	PBO as-needed		Completed suicide	30/5[a]	POS
12023A (alcohol dependence)	NMF 20mg as-needed		Sudden death	46/3	NR
12013A (alcohol dependence)	NMF 20mg as-needed		Traumatic brain injury	128/15	NR
CPH-101-0701 (alcohol-use disorders)	PBO		Disseminated carcinoma	226/40	UNL
IX-317-003 (interstitial cystitis)	unknown		Metastatic prostatic carcinoma	473/150	NR
IX-302-003 (pruritus)	NMF 1mg/day		Complications due to long-standing post-stroke status and renal failure	32/21	NR
IX-319-003 (rheumatoid arthritis)	NMF 20mg b.i.d		Acute myocardial infarction	4 weeks/0	NR

M = man; IMP = investigational medicinal product; NR = not related; POS = possibly related; UNL = unlikely; b.i.d. = bis in die (twice daily)

a No information is available on IMP intake for the 5 days prior to the event.

8.3.5.1. Alcohol dependence pool

Four deaths occurred in the ADP, as follows:

- 2 suicides – both in placebo recipients
- 1 traumatic brain injury – a patient in the nalmefene group died as the result of a car accident in which he was a passenger
- 1 sudden death – a [information redacted] subject in the nalmefene group. The patient was found dead in bed on Day 46, having been well the evening before. His last known intake of nalmefene was 3 days prior to his death. The patient had reduced his alcohol intake from a high DRL to a low DRL. The patient was a smoker, but had no other relevant medical history and he was on no concomitant medication. No laboratory or ECG abnormalities had been noted during the study. No autopsy was performed and no cause of death was established.

Only one death, a suicide in a placebo recipient, was thought by the investigator to be potentially related to study-drug.

8.3.5.2. Other studies

Four deaths occurred in other studies, as listed below:

- 2 carcinomas – one in a placebo recipient in study CPH-101-0701, and one in a subject receiving unknown treatment in a study of interstitial cystitis, IX-317-003.
- 1 died from complications in the setting of a stroke, seizures and renal failure, after receiving nalmefene at the low dose of 1mg per day in a pruritus study, IX-302-003; the death occurred 21 days after receiving nalmefene, which had only been administered for 11 days.
- 1 acute myocardial infarction in a nalmefene recipient who had been receiving 20mg BD in a study of rheumatoid arthritis, IX-319-03.

None of these deaths was thought to be related to study drug, and there were no consistent patterns across the causes of death. Study of the individual patient narratives did not raise new concerns.

8.4. Laboratory tests

Because subjects in the ADP had at least moderate alcohol intake at baseline, several laboratory parameters had mean values that were outside the reference range, even in the placebo group. Parameters for which this was noted are summarised below.

Table 158. Mean Laboratory Values Above the Upper Limit of Reference Range (APTS) – Alcohol Dependence Pool

Laboratory Test (unit)	ULN	PBO		NMF	
		Range[a] [min;max]	Weeks >ULN[b]	Range[a] [min;max]	Weeks >ULN[b]
MCV (fL)	97	[96.2; 97.5]	BL,52,Last	[96.1; 97.3]	BL,52
Albumin (g/L)	46[c]	[45.3; 46.5]	BL,24	[46.0; 46.4]	All
CDT relative (%)	2.47	[2.50; 2.66]	All	[2.30; 2.61]	BL,12,52,Last
Total cholesterol (mmol/L)	5.18	[5.42; 5.76]	All	[5.27; 5.61]	All
Fasting total cholesterol (mmol/L)	5.18	[5.55; 5.76]	All	[5.30; 5.53]	All
GGT (IU/L)	42	[65.8; 88.2]	All	[46.8; 78.8]	All

ULN = upper limit of the reference range; BL = baseline, Last = last assessment; MVC = mean corpuscular volume; CDT = carbohydrate-deficient transferrin; GGT = γ-glutamyl transferase

Cross-reference: Tables 177 and 179 and Listing 11

a The minimum and maximum mean laboratory values observed at any assessment

b The assessment time points at which the mean laboratory values were above the reference range

c For patients 66 to 90 years of age

In the sections that follow, emphasis is placed on the Alcohol Dependence Pool, where the proposed dose was used for the proposed indication.

8.4.1. Liver function

Involvement in the study led to a reduction in alcohol intake in both the placebo and nalmefene groups, and there were associated improvements in some liver function parameters in both groups. These changes are reflected in the mean levels for GGT, ALAT and ASAT, as summarised in the table below, and in the number of patients with values of potential clinical significance (PCS), as shown in the subsequent table. Values of PCS were more common in the placebo group, reflecting the slightly higher alcohol intake in this group.

Against this background of general improvement in liver function tests (LFTs), individual cases of an adverse hepatic reaction to nalmefene could be difficult to discern. A listing of shifting values (as shown in the third table below) is more informative than mean changes. In general, shifts from non-concerning values to PCS values were similar in both treatment groups, with shifts to concerning GGT values more common in the placebo group, and shifts to concerning ALAT or ASAT slightly more common in the nalmefene group.

Table 159. Changes from Baseline in Mean ALAT, ASAT and GGT Values (IU/L) (APTS) – Alcohol Dependence Pool

Parameter (Unit) & Treatment	Baseline			Change from Baseline			
	n	Mean	SD	Week	n	Mean	SD
GGT (IU/L)							
PBO	768	84.70	127	12	622	2.22	240
				24	526	-8.82	95
				36	116	-6.63	114
				52	111	7.82	106
NMF	1113	78.76	125	12	821	-13.33	91
				24	691	-15.03	82
				36	327	-14.81	92
				52	305	-10.04	54
ALAT (IU/L)							
PBO	768	33.57	24	12	616	-0.14	25
				24	521	0.10	22
				36	116	-0.76	25
				52	110	0.82	23
NMF	1114	33.74	22	12	815	0.21	31
				24	688	-1.57	24
				36	324	-1.72	29
				52	304	-1.59	24
ASAT (IU/L)							
PBO	768	33.09	20	12	615	0.11	22
				24	520	0.15	22
				36	116	2.73	34
				52	110	3.31	22
NMF	1114	32.71	18	12	815	0.97	29
				24	687	-1.07	21
				36	322	-0.93	25
				52	304	-1.28	18

SD = standard deviation; ALAT = S-alanine aminotransferase; ASAT = S-aspartate aminotransferase; GGT = γ-glutamyl transferase

Cross-reference: Tables 177 and 178 and Listing 11

Table 160. Post-baseline PCS High Liver Test and INR Values – Alcohol Dependence Pool

Laboratory Test	PBO			NMF		
	N	n	(%)	N	n	(%)
GGT	704	81	(11.5)	992	74	(7.5)
ALAT	704	17	(2.4)	992	32	(3.2)
ASAT	703	16	(2.3)	992	26	(2.6)
Alkaline phosphatase	704	1	(0.1)	992	1	(0.1)
Bilirubin, Direct	699	4	(0.6)	970	2	(0.2)
INR	695	1	(0.1)	969	5	(0.5)

N = number of patients with assessments; n = number of patients with PCS high values

Cross-reference: Table 187 and Listing 11

Table 161. Shift in PCS Status from Baseline to Worst Assessment of GGT, ALAT and ASAT (APTS) – Alcohol Dependence Pool

		PBO			NMF		
Test	Shift	N	n	%	N	n	%
GGT	non-PCS to PCS high	637	36	5.7	919	32	3.5
	PCS high to non-PCS	65	20	30.8	72	30	41.7
ALAT	non-PCS to PCS high	699	16	2.3	991	31	3.1
	PCS high to non-PCS[a]	3	2	66.7	1	0	0.0
ASAT	non-PCS to PCS high	700	15	2.1	992	26	2.6
	PCS high to non-PCS[a]	1	0	0.0	0	0	0.0

N = number of patients with assessments; n = number of patients with PCS high values
Non-PCS = value does not meet the PCS criteria

Cross-reference: Table 190 and Listing 11

a Patients with PCS high ALAT or ASAT values at baseline were not eligible for inclusion in Studies 12014A, 12023A, and 12013A

Subjects meeting the criteria for suspected drug-induced liver injury (DILI) are summarised below. No nalmefene subjects had concurrent elevation of ASAT or ALAT ≥ 3x upper limit of normal (ULN) and bilirubin ≥ 2x ULN, but this combination was seen in one placebo recipient. The occurrence of ASAT or ALAT ≥3x ULN was similar in the two treatment groups.

Table 162. Summary of Post-baseline Potential Signals of Drug-Induced Liver Injury (APTS) – Alcohol Dependence Pool

Criterion	PBO n (%)	NMF n (%)
Number of patients	693	965
ASAT or ALAT>=3xULN	27 (3.9%)	37 (3.8%)
ASAT or ALAT>=5xULN	6 (0.9%)	15 (1.6%)
ASAT or ALAT>=10xULN	0 (0.0%)	1 (0.1%)
ASAT or ALAT>=20xULN	0 (0.0%)	0 (0.0%)
TBL>=2xULN	2 (0.3%)	3 (0.3%)
AP>=1.5xULN	9 (1.3%)	1 (0.1%)
ASAT or ALAT>=3xULN and TBL>=2xULN	1 (0.1%)	0 (0.0%)
ASAT or ALAT>=3xULN and TBL>=2xULN and AP<=1.5xULN	0 (0.0%)	0 (0.0%)
ASAT or ALAT>=3xULN and (TBL>=2xULN or INR>=1.5xULN) and AP<=1.5xULN	0 (0.0%)	0 (0.0%)

ASAT=Aspartate Aminotransferase, ALAT=Alanine Aminotransferase, TBL=Total Bilirubin, AP=Alkaline Phosphatase, INR=Prothrombin Intl. Normalized Ratio

Overall, this data does not suggest that nalmefene is associated with an increased risk of hepatic disease or DILI.

8.4.2. Kidney function and other clinical chemistry

The Sponsor listed all laboratory values of potential clinical concern in a multi-page table reproduced below; this table includes urea and electrolytes, as well as liver function tests and basic haematology. Abnormalities in renal function and electrolytes were rare, and similar in the two treatment groups.

Post-baseline creatinine levels that were high and in the PCS range were seen in 0.3% of placebo recipients and 0.3% of nalmefene recipients. High urea of PCS was seen in 0.4% of placebo recipients and 0.5% of nalmefene recipients. Overall, there is no evidence of renal toxicity.

Table 163. Post-baseline PCS Laboratory Values (APTS) – Alcohol Dependence Pool

		PBO			NMF	
Parameter (Unit)	n	PCS	(%)	n	PCS	(%)
Number of patients	787			1130		
SUB CATEGORY: ELECTROLYTES						
S-Bicarbonate (mmol/L)						
High	701	0	(0.0)	990	0	(0.0)
Low	701	0	(0.0)	990	0	(0.0)
PCS Criteria: [NALMEFENE] LOW:<=12 mmol/L HIGH:>=38 mmol/L LOW:<=12 mmol/L HIGH:>=38 mmol/L LOW:<=12 mmol/L HIGH:>=38 mmol/L						
S-Calcium (mmol/L)						
High	704	1	(0.1)	992	0	(0.0)
Low	704	0	(0.0)	992	0	(0.0)
PCS Criteria: [NALMEFENE] LOW:<=1.8 mmol/L HIGH:>=3 mmol/L LOW:<=1.8 mmol/L HIGH:>=3 mmol/L LOW:<=1.8 mmol/L HIGH:>=3 mmol/L						
S-Potassium (mmol/L)						
High	704	3	(0.4)	993	2	(0.2)
Low	704	2	(0.3)	993	2	(0.2)
PCS Criteria: [NALMEFENE] LOW:<=3 mmol/L HIGH:>=6 mmol/L LOW:<=3 mmol/L HIGH:>=6 mmol/L LOW:<=3 mmol/L HIGH:>=6 mmol/L						
S-Sodium (mmol/L)						
High	704	5	(0.7)	993	10	(1.0)
Low	704	0	(0.0)	993	1	(0.1)
PCS Criteria: [NALMEFENE] LOW:<=125 mmol/L HIGH:>=155 mmol/L LOW:<=125 mmol/L HIGH:>=155 mmol/L LOW:<=125 mmol/L HIGH:>=155 mmol/L						
SUB CATEGORY: ENDOCRINE, METABOLIC						
S-Albumin (g/L)						
High	704	0	(0.0)	992	0	(0.0)
Low	704	1	(0.1)	992	1	(0.1)
PCS Criteria: [NALMEFENE] LOW:<=27 g/L HIGH: N/A LOW:<=27 g/L HIGH: N/A LOW:<=27 g/L HIGH: N/A						
S-Glucose (mmol/L)						
High	704	28	(4.0)	992	46	(4.6)
Low	704	11	(1.6)	992	11	(1.1)
PCS Criteria: [NALMEFENE] LOW:<=3.5 mmol/L HIGH:>=9 mmol/L LOW:<=3.5 mmol/L HIGH:>=8.1 mmol/L LOW:<=3.5 mmol/L HIGH:>=9 mmol/L LOW:<=3.5 mmol/L HIGH:>=8.1 mmol/L LOW:<=3.5 mmol/L HIGH:>=9 mmol/L LOW:<=3.5 mmol/L HIGH:>=8.1 mmol/L						
S-Prolactin (mIU/L)						
High	703	4	(0.6)	990	3	(0.3)
Low	703	0	(0.0)	990	0	(0.0)
PCS Criteria: [NALMEFENE] LOW: N/A HIGH:>=1350 mIU/L LOW: N/A HIGH:>=1350 mIU/L LOW: N/A HIGH:>=1350 mIU/L						
SUB CATEGORY: HAEMATOLOGY						
B-Basophils/Leukocytes (%)						
High	700	0	(0.0)	986	0	(0.0)
Low	700	0	(0.0)	986	0	(0.0)
PCS Criteria: [NALMEFENE] LOW: N/A HIGH:>=10 % LOW: N/A HIGH:>=10 % LOW: N/A HIGH:>=10 %						
B-Hemoglobin (g/dL)						
High	703	3	(0.4)	992	3	(0.3)
Low	703	7	(1.0)	992	5	(0.5)
PCS Criteria: [NALMEFENE] LOW:Women<=9.5 g/dL - Men<=11.5 g/dL HIGH:Women>=16.5 g/dL - Men>=18.5 g/dL LOW:Women<=9.5 g/dL - Men<=11.5 g/dL HIGH:Women>=16.5 g/dL - Men>=18.5 g/dL LOW:Women<=9.5 g/dL - Men<=11.5 g/dL HIGH:Women>=16.5 g/dL - Men>=18.5 g/dL						
B-Leukocytes (10E9/L)						
High	703	4	(0.6)	991	4	(0.4)
Low	703	2	(0.3)	991	4	(0.4)
PCS Criteria: [NALMEFENE] LOW:<=2.8 10E9/L HIGH:>=16 10E9/L LOW:<=2.8 10E9/L HIGH:>=16 10E9/L LOW:<=2.8 10E9/L HIGH:>=16 10E9/L						
B-Lymphocytes/Leukocytes (%)						
High	700	0	(0.0)	986	0	(0.0)
Low	700	8	(1.1)	986	19	(1.9)
PCS Criteria: [NALMEFENE] LOW:<=10 % HIGH:>=75 % LOW:<=10 % HIGH:>=75 % LOW:<=10 % HIGH:>=75 %						
B-Mean Corpuscular Volume (fL)						
High	697	2	(0.3)	973	5	(0.5)
Low	697	0	(0.0)	973	0	(0.0)
PCS Criteria: [NALMEFENE] LOW:<=0.8 x LLN fL HIGH:>=1.2 x ULN fL LOW:<=0.8 x LLN fL HIGH:>=1.2 x ULN fL LOW:<=0.8 x LLN fL HIGH:>=1.2 x ULN fL						
B-Monocytes/Leukocytes (%)						

Table 163 (continued). Post-baseline PCS Laboratory Values (APTS) – Alcohol Dependence Pool

Parameter (Unit)		PBO n	PBO PCS	PBO (%)	NMF n	NMF PCS	NMF (%)
	High	700	7	(1.0)	986	28	(2.8)
	Low	700	0	(0.0)	986	0	(0.0)
PCS Criteria: [NALMEFENE] LOW: N/A HIGH:>=15 % LOW: N/A HIGH:>=15 % LOW: N/A HIGH:>=15 %							
B-Neutrophils/Leukocytes (%)							
	High	700	8	(1.1)	986	11	(1.1)
	Low	700	0	(0.0)	986	2	(0.2)
PCS Criteria: [NALMEFENE] LOW:<=20 % HIGH:>=85 % LOW:<=20 % HIGH:>=85 % LOW:<=20 % HIGH:>=85 %							
B-Platelet (10E9/L)							
	High	699	5	(0.7)	985	0	(0.0)
	Low	699	2	(0.3)	985	5	(0.5)
PCS Criteria: [NALMEFENE] LOW:<=75 10E9/L HIGH:>=600 10E9/L LOW:<=75 10E9/L HIGH:>=600 10E9/L LOW:<=75 10E9/L HIGH:>=600 10E9/L							
Eosinophils/Leukocytes (%)							
	High	700	11	(1.6)	986	11	(1.1)
	Low	700	0	(0.0)	986	0	(0.0)
PCS Criteria: [NALMEFENE] LOW: N/A HIGH:>=10 % LOW: N/A HIGH:>=10 % LOW: N/A HIGH:>=10 %							
S-Prothrombin Intl. Normalized Ratio							
	High	695	1	(0.1)	969	5	(0.5)
	Low	695	0	(0.0)	969	0	(0.0)
PCS Criteria: [NALMEFENE] LOW: N/A HIGH:>=2 % LOW: N/A HIGH:>=2 % LOW: N/A HIGH:>=2 %							
SUB CATEGORY: INFECTION							
S-C-Reactive Protein (mg/L)							
	High	704	14	(2.0)	992	23	(2.3)
	Low	704	0	(0.0)	992	0	(0.0)
PCS Criteria: [NALMEFENE] LOW: N/A HIGH:>=25 mg/L LOW: N/A HIGH:>=25 mg/L LOW: N/A HIGH:>=25 mg/L							
SUB CATEGORY: KIDNEY							
S-Creatinine (µmol/L)							
	High	704	2	(0.3)	992	3	(0.3)
	Low	704	0	(0.0)	992	0	(0.0)

Parameter (Unit)		PBO n	PBO PCS	PBO (%)	NMF n	NMF PCS	NMF (%)
PCS Criteria: [NALMEFENE] LOW:<=0.5 x LLN µmol/L HIGH:>=1.5 x ULN µmol/L LOW:<=0.5 x LLN µmol/L HIGH:>=1.5 x ULN µmol/L LOW:<=0.5 x LLN µmol/L HIGH:>=1.5 x ULN µmol/L							
S-Urea (mmol/L)							
	High	704	3	(0.4)	992	5	(0.5)
	Low	704	0	(0.0)	992	0	(0.0)
PCS Criteria: [NALMEFENE] LOW: N/A HIGH:>=11 mmol/L LOW: N/A HIGH:>=11 mmol/L LOW: N/A HIGH:>=11 mmol/L							
SUB CATEGORY: LIPIDS							
S-Cholesterol (mmol/L)							
	High	704	167	(23.7)	992	230	(23.2)
	Low	704	0	(0.0)	992	0	(0.0)
PCS Criteria: [NALMEFENE] LOW: N/A HIGH:>=6.2 mmol/L LOW: N/A HIGH:>=7.8 mmol/L LOW: N/A HIGH:>=6.2 mmol/L LOW: N/A HIGH:>=7.8 mmol/L LOW: N/A HIGH:>=6.2 mmol/L LOW: N/A HIGH:>=7.8 mmol/L							
S-Triglycerides (mmol/L)							
	High	704	40	(5.7)	992	60	(6.0)
	Low	704	0	(0.0)	992	0	(0.0)
PCS Criteria: [NALMEFENE] LOW: N/A HIGH:>=5.65 mmol/L LOW: N/A HIGH:>=4.2 mmol/L LOW: N/A HIGH:>=5.65 mmol/L LOW: N/A HIGH:>=4.2 mmol/L LOW: N/A HIGH:>=5.65 mmol/L LOW: N/A HIGH:>=4.2 mmol/L							
SUB CATEGORY: LIVER							
S-Alanine Aminotransferase (IU/L)							
	High	704	17	(2.4)	992	32	(3.2)
	Low	704	0	(0.0)	992	0	(0.0)
PCS Criteria: [NALMEFENE] LOW: N/A HIGH:>=3 x ULN IU/L LOW: N/A HIGH:>=3 x ULN IU/L LOW: N/A HIGH:>=3 x ULN IU/L							
S-Alkaline Phosphatase (IU/L)							
	High	704	1	(0.1)	992	1	(0.1)
	Low	704	0	(0.0)	992	0	(0.0)
PCS Criteria: [NALMEFENE] LOW: N/A HIGH:>=3 x ULN IU/L LOW: N/A HIGH:>=3 x ULN IU/L LOW: N/A HIGH:>=3 x ULN IU/L							
S-Aspartate Aminotransferase (IU/L)							
	High	703	16	(2.3)	992	26	(2.6)
	Low	703	0	(0.0)	992	0	(0.0)

Table 163 (continued). Post-baseline PCS Laboratory Values (APTS) – Alcohol Dependence Pool

		PBO			NMF	
Parameter (Unit)	n	PCS	(%)	n	PCS	(%)
PCS Criteria: [NALMEFENE] LOW: N/A HIGH:>=3 x ULN IU/L LOW: N/A HIGH:>=3 x ULN IU/L LOW: N/A HIGH:>=3 x ULN IU/L						
S-Bilirubin (µmol/L)						
High	703	2	(0.3)	990	3	(0.3)
Low	703	0	(0.0)	990	0	(0.0)
PCS Criteria: [NALMEFENE] LOW: N/A HIGH:>=2 x ULN µmol/L LOW: N/A HIGH:>=2 x ULN µmol/L LOW: N/A HIGH:>=2 x ULN µmol/L						
S-Bilirubin, Direct (µmol/L)						
High	699	4	(0.6)	970	2	(0.2)
Low	699	0	(0.0)	970	0	(0.0)
PCS Criteria: [NALMEFENE] LOW: N/A HIGH:>=12 µmol/L LOW: N/A HIGH:>=12 µmol/L LOW: N/A HIGH:>=12 µmol/L						
S-Gamma Glutamyl Transferase (IU/L)						
High	704	81	(11.5)	992	74	(7.5)
Low	704	0	(0.0)	992	0	(0.0)
PCS Criteria: [NALMEFENE] LOW: N/A HIGH:>=200 IU/L LOW: N/A HIGH:>=200 IU/L LOW: N/A HIGH:>=200 IU/L						

Nalmefene Integrated ST_LA03_POP3 13OCT2011:16:46:57 1001/120/83/147 - TGML/SAD Build Numbers

8.4.3. Glucose and lipids

Abnormalities of lipids are common in subjects with excess alcohol intake, and mean cholesterol levels were outside the normal range in both the treatment groups. Lipid values of PCS were seen at a similar incidence in the two treatment groups. There was a very slight excess of high glucose levels in the nalmefene group, which seems unlikely to represent a drug effect.

Table 164. Post-baseline PCS Glucose and Lipid Values (APTS) – Alcohol Dependence Pool

Laboratory Test	PBO			NMF		
	N	n	(%)	N	n	(%)
PCS high glucose						
Fasting and non-fasting	704	28	(4.0)	992	48	(4.8)
Fasting	411	13	(3.2)	619	34	(5.5)
PCS low glucose						
Fasting and non-fasting	704	11	(1.6)	992	11	(1.1)
Fasting	411	4	(1.0)	619	6	(1.0)
PCS high total cholesterol						
Fasting and non-fasting	704	167	(23.7)	992	230	(23.2)
Fasting	411	148	(36.0)	618	212	(34.3)
PCS high triglycerides						
Fasting and non-fasting	704	40	(5.7)	992	60	(6.0)
Fasting	411	27	(6.6)	618	38	(6.1)

N = number of patients with assessments; n = number of patients with PCS high values

Cross-reference: Tables 187 and 188 and Listing 11

8.4.4. Haematology

Haematology results for the Alcohol Dependence Pool and other data pools were not initially reported in a convenient format, but the Sponsor provided the table below in response to a specific request for a summary table. The request read as follows: "The Sponsor should also provide a summary table for haematological indices of potential clinical concern." The provided table only refers to mean values in each treatment group. Mean values would be relatively insensitive to the potential occurrence of individual values of clinical concern, so the provided table does not directly address the first-round request.

Some haematological indices were included above in Table 158, but this table also refers to mean values, instead of individual values of concern.

Some haematological indices were included in the Sponsor's "Table 187 Summary of Post-baseline PCS Laboratory Values (APTS) – Alcohol Dependence Pool" from the Summary of

Clinical Safety, but this was a multi-page table inconvenient for reproduction here. A review of the haematological sections of that table did not suggest that nalmefene poses a significant risk of haematological toxicity.

Table 165. Haematology Variables: Mean Values Above or Below the Reference Range – Alcohol Dependence Pool

Laboratory Test (unit)	Reference Range	PBO		NMF	
		Range[a] [min;max]	Weeks out of Reference Range[b]	Range[a] [min;max]	Weeks out of Reference Range[b]
Haemoglobin (g/dL)	≥11 – ≤16.1[d]	[14.5; 14.7]	none	[14.6; 14.7]	none
MCV (fL)	≥79 – ≤97[f]	[96.2; 97.5]	BL, 52, Last	[96.1; 97.3]	BL, 52
Total Leukocytes (10E9/L)	≥4.1 – ≤12.3	[6.7; 7.3]	none	[6.9; 6.7]	none
- Neutrophils/Leukocytes (%)	≥40.9 – ≤77	[59.0; 63.3]	none	[60.4; 62.2]	none
- Eosinophils/Leukocytes (%)	0 - – ≤6[e]	[2.2; 2.7]	none	[2.2; 2.2]	none
- Basophils/Leukocytes (%)	0 - – ≤2.4[c]	[0.40; 0.45]	none	[0.42; 0.45]	none
- Lymphocytes/Leukocytes (%)	≥15.5 – ≤46.6	[27.2; 30.6]	none	[28.1; 30.0]	none
- Monocytes/Leukocytes (%)	≥3.1 – ≤12.5	[6.9; 7.4]	none	[6.9; 7.4]	none
Platelets (10E9/L)	≥140 – ≤450	[228; 238]	none	[223; 227]	none

[a] The minimum and maximum mean laboratory values observed at any assessment
[b] The assessment time points at which the mean laboratory values were above or below the reference range
[c] For subjects 1 to 110 years old
[d] For women 66 to 110 years old
[e] For subjects 19 to 110 years old
[f] For men

8.4.5. Laboratory tests in minor studies

Laboratory results in the AUDP were reported individually for each study, and were not pooled by the Sponsor for convenient reproduction in this report. There were no substantial or clinically relevant differences in the incidence of abnormal laboratory values between nalmefene recipients and placebo recipients.

In studies performed for other indications, no specific safety signals arose from laboratory monitoring.

In the clinical pharmacology program, laboratory monitoring did reveal any significant safety issues. For a small subset of these studies (5 of 17 studies), the abnormal laboratory results are summarised below.

Table 166. PCS Laboratory Values in the Nalmefene Group (ASTS) – Studies 12417A, 12393A, 13505A and 21

	Study				
	12417A[a]	12393A[b]	13513A	13505A	21[c]
Laboratory Test	N = 24	N = 6	N = 46	N = 13	N = 24
Haematology					
Eosinophils/leukocytes	1		1	1	5
Lymphocytes/leukocytes	3			1	
Monocytes/leukocytes		1		1	
Neutrophils/leukocytes	1				
Erythrocytes	5		1		
Haematocrit	5		1	2	3
Haemoglobin	2				
Platelets	7				
Red blood cell count					2
Liver					
Alanine aminotransferase					8
Aspartate aminotransferase	1				10
Alkaline phosphatase					7
Bilirubin	7			2	2
Metabolic					
Glucose	7	1		1	7
Cholesterol				1	
Other					
Creatine phosphokinase					
Bicarbonate					
Blood urea nitrogen					1
Potassium	1			1	2
Protein					
Uric acid					
Casts (urine analysis)					

a Study in subjects with mild, moderate, or no hepatic impairment. The vast majority of patients who had PCS laboratory values also had moderate hepatic impairment.
b One subject had occult blood in urine.
c Study in subjects with mild, moderate, severe, or no hepatic impairment. The vast majority of patients who had PCS laboratory values also had mild, moderate, or severe hepatic impairment.

8.4.6. Electrocardiograph

8.4.6.1. Alcohol dependence pool

As summarised in the tables below, mean changes in ECG parameters were minor, and they were similar in the placebo and nalmefene groups. The incidence of ECG parameters of potential clinical significance was also low and similar in the two groups. The data do not suggest that nalmefene has any important adverse effects on the ECG or cardiac function.

Table 167. Changes from Baseline in ECG Parameters (APTS) – Alcohol Dependence Pool

Parameter (Unit)	Treatment Group	Week	n	Mean	SD	Min	Max	Median
PR Duration Mean (msec)	PBO	12	616	1.35	13.35	-46.00	57.00	2.00
		24	523	0.32	14.25	-46.00	69.00	0.00
		36	115	2.12	14.07	-45.00	46.00	2.00
		52	110	2.04	16.35	-37.00	77.00	2.00
		Last	691	-0.19	14.04	-46.00	77.00	0.00
	NMF	12	819	1.17	13.29	-48.00	56.00	1.00
		24	687	1.30	14.03	-49.00	79.00	1.00
		36	323	1.32	14.99	-60.00	63.00	1.00
		52	303	-0.11	16.01	-67.00	59.00	0.00
		Last	976	0.62	14.56	-67.00	79.00	1.00
QRS Duration (msec)	PBO	12	616	0.03	7.42	-60.00	22.00	0.00
		24	523	0.80	7.17	-22.00	33.00	1.00
		36	115	2.52	6.32	-18.00	17.00	2.00
		52	110	3.14	6.64	-14.00	24.00	4.00
		Last	691	1.00	7.19	-24.00	33.00	1.00
	NMF	12	819	0.68	7.31	-41.00	30.00	1.00
		24	687	1.20	7.62	-39.00	29.00	1.00
		36	323	1.76	7.47	-27.00	26.00	2.00
		52	303	2.96	7.37	-18.00	25.00	2.00
		Last	976	1.49	7.44	-39.00	29.00	1.00
QT Duration Mean (msec)	PBO	12	610	0.42	25.67	-86.00	80.00	-1.00
		24	516	-1.03	25.93	-78.00	73.00	-0.50
		36	115	2.27	27.44	-52.00	77.00	0.00
		52	110	0.53	27.48	-65.00	93.00	-1.50
		Last	682	-0.57	26.06	-101.0	93.00	-1.00
	NMF	12	812	-2.47	25.24	-111.0	75.00	-3.00
		24	680	-3.01	25.33	-101.0	91.00	-3.00
		36	323	-2.62	25.50	-83.00	80.00	-3.00
		52	302	-4.17	26.73	-74.00	76.00	-3.00
		Last	968	-3.41	25.98	-111.0	91.00	-3.00
QTcB (msec)	PBO	12	610	1.34	21.76	-67.00	67.00	2.00
		24	516	2.59	22.17	-74.00	75.00	2.00
		36	115	5.50	20.19	-41.00	50.00	6.00
		52	110	4.04	20.92	-55.00	48.00	8.00
		Last	682	2.52	21.38	-74.00	65.00	2.00
	NMF	12	812	2.05	21.11	-66.00	84.00	1.00
		24	680	1.77	21.99	-62.00	97.00	1.00
		36	323	1.37	21.52	-81.00	68.00	2.00
		52	302	4.51	23.57	-62.00	70.00	4.00
		Last	968	2.36	22.58	-62.00	78.00	2.00
QTcF (msec)	PBO	12	610	1.07	17.09	-58.00	60.00	1.00

Table 168. Changes from Baseline in ECG Parameters (APTS) – Alcohol Dependence Pool

Parameter (Unit)	Treatment Group	Week	n	Mean	SD	Min	Max	Median
		24	516	1.41	17.63	-46.00	51.00	0.50
		36	115	4.36	15.95	-29.00	45.00	3.00
		52	110	2.82	16.49	-48.00	43.00	2.50
		Last	682	1.46	16.63	-48.00	51.00	1.00
	NMF	12	812	0.42	16.78	-48.00	78.00	1.00
		24	680	0.06	17.16	-57.00	78.00	0.00
		36	323	-0.02	17.47	-70.00	56.00	0.00
		52	302	1.48	18.40	-64.00	50.00	1.50
		Last	968	0.32	17.96	-68.00	78.00	0.00
RR Duration Mean (msec)	PBO	12	616	-3.35	146.44	-406.0	628.00	-4.50
		24	523	-14.84	142.86	-477.0	541.00	-16.00
		36	115	-13.16	150.72	-318.0	423.00	-10.00
		52	110	-16.14	149.47	-334.0	384.00	-23.00
		Last	691	-11.84	146.11	-578.0	541.00	-17.00
	NMF	12	819	-19.07	139.16	-584.0	460.00	-13.00
		24	687	-20.01	141.57	-438.0	524.00	-18.00
		36	323	-18.80	138.04	-484.0	414.00	-19.00
		52	303	-38.98	154.08	-654.0	434.00	-38.00
		Last	976	-24.63	143.86	-654.0	468.00	-16.00
Ventricular Rate Mean (bpm)	PBO	12	616	0.08	12.05	-39.00	45.00	0.00
		24	523	0.99	12.16	-51.00	49.00	1.00
		36	115	1.04	13.17	-45.00	37.00	1.00
		52	110	1.04	13.05	-50.00	32.00	2.00
		Last	691	1.01	12.69	-50.00	73.00	1.00
	NMF	12	819	1.84	11.86	-33.00	55.00	1.00
		24	687	1.81	11.84	-44.00	41.00	2.00
		36	323	1.38	11.21	-36.00	41.00	2.00
		52	303	3.13	12.40	-37.00	36.00	3.00
		Last	976	2.16	11.97	-44.00	55.00	1.00

Week: Last=Last post-baseline observation

Table 169. Post-baseline PCS ECG Parameters (APTS) – Alcohol Dependence Pool

	PBO			NMF		
Parameter (Unit)	n	PCS	(%)	n	PCS	(%)
Number of patients	700			987		
PR Duration Mean (msec)						
High	696	1	(0.1)	983	3	(0.3)
Low	696	0	(0.0)	983	0	(0.0)
QRS Duration (msec)						
High	696	6	(0.9)	983	2	(0.2)
Low	696	0	(0.0)	983	0	(0.0)
QT Duration Mean (msec)						
High	686	0	(0.0)	975	1	(0.1)
Low	686	0	(0.0)	975	0	(0.0)
QTcB (msec)						
High	686	5	(0.7)	975	14	(1.4)
Low	686	0	(0.0)	975	0	(0.0)
QTcF (msec)						
High	686	1	(0.1)	975	4	(0.4)
Low	686	0	(0.0)	975	0	(0.0)
RR Duration Mean (msec)						
High	696	33	(4.7)	983	46	(4.7)
Low	696	2	(0.3)	983	1	(0.1)
Ventricular Rate Mean (bpm)						
High	696	2	(0.3)	983	1	(0.1)
Low	696	10	(1.4)	983	12	(1.2)

8.4.6.2. Thorough QTc study

The thorough QTc study (Study BTT31-CD005) showed that nalmefene does not cause clinically significant QT prolongation at doses of 20mg daily or 80mg daily. Some statistical differences between nalmefene and placebo were noted at some time points, but the differences were <10ms. The maximum mean change from baseline in QTcI was 5.4ms (90% CI [max of lower bounds; max of upper bounds] = [1.52; 9.37]) for 20mg/day nalmefene and 5.6ms (90% CI [max lower; max upper] = [1.61; 9.52]).

8.4.6.3. Other studies

No clinically important ECG patterns were noted in the AUDP, or in studies performed for other indications. ECG changes in the clinical pharmacology studies were generally minor, apart from one SAE in a nalmefene recipient in Study 21; this patient developed reversible ischaemic changes in his ECG, which resolved without treatment.

8.5. Seizures

Seizures were rare in the major studies, and there was no evidence of any important differences between nalmefene and placebo recipients, as shown in the table below.

Table 170. Convulsions: TEAEs by SOC and Preferred Term (APTS) – Alcohol Dependence Pool

	PBO		NMF	
System Organ Class Preferred Term	n	(%)	n	(%)
Number of Patients	797		1144	
Men	570	(71.5)	843	(73.7)
Women	227	(28.5)	301	(26.3)
Patients with TEAEs	4	(0.5)	5	(0.4)
NERVOUS SYSTEM DISORDERS	4	(0.5)	5	(0.4)
Alcoholic Seizure	1	(0.1)	1	(< 0.1)
Convulsion	2	(0.3)	3	(0.3)
Epilepsy	1	(0.1)	1	(< 0.1)

Events captured using the MedDRA SMQ: Convulsions
Patients are only counted once for each SOC/Preferred Term
Dictionary: MedDRA 13.0
Nalmefene Integrated ST_AE06_POP3_conv 28OCT2011:15:23:51 1001/120/83/147 - TGML/SAD Build Numbers

8.6. Vital signs

8.6.1. Alcohol dependence pool

Mean changes in vital signs were minor, and did not lead to mean values outside the reference range. Instead, there was a trend to normalisation of blood pressure considered likely to be related to a reduced alcohol intake. At baseline, the prevalence of diastolic and systolic blood pressures above the reference ranges was 36% and 29%, respectively, in the placebo group. In the nalmefene group, the prevalances were 41% and 26%, respectively. These proportions decreased over time. For systolic blood pressure, the decrease from baseline to Week 24 in the proportion of patients with high values was 5% in both treatment groups.

Individual vital signs of PCS are summarised in the table below. No major differences were noted between the placebo and nalmefene groups.

Table 171. Summary of Post-baseline PCS Vital Signs and Weight Changes (APTS) – Alcohol Dependence Pool

	PBO			NMF		
Parameter (Unit)	n	PCS	(%)	n	PCS	(%)
Number of patients	787			1128		
Diastolic Blood Pressure Sitting (mmHg), after 5 minutes						
High	786	23	(2.9)	1128	45	(4.0)
Low	786	3	(0.4)	1128	3	(0.3)
Pulse Rate Sitting (bpm), after 5 minutes						
High	786	3	(0.4)	1128	8	(0.7)
Low	786	6	(0.8)	1128	6	(0.5)
Systolic Blood Pressure Sitting (mmHg), after 5 minutes						
High	786	19	(2.4)	1128	26	(2.3)
Low	786	7	(0.9)	1128	7	(0.6)
Weight (kg)						
High	786	49	(6.2)	1128	97	(8.6)
Low	786	56	(7.1)	1128	94	(8.3)

8.6.2. Other studies

Vital signs were reported in individual studies within the Alcohol-use Disorders pool, Phase I studies, and studies performed for other indications, but the data was not pooled in a convenient summary format. A review of the individual studies described in this report did not raise any concerns about the effect of nalmefene on vital signs.

8.7. Post marketing experience

No data are available related to the post marketing use of oral nalmefene at or near the proposed dose. Previous parenteral use of nalmefene of reversal of opioid overdose has not raised any significant safety concerns, but the doses used are much lower than that proposed for use in Alcohol Dependence.

8.8. Safety issues with the potential for major regulatory impact

8.8.1. Liver toxicity

There is no evidence that nalmefene causes serious liver toxicity. On average, liver function tests improved in the pivotal studies, with marginal superiority in the nalmefene groups compared to placebo.

8.8.2. Haematological toxicity

There is no evidence that nalmefene is associated with a significant risk of haematological toxicity.

8.8.3. Serious skin reactions

Overall, in the Alcohol Dependence Pool, skin and subcutaneous reactions were seen more often in placebo recipients (6.2%) than nalmefene recipients (4.5%). There was no evidence of an increased incidence of serious skin reactions.

8.8.4. Cardiovascular safety

Based on vital signs and a through QTc study, there is no evidence that nalmefene poses serious safety concerns in relation to the cardiovascular system.

8.8.5. Unwanted immunological events

In the Alcohol Dependence Pool, the AE of "Drug Hypersensitivity" was reported in 9 placebo recipients (1.1%) and 7 nalmefene recipients (0.6%). Overall, there was no evidence that the use of nalmefene is associated with a significant excess of unwanted immunological events.

8.9. Other safety issues

8.9.1. Safety in special populations

When AEs were assessed in subgroups defined on the basis of gender and age, the overall safety profile was similar to that seen in the general population, but a number of AEs related to tolerability were more common in women than in men, for both the placebo and nalmefene groups. These included nausea, dizziness, headache, vomiting, influenza, irritability, and malaise.

Nalmefene has not been assessed in the paediatric population, and should not be used in children.

Table 172. TEAEs with an Incidence of ≥ 3% by Gender (APTS) – Alcohol Dependence Pool

	Women PBO		Women NMF		Men PBO		Men NMF	
Preferred Term	n	(%)	n	(%)	n	(%)	n	(%)
Number of Patients	227		301		570		843	
Patients with TEAEs	156	(68.7)	234	(77.7)	344	(60.4)	621	(73.7)
Nausea	23	(10.1)	93	(30.9)	24	(4.2)	160	(19.0)
Dizziness	19	(8.4)	78	(25.9)	25	(4.4)	130	(15.4)
Insomnia	11	(4.8)	39	(13.0)	32	(5.6)	114	(13.5)
Headache	21	(9.3)	54	(17.9)	45	(7.9)	87	(10.3)
Nasopharyngitis	24	(10.6)	29	(9.6)	49	(8.6)	78	(9.3)
Fatigue	14	(6.2)	26	(8.6)	23	(4.0)	69	(8.2)
Vomiting	7	(3.1)	41	(13.6)	11	(1.9)	59	(7.0)
Somnolence	7	(3.1)	13	(4.3)	16	(2.8)	46	(5.5)
Decreased appetite	4	(1.8)	16	(5.3)	5	(0.9)	40	(4.7)
Diarrhoea	12	(5.3)	10	(3.3)	23	(4.0)	33	(3.9)
Hyperhidrosis	3	(1.3)	16	(5.3)	5	(0.9)	33	(3.9)
Tremor	1	(0.4)	10	(3.3)	10	(1.8)	30	(3.6)
Anxiety	13	(5.7)	13	(4.3)	14	(2.5)	29	(3.4)
Hypertension	8	(3.5)	6	(2.0)	17	(3.0)	29	(3.4)
Back pain	5	(2.2)	10	(3.3)	18	(3.2)	27	(3.2)
Dry mouth	3	(1.3)	8	(2.7)	9	(1.6)	26	(3.1)
Sleep disorder	2	(0.9)	13	(4.3)	3	(0.5)	25	(3.0)
Asthenia	1	(0.4)	10	(3.3)	4	(0.7)	23	(2.7)
Accidental overdose	8	(3.5)	3	(1.0)	26	(4.6)	22	(2.6)
Dysgeusia	2	(0.9)	10	(3.3)	7	(1.2)	16	(1.9)
Fall	4	(1.8)	2	(0.7)	19	(3.3)	15	(1.8)
Abdominal pain	3	(1.3)	10	(3.3)	8	(1.4)	14	(1.7)
Influenza	8	(3.5)	13	(4.3)	11	(1.9)	12	(1.4)
Malaise	2	(0.9)	11	(3.7)	0	(0.0)	12	(1.4)
Bronchitis	7	(3.1)	4	(1.3)	5	(0.9)	11	(1.3)
Irritability	3	(1.3)	10	(3.3)	4	(0.7)	7	(0.8)

{SS} Sex Specific
Dictionary: MedDRA 13.0

Table 173. TEAEs with an Incidence of ≥ 5% by Age (APTS) – Alcohol Dependence Pool

	age<65 PBO		age<65 NMF		age>=65 PBO		age>=65 NMF	
Preferred Term	n	(%)	n	(%)	n	(%)	n	(%)
Number of Patients	756		1074		41		70	
Patients with TEAEs	474	(62.7)	803	(74.8)	26	(63.4)	52	(74.3)
Nausea	45	(6.0)	235	(21.9)	2	(4.9)	18	(25.7)
Dizziness	42	(5.6)	187	(17.4)	2	(4.9)	21	(30.0)
Insomnia	40	(5.3)	141	(13.1)	3	(7.3)	12	(17.1)
Headache	59	(7.8)	138	(12.8)	7	(17.1)	3	(4.3)
Nasopharyngitis	68	(9.0)	103	(9.6)	5	(12.2)	4	(5.7)
Vomiting	18	(2.4)	94	(8.8)	0	(0.0)	6	(8.6)
Fatigue	34	(4.5)	89	(8.3)	3	(7.3)	6	(8.6)
Somnolence	21	(2.8)	57	(5.3)	2	(4.9)	2	(2.9)
Decreased appetite	9	(1.2)	51	(4.7)	0	(0.0)	5	(7.1)
Hyperhidrosis	7	(0.9)	44	(4.1)	1	(2.4)	5	(7.1)
Diarrhoea	33	(4.4)	39	(3.6)	2	(4.9)	4	(5.7)
Anxiety	26	(3.4)	38	(3.5)	1	(2.4)	4	(5.7)
Tremor	10	(1.3)	36	(3.4)	1	(2.4)	4	(5.7)
Back pain	22	(2.9)	32	(3.0)	1	(2.4)	5	(7.1)
Tachycardia	7	(0.9)	29	(2.7)	1	(2.4)	4	(5.7)
Paraesthesia	2	(0.3)	15	(1.4)	0	(0.0)	4	(5.7)

{SS} Sex Specific
Dictionary: MedDRA 13.0

8.9.2. Safety in overdose

The safety of nalmefene when taken at very high doses is unknown, but three studies have explored the use of nalmefene at doses substantially higher than that proposed:

- In Study CPH-101-0600 in pathological gambling, 52 patients were randomised to 100mg/day for 16 weeks.
- In Study IX-318-003-EXT in interstitial cystitis, 20 patients received 120mg/day for more than 2 years.
- In Study SP-N0408 in nicotine dependence, 12 patients received 80mg/day for 5 weeks following a 2-week titration period.

No major safety concerns arose at these high doses.

The highest single dose administered was 500mg, given as an oral dose in a study of opioid addiction. This dose produced no major sequelae or significant changes in vital signs.

8.9.3. Safety related to drug-drug interactions and other interactions

Based on the fact that nalmefene is an opioid antagonist, it would be expected to interact with recreational and therapeutic opioids. This raises a number of manageable safety concerns. For subjects who are dependent on opioids, introduction of nalmefene could induce a withdrawal syndrome, so it should be avoided in this context. For subjects requiring acute use of opioids for pain management, nalmefene could antagonise the analgesic effects of opioids leading to poor pain control, and it should be discontinued when opioids are prescribed.

Apart from this expected interaction, there was no evidence of any significant safety issues arising from the concurrent use of nalmefene with other medications.

8.10. Evaluator's conclusions based on the CER round 1

Overall, the safety profile of nalmefene is acceptable. Its use is associated with an increased incidence of a number of symptoms that reflect tolerability rather than safety issues. These include nausea, dizziness, headache, vomiting, and malaise. The therapeutic index appears to be broad, and higher doses have been used in previous studies for other indications, without major

problems. The side effects associated with nalmefene usage are likely to have caused some unblinding.

9. First round benefit-risk assessment

9.1. First round assessment of benefits

The benefits of nalmefene in the proposed usage are uncertain but, based on the more reliable of the two pivotal studies, may consist of the following:

- about 1-2 heavy-drinking days per month might be converted to moderate-drinking days
- about 3.5-5g of alcohol might be avoided per day, but the alcohol that is consumed could have an increased AUC, negating much of this small benefit

9.2. First round assessment of risks

Nalmefene does not appear to pose major safety concerns, and analysis of the safety data primarily points to tolerability issues.

The risks of nalmefene in the proposed usage are:

- busy clinicians could trust the nalmefene to reduce alcohol intake, and cut back on effective psychosocial treatments, leading to increased alcohol intake compared to standard care.
- subjects taking nalmefene are at increased risk of dizziness, nausea, fatigue and insomnia

9.3. First round assessment of benefit-risk balance

The benefit-risk balance of nalmefene, given the proposed usage, is unclear and possibly unfavourable. The efficacy of the drug was marginal in the original target population, was only slightly better in a post hoc subgroup of high-risk drinkers, and could have been over-estimated because of methodological flaws in the submitted studies. Efficacy in a trial setting might also translate poorly into clinical practice. There are potential PK interactions between nalmefene and ethanol that appear to increase exposure to ethanol, compromising the therapeutic intent of the drug. Finally, the proposed PI is misleading.

10. First round recommendation regarding authorisation

The application to register nalmefene for the treatment of alcohol dependence should be rejected.

Should the Sponsor convince the TGA to approve the application despite the recommendations of this report, the PI should be modified extensively to emphasise the prospective primary endpoints of each Lundbeck study, with post hoc subgroup given a secondary emphasis. Post hoc endpoints should be presented descriptively rather than with p-values.

Other changes to the PI should be made as specified below.

11. Clinical questions to the CER round 1

11.1. Pharmacokinetics

How much does concurrent administration of nalmefene and alcohol increase exposure to alcohol?

The potential PK interaction between nalmefene and ethanol should be clarified with a new, adequately powered study.

What is the absolute bioavailability of nalmefene?

Ideally, a direct bioavailability study should be performed comparing the proposed oral formulation with intravenous administration.

How does the PK profile of nalmefene vary in the target population of subjects with excessive alcohol consumption?

The PK of nalmefene in the target population – subjects with excessive alcohol consumption – should be evaluated in a new PK study.

11.2. Pharmacodynamics

No questions.

11.3. Efficacy

What was the magnitude of the effective reduction in alcohol intake in the pivotal studies after allowing for the potential drug interaction between nalmefene and alcohol?

Combining nalmefene with ethanol appears to increase the AUC for ethanol by ~9%, with a 90%CI consistent with an increase of up to 21% (ratio 1.086, 90%CI 0.977 to 1.208), as shown in the table below. The existing efficacy results should be adjusted for the potential PK interaction between nalmefene and ethanol, and resubmitted. Two mean values for TAC-reduction should be estimated, one based on the mean increase in exposure (8.6%) and one based on the upper limit of the 90% CI for the increase in exposure (20.8%). Alternatively, a new study of this interaction should be performed to provide more accurate estimates of the strength of the interaction.

Table 174. PK Parameters of Ethanol following a Single Oral Dose of 0.6g/kg Ethanol Administered to Subjects Exposed to Nalmefene 20mg or Placebo – Study 13513A

Parameter	Ethanol plus Nalmefene (20mg) or Placebo Tablet	Ethanol		Nalmefene plus Ethanol *versus* Placebo Tablet plus Ethanol (all subjects) Ratio (90% CI)
		Men	Women	
AUC_{0-t} (mmol·h/L)	NE	42.0 (25.6)[a]	52.1 (30.9)[a]	1.086 (0.977, 1.208)
	PE	38.4 (35.4)[b]	46.8 (37.2)[c]	
C_{max} (mmol/L)	NE	14.7 (19.5)[a]	18.1 (31.5)[a]	1.003 (0.913, 1.101)
	PE	15.2 (<2.2, 19.1)[b]	18.5 (24.8)[c]	
t_{max} (h)	NE	2.25 (1.50, 3.00)[a]	2.50 (1.50, 4.00)[a]	
	PE	2.00 (1.50, 3.00)[d]	2.00 (1.50, 3.00)[c]	

Arithmetic mean (CV%) data are presented for AUC_{0-t} and C_{max} (NE for men and women and PE for women).
Median (min, max) data are presented for t_{max} and C_{max} (PE for men).
NE = nalmefene plus ethanol; PE = placebo tablet plus ethanol; N = number of subjects; CI = confidence interval

a N=16 c N=21
b N=18 d N=17

How many subjects in each of the three Lundbeck studies had a major protocol deviation?

This data was not clearly summarised in the study reports. For each study, the Sponsor should provide a single-page table summarising the incidence of major and minor protocol deviations, by category.

What was the drinking behaviour of subjects who withdrew from the pivotal studies?

The withdrawal rates in the pivotal studies were high, particularly in Study 12014A, raising substantial concerns about withdrawal bias. The Sponsor should clarify what is known about the drinking behaviour of subjects who withdrew from Studies 12014A and 12023A after the point of withdrawal. If no such data is available, this should be stated clearly.

How many subjects in the pivotal studies guessed they were receiving active treatment?

In the pivotal Lundbeck studies, was a bittering agent used in the placebo tablets?

Nalmefene is reported to have a bitter taste, and also produces some side effects. Unblinding in the pivotal studies was not assessed or reported, and the Lundbeck study reports do not mention use of a bittering agent. (The major Biotie studies did employ a bittering agent in the placebo tablets.) The capacity for unblinding in the pivotal studies should be tested in a new, placebo-controlled study of drinkers, who should be asked to guess their assigned treatment during and after a period of using nalmefene or placebo in the same manner as in the pivotal studies. The period of blinded treatment should be long enough that subjects have a chance to encounter the typical spectrum of nalmefene side effects (at least 2-4 weeks). These results would then allow interpretation of the efficacy results in the pivotal studies.

Do the results obtained in post hoc analyses of the pivotal studies fairly reflect the likely efficacy of nalmefene when used prospectively in high-risk drinkers?

A new, prospective placebo-controlled study of subjects with high or very high DRL at randomisation should be performed, allowing prospective confirmation of adequate efficacy in this subgroup, which has so far only been identified post hoc.

In real clinical practice, outside the artificial context of a clinical trial, how does the availability of a pharmacological treatment for alcohol dependence affect the thoroughness with which non-pharmacological measures are provided by busy clinicians?

Most subjects in the pivotal studies showed a substantial response to non-pharmacological measures. The additional clinical benefit of nalmefene demonstrated in the pivotal studies was marginal, even in an artificial setting where all subjects received the same non-pharmacological measures. If, in real clinical practice, nalmefene partially displaced non-pharmacological approaches rather than being provided in addition to non-pharmacological approaches, that marginal benefit could be negated, and the availability of nalmefene could even lead to worse outcomes by appearing to give clinicians a treatment option that is quicker and easier than time-consuming counselling. What evidence does the Sponsor have that this will not occur? What post-marketing monitoring does the Sponsor propose to assess for this effect?

11.4. Safety

The Sponsor should explain the discrepancy between Table 105 of the Summary of Clinical Safety (excerpt below), and Table 106 from the same report (reproduced below). The number of placebo recipients with a severe related TEAE differs in the two tables.

Table 105 Related TEAEs by SOC, Preferred Term, and Intensity (APTS) - Alcohol Dependence Pool

	Mild		Moderate		Severe	
System Organ Class and Preferred Term	n	(%)	n	(%)	n	(%)
PBO, Number of Patients=797, Men=570, Women=227	137	(17.2)	118	(14.8)	34	(4.3)

Table 106 Related, Severe TEAEs with an Incidence of 1% or More in Either Treatment Group (APTS) - Alcohol Dependence Pool

	PBO		NMF	
Preferred Term	n	(%)	n	(%)
Number of Patients	797		1144	
Patients with Related, severe TEAEs	33	(4.1)	120	(10.5)
Dizziness	3	(0.4)	27	(2.4)
Nausea	4	(0.5)	26	(2.3)
Insomnia	0	(0.0)	24	(2.1)
Vomiting	0	(0.0)	15	(1.3)

The Sponsor should also provide a summary table for haematological indices of potential clinical concern.

12. Second round evaluation of clinical data

12.1. List of issues and questions

The list of issues outlined below follows the sequence in the Sponsor's two Section 31 Responses (S31Rs) - in particular, the first S31R, dated 4th June, 2014, and the second S31R, dated 30th June. The first S31R contained questions numbered 1 to 12 (AU Response to TGA Section 31 Request_clinical_Apr 2014.pdf). The second S31R contained new questions and responses, originally numbered 1 and 2 (AU Response to TGA Section 31 Request_clinical_additional questions_June 2014.pdf). To avoid duplication of numbers, this document (Second Round Clinical Evaluation Report, CER2) has renumbered the later questions. Questions 1 to 12 were already numbered as such in the Sponsor's first S31R and keep their numbers here; the subsequent items discussed in the second S31R (Questions 1 and 2) have been reassigned numbers (Questions 13 and 14) in this document. Suggested edits to the PI are discussed.

Some areas of disagreement were not directly addressed in either S31R, but were instead submitted as annotations to a copy of the original First Round Clinical Evaluation Report

(CER1). Where the comments were substantive, they have been discussed at the end of the list of questions; where they were minor, they have simply been incorporated into the second-round report.

A review of the Sponsor's annotations revealed a couple of errata in the first-round report, discussed at the end of this section. Any known errors have been corrected in a revised version of CER1, along with some minor corrections of typographical errors. Any changes of significance have been marked with footnotes in the main body of this report; the changes do not modify the overall conclusions reached.

12.1.1. Question 1

12.1.1.1. Question

How much does concurrent administration of nalmefene and alcohol increase exposure to alcohol?

The potential PK interaction between nalmefene and ethanol should be clarified with a new, adequately powered study.

12.1.1.2. Background

This issue is discussed extensively elsewhere in this report. Basically, the only PK interaction study performed with alcohol and nalmefene demonstrated a small, statistically insignificant effect whereby alcohol exposure appeared to be increased by 9% when consumed with concurrent nalmefene. Although the effect was uncertain, and could be overturned with further study, it is noteworthy that, if the observed results were reproduced in clinical practice, this could be enough to negate the observed treatment effect in the pivotal studies. The uncertainty bounds surrounding the estimate of the PK effect are sufficient that it remains possible that there is no PK effect, but the estimated PK uncertainty adds to the uncertainty of the efficacy results, so that a proper accounting of all the residual uncertainty in the Sponsor's submission would make the efficacy of nalmefene even less statistically robust than it already is. That is, conventional considerations of statistical significance and the traditional p-value threshold of $p < 0.05$ require at least 95% certainty that an efficacy result has not arisen by chance, but there is already at least a 50% chance that the PK effect of nalmefene increases alcohol exposure by an amount similar to the claimed reduction in exposure. Worse, the 90%CI for the interaction includes the possibility that alcohol exposure could even be increased by 21% when alcohol is taken with nalmefene.

12.1.1.3. Sponsor's response

No increase in alcohol exposure is seen when nalmefene is given concurrently with alcohol. This was demonstrated in the pharmacokinetic/pharmacodynamics interaction Study 13513A which was an adequately powered study. As acknowledged in the clinical evaluation report, the estimate of alcohol exposure is completely within the defined range for declaring bioequivalence. By establishing bioequivalence, there is *no clinically relevant difference* in the exposure to alcohol when nalmefene or placebo is given concurrently with alcohol. [Emphasis added.]

12.1.1.4. Discussion

In the context of the current submission, which seeks to register a drug that, at best, minimally reduces alcohol intake and thereby reduces alcohol exposure, the italicised sections in the Sponsor's response cited above expose a double standard. The sponsor is quite prepared to equate 9% with "no increase" or "no clinically relevant difference" even though the mean PK change is similar in magnitude (but opposite in sign) to the observed treatment effect in the primary analysis of the pivotal studies. One could as easily dismiss the treatment effect itself as constituting "no clinically relevant difference" because it, too, is within the conventional bioequivalence range. That is, the Sponsor is quite prepared to dismiss a 9% increase in

exposure mediated through PK means as insignificant while claiming a similar decrease mediated by PD means as clinically worthwhile. The same magnitude of change in exposure is considered important when it suits their argument but negligible when it does not.

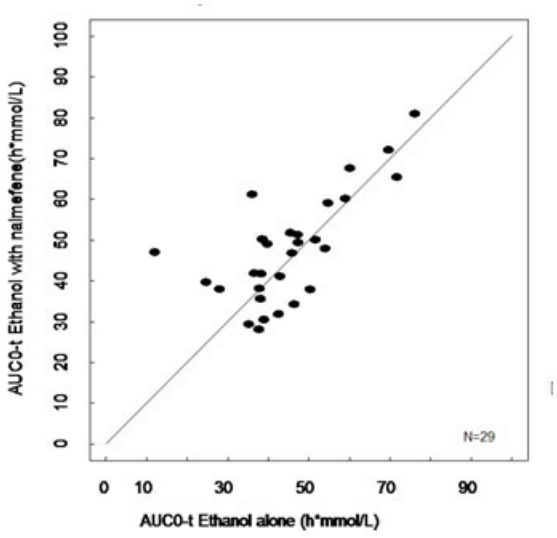
For many drugs, the therapeutic index is so broad that doses in modest excess of those intended are completely harmless. For many other drugs, the dose is titrated against the desired effect so that a consistent change in exposure can be compensated by changing the dose. The rationale behind the conventional bioequivalence range of 0.8 to 1.25 rests on the assumption that, for most drugs, a modest change in exposure is of little clinical importance. This not true of alcohol, however, where even small changes in exposure could, theoretically, increase the risk of alcohol-related complications. Indeed, the Sponsor's submission rests entirely on the assumption that reducing exposure by a few percent over and above the placebo effect is a worthwhile goal, and the Sponsor devotes a whole section of their submission to a model in which minor changes in alcohol exposure are translated through a hundred thousand patient-years of nalmefene use into estimated mortality benefits. One could use the same modelling process to derive a mortality cost of the PK interaction, complete with a 95%CI based on a possible 21% increase in exposure.

Having dismissed a 9% increase in exposure as insignificant, the Sponsor then proceeds to argue: "There is no scientific basis for a metabolic interference between alcohol and nalmefene." This is a reasonable inference, based on what is known of the two drugs, but clinical submissions for new therapeutic agents generally require that claims be supported by direct evidence, not from indirect inferences. Not all drug interactions can be reliably estimated from first principles, which is why PK studies are necessary in the first place.

The Sponsor also points out that the pharmacokinetics of alcohol are variable, and that the individual PK data in the interaction study were randomly distributed with no overall convincing trend suggesting a true PK effect.

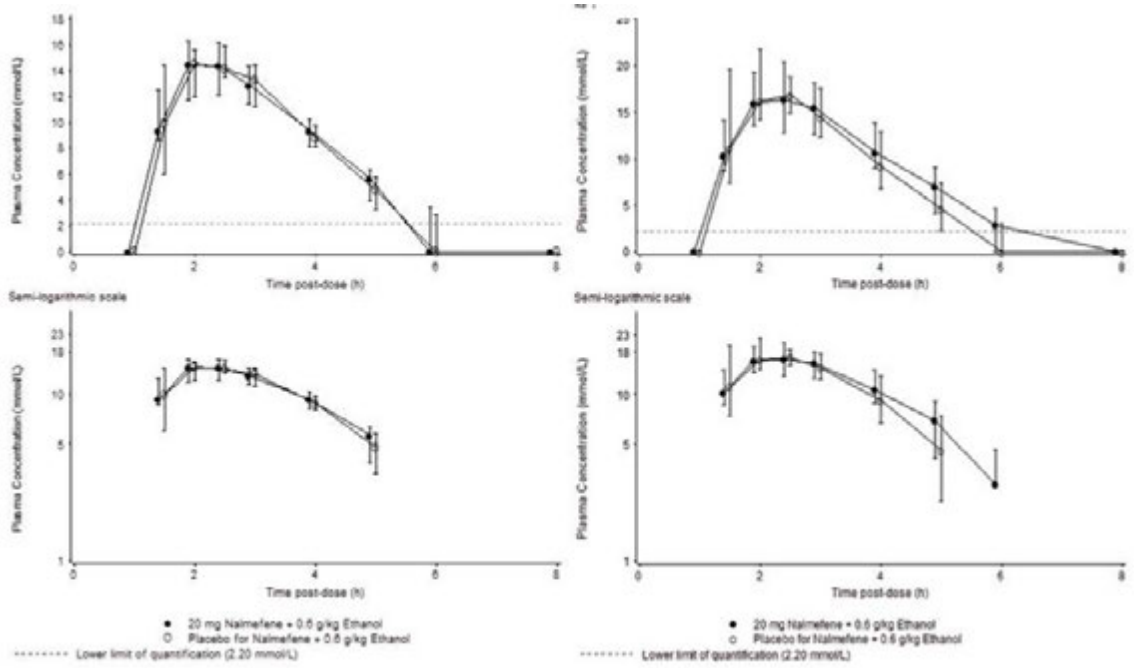
For instance, in a plot of alcohol AUC0-t with and without nalmefene, the results appeared to be randomly distributed about unity.

Figure 44. Alcohol AUC0-t with Nalmefene versus Alcohol AUC0-t without Nalmefene – Study 13513A



Also, in plots of alcohol concentration versus time, with and without nalmefene, it is clear that the variability and uncertainty around the mean estimate is greater than the difference between the two conditions.

Figure 45. Concentration-time Profiles for Alcohol with and without Nalmefene



The evaluator accepts that the PK result is likely to represent random variation that may not be reproducible in subsequent PK studies. The evaluator also accepts that the observed mean ratio for AUC is a point estimate within a broader confidence interval that includes unity, as well as the possibility of a favourable PK effect. Nonetheless, there is currently no strong, reliable, empirical evidence that the mean PK interaction is small in comparison to the therapeutic effect.

If the observed therapeutic effect of nalmefene were strong enough, then this would not be an issue, and some uncertainty about the PK interaction would be acceptable, but as it stands the mean PK effect observed in Study 13513A is similar in magnitude but opposite in sign to the apparent therapeutic effect in the pivotal studies. It would be therefore be of interest to reassess the efficacy data with this in mind (see Question 4, below), if only to clarify the residual uncertainty surrounding the efficacy estimate, which is why the analysis was requested.

12.1.1.5. Conclusion

The Sponsor has not resolved this issue satisfactorily. The PK interaction remains a potentially serious issue because the observed interaction is similar in magnitude to the therapeutic effect and, if it were maintained in general use, could counteract the claimed benefits of nalmefene. The Sponsor's attempt to dismiss this issue by reference to the traditional bioequivalence range ignores the fact that their whole submission rests on the idea that small changes in alcohol exposure are important.

12.1.2. Question 2

12.1.2.1. Question

What is the absolute bioavailability of nalmefene?

Ideally, a direct bioavailability study should be performed comparing the proposed oral formulation with intravenous administration.

12.1.2.2. Sponsor's response

"It is acknowledged that the individual values of absolute bioavailability in Study R7 (63, 120, 60, and 72%) are different from the mean estimate of the absolute bioavailability (41%) given in the Package Insert (PI)."

The Sponsor presents a theoretical argument to the effect that their PK modelling of the population-PK results should produce a more reliable estimate of bioavailability than the individual estimates obtained from Study R7. Given that Study R7 only involved 4 patients, this seems plausible.

Overall, this issue is relatively unimportant compared to the many other issues discussed in this document. It remains the case, though, that no formal bioavailability study has been performed that confirms the Sponsor's indirect estimate of 41%. It would be appropriate for the PI to acknowledge this.

12.1.3. Question 3

12.1.3.1. Question

How does the PK profile of nalmefene vary in the target population of subjects with excessive alcohol consumption?

The PK of nalmefene in the target population – subjects with excessive alcohol consumption – should be evaluated in a new PK study.

12.1.3.2. Sponsor's Response

"The pharmacokinetic profile of nalmefene in the target population, that is, in patients with high or very high DRL, was not obtained in the Lundbeck clinical studies. The Applicant acknowledges that it would have been standard procedure for most drug development programmes. However, it has to be taken into account that the IMP was to be taken 'as needed', as opposed to a fixed-dose design with daily regular dosing. In addition, the pharmacokinetics of nalmefene in the target population is not expected to be different to a clinically relevant degree from the pharmacokinetics in healthy subjects. As nalmefene is extensively metabolised in the liver, the only comorbidities in the target population that would be expected to have potential impact on the pharmacokinetics of nalmefene at baseline and at randomisation would be related to those comorbidities involving hepatic impairment."

The Sponsor points out that most subjects in the target population would be expected to have normal liver function. For those with mild-to-moderate hepatic impairment, the PK results in Study 12417A are relevant:

"In Study 12417A, the systemic exposure following a single oral nalmefene dose, based on AUC0-inf, was statistically significantly larger for subjects with mild or moderate hepatic impairment than in healthy subjects with normal hepatic function (1.5 and 2.9 times, respectively). Given the wide safety margin of nalmefene, the difference in exposure between the groups did not raise any safety concerns."

The Sponsor also notes that nalmefene is contraindicated in the presence of severe hepatic impairment.

12.1.3.3. Conclusion

The lack of PK studies directly assessing the target population represents a relatively minor deficiency in the Sponsor's submission, and the evaluator agrees it is unlikely that major safety issues would arise from unexpected PK changes in this population. On the other hand, it would have been more appropriate to perform such studies, and accepting the omission sets an undesirable precedent.

12.1.4. Question 4

12.1.4.1. Question

What was the magnitude of the effective reduction in alcohol intake in the pivotal studies after allowing for the potential drug interaction between nalmefene and alcohol?

12.1.4.2. Sponsor's Response

"It has been established that there is no increase in alcohol exposure when nalmefene is given concurrently with alcohol."

The Sponsor's claim is false. The estimate for the increase in exposure was 9% (with a 95%CI reaching as high as 21%). While it is traditional to ignore PK interactions of this magnitude, it is not the case that a 9% increase is "no increase", and the traditional reasons for ignoring small PK interactions do not apply in a situation where the sole purpose of administering nalmefene is to produce a small decrease in alcohol exposure.

The Sponsor also mounts an argument that nalmefene could not increase exposure to alcohol because the efficacy results showed reduced alcohol consumption. For instance, the Sponsor writes, as evidence of a lack of a PK effect:

"Efficacy Variables Derived from TLFB Data There was a greater proportion of responders in the nalmefene group than in the placebo group where response was defined as a 2-category downward shift from baseline in DRL, as a shift to low DRL or below, as a ≥70% reduction in TAC compared with baseline, as 0 to 4 HDDs/month or as a ≥70% reduction from baseline in the number of HDDs."

This logic is questionable. The Sponsor is explicitly citing TLFB data as evidence that no PK interaction has occurred, even though there is no conceivable way that TLFB data could detect a PK interaction (except indirectly, if subjects reduced their number of drinks because each drink provided more alcohol exposure – but if this happened to any appreciable extent then favourable efficacy results could be counted as evidence for a PK interaction, not as evidence against an interaction). The original concern was that, due to a PK interaction, each drink of alcohol taken with nalmefene could have produced 109% of the exposure that would have been produced by the same drink taken without nalmefene. Even if this concern were fully justified, the diary data would still record that a single drink had been consumed regardless of the subsequent exposure, so referring to TLFB in this context is pointless.

For other variables, such as Clinician's Global Impression, it is also doubtful that a previously unsuspected PK interaction would be detected by examining the efficacy results. If a patient reduced intake by ~9% and also suffered a PK-mediated increase in exposure of ~9%, such that their overall alcohol exposure was unchanged, this might be interpreted by the clinician as an improvement because the stated number of drinks consumed would be less.

For one set of efficacy variables, liver function tests, there is some potential for detecting the overall balance between PD and PK effects. The Sponsor states:

"When considering efficacy variables independent of the timeline followback (TLFB) data, such as ALT and GGT, that is, indicators of potential liver damage, there were greater improvements in ALT and GGT at Week 24 in the nalmefene group than in the placebo group. The difference was in favour ($p<0.05$) of nalmefene for both ALT and GGT in Study 12014A, and for ALT in Study 12023A (Panel 7 and Panel 8). That means that there is no evidence to support the Evaluator's assumption that nalmefene increased alcohol exposure as greater exposure to alcohol due to nalmefene would not be expected to lead to greater reductions in ALT and GGT relative to placebo."

This logic is also questionable, because the net effect on ALT and GGT could represent the balance between a favourable PD effect and an adverse PK effect; an improvement in LFTs does not necessarily mean that there was no adverse PK effect. Even if it is accepted that the overall balance between PD and PK effects produced by nalmefene was favourable, as revealed in improved LFTs, it is very difficult to translate from mean changes in liver enzymes back into standard, clinically meaningful efficacy variables such as TAC and HDDs. For instance, if a subject reduced their consumption by 40 standard drinks in a month, and a PK effect effectively increased their exposure by the equivalent of 20 standard drinks, then the net effect on LFTs

could be favourable, but it would still be the case that the recorded change in intake overestimated the net benefit produced by nalmefene. Also, it should be noted that the LFT results cited come from a post hoc subgroup analysis, so their reproducibility in a prospective study remains unknown.

Table 175. ALT results in the Post Hoc High-risk Subgroup

Panel 7 ALT at Week 24 (FAS, MMRM) - Patients with a *High* or *Very High* DRL at Baseline and Randomisation – Studies 12014A and 12023A

Study Treatment Group	Geometric Mean at Baseline		Geometric Mean at Week 24		Ratio to PBO		
	N	Mean	N	Mean	Ratio	95% CI	p-value
12014A							
PBO	166	29.3	110	29.6			
NMF	171	29.4	87	24.7	0.83	[0.75; 0.93]	0.001
12023A							
PBO	153	29.0	108	31.5			
NMF	148	29.3	100	26.8	0.85	[0.75; 0.96]	0.010

Table 176. GGT results in the Post Hoc High-risk Subgroup

Panel 8 GGT at Week 24 (FAS, MMRM) – Patients with a *High* or *Very High* DRL at Baseline and Randomisation – Studies 12014A and 12023A

Study Treatment Group	Geometric Mean at Baseline		Geometric Mean at Week 24		Ratio to PBO		
	N	Mean	N	Mean	Ratio	95% CI	p-value
12014A							
PBO	167	60.1	112	53.9			
NMF	171	55.7	87	39.5	0.73	[0.64; 0.84]	<0.001
12023A							
PBO	153	54.9	108	52.4			
NMF	148	55.9	100	47.3	0.90	[0.80; 1.07]	0.244

12.1.4.3. Conclusion

This issue remains unresolved, and the Sponsor has refused to perform the requested analysis. It should be noted that, if a 9% increase in exposure really were equivalent to "no increase", then the Sponsor would have nothing to lose by performing the requested analysis. It is difficult to escape the conclusion that the Sponsor has avoided answering this question because a repeat analysis of the efficacy results that incorporated the PK results would reveal the efficacy results to be marginal and uncertain.

12.1.5. Question 5

12.1.5.1. Question

How many subjects in each of the three Lundbeck studies had a major protocol deviation?

This data was not clearly summarised in the study reports. For each study, the Sponsor should provide a single-page table summarising the incidence of major and minor protocol deviations, by category.

12.1.5.2. Sponsor's response

The Sponsor has provided summary tables for protocol deviations and these have been incorporated into this report. Unfortunately, the tables did not differentiate between major and minor protocol deviations.

12.1.5.3. Conclusion

The number of protocol deviations was acceptable.

12.1.6. Question 6

12.1.6.1. Question

What was the drinking behaviour of subjects who withdrew from the pivotal studies? The withdrawal rates in the pivotal studies were high, particularly in Study 12014A, raising substantial concerns about withdrawal bias. The Sponsor should clarify what is known about the drinking behaviour of subjects who withdrew from Studies 12014A and 12023A after the point of withdrawal. If no such data is available, this should be stated clearly.

12.1.6.2. Background

This question was asked because of concerns about withdrawal bias, as discussed in more detail elsewhere.

12.1.6.3. Sponsor's response

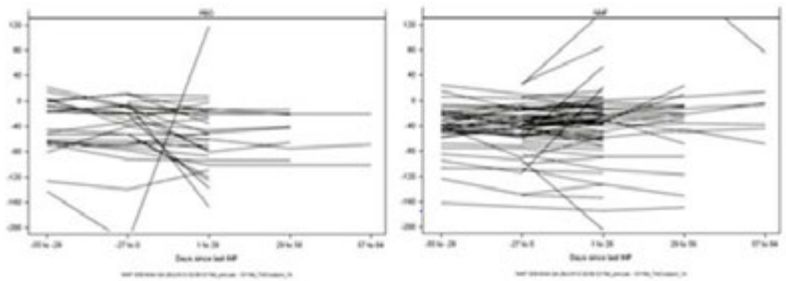
"Information on the drinking behaviour after the patients withdrew from the studies has not been collected."

In the S31R, the Sponsor conceded that the drinking behaviour of withdrawing subjects is unknown, but then discussed four factors that they felt limited the scope for withdrawal bias.

'The Applicant will address the following topics in the response: alcohol consumption during the study for patients who withdrew from the pivotal studies adverse event profile and lack of rebound effect when discontinuing nalmefene (Run-out Period) potential withdrawal bias assessed by sensitivity analyses the proportions of withdrawals in the Lundbeck-sponsored clinical studies are comparable with proportions of withdrawals in other published studies."

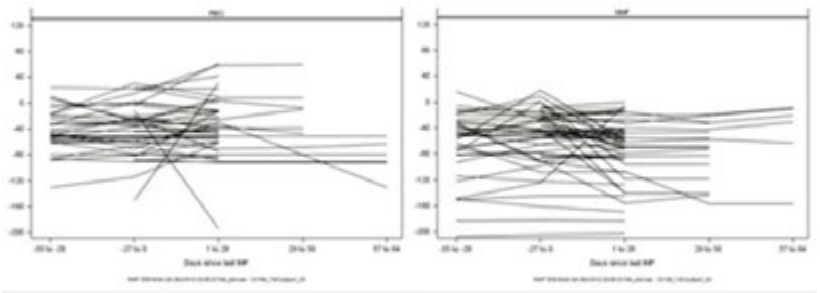
With regard to alcohol consumption of withdrawing subjects, the Sponsor provided graphs showing alcohol consumption in relation to the last dose of IMP, up to the point that subjects withdrew. The figures below shows the results for TAC, and similar results were observed with HDDs. The first figure relates to Lundbeck14, and the second to Lundbeck23. A couple of observations can be made about the Lundbeck14 results. In several placebo recipients, the data ends on a downward trend and in one placebo recipient there was a sharp upward trend in drinking prior to withdrawal. In nalmefene recipients, some subjects showed minor downward trends in drinking prior to withdrawal, but a few showed upward trends, including four subjects with marked upwards trends. In most subjects from both treatment groups, there was no clear change in drinking behaviour in the lead-up to withdrawal. The overall mean trend prior to withdrawal is unclear from these graphs, as is the difference, if any, between treatment groups, but visual inspection is not particularly reassuring with respect to the possibility of withdrawal bias: marked upward trends appear more common with nalmefene.

Figure 46. Changes from baseline in TAC (g/day) in patients who withdrew by time since last dose of IMP (PBO, n = 30, NMF, n = 63) – Study 12014A



For Lundbeck23, withdrawal rates were more similar in the two treatment groups. Visual inspection of the graphs suggests that several placebo recipients were increasing their intake prior to quitting, and several nalmefene recipients were decreasing their intake, which is somewhat reassuring with respect to the possibility of withdrawal bias.

Figure 47. Changes from baseline in TAC (g/day) in patients who withdrew by time since last dose of IMP (PBO, n = 43, NMF, n = 53) – Study 12023A



Importantly, though, such an analysis reveals nothing about the drinking behaviour of subjects after they withdrew. The concern remains that subjects who found it arduous to curtail their drinking were more likely to quit the study, particularly if they also experienced side effects, leading to progressive enrichment of the continuing cohort with more motivated subjects. This enrichment would not necessarily be revealed in data collected before subjects quit, while they were still motivated enough to continue.

The Sponsor also notes: "There were no differences in the baseline characteristics between the patients who withdrew and the patients who completed each study. This is consistent with the results of the analyses of time to withdrawal for any reason performed for subpopulations based on demographic and disease."

This implies that there is no simple, predictable relationship between baseline disease characteristics and subsequent withdrawal. It does not rule out the possibility that withdrawing subjects were different from continuing subjects.

With regard to the Sponsor's second discussion point, the lack of an apparent rebound effect on discontinuing nalmefene, the Sponsor writes: "Data from the 4-week Run-out Period in Study 12014A (Panel 18) and Study 12023A (Panel 19) do not suggest that the patients in the nalmefene-placebo group increase their alcohol consumption after they received placebo in the Run-out Period (ROP). The patients who had been treated with nalmefene in the Main Treatment Period continued to have reduced alcohol consumption in the ROP."

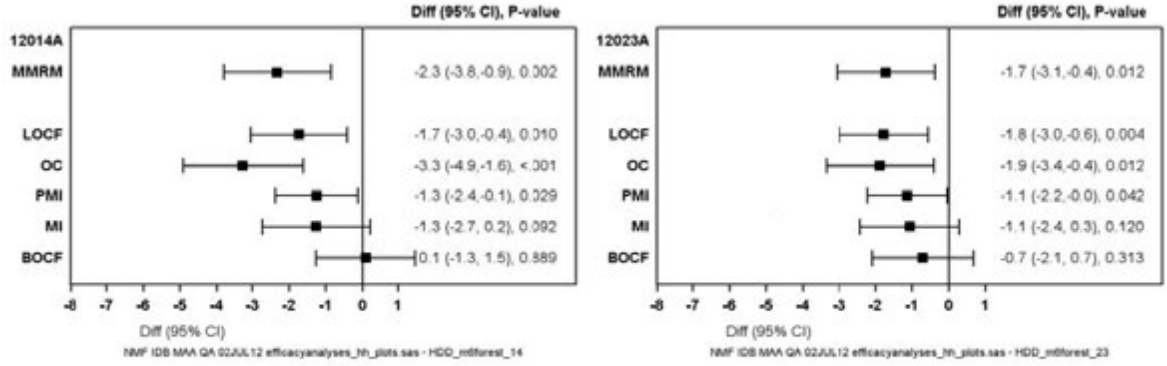
This observation rests on the assumption that the observed behaviour after ceasing nalmefene in the ROP can be used to infer the behaviour of discontinuing subjects. For such an inference to be relevant, it would need to be known that subjects randomised to placebo in the ROP were comparable to patients who quit the study. At the core of the concern about withdrawal bias is the suspicion that the groups are not comparable, so it is not helpful in this context to assume that they are. The Sponsor's point is irrelevant.

The Sponsor's third discussion point refers to sensitivity analyses purported to assess the potential withdrawal bias. These analyses have already been discussed in the main body of this evaluation report, but the key results cited in the S31R are shown in the figures below. It is notable that Lundbeck23, which was potentially less susceptible to withdrawal bias, produced less impressive results for HDDs (upper right figure) and failed to achieve significance for TAC (lower right figure) for most methods of imputation. Both studies failed to achieve a significant result for either co-primary endpoint using the most conservative, pessimistic imputation method (BOCF), and both studies produced better results using an observed cases (OC) result. Other imputation methods produced intermediate results. All of this is consistent with, but not proof of, the hypothesis that the treatment groups were progressively enriched with better-motivated subjects.

Although some of the imputation methods (such as MI) attempted to model drinking behaviour of withdrawing subjects using a variety of statistical approaches, it is completely unknown whether such modelling is successful, because, as the Sponsor concedes, there is no data

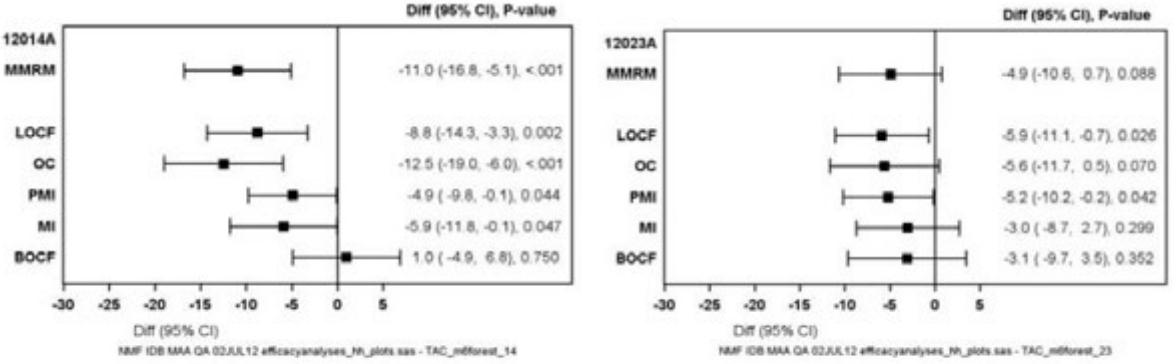
available to indicate how subjects actually behaved after they quit. Such modelling relies on the assumption that behaviour prior to quitting a study is a good predictor of behaviour after quitting, which remains an untested assumption.

Figure 48. Sensitivity Analyses – Changes from Baseline to Month 6 in HDDs (days/month) – Studies 12014A and 12023A



MMRM is the co-primary efficacy analysis and is included for comparison.
Negative values indicate a greater reduction in the number of HDDs and TAC in the nalmefene group than in the placebo group.

Figure 49. Sensitivity Analyses – Changes from Baseline to Month 6 in TAC (g/day) – Studies 12014A and 12023A



MMRM is the co-primary efficacy analysis and is included for comparison.
Negative values indicate a greater reduction in the number of HDDs and TAC in the nalmefene group than in the placebo group.

One imputation method that appears less reliant on assumptions about future drinking is the PMI method, described by the Sponsor as follows:

"The PMI approach imputed the reduction at Month 1 in the placebo group to all time points with missing data, thus this approach is considered conservative and it is unlikely to be biased in favour of nalmefene, as it does not take into account the continued reduction in alcohol consumption after Month 1 in the placebo group."

This method narrowly achieved significance for both co-primary endpoints in both pivotal studies, without relying on post hoc selection of a subgroup. Potentially, this amounts to the Sponsor's strongest prospective evidence that nalmefene has significant efficacy – except that the PMI method was not designated as the primary imputation method, and the results would not remain significant if appropriately corrected for the use of multiple statistical methods. Also, it is potentially misleading to emphasise the fact that the PMI method "does not take into account the continued reduction in alcohol consumption after Month 1 in the placebo group", as stated above. The placebo group showed a major reduction in drinking during the first month on treatment, and only minimal improvements thereafter. Signing up to a study of this nature

clearly provided an initial motivational boost for most subjects. The PMI method locked in the first month of improvement for subjects who quit, regardless of whether their quitting reflected declining motivation to continue to reduce their alcohol intake.

The fourth discussion point relates to how the withdrawal rate in the pivotal Lundbeck studies compares to other studies reported in the literature (see the table over the page). In this context, the Lundbeck studies appear broadly comparable to other studies of alcohol dependence, but it should be noted that the Sponsor's S31R table (over the page) does not reveal the withdrawal rate in each treatment group (see the table below, from the original submission). The problem with Lundbeck14 was not just that the withdrawal rate was high but also that it was unbalanced, being particularly high in the active group – almost twice that seen in the placebo group. Even if other studies were similarly affected by high or unbalanced withdrawal, the fact that other clinical studies have faced a similar problem does not alter the likelihood of withdrawal bias in the Sponsor's studies; it simply raises the possibility that withdrawal bias is a widespread problem in studies of this nature.

Table 177. Withdrawals in the Pivotal Lundbeck Studies

	12014A		12023A	
	PBO n (%)	NMF n (%)	PBO n (%)	NMF n (%)
FAS	289 (100)	290 (100)	326 (100)	329 (100)
Patients completed[a]	213 (74)	152 (52)	229 (70)	212 (64)
Patients withdrawn	76 (26)	138 (48)	97 (30)	117 (36)
Primary Reason				
Adverse events	17 (6)	57 (20)	4 (1)	12 (4)
Lack of efficacy	19 (7)	17 (6)	13 (4)	7 (2)
Non-compliance		7 (2)	3 (1)	7 (2)
Protocol violation	4 (1)	10 (3)	25 (8)	19 (6)
Withdrawal of consent	25 (9)	31 (11)	32 (10)	43 (13)
Lost to follow-up	7 (2)	12 (4)	11 (3)	12 (4)
Other	4 (1)	4 (1)	9 (3)	17 (5)

a Patients with TLFB data at Month 6

Table 178. Withdrawals in Studies of Patients with Alcohol Dependence

Study	Withdrawals (%)	Withdrawal of Consent (%)	Lost to Follow-up (%)
Published Studies of 12 to 16 weeks duration			
Guardia et al 2002[12]	27.2	9.9*	12.9***
Kranzler et al 2004[13]	22.2	5.7	9.2
Anton et al 2006[11]	34.9	26.0**	5.6
Anton et al 2008[17]	33.9	Not specified	8.5
Johnson et al 2003[15]	34.8	10.1	16.5
Johnson et al 2007[16]	31.0	10	6.2
Fertig et al 2012[18]	28.5	Not specified	Not specified
Published Studies of 6 months duration			
Garbutt et al 2005[14]	39.7	12.0	13.2
* referred to as "treatment refusal"			
** "None" was the main reason given for patients who withdrew			
*** referred to as patients who did not take medication or did not attend the physician's weekly assessments over 2 consecutive weeks			
Lundbeck-sponsored Studies			
At Month 4 (FAS)			
12014A	27.5	7.1	1.7
12023A	24.0	7.8	3.1
12013A	17.0	10.0	0.7
At Month 6 (FAS)			
12014A	37.0	9.7	3.3
12023A	32.7	11.5	3.5
12013A	22.1	12.7	1.3

12.1.6.4. Conclusion

It seems likely that withdrawal bias has inflated the apparent treatment effect of nalmefene in the Sponsor's pivotal studies, but it remains unclear whether this is a substantial problem or merely a theoretical concern. The potential for withdrawal bias was less pronounced in Lundbeck23, which produced weaker results than Lundbeck14; for this reason, the Lundbeck23 results may better reflect the true efficacy of nalmefene.

Sensitivity analyses using a variety of imputation methods reveal a range of different outcomes, ranging from poor results with the most pessimistic imputation methods to better results with the most optimistic methods. The PMI method, which replaces missing data with the 1-month placebo results, managed to produce positive outcomes for both co-primary endpoints in both pivotal studies. If there were no other concerns (such as PK effects and unblinding), this would come close to demonstrating that nalmefene had significant efficacy, but it is important to note that the PMI method locks in most of the placebo improvements, which have already occurred by Month 1, and would not remain significant if corrected for the use of multiple methods.

12.1.7. Question 7

12.1.7.1. Question

7a) How many subjects in the pivotal studies guessed they were receiving active treatment?

7b) In the pivotal Lundbeck studies, was a bittering agent used in the placebo tablets?

Nalmefene is reported to have a bitter taste, and also produces some side effects. Unblinding in the pivotal studies was not assessed or reported, and the Lundbeck study reports do not mention use of a bittering agent. (The major Biotie studies did employ a bittering agent in the placebo tablets.)

7c) The capacity for unblinding in the pivotal studies should be tested in a new, placebo-controlled study of drinkers, who should be asked to guess their assigned treatment during and after a period of using nalmefene or placebo in the same manner as in the pivotal studies. The period of blinded treatment should be long enough that subjects have a chance to encounter the typical spectrum of nalmefene side effects (at least 2-4 weeks). These results would then allow interpretation of the efficacy results in the pivotal studies.

12.1.7.2. Sponsor's response

The Sponsor's response to this important issue was very brief. It reproduced in its entirety below, broken into 3 sections.

7a) "All tablets (placebo as well as nalmefene tablets) used in the Lundbeck pivotal studies were Opadry coated to mask the bitter taste of nalmefene. In addition, the placebo tablets used in the pivotal studies contained the bittering agent, denatonium benzoate (see batch documentation in Appendix II), as an additional means to maintain the blinding due to the bitter taste of nalmefene. Thus, the bitter taste of nalmefene could not have led patients to guess that they were on active treatment. The Applicant acknowledges that this information was not included in the integrated study reports."

7b) "During the 4-week Run-out Period (ROP) in Studies 12014A and 12023A, there were few withdrawals among patients who were randomised to nalmefene in the Main Treatment Period (MTP) and then re-randomised to placebo in the ROP. This indicates that the patient were not able to guess that they were randomised to nalmefene in the MTP, and that they then received placebo in the ROP, since it would have been expected that more patients would withdraw during the ROP if they had guessed that they received placebo. Thus, there were no signs that patients in the nalmefene-placebo group guessed which treatment they were randomised to during the ROP."

7c) "In conclusion, according to the Applicant there is no need for testing the capacity for unblinding."

12.1.7.3. Discussion

It is somewhat reassuring to read that a bittering agent was used; without such an agent, all Lundbeck studies would need to be rejected outright. Unfortunately, the fact that the most obvious source of unblinding was ameliorated does not remove the possibility that unblinding occurred through other means.

The evidence against potential unblinding that the Sponsor cites in part b) of their response is extraordinarily weak. Basically, the Sponsor notes that patients who tolerated nalmefene and remained in the study for six months also generally remained in the study for one additional month when they were switched to placebo in the ROP. The Sponsor's assumption is that subjects would have quit if they had realised they were now receiving placebo, despite the following facts:

- these subjects had already consented to being involved in a study in which a randomised switch to placebo in the ROP was part of the protocol;
- these were the most compliant subjects from the whole study cohort, having remained in the study for the longest and having been enriched for compliance by withdrawal of up to half the active treatment group;
- they had already put up with the side effects of nalmefene for six months, and now only had to put up with the minor inconvenience of taking a placebo tablet on an as-needed basis.

The assumption that many of these subjects would have quit on realising they were receiving placebo is therefore simply not plausible. Also, note that the switch from active treatment to placebo, in subjects who have already developed tolerance to a drug, is generally less noticeable

than commencing active drug. The time period that the Sponsor focuses on is therefore the least likely period to furnish evidence of unblinding.

It seems particularly odd that the Sponsor is prepared to try to read an unblinding signal from the willingness of this super-compliant group to take innocuous placebo for one month at the tail end of the pivotal studies while simultaneously rejecting the possibility of reading an unblinding signal from any of the following:

- Nalmefene subjects withdrew from the pivotal studies at a substantially higher rate than placebo subjects (nearly twice as often in the FAS of Lundbeck14)
- Withdrawals from the pivotal studies were attributed to AEs in 6% of the placebo group and 20% of the nalmefene groups, an excess of 14%.
- On the first day of treatment in clinical studies, TEAES with placebo occurred in 17.7% of subjects, compared to 40.8% in the nalmefene group, an absolute excess of 23.1%.
- Overall, nausea, dizziness, insomnia, headache, vomiting, fatigue and somnolence all occurred with a clear excess in the active group (see table below).
- Phase 1 studies (such as the PK interaction study shown in the second table below) revealed that most subjects receiving nalmefene reported side effects after a single dose, compared to very few subjects after receiving placebo (81% vs 15%).

Table 179. Frequent Adverse Events by Preferred Term (APTS) – Alcohol Dependence Pool

	PBO		NMF	
Preferred Term	n	(%)	n	(%)
Number of Patients	797		1144	
Patients with TEAEs	500	(62.7)	855	(74.7)
Nausea	47	(5.9)	253	(22.1)
Dizziness	44	(5.5)	208	(18.2)
Insomnia	43	(5.4)	153	(13.4)
Headache	66	(8.3)	141	(12.3)
Nasopharyngitis	73	(9.2)	107	(9.4)
Vomiting	18	(2.3)	100	(8.7)
Fatigue	37	(4.6)	95	(8.3)
Somnolence	23	(2.9)	59	(5.2)

Table 180. Summary of Adverse Events (All Causes, Safety Set)

	Pre-treatment (N=46)	20 mg Nalmefene + Ethanol (N=43)	20 mg Nalmefene + Placebo (to Ethanol) (N=43)	Placebo (to Nalmefene) + Ethanol (N=42)	Placebo (to Nalmefene) + Placebo (to Ethanol) (N=41)	Overall (N=46)
Subjects with adverse events	1 (2%)	43 (100%)	35 (81%)	42 (100%)	6 (15%)	46 (100%)
Subjects with SAEs	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Subjects with adverse events leading to withdrawal	0 (0%)	2 (5%)	2 (5%)	0 (%)	0 (0%)	4 (9%)
Total number of adverse events	1	224	152	80	7	464

Source: Table 51
N = number of subjects

Interestingly, the Sponsor raised a different argument against the possibility of unblinding in one of their annotations to CER1. The Sponsor noted that many placebo recipients reduced their drinking while on placebo, and the Sponsor interpreted this as a lack of unblinding in the placebo group. There are many flaws with this argument, but the most important is that unblinding in the placebo group bears little relation to unblinding in the active group; this argument is discussed.

12.1.7.4. Conclusions

The Sponsor has confirmed that a bittering agent was used in the pivotal studies, which removes the most obvious potential source of unblinding. Description of the study protocols has been amended to include this new information.

The Sponsor appears unwilling to address the possibility of unblinding in the pivotal studies; instead, the Sponsor has attempted to draw inferences from the lack of a clear unblinding signal in the ROP of the pivotal studies whilst simultaneously ignoring several strong signals from the main study period and from Phase 1 studies.

The Sponsor made no attempt to assess unblinding during the pivotal studies and has declined the request to assess such unblinding now, in a follow-up study. In the pivotal studies, all the Sponsor had to do to allow an estimate of unblinding was to ask patients to guess their assigned treatment – this would have represented a trivial logistical challenge compared to all of the complex endpoints they did quantify. The technique of assessing unblinding by asking patients to guess their treatment assignment has been known to researchers for decades, and the Sponsor's lack of interest in this question could be interpreted as a sign that the studies were conceived with commercial rather than scientific goals in mind. Their failure to engage in a serious discussion of the issue in response to the first-round evaluation also potentially reveals something of their priorities.

12.1.8. Question 8

12.1.8.1. Question

Do the results obtained in post hoc analyses of the pivotal studies fairly reflect the likely efficacy of nalmefene when used prospectively in high-risk drinkers?

A new, prospective placebo-controlled study of subjects with high or very high DRL at randomisation should be performed, allowing prospective confirmation of adequate efficacy in this subgroup, which has so far only been identified post hoc.

12.1.8.2. Background

This issue has been discussed extensively throughout this evaluation. Essentially, when a subgroup is chosen post hoc, in response to favourable results, it is inappropriate to apply statistical tests that were designed for prospective hypothesis testing. This has nothing to do with the biological plausibility or clinical importance of the pharmacological effect in the subgroup in question; it is a basic statistical principle.

12.1.8.3. Sponsor's response

The Sponsor defends the focus on high-risk drinkers on biological and clinical grounds, proposing five main arguments in defence of the post hoc approach:

- 'The proposed target population (that is, patients with alcohol dependence with a high or very high drinking risk level [DRL] at baseline and at randomisation) is valid and in need of treatment. The proposed target population is a clinically relevant entity.
- The nalmefene treatment effect was more pronounced in the target population than in the total population of patients with alcohol dependence.
- There was no imbalance in baseline characteristics between the treatment groups in each of the studies, which indicates reliability of individual study results.
- The larger treatment effect of nalmefene versus placebo in the target population was replicated across 4 clinical studies.
- The Applicant also wishes to draw attention to the fact that an independent Scientific Advisory Group (SAG), mandated by the Committee for Medicinal Products for Human Use (CHMP) was asked to comment on the treatment goal for nalmefene during the evaluation of

the Marketing Authorisation Application by the EMA. The SAG recognised that the proposed target population was a clinically relevant entity, and endorsed the validity of the post hoc analysis.'

The first point is true but irrelevant. The post hoc target group is in need of treatment, but unfortunately the Sponsor did not identify them prospectively – nor has the Sponsor mounted any argument, even post hoc, in defence of the idea that medium-DRL subjects are not a "clinically relevant entity". The Sponsor's rejection of medium-DRL subjects cannot be defended on the basis that this group does not require treatment, because such subjects still face alcohol-related harm. The Sponsor initially also thought them worthy of targeting in the pivotal studies. The medium-DRL subjects appear to have been rejected solely in response to poor efficacy in this group.

The second point – the superiority of the results in the post hoc group compared to the original target group – clearly provides the Sponsor's motive for focussing on the post hoc analysis, but it merely underscores the fact that the results for this group were not representative of the broader results. Obviously, if results in this group had been less favourable than in the overall cohort, it is unlikely that the Sponsor would have chosen to emphasize them. The superiority of the results in this group provides a clear basis for studying them further in a prospective fashion, but does not legitimise the Sponsor's post hoc approach.

The third point is a non sequiter. Pointing out that one methodological flaw (baseline mismatch) is absent in a study does not provide any defence against claims of a different methodological flaw (abandoning prospective endpoints in favour of post hoc endpoints).

The fourth point has some merit. The fact that four studies (three Lundbeck studies and a major Biotie study) tended to show better results in high-risk drinkers makes it plausible that nalmefene could have clinical efficacy in high-risk drinkers. If the Sponsor had performed Phase II studies that led them to anticipate this problem, they could have designed a clear Phase III test of a sensible hypothesis, and they would not have had to resort to post hoc analysis. Indeed, if any one of the four studies had been designed in response to the other three, with high-risk drinkers at randomisation declared prospectively as the target population, the Sponsor would now have a much better case. (It should be noted, though, that even if the post hoc results of the pivotal studies were re-interpreted as primary, prospective results, there would still be residual concerns about the size of the efficacy effect, given that it has potentially been inflated by withdrawal bias and unblinding.)

The Sponsor's response also referred to discussions by the European Scientific Advisory Group (SAG), referring to the minutes of an EMA meeting. The relevant portion of the minutes of the EMA meeting is reproduced below:

> *Question: Is the study population representative of the population for whom nalmefene is proposed to be prescribed (or should be prescribed if the answer to question 2 is yes)? [Question 2 related to whether reduction in alcohol was a worthwhile goal]. The respective EMA Guideline recommends that, for development of medicinal products for the treatment of alcohol dependence, patients with a high or very high level of TAC at baseline should be investigated in clinical trials in order to be clearly representative for moderate to severely alcohol dependent patients. The appropriateness of post-hoc defining the target population as the subgroup of patients that is more likely to benefit from nalmefene treatment (i.e. those with a high or very high DRL at baseline and randomisation) is questioned.*
>
> *Response: Yes. The group considers that the study population is representative of the population for whom nalmefene is proposed to be prescribed. Based on the data provided, HDRL/ VHDRL patients are more likely to be the target population who could benefit from nalmefene treatment. The group recognises the validity of the post-hoc findings defining the target population. It is acknowledged that post-hoc analyses are not ideal although*

they are commonly used in clinical trials for psychiatric drugs given the high dropout rates encountered in this population.

In a slightly different context (Question 13, below), the Sponsor described the SAG position as follows:

*"**The SAG recognised the validity of the post-hoc analysis defining the target population**. Whilst it was acknowledged that post-hoc analyses are not ideal, it was stated that they are commonly used in clinical trials for psychiatric drugs, **given the high dropout rates** encountered in these populations. The SAG also acknowledged that the reduction in alcohol consumption is an appropriate goal in a subgroup of alcohol dependent patients with high or very high drinking risk level (HDRL, VHDRL) without physiological signs of withdrawal and not requiring any immediate detoxification procedure. To avoid misleading clinicians and to minimise off-label use, the group emphasized that the therapeutic indications should clearly instruct physicians (including general practitioners) to easily recognise the patients who could be the target of the drug." [Emphasis added.]*

Four comments can be made about the SAG position on post hoc analyses.

Firstly, it should be noted that there is a big difference between using a post hoc analysis to identify a target population, which the SAG accepted, and using such an analysis as the primary evidence of efficacy or citing the statistical results of a post hoc analysis with uncorrected p-values, as though the results had been obtained prospectively. It is reasonable to propose that the best chance of establishing efficacy for nalmefene will be in the high-risk subgroup, but much more debateable to propose that substantial efficacy has already been established.

Secondly, it seems inaccurate in the context of the pivotal studies to state that a post hoc analysis was necessary because of a high dropout rate. Although the dropout rate was high, particularly in Lundbeck14, the main reasons that the Sponsor found a post hoc analysis to be necessary were: 1) the primary, prospective analysis was disappointing; and 2) in retrospect, the study design was poor, failing to anticipate changes in drinking between Baseline and Randomisation.

Thirdly, just because one regulatory agency is prepared to lower their evidentiary standards to include post hoc analyses from negative studies, it does not mean that the TGA should be obliged to do the same.

Fourthly, the EMA minutes do not indicate that the statistical issues surrounding post hoc revision of endpoints were ever the focus of detailed mathematical discussion – there is no mention of p-values or the principle of adjusting for multiple endpoints or the difficulties of applying probability calculations to non-random post hoc selection of favourable data.

Further discussion of the problems with post hoc analyses is contained elsewhere in this report and the issue is also the subject of an independent statistician's report. Like the evaluator, the independent statistician did not share the Sponsor's view that post hoc analyses were appropriate to use as the pivotal Phase III evidence of efficacy for a new agent.

12.1.8.4. Conclusion

The Sponsor and the evaluator have fundamentally different views on the acceptability and applicability of post hoc analyses. Post hoc analyses can be very useful for generating hypotheses, but cannot provide robust statistical confirmation of hypotheses. The Sponsor has produced the plausible hypothesis that nalmefene may reduce alcohol intake in high-risk drinkers who continue to drink despite initial non-pharmacological measures, but this hypothesis awaits prospective confirmation.

At the very least, the Sponsor has an obligation to clinicians to describe their studies accurately, so the post hoc nature of the results should be made much more apparent in the PI, and the

actual primary prospective endpoints should be included in the PI. Where p-values are used in the current version of the PI, they should either be corrected for the use of multiple endpoints or discarded.

12.1.9. Question 9

12.1.9.1. Question

In real clinical practice, outside the artificial context of a clinical trial, how does the availability of a pharmacological treatment for alcohol dependence affect the thoroughness with which non-pharmacological measures are provided by busy clinicians?

Most subjects in the pivotal studies showed a substantial response to non-pharmacological measures. The additional clinical benefit of nalmefene demonstrated in the pivotal studies was marginal, even in an artificial setting where all subjects received the same nonpharmacological measures. If, in real clinical practice, nalmefene partially displaced nonpharmacological approaches rather than being provided in addition to non-pharmacological approaches, that marginal benefit could be negated, and the availability of nalmefene could even lead to worse outcomes by appearing to give clinicians a treatment option that is quicker and easier than time-consuming counselling. What evidence does the Sponsor have that this will not occur? What post-marketing monitoring does the Sponsor propose to assess for this effect?

12.1.9.2. Sponsor's response

The Sponsor's response to this question is complex, but incorporates a number of issues as cited below:

- "The type of psychosocial intervention for patients eligible for nalmefene is neither a structured type of therapy nor a time consuming approach. '
- 'The revised wording of the indication proposed by the Applicant as part of this Section 31 response takes into account the 2-step approach, and that nalmefene is not intended for all patients, and it clearly illustrates that the psychosocial support is the backbone of the alcohol reduction approach.'
- 'The Applicant also plans an educational campaign, specifically targeted at prescribers who are currently unfamiliar with treatment of alcohol dependence. '
- 'The Applicant is also setting up a Post-Authorisation Safety Study (PASS) in Europe that will be conducted in a real-life setting. The study will investigate the patterns of use of nalmefene in patients treated with nalmefene in routine clinical practice."

12.1.9.3. Conclusion

In general, the Sponsor's response to this question provides some reassurance that the Sponsor would at least try to encourage clinicians not to abandon non-pharmacological measures when prescribing nalmefene. It remains possible that clinicians will be tempted to take shortcuts in the treatment of alcohol-dependent patients, but the addition of one new agent for alcohol dependence is unlikely to aggravate this problem. The evaluator concedes that there is no direct evidence of a displacement effect, while nonetheless noting that it would be difficult to detect such an effect even if it existed.

It is of concern, though, that the revised wording of the indication favoured by the Sponsor is weaker than that proposed in the first-round Clinical Evaluation Report (CER1), as further discussed. Ironically, in their revisions to the wording of the indication, the Sponsor seeks to displace mention of psychosocial measures from the pre-treatment phase of patient selection, which undermines their argument that actual psychosocial measures will not be displaced by nalmefene in practice.

12.1.10. Question 10

12.1.10.1. Question

The Sponsor was asked to explain a minor discrepancy between two tables in the number of severe, related TEAEs.

12.1.10.2. Sponsor's response

In Table 105, adverse events for which information about severity was missing were classified as severe. In the Alcohol Dependence Pool, there was one patient with a related adverse event (preferred term: accidental overdose) for which information about severity was missing.

In Table 106, only patients with related adverse events with information about severity were included, which explains the discrepancy between Tables 105 and 106.

12.1.10.3. Conclusion

This issue is resolved. The discrepancy arose because of two different approaches to missing data.

12.1.11. Question 11

12.1.11.1. Question

The Sponsor should also provide a summary table for haematological indices of potential clinical concern.

12.1.11.2. Sponsor's response

The Sponsor has provided summary tables, which have now been incorporated into the body of the evaluation report, but the provided tables refer to mean values across each treatment group. A few patients with extreme values of clinical concern would not necessarily shift the mean of the whole group to a concerning value, so the provided table was of little use.

12.1.11.3. Conclusion

This issue is partially resolved. Ideally, the sponsor should provide a table listing the incidence of haematological values of clinical concern, but this issue is minor compared to the other concerns raised in this report.

12.1.12. Question 12

12.1.12.1. Question

As noted in Section 11.1, the PI should be extensively rewritten, particularly the efficacy section.

12.1.12.2. Sponsor's response

"The Applicant acknowledges the necessity to update the originally proposed text of the PI and CMI for nalmefene."

The Sponsor's detailed responses to individual criticisms of the original PI were not included in the first S31R, but were instead included in a second S31R. These responses are addressed in detail.

12.1.12.3. Conclusion

The Sponsor's new version of the proposed PI is discussed. Many of the minor issues noted in the original PI have been satisfactorily resolved in the new version. Most of the major issues remain unresolved, though, and the new version of the proposed PI is still unacceptable. In particular, the description of the pivotal studies fails to report results for the primary endpoints, and the p-values included in the PI have not been corrected for the use of multiple endpoints.

12.1.13. Question 13

12.1.13.1. Question

"Please clarify the following point. In the individual pivotal study reports, prominence was given to subgroup analysis of subjects with High or Very High DRL at Baseline [and Randomisation], but there was little or no mention of a corresponding analysis of subjects with Medium DRL at Baseline [and Randomisation]. In the clinical evaluator's description of the individual studies it was commented that this was a serious omission. In the S31 response the sponsor suggested that the evaluator comment was mistaken because, later, in a pooled analysis of both studies, results in the medium DRL subgroup were included in a couple of tables. The evaluator's comments were not related to the section that dealt with pooled data and were clearly describing the individual studies, so as far as the clinical evaluator can tell their original comments were appropriate and the sponsor's 'correction' fails to acknowledge the imbalance in their presentation of the data."

12.1.13.2. Background

The first-round Clinical Evaluation Report (CER1) contained the following comment within a description of the statistical methods of the two pivotal studies.

"Subgroup analyses were performed for the patients with a high or very high DRL at Baseline, using an MMRM approach similar to that used for the co-primary efficacy analyses, as well as with a post hoc ANCOVA using LOCF. This important subgroup is considered separately in Section 7.1.14.

An important omission in the Sponsor's submission was a corresponding subgroup analysis in patients with medium DRL at Baseline."[22]

Note that this section of CER1 acknowledges the importance of subjects with high/very-high DRL, and merely states that lack of a "corresponding subgroup analysis" of medium-DRL patients was "an important omission"; that is, the Sponsor's approach to medium-DRL and high/very-high-DRL patients was not the same, and this was important to note.

The Sponsor annotated this sentence with the comment:

> *Correction: The analyses in patients with medium DRL at baseline were included in module 2.7.3, Summary of Clinical Efficacy, Panels 116 and 117.*

The Sponsor's use of the term "Correction" in this context implies that the original comment in the CER1 was mistaken. The flagged CER1 comment appeared in a section clearly dedicated to the prospective Statistical Analysis Plans (SAPs) for the pivotal studies, and a subsequent electronic search of the individual study reports confirmed that, as claimed in CER1, there was no subgroup analysis of the medium DRL group that corresponded to the analysis of the high/very-high-DRL group. The fact that a later post hoc analysis of pooled medium-DRL subjects appeared in a couple of tables in the Summary of Clinical Efficacy (SCE) does not mean that the original CER1 comment was inaccurate; the SCE analysis of the medium DRL group does not "correspond" to the much more extensive analysis of the high/very-high-DRL group.

[22] The evaluator points out that in the context of the sentence flagged by the sponsor as an error, the various ways of defining high risk or medium risk - such as at Baseline only or at Baseline and Randomisation - are not at issue; what is being challenged is the marked asymmetry of the approach taken by the Sponsor to medium-risk vs high-risk subjects. Nonetheless, it should be noted that the original CER1 comment specifically referred to subgroup analysis of high/very-high DRL subjects at Baseline (HDAB), not to those with high/very-high DRL at Baseline and Randomisation (HDAR), as implied in the text of the TGA's question. That is, a tangenial and potentially confusing issue was added as the TGA passed on the Evaluator's concerns to the Sponsor. The overall emphasis of the Sponsor's report focussed heavily on the HDAR, which represents an approach that chooses a subgroup along the DRL risk spectrum and then again along the temporal dimension of when that risk needed to be demonstrated, but this issue is considered elsewhere.

The main reason that the asymmetrical analysis of these two subgroups was originally flagged as important in CER1 is that results in the overall patient population, which included medium-DRL subjects, were inferior to those in the high/very-high-DRL subgroup, despite the fact that subgroup analyses usually have lower statistical power than primary analyses in the total population. This divergence requires an explanation. Many hypotheses for the divergence could be entertained, but it is important to at least consider whether the lack of statistical significance in the overall population could be due to adverse results in the medium DRL group offsetting the modest benefit observed in the high-risk group. In this context, the lack of an analysis of the medium DRL subgroup is indeed an omission, and an important one, so the Sponsor's suggestion that the CER1 stood in need of "correction" appears unjustified.

Also, it should be noted that in many other contexts, medium-risk subjects represent a group of patients that are easier to treat than high-risk subjects. For instance, mild-to-moderate epilepsy may be more responsive to anticonvulsants than severe epilepsy, and mild-to-moderate Parkinsonism may be more medication-responsive than end-stage parkinsonism. Treatment of early multiple sclerosis may produce more benefit than later stages of multiple sclerosis, and so on. On the other hand, in still other contexts, higher risk subjects stand to gain the most from medication: treatment of severe hypercholesterolaemia is likely to prevent more strokes than treatment of borderline hypercholesterolaemia. The point is that it is of general interest to compare efficacy results in medium-risk subjects with those obtained in high-risk subjects, even when charges of selecting a favourable data set are not being considered, and it is appropriate to note the lack of such an analysis in a Clinical Evaluation Report. If the results in medium-risk subjects were greatly inferior to those in high-risk subjects, this would be an important point for clinicians to be aware of when deciding to prescribe for an individual patient.

12.1.13.3. Sponsor's response

The S31R comments on this issue seek to justify the greater emphasis given to the high/very-high DRL subgroup than the medium DRL subgroup in the original SAP. For instance, the Sponsor notes:

"Prior to starting the clinical phase III programme, the Applicant sought scientific advice from the EMA in 2008 (Appendix I). In the EMA scientific advice it was stated that: included patients should have a high or very high total alcohol consumption at baseline in order to be really representative for alcohol dependence in the general population However, the Applicant decided to include patients with a medium DRL in the 6-month studies (Studies 12014A and 12023A), as patients with a medium DRL also have significant risk of alcohol-related harm and thus are expected to benefit from a reduction in their alcohol consumption."

The sponsor also argues, after noting that subjects often curtailed their drinking between Baseline and Randomisation:

"Therefore, in the CHMP Day 120 List of Questions, the EMA requested analyses of patients who maintained a high or very high DRL at randomisation in order to substantiate the clinical efficacy and the clinical relevance of nalmefene, and most particularly in order to define a population where the benefit of nalmefene would be greatest (see European Public Assessment Report – Selincro). Consequently, the EU MAA was updated to also include patients with a high or very high DRL at baseline and randomization; the subgroup of patients with medium DRL was not considered."

For the most part, these and similar comments are reasonable and the evaluator agrees that the high-risk subgroup is an appropriate treatment target. Post hoc analyses of the Sponsor's

pivotal studies certainly suggest that efficacy in medium risk patients is particularly poor (see the next section).[23]

Most of this discussion, though, is simply tangential to the original concerns about a post-hoc defined target population. That a plausible argument can be mounted to target high-risk subjects is of interest, and certainly justifies future prospective studies in this subgroup, but the Sponsor's position is weakened by the fact that they did not find the same arguments compelling or obvious when they planned their studies. If it is now obvious that the correct target group is high/very-high DRL subjects who maintain a high or very-high DRL at Randomisation, then this should have been the hypothesis taken into the Phase III pivotal studies.

The fact that a post hoc rationalisation sounds biologically plausible does not change the fact that post hoc statistical evidence is weak and unreliable. If the results in the medium-DRL group had been favourable, and results in the high/very-high-DRL group had been unfavourable, the Sponsor could have mounted an (ostensibly plausible) argument that the true and obvious target for nalmefene was actually the medium-DRL group, whose alcoholism was not so intractable as the heavier drinkers, so they still had the chance of responding to pharmacological measures. The only way to avoid such claims and to achieve a statistically robust result is to perform adequate Phase II studies to identify an appropriate target group, and then to follow these up with appropriate Phase III studies that incorporate a mature understanding of the role of the drug and do not require post hoc revisions of the target population.

The Sponsor has rejected the idea that post hoc analyses are questionable, citing the Scientific Advisory Group (SAG) involved in the EMA evaluation. This has already been discussed above.

12.1.13.4. Conclusion

With respect to the question of whether subgroup analysis of the medium-DRL subgroup was omitted from the original study reports, it appears that the original comment in CER1 was accurate: “An important omission in the Sponsor’s submission was a corresponding subgroup analysis in patients with medium DRL at Baseline.”

Such an analysis of the medium DRL subgroup is included in the next section.

12.1.14. Question 14

A digital search of the individual study reports (e.g. 12014a-study-report-body.pdf) finds numerous uses of the word 'subgroup', but always in the context of a High or Very High DRL subgroup. The clinical evaluator did not find a similar analysis of the Medium DRL subgroup. Also, the study synopsis at the start of the study reports explicitly mentions subgroup analysis of one group but not the other. The clinical evaluator requests to know whether he is mistaken in concluding that the individual pivotal study reports contained subgroup analyses of the High/Very-High DRL group without a corresponding analysis of the Medium DRL subgroup. The clinical evaluator notes that it is possible that such a subgroup analysis is buried in an appendix, and would be happy to mention this if it can be located, but even then this would not count as 'corresponding' treatment of the two subgroups.

[23] The evaluator points out that EMA advice directly cited by the sponsor does not raise the issue of patients who change their drinking habits between Baseline and Randomisation, but merely refers to subjects with high risk at Baseline. It seems obvious now, in retrospect, that subjects who respond well to non-pharmacological interventions (including the simple but powerful intervention of signing up to a study and documenting their alcohol intake) are not suitable subjects for a pivotal study based on reducing alcohol intake with pharmacological measures, so it would have been appropriate to define the target population prospectively as subjects with at least high DRL at randomisation. Unfortunately, the Sponsor did not realise this until the studies had been completed.

12.1.14.1. Sponsor's response

This question, posed by the TGA, seeks to clarify one of the Sponsor's "corrections" of CER1, as discussed above for Question 13. Rather than conceding that a correction was probably not necessary, after all, the Sponsor responded with the following comment:

"Analysis of the data for the subgroup of patients with a medium DRL at baseline was not included in the individual CSRs."

Note that this non-inclusion is exactly what CER1 was referring to with the comment:

"An important omission in the Sponsor's submission was a corresponding subgroup analysis in patients with medium DRL at Baseline."

The remainder of the Sponsor's response to this question consisted of a subgroup analysis of the subgroup with medium DRL at baseline. The first table below shows the subgroup results in each pivotal study for HDDs, and the subsequent table shows the results for TAC. In Lundbeck14, at 6 months, the mean reduction in HDDs was slightly better in the nalmefene group (-6.91 days), than in the placebo group (-5.46 days), but this difference was not significant and at other time points the two treatment groups were very similar. In Lundbeck23 (which was potentially less susceptible to withdrawal bias), there was a trend in favour of placebo at all time points - including six months, when the reduction in the placebo group (-8.91 days) was slightly better than that seen with nalmefene (-8.41 days).

Table 181. Adjusted Changes from Baseline in Monthly HDDs (FAS, MMRM) – Patients with *Medium* DRL at Baseline

		Baseline						Comparison to PBO				
									95% CI			
Study	Treatment Group	N	Mean	Month	N	Mean	SE	Diff.	SE	Lower	Upper	p-value
12014A	PBO	59	11.77	1	59	-3.70	0.80					
				2	56	-5.21	0.80					
				3	55	-5.36	0.81					
				4	53	-6.19	0.77					
				5	52	-6.37	0.80					
				6	46	-5.46	0.84					
	NMF	68	11.40	1	68	-4.77	0.75	-1.07	1.04	-3.14	0.99	0.3048
				2	60	-5.75	0.77	-0.54	1.05	-2.63	1.55	0.6086
				3	54	-6.17	0.80	-0.81	1.09	-2.97	1.34	0.4559
				4	47	-6.63	0.78	-0.44	1.04	-2.51	1.63	0.6735
				5	43	-6.49	0.83	-0.12	1.10	-2.31	2.06	0.9112
				6	41	-6.91	0.86	-1.45	1.15	-3.73	0.83	0.2103
12023A	PBO	79	11.23	1	79	-7.46	0.77					
				2	70	-8.41	0.76					
				3	66	-8.36	0.83					
				4	59	-8.64	0.82					
				5	58	-8.83	0.85					
				6	54	-8.91	0.79					
	NMF	64	12.38	1	64	-6.66	0.81	0.80	0.92	-1.03	2.63	0.3865
				2	58	-7.58	0.81	0.82	0.92	-0.99	2.64	0.3714
				3	47	-6.86	0.90	1.50	1.05	-0.58	3.58	0.1564
				4	43	-7.51	0.90	1.13	1.04	-0.93	3.18	0.2799
				5	38	-7.36	0.95	1.47	1.12	-0.74	3.69	0.1913
				6	36	-8.41	0.87	0.50	0.98	-1.43	2.43	0.6080

For TAC, the results were qualitatively similar: there was a trend in favour of nalmefene in Lundbeck14, and a trend in favour of placebo for Lundbeck23.

Table 182. Adjusted Changes from Baseline in Monthly TAC (FAS, MMRM) – Patients with *Medium* DRL at Baseline

		Baseline						Comparison to PBO				
										95% CI		
Study	Treatment Group	N	Mean	Month	N	Mean	SE	Diff.	SE	Lower	Upper	p-value
12014A	PBO	59	44.3	1	59	-10.5	3.2					
				2	56	-16.1	4.3					
				3	55	-16.7	3.4					
				4	53	-20.4	3.3					
				5	52	-20.8	3.3					
				6	48	-17.2	3.6					
	NMF	68	45.2	1	68	-15.5	2.9	-5.0	3.8	-12.5	2.4	0.1819
				2	60	-16.8	4.1	-0.8	5.5	-11.7	10.2	0.8919
				3	54	-19.9	3.3	-3.2	4.2	-11.5	5.2	0.4530
				4	47	-22.3	3.2	-1.9	4.0	-9.9	6.2	0.6444
				5	43	-22.1	3.3	-1.3	4.2	-9.5	7.0	0.7614
				6	41	-24.2	3.7	-7.0	4.7	-16.3	2.4	0.1425
12023A	PBO	79	48.0	1	79	-26.3	3.1					
				2	70	-31.6	3.0					
				3	66	-31.0	3.2					
				4	59	-32.1	3.3					
				5	58	-32.7	3.3					
				6	54	-33.8	3.1					
	NMF	64	48.7	1	64	-24.5	3.2	1.8	3.1	-4.4	8.0	0.5651
				2	58	-27.9	3.2	3.7	3.1	-2.4	9.8	0.2300
				3	47	-26.7	3.5	4.3	3.6	-2.7	11.3	0.2274
				4	43	-28.4	3.6	3.7	3.7	-3.5	11.0	0.3126
				5	38	-30.1	3.5	2.6	3.6	-4.6	9.8	0.4691
				6	36	-33.1	3.4	0.7	3.3	-5.9	7.3	0.8333

NMF IDB MAA Other Countries FINAL MAA AU TGA_Q round 1 Additional requests TAC_Medium_M_1423
25JUN2014:08:53:23 SADs No.: 120/83/147 TGML v.6

The poor results in the medium-DRL subgroup were not expected prospectively, and remain largely unexplained by the Sponsor, apart from the fact that many subjects responded to enrolment in the studies by curtailing their intake even prior to Randomisation. This non-pharmacological response may have diluted the efficacy signal, making it harder to demonstrate efficacy, but it is of some concern that not even a consistent favourable trend was seen in the medium-DRL subgroup, because it would have been expected that their milder alcoholism would have been relatively easy to treat.

In this section of the S31R, the Sponsor also provided a subgroup analysis for subjects who did not fall into their post hoc target subgroup (subjects with at least high DRL at Baseline and Randomisation). Such subjects included those who only had medium DRL at Baseline (as assessed in the tables above) as well as those who had at least high DRL at Baseline but responded to enrolment in the study by reducing their intake below high DRL by the time of Randomisation.

In this subgroup, the complement of the Sponsor's proposed post hoc target group, there was no overall treatment effect, as shown in the tables below.

Table 183. Adjusted Changes from Baseline in Monthly HDDs (FAS, MMRM) – Patients Below a *High* DRL at Baseline or Randomisation

		Baseline					Comparison to PBO				
									95% CI		
Study	Treatment Group	N	Mean	Month	N	Mean	SE	Diff.	SE	Lower	Upper	p-value
12014A	PBO	122	14.75	1	122	-6.31	0.57					
				2	113	-7.35	0.57					
				3	111	-8.26	0.56					
				4	106	-8.82	0.54					
				5	104	-8.85	0.57					
				6	99	-8.21	0.64					
	NMF	119	14.24	1	119	-6.82	0.56	-0.51	0.75	-1.98	0.97	0.4987
				2	105	-7.88	0.57	-0.53	0.75	-2.01	0.95	0.4793
				3	94	-8.40	0.58	-0.14	0.75	-1.61	1.34	0.8556
				4	78	-8.54	0.58	0.27	0.74	-1.19	1.74	0.7117
				5	71	-8.49	0.62	0.36	0.80	-1.21	1.93	0.6529
				6	67	-8.46	0.71	-0.25	0.91	-2.05	1.54	0.7821
12023A	PBO	171	15.32	1	171	-11.37	0.53					
				2	151	-11.86	0.54					
				3	142	-11.87	0.59					
				4	131	-11.96	0.62					
				5	127	-11.85	0.66					
				6	118	-12.20	0.60					
	NMF	181	17.47	1	181	-11.51	0.50	-0.14	0.58	-1.28	0.99	0.8059
				2	161	-12.21	0.52	-0.36	0.60	-1.55	0.83	0.5529
				3	139	-11.51	0.58	0.35	0.70	-1.02	1.73	0.6137
				4	129	-11.36	0.61	0.60	0.74	-0.86	2.06	0.4216
				5	117	-11.54	0.66	0.31	0.82	-1.31	1.93	0.7106
				6	109	-12.33	0.60	-0.13	0.72	-1.55	1.30	0.8592

NMF IDB MAA Other Countries FINAL MAA AU TGA_0 round 1 Additional requests HDD_CTarget_M_1423
25JUN2014:08:53:38 SADs No.: 120/83/147 TGML v.6

Table 184. Adjusted Changes from Baseline in Monthly TAC (FAS, MMRM) – Patients Below a *High* DRL at Baseline or Randomisation

		Baseline					Comparison to PBO				
									95% CI		
Study	Treatment Group	N	Mean	Month	N	Mean	SE	Diff.	SE	Lower	Upper	p-value
12014A	PBO	122	65.3	1	122	-28.7	2.1					
				2	113	-32.8	2.6					
				3	111	-36.0	2.1					
				4	106	-39.1	2.0					
				5	104	-38.9	2.2					
				6	99	-37.1	2.3					
	NMF	119	58.3	1	119	-32.2	2.1	-3.5	2.7	-8.8	1.8	0.1974
				2	105	-35.1	2.6	-2.3	3.5	-9.2	4.6	0.5162
				3	94	-37.6	2.2	-1.6	2.9	-7.3	4.0	0.5654
				4	78	-38.5	2.2	0.7	2.8	-4.9	6.2	0.8129
				5	71	-37.7	2.4	1.2	3.0	-4.8	7.2	0.7005
				6	67	-39.5	2.6	-2.4	3.3	-8.9	4.2	0.4781
12023A	PBO	171	72.6	1	171	-51.1	2.2					
				2	151	-54.3	2.3					
				3	142	-55.0	2.4					
				4	131	-56.4	2.4					
				5	127	-54.8	2.6					
				6	118	-56.0	2.5					
	NMF	181	76.3	1	181	-52.5	2.1	-1.4	2.3	-6.0	3.2	0.5510
				2	161	-53.9	2.3	0.4	2.6	-4.7	5.6	0.8756
				3	139	-53.5	2.3	1.6	2.7	-3.8	6.9	0.5685
				4	129	-52.6	2.4	3.8	2.8	-1.8	9.4	0.1794
				5	117	-53.7	2.6	1.1	3.1	-4.9	7.1	0.7169
				6	109	-55.0	2.5	1.0	3.0	-4.9	6.9	0.7378

NMF IDB MAA Other Countries FINAL MAA AU TGA_0 round 1 Additional requests TAC_CTarget_M_1423
25JUN2014:08:53:46 SADs No.: 120/83/147 TGML v.6

12.1.14.2. Conclusions

The Sponsor's response indirectly confirms that CER1 did not need a correction.

The subgroup analysis of medium DRL subjects shows that efficacy in this subgroup is minimal. This is not something that could have been predicted from the proposed MOA of nalmefene, and it remains unexplained by the Sponsor.

12.2. Independent statistician's report

Because of continued disagreement between the Sponsor and evaluator over the validity of the Sponsor's post hoc statistical results, the TGA sought advice from an independent statistician.

Clarification was sought on two specific issues, the PK interaction between nalmefene and alcohol, and the validity of post hoc analyses.

Basically, the statistician has confirmed the following points already raised in this document:

- the potential PK interaction between alcohol and nalmefene is of concern, increases the uncertainty surrounding efficacy estimates, and potentially negates the treatment effect;
- the Sponsor's post hoc approach was statistically invalid;
- the Sponsor abandoned their prospective Statistical Analysis Plan (SAP) and proceeded to perform analyses that were explicitly declared to be inappropriate in the SAP;
- even in the Sponsor's preferred post hoc subgroup, the size of the apparent treatment effect was small compared to pre-study estimates of a clinically meaningful effect, such as the Sponsor's initial power calculations and WHO recommendations;
- the main efficacy parameters were potentially subject to recall bias, and may have been influenced by 'the act of treatment' more than the pharmacological effects of treatment; this relates directly to the evaluator's concerns about unblinding.

The evaluator basically agrees with the independent statistician's conclusions, which are concordant with discussions elsewhere in this document, and so no additional commentary is needed.

The statistician also flagged some other issues, as follows:

- In both pivotal clinical trials use of medication was as needed. Results suggest that there was high non-compliance with nalmefene compared to placebo, subjects drank without taking nalmefene on 13-22% of days compared to 11-13% of days with placebo.
- Both co-primary end-points were subject to recall bias (HDD and TAC). Months with less than 7 days of data were discarded, however, these were potentially months with a high number of high drinking days which would have effected recall.
- A prospective clinical trial in the high to very high drinking level target groups may want to consider a counselling only comparison arm to determine the effect of the 'act of taking a treatment' in this population. Additionally, this might inform whether counselling is provided less often after treatment than with no treatment.

12.3. Errata in first-round clinical evaluation report

A review of the Sponsor's annotations to CER1 found a mistake that appeared during editing of CER1. In a section assessing the potential PK interaction between alcohol and nalmefene, the following comment appeared:

'To put this in context, the estimated treatment effect for TAC was 5g/d in the pooled pivotal population and the mean baseline TAC was 89 g/d.'

The sentence was not initially intended to refer to the pooled population, and earlier versions of the sentence made it clear that the 5g/d estimate was based on consideration of both pivotal studies, with rejection of the Lundbeck14 results because of the high and unequal withdrawal rates in that study. "Consideration of both studies" was inadvertently transformed into "the pooled pivotal population", which the Sponsor justifiably flagged as an error.

The comment has now been revised to read as follows:

'To put this in context, the estimated treatment effect for TAC was 5g/d in Lundbeck23, the pivotal study least compromised by unequal withdrawals, and the mean baseline TAC was 89 g/d.'

The Sponsor has also clarified the nature of the scientific advice received from the EMA prior to commencement of the pivotal studies. CER1 stated that the Sponsor had designed their studies in accord with the EMA Guidelines on studies in alcohol dependence, published in 2008, but the Sponsor actually received advice from the EMA in 2008 prior to publication of the official Guidelines. This is a minor issue, because the EMA advice and the EMA Guidelines both reflect the EMA opinion on the same issue and in the same year, but it might account for some minor differences between the EMA recommendations and the Sponsor's study designs.

12.4. Sponsor's annotations of first-round clinical evaluation report

The Sponsor has extensively annotated CER1, marking what they considered to be errors. The Sponsor also wrote a letter to the TGA expressing their concerns:

> *We would kindly like to draw your attention to the fact that Lundbeck's review of the report has revealed a large number of inaccuracies and, therefore, we have made extensive comments on the content of the report.*

They listed the following major concerns:

- *Interpretation of the results of the pharmacokinetic/pharmacodynamic interaction study in healthy volunteers (Study 135 13A) where the interaction between alcohol and nalmefene was assessed. The evaluation has inappropriately applied estimates from this study to the results of the phase III studies in relation to the primary endpoint TAC.*
- *Interpretation of clinical relevance and scientific validity of the target population represented by the post hoc analysis. The evaluation places emphasis on the efficacy data related to the total population, while the target population, i.e. the population intended for treatment is essentially dismissed and not considered.*
- *Interpretation of pre-randomisation activities in the pivotal trials and subsequent influence on the evaluation of the results and impact on the proposed indication statement.*
- *The First Round Benefit-Risk Assessment, appears to be based primarily on only one of the pivotal studies, again does not take into consideration the proposed target population in its overview of the efficacy of nalmefene and completely ignores the enormous unmet clinical need that exists for a medication to treat alcohol dependence.*

The evaluator agrees that each of these issues represents a substantial point of disagreement. The first two of these issues were referred to an independent statistician who agreed with the concerns of CER1. The third issue is based on the Sponsor's reluctance to consider pre-Randomisation activities as a "psychosocial intervention", despite the fact that these activities produced a major change in drinking behaviour, one that dwarfed the apparent treatment effect.

With respect to the fourth listed issue, the Sponsor objected to the fact that, in CER1, relatively brief treatment was given to pooled efficacy analysis of the pivotal studies, and greater emphasis was given to Lundbeck23 than to Lundbeck14 in considering the Benefit-Risk balance of nalmefene. This reflects the evaluator's belief that Lundbeck14 was severely compromised by excessive and unequal withdrawals, and produced results that were not concordant with Lundbeck23 – the difference in mean TAC reduction between the two studies was greater than the treatment effect in Lundbeck23. Pooling the more compromised study (Lundbeck14) with the less compromised study (Lundbeck23) merely inflates the apparent statistical significance of the results without increasing the overall robustness of the analysis.

One annotation suggesting that a "Correction" was needed is discussed elsewhere. A couple of minor errors flagged by the Sponsor are conceded above as errata; these changes do not have any impact on the overall conclusions reached.

In most cases, the Sponsor's annotations consisted of a declaration that one or other methodological criticism was unwarranted, often with reference to the relevant part of the Sponsor's Section 31 Response. This means that the issues under contention have already been discussed in the various sections above.

A consistent and repeated theme in the annotations was the Sponsor's objection to the fact that CER1 refused to promote the Sponsor's preferred post hoc analysis of the HDAR subgroup to the same central position it occupied in the Sponsor's PI and Summary of Clinical Efficacy (SCE). This reflects the evaluator's belief that the primary material being submitted for evaluation is the set of studies conceived and conducted prospectively. Despite being labelled post hoc as "the target group", and thus having direct relevance to the Sponsor's proposed indication and PI, the HDAR subgroup has no particular privilege as a source of efficacy data compared to the rest of the study population. In this respect, it is the PI and SCE that have drifted away from the source material and sought to redefine the debate, not the Clinical Evaluation Report. Also, the Sponsor has used a range of definitions of the target group, with initial versions of the indication in the Briefing Document indicating a much broader target population than the one chosen for estimation of the treatment effect in the PI. In deference to the Sponsor's concerns, however, this post hoc subgroup has been given its own section in this, the second-round CER (CER2), and the second-round benefit-risk assessment explicitly considers this post hoc subgroup.

Another recurring theme in the Sponsor's comments was that potential methodological flaws were dismissed whenever there was evidence in the literature that the methodology chosen was broadly appropriate. For instance, the TLFB method has been endorsed by regulatory authorities as a suitable method for trying to gather accurate drinking data, and for trying to cope with the fact that heavy drinkers are poorly compliant with record-keeping. This may have led the Sponsor to have greater confidence in the efficacy results than is justified.

For instance, in response to the CER1 comment 'The TLFB and all other methods relying on patient reports are inherently prone to bias, because subjects who are embarrassed about their drinking may be motivated to reduce their estimates," the Sponsor writes in a side-note:

The use of TLFB is a validated method for collecting information on alcohol consumption. HDD and TAC were approved by the CHMP scientific advice as co-primary endpoints (derived from TFLB). TFLB and the co-primary endpoints were later endorsed by the EMA guideline on the Development of Medicinal Products for Treatment of Alcohol Dependence.

These and similar annotations are true, but tangential to the point being discussed. The implication from the Sponsor's comment appears to be that validation by the CHMP somehow protects the TLFB method from recall bias and the effects of unblinding.

Overall, on reviewing the Sponsor's various criticisms and annotations, the evaluator's original concerns – raised in CER1 and subsequently endorsed by an independent statistician – remain undiminished.

13. Second round benefit-risk assessment

Consideration of the Sponsor's Section 31 Response does not substantially modify the benefit-risk assessment of nalmefene, as originally assessed in CER1.

CER1 suggested the following potential benefits of nalmefene (based on the prospective results of Lundbeck23, the pivotal study least susceptible to withdrawal bias, and Lundbeck13, which was broadly concordant with Lundbeck23, but without any correction for possible biases): about 1-2 heavy-drinking days per month might be converted to moderate-drinking days about 3.5-5g of alcohol might be avoided per day, but the alcohol that is consumed could have an increased AUC, negating much of this small benefit CER1 also suggested the following risks: busy clinicians could trust the nalmefene to reduce alcohol intake, and cut back on effective

psychosocial treatments, leading to increased alcohol intake compared to standard care. subjects taking nalmefene are at increased risk of dizziness, nausea, fatigue and insomnia.

Of the potential benefits listed above, the size of the therapeutic effect could prove to be greater than that cited if the Sponsor's results in their favoured subgroup turned out to be representative of the results obtained in similar patients prospectively. The reduction in total alcohol consumption could be approximately 10g, based on the more reliable pivotal study (or 18g in the less reliable study) – assuming that there has been no inflation of this estimate through withdrawal bias, or unblinding. About 2.7 to 3.7 heavy drinking days might also be converted to moderate drinking days.

Even in this favoured subgroup, the benefit demonstrated would be clinically modest, and less than one drinking risk level in the WHO schema. It would also be substantially less than the benefit anticipated in the Sponsor's power calculations.

On the other hand, the benefit of nalmefene could be substantially less than the Sponsor's estimate if unblinding had a significant effect on subjects drinking reports, which seems very likely, or if the results have been inflated by withdrawal bias.

Of the risks listed above, the first could potentially be down-graded. The Sponsor has provided some argument to the effect that clinicians will not take the shortcut of prescribing nalmefene instead of providing full psychosocial support, and it must be conceded that there is no clear evidence that this will actually represent a major risk in practice. It is of concern, however, that the Sponsor's Section 31 Response has suggested a new wording for the indication that de-emphasises the need for psychosocial input prior to prescription, and explicitly rejects the evaluator's suggestion that psychosocial treatment should be tried as the firstline approach.

The third risk listed above, the undesirable, precedent-setting nature of the Sponsor's PI, seems understated in retrospect. It would be completely unacceptable to approve a PI in which the primary results of the pivotal studies were suppressed, and p-values were cited that were known to be invalid.

14. Second round recommendation regarding authorisation

The application to register nalmefene should be rejected.

Should the sponsor choose to perform additional efficacy studies of nalmefene, with the intention of overcoming the deficiencies in the current submission, the following methodological features should be incorporated into their study program:

- a single unambiguous primary endpoint should be declared prospectively in a clearly defined target group, without any post hoc revisions;
- all additional (secondary and tertiary) endpoints should be declared prospectively and either ranked for hierarchical testing or adjusted for multiplicity;
- subjects who reduce drinking below high DRL prior to randomisation should be excluded;
- blinding should be assessed by asking subjects to guess their assigned treatment;
- strong consideration should be given to a non nalmefene active treatment arm using an agent known to produce a similar incidence of side effects; this agent need not have any particular effect on alcohol addiction but would allow estimation of the potential for side effects to alter drinking reports;
- strong consideration should be given to a psychosocial treatment arm (counselling-only arm) to allow assessment of the effect of psychosocial treatment as distinct from patient responses to the act of taking a tablet;

- some attempt should be made to follow the drinking habits of subjects after they withdraw (it is conceded that this could be difficult);
- the PI should restrict itself to the reporting of prospectively declared endpoints, which should be reported in full.

15. Sponsor's response to second round recommendations

Following receipt of the second round evaluation report and the independent statistician's report, and in recognition of the considerable differences remaining between the sponsor and the clinical evaluator, the sponsor requested an opportunity to respond to the Delegate. A comprehensive response document addressing key issues and an expert statement from a reputable Australian biostatistician were submitted for consideration by the Delegate prior to the request for ACPM advice. The response document was also submitted as an appendix in the sponsor's pre ACPM response.

15.1. Alcohol interaction study

The response document included further information to address the clinical evaluator's concern that nalmefene may increase exposure to alcohol based on the data obtained in the interaction study in healthy volunteers (Study 13513A). The results from the study showed that the 90% and 95% CI for alcohol AUC_{0-t} were fully contained within the range of 0.80 to 1.25. Excluding 2 outlier estimates of alcohol AUC_{0-t} due to few available data points changed the estimate to 1.0306 (90% CI [0.9604; 1.1060]). In addition, nalmefene did not inhibit alcohol dehydrogenase *in vitro*. Thus, even though the Applicant acknowledges that the data points measured in the interaction study did not allow for an exact calculation of alcohol C_{max} and AUC_{0-t} for a precise evaluation of a potential effect of nalmefene on alcohol exposure, the sponsor considers it unlikely that nalmefene increases the exposure of alcohol. As stated in the Delegate's overview, the sponsor's response to the alcohol interaction issue in the Clinical Evaluation Report Round 2 has:

> *satisfactorily responded to the evaluator's concerns that nalmefene may increase exposure to alcohol in individuals who do not reduce alcohol consumption while taking nalmefene.*

Thus, the current approved PI text states:

> *There is no clinically relevant pharmacokinetic drug-drug interaction between nalmefene and alcohol.*

16. References

The following references are cited directly in this report:

- Sadock BJ, Sadock VA. Kaplan and Sadock's Comprehensive textbook of psychiatry. Philadelphia: Lippincott Williams & Wilkins, 2004.
- Miller WR, Walters ST, Bennett ME. How effective is alcoholism treatment in the United States? J Stud Alcohol 2001; 62: 211-220.
- Gual A, Lligoña A, Colom J. Five-year outcome in alcohol dependence. A naturalistic study of 850 patients in Catalonia. Alcohol. 1999; 34: 183-192.
- Sobell LC, Sobell MB. Timeline Follow-back: a technique for assessing self-reported ethanol consumption. In: Litten RZ, Allen JP, editors. Measuring alcohol consumption: psychosocial and biological methods. Totowa, NJ, US: Humana Press; 1992. p 41-72.

The references below were cited in the Clinical Overview, and consulted where necessary in the preparation of this report.

1. Leshner AI. Addiction is a brain disease, and it matters. Science. 1997; 278: 45-47.
2. American Psychiatric Association (APA). Diagnostic and Statistical Manual of Mental Disorders. Fourth Ed., Text Revision (DSM-IV-TR). Washington DC: American Psychiatric Association; 2000.
3. Edwards G, Gross MM. Alcohol dependence: provisional description of a clinical syndrome. Br Med J. 1976; 1: 1058-1061.
4. World Health Organization (WHO). International Statistical Classification of Diseases and Related Health Problems. 10th Revision (ICD-10). Geneva, Switzerland: World Health Organization; 1992.
5. Gunzerath L, Hewitt BG, Li TK, Warren KR. Alcohol research: past, present, and future. Ann N Y Acad Sci. 2011; 1216: 1-23.
6. Prescott CA, Kendler KS. Genetic and environmental contributions to alcohol abuse and dependence in a population-based sample of male twins. Am J Psychiatry. 1999; 156: 34-40.
7. Caetano R, Tam T, Greenfield T, Cherpitel C, Midanik L. DSM-IV alcohol dependence and drinking in the U.S. population: a risk analysis. Ann Epidemiol. 1997; 7: 542-549.
8. Dawson DA, Goldstein RB, Chou SP, Ruan WJ, Grant BF. Age at first drink and the first incidence of adult-onset DSM-IV alcohol use disorders. Alcohol Clin Exp Res. 2008; 32: 2149- 2160.
9. Regier DA, Farmer ME, Rae DS, Locke BZ, Keith SJ, Judd LL. Comorbidity of mental disorders with alcohol and other drug abuse. JAMA. 1990; 264: 2511-2518.
10. Hasin DS, Stinson FS, Ogburn E, Grant BF. Prevalence, correlates, disability, and comorbidity of DSM-IV alcohol abuse and dependence in the United States: results from the National Epidemiologic Survey on Alcohol and Related Conditions. Arch Gen Psychiatry. 2007; 64: 830- 842.
11. Cargiulo T. Understanding the health impact of alcohol dependence. Am J Health Syst Pharm. 2007; 64 (5 Suppl 3): S5-11.
12. Anderson P, Baumberg B, editors. Alcohol in Europe. A public health perspective. A report for the European Commission. London, UK: Institute of Alcohol Studies; 2006.

13. Glanz J, Grant B, Monteiro M, Tabakoff B, WHO/ISBRA Study on State and Trait Markers of Alcohol Use and Dependence Investigators. WHO/ISBRA Study on State and Trait Markers of Alcohol Use and Dependence: analysis of demographic, behavioral, physiologic, and drinking variables that contribute to dependence and seeking treatment. International Society on Biomedical Research on Alcoholism. Alcohol Clin Exp Res. 2002; 26: 1047-1061.
14. World Health Organization (WHO). Global status report on alcohol and health. 2011.
15. Mäkelä P, Gmel G, Grittner U, Kuendig H, Kuntsche S, Bloomfield K, Room R. Drinking patterns and their gender differences in Europe. Alcohol Alcohol Suppl. 2006; 41: i8-i18.
16. World Health Organization (WHO). Global strategy to reduce the harmful use of alcohol. 2010.
17. World Health Organization (WHO). Global health risks: mortality and burden of disease attributable to selected major risks. 2009.
18. Rehm J, Baliunas D, Borges GL, Graham K, Irving H, Kehoe T, et al. The relation between different dimensions of alcohol consumption and burden of disease: an overview. Addiction. 2010; 105: 817-843.
19. Rehm J, Room R, Graham K, Monteiro M, Gmel G, Sempos CT. The relationship of average volume of alcohol consumption and patterns of drinking to burden of disease: an overview. Addiction. 2003; 98: 1209-1228.
20. Guiraud V, Amor MB, Mas JL, Touzé E. Triggers of ischemic stroke: a systematic review. Stroke.2010; 41: 2669-2677.
21. Klatsky AL, Armstrong MA. Alcohol use, other traits, and risk of unnatural death: a prospective study. Alcohol Clin Exp Res. 1993; 17: 1156-1162.
22. Mukamal KJ, Kawachi I, Miller M, Rimm EB. Drinking frequency and quantity and risk of suicide among men. Soc Psychiatry Psychiatr Epidemiol. 2007; 42: 153-160.
23. Rehm J, Gmel G. Patterns of alcohol consumption and social consequences. Results from an 8-year follow-up study in Switzerland. Addiction. 1999; 94: 899-912.
24. English DR, Holman CDJ, Milne E, Winter MG, Hulse GK, Codde JP, et al. The quantification of drug caused morbidity and mortality in Australia. Canberra, Australia: Australian Government Publishing Service; 1995.
25. European Medicines Agency (EMA). Guideline on the development of medicinal products for the treatment of alcohol dependence. EMA/CHMP/EWP/20097/2008. 2010.
26. World Health Organization (WHO). International guide for monitoring alcohol consumption and related harm. 2000.
27. Alonso J, Angermeyer MC, Bernert S, Bruffaerts R, Brugha TS, Bryson H, et al. Use of mental health services in Europe: results from the European Study of the Epidemiology of Mental Disorders (ESEMeD) project. Acta Psychiatr Scand Suppl. 2004; 109 (suppl 420): 47-54.
28. Kadden R, Carroll K, Donovan D, Cooney N, Monti P, Abrahams D, et al. Cognitive behavioural coping skills therapy manual. A clinical research guide for therapists treating individuals with alcohol abuse and dependence. Rockville, MD: National Institute on Alcohol Abuse and Alcoholism; Reprinted 2003.
29. Miller WR, Zweben A, DiClemente CC, Rychtarik RG. Motivational enhancement therapy manual. A clinical research guide for therapists treating individuals with alcohol abuse and dependence. Rockville, MD: National Institute on Alcohol Abuse and Alcoholism; Reprinted 1999.

30. Nowinski J, Baker S, Carroll K. Twelve step facilitation therapy manual. Rockville, MD: National Institute on Alcohol Abuse and Alcoholism. A clinical research guide for therapists treating individuals with alcohol abuse and dependence; Reprinted 1999.

31. Project MATCH Research Group. Matching alcoholism treatments to client heterogeneity: project MATCH posttreatment drinking outcomes. J Stud Alcohol. 1997; 58: 7-29.

32. Pettinati HM, Weiss RD, Miller WR, Donovan D, Ernst DB, Raunsaville BJ, editors. Medical management treatment manual. A clinical research guide for medically trained clinicians providing pharmacotherapy as part of the treatment for alcohol dependence. Combine Monograph Series, volume 2. Bethesda, MD: U.S. Department of Health and Human Services; 2004.

33. Johnson BA, Ait-Daoud N, Bowden CL, DiClemente CC, Roache JD, Lawson K, et al. Oral topiramate for treatment of alcohol dependence: a randomised controlled trial. Lancet. 2003; 361: 1677-1685.

34. Bien TH, Miller WR, Tonigan JS. Brief interventions for alcohol problems: a review. Addiction. 1993; 88: 315-336.

35. O'Malley SS, O'Connor PG. Medications for unhealthy alcohol use. Alc Res Health. 2011; 33: 300-312.

36. Bordin ES. The generalizability of the psychoanalytic concept of the working alliance. Psychotherapy. 1979; 16: 252-260.

37. Schuckit MA. Alcohol-use disorders. Lancet. 2009; 373: 492-501.

38. Hilsenroth MJ, Cromer TD. Clinician interventions related to alliance during the initial interview and psychological assessment. Psychotherapy. 2007; 44: 205-218.

39. Marlatt GA, Witkiewitz K. Harm reduction approaches to alcohol use: health promotion, prevention, and treatment. Addict Behav. 2002; 27: 867-886.

40. Luquiens A, Reynaud M, Aubin HJ. Is controlled drinking an acceptable goal in the treatment of alcohol dependence? A survey of French alcohol specialists. Alcohol Alcohol. 2011; 46: 586-591.

41. Heather N, Adamson SJ, Raistrick D, Slegg GP, UKATT Research Team. Initial preference for drinking goal in the treatment of alcohol problems: I. Baseline differences between abstinence and non-abstinence groups. Alcohol Alcohol. 2010; 45: 128-135.

42. Adamson SJ, Heather N, Morton V, Raistrick D, UKATT Research Team. Initial preference for drinking goal in the treatment of alcohol problems: II. Treatment outcomes. Alcohol Alcohol. 2010; 45: 136-142.

43. Anderson P, Gual A, Colom J, editors. Alcohol and primary health care. Clinical guidelines on identification and brief interventions. Barcelona, Spain: Department of health of the government of Catalonia; 2005.

44. Leon DA, Chenet L, Shkolnikov VM, Zakharov S, Shapiro J, Rakhmanova G, et al. Huge variation in Russian mortality rates 1984–94: artefact, alcohol, or what? Lancet. 1997; 350: 383- 388.

45. Zatonski WA, Sulkowska U, Manczuk M, Rehm J, Boffetta P, Lowenfels AB, et al. Liver cirrhosis mortality in Europe, with special attention to Central and Eastern Europe. Eur Addict Res. 2010; 16: 193-201.

46. Xin X, He J, Frontini MG, Ogden LG, Motsamai OI, Whelton PK. Effects of alcohol reduction on blood pressure: a meta-analysis of randomized controlled trials. Hypertension. 2001; 38: 1112- 1117.

47. Sadock BJ, Sadock VA. Kaplan and Sadock's Comprehensive textbook of psychiatry. Philadelphia: Lippincott Williams & Wilkins, 2004.

48. Miller WR, Walters ST, Bennett ME. How effective is alcoholism treatment in the United States? J Stud Alcohol 2001; 62: 211-220.

49. Gual A, Lligoña A, Colom J. Five-year outcome in alcohol dependence. A naturalistic study of 850 patients in Catalonia. Alcohol Alcohol. 1999; 34: 183-192.

50. Sobell MB, Sobell LC. It is time for low-risk drinking goals to come out of the closet. Addiction. 2011; 106: 1715-1717.

51. World Health Organization (WHO). European action plan to reduce the harmful use of alcohol 2012–2020. EUR/RC61/13. 2011.

52. Commission of the European Communities. Communication: An EU strategy to support Member States in reducing alcohol related harm. 2006.

53. Rosenthal RN. Current and future drug therapies for alcohol dependence. J Clin Psychopharmacol. 2006; 26: s20-s29.

54. Gastfriend DR, Garbutt JC, Pettinati HM, Forman RF. Reduction in heavy drinking as a treatment outcome in alcohol dependence. J Subst Abuse Treat. 2007; 33: 71-80.

55. ICH. ICH Harmonised Tripartite Guideline E6: Guideline for Good Clinical Practice. 1996.

56. United States Food and Drug Administration (FDA). Guidance for Industry: Food-effect bioavailability and fed bioequivalence studies. December 2002.

57. Dixon R, Gentile J, Hsu HB, Hsiao J, Howes J, Garg D, et al. Nalmefene: safety and kinetics after single and multiple oral doses of a new opioid antagonist. J Clin Pharmacol. 1987; 27: 233-239.

58. Liao RF, Zeng ZP, Wen YG. Pharmacokinetics of nalmefene after a single or multiple intravenous doses in Chinese healthy volunteers. Nan Fang Yi Ke Da Xue Xue Bao. 2008; 28: 1816-1819.

59. Kim S, Wagner HN Jr, Villemagne VL, Kao PF, Dannals RF, Ravert HT, et al. Longer occupancy of opioid receptors by nalmefene compared to naloxone as measured in vivo by a dual-detector system. J Nucl Med. 1997; 38: 1726-1731.

60. Gal TJ, DiFazio CA, Dixon R. Prolonged blockade of opioid effect with oral nalmefene. Clin Pharmacol Ther. 1986; 40: 537-542.

61. Jones HE, Johnson RE, Fudala PJ, Henningfield JE, Heishman SJ. Nalmefene: blockade of intravenous morphine challenge effects in opioid abusing humans. Drug Alcohol Depend. 2000; 60: 29 - 37.

62. Fudala PJ, Heishman SJ, Henningfield JE, Johnson RE. Human pharmacology and abuse potential of nalmefene. Clin Pharmacol Ther. 1991; 49: 300-306.

63. Grant SA, Millar K, Kenny GN. Blood alcohol concentration and psychomotor effects. Br J Anaesth. 2000; 85: 401-406.

64. European Commission (EC) [Internet]. Road Safety: The Legal Limit.

65. Sobell LC, Sobell MB. Alcohol timeline followback (TFLB). In: Rush AJ, First MB, Blacker D, editors. Handbook of psychiatric measures. Arlington, VA: American Psychiatric Publishing, Inc.; 2008.

66. Maisto SA, Sobell MB, Cooper AM, Sobell LC. Test-retest reliability of retrospective self-reports in the three populations of alcohol abusers. J Behav Assess. 1979; 1: 315-326.

67. Babor TF, Steinberg K, Anton R, Del Boca D. Talk is cheap: measuring drinking outcomes in clinical trials. J Stud Alcohol. 2000; 61: 55-63.

68. Volpicelli JR, Pettinati HM, McLellan AT, O'Brien CP. Combining medication and psychosocial treatments for addictions; the BRENDA Approach. New York, NY: The Guilford Press; 2001.

69. Starosta AN, Leeman RF, Volpicelli JR. The BRENDA model: integrating psychosocial treatment and pharmacotherapy for the treatment of alcohol use disorders. J Psychiatr Pract. 2006; 12: 80- 89.

70. Siddiqui O, Hung HMJ, O'Neill R. MMRM vs LOCF: A comprehensive comparison based on simulation study and 25 NDA datasets. J Biopharm Stat 2009; 19: 227-246.

71. Little R, Yau L. Intent-to-treat analysis for longitudinal studies with drop-outs. Biometrics.1996; 52: 1324-1333.

72. Dawson DA, Grant BF, Stinson FS, Chou PS, Huang B, Ruan WJ. Recovery from DSM-IV alcohol dependence: United States, 2001-2002. Addiction. 2005; 100: 281-292.

73. Bischof G, Rumpf HJ, Meyer C, Hapke U, John U. Influence of psychiatric comorbidity in alcohol-dependent subjects in a representative population survey on treatment utilization and natural recovery. Addiction. 2005; 100: 405-413.

74. Project MATCH Research Group. Matching alcoholism treatments to client heterogeneity: project MATCH three-year drinking outcomes. Alcohol Clin Exp Res. 1998; 22: 1300-1311.

75. Epstein EE, Drapkin ML, Yusko DA, Cook SM, McCrady BS, Jensen NK. Is alcohol assessment therapeutic? Pretreatment change in drinking among alcohol-dependent women. J Stud Alcohol. 2005; 66: 369-378.

76. Litten RZ, Fertig JB, Falk DE, Ryan ML, Mattson ME, Collins JF, et al. A double-blind, placebocontrolled trial to assess the efficacy of Quetiapine Fumarate XR in very heavy-drinking alcoholdependent patients. Alcohol Clin Exp Res. 2011; 1-11.

77. Anton RF, O'Malley SS, Ciraulo DA, Cisler RA, Couper D, Donovan DM, et al. Combined pharmacotherapies and behavioral interventions for alcohol dependence: the COMBINE study: a randomized controlled trial. JAMA. 2006; 295: 2003-2017.

78. Karhuvaara S, Simojoki K, Virta A, Rosberg M, Löyttyniemi E, Nurminen T, et al. Targeted nalmefene with simple medical management in the treatment of heavy drinkers: a randomized double-blind placebo-controlled multicenter study. Alcohol Clin Exp Res. 2007; 31: 1179-1187.

79. Miller WR, Tonigan JS, Longabaugh R. The drinker inventory of consequences (DrInC): An instrument for assessing adverse consequences of alcohol abuse. Project MATCH Monograph Series, Vol. 4. Rockville, MD: National Institute on Alcohol Abuse and Alcoholism; Reprinted 2006.

80. National Institute for Health and Clinical Excellence (NICE) [Internet]. Alcohol dependence and harmful alcohol use: Appendix 17d – pharmacological interventions forest plots. 2011.

81. Leucht S, Hierl S, Kissling W, Dold M, Davis JM. Putting the efficacy of psychiatric and general medicine medication into perspective: review of meta-analyses. Br J Psych 2012; 200:97-106.

82. Günther OH, Roick C, Angermeyer MC, König HH. Responsiveness of EQ-5D utility indices in alcohol-dependent patients. Drug Alcohol Depend. 2008; 92:291-5.

83. Grant JE, Odlaug BL, Potenza MN, Hollander E, Kim SW. Nalmefene in the treatment of pathological gambling: multicentre, double-blind, placebo-controlled study. Br J Psychiatry 2010; 197: 330-331.

Therapeutic Goods Administration

PO Box 100 Woden ACT 2606 Australia
Email: info@tga.gov.au Phone: 1800 020 653 Fax: 02 6232 8605
https://www.tga.gov.au