ORIGINAL RESEARCH

Comparison of modafinil and pitolisant in narcolepsy: a non-inferiority meta-analytical approach

Philippe Lehert, Dr Ir PhD[1,2], Cassandra Szoeke MD, GAICD, PhD, FRACP, AMA(M), MBBS[2]

[1]Faculty of Economics, University of Louvain, Mons, Belgium; [2]Faculty of Medicine, University of Melbourne, Melbourne, Victoria, Australia

Abstract

Aim: To assess the non-inferiority of pitolisant, a new compound for the relief of excessive daytime sleepiness (EDS) and cataplexy in narcolepsy, compared with modafinil.

Methods: Randomized controlled trials (RCTs) in narcolepsy were searched systematically. Network meta-analysis (NMA) compared the efficacy and safety of pitolisant and modafinil. The main endpoints are Epworth Sleepiness Scale (ESS), Maintenance of Wakefulness Test (MWT), the number of cataplexies, and overall safety.

Results: Of 312 articles after removing duplicates, 10 RCTs were eligible for NMA. For ESS, a non-significant superior beneficial decrease (−0.69, [−2.18, 0.79]) showed non-inferiority of pitolisant (non-inferiority margin [NIM]=1, p=0.015). An MWT beneficial increase (2.12 minutes [−0.95, 5.19]; p=0.18) showed non-inferiority of pitolisant (NIM=−1). For cataplexy, the mean beneficial effect of pitolisant was significant, providing evidence of pitolisant superiority in addition to non-inferiority. The risk ratio (RR) of treatment-suspected adverse events for pitolisant/modafinil was 0.86 [0.44, 1.24] favoring pitolisant, confirming non-inferiority considering a safety margin of RR=1.25 (tolerance of 25%).

Conclusions: Pitolisant is non-inferior to modafinil in relieving EDS, but superior to modafinil in reducing cataplexy, outranking modafinil in narcolepsy type-1 patients. Despite a slight superiority of pitolisant in EDS relief, both drugs perform equally in narcolepsy type-2 patients.

Keywords: cataplexy, daytime sleepiness, daytime wakefulness, Epworth Sleepiness Scale, modafinil, narcolepsy, network meta-analysis, pitolisant.

Introduction

Narcolepsy is a chronic neurological sleep disorder, the key characteristics of which are overwhelming sleepiness during the daytime (excessive daytime sleepiness [EDS]), cataplexy, sleep-related hallucinations, sleep paralysis, and disturbed nighttime sleep. Due to the fact that narcolepsy has a marked impact on patients' health-related quality of life, affected patients often show signs of psychosocial distress.[1,2] In addition, there is an elevated risk of comorbidity and mortality in patients with narcolepsy.[3,4] According to the International Classification of Sleep Disorders, there are two main types of narcolepsy: type-1 (previously referred to as narcolepsy with cataplexy) and type-2 (narcolepsy without cataplexy).[5]

The efficacy, safety, and benefit/risk (BR) ratio of pharmacotherapy for narcolepsy have recently been evaluated in a network meta-analysis (NMA) of published randomized controlled trials (RCTs).[6] With regard to EDS, modafinil, 200–400 mg/day, sodium oxybate, 9 g/day, and pitolisant, up to 40 mg/day, demonstrated similar clinical efficacy and were significantly more effective than placebo. With regard to cataplexy, sodium oxybate and pitolisant, but not modafinil, were shown to have similar beneficial efficacy. The meta-analysis showed that pitolisant had a marginally better safety profile and the best BR ratio. However, differences between these treatments did not reach a strict significant level of superiority.

Instead of testing the superiority of treatments, another approach is to assess non-inferiority when determining the extent to which a treatment is as good as another without necessarily being better. When using an active treatment as a comparative control instead of placebo, non-inferiority is often a more realistic conjecture, as the difference in efficacy between active treatments might be small. Modafinil was the first treatment registered for narcolepsy and, as such,

provides the largest body of evidence for efficacy gained throughout an extensive research program and numerous RCTs. The aim of the present study, conducted on the basis of regulatory requirements, was to assess the non-inferiority of pitolisant *versus* modafinil used as control. Due to the fact that data for the current non-inferiority meta-analysis are derived from the same studies, which were evaluated in the previous superiority meta-analysis,[6] and many of the methodological and statistical aspects are identical, some aspects of this article overlap for the sake of accuracy and comparison.

Materials and methods
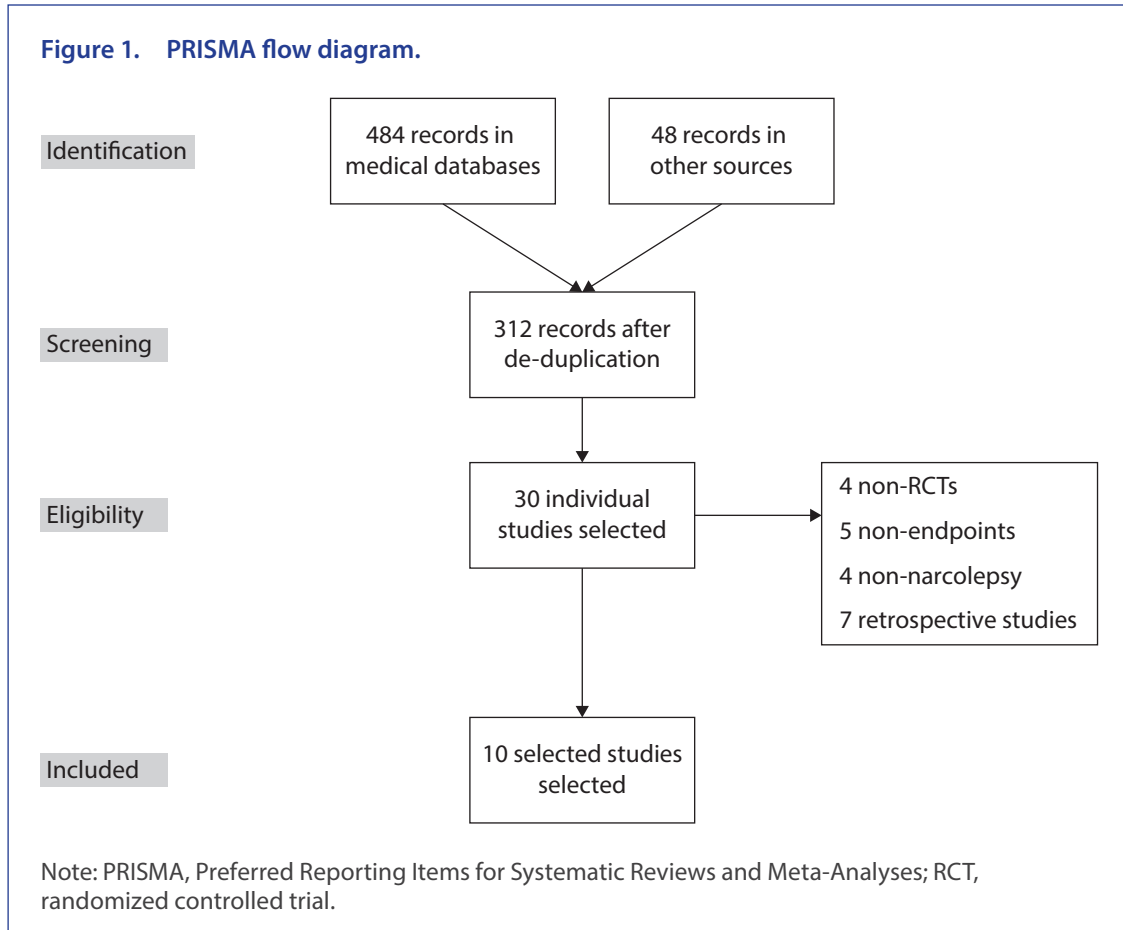
Protocol and registration

The protocols for this systematic literature review and meta-analysis adhered to Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Figure 1).[7] Prior to commencing the statistical analyses, the statistical analysis plan was locked and registered (ID: 167087) in the Prospective Register of Systematic Reviews (www.crd.york.ac.uk/PROSPERO/) database from the UK National Health Service National Institute for Health Research.

Study selection, compared intervention, and patient eligibility

As shown in previous research, there were no efficacy differences regarding EDS between the two modafinil doses (200 mg/day or 400 mg/day);[8,9] thus, the efficacy of modafinil was assessed irrespective of the dosage. Pitolisant has been evaluated at maximum doses of 20 mg/day and 40 mg/day, with the latter dosage considered to be superior.[6] Thus, all of our analyses compared only pitolisant titrated up to a maximum dose of 40 mg/day (designated as PIT) and modafinil up to 400 mg/day (MDF), or placebo. Any clinical study which compared one of these two active treatments, either directly or *versus* placebo, was selected, under the condition that it reported at least one of the following efficacy/safety outcome measures during treatment exposure: Epworth Excessive Sleepiness Score (ESS), Maintenance of Wakefulness Test (MWT), number of cataplexy attacks, and/or adverse events (AEs). Adult patients (male or female and of any age) with/without cataplexy were eligible for inclusion.

Information sources and literature search

An extensive search of published literature (full articles, abstracts, and books), in any language and regardless

Figure 1. PRISMA flow diagram.

Note: PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses; RCT, randomized controlled trial.

of publication date, regarding the efficacy/safety of pharmacotherapy for narcolepsy was conducted; all cited references underwent manual checking. The keywords 'modafinil' or 'armodafinil' and 'narcolepsy,' 'pitolisant,' 'modafinil,' and 'narcolepsy' were used to search literature in the following electronic databases: MEDLINE/PubMed, Database of Abstracts of Reviews of Effects (Cochrane Library), Cochrane Central Register of Controlled Trials (Cochrane Library), Cochrane Database of Systematic Reviews (Cochrane Library), ClinicalTrials.gov, World Health Organization International Clinical Trials Registry Platform search portal, and US Food and Drug Administration and European Medicines Agency websites. Published RCTs were selected if they included the treatment of narcolepsy in adult patients. After retrieving and reviewing an initial abstract list of potential articles, individual publications that met the study inclusion criteria were reviewed independently in full by the study authors, and all data from each publication were reviewed in a systematic manner.

Data items

Data collected for individual studies included the year of publication, study design (e.g. randomization and concealment procedures), patient sample size and disposition, intent-to-treat population, and outcomes (e.g. ESS, MWT, cataplexy, and AEs). Study heterogeneity was also reviewed/evaluated.

Risk of bias in individual studies

Three validity domains (internal, external, and statistical validity) were used to assess the methodological quality of selected studies. Two reviewers conducted the domain-based assessments (17 items), with any inter-reviewer discordance being discussed/resolved via consensus and summarized using an internal validity score (seven items), an external validity score (five items), and a statistical validity score (five items). In cases where individual studies had four or more inadequate items, we conducted a sensitivity analysis with/without these items; when both selections generated the same conclusion, the results were considered to be reliable.

Endpoints

EDS was evaluated using ESS and MWT. The Weekly Rate of Cataplexy (WRC) was used to measure cataplexy. These three separately analyzed endpoints constitute the most important and reported endpoints in cataplexy studies. We also provide aggregate (composite) scores to synthesize the various parameters of efficacy. Based on the Z scores for the three endpoints: ESS_z, MWT_z, WRC_z, we defined the narcolepsy type-2 score (NS2) as the mean of ESS_z and MWT_z, and the narcolepsy type-1 score (NS1) as the mean of NS2 and WRC_z. The NS2 score only aggregates EDS symptoms (in combining ESS and MWT), related with narcolepsy type-2 patients characterized only by EDS symptoms, whereas the NS2 score accounts both for EDS and cataplexy and is thus appropriate for narcolepsy type-2 patients.

The incidence of treatment-emergent adverse events (TEAEs) was used to estimate treatment safety. TEAEs were categorized as: (A) central nervous system (CNS): anxiety, confusion, dizziness, nervousness, psychiatric disorders, sleep disorders; (B) gastrointestinal: abdominal pain, anorexia, constipation, diarrhea, dry mouth, dyspepsia, gastrointestinal pain, nausea, vomiting; and (C) other: asthenia, fatigue, headache, hypothermia, infection, pain, pyrexia, weakness. The overall safety score (OSS), defined as the incidence rate of TEAEs during the treatment exposure period, was the key safety endpoint.

Finally, we attempted to measure the BR ratio through a composite score[6] evaluating the extent to which the observed efficacy (measured by NS1 and NS2 scores) of a treatment compensates the expected generated adverse reactions (measured by the OSS score). Although simple ratios like NS1/OSS and NS2/OSS are possible, correlation is often expected between efficacy and safety; thus, we considered the unit-less BR ratios, BR1 and BR2, defined as the residual values of the linear fit between OSS and NS1 and NS2, respectively.

Non-inferiority

When comparing active treatments, the non-inferiority approach is often more appropriate than testing for superiority, thus justifying the present meta-analysis to be considered as an extension of our first meta-analysis.[6] A non-inferiority test assesses the extent to which a test drug cannot be less than a prespecified threshold (called the non-inferiority margin, NIM) compared with a control drug. Before the analysis, all NIMs were prespecified in the protocol of this meta-analysis: ESS is a scale varying from 0 to 24, and a value of 1 was considered as the NIM. For MWT, we considered that 1 minute was the minimum clinically relevant difference and this value was associated with the NIM. For cataplexy, the minimum clinically relevant difference of 1 WRC was also defined as the prespecified NIM. For safety, the non-inferiority of PIT will be determined based on an NIM of the ratio of the OSS PIT/MDF predetermined to a maximum of an excess of 25% (NIM=1.25). The NIM for Z-scores (narcolepsy scores [NSs] and BR ratio) was −0.2 in conformity with a Cohen effect size empirical value of 0.2 assimilated to the smallest clinically relevant effect size.[10] The overall non-inferiority of PIT compared with MDF was confirmed when non-inferiority was demonstrated for the three main endpoints (ESS, MWT, and WRC).

Summary measures

With the exception of endpoints with heterogeneous noncombinable parameters (e.g. cataplexy), comparisons of continuous variables, including ESS, MWT, and Z-scores, were performed using weighted-mean differences. Due to the fact that, in the reviewed studies, cataplexy was reported using a range of different nonconvertible statistics, we used standardized mean differences. With regard to safety, the incidence risk ratio (IRR) was used to compare TEAE rates.

Synthesis of results

A classical meta-analysis comparing two drugs was not appropriate because direct comparisons between the two treatments constituted a small proportion of the studies, with most having assessed them *versus* placebo. An NMA was needed to account both for the direct comparisons and for indirect comparisons with placebo. In the network evidence graph,[11,12] each node in the network is associated with a treatment. An overlap (edge) for any two treatments represents a direct comparison, the degree of the overlap weighted by the inverse of the standard error of the treatment effect.

Before meta-analytical estimation, the heterogeneity of reports required conversion; using a heuristic approximation,[13] medians and quartiles were converted into means. Using knowledge of mean changes and observed *t*- or *p*-values allowed an estimation of non-reported SDs. Based on the assumption that the correlation between baseline and final values was R≅0.5 allowed final values to be assimilated to mean changes. Where possible, data which were not presented in tabular form in reviewed studies were approximated from graphs. While accounting for carryover effects,[14] results from crossover and parallel studies were mixed and corrected appropriately.

We fitted the data with both the random model (differences expected among studies) and the fixed model (no difference between studies). NMA has been shown to be appropriate for multiple comparisons.[11] For the expected multiarm corrections, the weight reduction approach, which is identical to the standard regression method (reduction of the design matrix dimension until it is invertible), was used to correct correlated pairwise comparisons. For the assessment-of-model fit, the generalized Cochran Qt[12] was divided into Qd (which measures the inconsistency between net estimates), Dt (based on a complete design-by-treatment interaction random-effects model), and the direct differences (Dd and Qh), which evaluate heterogeneity across different studies. Heterogeneity between studies was assessed through Chi-squared test and evaluated with the I^2 statistic. Results of the fixed model were used when $I^2 <50\%$ and the total heterogeneity was not significant.

P-scores, used to rank treatments, assessed the amount of confidence that any specific treatment was better than any other treatment, averaged across all comparator treatments,[11] equally with the surface under the cumulative ranking curve (defined as the treatment rank within the range of treatments).[15,16] R statistical packages (version 3.2.4), and the meta-library, Netmeta,[12] were used for all statistical analyses.

Risk of bias across studies

Publication bias was assessed using funnel plots; direct comparisons were conducted for each endpoint and treatment. As all included studies were placebo-controlled, placebo was able to be used as the unique control in our assessments.

Results

Study selection

An extensive search of the literature for published RCTs, which included modafinil and armodafinil for the treatment of narcolepsy, identified 484 articles from medical databases and 48 from other sources, with 312 articles remaining after removing duplicates. Thirty original clinical studies were found, out of which we excluded non-RCTs (4 studies), studies with no assessment of at least one efficacy or safety endpoint (5), retrospective studies (7), or those which did not involve patients with narcolepsy (4). A total of 10 RCTs were assessed and considered to be eligible for inclusion in the NMA (Figure 1 and Table 1);[8,9,17–24] all of the eligible studies had a placebo arm. Nine studies, which included pitolisant, were identified, resulting in the selection of two RCTs.[23,24] Eight studies compared MDF with placebo, one compared MDF and pitolisant with placebo, and one compared pitolisant with placebo (Table 2). ESS and MWT were used for all these studies. Cataplexy was evaluated in four studies.[17,20,23,24] Table 3 summarizes the RCTs that were excluded from this meta-analysis.[25–28]

Risk of bias within studies

The methodological quality of each clinical study was assessed (Supplementary Table 15; available at: https://www.drugsincontext.com/wp-content/uploads/2020/06/dic.2020-6-2-Suppl.pdf). All included studies had acceptable internal, external, and statistical validity. In the clinical study reported by Moldofsky and colleagues,[19] both selected study arms were preceded by a 16-week open-label MDF treatment period, potentially impacting the placebo group.[19] In the study by Black and Houghton, all patients received MDF at the established dose prior to randomization, and the abrupt withdrawal of patients from MDF treatment potentially resulted in an artificially worsened placebo group when patients changed treatment arms.[21]

Study characteristics

All evaluated clinical trials were short-term studies (2–12 week duration). In contrast, all long-term studies were open-label, non-comparative studies. For the included studies, design features and key characteristics are shown in Table 1. Treatments and endpoints assessed for each included study are presented in Table 2.

Main findings from individual studies

The approval of MDF was based on two 9-week double-blind, placebo-controlled studies,[8,9] in which no efficacy difference (for EDS) was observed between the two administered doses of MDF; moreover, no effect on cataplexy was reported in either trial, a finding which has also since been confirmed by meta-analysis.[29] In studies where direct comparisons were done between PIT and MDF, PIT was superior to placebo and similar

Table 1.   Included publications following review of published RCTs on drug treatment for narcolepsy.

Study	Tested drugs	Design	Treatment duration	Sample size	Endpoints of interest	Comments
Billiard et al. 1994[17]	Modafinil, 300 mg/d Placebo	RCT, 2-way 4-wk, cross-over	4 wks, 2 wk placebo washout (WO)	N=50	MWT, cataplexy, sleep attacks, inadvertent naps	No ESS. Safety not documented. Selected only for MWT and cataplexy
Broughton et al. 1997[18]	Modafinil, 200 mg/d Modafinil, 400 mg/d Placebo	DB, cross-over RCT 3 x 2 wks	3 x 2 wks No WO period	N=75	MWT (primary endpoint), ESS, sleep attacks, inadvertent naps. Safety = AE	Safety data poorly documented
US-MDF 1998[8]	Modafinil, 200 mg/d Modafinil, 400 mg/d Placebo	DB-RCT 3 parallel groups	9 wks	N=283 n=92 placebo n=96 MDF200 n=95 MDF400	20-min MWT and CGI (primary endpoint) ESS MSLT Sleep attacks on daily basis Safety AE	No data on cataplexy
US-MDF 2000[9]	Modafinil, 200 mg/d Modafinil, 400 mg/d Placebo	DB–RCT 3 parallel groups	9 wks	N=271 n=89/MDF200 n=89/MDF400 n=93/placebo	ESS 20 min MWT, sleep attacks, inadvertent naps CGI Safety AE	No data on cataplexy
Moldofsky et al. 2000[19]	Modafinil, 300–500 mg/d Placebo	DB, placebo-controlled, 2 wks After 16-wk MDF OL.	16 wks OL. 2 wks DB	N=63	40 min MWT ESS Daily number of cataplectic attacks Number of periods of severe sleepiness, voluntary sleep episodes (naps), and sleep attacks	Study assessing the treatment interruption and withdrawal symptoms after 16-wk OL. Safety not documented. Study selected only for efficacy
Harsh et al. 2006[20]	Armodafinil, 150 mg/d Armodafinil, 250 mg/d Placebo	DB, RCT 3 parallel groups	12 wks	N=196 n=64/ADF 150 n=67/ADF 250 n=63/placebo	20 min MWT (primary endpoint) ESS Cataplexy CGI Cognitive tests (CDR) Fatigue inventory Safety	Safety only most frequent AE (>5%)

(Continued)

Table 1.　(Continued)

Study	Tested drugs	Design	Treatment duration	Sample size	Endpoints of interest	Comments
Black and Houghton 2006[21]	Placebo Modafinil, 200–600 mg/d Sodium oxybate, 6–9 g/d Sodium oxybate 6–9 g/d + Modafinil 200-600 mg/d	DB, RCT 4 parallel groups	8 wks	N=222 n=55/placebo n=63/MDF n=50/X n=54/X + MDF	20 min MWT (primary endpoint) ESS CGI Sleep attacks Safety	No data on cataplexy
Saletu et al. 2005[22]	Modafinil fixed titration at 3 wks (200 mg/dW1, 300 mg/d,W2, 400 mg/d W3) Placebo	DB, RCT Placebo-controlled crossover	3 wks 1 wk	N=16 matched with 16 control HV	ESS MSLT, EEG AE	Safety data poorly documented
Dauvilliers et al. 2013 (HARMONY I)[23]	Pitolisant up to 40 mg/d Modafinil up to 400 mg/d Placebo	DB, RCT 3 parallel groups	8 wks	N=94 n=31/pitolisant n=33/MDF n=30/placebo	ESS (primary endpoint), % of responders, 20 min MWT Cataplexy and Sleep attacks CGI, Safety AE	
Szakacs et al. 2017 (HARMONY CTP)[24]	Pitolisant up to 40 mg/d Placebo	DB, RCT 2 parallel groups	7 wks	N=105 n=54/pitolisant n=51/placebo	Weekly rate of cataplexy (primary endpoint) ESS, % of responders 40 min MWT CGI, Patient Global Opinion, Safety AE	Patients included with at least 3 cataplexy per wk

ADF, armodafinil; AE, adverse event; CGI, Clinical Global Impression; CTP, cataplexy; DB, double blind; ESS, Epworth Sleepiness Scale; MDF, modafinil up to 400 mg/day; MSLT, Multiple Sleep Latency Test; MWT, Maintenance Wakefulness Test; OL, open label; PIT, pitolisant up to 40 mg/day; RCT, randomized controlled trial.

Table 2.   Comparison of treatments and studied endpoints within studies.

Study	Placebo	MDF	PIT	ESS	MWT	CTP	AE
Billiard et al. 1994[17]	*	*			+	+	
Broughton et al. 1997[18]	*	*		+	+	+	+
US-MDF 1998[8]	*	*		+	+		+
US-MDF 2000[9]	*	*		+	+	+	+
Moldofsky et al. 2000[19]	*	*		+	+		
Saletu et al. 2005[22]	*	*		+	+		+
Harsh et al. 2006[20]	*	*		+	+	+	+
Black & Houghton 2006[21]	*	*		+	+		+
Dauvilliers et al. 2013[23]	*	*	*	+	+	+	+
Szakacs et al. 2017[24]	*		*	+	+	+	+

*Treatments compared within studies.
+Endpoints evaluated within studies.
AE, adverse event; CTP, cataplexy; ESS, Epworth Sleepiness Scale; MDF, modafinil up to 400 mg/day; MWT, Maintenance Wakefulness Test; PIT, pitolisant up to 40 mg/day.

to MDF for both EDS and cataplexy,[23] and a second study,[24] in patients with cataplexy, confirmed the beneficial effect of PIT for the reduction of cataplexy and EDS.

Synthesis of results: efficacy assessment

A network evidence graph is shown in Figure 2. We first report results of the clinical endpoints (ESS, MWT, cataplexy, and safety), summarized in Table 4 and Figure 3. Detailed computation for each endpoint is referenced in the Supplementary File.

For ESS (9 studies, Supplementary File, section 3), we observed slight heterogeneity across studies ($p=0.05$), and homogeneity between designs ($p=0.40$). The two compared drugs achieved significant mean differences when compared to placebo: PIT (−3.39, 95% confidence interval [CI]: −4.76 to −2.04; $p<0.001$) and MDF (−2.70, [−3.35 to −2.04; $p<0.001$). A non-significantly superior beneficial decrease of −0.69 ([−2.18 to 0.79]; $p=0.36$) was observed for PIT. The CI upper limit for the difference (0.79) being inferior to the predetermined NIM (=1 provides evidence of the non-inferiority of PIT compared with MDF ($p=0.015$).

The MWT (10 studies, Supplementary File, section 4) evaluated the mean increase from baseline in time (minutes). We found a non-significant heterogeneity across studies ($p=0.76$, $I^2=0\%$), and no between-design inconsistency ($p=0.27$). Compared with placebo, significant relative benefits were observed for PIT (4.8 min [1.7–7.9]; $p\leq0.001$) and MDF (2.7 min, [1.9–3.4]; $p<0.02$). A non-significantly superior beneficial decrease of −2.12 min ([−0.95 to 5.19]; $p=0.18$) was observed for PIT. The lower limit of the CI of the difference (−0.95) being inferior to the predetermined NIM (=−1 provides evidence of the non-inferiority of PIT compared with MDF ($p=0.04$).

Cataplexy was reported in four studies (Supplementary File, Section 5); we used the standardized mean difference (SMD)

Figure 2.   Network tree.
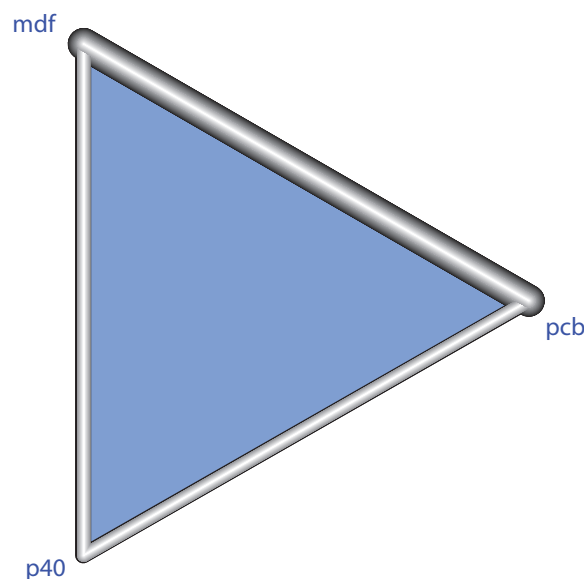
A network meta-analysis was needed to account both for the direct comparisons, but also indirect comparisons between modafinil and pitolisant. In this network evidence graph, each node in the network is associated with a treatment (pcb=placebo, p40=pitolisant up to 40 mg, mdf: modafinil up to 400 mg). An overlap (edge) between any two treatments represents a direct comparison, the thickness of the overlap proportional to the inverse of the standard error of the treatment effect.

and converted it by linear calibration into the decrease of weekly rate of cataplexies (DWRC). A significant reduction was only achieved for PIT (−0.54. [−0.85 to 0.23]; $p<0.001$; DWRC=−5.9), with no significant difference for MDF (0.049

Table 3. Publications excluded from the network meta-analysis.

Study	Tested drugs	Design	Treatment duration	Sample size	Endpoints	Comments
Laffont et al. 1988. (MOD 024)[25]	Modafinil, 200 mg/d Placebo	DB, RCT crossover 2 x 2wks	2 wks	N=10	No data on ESS, MWT, cataplexy No data on safety reported	Not published, only as an abstract. No data on ESS, MWT, or cataplexy. Safety not documented
Boivin et al. 1993[26]	Modafinil, 300 mg/d Placebo	DB, RCT 4-wk crossover	4 wks	N=10	PSG, EMG (Periodic Leg Movement index) EDS on 10 points VAS (no ESS) Cognitive test (FCRTT) Daily number of sleep attacks	No data on ESS, MWT or cataplexy Safety not documented
Besset et al. 1993[27]	Modafinil, 300 mg/d Placebo	DB, RCT 4-wk crossover	4 wks	N=16	Stanford scale instead of ESS. Attention Safety (poor data) PSG (REM)	No data on ESS, MWT or cataplexy Safety data poorly documented
Kollb-Sielecka et al. 2017 (HARMONY Ibis)[28]	Pitolisant up to 20 mg/d Modafinil up to 400 mg/d Placebo	DB-RCT 3 parallel groups	8 wks	N= 165 n=67/ pitolisant n=65/MDF n=33/ placebo	ESS (primary endpoint), % of responders 40 min MWT Cataplexy and sleep attacks CGI Safety AE	

DB, double blind; CGI, Clinical Global Impression; EDS, excessive daytime sleepiness; EMG, electromyography; ESS, Epworth Sleepiness Scale; FCRTT, Four Choice Reaction Time Test; MWT, Maintenance Wakefulness Test; PSG, polysomnography; RCT, randomized controlled trial; REM, rapid eye movement; VAS, visual analog scale.

[−0.32 to 0.21], $p=0.38$). There was no marked or significant heterogeneity across studies ($p=0.89$) or between-design inconsistency ($p=0.34$). The mean beneficial effect of PIT was significant −0.49 ([−0.86 to −0.12], DWRC=−4.3; $p=0.012$), providing evidence of the superiority of PIT compared with MDF (non-inferiority being demonstrated *de facto*).

Safety was compared in 9 studies (Supplementary File, Section 10) based on OSS. The safest intervention was confirmed to be placebo, followed by PIT (IRR=1.42 [1.01 to 2.01], $p=0.045$) and MDF (IRR=1.56 [1.30 to 1.89], $p<0.001$). There was significant heterogeneity across studies ($p<0.001$), and between-design consistency ($p=0.32$), requiring the random model to be used for difference estimation. The mean ratio PIT/MDF was 0.86 [0.44 to 1.24] in favor of PIT, and non-inferiority was concluded. Supportive analyses conducted on safety subgroups (including headache, CNS symptoms, and gastrointestinal symptoms [Supplementary File, sections 7–9]) provided similar results.

In addition to these endpoints, we also compared the treatments using aggregated scores: NS1 and NS2 scores were defined as efficacy scores adapted for narcolepsy type-1 (NT1) and type-2 (NT2) patients, respectively. NS1 was calculated from 10 studies

(Supplementary File, Section 6). For NT1 patients, a significant improvement was observed for PIT (0.87 [0.58 to 1.16]; $p<0.001$) and MDF (0.41 [0.28 to 0.54]; $p<0.001$). There was homogeneity across studies ($p=0.08$) and for between-design inconsistency ($p=0.72$). We observed a significant benefit of 0.46 ([0.06, to −0.96], $p=0.03$) inducing non-inferiority. For NT2 patients (Supplementary File, Section 12, for which cataplexy was not accounted for in the efficacy score), a significant improvement in NS2 was observed for PIT (0.56 [0.27 to 0.85]; $p<0.001$) and MDF (0.38 [0.25 to 0.61]; $p<0.001$). The mean difference (PIT–MDF) in favor of PIT was 0.18 ([−0.16 to 0.50]; $p=0.52$) which was considered as non-inferior on the basis of a NIM of 0.2.

The BR ratio, evaluated by BR1 and BR2 scores, was compared in 9 studies (Table 4, Supplementary Section 11). A linear relationship between NS1 (efficacy) and OSS (safety) was found with Intercept = −0.98 [−1.31 to 0.01] ($p=0.12$) and slope = 0.27 [0.02 to 0.47]; $p=0.008$), its residuals defining the B/R ratio for each study. For NT1 patients, a significant BR1 value was observed for PIT (0.84 [0.55 to 1.12]; $p<0.001$) and MDF (0.35 [0.23 to 0.48]; $p<0.001$). We found homogeneity across studies ($p=0.07$) and homogeneous between-design inconsistency ($p=0.72$). The mean observed

Table 4.   Characteristics and tests for each analysis.

	ESS	MWT	Cataplexy[a]	NS1[b]	OSS[c]	BR1[d]	NS2[e]	BR2[f]
Studies (n)	9	10	4	10	9	9	10	9
Pairwise computations (n)	11	12	6	12	11	11	12	11
I^2 (%)	45	0.0	0	0.01	0.01	0.01	0.01	0.01
Difference with Placebo[g]	MD	MD	MD	MD	RR	MD	MD	MD
MDF	−2.7***	2.7*	−0^{ns}	0.41**	1.59**	0.35**	0.38**	0.30***
PIT	−3.4***	4.8***	−5.9***	0.87***	1.38^{ns}	0.84***	0.56***	0.54***
Tests								
- Within Q_h	0.05	0.76	0.89	0.08	0.01	0.03	0.04	0.06
- Between Q_i	0.41	0.27	0.34	0.72	0.32	0.62	0.51	0.5
***p*-scores:**								
- MDF	0.59	0.54	0.32	0.59	0.17	0.58	0.39	0.4
- PIT	0.91	0.95	1	0.91	0.41	0.91	0.11	0.11
- Placebo	0.00	0.02	0.18	0	0.92	0.01	1	0.99
Difference PIT–MDF	−0.69	2.12	−0.49	0.46	0.86	0.49	0.15	0.24
- 95% CI	−2.18, 0.79	−0.95, 5.19	−0.86, −0.12	−0.11, 0.49	0.44, 1.24	0.08, 1.03	−0.15, 0.45	−0.19, 0.70
- $p^{(h)}$	0.015 (0.36)	0.04 (0.18)	<0.001 (0.012)	0.004 (0.22)	0.66 (0.04)	0.021 (<0.001)	0.32	0.25

Number of studies, number of pairwise computations, heterogeneity index (I^2), and tests of within-design (Q_h, measuring heterogeneity between studies), and between-designs (Q_i, measuring between design inconsistency) for the following endpoints: ESS, MWT, cataplexy, narcolepsy Z-Score, safety, and benefit/risk ratio.
(a) weekly reduction of cataplexy rate (CTP); (b) NS1 = Narcolepsy Score aggregating efficacy for EDS and cataplexy, thus appropriate for Type 1 Narcolepsy patients; (c) overall safety score (OSS); (d) BR1 = benefit/risk ratio applicable for Narcolepsy type 1 patients combining EDS and Cataplexy, calculated as the residual of the linear fit of the NS1 by the OSS.; (e) NS2 = Narcolepsy score limited to efficacy on EDS and appropriate for Narcolepsy type 2 patients; (f) BR2 = Benefit/Risk ratio based on efficacy limited to EDS, applicable to Narcolepsy type 2 patients and calculated as the residual of the linear fit of EDS Z score by OSS. (g) Differences between Modafinil or Pitolisant with placebo expressed as Mean differences (MD) or Risk Ratios (RR) significance abbreviated as ns (p>.05), * (p<.05), ** (0.001<p<.01), *** (p<.001). (h) The p-value of the difference is associated with the non-inferiority test compared with the Null hypothesis that the difference is at least as large as the prespecified NIM. The p-value enclosed within parentheses correspond to a superiority test of pitolisant on Modafinil.
ESS, Epworth Sleepiness Scale; MDF, modafinil; MWT, Maintenance Wakefulness Test; PIT, pitolisant.

Figure 3. Forest plot for all the compared endpoints.
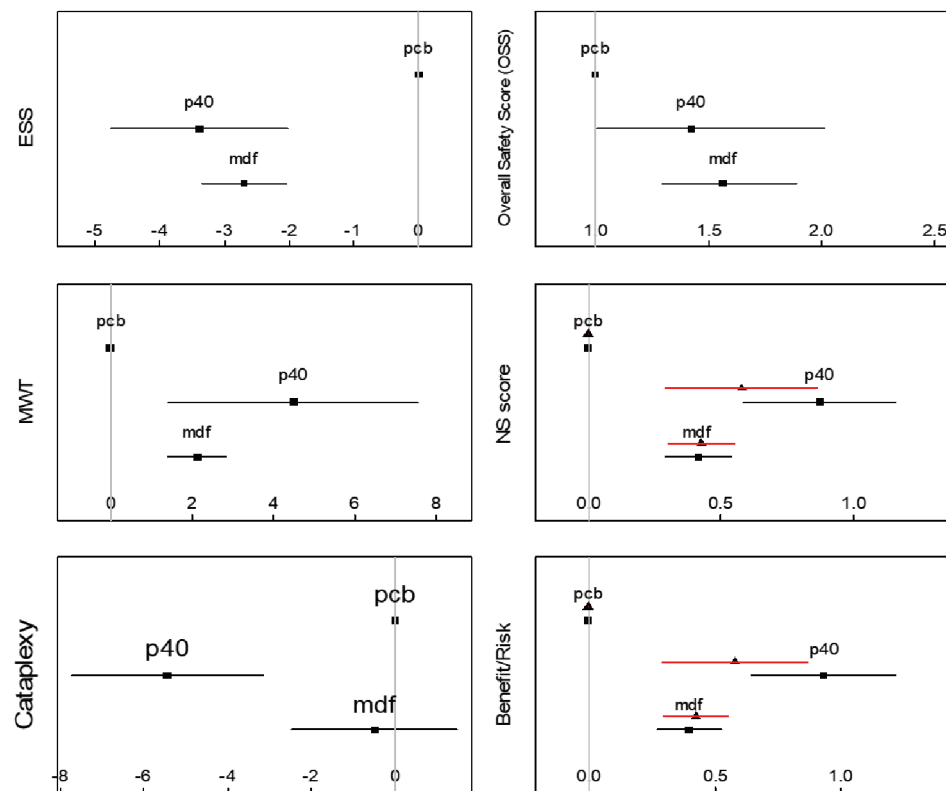
Measures: ESS (mean change from baseline in ESS scores), MWT (mean change from baseline in minutes), Cataplexy (mean change from baseline in weekly rate of cataplexies), Overall Safety Score (relative risk, defined as the ratio between the number of treatment emergent adverse events, nTEAEs, on the considered drug on nTEAE on the placebo arm). For the NS scores and risk benefits sub-graphics, black squares represent NS1 score and corresponding benefit/risk adapted for narcolepsy type 1 patients, whereas black triangles represent similar values for NS2 and corresponding benefit/risk for narcolepsy type 2 patients.

ESS, Epworth Sleepiness Scale; MWT, Maintenance of Wakefulness Test; mdf, modafinil up to 400 mg; pcb, placebo; p40, pitolisant up to 40 mg

difference (PIT–MDF) in favor of PIT was a significant benefit of 0.49 ([0.08 to 1.03]; p=0.02). For NT2 patients (Supplementary Section 13), a significant BR2 improvement was observed for PIT (0.54 [0.09 to 0.98]; p<0.001) and MDF (0.30 [0.08 to 0.53]; p<0.001). We found heterogeneity across studies (p=0.001) and homogeneity between-design consistency (p=0.62). The non-significantly superior benefit of PIT was 0.24 ([−0.19 to 0.70]), considered as non-inferior based on the NIM of 0.2.

Risk of bias across studies

Funnel plots were used to assess publication bias. Evaluation of all funnel plots for individual endpoints (ESS, MWT, cataplexy) and treatments, particularly when compared with placebo, provided no evidence of an asymmetrical distribution of the points representing the studies (Supplementary File, Section 14).

Discussion

Justification

This meta-analysis, conducted in line with regulatory requirements, compared pitolisant, a novel compound recently approved for the treatment of narcolepsy, with MDF, an established treatment with proven efficacy, which we considered as a control treatment. A classical meta-analysis comparing two drugs was not appropriate because direct comparisons between the two treatments constitute only a small proportion of the studies, which assessed these treatments compared with placebo. Thus, an NMA was needed to account for direct comparisons in addition to indirect comparisons with placebo. A recent NMA [6] provided a larger multi-treatment comparison in comparing all existing narcolepsy treatments. From that analysis, MDF, 200–400 mg/

day, sodium oxybate, 9 g/day, and pitolisant, up to 40 mg/day (PIT), were found to have similar clinical efficacy and were significantly more effective than placebo for EDS. Only sodium oxybate, 9 g/day, and PIT were shown to have a comparable beneficial effect on cataplexy. Overall, PIT was shown to have a slightly better safety profile and the highest BR ratio. However, the difference between the three drugs was not statistically significant, with only *p*-scores allowing any differentiation. In that sense, comparing active treatments requires a non-inferiority approach instead of strict superiority.

Methodological issues and limitations

The results presented are dependent on those reported in the literature. In some cases, only summary statistical data were reported. In addition, the initial condition of patients was often accounted for by mean change. However, adjustment for baseline cannot be conducted based on individual values. Therefore, approximations were needed in our analysis, due to the necessary estimate of standard deviations that were not reported in many of the publications. Some conversions were also needed when only medians were available. For cataplexy, only four studies provided values, which were sometimes reported in graphics, and heterogeneous measures were reported in terms of days, numbers, or rates of cataplexy.

EDS and cataplexy were considered to be the main symptoms of narcolepsy, whereas other symptoms (e.g. hallucinations or sleep paralysis, disturbed nocturnal sleep) that are irregularly documented in trials were not analyzed. Safety was often poorly documented in the publications, particularly for older trials.

Our network approach was limited to the comparison of MDF, PIT, and placebo, and only studies with at least a comparison between two of these treatments were used. The largest possible selection should have used studies with at least one of these treatments. Only three studies that compared sodium oxybate to MDF and placebo could have been added; however, this option would have uselessly increased the number of indirect comparisons and necessitated more hypotheses on the network consistency.

Study strengths

Clinical discussions and comparisons are mainly based on usual and recognized clinical endpoints. In narcolepsy, ESS, MWT, and cataplexy rate are almost always used. We first compared the tested treatments separately for each of these endpoints and our discussion is mainly based on these comparisons. However, the efficacy of a treatment is generally not restricted to one single symptom relief but should consider all existing symptoms in the pathology under assessment. In that sense, EDS and cataplexy constitute the essential symptoms of narcolepsy; thus, an aggregated score constitutes a useful approach for assessing global efficacy and, at the same time, a unique endpoint to reduce the multiplicity of test and the resulting inflation of statistical type-1 errors. Combining

variables like ESS, MWT, and cataplexy to produce a Z score might appear to be somewhat artificial at first glance, as they measure different endpoints with non-comparable units. However, a Z score has the well-known property of combining these values in standardized units to produce a unique evaluation of efficacy and even a measure of the clinical relevance of the effect. Finally, efficacy and safety are often assessed separately, although they should normally be integrated as a BR ratio, constituting the best estimate of the overall clinical utility of an intervention. The findings of this study confirmed the direct association between safety, measured by OSS, and efficacy, measured by the NS, justifying the residual of the linear regression of NS by OSS as a measure of efficacy and safety. Finally, we accounted for the existence of two different types of narcoleptic patients (NT1 and NT2), and we developed efficacy (NS1, NS2) and BR ratios (BR1, BR2) specific for these two patient types.

Clinical findings

The main clinical measurements (ESS, MWT, and number of cataplexies) were our principal concern and were studied separately. MDF was the most extensively investigated intervention, and was characterized by a significant improvement in EDS, based on ESS and MWT, with no significant effect on cataplexy, and an acceptable safety profile. This aspect of our results confirms earlier meta-analytical results.[29] The comparison between MDF and PIT was first conducted on mean change in ESS, for which PIT performed better but without reaching statistical significance; non-inferiority of PIT was demonstrated compared with MDF, based on the smallest possible NIM of 1. The comparison between PIT and MDF has been investigated previously,[23] without finding non-inferiority. However, in the current NMA, we demonstrate non-inferiority using a strengthened NIM (–1) due to the much larger sample size involved in our meta-analytical calculation.

For MWT, measuring the mean changes in time from baseline, similar better performance for PIT was observed, albeit without reaching statistical significance. The non-inferiority of PIT compared with MDF was demonstrated at a predetermined margin of 1 minute. Finally, for cataplexy, a strict statistically significant benefit was found for PIT, corresponding to a decrease of 4.3 cataplexy crises per week compared with MDF. The separate examination of these endpoints was completed by the comparison of NS1 and NS2, specific to NT1 and NT2 patients, respectively. We demonstrated strict significant superiority of PIT *versus* MDF for NS1 (adapted for NT1 patients) and non-inferiority of PIT for NS2 (adapted for NT2 patients).

We also compared safety by measuring OSS (defined as the TEAE incidence rate during the exposure period). Compared with placebo, the IRRs for PIT and MDF were 1.42 and 1.56, respectively, with 14% fewer TEAEs for PIT and non-inferiority when assuming the predetermined maximum margin of 25% (IRR=1.25).

The BR ratio constitutes a useful evaluation to compare the utility of alternative treatments. In that sense, compared with MDF, we demonstrated strict statistical superiority of PIT on the BR1 score, although only non-inferiority of PIT was concluded based on the BR2 score (applicable to NT2 patients). In addition to the findings of non-inferiority and superiority described earlier, *p*-scores constitute relevant indices to clinical practice[12] and provide evidence of better results with PIT for all the studied endpoints and, in particular, for the BR ratio.

Finally, the BR ratio, accounting for both efficacy and safety, constitutes the best estimate of clinical utility. Table 4 provides *p*-scores of comparison between placebo, MDF, and PIT. For the three efficacy endpoints (ESS, MWT, cataplexy) and their aggregate scores (NS, EDS, and BR ratio both for narcolepsy type-1 and type-2 patients), P-score ranking was constantly placebo < MDF < PIT whereas, for safety, the order placebo > PIT > MDF was observed for the OSS and all of its components.

Conclusion

Based on the largest number of randomized trials, modafinil is confirmed as the most extensively investigated intervention, characterized by a significant improvement in EDS (based on ESS and MWT), and an acceptable safety profile, but with no significant effect on cataplexy. Pitolisant is non-inferior to modafinil in relieving EDS, but superior to modafinil in reducing the occurrence of cataplexies. Higher P-scores for safety and BR ratios provide evidence that pitolisant outranks modafinil when prescribed to narcolepsy type-1 patients. Despite the slight superiority of pitolisant over modafinil in EDS relief, both drugs can be considered as equally performing for the treatment of narcolepsy type-2 patients.

References

1. Daniels E, King MA, Smith IE, et al. Health-related quality of life in narcolepsy. *J Sleep Res*. 2001;10(1):75–81. https://doi.org/10.1046/j.1365-2869.2001.00234.x
2. Ingravallo F, Vignatelli L, Brini M, et al. Medico-legal assessment of disability in narcolepsy: an interobserver reliability study. *J Sleep Res*. 2008;17(1):111–119. https://doi.org/10.1111/j.1365-2869.2008.00630.x
3. Jennum P, Ibsen R, Knudsen S, et al. Comorbidity and mortality of narcolepsy: a controlled retro- and prospective national study. *Sleep*. 2013; 36(6):835–840. https://doi.org/10.5665/sleep.2706

4.   Ohayon MM, Black J, Lai C, et al. Increased mortality in narcolepsy. *Sleep*. 2014;37(3):439–444. https://doi.org/10.5665/sleep.3470

5.   American Academy of Sleep Medicine. *International Classification of Sleep Disorders*. 3rd ed. Darien, IL: American Academy of Sleep Medicine; 2014.

6.   Lehert P, Falissard B. Multiple treatment comparison in narcolepsy: a network meta-analysis. *Sleep*. 2018;41(12):zsy185. https://doi.org/10.1093/sleep/zsy185

7.   Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA Statement. *PLoS Med*. 2009;6(7):e1000097. https://doi.org/10.1371/journal.pmed.1000097

8.   US Modafinil in Narcolepsy Multicenter Study Group. Randomized trial of modafinil for the treatment of pathological somnolence in narcolepsy. *Ann Neurol*. 1998;43(1):88–97. https://doi.org/10.1002/ana.410430115

9.   US Modafinil in Narcolepsy Multicenter Study Group. Randomized trial of modafinil as a treatment for the excessive daytime somnolence of narcolepsy: US Modafinil in Narcolepsy Multicenter Study Group. *Neurology*. 2000;54(5):1166–1175. https://doi.org/10.1212/wnl.54.5.1166

10. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed. New York, NY: Academic Press; 1988.

11. Rücker G. Network meta-analysis, electrical networks and graph theory. *Res Synth Methods*. 2012;3(4):312–324. https://doi.org/10.1002/jrsm.1058

12. Rücker G, Schwarzer G. Ranking treatments in frequentist network meta-analysis works without resampling methods. *BMC Med Res Methodol*. 2015;15:58. https://doi.org/10.1186/s12874-015-0060-8

13. Hozo S, Djulbegovic B, Hozo I. Estimating the mean and variance from the median, range, and the size of a sample. *BMC Med Res Methodol*. 2005;5:13. https://doi.org/10.1186/1471-2288-5-13

14. Curtin F, Elbourne D, Altman DG. Meta-analysis combining parallel and cross-over clinical trials. III: the issue of carry-over. *Stat Med*. 2002;21(15):2161–2173. https://doi.org/10.1002/sim.1207

15. Higgins JPT, Jackson D, Barrett JK, et al. Consistency and inconsistency in network meta-analysis: concepts and models for multi-arm studies. *Res Synth Methods*. 2012;3(2):98–110. https://doi.org/10.1002/jrsm.1044

16. Salanti G, Ades AE, Ioannidis JP. Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: an overview and tutorial. *J Clin Epidemiol*. 2011;64(2):163–171. https://doi.org/10.1016/j.jclinepi.2010.03.016

17. Billiard M, Besset A, Montplaisir J, et al. Modafinil: a double-blind, multicentre study. *Sleep*. 1994;17(8 Suppl):S107–S112. https://doi.org/10.1093/sleep/17.suppl_8.s107

18. Broughton RJ, Fleming JA, George CF, et al. Randomized, double-blind, placebo-controlled crossover trial of modafinil in the treatment of excessive daytime sleepiness in narcolepsy. *Neurology*. 1997;49(2):444–451. https://doi.org/10.1212/wnl.49.2.444

19. Moldofsky H, Broughton RJ, Hill JD. A randomized trial of the long-term continued efficacy and safety of modafinil in narcolepsy. *Sleep Med*. 2000;1(2):109–116. https://doi.org/10.1016/s1389-9457(99)00014-3

20. Harsh JR, Hayduk R, Rosenberg R, et al. The efficacy and safety of armodafinil as treatment for adults with excessive sleepiness associated with narcolepsy. *Curr Med Res Opin*. 2006;22(4):761–774. https://doi.org/10.1185/030079906X100050

21. Black J, Houghton WC. Sodium oxybate improves excessive daytime sleepiness in narcolepsy. *Sleep*. 2006;29:939–946. https://doi.org/10.1093/sleep/29.7.939

22. Saletu MT, Anderer P, Saletu-Zyhlarz GM, et al. EEG-mapping differences between narcolepsy patients and controls and subsequent double-blind, placebo-controlled studies with modafinil. *Eur Arch Psychiatry Clin Neurosci*. 2005;255(1):20–32. https://doi.org/10.1007/s00406-004-0530-1

23. Dauvilliers Y, Bassetti C, Lammers GJ, et al. Pitolisant versus placebo or modafinil in patients with narcolepsy: a double-blind, randomised trial. *Lancet Neurol*. 2013;12(11):1068–1075. https://doi.org/10.1016/S1474-4422(13)70225-4

24. Szakacs Z, Dauvilliers Y, Mikhaylov V, et al. Safety and efficacy of pitolisant on cataplexy in patients with narcolepsy: a randomised, double-blind, placebo-controlled trial. *Lancet Neurol*. 2017;16(3):200–207. https://doi.org/10.1016/S1474-4422(16)30333-7

25. Laffont F, Cathala HP, Waisboard P, et al. Effects of modafinil in narcoleptic patients. Double-blind cross-over study. In 9th European Congress of Sleep Research, Jerusalem, Israel. September 4–9, 1988:215.

26. Boivin DB, Montplaisir J, Petit D, et al. Effects of modafinil on the symptomatology of human narcolepsy. *Clin Neuropharmacol*. 1993;16(1):46–53. https://doi.org/10.1097/00002826-199302000-00005

27. Besset A, Tafti M, Villemin E, et al. The effects of modafinil (300 mg) on sleep, sleepiness and arousal in narcoleptic patients. *Neurophysiol Clin*. 1993;23(1):47–60. https://doi.org/10.1016/s0987-7053(05)80282-5

28. Kollb-Sielecka M, Demolis P, Emmerich J, et al. The European Medicines Agency review of pitolisant for treatment of narcolepsy: summary of the scientific assessment by the committee for medicinal products for human use. *Sleep Med*. 2017;33:125–129. https://doi.org/10.1016/j.sleep.2017.01.002

29. Golicki D, Bala MM, Niewada M, Wierzbicka A. Modafinil for narcolepsy: systematic review and meta-analysis. *Med Sci Monit*. 2010;16(8):177–186.