Antidepressant effect in older depressed patients: the lessons of two agomelatine's trials.

Guy M. Goodwin, FMedSci	University Department of Psychiatry, Warneford Hospital, Oxford OX3 7JX
Pierre Thomas, MD	Service de Psychiatrie Generale C.H.R.U. de Lille 6, rue du Pr Laguesse 59037 Lille Cedex - France
Reinhard Heun, MD	Radbourne Unit Royal Derby Hospital Uttoxeter Road Derby DE22 3WQ United Kingdom
Patrice Boyer, MD	University of Ottawa (Ontario) and of Paris Diderot (France)
Françoise Picarel-Blanchot, PhD	Institut de Recherches Internationales Servier (IRIS) 50 rue Carnot 92284 Suresnes Cedex, France
Christian de Bodinat, MD	Institut de Recherches Internationales Servier (IRIS) 50 rue Carnot 92284 Suresnes Cedex, France

Short title: Agomelatine and the older MDD patients in two studies

Corresponding author: Professor G.M. Goodwin University of Oxf ord Department of Psychiatry Warneford Hospital Headington Oxford OX3 7JX. email: guy.goodwin@psych.ox.ac.uk – fax: 44 1865 204198

Funding support: This study was sponsored by Servier (Suresnes, France)

Conflict of interests: GMG holds a grant from Wellcome Trust, shares in P1vital and has served in the last 2 years as consultant, advisor or CME speaker for AstraZeneca, Merck, Cephalon/Teva,, Eli Lilly, Lundbeck, Medscape, Otsuka, P1Vital, Pfizer, Servier, Sunovion, Takeda. He is an NIHR Senior Investigator ; the views expressed are his and not necessarily those of the NHS, the NIHR or the Department of Health.; **RH** has received research grants, speaker's fees and/or consultancy honoraria from AstraZeneca, Bayer, BMS, Eisai, Novartis, Pfizer, Servier; **PB** has received honoraria from Servier and Pierre Fabre for advice and for speaking or organising training ;**FPB and CDB** are employees at Servier.

The authors had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Presentation at a meeting: No

Number of Words text:	4354	Number of Words abstract:	123
Number of Tables:	4	Number of Figures:	1

Trial registration name:

Study 1: Not applicable - **Study 2**: Efficacy and safety of agomelatine oral administration (25 to 50 mg/day) in elderly patients suffering from major depressive disorder: an 8-week, randomised, double-blind, flexible-dose, parallel groups, placebo-controlled, international, multicentre study followed by an extension double-blind treatment period of 16 weeks

Trial registration number:

Study 1: Not applicable - **Study 2**: ISRCTN57507360

ABSTRACT

The present paper reports in parallel the findings of the two phase III trials that evaluated the efficacy of agomelatine in older depressed patients. It describes how the particular methodological innovations (particularly in relation to patient selection, design and accuracy of diagnosis of depression) introduced in study 2 have improved the quality recruitment of patients and the assay sensitivity.

Study 1 lacked assay sensitivity, and among the many differences with study 2, the inclusion of unexpected mildly ill patients could have inflated placebo response. The increased demands on investigators in study 2 appear to have reduced the placebo effect and showed a robust benefit of agomelatine.

The two agomelatine studies offer the opportunity to discuss hypotheses that have been raised to explain the low level of response of older patients to available antidepressants.

Keywords: Agomelatine – antidepressant –depression - old- innovations

INTRODUCTION

The pharmacotherapy of depression in older patients is particularly challenging due to a significant variability in treatment response and because of the inherent heterogeneity of the older population. To date, only sparse clinically relevant results from placebo-controlled trials have been obtained in older patients with acute episodes of major depressive disorder (MDD) (Allard et al., 2004;Bose et al., 2008;Katona et al., 2012;Rapaport et al., 2003;Rapaport et al., 2009;Raskin et al., 2007;Raskin et al., 2008;Roose et al., 2004;Rossini et al., 2005;Schatzberg and Roose, 2006;Schneider et al., 2003;Sheikh et al., 2004) and, when condensing results with meta-analytic approaches, only a modest efficacy of antidepressant treatment is observed (Nelson et al., 2008;Tedeschini et al., 2011).

A clear-cut demonstration of the antidepressant efficacy of the compound agomelatine in the treatment of older patients has also been difficult to obtain. An initial trial (study 1) lacked assay sensitivity due to an unexpectedly high response rate in the placebo arm. Promisingly, a *post-hoc* analysis supported the potential of the drug in a sub-group of patients with more severe depressive symptoms at inclusion. We recently reported a number of methodological innovations which ensure that a patient population displays adequate severity not just on ratings of symptom severity but also on measures of functional impairment (Goodwin et al., 2013). Those modifications were applied in a second agomelatine trial conducted in older patients (study 2), and conclusive results were obtained (Heun et al., 2013).

The present paper reports in parallel the findings of studies 1 and 2, and describes how the particular methodological innovations introduced in study 2 have improved the quality recruitment of patients and the assay sensitivity. We anticipate these innovations can minimize the risks of non-conclusive studies quite generally in the older depressed population.

METHODS

The studies were both conducted in phase III of clinical development and were international, double-blind, randomised, and placebo-controlled. Study 1 was conducted in 59 centers in 7 countries (Sweden, Finland, France, Portugal, Australia, Canada and South Africa) from November 1999 to November 2001, and study 2 in 27 centers in 5 countries (Argentina, Finland, Mexico, Portugal and Romania) from November 2009 to October 2011.

The two studies were run in accordance with the principles of Good Clinical Practice E6 of the International Conference of Harmonisation (CPMP/ICH/135/95) and the Declaration of Helsinki, Finland (1964, 1996). The two studies were approved by the relevant local ethics committees and included only patients having given their written informed consent. The study 1 was not registered on a clinical trials registration site as it was not mandatory at that time. The main study results are in the public domain (EPAR 2008).

Patients

Eligible patients in study 1 were male or female physically healthy community patients or patients hospitalized for the current episode, aged over 60 years. In study 2, outpatients aged over 65 years were recruited. Patients had to be diagnosed with a single or recurrent episode (study 1) or a recurrent depression (study 2) and a current major depressive episode assessed as moderate or severe, according to *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision* (DSM-IV-TR) criteria (DSM-IV criteria were used for study 1)

In both studies, patients were selected on the basis of an episode of MDD, with or without melancholic features, without psychotic and catatonic features, and without seasonal pattern. In study 1, MDD patients with seasonal pattern can be recruited. The duration of the ongoing

episode was to be ≤ 6 months for study 1 and must have lasted at least 4 weeks (and no more than 12 months) for study 2. The diagnosis was to be documented using the brief structured Mini International Neuropsychiatric Interview (MINI) (Sheehan et al., 1998).

Exclusion criteria, including non authorized disorders and concomitant therapies were almost identical for both studies, and have been previously described for study 2 (Heun et al., 2013). In study 1, patients with a Mental Status Examination score (MMSE) scores that ranged between 25 and 30 (Folstein et al., 1975) could be included. In study 2, a MMSE score ≥ 27 was requested. Transcranial magnetic stimulation was authorized in study 1 but not in study 2. Authorized concomitant medications and washout time for medications were the same and have been described (Heun et al., 2013).

Key Study differences.

In study 1, patients were selected with a MADRS total score ≥ 24, whereas, in study 2, patients were required to have a HAM-D 17-item total score ≥ 22. However, additional criteria were requested in study 2 : a Clinical Global Impression (CGI) (Guy, 1976) Severity of illness (CGI)-S) ≥ 4, a Hospital Anxiety Depression Scale (HAD) depression sub-score ≥ 11, and must have completed the Geriatric Depression Scale (GDS) (Yesavage, 1988). The Sheehan Disability Scale (SDS) (Sheehan et al., 1996) had to be completed by patients at selection. Finally, study 2 required that subjects with a decrease of greater than 20% on the HAM-D_{17} between selection and inclusion were excluded. A summary of the main similarities and differences between patients selection in the two studies is given in Table 1.

Designs of the two studies

The two specifically designed studies evaluated the efficacy of agomelatine either at a fixed 25mg dose (study 1) or at one step uptitration 25-50mg dose (study 2).

After an initial run-in selection period without treatment, patients were randomized to a 6- (study 1) or 8-week (study 2) double-blind, placebo-controlled treatment period. After this period, patients could continue for a further optional 18- (study 1) or 16- week (study 2) double-blind extension period, and a one-week follow-up period without treatment. The primary analysis, reported in the present paper, was performed over the 6 or 8 weeks period. Data of study 2 were previously reported (Heun et al., 2013).

In each study, products received were of identical appearance whatever the treatment arm, and whatever the dose for study 2, in order to protect the blinding vis-à-vis the patient and the investigator.

Key Study differences. In study 2, the agomelatine dose at week 2 was either maintained at 25 mg/day or increased to 50 mg/day in patients with insufficient improvement based on blinded criteria. During study 1, all patients took orally 1 tablet once a day in the evening, irrespective the treatment allocated, whereas during study 2, all patients took orally 2 tablets once a day at bedtime– either 2 x 25mg agomelatine, or 1 x 25mg agomelatine and 1 x placebo (or 2x placebo).

In study 1, eligible patients were assigned to agomelatine or placebo treatment according to a balanced randomisation with stratification according to sex, age ([60-75[/ $\geq$ 75), geographic zone (Northern Europe, Southern Europe, Australia, America, and South Africa) and type of centre (General psychiatry or gerontopsychiatry) with an Interactive Response System. This randomisation was adaptative using a Pocock and Simon minimization randomization method to balance treatments according to the four stratification factors, based on a non-deterministic scheme. In study 2, randomization was unbalanced with a 2 to 1 ratio, and stratified on the centre

and on the age of patients ([65-75[/ ≥ 75) using an Interactive Response System. A summary of the main similarities and differences between the two studies' designs is given in Table 1.

Assessments

Primary efficacy assessment

The primary objective of study 1 was to assess the short-term efficacy of agomelatine using the Montgomery and Åsberg Depression Rating Scale (MADRS) (Montgomery and Asberg, 1979). In study 2, the primary outcome measure was the HAM-D_{17} items (HAMILTON, 1960) total score.

Secondary efficacy assessment

Secondary efficacy assessments included the HAM-D_{17} items (study 1), the Clinical Global Impression–Severity of illness (CGI-S) and Improvement of illness (CGI-I) scale (Guy, 1976) for both studies, the Hamilton Anxiety Rating Scale (HAM-A) (HAMILTON, 1959) for study 1 ; the Sheehan Disability Scale (SDS) (Sheehan et al., 1996) for study 2.

Safety and tolerability

The tolerability and safety evaluations in both studies were based on emergent adverse events (AEs), vital signs (supine SBP and DBP, supine heart rate, weight), biochemistry and haematology parameters and ECG abnormalities. When premature discontinuation of treatment was due to an AE, the information related to the outcome of the event was collected.

Statistical Analysis

In both studies, efficacy analyses were performed in all randomized patients having taken at least one dose of study treatment and having at least one post-baseline assessment for the primary efficacy outcome (Full Analysis Set: FAS).

In study 1, agomelatine-placebo differences were examined on the last post-baseline value of MADRS total score using a two sided Student's t-test for independent samples and using a four-way analysis of variance on factor treatment with adjustment for class of age, sex, and geographic area and without interaction. In study 2, agomelatine-placebo differences were examined on the last post-baseline value of HAM-D_{17} total score by using a 3-way analysis-of-covariance model on factor treatment, with center (random effect), classes of age (fixed effects), and baseline HAM-D_{17} total score as covariates and without interaction.

A sensitivity analysis to the method of handling missing values was performed in study 2. Treatment groups were compared on the value at week 8, using a mixed-effects for repeated measures model (MMRM), including terms for the fixed effects of treatment, class of age, baseline HAM-D_{17} total score, visit, and interaction term of treatment and visit, and for the random effect of center. The analysis fitted an unstructured covariance matrix.

Additional analyses were conducted by using a $\chi 2$ test to assess agomelatine-placebo differences in response to treatment (at least 50% decrease from baseline on MADRS total score for study 1 and on HAM-D_{17} total score for study 2) taking into account the last post-baseline value.

The agomelatine-placebo differences were also studied on the last post-baseline value of CGI-S score and last value of CGI-I score, using a two-sided Student t-test for independent samples.

For study 2, treatment groups were also compared on SDS work, social life, and family life scores, taking into account the last post-baseline value using a two-sided Student t-test (post hoc analysis).

All the previous efficacy analyses (except for the analysis of variance on MADRS total score in study 1 and the analysis of SDS scores in study 2) were also performed in the subset of more severely depressed patients; i.e. in study 1: patients with baseline MADRS total score ≥ 30 plus CGI-S ≥ 5; in study 2: patients with both HAM-D_{17} score ≥ 25 and CGI-S ≥ 5 at baseline.

For both studies, descriptive statistics were provided for safety measurement during the double-blind treatment period, in all included patients having taken at least one dose of study treatment.

Statistical analysis was performed on SAS® software, version 8.2 (study 1) and 9.1.3 (study 2) (Cary, North Carolina). The type I error was set at 5%.

RESULTS

Patient Characteristics

Dispositions of included and randomised patients in the two studies are depicted in **figure 1**. Of the 282 patients selected for study 1, 220 patients were included and randomly allocated to receive agomelatine (109 patients) or placebo (109 patients) for 6 weeks. Of the 271 patients selected for study 2, 222 patients were included and randomly allocated to receive agomelatine (151 patients) or placebo (71 patients) for 8 weeks.

There were no clinically relevant differences in demographic characteristics between the randomized groups at baseline in either study (***Table 2***). Compared to study 1, the duration of illness episode was somewhat greater in study 2, while the percentage of patient having previous psychotropic treatments was lower. The level of severity of the patients randomized in double blind treatment periods of studies 1 and 2 was not different between the groups at baseline in both studies. The MADRS scores (study 1), HAM-D_{17} scores (study 2) or CGI-S scores at baseline were similar in the two studies; though a somewhat greater level of HAM-D_{17} total score was found in patients of study 2.

The proportion of severely depressed patients was greater in study 2 (138 patients, 62.2% of the FAS) than in study 1 (86 patients, 40.6% of the FAS).

Efficacy

Whole study population

In study 1, the results over the 6-week period showed no statistically significant between-group difference in MADRS total score at last post-baseline evaluation without and with

adjustment for age, sex and geographic zone (***Table 3).*** Response rate was also not significantly different between groups. No statistically significant between-group difference was found for HAM-D_{17} total score at last post-baseline evaluation (***Table 3***).

No statistically significant differences between treatment groups were found for most secondary efficacy measurements, including CGI scores, the mean HAM-A total score, psychic anxiety score and somatic anxiety score in the FAS.

In study 2, agomelatine was associated with a statistically significant decrease in HAM-D_{17} total score at the last post-baseline value (Ancova: placebo *minus* agomelatine difference of 2.67 ± 1.06 points, p = 0.013) (***Table 3***). The MMRM sensitivity analysis provided consistent results (placebo *minus* agomelatine difference of 2.76 ± 1.02 (95% CI = [0.75 ; 4.78], p = 0.007). There was a significantly higher response rate on agomelatine (59.46%) *versus* placebo (38.57%) (p = 0.004) with a clinically relevant difference in favour of agomelatine of 20.89% (***Table 3***).

For the three SDS scores, patients reported significantly less symptom-related impairments on agomelatine than on placebo at work/daily activities (mean change from baseline to last-post baseline value: -3.1 ± 2.6 in the agomelatine group *versus* -2.0 ± 2.9 in the placebo group, p<0.001), in social life (-3.4 ± 2.8 in the agomelatine group *versus* -2.6 ± 2.8 in the placebo group, p=0.004), and in family life (-3.2 ± 2.9 in the agomelatine group *versus* -2.1 ± 2.5 in the placebo group, p=0.002).

In study 1, both CGI mean scores were similar in the agomelatine group at the last post-baseline value in the FAS (3.2 ± 1.5 for severity of illness, and 2.6 ± 1.3 for global improvement) and in the placebo group (3.3 ± 1.5 and 2.6 ± 1.3, respectively). The placebo *minus* agomelatine difference did not reach statistical significance, neither for CGI-S nor for CGI-I (***Table 3***).

In study 2, both CGI mean scores were lower in the agomelatine group (3.0 ± 1.3 for severity, and 2.2 ± 1.2 for global improvement) than in the placebo group (3.5 ± 1.3, and 2.6 ± 1.2, respectively). The placebo *minus* agomelatine difference was significant for CGI-S (0.48 ± 0.19; p=0.010) and for CGI-I (0.36 ± 0.17; p=0.034) (***Table 3***). The percentage of responders according to CGI-I was significantly higher in the agomelatine group (71.0%) than in the placebo group (50.0%) (p = 0.003).

Sub-population of more severely depressed patients

In both studies, agomelatine was associated with a statistically significant antidepressant effect *vs.* placebo in the primary criteria for efficacy. In study 1, at the last post-baseline value, the placebo *minus* agomelatine difference on MADRS total score was of 5.58 ± 2.32 points, p = 0.018). Significant between group differences in favour of agomelatine were also found on the HAM-D_{17} scale (***Table 4***). A significant placebo *minus* agomelatine difference was noted for CGI-S (p = 0.014) and CGI-I scores (p = 0.012) (***Table 4***).

In study 2, the analysis of the subset of severely depressed patients provided results consistent with the analysis in the whole population. The placebo-agomelatine difference was 3.79 ± 1.37 points (p=0.007) and the response rate by HAM-D_{17} was significantly higher in patients on agomelatine than on placebo (p=0.002). For CGI-I scores, a significant placebo *minus* agomelatine difference was noted (p = 0.006) (***Table 4***).

Tolerability

Some patients withdrew as a result of AEs during the short-term study periods: in study 1, 8 patients (7.3%) in the agomelatine group and 2 patients (1.8 %) in the placebo group; in study 2, 12 patients (7.9%) in the agomelatine group and 5 patients (7.0%) in the placebo group. In both

studies, the majority of the AEs were mild to moderate and the percentage of patients with at least one emergent AE was similar (study 1, agomelatine: 47.7% and placebo: 52.7%; study 2, agomelatine: 52.3% and placebo: 36.6%).

In both studies, patients reporting at least one severe emergent AE were comparably distributed between the agomelatine and the placebo group. The most common emergent AEs reported in the agomelatine group were generally the same as for placebo, and included somnolence, headache, dry mouth and diarrhoea.

During both studies, there were no clinically relevant mean changes in laboratory parameters and vital signs.

In study 2, two patients in agomelatine group (one patient on agomelatine 25mg and one on agomelatine 50mg) with normal liver enzymes values at baseline had emergent potentially clinically significant abnormal transaminases (> 3 ULN), probably related to treatment. All values normalized after agomelatine discontinuation.

No death occurred during both studies.

DISCUSSION

This is the second such formal publication comparing two agomelatine studies, the second of which was modified to remedy the factors we hypothesized to account for failure of the first (Goodwin et al., 2013). In both cases, the second study showed statistically significant drug-placebo differences. The different results observed between the two clinical placebo-controlled studies conducted with agomelatine in the older MDD population illustrate the fact that pharmacotherapy of these depressed patients is particularly challenging. This appears to be due to a significant variability in treatment response in populations that are very similar in clinical baseline characteristics. The significant elements of the improved methodology are of general interest to the field.

The results obtained in study 1, which failed to show a significant difference between the treatment groups in the entire population of study 1, were driven by a very high rate of response on placebo after 6 weeks (approx. 52% by MADRS scale). It is notable that response on placebo usually is about 40% (by HAM-D_{17} scale) or less in placebo-controlled studies in the older population- while response rate to agomelatine (46%) falls in the range of what is commonly seen with active treatments in these patients (Nelson et al., 2008;Tedeschini et al., 2011).

An extensive literature has suggested that inclusion of more severely ill patients with depression is predictive of a better response to antidepressants relative to placebo (Blom et al., 2007;Fournier et al., 2010;Khan et al., 2002;Khan and Schwartz, 2005;Kilts et al., 2009;Montgomery and Kasper, 2007). Analysis of study 1 in the subset of more severely depressed patients illustrated this statement. In addition, the analysis of HAM-D_{17} scores showed similar decrease in this population in both study 1 and 2 with a significant placebo-agomelatine difference of around 4 points. While this finding supported the potential of agomelatine, we have

previously demonstrated that simply pushing up the inclusion threshold on a single rating of symptoms to higher and higher levels is not the simple solution to improve selection and capture a population of patients with adequate severity (Goodwin et al., 2013). This approach reduces study feasibility and generalizability, can lead to inflation of scores at baseline and contrasts with how severity is actually estimated in clinical practice, where the focus is more on functional impairment.

To increase the assay sensitivity of the trial, key modifications of the protocols were added in the second study. We previously proposed to use measures, additional to a minimum entry score on a specific scale of depression in order to improve the quality of the diagnosis of well-characterized patients. Thus in addition to a classical minimum entry score on HAM-D_{17}, patients were required to have a CGI-S of at least 4 and third, the self-rating questionnaire HAD depression was included to avoid discordance between the clinician evaluation and the self perception by the patient (Fava et al., 2003;Kobak et al., 2007). The proposed cut-off of 11 is recognized as the threshold for defining a depressed state (Zigmond and Snaith, 1983). The Sheehan questionnaire (Leon et al., 1992) and the GDS aimed at checking the patients' characteristics, and were involved to confirm diagnosis criteria at selection. The Sheehan questionnaire (Leon et al., 1992), was employed to assess impairment in social, occupational or other important areas of functioning that are caused by depression. The GDS score of patients allowed confirmation that patients felt depressed. In addition, patients had to have a MMSE score ≥ 27 to check that no patient had relevant cognitive impairment or dementia.

Thus, by cross-checking the diagnosis from 3 different sources (diagnostic criteria; DSM, GDS, SDS), ratings by the investigator (HAM-D_{17} and CGI) and self-evaluation by the patients (HAD), more unsuitable mildly ill patients were probably excluded. This must be inferred because the apparent similarity of the patient samples on conventional measures at baseline is

evident in Table 2. Thus, simply using the inclusion threshold on a single scale (MADRS) in study 1 may have reduced the quality of patient recruitment; for example, 5% of patients included with a MADRS baseline score ≥24 had a baseline HAM-D score < 16. By contrast, in the subset of patients with MADRS ≥24 *and* HAM-D≥22 at baseline, there was a trend for a significant between-group difference in MADRS total score at last post-baseline evaluation (E[SE]= 3.22 (1.80), [-0.35 ;6.79] p=0.077, post-hoc analysis). The recruitment of a population of patients with few previous psychotropic treatments may also be a factor that influenced the differences in outcomes of the two trials. Though previous treatments were not an inclusion/exclusion criteria, the rate of patients with previous psychotropic treatments within the year prior to selection in study 2 (agomelatine: 37.1%; placebo: 43.7%) is numerically lower than the rate of patients with previous psychotropic treatments prior to inclusion in study 1 (agomelatine: 53.2%; placebo: 62.4%). Whatever, the exact mechanism, the responder rate observed on placebo was 38.6% after 8 weeks of treatment, and the benefit of agomelatine could be demonstrated. Our interpretation is that, although compared to the population of patients included in study 1, patients included in study 2 were not more impaired (levels of severity recorded at baseline according to CGI-S was not different between the two studies), they were *better characterized*; we propose that this leads to a higher proportion of patients responding to treatment, with numerically lower CGI-S scores *vs*. placebo at the end of the study period.

Other changes in the protocol of study 2 could be evoked, but they are unlikely to explain the different outcomes between the two studies. First, changing the main evaluation criterion has no effect on trial sensitivity. In study 1, where the HAM-D_{17} was used as secondary criterion, non significant results were observed as well. Second, in acute MDD studies, flexible dosing approximately doubles the likelihood of drug-placebo separation (Khan et al., 2003). However,

only 32 patients taking agomelatine had a dose increase in study 2 and the efficiency and the safety profile- of agomelatine observed in a majority of older patients treated with the daily dose of 25mg (78.8% of the FAS) replicate the outcome seen in adult treated patients. Therefore the change in dose regimen did not account for the different outcome observed in the two studies. Third, there were twice as many centres in study 1 (59 centers in 7 countries) than in study 2 (27 centers in 5 countries). While enrolling participants of more diverse demographics from a broader spectrum of geographical locations and various clinical settings, increase generalizability of the finding (ICH E9, 1999) it can also add to variance (Dechartres et al., 2011).

The two agomelatine studies also offer the opportunity to discuss some hypotheses that have been raised to explain the low level of response of older patients to available antidepressants. For example, contrary to a widely-held belief (Iosifescu et al., 2003), our findings do not show that the presence of coexisting illnesses and medication explain a weaker antidepressant efficacy; our conclusion has been reached by other authors (Sheikh et al., 2004;Harpole et al., 2005). In the positive study 2, there was a high level of comorbidities (91.9%) and the majority of patients had 1 or more concomitant medical illness and were treated with multiple medications. Clearly, this did not negatively influence the antidepressant action of agomelatine. It has been also hypothesized that a greater duration of the depressive episode in the older patient (Pettit et al., 2009) may result in a lower response to antidepressant treatments. However, compared to study 1, the increased duration of the current episode in the population of study 2 (5 months) has not prejudiced treatment efficacy. Of course, patients with duration of the presenting episode over 1 year may show lower response rate to treatment, and this requires further clarification. Finally, there is a belief that older patients take longer to respond to antidepressant treatment because of the physiological effects of aging and/or the more treatment-resistant nature of mood disorders

over time (Lebowitz et al., 1997). This idea has led to the conclusion that 12-weeks is the minimum trial duration necessary to identify response in an older person. In fact, the mean age of patients is marginally higher in the positive study 2 than in study 1 (67 *vs*. 71 years) with a higher rate of patients aged more than 75 years in study 2 (31% versus 9% in study 1). Therefore, if they are adequately characterized, an 8-week period is sufficient to detect a benefit for older patients. The advantage of agomelatine over placebo is even seen as early as 6 weeks after treatment initiation (with a placebo-agomelatine difference of 2.39 points by HAM-D_{17}, p=0.012; *post hoc* analysis). A *post-hoc* analysis of a subset of patients aged over 75 years in study 2 provided the evidence that the response to a 8-week treatment remained close to those observed in patients aged less than 75 years (around 60.0%) (Heun et al., 2013).

The present comparison between two agomelatine trials has some limitations as neither study was designed with this comparison in mind, but to assess drug efficacy. While the variables we have focused on are of interest, we cannot rule out other potentially significant differences in the design of both studies, or the play of chance, that could also have contributed some variance.

There have been 10 years between the two agomelatine studies in the older depressed population, a period of time during which a steady increase in the placebo effect has been reported in clinical trials (Walsh et al., 2002;Walsh and Sysko, 2005). This has reduced the possibility of demonstrating the benefits of antidepressants. Concomitantly, there is only a modest efficacy of antidepressants tested in trials during this period of time in older patients, with heterogeneous results across studies (Tedeschini et al., 2011). Clearly, the increased demands on investigators in study 2 have permitted us to reduce the placebo effect, and may have ensured a

recruited patient sample closer to those patients normally seeking treatment in clinical practice. We propose that systematic attention to patient characterisation may minimize the risks of inconclusive studies in the older depressed population.

Legends to figures

Figure 1: Disposition of included and randomised patients

References

Allard P, Gram L, Timdahl K, Behnke K, Hanson M, Sogaard J (2004). Efficacy and tolerability of venlafaxine in geriatric outpatients with major depression: a double-blind, randomised 6-month comparative trial with citalopram. Int J Geriatr Psychiatry 19:1123-1130

Blom MB, Spinhoven P, Hoffman T, Jonker K, Hoencamp E, Haffmans PM, van Dyck R (2007). Severity and duration of depression, not personality factors, predict short term outcome in the treatment of major depression. J Affect Disord 104:119-126

Bose A, Li D, Gandhi C (2008). Escitalopram in the acute treatment of depressed patients aged 60 years or older. Am J Geriatr Psychiatry 16:14-20

Dechartres A, Boutron I, Trinquart L, Charles P, Ravaud P (2011). Single-center trials show larger treatment effects than multicenter trials: evidence from a meta-epidemiologic study. Ann Intern Med 155:39-51

Fava M, Evins AE, Dorer DJ, Schoenfeld DA (2003). The problem of the placebo response in clinical trials for psychiatric disorders: culprits, possible remedies, and a novel study design approach. Psychother Psychosom 72:115-127

Folstein MF, Folstein SE, McHugh PR (1975). "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 12:189-198

Fournier JC, DeRubeis RJ, Hollon SD, Dimidjian S, Amsterdam JD, Shelton RC, Fawcett J (2010). Antidepressant drug effects and depression severity: a patient-level meta-analysis. JAMA 303:47-53

Goodwin GM, Boyer P, Emsley R, Rouillon F, de BC (2013). Is it time to shift to better characterization of patients in trials assessing novel antidepressants? An example of two relapse prevention studies with agomelatine. Int Clin Psychopharmacol 28:20-28

Guy W (1976). ECDEU Assessment Manual for Psychopharmacology. US Department of Health, Education and Welfare publication ADM 76-338 National Institute of Mental Health, Rockville MD217-222

HAMILTON M (1959). The assessment of anxiety states by rating. Br J Med Psychol 32:50-55

HAMILTON M (1960). A rating scale for depression. J Neurol Neurosurg Psychiatry 23:56-62

Harpole LH, Williams JW, Jr., Olsen MK, Stechuchak KM, Oddone E, Callahan CM, Katon WJ, Lin EH, Grypma LM, Unutzer J (2005). Improving depression outcomes in older adults with comorbid medical illness. Gen Hosp Psychiatry 27:4-12

Heun R, Ahokas A, Boyer P, Gimenez-Montesinos N, Pontes-Soares F, Olivier V (2013). The efficacy of agomelatine in elderly patients with recurrent major depressive disorder: a placebo-controlled study. J Clin Psychiatry 74:587-594

ICH E9 (1999). ICH Harmonised Tripartite Guideline. Statistical principles for clinical trials. International Conference on Harmonisation E9 Expert Working Group. Stat Med 18:1905-1942

Iosifescu DV, Nierenberg AA, Alpert JE, Smith M, Bitran S, Dording C, Fava M (2003). The impact of medical comorbidity on acute treatment in major depressive disorder. Am J Psychiatry 160:2122-2127

Katona C, Hansen T, Olsen CK (2012). A randomized, double-blind, placebo-controlled, duloxetine-referenced, fixed-dose study comparing the efficacy and safety of Lu AA21004 in elderly patients with major depressive disorder. Int Clin Psychopharmacol 27:215-223

Khan A, Khan SR, Walens G, Kolts R, Giller EL (2003). Frequency of positive studies among fixed and flexible dose antidepressant clinical trials: an analysis of the food and drug administration summary basis of approval reports. Neuropsychopharmacology 28:552-557

Khan A, Leventhal RM, Khan SR, Brown WA (2002). Severity of depression and response to antidepressants and placebo: an analysis of the Food and Drug Administration database. J Clin Psychopharmacol 22:40-45

Khan A, Schwartz K (2005). Study designs and outcomes in antidepressant clinical trials. Essent Psychopharmacol 6:221-226

Kilts CD, Wade AG, Andersen HF, Schlaepfer TE (2009). Baseline severity of depression predicts antidepressant drug response relative to escitalopram. Expert Opin Pharmacother 10:927-936

Kobak KA, Kane JM, Thase ME, Nierenberg AA (2007). Why do clinical trials fail? The problem of measurement error in clinical trials: time to test new paradigms? J Clin Psychopharmacol 27:1-5

Lebowitz BD, Pearson JL, Schneider LS, Reynolds CF, III, Alexopoulos GS, Bruce ML, Conwell Y, Katz IR, Meyers BS, Morrison MF, Mossey J, Niederehe G, Parmelee P (1997). Diagnosis and treatment of depression in late life. Consensus statement update. JAMA 278:1186-1190

Leon AC, Shear MK, Portera L, Klerman GL (1992). Assessing impairment in patients with panic disorder: the Sheehan Disability Scale. Soc Psychiatry Psychiatr Epidemiol 27:78-82

Montgomery SA, Asberg M (1979). A new depression scale designed to be sensitive to change. Br J Psychiatry 134:382-389

Montgomery SA, Kasper S (2007). Severe depression and antidepressants: focus on a pooled analysis of placebo-controlled studies on agomelatine. Int Clin Psychopharmacol 22:283-291

Nelson JC, Delucchi K, Schneider LS (2008). Efficacy of second generation antidepressants in late-life depression: a meta-analysis of the evidence. Am J Geriatr Psychiatry 16:558-567

Pettit JW, Lewinsohn PM, Roberts RE, Seeley JR, Monteith L (2009). The long-term course of depression: development of an empirical index and identification of early adult outcomes. Psychol Med 39:403-412

Rapaport MH, Lydiard RB, Pitts CD, Schaefer D, Bartolic EI, Iyengar M, Carfagno M, Lipschitz A (2009). Low doses of controlled-release paroxetine in the treatment of late-life depression: a randomized, placebo-controlled trial. J Clin Psychiatry 70:46-57

Rapaport MH, Schneider LS, Dunner DL, Davies JT, Pitts CD (2003). Efficacy of controlled-release paroxetine in the treatment of late-life depression. J Clin Psychiatry 64:1065-1074

Raskin J, Wiltse CG, Siegal A, Sheikh J, Xu J, Dinkel JJ, Rotz BT, Mohs RC (2007). Efficacy of duloxetine on cognition, depression, and pain in elderly patients with major depressive disorder: an 8-week, double-blind, placebo-controlled trial. Am J Psychiatry 164:900-909

Raskin J, Xu JY, Kajdasz DK (2008). Time to response for duloxetine 60 mg once daily versus placebo in elderly patients with major depressive disorder. Int Psychogeriatr 20:309-327

Roose SP, Sackeim HA, Krishnan KR, Pollock BG, Alexopoulos G, Lavretsky H, Katz IR, Hakkarainen H (2004). Antidepressant pharmacotherapy in the treatment of depression in the very old: a randomized, placebo-controlled trial. Am J Psychiatry 161:2050-2059

Rossini D, Serretti A, Franchini L, Mandelli L, Smeraldi E, De Ronchi D, Zanardi R (2005). Sertraline versus fluvoxamine in the treatment of elderly patients with major depression: a double-blind, randomized trial. J Clin Psychopharmacol 25:471-475

Schatzberg A, Roose S (2006). A double-blind, placebo-controlled study of venlafaxine and fluoxetine in geriatric outpatients with major depression. Am J Geriatr Psychiatry 14:361-370

Schneider LS, Nelson JC, Clary CM, Newhouse P, Krishnan KR, Shiovitz T, Weihs K (2003). An 8-week multicenter, parallel-group, double-blind, placebo-controlled study of sertraline in elderly outpatients with major depression. Am J Psychiatry 160:1277-1285

Sheehan DV, Harnett-Sheehan K, Raj BA (1996). The measurement of disability. Int Clin Psychopharmacol 11 Suppl 3:89-95

Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, Hergueta T, Baker R, Dunbar GC (1998). The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. J Clin Psychiatry 59 Suppl 20:22-33

Sheikh JI, Cassidy EL, Doraiswamy PM, Salomon RM, Hornig M, Holland PJ, Mandel FS, Clary CM, Burt T (2004). Efficacy, safety, and tolerability of sertraline in patients with late-life depression and comorbid medical illness. J Am Geriatr Soc 52:86-92

Tedeschini E, Levkovitz Y, Iovieno N, Ameral VE, Nelson JC, Papakostas GI (2011). Efficacy of antidepressants for late-life depression: a meta-analysis and meta-regression of placebo-controlled randomized trials. J Clin Psychiatry 72:1660-1668

Walsh BT, Seidman SN, Sysko R, Gould M (2002). Placebo response in studies of major depression: variable, substantial, and growing. JAMA 287:1840-1847

Walsh BT, Sysko R (2005). Placebo control groups in trials of major depressive disorder among older patients. J Clin Psychopharmacol 25:S29-S33

Yesavage JA (1988). Geriatric Depression Scale. Psychopharmacol Bull 24:709-711

Zigmond AS, Snaith RP (1983). The hospital anxiety and depression scale. Acta Psychiatr Scand 67:361-370