Head-to-head comparison of 1-year aripiprazole long-acting injectable (LAI) versus paliperidone LAI in comorbid psychosis and substance use disorder: impact on clinical status, substance craving, and quality of life

Ilaria Cuomo¹,²
Georgios D Kotzalidis¹
Simone de Persis²
Daria Piacentino¹
Filippo Perrini¹
Emanuela Amici²
Sergio De Filippis¹

¹Department of Neurosciences, Mental Health, and Sensory Organs (NEMOS), Sapienza School of Medicine and Psychology, Sant’Andrea Hospital, Rome, Italy; ²Department of Neuropsychiatry, Villa von Siebenthal Neuropsychiatric Hospital and Clinic, Genzano di Roma, Rome, Italy

Background: To overcome nonadherence in patients with psychosis switch to long-acting injectable (LAI) antipsychotic formulations is adopted. Most oral versus LAI comparisons showed similar antipsychotic responses. Psychoses often overlap with substance use disorder (SUD). Head-to-head LAI comparisons have hitherto focused only on non-comorbid populations.

Objective: The objective of this study was to compare two LAIs, administered for 12 months, in initially hospitalized patients with psychosis comorbid with SUD in their clinical and quality of life (QoL) outcomes.

Patients and methods: Inpatients were recruited during 2016 and switched randomly to 400 mg intramuscular aripiprazole monohydrate (AM) (N=50) or to 100 mg intramuscular paliperidone palmitate (PP) once-monthly (N=51); patients were discharged and followed up for 12 months. Patients were rated at baseline and after 1 year through the Clinical Global Impression–severity (CGIs), substance craving intensity was rated through a visual analog scale for substance craving, and QoL through the World Health Organization (WHOQOL–BREF) scale. We addressed confounders with backward stepwise logistic regression and three-way analysis of variance.

Results: PP were older and had more cases of schizophrenia spectrum and less bipolar disorders than AM, but AM had a stronger craving for substances at baseline. Both LAIs were associated with significant improvements in all outcomes, with AM displaying stronger effect sizes than PP. The two groups did not differ on baseline WHOQOL–BREF scores in any domain, but at the 1-year follow-up, AM fared better on all domains. The two groups did not differ in final severity, but PP scored higher than AM in craving at the 1-year endpoint.

Limitation: The CGIs is not a refined tool for severity and the substance craving may be subject to recall bias.

Conclusion: 1-year AM and PP was followed by improved clinical status and QoL and reduced substance craving in a population with psychosis and SUD comorbidity. AM, compared to PP, improved craving and QoL at the 1-year follow-up.

Keywords: psychosis, schizophrenia spectrum and other psychotic disorders, bipolar disorder, long-acting injectable antipsychotic drugs, aripiprazole once-monthly, paliperidone once-monthly
Introduction

About 35% of patients with schizophrenia have current substance use disorder (SUD) comorbidity, while lifetime comorbidity is 80%. SUD comorbidity reduces response to treatments in patients with schizophrenia and predicts relapse, since it complicates adherence, thus becoming a treatment priority in dealing with patients with schizophrenia. One way to overcome treatment nonadherence in patients with schizophrenia is to administer long-acting injectable (LAI) antipsychotics.

It is for complex reasons that currently the most prescribed LAIs are aripiprazole monohydrate (AM) and paliperidone palmitate (PP). Risperidone microspheres are quite costly to produce, for olanzapine pamoate there are concerns about postinjection syndrome as a side effect whose threat requires the patient to remain under observation in a protected environment for further 3 hours, while newly approved formulations like aripiprazole lauroxil and PP 3 months have not been launched in all countries. Hence, the most recent studies have focused on the efficacy, effectiveness, quality of life (QoL), functioning, and pharmacoeconomy of AM and PP and there have been head-to-head comparisons between the two. A first study on cost-effectiveness was conducted in a hypothetical cohort of patients and obtained 1-year estimates favoring AM. The second was a real sample, 71-site, 10-country study called QUALIFY, comparing AM and PP on QoL, clinical severity as assessed through the Clinical Global Impressions scale – severity (CGIs), and side effect profile after 6 months, and again favored AM. The subsequent two studies were regional and published as Congress Abstracts, but the detailed papers have still to appear. One, dealing with data from the UK, declared to have drawn data from QUALIFY, while the other – Spanish – did not, but used the same methodology and assessment time points. As expected, also in these studies AM superseded PP. A further study also used QUALIFY data to calculate expenses if the patients were in the US health care system (actually they were not, they were from the European Union [EU] or Canada) and found AM to be more convenient and effective than PP. Subsequently, the authors of QUALIFY in different order reported on physician-rated readiness for work, and patient-rated subjective well-being and tolerability/QoL; they found AM to be better. Then, they focused on side effects, and focused not on akathisia, a side effect known to be related to the use of aripiprazole, but rather on sexual dysfunction, which depends on serum prolactin and has been found to be lower with aripiprazole compared to other antipsychotics; they found AM to reduce prolactin and PP to elevate it and AM to improve sexual dysfunction when present, and that improved sexuality paralleled improved QoL. At last, the same authors who focused on readiness for work, subdivided their sample as yes or no. There were 49 patients with a yes at baseline and at 6 months, 118 with a no at both time points, and 41 with a no at baseline and a yes at 6 months. Patients who shifted from no to yes had a stronger improvement in QoL. The authors failed to report that there were significantly more people who converted to yes in the AM group than in the PP (chi-square =8.656; p=0.0032). So summarizing, we might state that there exists just one study conducted on patients and comparing head-to-head AM and PP that has yielded a database which has been the subject of subsequent, post hoc analyses.

There have also been two comparative analyses of double-blind studies versus placebo using AM or PP. The comparative analyses included respectively two studies with AM and one with PP and yielded no differences in efficacy or safety, with one AM study versus five (in fact, four) PP, and one aripiprazole lauroxil versus three PP studies, all versus placebo, that compared the LAI with placebo, and were used to measure efficacy and safety outcomes and compare them by employing meta-analytical procedures. The first indirect treatment comparison included Kane et al as the only AM study and four PP randomized controlled trials (RCTs), one of which was a post hoc analysis. The authors found an advantage of AM over PP in efficacy and for dropout due to inefficacy. The other indirect treatment comparison included the only aripiprazole lauroxil study and three PP RCTs versus placebo. This indirect analysis compared various doses of aripiprazole and PP and reported no significant differences in efficacy or tolerability measures. One might ask why no investigator to date has compared AM to aripiprazole lauroxil. Another comparative analysis used a mixed treatment comparison competing risks model to assess efficacy and binary models to assess safety. This work compared all RCTs of all LAIs, including two studies of AM versus placebo or versus suboptimal AM dose and one of PP versus risperidone microspheres, finding no significant differences in efficacy or safety and only a numerical advantage for AM in dropout rates. A systematic review that included only real-world data, that is, naturalistic studies and pragmatic RCTs, found PP to fare better than oral antipsychotics and similarly to AM as for efficacy, but less well as far as QoL is concerned, but the only possible comparison in this respect was the Naber et al study, the authorship of which was shared by Lundbeck/Otsuka people; the authors of the systematic review, who shared their...
work with a Janssen affiliate, mitigated the conclusions of Naber et al.9

Pharmacoeconomic studies may be based on real data or on simulations and models. After the above cited Citrome et al9 analysis that favored AM, two EU studies reported PP to dominate over AM,31,32 while a UK study found the opposite.33 Not surprisingly, among authors of papers finding an advantage for AM were people affiliated to Lundbeck or Otsuka, while in those finding an advantage for PP there were people from Janssen or Johnson & Johnson.

There are sound rationales for using both molecules in SUD. Paliperidone counteracts dizocilpine-induced blockade of the NMDA channel through inhibition of voltage-gated Ca2+ channels, in turn downplaying SIRT1 and allowing miRNA34 to turn back from suppression34 and restoring the activity of the Akt1/GSK3β signaling pathway,35 as well as counteracting dizocilpine-mediated increase of extracellular glutamate in rat prefrontal cortex.36 All this would render paliperidone suitable for treating SUD of NMDA inhibiting dissociative agents, and in fact, it was shown to suppress ketamine consumption and psychosis in a 34-year-old man.37 Also, aripiprazole was shown to counteract behavioral effects of not only the NMDA antagonists dizocilpine38–40 and phencyclidine41 in rodents but also those of cocaine42–44 and amphetamine,45–47 thus rendering it suitable for a wide range of SUDs.

Clinical trials in SUD patients of the two drugs showed some promise. Switching from other antipsychotics to aripiprazole showed symptom improvement in patients with comorbid SUD and bipolar or schizoaffective disorder.48 Aripiprazole reduced cocaine craving in patients with schizophrenia comorbid with SUD49,50 and with cocaine use disorder.51 However, in one study, despite decreasing craving, aripiprazole increased smoked cocaine self-administration in patients with cocaine use disorder who did not want to quit52 and in another, although it delayed relapse in cocaine abstainers, it increased craving compared to placebo.53 One meta-analysis confirmed,54 but not another did not,55 that aripiprazole reduced craving in cocaine use disorder. Aripiprazole was shown to reduce alcohol craving and clinical severity in alcohol use disorder.56,57 Regarding its use in amphetamine use disorder, aripiprazole did worse than methylphenidate and placebo in one study58 and not better than placebo in two others.59,60 However, it improved amphetamine-induced psychosis in two studies, although less than risperidone.61,62 There have been no studies on paliperidone in SUD-only patients, but just the aforementioned case report in ketamine addiction.37

Of the hitherto published studies that investigated each of the two LAIs we consider here in populations with comorbid psychosis and SUD, one regards AM and three regard PP. In a poster presentation not yet published after being peer-reviewed, Montemiro et al53 investigated AM in 18 patients with psychosis and comorbid SUD and reported improved psychosis and global functioning along with quitting substance use (12 patients, 67%). The other studies were conducted with PP. A group including Janssen-affiliated investigators found that those patients with schizophrenia, both with and without SUD comorbidity, who received PP, had significantly lower treatment failures and time to treatment failure than their counterparts who received oral antipsychotic medication.64 The other two papers including Janssen affiliates among their authors were focused on pharmacoeconomic issues. The first was a cost/service utilization study which compared PP with oral antipsychotics in veterans with schizophrenia and comorbid SUD and found a significant advantage of PP over oral atypical antipsychotic agents in total medical cost savings secondary to reduced hospitalization rates and lower rates of substance use-related health care resource utilization.65 The other examined treatment adherence and costs of PP versus oral antipsychotics in a Medicare-benefitting population with comorbid schizophrenia and SUD and reported higher treatment adherence and persistence, and less inpatient days in the PP group compared to the oral antipsychotics group, and similar costs.66 Considering globally these data, we deemed useful to use these two LAIs in populations affected from psychosis and SUD and to compare their impact on QoL, craving, and clinical status. Thus, we aimed to compare these outcomes in initially hospitalized patients with a psychotic disorder (schizophrenia spectrum disorders or psychotic bipolar disorder) and comorbid SUD, subsequently discharged and followed up for 1 year in a naturalistic setting. Given that there is a strong suspicion for sponsor bias in most studies in this field, we decided to conduct a non-sponsored investigation.

**Materials and methods**

**Patients**

All 125 inpatients, any sex, aged between 18 and 65 years, affected by DSDM, fifth edition (DSM-5) schizophrenia spectrum and other psychotic disorders or bipolar disorder with psychotic features comorbid with DSM-5 SUD were recruited during January 1, 2016 to December 31, 2016 at the Psychiatry Department of the Villa von Siebenthal Neuropsychiatric Hospital. Inpatients were given the option to switch from their oral therapy to an LAI; those who accepted were randomly...
assigned to 400 mg intramuscular (im) AM once-monthly (N=50, 41 men, 9 women, mean age =32.92; SD =12.115, range 19–57 years) or to 100 mg im PP once-monthly (N=51, 40 men, 11 women, mean age =36.92; SD =10.737, range 19–52 years) according to a block randomization design. Patients were started on treatment and remained hospitalized until stabilization. Thereafter, they were discharged and followed up at the outpatient facility of the clinic. A new hospitalization was considered to represent relapse (defined as at least one of the following: psychiatric hospitalization for psychotic symptoms; an increase on the CGIs of at least 2 points at two consecutive scheduled visits; clinically significant and deliberate self-harm or suicidal/homicidal behavior/ideation or aggression). All patients were reassessed for clinical/QoL measures at the 12-month endpoint.

Exclusion criteria included risk of suicidal or aggressive behavior, major medical conditions (diabetes, dementia, and other neurological diseases, cardiovascular or renal failure, and major respiratory diseases), use of LAI antipsychotics in the last 2 months, use of oral antipsychotics, involuntary admission to a psychiatric hospital, and pregnancy or lactation and women who did not use contraception.

Medications
People who were randomized to PP and were naïve to benzodiazepines received two oral doses of either risperidone 4–6 mg or paliperidone 6–9 mg on the week preceding the first im administration of PP. At baseline, patients randomized to PP received an intradeltoid injection of Xeplion (paliperidone palmitate, Janssen, Beerse, Belgium) 150 mg and 1 week later Xeplion 100 mg in the contralateral deltoid. Thereafter, monthly im Xeplion injections were administered in the gluteus muscle or in the deltoid, according to patient preference. Doses were adjusted to patient’s weight and clinical needs, but averaged 100 mg/month. Patients randomized to AM received intradeltoid or intragluteal injections of 400 mg/month Abilify Maintena (aripiprazole once monthly, Otsuka, Tokyo, Japan) at baseline and 10–20 mg/day oral aripiprazole for 2 weeks. Thereafter, they received 400 mg/month im AM in their deltoid or gluteus; a reduction to 300 mg/month was envisaged in case of adverse event development.

Other medications
During the course of their LAI treatment, patients were allowed to use benzodiazepines p.r.n. (to a maximum dose of 2 mg lorazepam equivalents) in case of increased anxiety or insomnia and non-opioid analgesics/antimigraine agents (ie, nonsteroidal anti-inflammatory drugs/salicylates, cyclooxygenase 2 inhibitors, triptans, and ergot alkaloids) in case of painful conditions. Patients were instructed to note in a diary any occurring adverse event along with its duration and severity (mild, moderate, severe) and any supplementary drug administration, with time of administration and dosage, and to report to the clinician at their subsequent visit.

Assessment
At baseline and at the 1-year endpoint we rated patients’ clinical status through the CGIs, their substance craving intensity through a visual analog scale for substance craving (VAScrav), and their QoL through the World Health Organization (WHOQOL-BREF) scale.

We used the first item of the CGI scale to rate the clinical severity of each patient. The CGI measures the severity of mental disease by asking the clinician to respond to just one question on a Likert scale ranging from 1 to 7, that is, “Considering globally your clinical experience of this particular patient population, how much is this patient mentally ill in this moment?” The item is rated as follows: 1= normal, not at all ill; 2= borderline mentally ill; 3= mildly ill; 4= moderately ill; 5= markedly ill; 6= severely ill; 7= among the most extremely ill patients. The rating is based on signs and symptoms observed and reported, on behavior and functioning over the 7 days preceding the assessment. Clearly, symptoms and behavior may fluctuate over 1 week; hence, the attributed score must reflect the mean of the last 7 days. This CGI item was found to possess better psychometric properties than the improvement item and showed good internal consistency and concurrent validity for nonpsychotic disorders. The CGI proved to be reliable in assessing symptoms and change in psychoses.

To evaluate intensity of craving, we used VAScrav; this is a 10-cm ruler or straight line with one extreme (0) meaning no craving and the other (10) extremely intense craving, allowing a continuous, non-discrete rating of the extent of craving, differently from Likert-type scales. Participants were invited to rate their peak craving the day before. VAScrav scales may refer specifically to individual substances or to any substance, in general. They are commonly used since Glassman’s et al 1984 report. Given their single-item nature, no reliability may be assessed, but an acceptable concurrent validity has been shown in one study.

QoL has been rated through the brief, abridged version of the WHOQOL-BREF scale. It is a self-rating scale consisting of 26 questions, each rated on a Likert scale 1–5, which are differently labeled, but range from the worst (1) to the best (5). Higher scores correspond to a better QoL.
The items belong to four domains, that is, physical health, psychological, social relationships, and environment. Scores range from 0 to 100 for each domain; the responses to each item are added in a given domain to obtain a raw score that is subsequently transformed to a 4–20 score and finally to a 0–100 score according to an established transformation table. Scores are considered in the context of each domain; they are not added to obtain a total score.

We used no other, disorder-specific tools, like the Positive and Negative Syndrome Scale or the Brief Psychiatric Rating Scale (BPRS) due to our population’s diagnostic heterogeneity. Ratings were conducted by IC, SdP, and SDF, all psychiatrists with a long-established inter-rater reliability (Fleiss’ kappa > 0.85 on both instruments).

The design of the study received approval from the Clinica von Siebenthal ethical committee. All patients signed free, informed consent for participating in the study after receiving sufficient information to understand study aims and procedures. Patients were explained the risk that their comorbid SUD harbored for treatment resistance and nonadherence. The study was conducted in accordance with the Principles of Human Rights, as adopted by the World Medical Association (WMA) at the 18th WMA General Assembly, Helsinki, Finland, June 1964 and subsequently amended by the 64th WMA General Assembly, Fortaleza, Brazil, October 2013.

Statistical analysis
We first used the Shapiro–Wilk normality test to test for normality of distribution of our data. Since data distribution was found to be normal (W=0.916, >0.01 threshold), except for CGIs data (W=0.581 for a threshold of 0.939 at p<0.05), we proceeded with parametric testing, save for CGI, for which we applied the nonparametric Mann–Whitney U test. We carried out Student’s t-test to compare intragroup and intergroup differences of continuous variables and the chi-square ($\chi^2$) test to compare intergroup differences on categorical variables. QoL scores for each domain are shown after adjusting for potential confounders (ie, demographic differences, diagnosis, comorbidity, and hospitalizations). The effect size (ES) of the two treatments was calculated according to Cohen and classified according to Cohen and Sawilowsky. The significant differences derived from the comparisons between variables in our descriptive statistics were further examined and confirmed by a forward stepwise logistic regression, which is a semi-automated process of building a model by successively adding variables based solely on the t-statistics of their estimated coefficients. More specifically, sociodemographic and clinical variables with a p-value <0.05 at the bivariate analysis were included in the multivariate analysis and the most parsimonious set of independent variables were selected through a forward procedure. $\chi^2$ with degrees of freedom and p-values were obtained to estimate the size of the association between type of treatment, on one hand, and patient sociodemographic characteristics, that is, gender, age, educational level, bipolar disorder type, episode type, and comorbid SUD, on the other hand, after the confounding effect of covariates was adjusted for, to the extent allowed by the data. Characteristics as “predictors” for treatment type were regarded as potential confounders and considered actual confounders if their distributions were substantially different in the two samples. The cutoff for statistical significance for logistic regression was set at p<0.05. The Nagelkerke’s pseudo-$R^2$ was calculated to determine the proportion of the variance in the dependent variable that is predictable from the independent variable(s). We conducted three-way analysis of variance (ANOVA) to investigate whether a three-way interaction existed between the three variables that were found to meet the level to stay in the logistic regression. Statistical analyses were conducted using the SPSS v.24 software (IBM Corporation, Armonk, NY, USA).

Results
Both drugs significantly improved baseline scores of all scales. AM and PP improved CGI-s from baseline to the 1-year follow-up (Mann–Whitney’s U test, Z=8.934; p<0.00001 and Z=8.70052; p<0.00001, respectively) with a Cohen’s d of 6.26 and 4.74, respectively, that is, a large ES according to Cohen and huge according to Sawilowsky. Both LAIs reduced craving intensity from baseline to 1 year, with AM t=20.395, p<0.00001; Cohen’s d=4.08, ES large or huge and PP with t=6.62, p<0.00001; Cohen’s d=1.31, ES large or very large. AM increased QoL in all domains with large/very large ESs (WHOQOL physical domain, baseline versus 1 year: t=−9.92, p<0.00001; Cohen’s d=1.98, ES large/very large; WHOQOL psychological domain, baseline versus 1 year: t=−9.55, p<0.00001; Cohen’s d=1.91, ES large/very large; WHOQOL social relationships domain, baseline versus 1 year: t=−9.55, p<0.00001; Cohen’s d=1.68, ES large/very large; WHOQOL environment domain, baseline versus 1 year: t=−11.43, p<0.00001; Cohen’s d=2.13, ES large/huge). Also, PP increased QoL in all domains (WHOQOL physical domain, baseline versus 1 year: t=−3.27, p<0.002; Cohen’s d=0.65, ES medium; WHOQOL psychological domain, baseline versus 1 year:
In the AM group, one patient was considered as remitter, while 33 were considered responders. In the PP group, there were no remitters and 29 responders. The two groups did not differ in the chi-square test ($\chi^2=0.213; p=0.645$).

To address possible confounders, we used forward stepwise logistic regression by entering in the model age, diagnosis (schizophrenia vs bipolar disorders), baseline CGIs, and baseline VAScrav. We found age, diagnosis, and VAScrav to be confounders, resulting in a distortion of the actual relationship between the independent and the dependent (treatment, ie, AM/PP) variables considered in the model (Table 2). Differently, CGIs at admission did not affect the relationship between the variables.

Nagelkerke’s pseudo-$R^2$ was 0.843, meaning that our regression model explained 84.3% of the variance of the dependent variable (AM/other treatments), that is, our regression model explained 84.3% of belonging to the AM or to the PP groups.

By performing three-way ANOVA, we found no significant interaction between and among age, diagnosis, and craving intensity in the determination of our results (Table 3).

Response/remission rates were based on the CGIs. Scores of 1 were considered to represent remission, whereas responders were considered patients who scored 1 or 2 at endpoint.
### Table 2 Forward stepwise logistic regression

<table>
<thead>
<tr>
<th>Model</th>
<th>Effects</th>
<th>-2 log likelihood of reduced model</th>
<th>$\chi^2$</th>
<th>df</th>
<th>$p$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Intercept</td>
<td>28.618</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Age (years)</td>
<td>90.469</td>
<td>61.851</td>
<td>29</td>
<td>0.000</td>
</tr>
<tr>
<td>2</td>
<td>Diagnosis</td>
<td>41.372</td>
<td>12.754</td>
<td>1</td>
<td>0.000</td>
</tr>
<tr>
<td>3</td>
<td>VAScrav</td>
<td>55.389</td>
<td>26.771</td>
<td>6</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Notes: Stepwise method: forward entry. Significant results are represented in bold characters. $\chi^2$ for data entry is based on the likelihood ratio test.

**Abbreviations:** df, degrees of freedom; VAScrav, visual analog scale for substance craving.

### Clinical and comorbid drug use disorders

In the AM group, 18 patients used alcohol, 38 cannabinoids, 20 cocaine, four MDMA, four ketamine, and three opioids (27 used multiple substances); six were diagnosed with schizophrenia, 16 with schizoaffective disorder, 10 with bipolar disorder I, and 18 with a first-episode psychosis (FEP).

In the PP group, 25 patients used alcohol, 23 cannabinoids, 10 cocaine, and two opioids (seven used multiple substances); 26 were diagnosed with schizophrenia, 15 with schizoaffective disorder, one with bipolar disorder I, and nine with an FEP. There were significantly more multiple drug use disorder cases in the aripiprazole LAI group than in the paliperidone LAI group ($\chi^2=8.72; p=0.003, <0.005$). Differences were not significant (ns) for alcohol ($\chi^2=0.71; p=0.400$, ns), cannabinoids ($\chi^2=2.51; p=0.11$, ns), cocaine ($\chi^2=2.74; p=0.098$, ns), and opioids ($\chi^2=0.21; p=0.65$, ns), while MDMA and ketamine use was absent in the paliperidone LAI group. The two groups were diagnostically heterogeneous. There were significantly more cases of schizophrenia in the paliperidone LAI group compared to the aripiprazole LAI group ($\chi^2=9.43; p=0.002, <0.01$) and more cases of bipolar disorder I in the aripiprazole LAI group compared to the paliperidone LAI group ($\chi^2=6.84; p=0.009, <0.01$), while the two groups did not differ for schizoaffective disorder ($\chi^2=0.04; p=0.837$, ns) and FEP ($\chi^2=2.52; p=0.112$, ns).

The two groups did not differ for additional comorbidities or for additional drug prescriptions. Additional prescriptions mainly regarded antiepileptics/mood stabilizers, benzodiazepines, and antidepressants (16 patients from the AM group and 14 from the PP group) and did not differ for dosing. Additional prescriptions present at baseline were reduced for both groups at the endpoint.

Regarding the effects of LAI treatment in patients with comorbid cannabinoid use disorder (all, N=38 aripiprazole; N=23 paliperidone), we performed Student’s t-tests, all two-tailed. To account for multiple comparisons, we used the following Bonferroni correction: 0.05/6=0.0083, cutoff for statistical significance. In this subgroup, AM was followed by significant improvements on the CGIs, VAScrav, and all WHOQOL domains at the P<0.0001 level, while PP was followed by significant improvements on the CGIs, the VAScrav, and only on the environment domain of the WHOQOL. People using cannabinoids and taking AM scored lower on craving intensity and higher in all WHOQOL domains after 1 year than patients who used cannabinoids and received PP. Patients using cannabinoids only were N=13 in the AM group and N=16 in the PP group. AM was followed by significant improvements at P<0.0001 on all considered measures, while PP was followed by improvements on only CGIs (P<0.00001) and VAScrav (P=0.004), and not on any WHOQOL domain. AM was followed after 1 year by significantly greater improvement on VAScrav than PP (P=0.0007) and only on two WHOQOL domains, that is, psychological (P=0.0004) and environment (P=0.006).

Comparisons between cannabinoid plus other substance users and cannabinoid only users showed no significant differences between AM and PP either at baseline or at the 1-year follow-up.

The 24 patients who did not accept to participate in the study did not differ in age, gender, socioeconomic or educational level, or marital status from the 101 who were included in the study. We are unable to compare their outcome data, since observation time points and treatments varied considerably.

During the entire year of observation, five patients on AM and 12 on PP met the criteria for relapse and needed short-term hospitalization (10 days on the average). Differences were ns (Table 1).

### Table 3 Three-way analysis of variance

<table>
<thead>
<tr>
<th>Origin</th>
<th>Type III df sum of squares</th>
<th>Mean square</th>
<th>F</th>
<th>$p$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correct model</td>
<td>6.833</td>
<td>0.976</td>
<td>4.915</td>
<td>0.000</td>
</tr>
<tr>
<td>Intercept</td>
<td>36.451</td>
<td>1</td>
<td>36.451</td>
<td>183.549</td>
</tr>
<tr>
<td>Age</td>
<td>1.317</td>
<td>1</td>
<td>1.317</td>
<td>6.632</td>
</tr>
<tr>
<td>Diagnosis</td>
<td>0.048</td>
<td>1</td>
<td>0.048</td>
<td>1.242</td>
</tr>
<tr>
<td>VAScrav</td>
<td>0.838</td>
<td>2</td>
<td>0.419</td>
<td>2.210</td>
</tr>
<tr>
<td>Age $\times$ diagnosis</td>
<td>0.305</td>
<td>1</td>
<td>0.305</td>
<td>1.355</td>
</tr>
<tr>
<td>Age $\times$ VAScrav</td>
<td>0.000</td>
<td>0</td>
<td>0.000</td>
<td>0</td>
</tr>
<tr>
<td>Diagnosis $\times$ VAScrav</td>
<td>0.387</td>
<td>2</td>
<td>0.193</td>
<td>0.974</td>
</tr>
<tr>
<td>Age $\times$ diagnosis $\times$ VAScrav</td>
<td>0.000</td>
<td>0</td>
<td>0.193</td>
<td>0.974</td>
</tr>
<tr>
<td>Error</td>
<td>18.667</td>
<td>94</td>
<td>0.199</td>
<td>0.199</td>
</tr>
<tr>
<td>Total</td>
<td>255,000</td>
<td>102</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>Adjusted total</td>
<td>25,500</td>
<td>101</td>
<td>0.000</td>
<td></td>
</tr>
</tbody>
</table>

Notes: $R^2=0.068$ (adjusted $R^2=0.233$). Significant results are represented in bold characters.

**Abbreviations:** df, degrees of freedom; VAScrav, visual analog scale for substance craving.
During the entire study period, there were few side effects reported, none of which led to dropout. However, the two akathisia cases developing after some time in the AM group led us to decrease the monthly dose from 400 to 300 mg; the side effect eventually subsided. Side effects were less with AM than PP and were typical of the respective oral formulations. In detail, the only side effect that developed in the AM group was akathisia, occurring in two patients, in one after 4 months and in the other after 5 months after establishing LAI administration. In the PP group, five patients developed galactorrhea, two after 3 months from LAI initiation, two others after 4 months, and one after 5 months. Hyperprolactinemia developed in four patients, each of whom had this symptom’s onset 2, 3, 4, and 6 months respectively after LAI initiation. Three of the patients who manifested hyperprolactinemia also developed galactorrhea. Two patients in the PP group developed increased substance craving (one cocaine, the other cannabinoid) after 4 and 5 months, respectively. In the same group, one patient developed avolition after 6 months and another sexual dysfunction after 7 months, but both side effects had subsided by the end of the observation period.

Discussion

In this naturalistic study, where we followed up for 1 year an about equal number of patients randomized to either AM LAI or PP LAI, we observed improvement with both drugs in clinical status, craving, and QoL. However, the effect of AM on craving and QoL was greater than that of PP.

This is the first head-to-head comparison of AM to PP in a population with psychoses and comorbid SUD. Our results are not affected by a possible sponsor bias as other studies with sponsor-affiliated author participation, but need to be discussed in the light of the different patient composition of the two samples. In fact, people randomized to AM were younger than people randomized to PP; furthermore, the latter had significantly more patients with schizophrenia spectrum disorders and less patients with bipolar disorder than the former. This did not interact with the type of drug use disorder, which did not differ significantly between the two groups. It is possible that people randomly assigned to PP were more chronic or socially impaired than those randomized to AM; however, their WHOQOL scores did not differ at baseline on any domain, whereas they were all pronouncedly higher for those taking AM compared to PP in all domains at the endpoint. This specificity of AM with respect to PP in improving QoL has been a finding also in another head-to-head comparison of AM and PP investigated in a large schizophrenia-only sample. This study, whose authors were affiliated to Otsuka or Lundbeck for their 84.62%, used a different scale than ours, that is, the Heinrichs–Carpenter QoL Scale; yet, it obtained similar results to ours in a different population. Interestingly, it matched our results also regarding side effects, in that they found more adverse effects with PP and akathisia with AM.

The effect of AM was not stronger than that of PP on the physical and social relationships domains of the WHOQOL in the sample that was comorbid with cannabis use disorder only; we may attribute this loss of the strength of significance to the reduced sample size of this subgroup.

Our results showed both drugs to be safe, with tolerable side effects. Each drug produced side effects that would have been expected by its oral formulation, that is, akathisia with aripiprazole and hyperprolactinemia/galactorrhea with paliperidone. Interestingly, there have been two cases of increased craving with PP; it should be noted that the AM group at baseline scored significantly higher on the craving intensity VAS scale than the PP group, whereas at the 1-year endpoint, the former scored significantly lower than the latter. Currently there are no reports that paliperidone could be involved with craving, while aripiprazole has been associated with reduced craving of alcohol in comorbid patients with schizoaffective or bipolar disorders and of cocaine in animals, probably due to a partial dopamine D_2 agonist effect that successfully addresses dopamine dysregulation.

We have no clue as to what did work for AM to supersede PP regarding QoL. However, it has been shown that dopaminergic agents may favorably impact QoL in patients with major depression. Since dopaminergic activity is related to reward and pleasure, it is possible that aripiprazole, which is a partial agonist of the dopaminergic D_2 receptors, by increasing the activity of the reward circuitry, could have enhanced QoL. Beyond the point that PP could, due to the latter’s exclusively antagonist effects on the same dopaminergic receptors.

Our forward stepwise logistic regression found age, diagnosis, and VAScrav to be confounders, as they distorted the relationships between these variables and AM or PP treatment, but baseline CGIs, that is, severity of mental illness did not affect these relationships. Our regression model explained 84.3% of the total variance, which can be considered satisfactory. Finally, the three-way ANOVA found no significant interaction between and among age, diagnosis, and craving intensity, that is, variables that differed between the two treatment groups at baseline, in the determination
of results. Hence, we might infer that our results were not affected by interactions.

A remarkable finding of our study was the complete adherence of our patients and the absence of dropouts. This needs replication.

Limitations
Assessment times were set wide apart, so we could not control the effect of the two LAIs on the course of the psychoses, but could do so only after 1 year. Furthermore, the assessment of symptom severity was carried out with only one generic instrument, the CGIs, that does not allow for parametric analyses, given that data tend to be skewed toward the extreme right of the range at intake and toward the left at endpoint. Perhaps, the BPRS would have been more suitable to provide stronger data on clinical change, since it has been successfully tested in people with dual diagnosis problems, but for many of our patients BPRS data were missing. Moreover, our choice to use a VAS measuring craving the day before could have exposed participants to recall bias. The open design of the study, despite randomization, may limit the validity of its data, but this is compensated by its naturalistic design.

Our study should not be interpreted as indicating the clinical superiority of AM over PP in patients with comorbid psychosis and SUD, but rather as an indication for further head-to-head comparisons to focus also on other variables than QoL and craving. In fact, in this as in another study, there was no clear-cut superiority of one LAI antipsychotic over the other, but only a clear difference favoring AM on QoL. At any rate, our industry non-sponsored study agrees partially with the findings of the Otsuka/Lundbeck-sponsored QUALIFY study, meaning that more non-sponsored studies are needed and that not all industry-sponsored studies produce biased results.

Conclusion
Our trial confirmed the superiority of AM over PP in improving QoL previously obtained in patients with schizophrenia, and also in patients with a comorbid psychotic and SUD, while it showed an anti-craving effect for all substances that were superior to that of PP. Further studies with more homogeneous groups at baseline and using more assessment scales are needed to confirm our results.

Acknowledgments
We gratefully acknowledge the contribution of the Librarians of the School of Medicine and Psychology of Sapienza University, Ms Mimma Ariano, Ms Felicia Proietti, Ms Ales Casciano, Ms Teresa Pioreschi, and Ms Susanna Rospo for rendering important bibliographical material accessible, as well as our Secretary Lucilla Martinelli for her assistance during the writing of this paper.

Author contributions
All authors contributed toward data analysis, drafting and revising the paper and agree to be accountable for all aspects of the work.

Disclosure
The authors report no conflicts of interest in this work.

References


