American Journal of Epidemiology **Submitted Manuscript**

Special Collection: Mental Health

Title: Antipsychotic drugs in first-episode psychosis: A target trial emulation in the FEP-CAUSAL Collaboration

Authors: Alejandro G. Szmulewicz1, Gonzalo Martínez-Alés1, Roger Logan1, Maria Ferrara2, Christian Kelly3, Diane Fredrikson4, Juan Gago5, Sarah Conderino5, Covadonga M Díaz-Caneja6, Joaquín Galvañ6, Lorna Thorpe5, Vinod Srihari2, Lakshmi Yatham4, Deepak K. Sarpal3, Ann K. Shinn7,8, Celso Arango6, Dost Öngür7,8, Miguel A. Hernán1,9, on behalf of the FEP-CAUSAL Collaboration

ORCiD IDs: Alejandro Szmulewicz: 0000-0002-2664-802X

Correspondence Address: Harvard TH Chan School of Public Health 677 Huntington Av, Boston, Massachusetts 02115

Joint Authorship: N/A

Affiliations: 1-CAUSALab, Department of Epidemiology, Harvard TH Chan School of Public Health, 2-USA Department of Psychiatry, Yale School of Medicine, USA 3-Department of Psychiatry, School of Medicine, University of Pittsburgh, Pittsburgh, PA, USA 4-University of British Columbia, Canada 5-Division of Epidemiology, Department of Population Health, New York University Grossman School of Medicine, USA 6-Department of Child and Adolescent Psychiatry, Institute of Psychiatry and Mental Health,
Hospital General Universitario Gregorio Marañón, IiSGM, CIBERSAM, ISCIII, School of Medicine, Universidad Complutense, Madrid, Spain.. 7-McLean Hospital, Belmont, Massachusetts, USA Department of Psychiatry, 8-Harvard Medical School, Boston, MA, USA 9-Department of Biostatistics, Harvard T.H. Chan School of Public Health

Key words: First Episode Psychosis, Psychotic Disorders, Schizophrenia, Antipsychotics, Comparative effectiveness, Tolerability, Target Trial Emulation, Benchmarking, Treatment

Acknowledgments[1]**:** None

Funding: Dr. Alejandro Szmulewicz was partly supported by a 2022 NARSAD Young Investigator Grant from the Brain & Behavior Research Fundation

Conflict of Interest: None

Disclaimer: None

Data Availability Statement: Upon reasonable request and credentialing.

[1] Study investigators, conference presentations, preprint publication information, thanks.

ABSTRACT

Good adherence to antipsychotic therapy helps prevent relapses in First Episode Psychosis (FEP). We used data from the FEP-CAUSAL Collaboration, an international consortium of observational cohorts to emulate a target trial comparing antipsychotics with treatment discontinuation as the primary outcome. Other outcomes included all-cause hospitalization. We benchmarked our results to estimates from EUFEST, a randomized trial conducted in the 2000s. We included 1097 patients with a psychotic disorder and less than 2 years since psychosis onset. Inverse probability weighting was used to control for confounding. The estimated 12-month risks of discontinuation for **aripiprazole, first-generation agents, olanzapine, paliperidone, quetiapine, and risperidone** (95% CI) were: 61.5**%** (52.5-70.6), 73.5**%** (60.5-84.9), 76.8**%** (67.2-85.3), 58.4**%** (40.4-77.4), 76.5**%** (62.1-88.5), and 74.4**%** (67.0-81.2) respectively. Compared with aripiprazole, the 12-month risk differences (95% CI) were -15.3**%** (-30.0, 0.0) for olanzapine, -12.8**%** (-25.7, -1.0) for risperidone, **and 3.0% (-21.5, 30.8) for paliperidone**. The 12-month risks of hospitalization were similar between agents. Our estimates support use of aripiprazole and paliperidone as first-line therapies for FEP. Benchmarking yielded similar results for discontinuation and absolute risks of hospitalization as in the original trial, suggesting that data from the FEP-CAUSAL Collaboration data sufficed to **approximately** remove confounding for these clinical questions.

Pharmacological therapy with antipsychotic agents is recommended for individuals with a first episode of psychosis (FEP). Because **in most cases** good adherence to therapy **may** prevent relapses, the decision to choose a particular antipsychotic needs to be informed by the comparative effectiveness and tolerability of the medication. In the early 2000s, the European First Episode Schizophrenia Trial (EUFEST) compared the 12-month risk of treatment discontinuation in FEP patients randomly assigned to either haloperidol (a first-generation antipsychotic agent) or to one of 4 second-generation agents: quetiapine, olanzapine, ziprasidone, or amisulpride (1). Second-generation antipsychotics were better tolerated than haloperidol, and no differences were found between second generation drugs.

Though EUFEST was a milestone in FEP research, it is not fully applicable to current FEP care: ziprasidone is not widely used, amisulpride is not approved in the US, and drugs that are commonly used (e.g., aripiprazole, risperidone, paliperidone) were not studied. EUFEST recruited nearly 500 FEP patients from over 14 countries and 49 centers over a 6-year period. Understandably, such an effort cannot take place every few years when clinical practice changes. More recent randomized trials have found it challenging to recruit adequate sample sizes of FEP patients (2,3). **Conversely, observational studies can benefit from a larger sample size. However, the few observational studies of antipsychotic effectiveness in individuals with first-episode psychosis only studied the most widely prescribed medications (olanzapine, risperidone, aripiprazole)(4) or were not able to adequately adjust for confounding due to disease severity (4,5).**

An alternative approach to obtain this information is to use observational data to explicitly emulate a pragmatic (hypothetical) randomized trial that would answer the causal question of interest—the target trial (6). An explicit emulation of a target trial eliminates design-induced biases related to selection and immortal time (7), though the observational estimates may still be biased if important confounders remain unmeasured. To increase confidence in estimates from observational data, one can emulate an existing randomized trial and then compare the observational estimates with those from the trial, a process referred to as benchmarking.(8)

Here, we used data from the FEP-CAUSAL Collaboration, an international consortium of observational cohorts of FEP patients in North America and Europe to emulate a target trial of first-line antipsychotic therapy. We then benchmarked our results to estimates from EUFEST.

METHODS

Our approach has two steps: 1) specify the protocol of the target trial and 2) emulate the target trial using the available observational data.(6)

Specification of the target trial

We designed the target trial emulation to follow closely the protocol from the EUFEST trial. The key components of the protocol are summarized in **Table 1**.

Eligibility criteria are age between 18-40 years, less than 2 years since psychosis onset, less than 6 weeks of antipsychotic use, and a diagnosis of schizophrenia, schizoaffective or schizophreniform disorder, or psychosis not otherwise specified.

The treatment strategies are initiation and continuation of monotherapy with (a) a first-generation antipsychotic agent (haloperidol 5-10 mg, fluphenazine 2.5-40 mg, or perphenazine 4-64 mg), (b) olanzapine 5-30 mg, (c) quetiapine 200-750 mg, (d) risperidone 2-6 mg, (e) aripiprazole 5-30 mg, and (f) paliperidone 3-12 mg. We group three high or medium potency first-generation antipsychotic agents together based on their similar mechanism of action and tolerability profile.(9) Individuals may initiate any other psychotropic agent (e.g., mood stabilizers, antidepressants) but no other antipsychotic agents under either strategy. **The use of a long-acting injectable formulation of the initially assigned antipsychotic agent is allowed, but not the use of a long-acting injectable formulation of a different antipsychotic.**

The primary outcome of interest is discontinuation of the treatment strategy, defined as 14 consecutive days with: (a) a dose outside the specified range assigned antipsychotic drug, or (b) the prescription of another antipsychotic agent. Other outcomes are all-cause psychiatric hospitalization and use of a concomitant mood stabilizer or antidepressant, as well as Clinical Global Impression-Severity (CGI-S) score and Global Assessment of Functioning (GAF) score at 12 months (though changes in clinical scores are difficult to detect over 1 year(10,11)).

For each eligible individual, follow-up starts at treatment assignment (baseline) and ends upon the outcome of interest, death, loss to follow-up (defined as 12 consecutive weeks without a clinical appointment), 1 year after baseline, or administrative end of follow-up (September 31, 2022), whichever occurs first.

The causal estimands of interest are the intention-to-treat effect and the per-protocol effect. For the intention-to-treat analysis of failure time outcomes (i.e., discontinuation,

hospitalization, co-prescription), we compare the 1-year risk under each strategy via risk differences. **The risks can be estimated nonparametrically by using a Kaplan-Meier estimator or parametrically by fitting a pooled logistic regression model with an indicator for assigned strategy, time of follow-up, and a product term between the strategy indicator and time of follow-up**. For continuous outcomes, we compare the mean scores between groups at 12 months.

For the per-protocol analysis, the analysis is restricted to those who initiate the assigned treatment and prognostic predictors of adherence are added as covariates to the **same outcome model.** Nonparametric bootstrapping with 500 samples can be used to calculate percentile-based 95% confidence intervals for risk estimates.

Emulation of the target trial

Data source

The FEP-CAUSAL Collaboration is an international consortium of observational cohorts of individuals with FEP. **Appendix S1** provides details on the information collected by the collaboration from FEP clinics in North America and Europe and **Table S1** provides information on study variables. This report includes data from OnTrack at McLean Hospital (Massachusetts), STEP at Connecticut Mental Health Center/Yale University-Yale (Connecticut), RAISE-ETP (US), NYU Langone Health (New York), Pittsburgh-STEP at UPMC (Pennsylvania), University of British Columbia (British Columbia, Canada) and 7 hospitals participating in AGES-CM (Madrid, Spain). Baseline data on sociodemographic (e.g., age, sex) and clinical factors (e.g., diagnoses, comorbidities) were extracted from the intake interview of the FEP clinics or the first interview with a mental health clinician. Longitudinal data were extracted from electronic health records. For each recorded clinical encounter, trained research assistants extracted data on medications use, substance use disorder, employment and educational statuses, hospitalizations, emergency room visits, and concomitant services received; psychiatrists and psychologists extracted information on symptomatic status and side-effects from the clinical notes. **Appendix S2 provides intraclass correlation coefficients to evaluate the reliability of the different symptom scales assessed by psychiatrists and psychologists from reading clinical notes.**

Data linkages were used to identify antipsychotic therapy from outpatient prescription data and hospitalizations. Most data extraction used clinical databases; one of the cohorts (RAISE-ETP) was a randomized trial launched by the National Institute of Mental Health and all data were collected from research questionnaires.

Target trial emulation

We used the observational data to emulate each protocol component of the target trial as closely as possible (**Table 1**). We identified patients who met the eligibility criteria (see **Table S2** for details on eligibility criteria implementation) and classified them into one of the six treatment groups according to the strategy that their data were compatible with at baseline.

We assumed treatment groups were exchangeable at baseline conditional on the baseline covariates age, race, sex, housing status, diagnosis, comorbid substance abuse, prior use of antipsychotics (including treatment failures), employment and educational status, and baseline symptomatic (Clinical Global Impression-Severity [CGI-S] score) and functional (Global Assessment of Functioning [GAF] score) statuses. We also included measures of frequency and intensity of substance use.

The analysis was identical to the per-protocol analysis of the target trial except that we emulated a sequence of target trials.(12) Specifically, we emulated a new target trial each week between August 1 2010 and September 1 2022. Each weekly trial included all patients who met the eligibility criteria and started one of the six antipsychotics of interest in that week. Each individual may participate in multiple trials.(7,13) **Sequential emulation accommodates the fact that individuals may meet the eligibility criteria at several times during the follow-up and results in more precise estimates than choosing just one of those times as time zero**. See **Appendix S3** for details on the structure of the analytical dataset. To explore the robustness of our findings, we conducted a number of sensitivity analyses. **First, we changed the outcome definition from all-cause psychiatric hospitalization to all-cause hospitalization to avoid relying on the valid recording of the main reason for hospitalization. Second, we changed the functional form of the continuous variables in our models (i.e., CGI-S, GAF, and intensity of substance use) to categories (Table S1) to explore the sensitivity of the estimates to different model specifications. Finally, we conducted our analysis by using only the first**

eligible week of each patient to assess the gain in statistical efficiency of sequential emulation.

Benchmarking

We emulated the protocol of the EUFEST trial with 3 differences.

First, because current clinical guidelines often recommend the use of second-generation drugs (largely based on EUFEST findings(14)), few patients in our study initiated haloperidol. Therefore, we combined the first-generation drugs haloperidol, fluphenazine, and perphenazine in a single group.

Second, in current clinical practice many psychiatrists may initiate a particular antipsychotic agent (e.g., paliperidone) with the a priori intention of switching to a long-acting injectable formulation during follow-up. However, switching to a long-acting injectable formulation of the same antipsychotic agent was not allowed in the protocol of EUFEST. Therefore, to emulate EUFEST, we modified the protocol of our target trial so that switching to a long-acting injectable formulation was not allowed.

Third, we used average hazard ratios estimated from a Cox model as effect measures for comparability with EUFEST. We considered the emulation successful if these estimates would have led to the same clinical decision as those from the index trial.

In sensitivity analyses, we standardized our estimates to the McLean population, which is similar to the EUFEST population in terms of distribution of risk factors (**Appendix S4**), and we conducted separate analyses by site to explore geographic variability (**Appendix S5**). **Finally, we included psychosis not otherwise specified (NOS) as an eligible diagnosis in our target trial emulation because it is common practice in the US to record this diagnosis at first encounter and update it to a more specific one when more clinical information is available. Psychosis NOS was not an eligible diagnosis in EUFEST, so in sensitivity analysis we repeated the analysis restricting to those with other diagnoses.**

RESULTS

Figure 1 shows a flowchart for the 1123 individuals who initiated one of the eligible antipsychotic drugs during the follow-up. Of 1236 eligible initiations (corresponding to 1097 unique individuals), 11.1% were first-generation drugs, 22.4% were aripiprazole, 18.0% were

olanzapine, 8.2% were quetiapine, 9.3% were paliperidone, and 31.0% were risperidone. **Table 2** shows the baseline characteristics of initiators. Initiators of first-generation drugs were less likely to be white and had higher mean scores of CGI-S. Aripiprazole initiators were more likely to be female and employed, and less likely to have a schizophrenia diagnosis at baseline as compared to initiators of other agents. Paliperidone initiators were more likely to have a diagnosis of schizophrenia. The median follow-up was 31.0 weeks (interquartile range: 12.0 – 52.0). The proportion of patients who were lost to follow up was 16.4%; 9 patients died during follow-up. **Most patients were eligible for only one trial during follow-up (92.3% of the sample), and patients who were eligible more than once had a higher mean CGI-S score at their second or third trial as compared to their first eligible trial. At the first trial, risperidone was the most used agent (33.1%) followed by aripiprazole (21.7%) and olanzapine (18.1%), at the second and third trial, aripiprazole was the most used agent (46.7%) followed by risperidone (20.0%) and paliperidone (13.3%).**

The intraclass correlation coefficient was 0.67 for the MIRECC-GAS symptom subscale, 0.83 for the CGI-S scale, 0.75 for the PANSS scale, and 0.68 for the BPRS scale (Appendix S2).

Discontinuation of treatment strategy

The nonstabilized IP weights had a mean of 5.9 (SD: 6.7), and a range from 1.0 to 100.3. **Figure 2** shows the standardized curves for time to discontinuation of treatment strategy. The 12-month risks (%) of discontinuation (95% CI) were 61.5% (52.5, 70.6) for aripiprazole, 73.5% (60.5, 84.9) for first-generation drugs, 76.8% (67.2, 85.3) for olanzapine, 58.4% (40.4, 77.4), for paliperidone, 76.5% (62.1, 88.5) for quetiapine, and 74.4% (67.0, 81.2) for risperidone. Compared with aripiprazole, the 12-month risk differences (95% CI) were -12.0% (-27.9, 5.0) for first-generation drugs, -15.3% (-30.0, 0.0) for olanzapine, -14.9% (-33.7, 1.4) for quetiapine, 3.0% (-21.5, 30.8) for paliperidone, and -12.8% (-25.7, -1.0) for risperidone. **Results were similar in sensitivity analyses (Table S3).**

All-cause psychiatric hospitalization

Figure 2 shows the cumulative incidences of all-cause psychiatric hospitalization. The 12-month risks (%) of all-cause hospitalization (95% CI) were 19.1% (12.9, 26.1) for aripiprazole, 19.2%

(12.4, 28.1) for FGA agents, 19.3% (12.6, 26.4) for olanzapine, 23.4% (12.1, 35.4) for paliperidone, 21.4% (11.8, 32.0) for quetiapine, and 28.1% (22.4, 34.2) for risperidone.

Compared with aripiprazole, the 12-month risk differences (95% CI) were -0.9% (-12.0, 12.0) for first-generation agents, -0.2% (-8.9, 8.9) for olanzapine, -2.2% (-15.5, 10.1) for quetiapine, -4.2% (-18.2, 8.9) for paliperidone, and -8.9% (-17.0, 0.0) for risperidone. **In sensitivity analysis, the risk differences were similar when considering all-cause hospitalization: 2.5% (-9.8, 11.6), 1.9% (-7.5, 11.7), -8.6% (-23.0, 4.6), -2.4% (-18.7, 11.8), and -9.3% (-19.1, 0.0), respectively, and when considering only the first trial per individual (Table S3).**

Other outcomes

Table 3 shows the mean CGI-S scores at 12 months and the percentage of patients on mood stabilizers and antidepressants.

Benchmarking against the EUFEST randomized trial

The absolute risk estimates of discontinuation and hospitalization, but not of co-prescription of antidepressants and mood stabilizers, were similar to those in EUFEST (**Table S4**). The proportion of patients that were lost-to-follow-up in EUFEST was 29.4%. Compared with first-generation drugs, the hazard ratios of discontinuation were 0.32 (0.24, 0.43) for olanzapine and 0.41 (0.28, 0.59) for quetiapine, similar to 0.28 (0.18, 0.43) and 0.52 (0.35, 0.76), respectively, in EUFEST. CGI-S scores were comparable and GAF scores were lower in our emulation than in EUFEST, but the differences between groups were largely compatible. Sensitivity analyses yielded results that were similar to those estimated in the main analyses **(Table S3** and **Appendix S4)**.

DISCUSSION

We used observational data to emulate a target trial of antipsychotic treatment in individuals with FEP. **We found that adherence to paliperidone and aripiprazole was higher than adherence to other second-generation agents,** and that the risk of psychiatric hospitalization was greater for risperidone compared with other agents. In line with previous studies,(3,15) these findings support the consideration of both aripiprazole and paliperidone as

first-line options. An advantage of both drugs is the possibility of transitioning to a long-acting injectable formulation.

The design of our target trial is an update of that of the EUFEST randomized trial, which compared therapeutic strategies that were common in the early 2000s. When we used our observational data to emulate a target trial with a design similar to that of EUFEST, we found similar hazard ratios of discontinuation for olanzapine and quetiapine vs. first-generation antipsychotics, as well as similar estimates of the 12-month risk of psychiatric hospitalizations and 12-month average CGI-S scores for first generation agents, olanzapine, and quetiapine. The larger sample size of our study resulted in more precise estimates.

We found two notable differences between results from the original EUFEST and estimated from our observational emulation. First, our observational emulation found that the risk of discontinuation in the olanzapine arm was about 40% higher as compared to the risk estimated in EUFEST. Second, the proportion of patients co-prescribed antidepressants was around 25% higher for quetiapine in the FEP-CAUSAL Collaboration. These discrepancies persisted even after standardizing our estimates to the patient characteristics of McLean OnTrack, the component of FEP-CAUSAL that most closely resembles the EUFEST population, which was older and with a higher proportion of individuals who were white, female, and employed or student. That is, even though the FEP-CAUSAL Collaboration collects high-quality data on treatments, outcomes, and confounders, a target trial emulation may not exactly replicate the index trial results because of differences in study population and variations in clinical practices. For example, current clinical guidelines recommend discontinuation of olanzapine if weight gain over 7% is detected (16), but this recommendation was not in place at the time EUFEST was conducted.

Our analysis has several limitations. First, like in any observational analysis, residual confounding may bias the results. However, we adjusted for key prognostic factors associated with treatment initiation (e.g., substance abuse, comorbidities, educational and working statuses, symptomatic assessment) and our estimates benchmarked those of a previous randomized trial. Second, misclassification of some outcomes is possible. The reason for psychiatric hospitalization was not always available or may have been recorded imperfectly. However, when we modified our outcome definition to include all-cause hospitalizations, results were largely unchanged, except for a higher risk of hospitalization in those initiating quetiapine. Also,

symptoms may be ascertained with error. However, our reliability analysis showed moderate to good reliability of our symptom measurements performed by chart review. Finally, model misspecification may have introduced bias in our effect estimates. However, our estimates were robust to changes in functional form of continuous covariates (i.e., from linear and quadratic to categorical).

Our estimates might not be generalizable to other FEP populations with a different distribution of variables such as age, prior treatment failures, or psychiatric comorbidities. However, the inclusion of different international cohorts as part of the FEP-CAUSAL Collaboration consortium allows for a wider representation of patients with FEP, which facilitates transportability analyses.

In summary, our estimates support the selection of aripiprazole and paliperidone as first-line therapies for the treatment of individuals with FEP. The compatibility with the randomized trial estimates suggests that the observational data from the FEP-CAUSAL Collaboration data sufficed to approximately remove confounding for these clinical questions.

REFERENCES

1. Kahn RS, Fleischhacker WW, Boter H, et al. Effectiveness of antipsychotic drugs in first-episode schizophrenia and schizophreniform disorder: an open randomised clinical trial. *Lancet*. 2008;371(9618):1085–1097.

2. San L, Arranz B, Perez V, et al. One-year, randomized, open trial comparing olanzapine, quetiapine, risperidone and ziprasidone effectiveness in antipsychotic-naive patients with a first-episode psychosis. *Psychiatry Res*. 2012;200(2–3):693–701.

3. Gómez-Revuelta M, Pelayo-Terán JM, Juncal-Ruiz M, et al. Antipsychotic Treatment Effectiveness in First Episode of Psychosis: PAFIP 3-Year Follow-Up Randomized Clinical Trials Comparing Haloperidol, Olanzapine, Risperidone, Aripiprazole, Quetiapine, and Ziprasidone. *Int J Neuropsychopharmacol*. 2020;23(4):217–229.

4. Whale R, Harris M, Kavanagh G, et al. Effectiveness of antipsychotics used in first-episode psychosis: a naturalistic cohort study. *BJPsych Open*. 2016;2(5):323–329.

5. Tiihonen J, Wahlbeck K, Lönnqvist J, et al. Effectiveness of antipsychotic treatments in a nationwide cohort of patients in community care after first hospitalisation due to schizophrenia and schizoaffective disorder: observational follow-up study. *BMJ*. 2006;333(7561):224.

6. Hernán MA, Robins JM. Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. *Am J Epidemiol*. 2016;183(8):758–764.

7. Hernán MA, Sauer BC, Hernández-Díaz S, et al. Specifying a target trial prevents immortal time bias and other self-inflicted injuries in observational analyses. *J Clin Epidemiol*. 2016;79:70–75.

8. Dahabreh IJ, Robins JM, Hernán MA. Benchmarking Observational Methods by Comparing Randomized Trials and Their Emulations. *Epidemiology*. 2020;31(5):614–619.

9. Stroup TS, McEvoy JP, Swartz MS, et al. The National Institute of Mental Health Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) project: schizophrenia trial design and protocol development. *Schizophr Bull*. 2003;29(1):15–31.

10. Crespo-Facorro B, Pérez-Iglesias R, Mata I, et al. Relapse prevention and remission attainment in first-episode non-affective psychosis. A randomized, controlled 1-year follow-up comparison of haloperidol, risperidone and olanzapine. *J Psychiatr Res*. 2011;45(6):763–769.

11. Crespo-Facorro B, de la Foz VO-G, Mata I, et al. Treatment of first-episode non-affective psychosis: a randomized comparison of aripiprazole, quetiapine and ziprasidone over 1 year. *Psychopharmacology (Berl)*. 2014;231(2):357–366.

12. García-Albéniz X, Hsu J, Hernán MA. The value of explicitly emulating a target trial when using real world evidence: an application to colorectal cancer screening. *Eur J Epidemiol*. 2017;32(6):495–500.

13. Danaei G, Rodríguez LAG, Cantero OF, et al. Observational data for comparative effectiveness research: an emulation of randomised trials of statins and primary prevention of coronary heart disease. *Stat Methods Med Res*. 2013;22(1):70–96.

14. Australian Clinical Guidelines for Early Psychosis - Second Edition Updated - Orygen, Revolution in Mind. (https://www.orygen.org.au/Training/Resources/Psychosis/Guidelines/Australian-Clinical-Guidelines-for-Early-Psychosis). (Accessed March 25, 2023)

15. Leucht S, Cipriani A, Spineli L, et al. Comparative efficacy and tolerability of 15 antipsychotic drugs in schizophrenia: a multiple-treatments meta-analysis. *Lancet*. 2013;382(9896):951–962.

16. Keepers GA, Fochtmann LJ, Anzia JM, et al. The American Psychiatric Association Practice Guideline for the Treatment of Patients With Schizophrenia. *Am J Psychiatry*. 2020;177(9):868–872.

Table 1. Specification of the protocol of a target trial and its emulation using data from the FEP-CAUSAL collaboration.

Protocol component	Index trial (EUFEST)	Target trial	Emulation using FEP-CAUSAL
Eligibility criteria	1) Individuals aged 18-40 between 2002-2006, 2) Diagnosis of schizophrenia, schizophreniform, or schizoaffective disorder validated using MINI questionnaire, 3) Less than 2 years since psychosis onset, 4) Antipsychotics used for less than 2 weeks in the previous year or 6 weeks at any time, 5) No known intolerance to study drugs and, 6) No contraindications to study drugs	Same as EUFEST, except 1) is 2012-2020 and 2) includes patients with psychosis NOS (clinical diagnoses were used)	Same as target trial. The use of a drug is assumed to imply that no contraindications existed for that drug.
Treatment strategies	Initiation and continuation of the following agents: 1) Haloperidol 1-4mg 2) Olanzapine 5-20mg 3) Quetiapine 200-750mg 4) Ziprasidone 40-160mg 5) Amisulpiride 200-800mg Individuals may initiate any other psychotropic agent (e.g., mood stabilizers, antidepressants) but no other antipsychotic agents under either strategy. Patients are not allowed to transition to a depot formulation.	Initiation and continuation of the following agents: 1) First generation agents: haloperidol 5-10 mg, fluphenazine 2.5-40 mg, perphenazine 4-64 mg. 2) Olanzapine 5-30 mg 3) Quetiapine 200-750 mg 4) Risperidone 2-6 mg 5) Aripiprazole 5-30 mg 6) Paliperidone 3-12 mg Individuals may initiate any other psychotropic agent (e.g., mood stabilizers, antidepressants) but no other antipsychotic agents under either strategy. Patients are	Same as target trial

		allowed to transition to a depot formulation of the same agent they initiated orally, but not to depot formulations for other antipsychotics.	
Treatment assignment	Patients were randomly assigned to one of the five treatment strategies and were aware of the drug they were randomized to	Same as EUFEST	Randomization was assumed conditional on the following baseline covariates: age, race, sex, housing status, diagnosis, comorbid substance abuse, prior use of antipsychotics, employment and educational status, and baseline symptomatic status (CGI-S score) and GAF score.
Outcomes	Discontinuation of treatment strategy, defined as dose out of range, or prescription of another antipsychotic for 14 consecutive days (including depot formulation). Other outcomes: Psychiatric hospitalizations, concomitant prescription of a mood stabilizer or antidepressant, mean CGI-S, GAF at 12 months.	Same as EUFEST. Except that a transition to a depot formulation of the *same* antipsychotic agent initiated orally was not considered a treatment discontinuation.	Same as target trial
Follow-up	Starts at treatment assignment and ends at the first to occur of loss to follow-up, death, 12 months, or outcome occurrence.	Same as EUFEST	Same as target trial
Causal	Intention-to-treat effect	Intention-to-treat effect, per-	Observational

contrast		protocol effect	analogue of the per-protocol effect
Statistical analysis	Intention-to-treat analysis: for failure time outcomes, comparison of 12-month risks (estimated via Kaplan-Meier or pooled logistic models) and hazard ratios (estimated through Cox proportional regression models); for continuous outcomes, comparison of means at 12 months. IP weighting to adjust for losses to follow-up. Per-protocol analysis: Same with standardization for baseline variables.	Same as EUFEST	Same as per-protocol analysis of target trial

Table 2. Baseline characteristics of initiators of antipsychotics, FEP-CAUSAL collaboration, 2012-2022. Numbers are percentage unless otherwise specified.

	First generation antipsychotics (n=137)	Aripiprazole (n=277)	Olanzapine (n=223)	Quetiapine (n=101)	Risperidone (n=383)	Paliperidone (n=115)
Sociodemographic characteristics						
Age (years, SD)	23.1 (3.5)	24.2 (4.6)	23.9 (4.6)	24.2 (5.0)	23.5 (3.9)	24.1 (3.8)
Female	27.7	38.6	19.7	30.7	20.6	20.0
White	43.6	57.2	57.2	53.6	51.4	48.7
Years of education (mean, SD)	11.8 (1.6)	12.3 (2.5)	12.1 (2.0)	11.6 (1.3)	11.9 (1.9)	11.8 (1.6)
Living alone	11.3	11.6	6.0	11.2	7.4	10.7
Employed or student	46.7	57.4	52.5	54.5	54.0	40.0
Diagnoses						
Schizophreniform	4.4	9.7	11.2	6.9	9.9	12.2
Schizoaffective	16.8	16.8	13.9	15.8	11.0	17.4
Schizophrenia	50.4	37.2	50.7	50.5	46.2	60.0
Psychosis NOS	28.5	26.4	22.9	25.7	29.2	9.6
Clinical characteristics						
Antipsychotic naïve	61.3	67.5	64.6	77.2	79.6	57.4
Substance use	**26.3**	**24.5**	**28.7**	**30.7**	**25.1**	**28.7**
CGI-S (mean, SD)	4.1 (0.8)	3.9 (1.0)	3.9 (0.9)	4.0 (0.8)	3.9 (1.0)	3.9 (1.0)
GAF score (mean, SD)	51.1 (11.5)	57.4 (10.6)	51.6 (11.4)	50.9 (11.9)	53.6 (11.8)	52.6 (12.2)
Overweight (>25 kg/m2)	27.0	18.4	10.8	18.8	16.2	16.5

SD: Standard Deviation; NOS: Not Otherwise Specified; CGI-S: Clinical Global Impression – Severity; GAF: Global Assessment of Functioning.

Table 3. Estimated effectiveness measures by antipsychotic treatment in first episode psychosis, FEP CAUSAL Collaboration, 2012-2022.

	First-generation agents	Olanzapine	Quetiapine	Risperidone	Paliperidone	Aripiprazole
	N = 137	N = 223	N = 101	N = 383	N = 115	N = 277
12-month risk of discontinuation (%) (95% CI)						
	73.5 (60.5, 84.9)	76.8 (67.2, 85.3)	76.5 (62.1, 88.5)	74.4 (67.0, 81.2)	58.4 (40.4, 77.4)	61.5 (52.5, 70.6)
12-month risk of hospitalization (%) (95% CI)						
	19.2 (12.4, 28.1)	19.3 (12.6, 26.4)	21.4 (11.8, 32.0)	28.1 (22.4, 34.2)	23.4 (12.1, 35.4)	19.1 (12.9, 26.1)
12-month mean CGI-S score (95% CI)						
	3.6 (3.3, 3.9)	3.3 (3.1, 3.5)	3.8 (3.4, 4.1)	3.2 (2.9, 3.5)	3.5 (3.2, 3.8)	3.6 (3.4, 3.8)
12-month mean GAF score (95% CI)						
	54.8 (49.7, 60.0)	58.0 (53.5, 61.0)	57.5 (50.1, 58.4)	61.0 (58.2, 63.5)	58.9 (52.2, 63.4)	57.2 (54.1, 58.5)
12-month co-prescription of (%) (95% CI)						
Antidepressants	43.6 (25.4, 61.8)	25.4 (15.1, 35.8)	35.2 (24.3, 46.2)	44.2 (25.0, 63.4)	19.7 (0.0, 39.7)	42.2 (29.5, 55.0)
Mood stabilizers	7.3 (0.0, 14.6)	11.6 (5.0, 18.2)	12.7 (6.8, 18.6)	8.7 (0.0, 17.5)	27.5 (6.3, 48.2)	1.5 (0.0, 2.6)

*** Adjusted for age, sex, race, housing status, baseline and time-varying CGI-S score, baseline and time-varying GAF score, years of education, antipsychotic use in the past, diagnosis (schizophreniform, schizophrenia, psychosis NOS, schizoaffective), Center, diagnosis of substance use disorder, frequency and type of substance use, employment and educational status at baseline and time-varying. Estimated risk differences were standardized to the joint distribution of the baseline covariates.**

Abbreviation: CI, confidence interval; CGI-S: Clinical Global Impression (Severity); GAF: Global Assessment of Functioning.

Figure legends.

Figure 1. Flowchart for selection of eligible persons from the FEP-CAUSAL Collaboration when emulating the EUFEST trial. Numbers in parentheses represent unique individuals in each group.

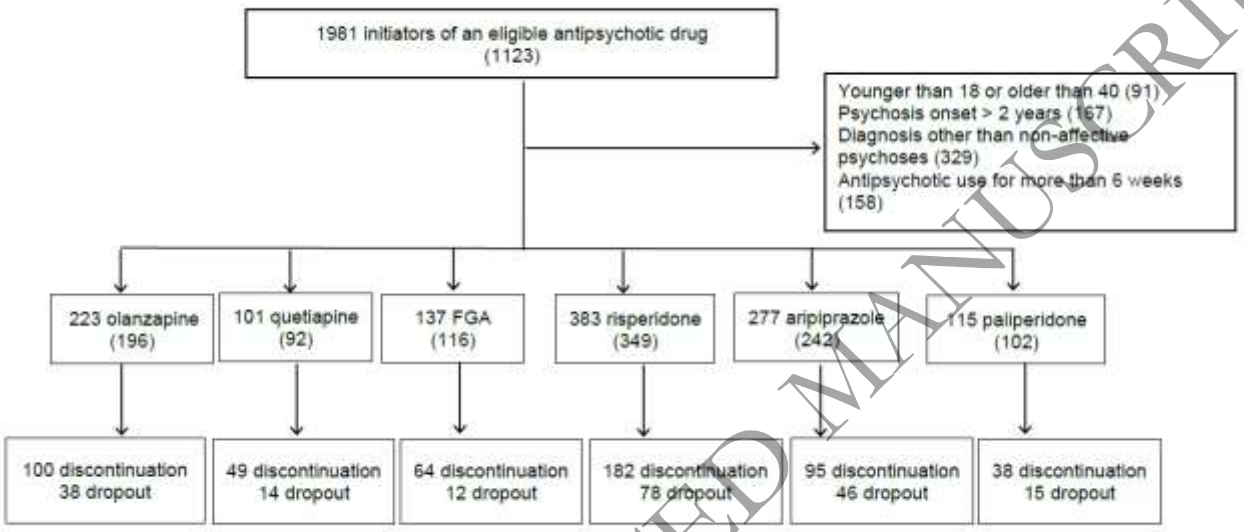

Figure 2. Estimated cumulative incidence of (a) all-cause treatment discontinuation and (b) all-cause psychiatric hospitalization by antipsychotic therapy, FEP-CAUSAL Collaboration (2012–2022).