**Association between poly-substance use and substance use disorder treatment noncompletion in patients** **admitted to multiple treatments between 2010-2019 in Chile**

Andrés González-Santa Cruz a.b,c; José Ruiz-Tagleb,c Mariel Mateo Pinones b,d; Álvaro Castillo-Carniglia b,c,e;

a Doctoral Program in Public Health, Institute of Population Health, Faculty of Medicine, Universidad de Chile, Independencia 939, Independencia, Santiago, Chile

b Millennium Nucleus for the Evaluation and Analysis of Drug Policies (nDP), Chile.

c Associate Professor, Department of Public Health, Facultad de Medicina y Ciencia, Universidad San Sebastian, Chile.

d School of Criminology and Criminal Justice, Griffith University, Australia

e Millennium Nucleus on Sociomedicine (SocioMed), Chile.

## Supplemental Section 1. Covariates

**Table S1. Description of covariates**

|  |  |
| --- | --- |
| **Covariate name** | **Description** |
| Biopsychosocial compromise | A clinical appraisal from professionals in the treatment team at admission to treatment, which considers the level of withdrawal symptoms, the motive of admission, motivation to change, the severity of SUD, number of previous treatments, number of physical complaints, and characteristics of the social environment (e.g., family functioning). Is coded by the authors as mild or moderate vs. severe (Ruiz-Tagle Maturana et al., 2023). |
| Age at Admission to Treatment (First Entry) | Age of the patient at the time of admission to treatment. The value of the first available treatment is considered. |
| Birth year | In calendar years (e.g., 1995). Obtained from the registered birth date. |
| Primary substance of the initial diagnosis | Indicated by patients as the starting substance of their drug use. SISTRAT allows the presentation of up to three starting substances, which is why we kept the most vulnerable category for patients who reported more than one value ("Base paste", "Cocaine hydrochloride", "Marijuana", "Alcohol", "Other"). On the other hand, other substances, such as hallucinogens, amphetamine-type stimulants, inhalants, opiates, sedatives, hypnotics and tranquilizers, and other substances (anabolic steroids, etc.), did not exceed 3% of the responses, and were therefore classified as "Other". |
| Psychiatric comorbidity (ICD-10) | According to the ICD-10 (International Classification of Diseases 10th Edition) classification. The system allowed for submitting up to three psychiatric conditions. If, at the time of dropout or other treatment termination, patients were not diagnosed but a condition is suspected by clinical staff (though not by a specialist), then the status is coded as "under study” or “incomplete assessment”. This status refers to cases under examination, so no information is available to confirm or rule out an accurate diagnosis. |
| Daily frequence of primary substance use at admission | Days of use in the last 30 days before treatment of the primary substance at admission. Patients can answer the following: Did not use, less than one day a week, one day a week or more, 2 to 3 days a week, 4 to 6 days a week and Daily. We coded as “Daily” vs. the rest of frequencies. |
| Occupational status | Coded from a single question asked whether the respondent worked at least one hour during the last week, excluding care work at home. If the respondent answered affirmatively, they were coded as “employed.” If the respondent answered negatively, their responses were categorized as follows:   * **Unemployed** for those who actively looking but unable to find work or reported being first-time job seekers. * **Inactive** or without any activity: This includes individuals studying but not working, permanently unable to work, pensioned or retired but not working, rentiers receiving income from properties or investments, those engaged in household chores, not seeking work, or other specified reasons. |
| Primary substance at admission to treatment | Identified and defined by the patient as the substance that motivates the consultation and admission, being the substance of greatest concern. |

Note. We consider the value of all covariates at the first treatment completed or not.

## Supplemental Section 2. Exploring data structure and regularity of repeated observations

The degree of irregularity of assessment times (times in which reported PSU was measured), in this case, in admission processes, is measured through the area under the curve (AUC) and log-transformed area under the curve. To inspect the degree of visit irregularity, the entire follow-up period was divided into adjacent and equally sized timeframes, and the number of bins changed. The AUC was obtained by plotting the mean proportions of individuals with 0 visits per bin against the mean proportions of individuals with > 1 visit per bin. AUC was log-transformed for better interpretation, in which 100 is very irregular, and 0 is equivalent to a repeated measures design (Lokku et al., 2020).

**Figure 2. Mean proportions of patients with 0,1, and >1 admissions per bin as bin width varies from 1 to 50% of the gap between visits.**

Imagen que contiene Diagrama

Descripción generada automáticamente

Note. Modified= Time in months from the first admission, and formatted to avoid overlapping between treatments, complete cases; Modified (imputation)= Time in months from the first admission, and formatted to avoid overlapping between treatments, imputed missing values; Original calendar= Formatted in calendar date, no further format.

According to Figure 2, when considering a bin width of 2% of the gap between treatments, the proportion of patients with two or more admissions was 0.33, decreasing to a proportion of 0.04 when considering a bin width of 6%. An AUC of 0.21 (log-transformed= 80) suggest there is likely a counting process behind visits rather than perfectly repeated and thus regular measures. Therefore, it seems reasonable to conduct an analysis to measure the association between reported PSU and non-completion statuses that considers irregularity in observations and an extent of informative assessment times, conditional on several covariates, including past observed outcomes, past assessment history, and baseline covariates.

## Supplemental Section 3. Counting process and alternative weighting schemes

Models such as marginal structural models, g-computation and targeted maximum likelihood estimation assume that observation times and gaps between them are not informative of the outcome of interest (Pullenayegum et al., 2023). The models adjust for the mentioned confounding variables. The study sample is based on a pseudo population in which the counting process (i.e., subsequent SUD treatment episodes and times between them) is static, hence, completely at random and ignorable (Carrero et al., 2023), based on generalized estimating equations and inverse probability weights given previously observed data (Cole & Hernán, 2008). We obtained these stabilized weights from a proportional intensity model in which we adjusted for baseline covariates, previous treatment outcomes, and previous polysubstance use (if any). The weights are represented through a proportional intensity model as: , where is a vector of covariates before , is a vector of regression coefficients, and the denominator is , a constant baseline hazard to stabilize weights (Pullenayegum, 2022). Auxiliary covariates may include confounders of the outcome model. Here, time is defined since the first individual admission in the study period (2010-2019), followed and divided by months (30.1 days). Inverse intensity weights to adjust for irregularity in the observation of treatment outcomes due to irregular admission to treatment patterns, were truncated at the 2.5th and 97.5th percentiles to mitigate the influence of extreme weights, thus, aiming for a more stable and reliable analysis. Covariates for weights are listed below: Treatment outcome of the previous treatment, previous biopsychosocial compromise (severe status), previous treatment duration (<90 days), previous treatment duration (in logarithmic scaled days), polysubstance use status of the previous treatment, age at admission to treatment, birth year, primary substance at admission to treatment (initial diagnosis; alcohol, cocaine hydrochloride, cocaine base paste, marijuana), psychiatric comorbidity under the International Classification of Diseases- 10th Revision (ICD-10) (confirmed comorbidity and diagnosis unknown or under study), daily frequency of primary substance use at admission, occupational status (inactive or unemployed), and primary substance at admission to treatment (alcohol, cocaine hydrochloride, cocaine base paste, and marijuana).

We analysed the visiting process of the patients readmitted to treatment as a function of several predictors. The intensity model identifies several significant factors of treatment visits. Previous treatment non-completion status (HR[*Hazard ratio*]= 1.17, 95% CI 1.13, 1.21), severe biopsychosocial compromise diagnosed in the previous treatment (HR= 1.06 95%CI 1.03, 1.10), less than 90 days in the previous treatment (HR= 1.11 95%CI 1.06, 1.16) a reduction of one unit in the log scale of days in the previous treatment (HR= 0.98 95%CI 0.96, 1.00), a one unit increase in age at admission to the initial treatment (HR= 1.26 95%CI 1.25, 1.27), and one unit increase in birth year (HR= 1.27 95%CI 1.26, 1.28) are associated with an increased likelihood of subsequent treatment visits. Specific substances at initial diagnosis such as cocaine base paste (HR= 1.17 95%CI 1.06, 1.30) and marijuana (HR= 1.17 95%CI 1.07, 1.29) (vs. other) and inactive (HR= 1.06 95%CI 1.02, 1.11) or unemployed (HR= 1.06 95%CI 1.03, 1.10) occupational status (vs. employed) also influence the time to returning for treatment, thus, being observed again in the database (See Table S2).

**Table S2. Specifications of the treatment (visit) process**

|  |  |
| --- | --- |
| **Term** | **Hazard ratio (95% CI)** |
| Treatment outcome of the previous treatment | 1.17 (1.13, 1.21) |
| Previous biopsychosocial compromise (severe) | 1.06 (1.03, 1.10) |
| Previous treatment duration (<90 days) | 1.11 (1.06, 1.16) |
| Previous treatment duration (in logarithmic scaled days) | 0.98 (0.96, 1.00) |
| Polysubstance use status of the previous treatment | 0.98 (0.95, 1.02) |
| Age at admission to treatment | 1.26 (1.25, 1.27) |
| Birth year | 1.27 (1.26, 1.28) |
| Primary substance (initial diagnosis), alcohol | 1.05 (0.95, 1.14) |
| Primary substance (initial diagnosis), cocaine | 1.12 (1.00, 1.25) |
| Primary substance (initial diagnosis), cocaine base paste | 1.17 (1.06, 1.30) |
| Primary substance (initial diagnosis), marijuana | 1.17 (1.07, 1.29) |
| Psychiatric comorbidity (diagnosis unknown or under study) | 1.03 (0.99, 1.08) |
| Psychiatric comorbidity (confirmed comorbidity) | 1.02 (0.99, 1.05) |
| Daily frequence of primary substance use at admission | 1.01 (0.98, 1.05) |
| Occupational status (inactive) | 1.06 (1.02, 1.11) |
| Occupational status (unemployed) | 1.06 (1.03, 1.10) |
| Primary substance at admission to treatment (alcohol) | 0.90 (0.80, 1.02) |
| Primary substance at admission to treatment (cocaine hydrochloride) | 0.90 (0.79, 1.01) |
| Primary substance at admission to treatment (cocaine base paste) | 0.91 (0.81, 1.03) |
| Primary substance at admission to treatment (marijuana) | 0.89 (0.78, 1.02) |

Note. 95%CI= 95% confidence intervals in parenthesis; Intensity model, in Andersen-Gill format Clustered by patients’ ID and stratified by treatment setting.

To tackle potential violations of the proportional hazards assumption for outcome assessment intensity, several strategies were tested: non-proportional hazards without variable transformations, proportional hazards with time-dependent transformations (recode multiple variables interacting with different functional forms of follow-up time), and stratifying follow-up times using the 'survSplit' method. Hence, three approaches were employed: (A) no correction for intensity proportionality violations through specifying time-varying coefficients (i.e., interacting with time), (B) correction using time-dependent coefficients via variable transformation and interaction with time, following the method outlined by Putter and colleagues (Putter et al., 2005), and (C) stratification of survival time among 30 different stratifications with better balance of. Each method considered cases where lag values of lagged variables (i.e., Treatment outcome of the previous treatment, Previous severe biopsychosocial compromise, Previous treatment duration of less than <90 days, Previous treatment duration in logarithmic scaled days, Polysubstance use status of the previous treatment) were 0 or 1 in case of missing binary values (i.e., accounting for potential missing values due to the absence of prior treatments within the study period), and ensuring a rigorous evaluation of model validity under potential assumption breaches. Continuous variables such as follow-up time and Previous treatment duration in logarithmic scaled days was fixed in 2.95 months and the natural logarithm of 90 days or 45 days if the rest of the lagged covariates were fixed in 0, respectively.

We recoded several covariates to account for its interaction with time through visual inspection of Schoenfeld residuals vs. time. This transformation aims to capture the changing impact of the previous treatment outcome over different time periods. We recoded the following: Treatment outcome of the previous treatment, previous biopsychosocial compromise, previous treatment duration (<90 days), primary substance (initial diagnosis), cocaine, and Psychiatric comorbidity (confirmed comorbidity).

Choosing the stratification involved considerations such as the distribution of events across different treatment settings and time intervals, which likely provided a good balance between granularity and statistical power, and improve model diagnostics such as the AIC (Akaike Information Criterion) and tests for proportional hazards (Keele, 2010; Zhang et al., 2018). Consequently, follow-up times were stratified into the following intervals: [0,10], (10,20], (20,30], (30,50], (50,70], and (70,135] for the second stratified model.

**Table S3. Specifications of the treatment (visit) process, in Hazard Ratios (HR) and under different lagged variables scenarios**

|  |  |  |
| --- | --- | --- |
| **Term** | **HR (95% CI) lag=0** | **HR (95% CI) lag=1** |
| Treatment outcome of the previous treatment | 0.65 (0.63, 0.67) | 1.30 (1.26, 1.34) |
| Previous treatment duration (in logarithmic scaled days) | 0.85 (0.84, 0.86) | 1.48 (1.44, 1.53) |
| Previous treatment duration (<90 days) | 0.69 (0.67, 0.72) | 2.63 (2.52, 2.74) |
| Previous biopsychosocial compromise (severe) | 0.88 (0.85, 0.90) | 1.54 (1.50, 1.57) |
| Polysubstance use status of the previous treatment | 0.62 (0.61, 0.64) | 1.26 (1.22, 1.30) |
| Age at admission to treatment | 1.05 (1.05, 1.06) | 1.06 (1.06, 1.06) |
| Birth year | 1.06 (1.06, 1.06) | 1.06 (1.06, 1.06) |
| Primary substance (initial diagnosis), cocaine | 1.01 (0.95, 1.08) | 1.05 (0.98, 1.12) |
| Primary substance (initial diagnosis), alcohol | 1.01 (0.96, 1.07) | 1.04 (0.99, 1.10) |
| Primary substance (initial diagnosis), cocaine base paste | 0.98 (0.92, 1.04) | 1.05 (0.99, 1.11) |
| Primary substance (initial diagnosis), marijuana | 1.04 (0.99, 1.10) | 1.03 (0.98, 1.09) |
| Psychiatric comorbidity (confirmed comorbidity) | 1.00 (0.99, 1.02) | 0.95 (0.94, 0.97) |
| Psychiatric comorbidity (diagnosis unknown or under study) | 1.00 (0.98, 1.03) | 0.74 (0.72, 0.76) |
| Daily frequence of primary substance use at admission | 1.02 (1.01, 1.04) | 0.95 (0.94, 0.97) |
| Occupational status (inactive) | 0.99 (0.96, 1.01) | 0.96 (0.94, 0.99) |
| Occupational status (unemployed) | 1.03 (1.01, 1.05) | 0.98 (0.96, 1.00) |
| Primary substance at admission to treatment (alcohol) | 0.99 (0.93, 1.06) | 1.03 (0.96, 1.09) |
| Primary substance at admission to treatment (cocaine hydrochloride) | 1.11 (1.04, 1.19) | 1.03 (0.97, 1.10) |
| Primary substance at admission to treatment (cocaine base paste) | 1.14 (1.07, 1.22) | 1.01 (0.95, 1.08) |
| Primary substance at admission to treatment (marijuana) | 1.04 (0.96, 1.12) | 1.05 (0.98, 1.12) |

Note. 95%CI= 95% confidence intervals in parenthesis; Intensity model, in Andersen-Gill format,

lag=0: Lagged covariates were fixed to 0 for binary variables and natural logarithm of 45 days in a hypothetical previous treatment stay for the first treatment; lag=1: Lagged covariates were fixed to 1 for binary variables and natural logarithm of 90 days in a hypothetical previous treatment stay for the first treatment.

**Table S4. Descriptive characterization of inverse intensity weights**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Visit intensity model** | **Min.** | **First quartile** | **Median** | **Mean** | **Third quartile** | **Max.** |
| No time-varying coefficients, lagged covariates fixed in 0 | 0.21 | 1.00 | 1.62 | 1.83 | 2.56 | 3.79 |
| No time-varying coefficients, lagged covariates fixed in 1 | 0.13 | 0.42 | 0.68 | 0.70 | 1.00 | 1.00 |
| With time-varying coefficients, lagged covariates fixed in 0 | 0.47 | 0.67 | 1.08 | 1.33 | 1.90 | 3.03 |
| With time-varying coefficients, lagged covariates fixed in 1 | 0.10 | 0.17 | 0.55 | 0.88 | 1.36 | 3.56 |
| Stratified by follow-up intervals, lagged covariates fixed in 0 | 0.29 | 0.60 | 0.86 | 0.99 | 1.18 | 3.02 |
| Stratified by follow-up intervals, lagged covariates fixed in 1 | 0.12 | 0.12 | 0.44 | 0.60 | 0.76 | 3.07 |

All models account for inverse intensity weights, which are crucial for adjusting the likelihood of treatment visit observations in the presence of irregular admission patterns.

## Supplemental Section 4. Model selection, alternatives accounting for overdispersion

In Poisson distribution, the variance equals the mean, implying that the theta value serves as the denominator of the squared mean and provides an indication of the degree of overdispersion present in the data. To evaluate the appropriate model for the data, a sequence of theta values ranging from 0.1 to and incrementing by 1,000. After finding an optimal value, a range of values from to incrementing by 100 were tested. Both Poisson and negative binomial models were compared using Quasi-likelihood Information criterion (QIC), a measure similar to Akaike Information Criterion for generalized estimating equations. The model that best fit the data (lowest QIC values) was selected based on this comparison. This approach highlights the methodology used to ascertain the presence of overdispersion and to determine the most suitable statistical model for handling such data deviations (Pan, 2001).

**Table S5. Relative risk of treatment non-completion status by reported polysubstance use**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Treatment setting** | **Model** | **QIC** | **RR** | **Sig** |
| Basic ambulatory |  |  |  |  |
|  | No time-varying coefficients, no weight | 19,452.9 | 1.02 (1.00, 1.05) | 0.0750 |
|  | No time-varying coefficients, no weight, NB | 19,452.9 | 1.02 (1.00, 1.05) | 0.0750 |
|  | No time-varying coefficients, lag=0 | 19,463.5 | 1.02 (0.99, 1.05) | 0.1665 |
|  | No time-varying coefficients, lag=0, NB | 19,463.5 | 1.02 (0.99, 1.05) | 0.1665 |
|  | No time-varying coefficients, lag=1 | 19,457.6 | 1.02 (1.00, 1.05) | 0.0838 |
|  | No time-varying coefficients, lag=1, NB | 19,457.6 | 1.02 (1.00, 1.05) | 0.0838 |
|  | With time-varying coefficients, no weight | 19,452.9 | 1.02 (1.00, 1.05) | 0.0750 |
|  | With time-varying coefficients, no weight, NB | 19,452.9 | 1.02 (1.00, 1.05) | 0.0750 |
|  | With time-varying coefficients, lag=0 | 19,477.3 | 1.03 (0.99, 1.06) | 0.1328 |
|  | With time-varying coefficients, lag=0, NB | 19,477.3 | 1.03 (0.99, 1.06) | 0.1328 |
|  | With time-varying coefficients, lag=1 | 19,514.7 | 1.03 (0.98, 1.07) | 0.2209 |
|  | With time-varying coefficients, lag=1, NB | 19,514.7 | 1.03 (0.98, 1.07) | 0.2209 |
|  | Stratified by follow-up intervals, no weight | 19,452.9 | 1.02 (1.00, 1.05) | 0.0750 |
|  | Stratified by follow-up intervals, no weight, NB | 19,452.9 | 1.02 (1.00, 1.05) | 0.0750 |
|  | Stratified by follow-up intervals, lag=0 | 19,465.3 | 1.04 (1.01, 1.07) | 0.0202 |
|  | Stratified by follow-up intervals, lag=0, NB | 19,465.3 | 1.04 (1.01, 1.07) | 0.0202 |
|  | Stratified by follow-up intervals, lag=1 | 19,502.8 | 1.03 (0.99, 1.07) | 0.1613 |
|  | Stratified by follow-up intervals, lag=1, NB | 19,502.8 | 1.03 (0.99, 1.07) | 0.1613 |
| General-population, intensive ambulatory |  |  |  |  |
|  | No correction | 22,259.1 | 1.04 (1.01, 1.07) | 0.0090 |
|  | No correction, NB | 22,259.1 | 1.04 (1.01, 1.07) | 0.0090 |
|  | No correction, lag=0 | 22,272.1 | 1.04 (1.01, 1.08) | 0.0112 |
|  | No correction, lag=0, NB | 22,272.1 | 1.04 (1.01, 1.08) | 0.0112 |
|  | No correction, lag=1 | 22,268.1 | 1.04 (1.01, 1.07) | 0.0127 |
|  | No correction, lag=1, NB | 22,268.1 | 1.04 (1.01, 1.07) | 0.0127 |
|  | With time-varying coefficients, no weight | 22,259.1 | 1.04 (1.01, 1.07) | 0.0090 |
|  | With time-varying coefficients, no weight, NB | 22,259.1 | 1.04 (1.01, 1.07) | 0.0090 |
|  | With time-varying coefficients, lag=0 | 22,280.5 | 1.05 (1.02, 1.09) | 0.0035 |
|  | With time-varying coefficients, lag=0, NB | 22,280.5 | 1.05 (1.02, 1.09) | 0.0035 |
|  | With time-varying coefficients, lag=1 | 22,318.8 | 1.05 (1.00, 1.09) | 0.0463 |
|  | With time-varying coefficients, lag=1, NB | 22,318.8 | 1.05 (1.00, 1.09) | 0.0463 |
|  | Stratified by follow-up intervals, no weight | 22,259.1 | 1.04 (1.01, 1.07) | 0.0090 |
|  | Stratified by follow-up intervals, no weight, NB | 22,259.1 | 1.04 (1.01, 1.07) | 0.0090 |
|  | Stratified by follow-up intervals, lag=0 | 22,265.5 | 1.02 (0.99, 1.05) | 0.1547 |
|  | Stratified by follow-up intervals, lag=0, NB | 22,265.5 | 1.02 (0.99, 1.05) | 0.1547 |
|  | Stratified by follow-up intervals, lag=1 | 22,276.1 | 1.01 (0.98, 1.05) | 0.4255 |
|  | Stratified by follow-up intervals, lag=1, NB | 22,276.1 | 1.01 (0.98, 1.05) | 0.4255 |
| General-population, residential |  |  |  |  |
|  | No correction | 9,817.2 | 0.97 (0.92, 1.02) | 0.1904 |
|  | No correction, NB | 9,817.2 | 0.97 (0.92, 1.02) | 0.1904 |
|  | No correction, lag=0 | 9,832.0 | 0.97 (0.92, 1.02) | 0.2528 |
|  | No correction, lag=0, NB | 9,832.0 | 0.97 (0.92, 1.02) | 0.2528 |
|  | No correction, lag=1 | 9,826.1 | 0.95 (0.90, 1.01) | 0.0991 |
|  | No correction, lag=1, NB | 9,826.1 | 0.95 (0.90, 1.01) | 0.0991 |
|  | With time-varying coefficients, no weight | 9,817.2 | 0.97 (0.92, 1.02) | 0.1904 |
|  | With time-varying coefficients, no weight, NB | 9,817.2 | 0.97 (0.92, 1.02) | 0.1904 |
|  | With time-varying coefficients, lag=0 | 9,828.8 | 0.98 (0.92, 1.03) | 0.4442 |
|  | With time-varying coefficients, lag=0, NB | 9,828.8 | 0.98 (0.92, 1.03) | 0.4442 |
|  | With time-varying coefficients, lag=1 | 9,865.7 | 1.00 (0.93, 1.07) | 0.9636 |
|  | With time-varying coefficients, lag=1, NB | 9,865.7 | 1.00 (0.93, 1.07) | 0.9636 |
|  | Stratified by follow-up intervals, no weight | 9,817.2 | 0.97 (0.92, 1.02) | 0.1904 |
|  | Stratified by follow-up intervals, no weight, NB | 9,817.2 | 0.97 (0.92, 1.02) | 0.1904 |
|  | Stratified by follow-up intervals, lag=0 | 9,836.6 | 0.97 (0.91, 1.04) | 0.4096 |
|  | Stratified by follow-up intervals, lag=0, NB | 9,836.6 | 0.97 (0.91, 1.04) | 0.4096 |
|  | Stratified by follow-up intervals, lag=1 | 9,883.8 | 1.01 (0.92, 1.10) | 0.9098 |
|  | Stratified by follow-up intervals, lag=1, NB | 9,883.8 | 1.01 (0.92, 1.10) | 0.9098 |
| Women-only, intensive ambulatory |  |  |  |  |
|  | No correction | 3,420.7 | 0.99 (0.92, 1.05) | 0.7075 |
|  | No correction, NB | 3,420.7 | 0.99 (0.92, 1.05) | 0.7075 |
|  | No correction, lag=0 | 3,431.1 | 0.99 (0.92, 1.07) | 0.8688 |
|  | No correction, lag=0, NB | 3,431.1 | 0.99 (0.92, 1.07) | 0.8688 |
|  | No correction, lag=1 | 3,426.7 | 0.99 (0.92, 1.06) | 0.7381 |
|  | No correction, lag=1, NB | 3,426.7 | 0.99 (0.92, 1.06) | 0.7381 |
|  | With time-varying coefficients, no weight | 3,420.7 | 0.99 (0.92, 1.05) | 0.7075 |
|  | With time-varying coefficients, no weight, NB | 3,420.7 | 0.99 (0.92, 1.05) | 0.7075 |
|  | With time-varying coefficients, lag=0 | 3,441.0 | 1.01 (0.92, 1.10) | 0.8228 |
|  | With time-varying coefficients, lag=0, NB | 3,441.0 | 1.01 (0.92, 1.10) | 0.8228 |
|  | With time-varying coefficients, lag=1 | 3,463.6 | 1.01 (0.91, 1.12) | 0.8719 |
|  | With time-varying coefficients, lag=1, NB | 3,463.6 | 1.01 (0.91, 1.12) | 0.8719 |
|  | Stratified by follow-up intervals, no weight | 3,420.7 | 0.99 (0.92, 1.05) | 0.7075 |
|  | Stratified by follow-up intervals, no weight, NB | 3,420.7 | 0.99 (0.92, 1.05) | 0.7075 |
|  | Stratified by follow-up intervals, lag=0 | 3,434.3 | 0.98 (0.91, 1.05) | 0.5343 |
|  | Stratified by follow-up intervals, lag=0, NB | 3,434.3 | 0.98 (0.91, 1.05) | 0.5343 |
|  | Stratified by follow-up intervals, lag=1 | 3,465.0 | 0.94 (0.87, 1.02) | 0.1571 |
|  | Stratified by follow-up intervals, lag=1, NB | 3,465.0 | 0.94 (0.87, 1.02) | 0.1571 |
| Women-only, residential |  |  |  |  |
|  | No correction | 4,823.9 | 1.14 (1.06, 1.23) | 0.0006 |
|  | No correction, NB | 4,823.9 | 1.14 (1.06, 1.23) | 0.0006 |
|  | No correction, lag=0 | 4,836.4 | 1.15 (1.06, 1.26) | 0.0009 |
|  | No correction, lag=0, NB | 4,836.4 | 1.15 (1.06, 1.26) | 0.0009 |
|  | No correction, lag=1 | 4,831.9 | 1.13 (1.04, 1.22) | 0.0027 |
|  | No correction, lag=1, NB | 4,831.9 | 1.13 (1.04, 1.22) | 0.0027 |
|  | With time-varying coefficients, no weight | 4,823.9 | 1.14 (1.06, 1.23) | 0.0006 |
|  | With time-varying coefficients, no weight, NB | 4,823.9 | 1.14 (1.06, 1.23) | 0.0006 |
|  | With time-varying coefficients, lag=0 | 4,837.1 | 1.11 (1.02, 1.21) | 0.0125 |
|  | With time-varying coefficients, lag=0, NB | 4,837.1 | 1.11 (1.02, 1.21) | 0.0125 |
|  | With time-varying coefficients, lag=1 | 4,857.5 | 1.13 (1.03, 1.25) | 0.0115 |
|  | With time-varying coefficients, lag=1, NB | 4,857.5 | 1.13 (1.03, 1.25) | 0.0115 |
|  | Stratified by follow-up intervals, no weight | 4,823.9 | 1.14 (1.06, 1.23) | 0.0006 |
|  | Stratified by follow-up intervals, no weight, NB | 4,823.9 | 1.14 (1.06, 1.23) | 0.0006 |
|  | Stratified by follow-up intervals, lag=0 | 4,839.3 | 1.11 (1.02, 1.20) | 0.0108 |
|  | Stratified by follow-up intervals, lag=1, NB | 4,839.3 | 1.11 (1.02, 1.20) | 0.0108 |
|  | Stratified by follow-up intervals, lag=1 | 4,870.0 | 1.09 (0.99, 1.20) | 0.0718 |
|  | Stratified by follow-up intervals, lag=1, NB | 4,870.0 | 1.09 (0.99, 1.20) | 0.0718 |

Note. NB= Negative binomial; QIC= Quasi-likelihood information criteria; lag=0: Lagged covariates were fixed to 0 for binary variables and natural logarithm of 45 days; lag=1: Lagged covariates were fixed to 1 for binary variables and natural logarithm of 90 days.

All models adjusted for the following covariates: biopsychosocial compromise (severe status) at admission to treatment, Age at admission to treatment, Birth year, Primary substance of the initial diagnosis (cocaine hydrochloride, cocaine base paste, marijuana, other substances), Psychiatric comorbidity (in study and with comorbidity), Daily frequence of primary substance use at admission, Occupational status (inactive and unemployed), Primary substance at admission to treatment (Cocaine hydrochloride, cocaine base paste, marijuana, other substances).

The magnitude and direction of the associations presented in the main analysis remained stable despite different weighting schemes and different distributions used, with intervals including the null across all of the models and scenarios of lagged variables for intensive ambulatory settings specific for women and residential settings for the general population. However, we observe that for the basic ambulatory settings, the model with stratified times of follow-up with lagged time-varying covariates fixed in 0 showed a non-overlapped association between PSU and treatment non-completion (RR 1.04 95% CI 1.01, 1.07), although very modest. In contrast, in intensive ambulatory settings for the general population, the association lose its strength, particularly among models where follow-up times were stratified by follow-up intervals, with its lower strength among lagged scenarios of time-varying covariates fixed in 1 (RR 1.01 95% CI 0.98, 1.05). For residential settings exclusive to women, the only model that overlapped with the null was the stratified by follow-up intervals, where lagged time-varying covariates were fixed at 1 (RR 1.09 95% CI 0.99, 1.20). However, both settings conserve its positive direction.

## Supplemental Section 5. The role of reporting PSU with and without alcohol as a secondary substance in treatment non-completion

A GEE model with a Poisson distribution was conducted to elucidate the association between reporting PSU with alcohol as a secondary substance, PSU without alcohol as a secondary substance, and reporting single substance use as the reference category, and the risk of treatment non-completion.

**Table S6. Association between polysubstance use at admission, with or without alcohol as a secondary substance, and treatment noncompletion** **status (dropout or spelled by misconduct) from Poisson Regression Multivariable Model with independence structure**

|  |  |  |
| --- | --- | --- |
| **Treatment setting** | **PSU status (ref.= No PSU)** | **RR (95%CI)** |
| Basic ambulatory | PSU with alcohol | 1.01 (0.98, 1.04) |
|  | PSU w/o alcohol | 1.08 (1.05, 1.12) |
| General-population, intensive ambulatory | PSU with alcohol | 1.02 (0.99, 1.05) |
|  | PSU w/o alcohol | 1.10 (1.07, 1.14) |
| General-population, residential | PSU with alcohol | 0.99 (0.94, 1.04) |
|  | PSU w/o alcohol | 0.89 (0.83, 0.94) |
| Women-only, intensive ambulatory | PSU with alcohol | 0.96 (0.89, 1.02) |
|  | PSU w/o alcohol | 1.07 (0.99, 1.15) |
| Women-only, residential | PSU with alcohol | 1.14 (1.06, 1.23) |
|  | PSU w/o alcohol | 1.14 (1.05, 1.24) |

Note. RR= Relative risk; 95%CI= 95% confidence intervals in parenthesis.

Model was adjusted for the following covariates: biopsychosocial compromise (severe status) at admission to treatment, Age at admission to treatment, Birth year, Primary substance of the initial diagnosis (cocaine hydrochloride, cocaine base paste, marijuana, and other substances), Psychiatric comorbidity (in study and with comorbidity), Daily frequence of primary substance use at admission, Occupational status (inactive and unemployed), Primary substance at admission to treatment (Cocaine hydrochloride, cocaine base paste, marijuana, and other substances).

As seen in Table S6, patients in residential settings exclusive to women showed associations for both patients reporting PSU with (RR= 1.14 95% CI 1.06, 1.23) and without (RR= 1.14 95% CI 1.05-1.24) alcohol. However, patients in intensive ambulatory settings for the general population only showed associations among people with alcohol use as a secondary substance (RR= 1.10 95% CI 1.07, 1.14). Still, people who reported PSU without alcohol showed positive associations but with confidence intervals that crossed the null (1.02 95% CI 0.99, 1.05). These associations also occurred for patients in basic ambulatory settings reporting PSU with alcohol (RR= 1.08 95% CI 1.05, 1.12) but did not occur in those without alcohol (1.01 95% CI 0.98-1.04). Interestingly, reporting PSU with alcohol as a secondary had a protective role among people in residential treatments for the general population (RR= 0.89 95% CI 0.83, 0.94).

BIBLIOGRAPHY

Carrero, J. J., Fu, E. L., Vestergaard, S. V., Jensen, S. K., Gasparini, A., Mahalingasivam, V., Bell, S., Birn, H., Heide-Jørgensen, U., Clase, C. M., Cleary, F., Coresh, J., Dekker, F. W., Gansevoort, R. T., Hemmelgarn, B. R., Jager, K. J., Jafar, T. H., Kovesdy, C. P., Sood, M. M., … Nitsch, D. (2023). Defining measures of kidney function in observational studies using routine health care data: Methodological and reporting considerations. *Kidney International*, *103*(1), 53–69. https://doi.org/10.1016/j.kint.2022.09.020

Cole, S. R., & Hernán, M. A. (2008). Constructing Inverse Probability Weights for Marginal Structural Models. *American Journal of Epidemiology*, *168*(6), 656–664. https://doi.org/10.1093/aje/kwn164

Keele, L. (2010). Proportionally Difficult: Testing for Nonproportional Hazards in Cox Models. *Political Analysis*, *18*(2), 189–205. https://doi.org/10.1093/pan/mpp044

Lokku, A., Lim, L. S., Birken, C. S., Pullenayegum, E. M., & on behalf of the TARGet Kids! Collaboration. (2020). Summarizing the extent of visit irregularity in longitudinal data. *BMC Medical Research Methodology*, *20*(1), 135. https://doi.org/10.1186/s12874-020-01023-w

Pan, W. (2001). Akaike’s Information Criterion in Generalized Estimating Equations. *Biometrics*, *57*(1), 120–125. https://doi.org/10.1111/j.0006-341X.2001.00120.x

Pullenayegum, E. (2022). *IrregLong: Analysis of Longitudinal Data with Irregular Observation Times* [Computer software]. https://cran.r-project.org/web/packages/IrregLong/index.html

Pullenayegum, E. M., Birken, C., Maguire, J., & TARGet Kids! Collaboration. (2023). Causal inference with longitudinal data subject to irregular assessment times. *Statistics in Medicine*, *42*(14), 2361–2393. https://doi.org/10.1002/sim.9727

Putter, H., Sasako, M., Hartgrink, H. H., van de Velde, C. J. H., & van Houwelingen, J. C. (2005). Long-term survival with non-proportional hazards: Results from the Dutch Gastric Cancer Trial. *Statistics in Medicine*, *24*(18), 2807–2821. https://doi.org/10.1002/sim.2143

Ruiz-Tagle Maturana, J., González-Santa Cruz, A., Rocha-Jiménez, T., & Castillo-Carniglia, Á. (2023). Does substance use disorder treatment completion reduce the risk of treatment readmission in Chile? *Drug and Alcohol Dependence*, *248*, 109907. https://doi.org/10.1016/j.drugalcdep.2023.109907

Zhang, Z., Reinikainen, J., Adeleke, K. A., Pieterse, M. E., & Groothuis-Oudshoorn, C. G. M. (2018). Time-varying covariates and coefficients in Cox regression models. *Annals of Translational Medicine*, *6*(7), 121–121. https://doi.org/10.21037/atm.2018.02.12