

## **Examining the causal genetic effects of substance use on aging**

Jared V. Balbona\*<sup>1</sup>, Paul Jeffries<sup>1</sup>, Aaron J. Gorelik<sup>2</sup>, Elliot C. Nelson<sup>1</sup>, Ryan Bogdan<sup>2</sup>,  
Arpana Agrawal<sup>1</sup>, Emma C. Johnson\*<sup>1</sup>

<sup>1</sup> Washington University School of Medicine in St. Louis, Dept. of Psychiatry

<sup>2</sup> Washington University in St. Louis, Department of Psychological & Brain Sciences

\*Corresponding Authors

Funding: R01DA054869, T32DA007261, K01DA051759

## **Abstract**

Substance use shortens lifespan, impedes health, and accelerates the biological aging process. We found widespread genetic correlations between alcohol, tobacco, cannabis, and opioid use and use disorders with indices of aging across the lifespan. There was evidence of tobacco and alcohol use and use disorders causally impacting physical, cognitive, and biological aging, with the effects of alcohol being more dependent on quantity of consumption; evidence of reverse causality was scant.

## Main Text

Chronic substance use and substance use disorders (SUDs)—psychiatric conditions marked by interpersonal, physiological, and psychological problems resulting from the loss of control over substance use—are among the leading causes of preventable death globally, placing a substantial burden on both individuals and society<sup>1</sup>. Despite this, not all individuals who use substances or have SUDs suffer premature mortality as a result. While many experience recovery and remission, others continue using substances maladaptively into later life, as evidenced by rising SUD rates among older adults<sup>2</sup>. As global life expectancy and substance availability have increased<sup>2</sup>, it is crucial to understand if and how substance involvement directly complicates biological aging and health in later life.

Psychoactive substances can directly and indirectly impact physical health, with evidence linking prolonged substance exposure to discrepancies between individuals' chronological and biologically estimated ages<sup>3</sup>. What is less well-understood, however, is the extent to which substance involvement *causally* modifies physical and biological aging (e.g., via neurotoxic effects) as opposed to both SUDs and accelerated aging resulting from shared predisposing factors (e.g., neighborhood disadvantage, medical stigma). Given the heritable nature of substance use and most aging metrics, genetically-informed study designs are well-suited to estimate the extent of shared genetic vulnerability to both domains and, after accounting for horizontal pleiotropy, to evaluate support for causal inference<sup>4</sup>.

In the present study, we assessed whether substance involvement (i.e., use and use disorders) is genetically correlated with physical, cognitive, and biological aging, and whether such associations reflect shared predisposition and/or causal influences of substance involvement. To this end, we calculated genetic correlations between 8 substance use and SUD phenotypes (drinks per week, problematic alcohol use, smoking initiation, tobacco use disorder, cannabis ever use, cannabis use disorder, opioid use disorder, and a multivariate SUD factor (mvSUD)) and 8 aging traits (Alzheimer Disease, frailty index, healthspan, parental lifespan, GrimAge, leukocyte telomere length, brain-age gap, and a multivariate aging factor (mvAge)) using Linkage Disequilibrium Score Regression<sup>5</sup> (LDSC; **Online Methods**). We then employed a series of Mendelian Randomization (MR) based approaches<sup>6</sup>, which utilize genetic variants as instrumental variables in order to estimate the causal effect of an “exposure” (e.g., SUD) on an “outcome” (e.g., aging; **Online Methods**). Briefly, MR leverages the random assortment of genes from parents to offspring, thereby ensuring that confounders are randomly distributed across the “treatment” and “control” groups (those who do and do not have a copy of the effect allele, respectively) in a manner analogous

to a randomized control trial. By reducing the potential for confounds in this way, MR-based approaches can disaggregate horizontal pleiotropy (the influence of a genetic locus on two separate, independent traits) from vertical pleiotropy (the direct causal effect of an exposure on an outcome; *Fig. 1*), allowing researchers to infer causality in situations where pleiotropy is expected and randomized control trials are not feasible.

**Genetic Correlations.** LDSC revealed widespread correlations between substance involvement and aging phenotypes, with genetic correlations ranging from  $r_g = -.50$  (tobacco use disorder and parental lifespan) to  $r_g = .55$  (smoking initiation and GrimAge), with a median absolute genetic correlation ( $|med_{r_g}|$ ) of .18 across all between-category trait pairs (*Fig. 2*). mvSUD was significantly associated with all aging indices ( $|med_{r_g}| = .36$ ) and mvAge was significantly associated with all substance involvement phenotypes, except cannabis ever use ( $|med_{r_g}| = .35$ ). Similarly, parental lifespan was significantly correlated with all substance use measures ( $|med_{r_g}| = .36$ ), while problematic alcohol use ( $|med_{r_g}| = .27$ ), smoking initiation ( $|med_{r_g}| = .36$ ), and tobacco use disorder ( $|med_{r_g}| = .36$ ) were each associated with 7 of 8 aging metrics. In contrast, despite cannabis use disorder showing moderate-to-high associations with most aging measures ( $|med_{r_g}| = .33$ ), cannabis ever use was only modestly correlated with parental lifespan ( $r_g = .10$ ) and no other aging index.

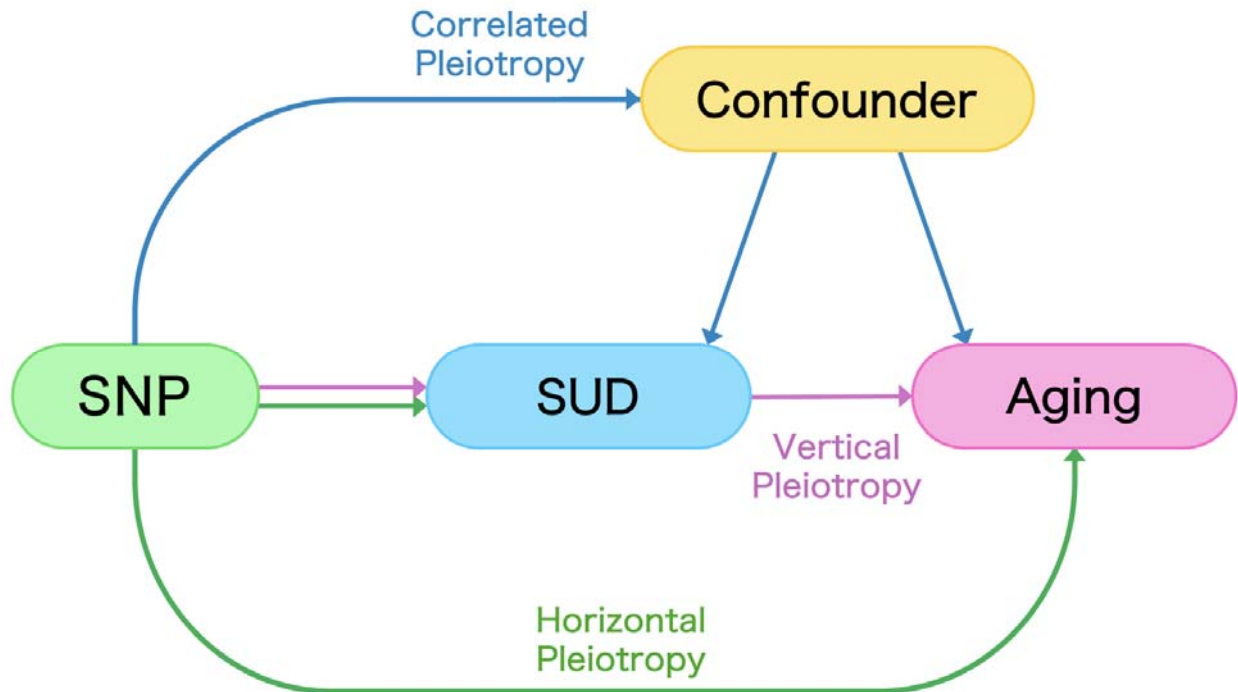
**Mendelian randomization.** For each of the significantly genetically correlated trait pairs, we evaluated evidence for genetic causation using Causal Analysis Using Summary Effect Estimates<sup>7</sup> (CAUSE) and Mendelian Randomization Pleiotropy RESidual Sum and Outlier<sup>8</sup> (MR-PRESSO), with the two approaches showing consistent agreement with one another despite their differing assumptions (*Fig. 3*). MR-PRESSO—which detects and removes potential pleiotropic outlier SNPs—found significant causal effects for smoking initiation and tobacco use disorder on 6 and 7 aging-related indices, respectively; problematic alcohol use and drinks per week, meanwhile, both showed causal effects on 3 aging metrics. All of these effects persisted after outlier SNPs were removed, indicating that they are likely due to vertical pleiotropy rather than outlier SNP-induced horizontal pleiotropy. Similarly, CAUSE results provided evidence that genetic liability for problematic alcohol use, smoking initiation, tobacco use disorder, and mvSUD may be causal for several aging outcomes, including frailty, GrimAge, and parental lifespan (*Supp Table 4*); across these traits, CAUSE found that a causal effect model fit the data significantly better than did a shared effects model (*Supp. Table 5*). Finally, we examined the effects of aging on substance use using both MR-PRESSO and CAUSE, finding little evidence of reverse causality from either approach.

Substance use and SUDs contribute to a majority of the leading causes of death and disability across the world (e.g., chronic obstructive pulmonary disease, cardiovascular disease, cancer)<sup>9</sup>. We found widespread evidence of genetic correlations between substance use/ SUDs and aging metrics, with MR results suggesting that both tobacco and alcohol use are causally implicated in shortening lifespan, deterring geriatric health, and modifying markers of biological youth. Notably, despite the causal effects of smoking initiation largely mirroring those of tobacco use disorder, we found that problematic alcohol use had a more significant impact on lifespan and mvAge than did drinks per week. These findings suggest that the effects of alcohol on aging may be quantity dependent, and thus highlight the importance of parsing the effects of quantitative substance exposure from disordered use to more fully understand their multifaceted impact on health. Furthermore, our scant evidence for reverse causality (i.e., aging impacting substance use) suggests that the reported increase<sup>2</sup> in SUD rates among older adults are not necessarily occurring in response to the physical and mental challenges of aging, as has been previously theorized<sup>10</sup>.

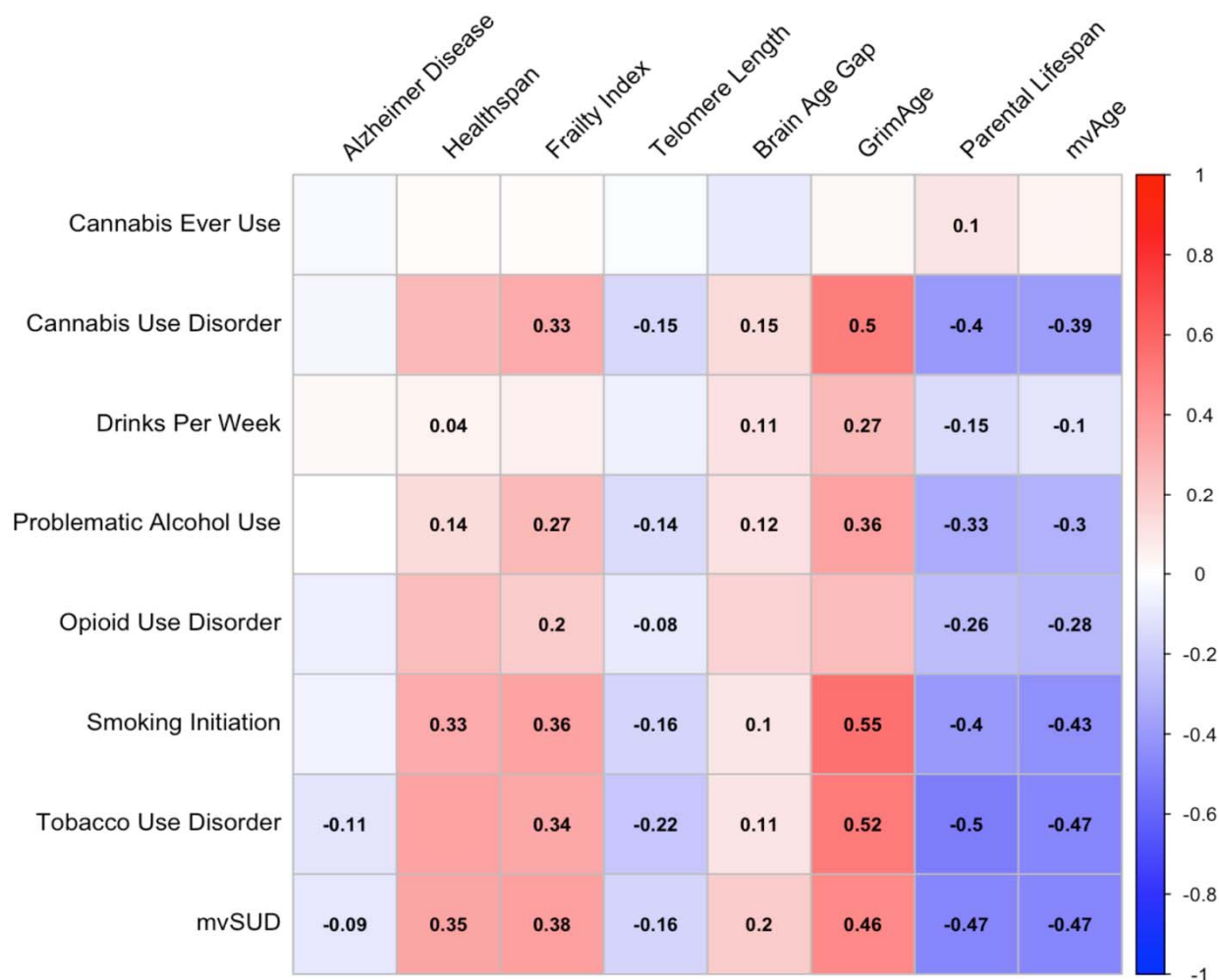
While our null MR findings (e.g., for cannabis, opioids, and some aging metrics) suggest that the observed genetic correlations between these SUD exposures and aging outcomes represent either common genetic mechanisms or third variable confounding, some caveats remain noteworthy. First, we cannot exclude the possibility that some null findings—particularly those involving cannabis use and Alzheimer Disease—were due in part to smaller GWAS sample sizes and a consequent lack of power to detect effects. Relatedly, we cannot dismiss the potential consequences of lower rates of recent SUDs in older individuals with health conditions (the “sick quitter” phenomenon<sup>11</sup>), survivorship/ selection bias, unmodeled interactive effects, or non-representative GWAS samples. Additionally, all aging-related summary statistics were derived from GWAS that only included individuals of European genetic ancestry, limiting the generalizability of these findings to other ancestries (**Online Methods**).

Stigma associated with the acknowledgement of SUDs and the consequent reluctance to query older adults about their substance use hinder effective and tailored interventions in this age group<sup>12</sup>. Our analyses suggest that prevention of substance use may reduce markers of biological aging and promote geriatric well-being and longevity. Additionally, these results underscore the importance of integrating genetic insights into public health strategies to effectively reduce the burden of SUDs across the lifespan.

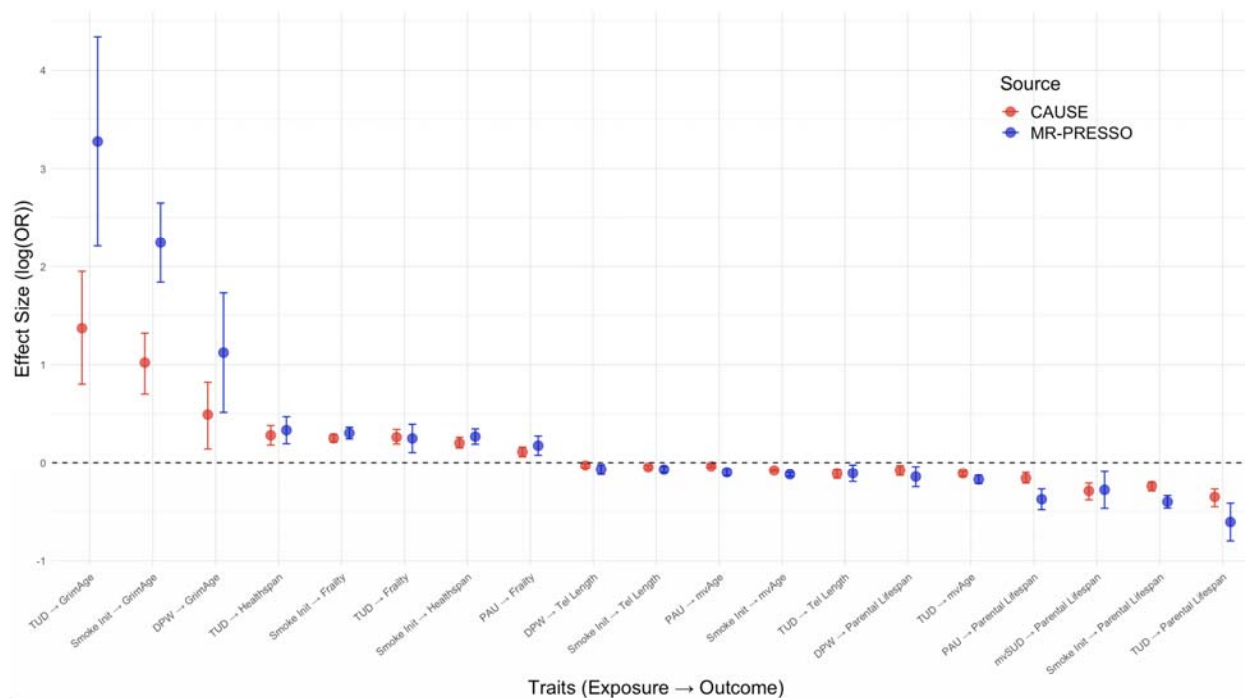
## FIGURES



**Figure 1.** Vertical pleiotropy occurs when a single nucleotide polymorphism (SNP) influences multiple traits through a direct causal pathway, where the SNP affects one trait that in turn affects another. In contrast, horizontal pleiotropy happens when a single SNP independently influences multiple traits through separate and distinct mechanisms, without one trait affecting the other. Correlated pleiotropy involves a situation where a single gene affects multiple traits that are correlated, often because the traits share common genetic or environmental influences.



**Figure 2.** Genome-wide genetic correlations between trait pairs, calculated using LDSC. Coefficients are provided for correlations reaching FDR-adjusted significance (all correlations are available in Supp Table 2). As shown, correlations were widespread, particularly for smoking and alcohol phenotypes, parental lifespan, mvSUD, and mvAge.



**Figure 3.** CAUSE and MR-PRESSO effect size estimates (log(OR)) for significant trait pairs, with significance being based upon an FDR-corrected  $p$ -value threshold. Results from both approaches are largely in agreement with one another, with the error bars (representing 95% confidence intervals) overlapping for nearly all trait pairs.



## ONLINE METHODS

### Data Sources and Phenotypes

Genetic association estimates were derived using publicly available summary statistics based on individuals of European ancestry (see *Supp. Table 1 for full descriptions for each trait/cohort*). In total, eight substance use traits (problematic alcohol use, drinks per week, smoking initiation, cannabis ever use, cannabis use disorder, opioid use disorder, tobacco use disorder, and a multivariate substance use factor (mvSUD)) and eight aging traits (telomere length, Alzheimer Disease, healthspan, brain age gap, parental lifespan, frailty index, GrimAge, and a multivariate aging factor (mvAge)) were analyzed.

### Study Design

All pairwise combinations of substance use and aging traits were first analyzed using LD score regression. Subsequently, causality between all significantly correlated pairs of traits were analyzed using both MR-PRESSO and CAUSE.

### Genetic Instrument Selection and QC

All included SNPs were non-palindromic, autosomal, and genome-wide significant ( $p \leq 5 \times 10^{-8}$ ). To exclude variants adjacent to the variants of interest, SNPs were clumped based on their linkage disequilibrium ( $r^2 = .01$ ;  $p < 5 \times 10^{-3}$ ) within a genomic window of 10,000 kb, with the 1000 Genomes European LD estimates serving as our reference panel. Additionally, each non-effect allele was aligned to the human genome reference sequence (build 37) to ensure that the effect of a given SNP is corresponding to the same allele for both the exposure and the outcome. Finally, the MR-PRESSO outlier test was used to exclude any potential outlier variants from the MR-PRESSO analyses.

### Linkage Disequilibrium Score (LDSC) Regression

LDSC regression was utilized to examine the genetic correlations between all pairs of traits, both within-category (i.e., aging-aging, SUD-SUD) and between-category (SUD-aging). Genetic correlation estimates from LDSC regression are not biased by sample overlap, but rather are able to estimate and account for the degree of overlap; as such, it is able to theoretically estimate the correlation between traits that is due to similarities in genetic architecture, rather than sampling artifacts. Because several of our traits are non-independent of one another (e.g., problematic alcohol use and drinks per week), an FDR-corrected  $p$ -value was employed to reduce the risk of type I errors.

### Mendelian Randomization Methods

A primary assumption of traditional MR-based approaches is that the genetic instruments should influence the outcome exclusively through the exposure of interest (i.e., the instrument exclusion restriction). However, this assumption is commonly violated due to the presence of horizontal pleiotropy, both when that the pleiotropy is correlated (i.e., when traits share a genetically influenced causal pathway) and when it is uncorrelated (when the genetic factor influences the exposure and outcome via distinct mechanisms). Therefore, the present study

implements a two-pronged MR approach to assessing the causal impact of substance use on aging.

First, MR-PRESSO was conducted using the MRPRESSO R package version 1.0 (<https://github.com/rondolab/MR-PRESSO>). Here, we present findings from the MR-PRESSO raw tests, except in cases in which the MR-PRESSO distortion test (indicating a significant influence of outlier variants) is FDR significant; in such cases, the MR-PRESSO outlier-corrected values are given instead. Additionally, we performed CAUSE using the R package cause version 1.2.0 (<https://github.com/jean997/cause>). CAUSE assumes that genetic relationships between traits are a mixture of both vertical and correlated pleiotropy, and thus evaluates the relative contributions of both. MR-PRESSO, meanwhile, detects and removes SNPs with evidence of horizontal pleiotropy, thus minimizing bias in its estimate of SNP-based vertical pleiotropy.

## References

1. GBD 2016 Alcohol and Drug Use Collaborators. The global burden of disease attributable to alcohol and drug use in 195 countries and territories, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet Psychiatry* **5**, 987–1012 (2018).
2. Weber, A., Lynch, A., Miskle, B., Arndt, S. & Acion, L. Older Adult Substance Use Treatment First-Time Admissions Between 2008 and 2018. *Am. J. Geriatr. Psychiatry* **30**, 1055–1063 (2022).
3. Bachi, K., Sierra, S., Volkow, N. D., Goldstein, R. Z. & Alia-Klein, N. Is biological aging accelerated in drug addiction? *Curr. Opin. Behav. Sci.* **13**, 34–39 (2017).
4. Bogdan, R., Hatoum, A. S., Johnson, E. C. & Agrawal, A. The Genetically Informed Neurobiology of Addiction (GINA) model. *Nat. Rev. Neurosci.* **24**, 40–57 (2023).
5. Bulik-Sullivan, B. K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* **47**, 291–295 (2015).
6. Lawlor, D. A., Harbord, R. M., Sterne, J. A. C., Timpson, N. & Davey Smith, G. Mendelian randomization: Using genes as instruments for making causal inferences in epidemiology. *Stat. Med.* **27**, 1133–1163 (2008).
7. Morrison, J., Knoblauch, N., Marcus, J. H., Stephens, M. & He, X. Mendelian randomization accounting for correlated and uncorrelated pleiotropic effects using genome-wide summary statistics. *Nat. Genet.* **52**, 740–747 (2020).
8. Verbanck, M., Chen, C.-Y., Neale, B. & Do, R. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. *Nat. Genet.* **50**, 693–698 (2018).
9. WHO. *World Health Statistics 2016 [OP]: Monitoring Health for the Sustainable Development Goals (SDGs)*. (World Health Organization, 2016).
10. Kuerbis, A., Sacco, P., Blazer, D. G. & Moore, A. A. Substance Abuse Among Older Adults. *Clin. Geriatr. Med.* **30**, 629–654 (2014).

11. Sarich, P. *et al.* A Prospective Study of Health Conditions Related to Alcohol Consumption Cessation Among 97,852 Drinkers Aged 45 and Over in Australia. *Alcohol. Clin. Exp. Res.* **43**, 710–721 (2019).
12. Chhatre, S., Cook, R., Mallik, E. & Jayadevappa, R. Trends in substance use admissions among older adults. *BMC Health Serv. Res.* **17**, 584 (2017).