Correct attribution of cause and effect is critical for prevention and treatment efforts. While randomized experiments or controlled trials are suitable for certain exposures, there are practical and ethical difficulties with the study of substance use and abuse in natural settings. However, there exist several promising methods for extracting causal inferences from non-experimental data. We consider methods that require: i) data collected from relatives, ii) genomic markers; iii) repeated measures, and iv) combinations of these approaches - many of which have yet to be developed or tested. The ABCD dataset of 11,874 youth aged 9-10 provides over 60,000 variables to which these methods could be applied. While substance use itself is, by design, at negligible levels in this young cohort, illustrative analyses of neuroimaging, behavioral and psychological correlates of substance use can already be performed. Example results of analyses that are already feasible are presented, and the potential for further, comprehensive investigations of substance use and its risk factors are discussed.