**Algorithm for the Labeling High-risk Participants**

September 20, y

Kevin Cummins

**Introduction**

The NCANDA project oversamples high-risk participants. The screener provides a brief assessment of the high-risk characteristics of participants. The clinical interview provides a more detailed assessment. SRI will be writing code to compute the risk status of each subject after their initial interview. This document provides an outline of the algorithm to be used.

High-risk individuals have externalizing or internalizing symptoms, early onset of drinking, or a family history of alcohol use disorders.

Some variables come from the blaise implementation of the ssaga and others come from the limesurvey interview data.

**Algorithm**

|  |  |
| --- | --- |
| Pseudo-code | Description |
| 1. Initialize a ‘*highrisk*’ variable  Let 1 indicate yes, high-risk  Let 0 indicate no, not high-risk  2. If any of these limesurvey variables equals 1 then code the person as high-risk(1):  *yfhi3a\_yfhi3a*  *pfhi3a\_pfhi3a*  *yfhi3a\_yfhi3f*  *pfhi3a\_pfhi3f*  *yfhi4a\_yfhi4a*  *pfhi4a\_pfhi4a*  *yfhi4a\_yfhi4f*  *pfhi4a\_pfhi4f*  3. Let *gparhx*=0 and add 1 to gparhx each time one of these expressions are met:   * *yfhi3a\_yfhi3b* or *pfhi3a\_pfhi3b* or *yfhi4a\_yfhi4b* or *pfhi4a\_pfhi4b* ==1 * *yfhi3a\_yfhi3c* or *pfhi3a\_pfhi3c* or *yfhi4a\_yfhi4c* or *pfhi4a\_pfhi4c* ==1 * *yfhi3a\_yfhi3g* or *pfhi3a\_pfhi3g* or *yfhi4a\_yfhi4g* or *pfhi4a\_pfhi4g* ==1 * *yfhi3a\_yfhi3h* or *pfhi3a\_pfhi3h* or *yfhi4a\_yfhi4h* or *pfhi4a\_pfhi4h* ==1   4. Code the person as high-risk(1) if *gparhx*>1  5. Code the person as high-risk (1) if:  *AL1AgeOns* =< 14  6. Initialize a ‘*extern*’ variable  Let 1 indicate yes  Let 0 indicate no  7. Recode the following variables such that 1=12, 2=14, 3=17, 4=AGE:  ASa\_ao2DK ASb\_ao2DK ASc1\_ao6DK ASc2\_ao6DK AS\_ao9DK AS\_ao10DK  AS1\_ao11DK  AS2\_ao11DK  AS1\_ao15DK  AS2\_ao15DK  AS1\_ao16DK  AS2\_ao16DK  ASa1\_ao14DK  ASa2\_ao14DK  ASc1\_ao14DK  ASc2\_ao14DK  AS1\_ao17DK  AS2\_ao17DK  AS1\_ao19DK  AS2\_ao19DK  AS1\_ao18DK  AS2\_ao18DK  AS1\_ao20DK  AS2\_ao20DK  8. When any of these expressions are true add 1 to “extern” variable:   * AS2A=5 & (ASa\_ao2 < AL1AgeOns or ASa\_ao2DK < AL1AgeOns) * AS2B=5 & [(ASb\_ao2 < AL1AgeOns or (ASb\_ao2DK < AL1AgeOns)] * AS6b > 1 & [(ASc1\_ao6 < AL1AgeOns) or (ASc2\_ao6 < AL1AgeOns) or (ASc1\_ao6DK < AL1AgeOns) or (ASc2\_ao6DK < AL1AgeOns)] * AS9=5 & [(AS\_ao9 < AL1AgeOns) or (AS\_ao9DK < AL1AgeOns)] * AS10a > 1 & [(AS\_ao10 < AL1AgeOns) or (AS\_ao10DK < AL1AgeOns)] * AS11 > 1 & [(AS1\_ao11 < AL1AgeOns) or (AS2\_ao11 < AL1AgeOns) or (AS1\_ao11DK < AL1AgeOns) or (AS2\_ao11DK < AL1AgeOns)] * AS15 > 1 & [(AS1\_ao15 < AL1AgeOns) or (AS2\_ao15 < AL1AgeOns) or (AS1\_ao15DK < AL1AgeOns) or (AS2\_ao15DK < AL1AgeOns)] * AS16 > 1 & [(AS1\_ao16 < AL1AgeOns) or (AS2\_ao16 < AL1AgeOns) or (AS1\_ao16DK < AL1AgeOns) or (AS2\_ao16DK < AL1AgeOns)] * AS14 > 1 & [(ASa1\_ao14 < AL1AgeOns) or (ASa2\_ao14 < AL1AgeOns) or (ASa1\_ao14DK < AL1AgeOns) or (ASa2\_ao14DK < AL1AgeOns)]   and  AS14b > 1 & [(ASc1\_ao14 < AL1AgeOns) or (ASc2\_ao14 < AL1AgeOns) or (ASc1\_ao14DK < AL1AgeOns) or (ASc2\_ao14DK < AL1AgeOns)]   * AS17a = 5 & [(AS1\_ao17 < AL1AgeOns) or (AS2\_ao17 < AL1AgeOns) or (AS1\_ao17DK < AL1AgeOns) or (AS2\_ao17DK < AL1AgeOns)] * AS19 > 1 & [(AS1\_ao19 < AL1AgeOns) or (AS2\_ao19 < AL1AgeOns) or (AS1\_ao19DK < AL1AgeOns) or (AS2\_ao19DK < AL1AgeOns)] * AS18b = 5 & [(AS1\_ao18 < AL1AgeOns) or (AS2\_ao18 < AL1AgeOns) or (AS1\_ao18DK < AL1AgeOns) or (AS2\_ao18DK < AL1AgeOns)] * AS20 > 1 & [(AS1\_ao20 < AL1AgeOns) or (AS2\_ao20 < AL1AgeOns) or (AS1\_ao20DK < AL1AgeOns) or (AS2\_ao20DK < AL1AgeOns)]   9. If “extern” > 2 then code highrisk=1  10. Initialize an ‘Intern’ variable  Let 1 indicate yes  Let 0 indicate no  11. When any of these expressions are met add 1 to “intern” variable:   * OC1=5 & (OC\_AO8 < AL1AgeOns) * OC9=5 & (OC\_ao16 < AL1AgeOns) * PN1x=5 or PN2a=5 or PN2b=5 or PN5 > 2 & [(PN\_ao8 < AL1AgeOns) or (PN\_ao8DK=1 & AL1AgeOns > 10) or (PN\_ao8DK=2 & AL1AgeOns > 20)] * DP4a=5 or DP4b=5 & [(DP3 < AL1AgeOns) or (DP3\_1=1 & AL1AgeOns > 10) or (DP3\_1=2 & AL1AgeOns > 20)] * DP11=5 or DP12=5 & [(DP3 < AL1AgeOns) or (DP3\_1=1 & AL1AgeOns > 10) or (DP3\_1=2 & AL1AgeOns > 20)] * DP15a=5 or DP15b=5 or DP15c=5 or DP15d=5 & [(DP3 < AL1AgeOns) or (DP3\_1=1 & AL1AgeOns > 10) or (DP3\_1=2 & AL1AgeOns > 20)]   12. If “intern” > 2 then code highrisk=1 | Initialize the high-risk variable  If either of the bio parents has a substance use disorder history the person is high-risk, per the youth or parent report.  Count the number of grandparents with a history of substance use disorders.  Code as high-risk if more than two grandparents have a history of substance use disorders, per the youth or parent report.  Code as high-risk (1) if there was early onset (<15) of drinking.  Initialize the Externalizing variable  Recode the “don’t know” timings so that it is easy to evaluate against age of drinking onset. AGE is the age at interview time.  Count the number of behaviors that occurred prior to drinking onset  If > 2 externalizing symptoms occurred prior to drinking then code as high risk  Initialize the Internalizing variable  Count the number of sections that occurred prior to drinking onset  If > 2 intern symptoms prior to drinking then code as high risk |