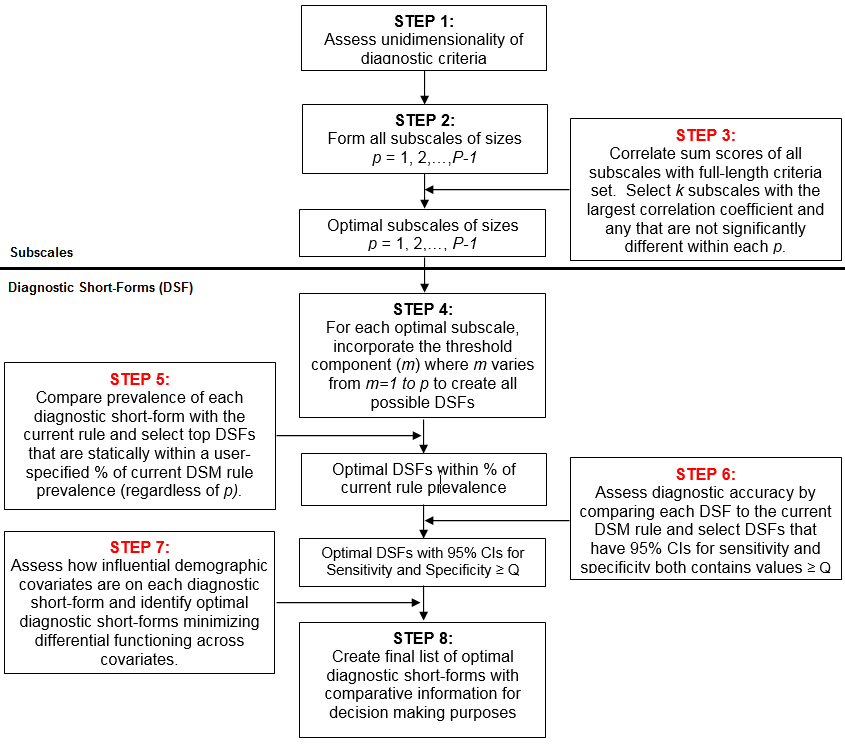
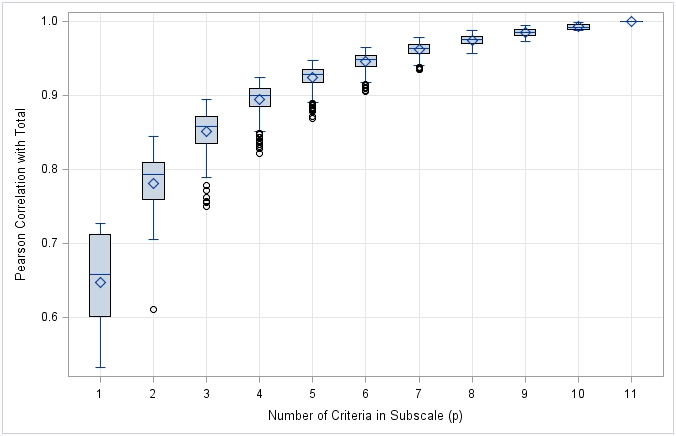
**Appendices**

**Appendix A:** Additional Tables and Figures

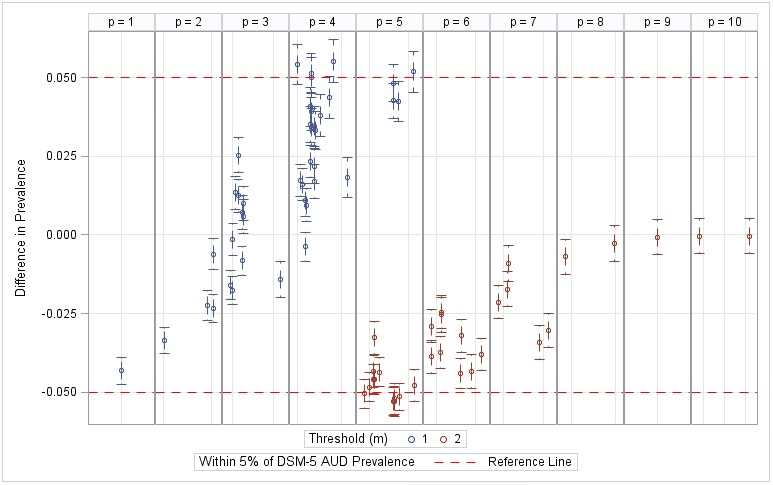
**Figure A.1.** Identifying Optimal Diagnostic Short-Forms: Application to DSM-5 AUD



**Figure A.2.** Pearson Correlations by for all 2,046 Subscales – Step 3

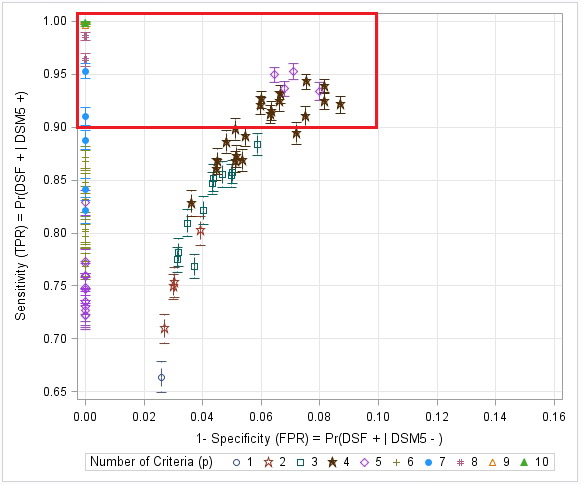


**Figure A.3.** Difference in Prevalence between DSM-5 AUD and Diagnostic Short-Forms\* (values greater than zero indicate DSF prevalence is higher than DSM-5 AUD prevalence, values less than zero indicate DSF prevalence is lower than DSM-5 AUD) – Step 5

******

*\* Only 72 diagnostic short-forms that fall within 5% of DSM-5 AUD prevalence (out of the 401 considered) are shown.*

**Figure A.4.** Sensitivity vs. 1-Specificity for 72 Diagnostic Short-Forms – Step 6

******

*\* Red line highlights the 25 diagnostic short-forms that are > 0.90 on both sensitivity and specificity.*

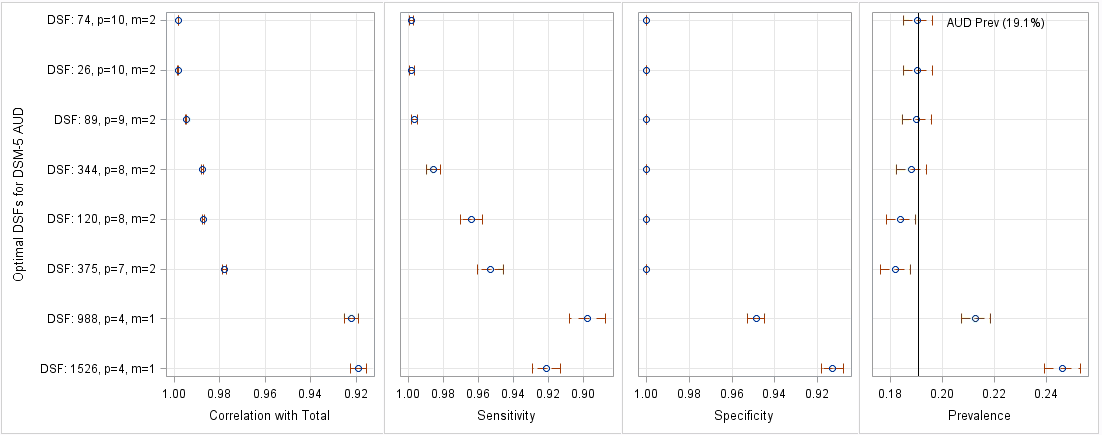
**Table A.1.** Diagnostic Short-forms and Differential Test Functioning – Step 7

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **DSF Labelab** |  |  | **Old vs. Young** | | **Male vs. Female** | | **Non-White vs. White** | | **Any Mood and/or Anxiety Disorder vs. None** | |
| **Est. (SE)** | **p-value** | **Est. (SE)** | **p-value** | **Est. (SE)** | **p-value** | **Est. (SE)** | **p-value** |
| 510 | 4 | 1 | -0.090 (0.079) | 0.257 | -0.251 (0.076) | **0.001** | 0.096 (0.082) | 0.245 | 0.007 (0.082) | 0.930 |
| 860 | 4 | 1 | -0.014 (0.061) | 0.822 | -0.099 (0.056) | 0.077 | -0.211 (0.059) | **<0.001** | -0.290 (0.066) | **<0.001** |
| 864 | 4 | 1 | -0.157 (0.068) | **0.021** | -0.023 (0.066) | 0.729 | -0.039 (0.066) | 0.552 | -0.327 (0.079) | **<0.001** |
| 866 | 4 | 1 | -0.213 (0.067) | **0.001** | -0.170 (0.068) | **0.012** | -0.235 (0.068) | **<0.001** | -0.375 (0.080) | **<0.001** |
| 884 | 4 | 1 | 0.031 (0.078) | 0.687 | -0.027 (0.066) | 0.681 | -0.224 (0.071) | **0.002** | -0.168 (0.074) | **0.024** |
| 892 | 4 | 1 | -0.092 (0.069) | 0.184 | -0.140 (0.067) | **0.038** | 0.051 (0.067) | 0.448 | -0.205 (0.077) | **0.008** |
| 894 | 4 | 1 | -0.179 (0.067) | **0.007** | -0.305 (0.068) | **<0.001** | -0.183 (0.068) | **0.007** | -0.312 (0.080) | **<0.001** |
| 896 | 4 | 1 | -0.274 (0.077) | **<0.001** | -0.122 (0.082) | 0.139 | -0.088 (0.075) | 0.243 | -0.335 (0.092) | **<0.001** |
| 988 | 4 | 1 | -0.023 (0.061) | 0.703 | 0.018 (0.056) | 0.753 | 0.068 (0.060) | 0.255 | -0.078 (0.061) | 0.198 |
| 992 | 4 | 1 | -0.309 (0.066) | **<0.001** | -0.011 (0.066) | 0.863 | -0.028 (0.065) | 0.662 | -0.141 (0.072) | 0.052 |
| 1020 | 4 | 1 | -0.256 (0.067) | **<0.001** | -0.156 (0.069) | **0.024** | 0.031 (0.067) | 0.642 | -0.114 (0.072) | 0.117 |
| 1154 | 4 | 1 | -0.056 (0.063) | 0.371 | -0.045 (0.061) | 0.462 | 0.136 (0.064) | **0.034** | -0.329 (0.069) | **<0.001** |
| 1406 | 4 | 1 | -0.081 (0.066) | 0.222 | -0.101 (0.062) | 0.107 | 0.071 (0.066) | 0.280 | -0.392 (0.073) | **<0.001** |
| 1526 | 4 | 1 | 0.025 (0.074) | 0.730 | -0.033 (0.065) | 0.618 | 0.031 (0.069) | 0.651 | -0.063 (0.070) | 0.371 |
| 861 | 5 | 1 | -0.065 (0.082) | 0.425 | -0.081 (0.079) | 0.306 | -0.028 (0.079) | 0.728 | -0.313 (0.097) | **0.001** |
| 863 | 5 | 1 | -0.199 (0.077) | **0.010** | -0.299 (0.078) | **<0.001** | -0.259 (0.078) | **0.001** | -0.348 (0.095) | **<0.001** |
| 989 | 5 | 1 | -0.272 (0.079) | **0.001** | -0.105 (0.078) | 0.179 | -0.007 (0.077) | 0.929 | -0.058 (0.085) | 0.495 |
| 1375 | 5 | 1 | -0.157 (0.073) | **0.032** | -0.026 (0.070) | 0.716 | 0.092 (0.076) | 0.225 | -0.406 (0.083) | **<0.001** |
| 345 | 7 | 2 | -0.069 (0.068) | 0.315 | -0.095 (0.061) | 0.120 | -0.232 (0.065) | **<0.001** | -0.163 (0.070) | **0.019** |
| 375 | 7 | 2 | -0.029 (0.093) | 0.755 | -0.007 (0.089) | 0.937 | -0.008 (0.088) | 0.928 | -0.150 (0.096) | 0.120 |
| 120 | 8 | 2 | -0.117 (0.106) | 0.272 | -0.073 (0.107) | 0.495 | 0.057 (0.106) | 0.588 | -0.089 (0.110) | 0.417 |

*a DSF label is the identifier given by the algorithm when all original subscales are enumerated and does not encode any specific information*

*b Four rules were bypassed from assessment due to very high sensitivity (≥ 97%) and specificity (100%) estimates (DSF 74: p=10, m=2, DSF 26: p=10, m=2, DSF 89: p=9, m=2, DSF 344: p=8, m=2)*

**Figure A.5.** Descriptive Statistics (estimates with 95% CIs) on Optimal Diagnostic Short-Forms for DSM-5 AUD – Step 8

****

*The number after each DSF is a label identifier given by the algorithm when all original subscales are enumerated and does not encode any specific information. P=# of criteria, M=threshold*

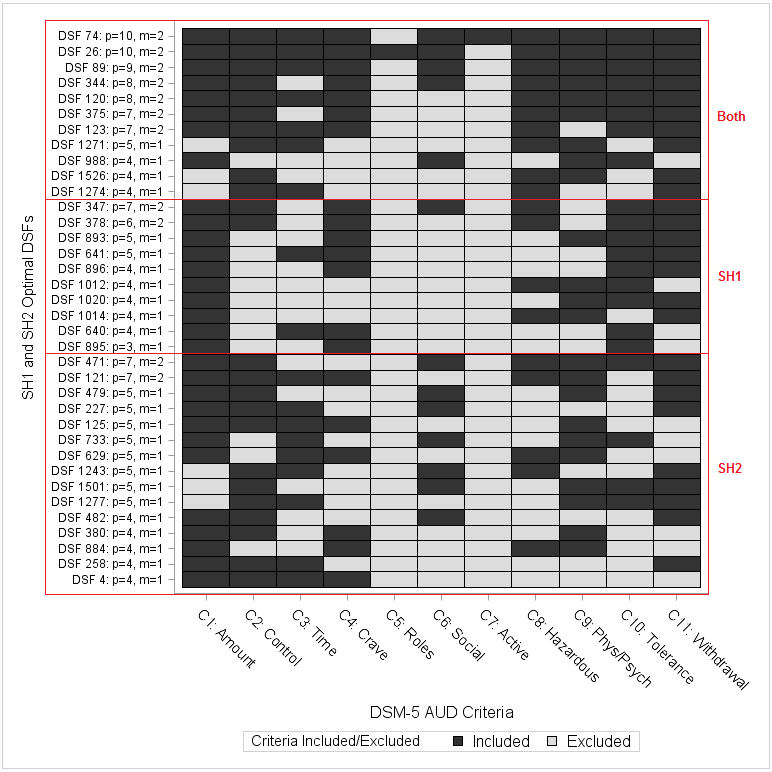
**Table A.2.** Version A: DSM-5 AUD Severity vs. *DSF 89* AUD Severity in NESARC-III

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **DSM-5 AUD Severity**  **(11 criteria)** |  | **DSF 89 (9 criteria)** | | | | |
| **None**  **(0-1)** | **Mild**  **(2-3)** | **Moderate**  **(4)** | **Severe**  **(5+)** | **Total** |
| **None**  **(0-1)** | **20,645**  **(100.0%)** | 0 | 0 | 0 | 20,645  (80.9%) |
| **Mild**  **(2-3)** | 18  (0.7%) | **2,606**  **(99.3%)** | 0 | 0 | 2,624  (10.0%) |
| **Moderate**  **(4-5)** | 0 | 39  (2.3%) | **704**  **(59.1%)** | 443  (38.5%) | 1,186  (4.4%) |
| **Severe**  **(6+)** | 0 | 0 | 4  (0.2%) | **1,319**  **(99.8%)** | 1,323  (4.6%) |
| **Total** | 20,663  (81.0%) | 2,645  (10.0%) | 708  (2.6%) | 1,762  (6.3%) | 25,778 |

**Table A.3.** Distribution of DSM-5 AUD Criteria and Selected Covariates in NESARC-III Split half 1, NESARC-III Split half 2 and Wave 2 NESARC

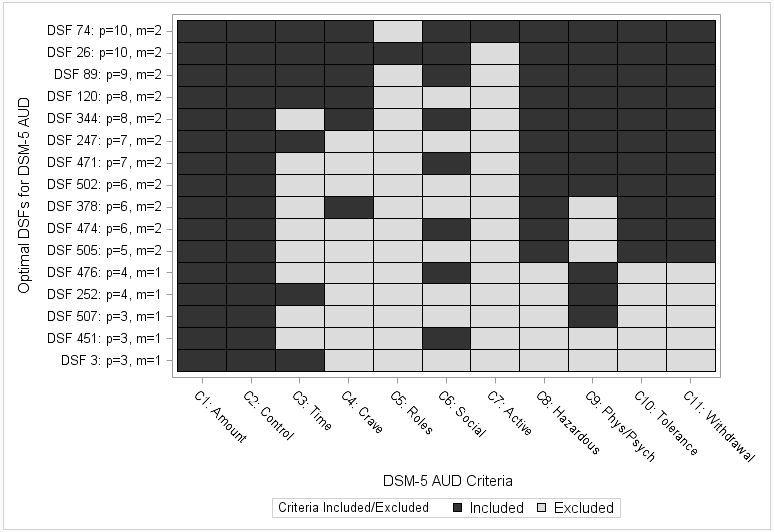
|  |  |  |  |
| --- | --- | --- | --- |
| **Criterion / Covariate** | **Split half 1**  **(n=12,889)** | **Split half 2**  **(n=12,889)** | **Wave 2 NESARC**  **(n=22,177)** |
| C1: Used larger amounts / longer | 15.1% | 14.3% | 14.6% |
| C2: Repeated attempts to quit/control | 13.2% | 12.7% | 12.4% |
| C3: Much time spent using | 4.8% | 5.1% | 3.0% |
| C4: Craving | 10.9% | 10.5% | 4.1% |
| C5: Neglected major roles to use | 1.7% | 2.0% | 1.1% |
| C6: Social/interpersonal problems | 6.3% | 6.3% | 2.5% |
| C7: Activities given up to use | 1.7% | 1.8% | 1.0% |
| C8: Hazardous use | 11.6% | 11.4% | 11.1% |
| C9: Physical/psychological problems | 7.0% | 7.3% | 5.3% |
| C10: Tolerance | 9.6% | 9.5% | 8.2% |
| C11: Withdrawal | 10.5% | 10.7% | 7.9% |
|  |  |  |  |
| Young (<= 44 years old) | 52.4% | 51.6% | 52.2% |
| Male | 51.1% | 50.3% | 52.1% |
| White | 81.0% | 80.6% | 75.0% |
| Any mood and/or anxiety disorders | 22.8% | 21.8% | 17.9% |

**Figure A.6.** Criteria Included/Excluded in Optimal DSFs in both Split Halves\*



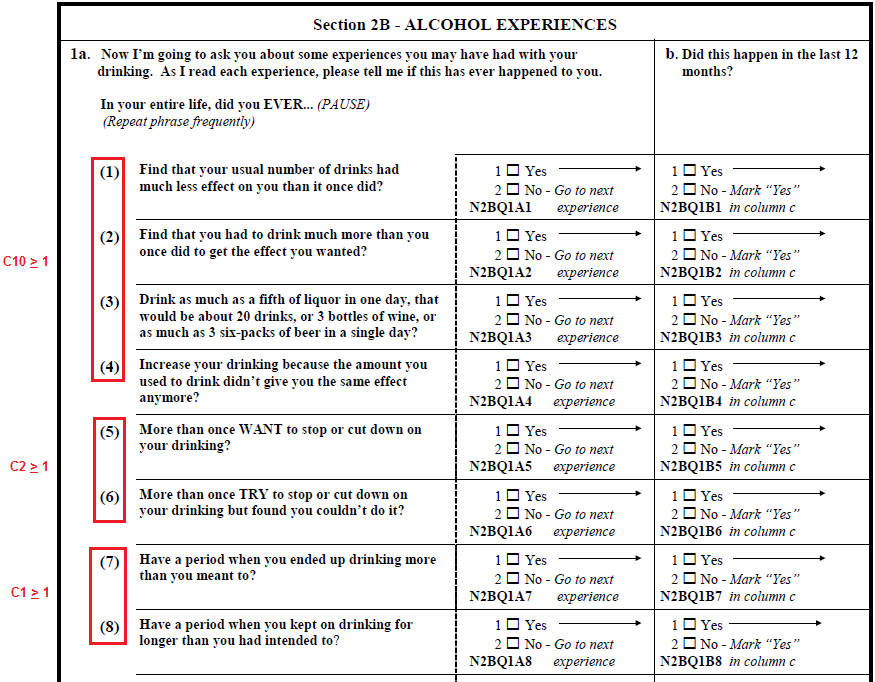
*\* DSFs ordered by p and combined Sensitivity, Specificity within split-half groups (Both, SH1, SH2)*

**Figure A.7.** Criteria Included/Excluded in Optimal Diagnostic Short-forms for DSM-5 AUD using Wave 2 NESARC

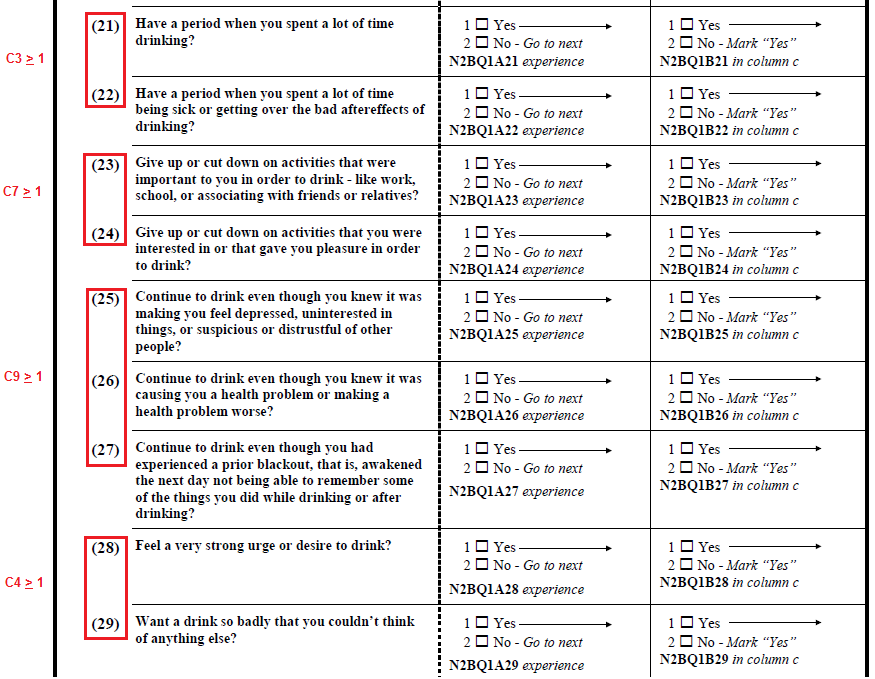


**Appendix B:** Survey questions from NESARC-III used to elicit each criterion

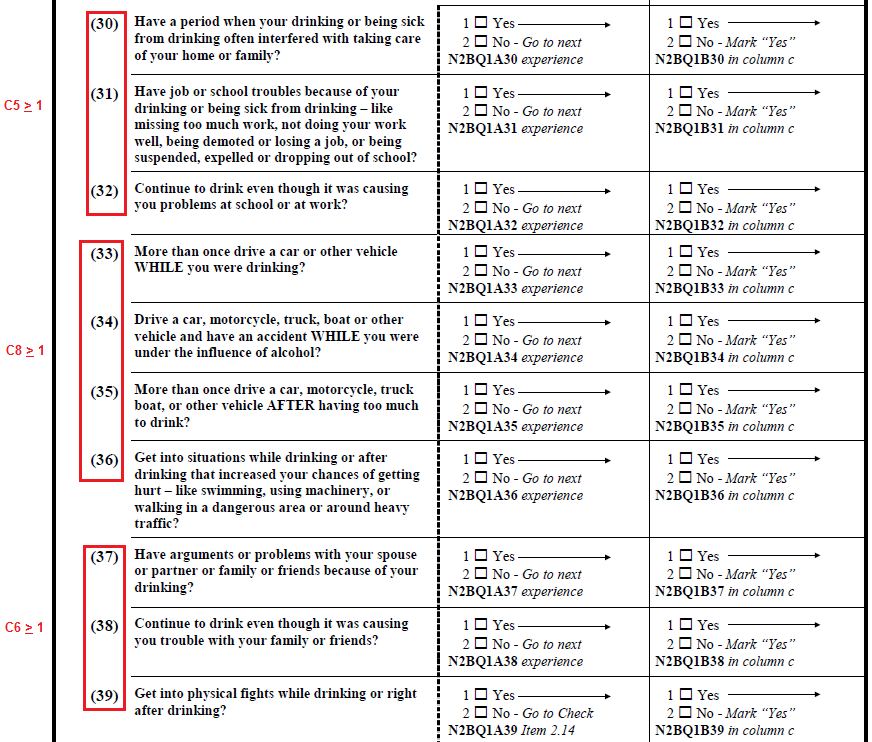
**Figure B.1** Survey Questions used to Elicit C1, C2 and C10



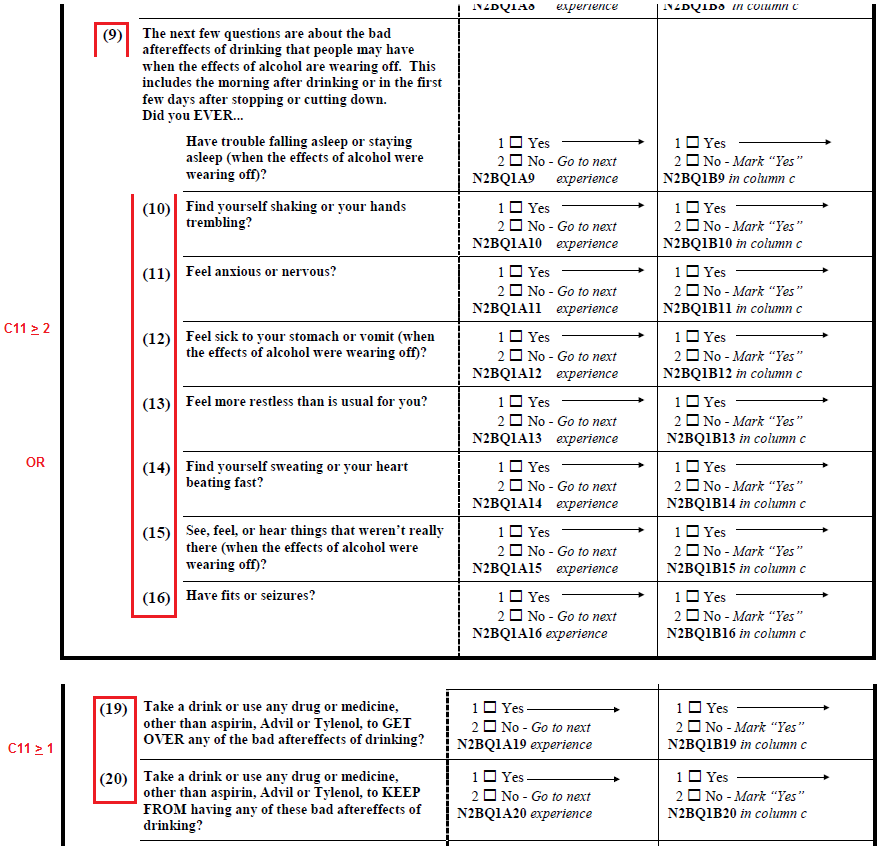
**Figure B.2** Survey Questions used to Elicit C3, C4, C7 and C9



**Figure B.3** Survey Questions used to Elicit C5, C6 and C8

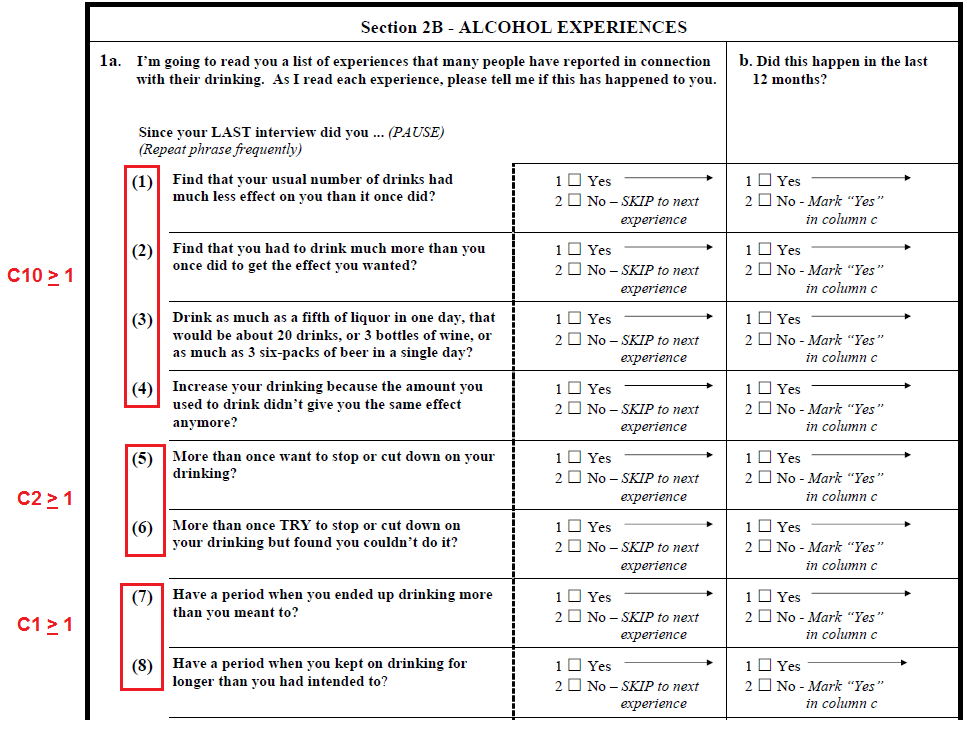


**Figure B.4** Survey Questions used to Elicit C11

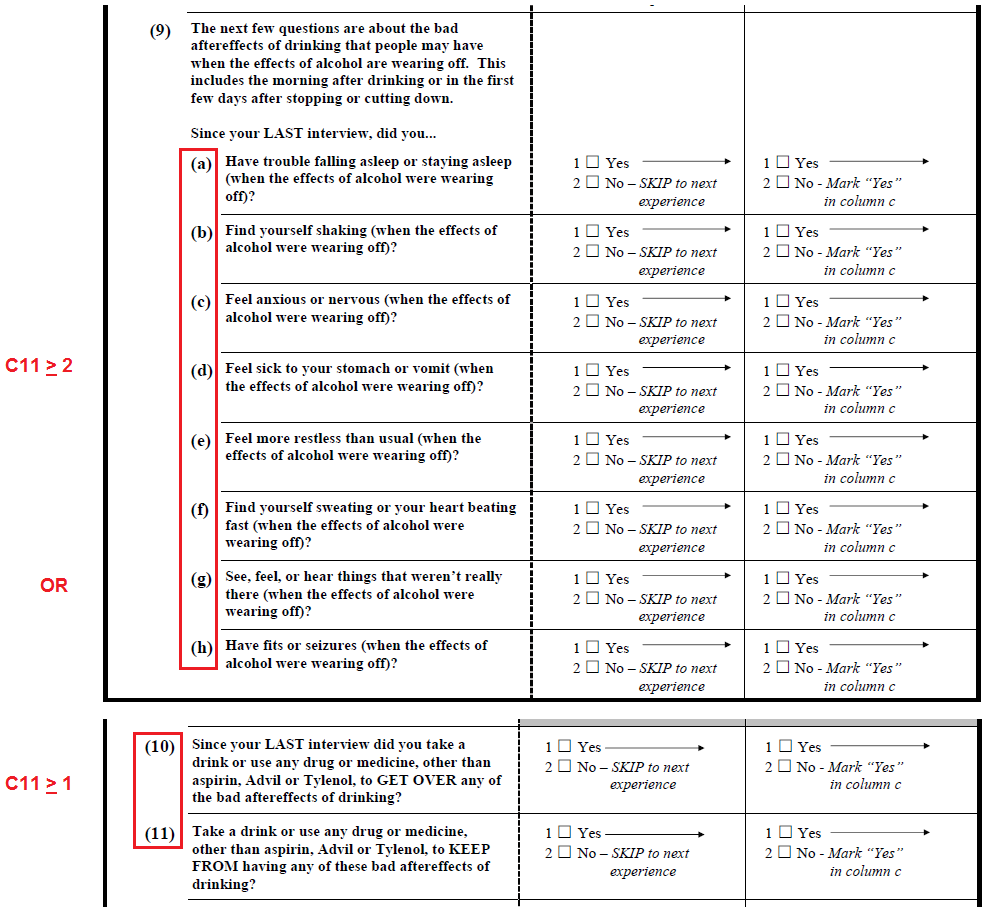


**Appendix C:** Survey questions from Wave 2 NESARC used to elicit each criterion

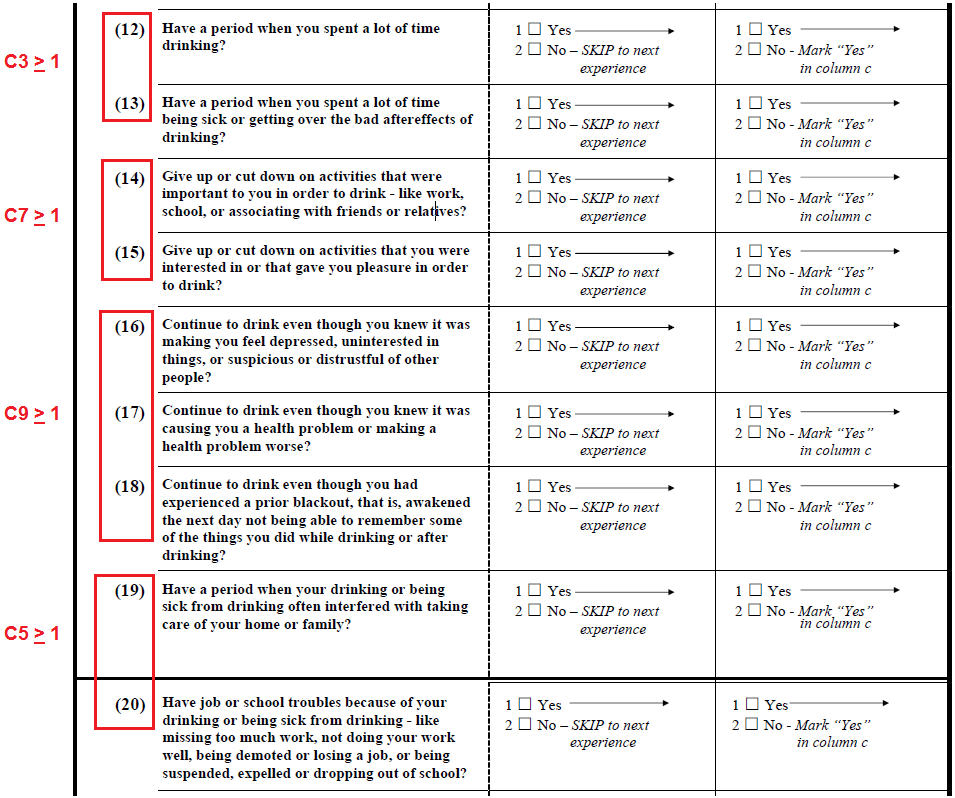
**Figure C.1** Survey Questions used to Elicit C1, C2 and C10



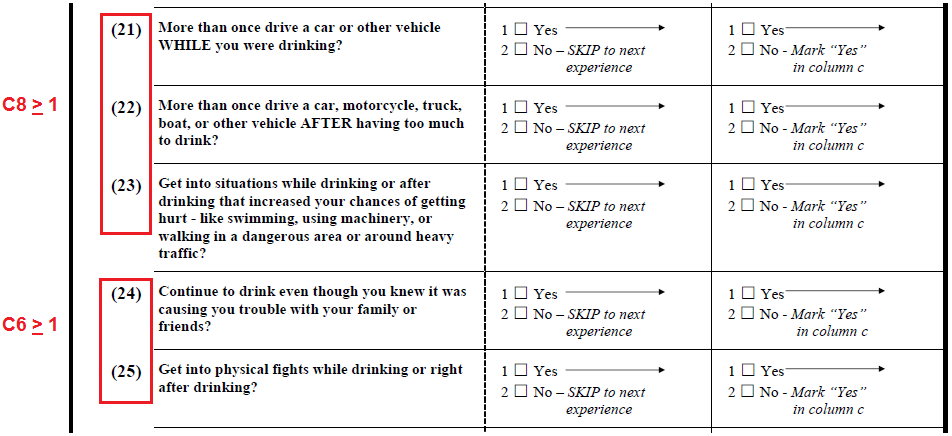
**Figure C.2** Survey Questions used to Elicit C11



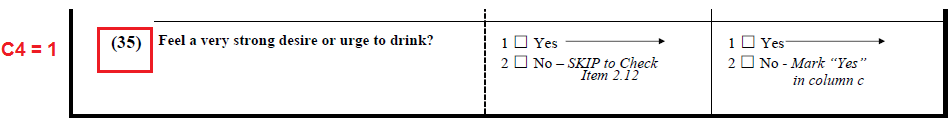
**Figure C.3** Survey Questions used to Elicit C3, C5, C7 and C9



**Figure C.4** Survey Questions used to Elicit C6 and C8



**Figure C.5** Survey Questions used to Elicit C4



**Appendix D:** SAS Macro

There are two macros available for download:

1. ***%DSF\_v2***
   1. This macro runs through each step as described in this article and provides a final dataset with the optimal diagnostic short-forms for the specified data set.
2. ***%DSF\_TF\_v2***
   1. This macro uses the datasets created in the previous macro to output tables and figures used to investigate findings

Email [cheriraffo@gmail.com](mailto:cheriraffo@gmail.com) for downloadable versions.