Consistency of reporting socio-demographic items and histories of physical and sexual abuse in DAANES

November 2006
Consistency of reporting socio-demographic items and histories of physical and sexual abuse in DAANES

By
James A. McRae, Jr., Ph.D.

Performance Measurement and Quality Improvement Division

November 2006

Copies of this report can be printed from http://www.dhs.state.mn.us/healthcare/studies.
**SUMMARY:** Data from the Minnesota Drug and Alcohol Abuse Normative Evaluation System (DAANES) indicate that providers are doing a very good job of reporting information on the gender and race of clients in treatment for chemical dependency but a very bad job of reporting whether the clients have been victims or perpetrators of physical or sexual abuse. This poor quality seriously hampers efforts to monitor the quality of treatment received by these clients. If, as seems likely, the reporting reflects a lack of knowledge by providers of histories of abuse, this lack of knowledge may hamper the effectiveness of treatment.

Providers of treatment for chemical dependency in Minnesota must report on various characteristics of their clients and the treatment process to the Department of Human Services (DHS) as part of the Drug and Alcohol Abuse Normative Evaluation System (DAANES). In order to study the outcomes of treatment, we link episodes of treatment for the same individuals over time. Many of the characteristics recorded in DAANES should be relatively stable over time. Variation in reports of these characteristics is likely to be due to errors in reporting by providers or errors in linking by DHS.

We linked episodes which had identical codes on at least two of (1) the first and third letters of the first and last name, (2) the last four digits of the Social Security Number, (3) the date of birth, and (4) the Person Master Index (PMI), which is a unique identifier for publicly funded clients. This linking showed that the 274,953 episodes of treatment in DAANES for 2000 through 2005 resulted from 151,766 people getting treatment. The number of episodes per person ranged from 1 to 25 with an average of 1.8. Although we did not perform the type of experiment that would enable us to quantify an error rate for the linking routine, we can say that it performed better than any other algorithm that we tried, including the well-regarded Link King program developed by the State of Washington.

Because we are using the linked data primarily to assess readmission rates for clients served by different providers, two desirable features of the data are that they be relatively recent and that we allow at least a year after discharge for the client to be readmitted. These two contradictory features led us to restrict attention to 2003.

In order to assess the stability of items over time, we cross-classified the measures obtained during the first three episodes for each person. Three measures that should show high levels of stability are gender, race, and ethnicity. Three measures that would vary considerably over the course of a lifetime are having military experience, being the victim of physical or sexual abuse, and being the perpetrator of physical or sexual abuse. Although these measures are likely to vary over one’s life, they vary in a distinct way:
once one has been a veteran, a victim of abuse, or a perpetrator of abuse, the status is not reversible.

For this analysis, gender is categorized male and female; race is categorized white, African American, American Indian or Alaskan native, Asian or Pacific Islander, mixed, and other; ethnicity is categorized Hispanic and not Hispanic; veteran status is categorized as none, non-combat, or combat; and being a perpetrator or victim of abuse are both categorized as no, physical abuse only, sexual abuse only, and both.

Differences in the stability of clients’ codes on these measures are striking. Among those with at least three episodes, we can calculate the percentage of clients who are coded the same way in all three episodes. Table 1 presents these percentages for the six measures. Interpreting these numbers is simple. For example, about 99 of 100 clients are coded the same gender in all three episodes; conversely, about 1% are coded male in some and female in others.

Table 1. Consistency of Coding Six Measures over Three Episodes

<table>
<thead>
<tr>
<th>Measure</th>
<th>Percent Consistent</th>
<th>Adjusted Percent Consistent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>98.9</td>
<td>97.3</td>
</tr>
<tr>
<td>Race</td>
<td>92.8</td>
<td>88.8</td>
</tr>
<tr>
<td>Ethnicity</td>
<td>97.3</td>
<td>72.3</td>
</tr>
<tr>
<td>Veteran status</td>
<td>95.2</td>
<td>66.1</td>
</tr>
<tr>
<td>Victim of abuse</td>
<td>57.3</td>
<td>33.5</td>
</tr>
<tr>
<td>Perpetrator of abuse</td>
<td>71.4</td>
<td>18.9</td>
</tr>
</tbody>
</table>

Consistency is substantial for four of these measures but much lower for the two that concern abuse. For gender, race, ethnicity, and veteran status, over 90% of clients have identical codes at all three times. Consistency is much lower for reports of being a victim (57%) or perpetrator (71%) of abuse. The fact that consistency is high for the first four measures suggests that the linking was relatively effective. Given that we did not link on any of the factors in the table, discrepancies should generally be as likely on the demographic variables as on any others. That they are not suggests that improper linking causes fewer errors than does misreporting.

We must consider one rather technical matter, though, before concluding that misreporting constitutes the primary problem. To illustrate this problem, consider dichotomous versions of two variables, say gender and race. Assume that gender is equally distributed with 50% male and 50% female, and race is distributed with 90% white and 10% black. If we measure both variables at time 1 and then simply assign people randomly at time 2 in a way that preserves the distributions, the percent consistent will be much higher for race than for gender. In fact, by chance alone, we expect that 82% of cases will be consistent on race, but only 50% will be consistent on gender. This results from the fact that our assignments of race are much more likely to be correct than are our assignments of gender, because most people are white and we will guess that most people are white. In fact, given that we must guess that 90% are white, making the maximum number of errors will result in correct guesses 80% of the time. Conversely,
making the maximum number of errors with gender will result in no correct guesses. Put simply, the odds are with us when a variable is unevenly split.

Addressing this rather technical problem requires a somewhat technical solution. The principle is simple: we must create a measure that assesses consistency above what would occur simply by chance. (Those who are not inclined to technical discussions should skip to the next paragraph.) A measure which accomplishes this task in the three-way cross-classification is,

\[
A = 100 \frac{\sum (f_{i=j=k} - F_{i=j=k})}{n - \sum F_{i=j=k}},
\]

where \(A\) is the Adjusted Percent Consistent, \(f\) is the observed frequency, \(F\) is the expected frequency under the model of independence, \(i\) indicates the category of the measure at Time 1, \(j\) indicates the category at Time 2, \(k\) indicates the category at Time 3, \(n\) is the total number of observations, and the subscript \(i=j=k\) indicates cells in which the measures are coded identically. For simplicity, we write \(\sum\) to indicate summing the cells over \(i, j,\) and \(k\). We multiply by 100 to move from the metric of proportions to the metric of percentages. In effect, the numerator measures the number of observations in consistent cells above what would be expected by chance, and the denominator measures the number of observations that are at risk of being consistent above what would be expected by chance. The measure ranges from 0 for items that are no more consistent than would occur by chance to 100 for those in which all observations are consistent.

The measures of adjusted percent consistent \((A)\) appear in the last column of Table 1. Agreement on gender and race remain high, but the measure for ethnicity is substantially smaller than the simple percent consistent. This suggests that much of the apparent agreement on ethnicity results from its distribution. It appears that some providers appear to code correctly only because they tend to guess that most clients are not Hispanic. It is worth noting, though, that an \(A\) of 72 indicates that most providers are doing a good job of getting this information. Similarly, although \(A\) is much smaller than the percent consistent for veteran status, it remains fairly high and indicates that a majority are coded consistently and presumably correctly.

The same cannot be said for being a victim or perpetrator of abuse. The apparent consistency of coding of these items results largely from the distributions of the variables. Once those distributions are considered, only a minority of additional cases are being coded consistently.

Given that \(A\) is lowest for those traits that are most likely to change over one’s life, it is possible that the actual values are changing between episodes. Since movement can only be in one direction—one can become a veteran, a victim of abuse, or an abuser, but one cannot become a non-veteran, a non-victim, or a non-abuser—this type of change would lead to a pattern of asymmetry in the cross-classifications. Table 2 classifies those who show change between the first and second episode according to whether the change is logically plausible or logically implausible.
Table 2. Number of Changes Between First and Second Episodes That Are Logically Plausible and Logically Implausible.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Logically Plausible Changes</th>
<th>Logically Implausible Changes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Veteran status</td>
<td>296</td>
<td>299</td>
</tr>
<tr>
<td>Victim of abuse</td>
<td>2,419</td>
<td>2,159</td>
</tr>
<tr>
<td>Perpetrator of abuse</td>
<td>1,468</td>
<td>1,449</td>
</tr>
</tbody>
</table>

There is a slight tendency in the data for change in the plausible direction, but this tendency is very slight. Under a random process, we would expect about equal numbers to move in both directions; the small deviations from equal change do little to dispel the notion that most of the reported change is random.

The analysis reported herein suggests that non-threatening demographic variables, especially those with visible clues to the correct coding, are reported quite reliably, although there is room for improvement. Unfortunately, analysis of reports of being a victim of abuse and being a perpetrator of abuse suggests that the room for improvement exceeds acceptable levels. The incorrect recording of data seriously impedes our ability to monitor the effectiveness of treatment for different groups. Given adjusted consistencies of 19% for perpetrating abuse and 34% for being a victim of abuse, attempting to determine outcomes of treatment for perpetrators or victims is an exercise in futility.

Perhaps more importantly, we know that having been abused elevates the risk of addiction substantially. Data of the 2004 Minnesota Student Survey for twelfth graders show that a history of abuse doubles the likelihood of needing treatment. Effectively addressing the problems with substances that these people experience is likely to involve addressing the histories of physical and sexual abuse. But those histories cannot effectively be addressed if they are not known. Assuming that the people who complete DAANES forms have access to the knowledge about clients gained during treatment, it appears that the histories of abuse are simply not being discovered.

In a similar vein, we know that having been in combat elevates the risk for post-traumatic stress, which is likely to lead to some addictive behavior and seriously complicate treatment. But the issue cannot be addressed if the history is not known.

We urge providers to get to know their clients better and to do a better job of conveying the information to DHS. Doing so will improve the quality of treatment and enable DHS to evaluate the effectiveness of treatment for diverse groups of clients more effectively.