### Smoking Status

<table>
<thead>
<tr>
<th>Age Cohorts</th>
<th>Older, Younger and Mid-age</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surveys</td>
<td>All Surveys for Younger and Mid-age</td>
</tr>
<tr>
<td>Derived Variable</td>
<td>SmokSt, SmokStShort, SmkAIHW</td>
</tr>
<tr>
<td>Definition</td>
<td>Tobacco smoking status</td>
</tr>
<tr>
<td>Source Items</td>
<td>SMOK1 (SMOK-007)</td>
</tr>
<tr>
<td>(Index Numbers)</td>
<td>SMOK2 (SMOK-018)</td>
</tr>
<tr>
<td>Statistical Form</td>
<td>Categorical variable</td>
</tr>
<tr>
<td>Index Number</td>
<td>SMOK-022</td>
</tr>
<tr>
<td>Prepared by</td>
<td>Anne Russell and Jean Ball</td>
</tr>
<tr>
<td>Endorsed</td>
<td>16 April 2003, Updated Oct 2012</td>
</tr>
</tbody>
</table>

Update Oct 2012

The Smoking Status variable, SmokSt, is present in all surveys of the 1946-51 cohort (Mid-age) and 1973-78 cohort (Younger). This variable is present in the first two surveys of the 1921-26 cohort (Older). The sixth survey of the 1921-26 cohort has a shorter version of this variable, SmokStShort. It has only four categories: Never smoked (1), Ex-smoker (2), Smokes occasionally (3), Smokes regularly (4).

In some surveys there is an extra category for the SmokSt variable. This category, 6, is for those women reporting smoking but not reporting the amount they smoked. They were classified smoking, indeterminate.

#### Source Items

**SMOK1** How often do you currently smoke cigarettes or any tobacco products?

<table>
<thead>
<tr>
<th>Code</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Daily</td>
</tr>
<tr>
<td>2</td>
<td>At least weekly (but not daily)</td>
</tr>
<tr>
<td>3</td>
<td>Less often than weekly</td>
</tr>
<tr>
<td>4</td>
<td>Not at all</td>
</tr>
</tbody>
</table>

**SMOK2** In your lifetime, would you have smoked at least 100 cigarettes (or equivalent)?

<table>
<thead>
<tr>
<th>Code</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Yes</td>
</tr>
<tr>
<td>2</td>
<td>No</td>
</tr>
</tbody>
</table>

#### Derived Variable

At the second ALSWH surveys of younger and mid-age women, survey items concerning tobacco smoking (above) were revised in line with contemporary recommendations by the Australian
Institute of Health and Welfare (AIHW). These two items were used to define categories for smoking status that were detailed in the National Health Data Dictionary (see over).
Smoking Status

<table>
<thead>
<tr>
<th>Code</th>
<th>AIHW Category</th>
</tr>
</thead>
</table>
| 1    | Never smoker<br>
A person who does not smoke now and has smoked fewer than 100 cigarettes or similar amount of other tobacco products in his/her lifetime |
| 2    | Ex-smoker<br>
A person who does not smoke at all now, but has smoked at least 100 cigarettes or a similar amount of other tobacco products in his/her lifetime. |
| 3    | Irregular smoker<br>
A person who smokes less than weekly |
| 4    | Weekly smoker<br>
A person who smokes at least weekly but not daily |
| 5    | Daily smoker:<br>
A person who smokes daily |

Note that the ALWSH data sets contain another measure of smoking status (SmokSt) which is similar to but not interchangeable with AIHW definition. Categories for SmokSt are:

<table>
<thead>
<tr>
<th>Code</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Never smoker</td>
</tr>
<tr>
<td>2</td>
<td>Ex-smoker</td>
</tr>
<tr>
<td>3</td>
<td>Current smoker – smokes less than 10 cigarettes per day</td>
</tr>
<tr>
<td>4</td>
<td>Current smoker – smokes 10-19 cigarettes per day</td>
</tr>
<tr>
<td>5</td>
<td>Current smoker – smokes 20 or more cigarettes per day</td>
</tr>
<tr>
<td>6</td>
<td>Smoker, indeterminate - the amount smoked was not given, but reported smoking.</td>
</tr>
</tbody>
</table>

Smoking status (SmkAIHW) is defined for women completing the full and short versions of Survey 2 in the Younger cohort (n=9 685) and women completing the full version only of Survey 2 in the Mid-age cohort (n=11 648). The distribution of SmkAIHW in the second survey of younger and mid-age women is shown in Table 1. Table 2 shows how combinations of responses to the relevant survey items were allocated to each category for smoking status.

### Table 1  Number and percent in AIHW smoking status categories at Survey 2

<table>
<thead>
<tr>
<th>Code</th>
<th>Smoking status - AIHW</th>
<th>Younger</th>
<th>Mid-age</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Number</td>
<td>Percent</td>
<td>Number</td>
<td>Percent</td>
</tr>
<tr>
<td>1</td>
<td>Never smoker</td>
<td>5 509</td>
<td>57.2</td>
<td>6 509</td>
<td>56.4</td>
</tr>
<tr>
<td>2</td>
<td>Ex-smoker</td>
<td>1 390</td>
<td>14.4</td>
<td>3 054</td>
<td>26.4</td>
</tr>
<tr>
<td>3</td>
<td>Irregular smoker</td>
<td>569</td>
<td>5.9</td>
<td>158</td>
<td>1.4</td>
</tr>
<tr>
<td>4</td>
<td>Weekly smoker</td>
<td>299</td>
<td>3.1</td>
<td>112</td>
<td>1.0</td>
</tr>
<tr>
<td>5</td>
<td>Daily smoker</td>
<td>1 873</td>
<td>19.4</td>
<td>1 712</td>
<td>14.8</td>
</tr>
<tr>
<td></td>
<td>Total classified</td>
<td>9 640</td>
<td>100.0</td>
<td>11 545</td>
<td>100.0</td>
</tr>
<tr>
<td>9</td>
<td>Missing</td>
<td>45</td>
<td>0.5</td>
<td>103</td>
<td>0.9</td>
</tr>
</tbody>
</table>
Table 2  Assignment to AIHW smoking status at Survey 2 by combinations of current smoking and lifetime consumption of at least 100 cigarettes

<table>
<thead>
<tr>
<th>Smoking status</th>
<th>Current smoking</th>
<th>Ever smoked 100 cigarettes</th>
<th>Younger</th>
<th>Mid-age</th>
</tr>
</thead>
<tbody>
<tr>
<td>Never smoker</td>
<td>Not at all</td>
<td>No</td>
<td>5 509</td>
<td>6 509</td>
</tr>
<tr>
<td>Ex-smoker</td>
<td>Not at all</td>
<td>Yes</td>
<td>1 390</td>
<td>3 054</td>
</tr>
<tr>
<td>Irregular smoker</td>
<td>Less often than weekly</td>
<td>Yes</td>
<td>412 142</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>No</td>
<td>151 14</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Missing</td>
<td>6 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weekly smoker</td>
<td>At least weekly (but not daily)</td>
<td>Yes</td>
<td>274 104</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>No</td>
<td>21 3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Missing</td>
<td>4 5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Daily smoker</td>
<td>Daily</td>
<td>Not applicable</td>
<td>1 797</td>
<td>1 119</td>
</tr>
<tr>
<td></td>
<td>Missing</td>
<td>76</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>589</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>Missing</td>
<td>Missing</td>
<td>25 82</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>4 4</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not at all</td>
<td>Missing</td>
<td>Missing</td>
<td>16 17</td>
<td></td>
</tr>
</tbody>
</table>

The SAS code defining smoking status at Survey 2 is:

```sas
/*Preliminary code for Mid-age */;
   if m2survey=2 then smkAIHW = . ;
else if m2survey = 1 then do ;
/* Smoking status based on AIHW definition */;
   if SMOK1 in (9,. ) then smkAIHW = 9 ;
else if SMOK1 = 4 then do ;
      if SMOK2 = 2 then smkAIHW = 1 ;
else if SMOK2 = 1 then smkAIHW = 2 ;
else if SMOK2 = . then smkAIHW = . ;
end ;
else if SMOK1 = 3 then smkAIHW = 3 ;
else if SMOK1 = 2 then smkAIHW = 4 ;
else if SMOK1 = 1 then smkAIHW = 5 ;
```

**Code for SMOKST**

The code for Smoking Status is not always the same in each wave and cohort. Below are listed the different pieces of code and the surveys they are used in.
The code for Smoking Status is not always the same in each wave and cohort. Below are listed the different pieces of code and the surveys they are used in.

Note that some surveys have a code for SMOKST of 6. This is a smoker whose quantity of smoking is unknown but they are not non-smokers.

Survey 1 Smoking Status, SMOKST, the three cohorts

This example has the Mid variables, and similarly for the other cohorts.

\[ \text{smkfreq} = m1q38; \]
\[ \text{smkday} = m1q40a; \]

```plaintext
/***********
***
*** WHA 1 smoking status ***
*** (for Young1, Mid1, Old1) ***
*** 1 = Non-smoker ***
*** 2 = Ex-smoker ***
*** 3 = smoker <10 per day ***
*** 4 = smoker 10-19 per day ***
*** 5 = smoker >=20 per day ***
*** 6 = smoker, qty unknown ***
***
*************************************************
if smkfreq=1 then smokst=1;
else if smkfreq=2 then smokst=2;
else if smkfreq in (3,4) then do;
   if 0<=smkday<10 then smokst=3;
   else if 10<=smkday<20 then smokst=4;
   else if smkday>=20 then smokst=5;
end;
else if smkfreq=. then smokst=.;
/*** For young 1 only, recode smokers with missing quantity to 6 ***/
if smkflag=1 then smokst=6;
```

Survey 2 for the OLD, 1921-26 cohort, only

```plaintext
/********************
***
*** Use smoking status from Old1 ***
*** and o2q28 to create o2smokst ***
***
******************************
data tmpold5;
length o2smokst 3;
merge tmpold4(in=a) o1stat(in=b);
   by id;
   if a and b;
if o1smokst=1 and o2q28 in (.,0) then o2smokst=1;
```

ALSWH Data Dictionary Supplement
Section 2 Core Survey Dataset
2.5 Lifestyle Variables
Smoking Status
Survey 2 MID, Survey 2 onwards Young
The Young cohort for surveys 2 onwards (to at least 6) and for Survey 2 only for the MID, 1946-51 cohort, uses this code. Note the 'p5' suffix is for Survey 5 but the code otherwise the same.

```plaintext
/****************************
***                     ***
*** WHA 1 smoking status ***
***           (Mid 5)     ***
***    1 = Non-smoker    ***
***    2 = Ex-smoker     ***
***    3 = smoker <10 per day ***
***    4 = smoker 10-19 per day ***
***    5 = smoker >=20 per day ***
***                     ***
*** Reads in variables: ***
***    smkfreq, smkday, smkstop ***
***                     ***
/****************************
if whasmk=1 then do ;
    if smkfreq=1 and 20<=smkday<=100 then p5smokst=5 ;
    else if smkfreq=1 and (10<=smkday<20) then p5smokst=4 ;
    else if smkfreq=1 and (1<=smkday<10) then p5smokst=3 ;
    else if 2<=smkfreq<=3 then p5smokst=3 ;
    else if smkfreq=4 and 1<=smkstop<=62 then p5smokst=2 ;
    else if smkfreq=4 and smkstop=. then p5smokst=1 ;
```

if o1smokst=1 and o2q28 in (3,4,5) then o2smokst=3 ;
if o1smokst=2 and o2q28 in (. ,0) then o2smokst=2 ;
else if o1smokst=2 and o2q28 in (1,2,3) then o2smokst=3 ;
else if o1smokst=2 and o2q28=4 then o2smokst=4 ;
else if o1smokst=2 and o2q28=5 then o2smokst=5 ;
if o1smokst=3 and o2q28=0 then o2smokst=2 ;
else if o1smokst=3 and o2q28 in (1,2,3) then o2smokst=3 ;
else if o1smokst=3 and o2q28=4 then o2smokst=4 ;
else if o1smokst=3 and o2q28=5 then o2smokst=5 ;
if o1smokst=4 and o2q28=0 then o2smokst=2 ;
else if o1smokst=4 and o2q28 in (1,2,3) then o2smokst=3 ;
else if o1smokst=4 and o2q28=4 then o2smokst=4 ;
else if o1smokst=4 and o2q28=5 then o2smokst=5 ;
if o1smokst=5 and o2q28=0 then o2smokst=2 ;
else if o1smokst=5 and o2q28 in (1,2,3) then o2smokst=3 ;
else if o1smokst=5 and o2q28=4 then o2smokst=4 ;
else if o1smokst=5 and o2q28=5 then o2smokst=5 ;
else if smkfreq=. then p5smokst=. ;
end ;

**MID 3 onwards**
The MID, 1946-51 cohort, Surveys 3 to 7 use the following code. Note that this is from Survey 5 and the variables have the ‘p5’ prefix.

```plaintext
/****************************
*** WHA 2 smoking status ***
*** 1 = Non-smoker ***
*** 2 = Ex-smoker ***
*** 3 = smoker <10 per day ***
*** 4 = smoker 10-19 per day ***
*** 5 = smoker >=20 per day ***
***
*** Reads in variables:
***
*** smkfreq, smkday, smk100, smkdly ***
***
*******************************************************************************/

if smkfreq=1 and (0<=smkday<10) then p5smokst=3 ;
else if smkfreq=1 and (10<=smkday<20) then p5smokst=4 ;
else if 2<=smkfreq<=3 then p5smokst=3 ;
else if smkfreq=4 and (smk100=1 or smkdly=1) then p5smokst=2 ;
else if smkfreq=4 and smk100=2 then p5smokst=1 ;
else if 20<=smkday<=70 then p5smokst=5 ;
else p5smokst=. ;
```

**References**