Use of an Industry Sector Pick List to Collect Industry of Employment in the Massachusetts Behavioral Risk Factor Surveillance System (BRFSS)

Technical Report 2010-1

February 2010

Occupational Health Surveillance Program
Massachusetts Department of Public Health

Acknowledgements

This report was prepared by Shona Fang, ScD, and Letitia Davis, ScD, EdM of the Occupational Health Surveillance Program. Special thanks go to the staff of the MDPH Health Survey Program, who oversees the Massachusetts Behavioral Risk Factor Surveillance System.

The work was funded in part through a cooperative agreement with the National Institute for Occupational safety and Health (U60/OH008490).

To obtain additional copies of this report, contact:

Massachusetts Department of Public Health
Bureau of Health Information, Statistics, Research and Evaluation
Occupational Health Surveillance Program
250 Washington Street, 6th Floor
Boston, MA 02018

617-624-5632

This report is also available on line at the MDPH's website:

www.mass.gov/dph/ohsp

Background

The Behavioral Risk Factor Surveillance System (BRFSS) is a state-based population health surveillance tool and widely relied-on source of information on a variety of health risk behaviors, clinical preventive health practices, health care access, chronic conditions, and emerging public health issues.[1] It is a continuous, population-based, random-digit-dial telephone survey of adults ages 18 and older conducted by state health departments in all states and territories in collaboration with the national Centers for Disease Control and Prevention (CDC). While there is some variation in the specific questions asked each year, the Massachusetts BRFSS includes a core set of questions developed by the CDC, optional state modules developed by the CDC, and state-added questions developed by programs within the Massachusetts Department of Public Health (MDPH). This survey has great potential to generate otherwise unavailable information about occupational health outcomes and risk factors at the state level. In recent years, Massachusetts, as well as a number of other states, has used it for this purpose.[2] Routine collection of information about occupation and/or industry in the BRFSS is potentially highly useful for other public health initiatives, most notably chronic disease prevention and control programs that are increasingly focusing on the worksite as a venue to promote health. Inclusion of employment information in this well-established surveillance system would also foster a broader public health vision of healthy work that addresses both health promotion and health protection in the workplace.

In 2007, Massachusetts included questions in the BRFSS on category of industry. Questions on work-related injuries and use of workers' compensation were also asked on the Massachusetts survey and by several other states. In the past (2001 and 2002), Massachusetts BRFSS had asked about industry and occupation along with questions specific to work-related asthma. Industry and occupation were collected as separate narrative text fields from those respondents who answered yes to questions about work-related asthma. Coding this information, however, proved to be resource intensive and was of limited success. Surveyors were not properly trained to ask these questions and many responses were too vague to be coded. Surveyor training is necessary to improve the usefulness of the narrative industry and occupation information collected. Future electronic industry and occupation coding software, currently under development at the National Institute for Occupational Safety and Health (NIOSH), will make for more efficient coding of narrative text.

Given past experience and limited resources for training surveyors and coding information collected on industry and occupation, Massachusetts elected to pilot the use of a pick-list to collect industry information in the 2007 survey. We chose to collect information on industry as opposed to occupation because industry is more useful for targeting worksite interventions. Moreover, the number of possible occupations far exceeds the number of industry sectors or subsectors, and thus asking about industry in the survey was more practical. We derived the industry pick list empirically from the industry distribution of the Massachusetts workforce, choosing the most common industries and categories that could be collapsed into the established North American Industry Classification System (NAICS) coding scheme of the U.S. Census Bureau. The pick list options were also selected and phrased to be easily interpretable by the general public.

This report summarizes findings from our analysis of the industry pick list, including an analysis of the narrative responses given for the option of "other". These findings demonstrate the utility of using an industry pick list in the BRFSS survey and its potential use in other surveys and data collection systems, including electronic medical records. Our findings also suggest how the industry pick list could be improved for future use.

[1] Centers for Disease Control and Prevention. Public health surveillance for behavioral risk factors in a changing environment: Recommendations from the Behavioral Risk Factors Surveillance Team. MMWR 2003;52(No. RR-9):[1-12]

[2] Fan ZJ, Bonauto DK, Foley MP, Silverstein BA. Under-reporting of work-related injury or illness to workers' compensation: individual and industry factors. J Occup Environ Med. 2006 Sep;48(9):914-22.

Methods

In 2007, the Massachusetts BRFSS surveyed 21,507 adult Massachusetts residents. To increase the number of respondents who belong to racial and/or ethnic minority groups, the cities of Boston, Worcester, Springfield, Lawrence, Lowell, Fall River, and New Bedford were oversampled. In order to reflect both the probability that an individual was selected to participate in the survey and differential participation by sex, age, and race-ethnicity the percentages in this report were weighted to the total Massachusetts population in 2007.

In order to collect data on a wider range of topics each year different portions of the sample population answer different sets (splits) of BRFSS questions. The Massachusetts BRFSS sample has consisted of three survey splits since 2000. Each split contains state-added questions or optional CDC modules about health topics relevant to Massachusetts residents. Questions on work-related injuries and use of workers' compensation were included in the second split of the survey, which included 4,799 respondents.

If the respondent was employed for any period of time in the past twelve months as part-time, full-time or self employed, the respondent was asked the following question and read the possible responses:

What kind of business or industry do you work in?

1 Construction
2 Manufacturing
3 Wholesale, Retail Sales
4 Finance, Insurance, Real Estate
5 Health Care
6 Education
7 Government
8 Other [Specify: _____]

We computed the distribution of the responses and manually reviewed the narratives for the category of "other". For the narratives, we first determined whether the narrative was an industry or an occupation. We then assigned an industry based on the 2007 NAICS coding system where possible. In some cases, an occupation was specific enough so that an industry could be assigned. In contrast, in some cases, a reported business or industry was too vague for an industry code to be assigned.

To evaluate the performance of the industry pick list, we determined which narrative responses should have been coded as one of the available items in the pick list. Subsequently, the weighted percent distribution of the industries, including the manually coded industries, was computed, along with the 95% confidence intervals. In addition, while data were not available to perform formal statistical comparisons, we present qualitative comparisons with industry distributions obtained from two other sources: the BLS Quarterly Census of Employment and Wages (QCEW) and the Current Population Survey (CPS).

Based on the narrative responses, we also determined which industries should be included in future industry pick lists used in Massachusetts.

Findings

A total of 2,616 individuals who were employed for any period time in the past twelve months were asked the industry question.

Table 1: Distribution of responses to the industry question in the 2007 MA BRFSS (n=2,616)

	No.	%*	(95% CI)
Construction	126	5.7	(4.2 – 7.1)
Manufacturing	184	7.1	(5.4 – 8.8)
Wholesale, retail sales	265	10.7	(8.9 – 12.6)
Finance, insurance, real estate	214	8.6	(7.1 – 10.1)
Healthcare	462	15.5	(13.6 – 17.3)
Education	356	12.3	(10.6 – 14.1)
Government	153	5.6	(4.3 – 7.0)
Other	856	33.3	(30.5 – 36.0)
Don't know	10	0.3	(0.1 – 0.5)
Refused	25	0.9	(0.4 – 1.4)

*Percent is weighted to Massachusetts population characteristics in 2007.

Close to 65% of the respondents chose one of the industries on the pick list. The most common industry selected among the available options in the pick list was Healthcare (15.5%) (Table 1).

Table 2: Characteristics of "other" responses to the industry question in the 2007 MA BRFSS (n=856)

	No.	%
Industry or establishment provided:	512	59.8
NAICS code assigned	377	44.0
NAICS code not assigned (industry too non-specific)	135	15.8
Occupation provided:	331	38.7
NAICS code assigned (occupation was detailed enough)	94	11.0
NAICS code not assigned	237	27.7
Total assigned a NAICS code:	471	55.0
Should have been classified into pick list:	79	9.2
Construction	2	0.2
Manufacturing	23	2.7
Retail sales (retail trade)	16	1.9
Finance, insurance, real estate	8	0.9
Healthcare (not including social assistance)	2	0.2
Education	11	1.3
Government (Public administration)	17	2.0
Other types of responses:		
Unclear if industry or occupation	10	1.2
Retired	3	0.4

Among the "other" responses, 60% were an industry/business name while 39% were an occupation/job title. Ultimately, 55% of the other responses were assigned a NAICS code (Table 2).

Some of these assigned codes were based on a given occupation, if the occupation was specific enough (n=94).

The text response was too vague to be assigned an industry for 43.5% of the "other" responses. Examples included "non-profit organization", "computers", "software", "arts", and "technology/high tech".

Among the 471 "other" responses assigned an industry code, only 79 should have been assigned one of the industries from the dropdown list by the respondent. The most common were manufacturing (n=23), government (n=17), and retail sales (n=16). The 79 represented only a small percentage of all responses (3%), which indicated that overall, the available options in the pick-list performed well.

Table 3: "Other" responses to the industry question in the 2007 MA BRFSS subsequently coded to an industry (n=471)

Industry (2-digit NAICS code)	No.	%
Goods producing		
Agriculture, forestry, fishing, and hunting (11)	4	0.8
Mining (21)	2	0.4
Construction (23)	2	0.4
Manufacturing (31-33)	23	4.9
Service Producing		
Trade, transportation, utilities	53	11.3
Utilities (22)	4	0.8
Retail trade (44-45)	16	3.4
Transportation and warehousing (48-49)	33	7.0
Information (51)	44	9.3
Financial activities	8	1.7
Finance and insurance (52)	4	0.8
Real estate and rental and leasing (53)	4	0.8
Professional and business services	142	30.1
Professional, scientific, and technical services (54)	122	25.9
Management of companies and enterprises (55)	1	0.2
Administrative & support, waste management, remediation services (56)	19	4.0
Education and health services	70	14.9
Educational services (61)	11	2.3
Healthcare and social assistance (62)	59	12.5
Leisure and hospitality	72	15.3
Arts, entertainment and recreation (71)	21	4.5
Accommodation and food services (72)	51	10.8
Other services (except public administration) (81)	34	7.2
Public administration (92)	17	3.6
Total coded	**471**	**100.0**

Table 4: Top three most common sub-sectors in the "other" category and associated narrative terms in the 2007 MA BRFSS (n=232)

	No.	%
Professional, scientific, and technical services	122	25.9
"Law/legal"	41	8.7
"Engineering"	15	3.2
"Human services"	11	2.3
"Biotech/pharmaceutical"	10	2.1
Healthcare and social assistance*	59	12.5
"Childcare"	24	5.1
"Social services"	17	3.6
"Nursing home/home care/elderly care"	8	1.7
Accommodation and food services	51	10.8
"Restaurant/bar/food services"	40	8.5
"Hotel"	8	1.7

*While "healthcare" was on the pick list, "healthcare and social assistance" was not.

Among the "other" responses assigned an industry, the most common sub-sector was Professional, scientific, and technical services (25.9%) (Table 3). The most common responses in this sub-sector were: "law/legal", "engineering", and "human services" (Table 4).

The next most common sub-sector was "Healthcare and social assistance" (12.5%), the majority of which were in the social assistance category (Table 3). Common texts for this sub-sector were "childcare", "social services", and "nursing home/home care/elderly care" (Table 4).

The third most common sub-sector was "Accommodation and food services" (11%) (Table 3). This most commonly included "restaurant/bar/food services" and "hotel" (Table 4).

Table 5: MA industry distributions according to MA BRFSS, QCEW, and CPS

	MA BRFSS			QCEW**	CPS**
	No	%*	(95% CI)	Percent	Percent
Goods producing					
Natural resources and mining (11; 21)	6	0.3	(<0.1 – 0.6)	0.2	0.3
Construction (23)	128	6.8	(5.1 – 8.5)	4.6	6.5
Manufacturing (31-33)	207	9.4	(7.4 – 11.4)	9.1	8.9
Service producing					
Trade, transportation, utilities (22; 44-45; 48-49)†	318	14.4	(12.1 – 16.6)	18.4	16.9
Information (51)	44	2.4	(1.4 – 3.3)	2.9	2.5
Financial activities (52-53) ‡	220	10.7	(8.9 – 12.5)	7.0	8.3
Professional and business services (54-56)	142	7.1	(5.5 – 8.6)	15.1	13.4
Educational services (61)	367	15.1	(13.0 – 17.1)	9.7	10.6
Healthcare and social assistance (62)‖	521	20.6	(18.2 – 23.0)	15.1	15.0
Leisure and hospitality (71-72)	72	4.5	(2.9 – 6.1)	9.5	8.9
Other services (81)	34	1.3	(0.6 – 2.1)	3.9	4.7
Public administration (92)§	170	7.5	(5.9 – 9.2)	7.9	4.0

*Weighted to population characteristics; **2007 annual averages;
†Includes "wholesale, retail sales"; ‡Includes "finance, insurance, real estate; ‖Includes "healthcare"; §Includes "government".
QCEW: Quarterly Census of Employment and Wages; CPS: Current Population Survey

After we assigned a NAICS code to the "other" responses where possible and re-calculated the industry distribution, 21% of the respondents were employed in ***healthcare and social assistance***, 15% in ***educational services***, and 14% in the combined sectors of ***trade, transportation, and utilities*** (Table 5).

With the exception of a few industries, in general, the industry distribution estimates obtained from the BRFSS were not strikingly different from those obtained from the QCEW or the CPS, which in turn were fairly comparable with each other, except for ***public administration*** (Table 5). However, ranking of industries differed among the three sources, with ***healthcare and social assistance*** being the leading industry according to the BRFSS and ***trade, transportation and utilities*** being the leading industry according to the QCEW and CPS.

Notably, industry groupings that were not offered as a choice in the pick list - ***professional and business services*** and ***leisure and hospitality*** - were most under-represented as compared to findings from the QCEW and CPS. It is possible that respondents in these industries selected one of the available industry options because a better choice was not available.

Potential differences in the estimates obtained from the BRFSS and those obtained from the CPS and QCEW may be explained by differences in the population coverage of the surveys and the census; the latter covering only jobs on non-farm payrolls that are covered by unemployment insurance, and the former two surveys in theory being inclusive of the entire non-institutionalized civilian population. The BRFSS does not restrict participation in the survey by employment in certain industries.

Limitations

As a telephone survey, the BRFSS excludes individuals without a landline telephone (usually individuals of lower incomes and younger individuals),[3] and those who do not speak English, Spanish, or Portuguese. Therefore BRFSS results may not be generalizable to populations without a landline or who do not speak these languages. For example, the finding that a greater proportion of the BRFSS respondents worked in ***healthcare and social assistance*** as well as ***educational services***, as compared to the CPS and QCEW, may be reflective of the demographics of individuals more likely to be included in a telephone-based survey. Notably, due to the decreased use of landlines and increase of cell phone-only households in recent years, the Massachusetts BRFSS piloted the use of cellular phones to reach underrepresented groups and to reduce non-response bias in 2008. In 2009, every state's BRFSS will include cell phone interviewing, however only core questions will be asked in cell phone interviews.

In addition, the findings in this analysis may be affected by non-response bias associated with the low response rate, as the 2007 Massachusetts BRFSS CASRO response rate[4] was 34.6%. However the low response rate in the BRFSS does not necessarily reflect poor data quality as suggested by some studies.[5]

Also of importance to note is that these BRFSS data are based on self-reported information from respondents and therefore may be subject to some subjectivity. For example, an employer may be classified under more than one industry and a respondent may choose one over the other depending

[3] Stephen Blumberg and Julian Luke. Coverage bias in traditional surveys of low-income and young adults. Public Opinion Quarterly, Vol 17, No. 5 2007, pp. 734-749.
[4] This response rate formula developed by the Council of American Survey Research Organizations (CASRO) results in an estimate that reflects telephone sampling efficiency and the degree of cooperation among eligible persons who were contacted. The formula assumes that numbers that are never contacted contain the same percentage of eligible households as the records whose eligibility status is known.
[5] Fahimi M, Link M, Mokdad A, Schwartz D, and Levy P. Tracking Chronic Disease and Risk Behavior Prevalence as Survey Participation Declines: Statistics From the Behavioral Risk Factor Surveillance System and Other National Surveys. Prev Chronic Dis. 2008 July; 5(3): A80.

on his or her own preference (e.g. a state hospital employee may consider him or herself to work in the healthcare industry or for government).

Despite the limitations, the results from this survey provide valuable information on the ability to collect information on industry using an empirically derived pick list.

Conclusions and Recommendations

An evaluation of the narrative "other" responses showed that only 9.2% of those responding "other" (table 2), which is only 3% of all the respondents to the industry question, should have chosen an industry from the pick list, suggesting that the pick list performed reasonably well. We also found that the text for the "other" responses most commonly fell into the broad industry sectors of ***professional and business services*** and ***leisure and hospitality,*** and the sub-sector ***social assistance***. The most commonly reported type of industries for these sectors were "***law/legal***"; "***engineering***"; "***human services***"; "***childcare***", "***social services***", and "***nursing home/home care/elderly care***" and "***restaurant/bar/food services***" and "***hotel***". Future industry pick lists in Massachusetts should incorporate these commonly reported industries.

Providing fewer, but more interpretable industry categories in a pick list improves the ability of respondents to select the correct industry of their employment. While we were not able to validate the information collected from the pick list, we selected industry options for the pick list to be easily understandable and interpreted by the general public. Future work should be conducted to validate information collected from industry pick lists, either through the use of an external source of data or by collecting narrative text responses in addition to pick list responses. In addition, cognitive testing of industry categories should be conducted to determine respondents' understanding of industry categories. Furthermore, interviewers and other individuals collecting information on industry should be properly trained to solicit such information and to understand differences between industry and occupation, as occupations were commonly found in the narrative text of the "other" response category.

Routine collection of industry information in health data systems is important for targeting occupational health initiatives and for a variety of other disease and injury prevention initiatives, including chronic disease prevention and control programs increasingly implemented through worksite wellness programs. While the pick list used by the Massachusetts BRFSS in 2007 worked well, improvements as described in this report should be made. Future efforts should be directed at routinely collecting industry information, evaluating mechanisms to collect such data, and assessing the quality of the data collected. The use of a relatively few broad categories in a pick list, rather than a more comprehensive list of specific categories or the collection of narrative text may result in a loss of information that can be useful for more focused targeting for work-related injury and illness surveillance and interventions. It is also possible that by providing fewer broad industry categories in the dropdown list, respondents may decide to choose another industry that is available on the pick list as a close alternative, rather than choosing "other", thereby misclassifying their industry of employment.

When data are collected for sentinel case surveillance, clinical purposes, or in larger population based surveys in which analysis at the detailed industry level is feasible, collection of more detailed industry information, either via detailed dropdown lists or narrative text, should be the optimal choice. However, use of a broadly defined pick list for collecting industry data may be a useful alternative in population-based studies when resources for coding are limited, and/or the data are not intended for focused interventions. In addition, such pick lists may be useful when the size of the survey sample is small and detailed data would be lost by the need to group data together.