Synthetic Estimates of Work Loss Disability for Each State and the District of Columbia

TSUKASA NAMEKATA, PhD, PAUL S. LEVY, ScD, and THOMAS W. O'ROURKE, PhD, MPH

THROUGH SAMPLE SURVEYS and registration systems, the National Center for Health Statistics has for many years been collecting data on morbidity and mortality, the utilization of health facilities, and the social impact of disease. These surveys of the U.S. population, often referred to as the National Health Survey, are not designed to serve any single health program nor to meet the needs for detailed local data. Instead, they generally result in estimates of health statistics for the United States as a whole or for each of its four broad geographic regions (1).

The National Center's survey programs do not provide estimates of health characteristics for local areas such as States or counties partly because the sample sizes in the programs are not adequate for these small areas and partly because the strata used in the sampling designs cannot be combined easily into estimates for these local areas since they often cut across State lines.

On the other hand, the National Center for Health Statistics, as well as other Federal agencies, has recognized the increasing need for good estimates of the health, social, and economic parameters of local areas and is now giving some priority to devising methods for producing such estimates. One such method, "synthetic estimation," which was developed originally at the National Center and later at the Bureau of the Census is the subject of this report.

Synthetic estimates, by State, of certain kinds of disability have been published by the National Center (2), but evaluation of these estimates has been difficult because

of the lack of any direct estimates against which they might be compared. However, a question on work loss disability appeared on the U.S. decennial census questionnaire for the first time in 1970, and direct estimates have been published for each State and for the United States as a whole. (A person is considered as having a "partial work disability" or a "partial work loss disability" if he or she has a health condition which limits the kind or amount of work that he or she can do but does not totally prevent the person from holding a job for a relatively long time. A person is considered to have "complete work disability" or "complete work loss disability" if the health condition prevents him or her from working at a job for a relatively long time.)

The data on work loss disability were obtained from the responses to question 28 in the 5 percent sample questionnaire in the 1970 decennial census. In our study, persons whose disability (complete or partial) was of less than 6 months' duration were excluded.

Qualitatively, synthetic estimates are indirect estimates of a characteristic of the population of a local area within a larger population and are obtained by estimating the characteristic for demographic subgroups of the larger population in combination with population

data on the distribution of the local population into these demographic subgroups. This method has been used experimentally by the National Center to produce estimates of certain disability parameters for States (2) and by the Census Bureau to produce estimates of unemployment and other parameters for States and counties (3,4).

Although it is relatively easy to produce synthetic estimates if the necessary data are available, it is much more difficult to evaluate the estimates in terms of their validity and reliability. Statistically, such estimates are biased and, unlike the usual estimates derived from samples surveys, cannot be evaluated on the basis of the survey data by straightforward procedures such as the calculation of sampling errors. In fact, when synthetic estimates are based on the estimates from large samples, it is likely that the sampling errors will be small and that the main sources of error will lie in the biases of these estimates.

Synthetic estimates can be created for each State from the published data on disability for demographic subgroups of the United States as a whole and from the 1970 census data on the distribution of the population of each State into subgroups. Our purpose has been to produce such synthetic

☐ Dr. Namekata is an instructor in occupational and environmental medicine, and Dr. Levy is associate professor in biometry, School of Public Health, University of Illinois Medical Center. Dr. O'Rourke is associate professor, Department of Health and Safety Education, University of Illinois, Champaign. Tearsheet requests to Dr. Paul S. Levy, Associate Professor, School of Public Health, University of Illinois Medical Center, P.O. Box 6998, Chicago, Ill. 60680.

estimates and to compare them against the direct estimate for the State.

Methods

Sources of data. The data used in this paper were obtained from the 1970 census reports *(5,6)*.

Kind of data	Source
On work disability at national level.	Reference 5, table 1.
On work disability at State level.	Reference 6, table 169.
On population at State level.	Reference 6, tables 20 and 48.

Synthetic estimation. To estimate the percentage of persons with work disability, a slight modification of the method of synthetic estimates developed by the National Center for Health Statistics was applied *(2)*. Following is the formula used to produce the estimates:

$$\widetilde{x}_s = \overset{k}{\underset{\alpha = 1}{\Sigma}} P_{s\alpha}\overline{x}'_\alpha$$

where

$\overline{x}'_\alpha$ = the proportion of persons with partial (or complete) work disability in cell α on the national level (from the census report PC2-6C),

$P_{s\alpha}$ = the proportion of persons in State s that are in cell α (from census reports on particular States),

k = the number of α-cells in the grid, and

$\widetilde{x}_s$ = synthetic estimate for State s.

In our study, a grid of 60 α-cells, including age (18-44 years, 45-54 years, and 55-64 years), race (white and Negro or black); sex (male and female), and type of residence (central cities, urban fringe, other urban, rural farm, and nonfarm), was used to create the synthetic estimates.

Thus, a synthetic estimate for an individual State was determined by the application of national rates descriptive of the health characteristics of certain population groups. Synthetic estimates for complete and partial work disability were calculated separately. These measure qualitatively two different concepts of disability.

Direct estimation. The direct estimate of work disability for State s is denoted by the ratio of the number of persons (whites and Negroes combined) 18 to 64 years of age in the State with work dis-

Table 1. Direct and synthetic estimates, absolute and percentage absolute differences, and mean square errors for partial work disability of more than 6 months for persons 18−64 years of age

State	Direct estimate	Synthetic estimate	Absolute difference	Percentage absolute difference	Mean square error
1. Alabama	6.24	5.92	0.32	5.4	0.10
2. Alaska	4.77	5.47	0.69	12.6	0.48
3. Arizona	6.70	5.68	1.02	18.0	1.04
4. Arkansas	7.31	6.11	1.19	19.5	1.43
5. California	5.96	5.53	0.42	7.6	0.18
6. Colorado	6.15	5.52	0.63	11.5	0.40
7. Connecticut	5.05	5.71	0.67	11.7	0.44
8. Delaware	4.99	5.55	0.56	10.0	0.31
9. District of Columbia	5.74	6.04	0.30	5.0	0.09
10. Florida	6.38	5.84	0.53	9.1	0.29
11. Georgia	6.31	5.75	0.57	9.9	0.32
12. Hawaii	4.14	5.28	1.14	21.6	1.30
13. Idaho	7.50	6.03	1.47	24.5	2.17
14. Illinois	5.33	5.75	0.42	7.3	0.18
15. Indiana	5.78	5.79	0.01	0.2	0.00
16. Iowa	6.31	5.95	0.35	6.0	0.13
17. Kansas	6.02	5.85	0.17	2.9	0.03
18. Kentucky	5.81	5.83	0.03	0.5	0.00
19. Louisiana	5.38	5.82	0.44	7.5	0.19
20. Maine	5.61	5.97	0.35	5.9	0.12
21. Maryland	5.14	5.60	0.47	8.3	0.22
22. Massachusetts	5.14	5.60	0.46	8.1	0.21
23. Michigan	6.02	5.66	0.36	6.4	0.13
24. Minnesota	6.09	5.72	0.37	6.5	0.14
25. Mississippi	6.19	6.03	0.16	2.6	0.03
26. Missouri	6.02	5.84	0.18	3.2	0.03
27. Montana	6.69	6.02	0.67	11.2	0.45
28. Nebraska	6.32	5.92	0.40	6.7	0.16
29. Nevada	5.73	5.67	0.06	1.1	0.00
30. New Hampshire	5.04	5.87	0.83	14.2	0.69
31. New Jersey	4.79	5.58	0.79	14.1	0.62
32. New Mexico	6.13	5.68	0.45	8.0	0.21
33. New York	4.64	5.78	1.14	19.8	1.31
34. North Carolina	5.94	5.85	0.09	1.5	0.01
35. North Dakota	5.55	6.08	0.53	8.8	0.28
36. Ohio	5.65	5.70	0.05	0.9	0.00
37. Oklahoma	7.14	5.90	1.24	21.0	1.54
38. Oregon	7.47	5.84	1.63	27.9	2.65
39. Pennsylvania	5.03	5.88	0.85	14.5	0.72
40. Rhode Island	5.23	5.67	0.44	7.7	0.19
41. South Carolina	5.36	5.79	0.43	7.5	0.19
42. South Dakota	6.19	6.09	0.10	1.6	0.01
43. Tennessee	5.92	5.91	0.01	0.1	0.00
44. Texas	5.82	5.72	0.10	1.7	0.01
45. Utah	6.44	5.41	1.03	19.1	1.06
46. Vermont	6.67	5.92	0.75	12.6	0.56
47. Virginia	4.99	5.68	0.69	12.2	0.48
48. Washington	6.66	5.71	0.95	16.6	0.90
49. West Virginia	5.38	6.07	0.69	11.3	0.47
50. Wisconsin	5.58	5.83	0.24	4.2	0.06
51. Wyoming	5.95	5.97	0.02	0.3	0.00

ability of more than 6 months to the number of whites and Negroes between 18 and 64 years of age. If $\overline{x}'_s$ is the direct estimate for State s, then,

$$\overline{x}'_s = \frac{\text{Number having work disability}}{\text{Number of whites and Negroes}} \times 100.$$

In calculating the direct estimates of work disability, we noted a difference in age intervals in two tables on work disability that appear in the 1970 census report. Table 1, "Persons with work disability" (5), provides data on persons from 18 to 64 years, while table 169, which includes disability data for each State (6), provides such data on persons from 16 to 64. To determine the persons from 18 to 64 years with work disability in each State, the following procedure was used:

$$f = \frac{\text{White population 16–17 years old for State } s}{\text{White population 16–24 years for State } s}$$

$\overline{X}'_{s,w\,16-24}$ = work disability among whites 16–24 years

$f\overline{X}'_{s,w\,16-24} = \overline{X}'_{s,w\,16-17}$

= work disability among whites 16–17 years

$\overline{X}'_{s,w\,16-64}$

$-\overline{X}'_{s,w\,16-17} = \overline{X}'_{s,w\,18-64}$

= work disability among whites between 18 and 64 years for State s.

In the same manner, the number of Negroes 18–64 years of age with work disability was obtained for State s.

Evaluation of the synthetic estimate.
The following methods were used to evaluate the synthetic estimate of work disability:

1. Distribution of absolute differences. The absolute difference is denoted by the absolute value of the difference between the synthetic estimate and the direct estimate:

$$|\overline{x}'_s - \widetilde{\overline{x}}_s|$$

where

$\overline{x}'_s$ = the direct estimate for State s.

$\widetilde{\overline{x}}_s$ = the synthetic estimate for State s.

2. Distribution of percentage absolute differences. The percentage absolute difference used by Levy in his evaluation study (7) is denoted by the absolute difference (as defined in the preceding para-

Table 2. Direct and synthetic estimates, absolute and percentage absolute differences, and mean square errors for complete work disability of more than 6 months for persons 18–64 years of age

State	Direct estimate	Synthetic estimate	Absolute difference	Percentage absolute difference	Mean square error
1. Alabama	5.81	4.75	1.05	22.2	1.11
2. Alaska	1.36	3.15	1.79	57.0	3.22
3. Arizona	4.41	3.76	0.65	17.3	0.42
4. Arkansas	6.86	4.90	1.96	40.0	3.84
5. California	4.17	3.45	0.72	20.7	0.51
6. Colorado	2.98	3.44	0.46	13.3	0.21
7. Connecticut	2.41	3.79	1.39	36.5	1.92
8. Delaware	3.07	3.77	0.71	18.7	0.50
9. District of Columbia	4.35	5.74	1.39	24.2	1.93
10. Florida	4.80	4.31	0.49	11.5	0.24
11. Georgia	5.40	4.43	0.97	21.9	0.94
12. Hawaii	2.16	2.83	0.67	23.6	0.45
13. Idaho	3.77	4.13	0.36	8.8	0.13
14. Illinois	3.38	3.95	0.56	14.3	0.32
15. Indiana	3.27	3.96	0.69	17.5	0.48
16. Iowa	2.95	4.01	1.06	26.4	1.12
17. Kansas	3.19	3.95	0.76	19.1	0.57
18. Kentucky	6.04	4.09	1.95	47.7	3.80
19. Louisiana	5.68	4.60	1.09	23.6	1.18
20. Maine	3.89	4.36	0.46	10.6	0.21
21. Maryland	3.42	3.87	0.44	11.5	0.20
22. Massachusetts	3.22	3.61	0.39	10.9	0.15
23. Michigan	3.83	3.86	0.03	0.8	0.00
24. Minnesota	2.70	3.68	0.99	26.8	0.97
25. Mississippi	6.94	5.21	1.74	33.3	3.01
26. Missouri	4.37	4.13	0.24	5.9	0.06
27. Montana	3.55	4.10	0.55	13.5	0.31
28. Nebraska	2.75	3.99	1.23	30.9	1.52
29. Nevada	2.71	3.59	0.87	24.4	0.76
30. New Hampshire	2.91	4.17	1.27	30.3	1.60
31. New Jersey	3.08	3.57	0.50	13.9	0.25
32. New Mexico	4.15	3.74	0.41	10.9	0.17
33. New York	3.65	4.08	0.43	10.5	0.18
34. North Carolina	4.89	4.53	0.36	7.9	0.13
35. North Dakota	2.85	4.01	1.16	29.0	1.35
36. Ohio	3.76	3.86	0.10	2.6	0.01
37. Oklahoma	5.37	4.17	1.20	28.7	1.43
38. Oregon	3.96	3.98	0.02	0.5	0.00
39. Pennsylvania	4.01	4.20	0.19	4.5	0.04
40. Rhode Island	3.48	3.59	0.12	3.2	0.01
41. South Carolina	5.44	4.64	0.80	17.3	0.64
42. South Dakota	2.74	4.12	1.37	33.3	1.88
43. Tennessee	5.38	4.44	0.93	21.0	0.87
44. Texas	3.66	3.97	0.31	7.9	0.10
45. Utah	3.04	3.26	0.22	6.6	0.05
46. Vermont	3.83	4.38	0.55	12.6	0.30
47. Virginia	3.87	4.08	0.21	5.2	0.04
48. Washington	3.46	3.73	0.27	7.2	0.07
49. West Virginia	7.42	4.68	2.74	58.6	7.53
50. Wisconsin	2.66	3.91	1.25	32.0	1.57
51. Wyoming	3.11	4.07	0.95	23.4	0.91

graph), divided by the synthetic estimate and expressed as a percentage:

$$\frac{|\overline{x}'_s - \widetilde{\overline{x}}_s|}{\widetilde{\overline{x}}_s} \times 100.$$

3. Distribution of mean square errors. As indicated by Gonzalez and Waksberg (3), the mean square error (MSE) of the synthetic estimate can be approximated by

$$(\widetilde{\overline{x}}_s - \overline{x}'_s)^2$$

4. Pearson product-moment correlation coefficient. The correlation coefficient was used to examine the degree of the relationship between the synthetic estimate and the direct estimate.

Limitations of the study. Since the study is dependent on the data on work disability from the 1970 U.S decennial census, its limitation is governed by the census procedure. The number of persons with work disability as determined by the census is based on the ratio estimates from the 5 percent sample inflated to represent the total population. Consequently, this estimate is subject to sampling and measurement errors. The sampling errors for most of the estimates are of a relatively low magnitude. The population data used in the study are based partly on the 100 percent sample and partly on the 15 percent sample of the 1970 census. The data cannot be used to estimate the total number of persons with work disability since the 1970 census question on work disability was asked only of the noninstitutional population 14 to 64 years of age.

Results

Direct estimates, synthetic estimates, absolute differences, percentage absolute differences, and mean square errors for partial work disability of more than 6 months from 50 States and the District of Columbia are shown in Table 1. The synthetic estimates ranged from 5.28 percent for Hawaii to 6.11 for Arkansas, while direct estimates fluctuated between 4.14 percent for Hawaii and 7.50 percent for Idaho. Discrepancies between the synthetic estimate and the direct estimate were small for the District of Columbia, Alabama, Indiana, Kentucky, Kansas, Missouri, Mississippi, Nevada, North Carolina, Ohio, South Dakota, Tennessee, Texas, Wisconsin, and Wyoming. On the other hand, the observed values of absolute differences, the percentage absolute differences, and the mean square errors show that the synthetic estimates were very different from the direct estimates for Arizona, Arkansas, Hawaii, Oklahoma, Utah, Idaho, Oregon, and New York.

Table 2 shows the results of complete work disability of more than 6 months for all States and the District of Columbia. The values of the synthetic estimates fell between 2.83 percent for Hawaii and 5.74 percent for the District of Columbia, while the direct estimates ranged widely from 1.36 percent for Alaska to 7.42 percent for West Virginia. The synthetic estimates for Michigan, Ohio, Oregon, Pennsylvania, Rhode Island, and Virginia were fairly close to the direct estimates. On the other hand, the synthetic estimates for Connecticut, Mississippi, South Dakota, Wisconsin, Alaska, Arkansas, Nebraska, New Hampshire, Kentucky, and West Virginia differed greatly from the direct estimate.

Table 3. Distributions of synthetic estimates and direct estimates of partial work disability

Estimated percentage of persons with partial work disability	Synthetic estimates		Direct estimates	
	f	Cumulative percentage	f	Cumulative percentage
4.2–4.4			1	2.0
4.5–4.7			1	3.9
4.8–5.0			7	17.6
5.1–5.3	1	2.0	4	25.5
5.4–5.6	8	17.6	7	39.2
5.7–5.9	31	78.4	8	54.9
6.0–6.2	11	100.0	10	74.5
6.3–6.5			5	84.3
6.6–6.8			4	92.1
6.9–7.1			1	94.1
7.2–7.4			1	96.0
7.5–7.7			2	100.0
7.8 and more				
Total	**51**		**51**	
Median		5.79		5.87
Mean		5.79		5.85
Standard deviation		0.19		0.73
Skewness		−0.39		0.18

Distributions of estimates. The frequency and cumulative percentage distributions of the synthetic and direct estimates for partial work disability are presented in table 3. The difference in the shape of the distribution between the synthetic estimate and the direct estimate was striking. The variability of the synthetic estimates was small when compared with that of the direct estimates, as can be seen in the standard deviations, which were 0.19 for the synthetic estimate and 0.73 for the direct estimate.

The same pattern as was observed for the distribution of partial work disability was seen for the distribution of complete work disability (table 4). The values of the synthetic estimate for 39 States clustered between 3.6 and 4.5, while the values of the direct estimate for 36 States were concen-

Table 4. Distributions of synthetic estimates and direct estimates of complete work disability

Estimated percentage of persons with partial work disability	Synthetic estimates		Direct estimates	
	f	Cumulative percentage	f	Cumulative percentage
Less than 2.0			1	2.0
2.1–2.5			2	5.9
2.6–3.0	1	2.0	10	25.5
3.1–3.5	4	9.8	10	45.1
3.6–4.0	22	52.9	11	66.6
4.1–4.5	17	86.3	5	76.4
4.6–5.0	5	96.1	2	80.4
5.1–5.5	1	98.0	4	88.2
5.6–6.0	1	100.0	3	94.1
6.1–6.5				
6.6–7.0			2	98.0
7.1 and more			1	100.0
Total	51		51	
Median		4.03		3.70
Mean		4.05		3.93
Standard deviation		0.50		1.28
Skewness		0.65		0.84

trated between 2.6 and 4.5. Accordingly, the standard deviation of the direct estimate was 2.5 times greater than that of the synthetic estimate.

Distribution of absolute differences. The difference between the mean and the median absolute differences was observed to be greater than zero for both partial and complete work disability, as shown in table 5. This result implies that the two distributions were asymmetrical and that their skewness was positive. Consequently, the median may be preferred to the mean as a descriptive measure of the central tendency in these two distributions. Thirty-three States (64.7 percent) have an absolute difference of less than 0.6 in partial work disability, as contrasted with 23 States (45.1 percent) that have an absolute difference of less than 0.6 in complete work disability.

Distribution of percentage absolute differences. Distributions of percentage absolute differences are presented in table 6. It can be seen that the values of the percentage absolute differences for partial work disability concentrated between 0.0 and 15.4 percent, while the corresponding distribution for complete work disability was bimodal, with a major mode at 12.5 percent and a minor mode at 22.5 percent. The median percentage absolute difference was 8.1 percent for partial work disability and 17.3 percent for complete work disability. The mean of the percentage absolute differences was 9.3 percent for partial work disability and 19.6 percent for complete work disability. The variability among States as expressed by the standard deviation of the percentage absolute difference was higher for complete work disability, with 13.3, than for partial work disability, with 6.8. According to the values of the median percentage absolute differences, the agreement between the synthetic estimate and the direct estimate was fairly good for partial work disability, but generally poor for complete work disability.

Distribution of mean square errors. For both partial and complete work disability, the distribution of mean square errors was positively skewed (table 7), and a considerable difference was observed between the average MSE and the median MSE. The median MSE was 0.29 for partial work disability and 0.47 for complete work disability, while the average MSE was 0.44 for partial and 0.97 for complete work disability. For partial work disability, only 18 (35 percent) of the 51 States had MSE's greater than the average MSE (table 1), and for complete work disability, only 17 (33 percent) of the 51 States had MSE's greater than the average MSE. Thus, for both types of disability, the median MSE was more discriptive than the average MSE

Table 5. Distributions of absolute differences between synthetic estimate and direct estimate of partial and complete work disability

Absolute differences	Partial work disability		Complete work disability	
	f	Cumulative percentage	f	Cumulative percentage
Less than 0.1	4	7.8	2	3.9
0.1–0.3	12	31.4	8	19.6
0.4–0.6	17	64.7	13	45.1
0.7–0.9	10	84.3	9	62.7
1.0–1.2	6	91.1	9	80.4
1.3–1.5	1	98.0	5	90.2
1.6–1.8	1	100.0	2	94.1
1.9–2.1			2	98.0
2.2 and more			1	100.0
Total	51		51	
Median		0.46		0.70
Mean		0.54		0.80
Standard deviation		0.39		0.57
Skewness		0.73		1.04

Table 6. Distributions of percentage absolute differences between synthetic estimate and direct estimate of partial and complete work disability

Percentage absolute differences	Partial work disability		Complete work disability	
	f	Cumulative percentage	f	Cumulative percentage
0−5.4	15	29.4	6	11.8
5.5−10.4	17	62.7	6	23.5
10.5−15.4	10	82.0	11	45.1
15.5−20.4	5	91.8	5	54.9
20.5−25.4	3	97.7	9	72.5
25.5−30.4	1	100.0	5	82.3
30.5−35.4			4	90.2
35.5 and more			5	100.0
Total	51	....	51	....
Median		8.1	....	17.3
Mean		9.3	....	19.6
Standard deviation		6.8	....	13.3
Skewness		0.68	....	0.95

as a measure of central tendency.

A measure of the validity of a set of synthetic estimates is shown by the ratio of the square root of the average mean square error divided by the mean of the set of synthetic estimates *(4)*. The mean for all States of the synthetic estimates was 5.79 for partial work disability and 4.05 for complete work disability. The ratio of the square root of the average MSE to the mean of the synthetic estimates was 11.4 percent for partial work disability and 24.2 percent for complete work disability. This result indicates again that the synthetic estimates of complete work disability had greater validity than those of partial work disability.

Correlation analysis. The degree of the relationship between the synthetic estimate and the direct estimate for the entire State was investigated by means of the Pearson product-moment correlation coefficient for raw data and rank-order numbers. As seen in the following table, a fairly high, direct relationship, 0.711, for complete work disability contrasts with a less than moderate, direct relationship, 0.444, for partial work disability.

Data used in calculations	Partial work disability (N=51)	Complete work disability (N=51)
Raw scores	0.444	0.711
Rank-order	0.379	0.687

Each correlation was significantly different from zero at the 0.01 level. The rank-order correlation was slightly lower both for partial and complete work disability than the correlation between the raw scores of the synthetic estimate and the direct estimate, but the same tendency as the result just mentioned was observed. The low correlation between the synthetic estimate and the direct estimate of partial work disability may be attributed to great differences between the shape of their distributions (table 3). Although the synthetic estimates of partial work disability were less biased than those of complete work disability, the synthetic estimates of partial work disability were distributed so densely within a small interval that they did not reflect the size of the direct estimate with any great sensitivity.

Discussion

As mentioned, the synthetic estimate of work disability for a State is determined by summing over all cells the product of the national estimate of the proportion of persons with work disability for a cell and the proportion of the total State population falling into that cell. Consequently, variability among the States with respect to their synthetic estimates of their populations' work disability is due entirely to differences among the States in the proportions falling into the cells used to create the synthetic estimates. Because of the way the synthetic estimates are constructed, they are subject to two types of bias that could affect their

Table 7. Distributions of mean square errors for synthetic estimates of partial and complete work disability

Mean square errors	Partial work disability		Complete work disability	
	f	Cumulative percentage	f	Cumulative percentage
Less than 0.1	12	23.5	7	13.7
0.1−0.5	26	74.5	21	54.9
0.6−1.0	7	88.2	7	68.6
1.1−1.5	4	96.1	6	80.4
1.6−2.0			5	90.2
2.1−2.5	1	98.0		
2.6−3.0	1	100.0		
3.1−3.5			2	94.1
3.6−4.0			2	98.0
4.1 and more			1	100.0
Total	51	....	51	....
Median		0.29	....	0.47
Mean		0.44	....	0.97
Standard deviation		0.57	....	1.34
AMSE		0.66	....	0.98
Skewness		1.95	....	2.72

accuracy: (a) the cells used in creating the estimates may not reflect all the sources of possible variability among the State with respect to work disability and (b) demographic characteristics may have different effects in different States. For example, because of local factors, black males may have a high prevalence of work disability in a particular State, but a relatively low prevalence in another.

To examine the effect of the demographic factors used for the synthetic estimation, two approaches may be taken. The first is to compute synthetic estimates by subgroups, such as race, sex and age, and then compare them with the direct estimates. The second approach is to compute synthetic estimates by applying different population models and then compare the synthetic estimates from one population model with those from the other population model. For example, disability data at the national level and population data at the State level may be broken down by the level of education completed or by socioeconomic status, instead of by type of residence. Since evaluation of the sensitivity of the synthetic estimates according to the types of α-cells used in computing them is essential, we are conducting a study from these two approaches.

Conclusions

Since the distributions of absolute differences, percentage absolute differences, and mean square errors were positively skewed, the median of these evaluation measurements was more descriptive as a measure of the central tendency and as a representative value of a group than the mean was. Consequently, caution should be exercised in the use of the average mean square error (AMSE) in evaluating synthetic estimates if there is reason to suspect that the distribution of differences between the synthetic estimate and the true value of the characteristics being estimated is skewed. For partial work disability, the agreement between the synthetic estimate and the direct estimate was fairly good, while for complete work disability, the agreement was fairly poor. Results of our correlation analysis imply that when synthetic estimation is used for ranking States according to a value of the synthetic estimate, the synthetic estimates of complete work disability probably will be more applicable than those of partial work disability.

References

1. Origin, program, and operation of the U.S. National Health Survey. PHS Publication No. 1000. Vital and Health Statistics, Ser. 1, No. 1. U.S. Government Printing Office, Washington, D.C., April 1965.
2. Synthetic State estimates of disability. PHS Publication No. 1759. U.S. Government Printing Office, Washington, D.C., 1968.
3. Gonzalez, M. E., and Waksberg, J.: Estimation of the error of synthetic estimates. Paper presented at the first meeting of the International Association of Survey Statisticians, Vienna, Austria, August 18–25, 1973.
4. Gonzalez, M. E.: Use and the evluation of synthetic estimates. Paper presented at the 133d annual meeting of the American Statistical Association, New York City, December 27–30, 1973.
5. 1970 census of population. Subject report: Persons with work disability. U.S. Bureau of the Census Report (PC (2)-6C. U.S. Government Printing Office, Washington, D.C., January 1973.
6. 1970 census of population. Vol. 1. Characteristics of the population, for 50 States and the District of Columbia. U.S. Government Printing Office, Washington, D.C., 1973.
7. Levy, P. S.: The use of mortality data in evaluating synthetic estimates. In Proceedings of the American Statistical Association 1971, Social Statistics Section. American Statistical Association, Washington, D.C., 1971, pp. 328–331.

SYNOPSIS

NAMEKATA, TSUKASA (University of Illinois Medical Center), LEVY PAUL S., and O'ROURKE, THOMAS W.: Synthetic estimates of work loss disability for each State and the District of Columbia. Public Health Reports, Vol. 90, November–December 1975, pp. 532–538.

A method of synthetic estimation of health characteristics for local areas, devised by the National Center of Health Statistics, was evaluated in a recent study. In the method, local data on population are combined with national data on a given health characteristic to produce an indirect estimate of that characteristic. The health characteristic selected in the study was that of complete and partial work loss disability. Therefore, synthetic estimates of complete and partial work loss disability were calculated for each State by combining the estimated rates of such disability for the United States, specific to a set of demographic domains, with the data relating to the distribution of each State into this set of demographic domains. The synthetic estimates of complete and partial work loss disability for each State were then compared with the direct estimates available from the 1970 decennial census.

For partial work loss disability, agreement between the synthetic and the direct estimates, as judged by the median percentage abolute difference, was fairly good; for complete work loss disability, agreement was rather poor. On the other hand, the correlation between the synthetic and the direct estimates was higher for complete work loss disability than for partial, partly because the synthetic estimates had a greater spread for complete work disability than for partial. The skewness of the distributions of the squared differences indicates that the evaluation based on median percentage absolute differences was more descriptive than the one based on mean square errors.