

Comparing Web Survey Samples of Schizophrenic and Bipolar Patients with Concurrent RDD and In-Person Samples

LinChiat Chang¹, Channing Stave², Corinne O'Brien², Fred Rappard², Jill Glathar¹, Joe Cronin¹

¹Opinion Research Corporation

²Pfizer

Abstract

Web panel samples of patients with mental disorders face multiple problems, including coverage error, self-selection bias, and physicians cannot validate the purported diagnosis. To evaluate these web samples, we administered the same measures of disability to schizophrenic and bipolar patients via three modes: web, telephone, and paper-and-pencil in the physical presence of a moderator. The telephone samples were recruited via RDD methodology, while the paper-and-pencil samples were recruited from medical databases typically used for qualitative research. Among schizophrenic patients, the web sample yielded significantly higher estimates of disability, and contained higher proportions of young, married, highly educated, White male respondents with full time employment. Mode differences among bipolar patients were much less pronounced. The significant differences in disability between modes were often eliminated after controlling for demographic differences. Comparison of web samples of patients and caregivers across 11 countries indicated that the U.S. web sample of schizophrenic patients produced uniquely discrepant estimates, while all other web samples showed satisfactory convergent validity.

KEY WORDS: Web Survey, Mode Comparison

The Issue of Sample Representativeness

Web panel samples of patients with chronic ailments face multiple problems. As with most web panels, these panels suffer from coverage error because not all patients have Internet access. These panels also suffer from self-selection biases because patients often voluntarily enroll to participate in web surveys. More uniquely, physicians cannot validate the diagnosis of the ailments claimed by these web respondents. Hence, to the extent that mental patients enrolled in web panels are different from those in the general population, this lack of representativeness could undermine inferences drawn from research based on web samples.

Patients with schizophrenia or bipolar patient from a specialty web panel was recruited to participate in the current study. These panelists had opted in to join a general consumer panel, and during initial profiling, had indicated they had schizophrenia or bipolar disorder. The specialty panel was built with respondents reporting one or more ailments on a list of over 200 chronic health conditions. Schizophrenia and bipolar disorder were included in that list of health conditions. Evidently, this specialty panel potentially faces the multiple problems described above: coverage error, self-selection bias, and lack of validation.

To assess the representativeness of these web samples, an RDD telephone survey was implemented in the U.S. to assess demographic characteristics and functionality of patients with schizophrenia or bipolar patient.

Concurrently, a small group of respondents were also recruited to physically show up at specific facilities to complete the questionnaire using paper and pencil, in the presence of a moderator. These respondents were recruited by going to support groups, National Alliance on Mental Illness (NAMI) chapters, posting flyers in doctors' offices and at clinics, and also recruiting directly from clinic databases. To validate their diagnosis, respondents were required to show proof when they came into the offices to complete the surveys. Many of the respondents actually brought medication with them that was prescribed for their purported mental disorder. In essence, this P&P mode was added to examine whether survey estimates would differ between the web samples vs. samples that were validated in person. Both the web and RDD surveys utilized national samples, whereas the P&P survey was conducted with respondents residing in or near 12 major cities: Atlanta, Baltimore, Birmingham, Chicago, Detroit, Knoxville, Los Angeles, Minneapolis, New York, Portland, San Diego, and Seattle.

Over the course of ten weeks, RDD methodology was employed to yield a sample of respondents with schizophrenia or bipolar patient. When compared against available norms from the National Institute of Mental Health (<http://www.nimh.nih.gov/publicat/numbers.cfm>), weighted RDD estimates exactly matched population parameters for the prevalence of schizophrenia (1.1%),

and slightly underestimated the prevalence of bipolar disorder (2.1% vs. 2.6%). These figures suggest that the RDD sample adequately matched available population samples from the other two modes. Among patients diagnosed with schizophrenia, there were significant differences between the web and RDD samples on all

Table 1: Demographic Composition of Schizophrenic Patient Samples

		<i>Proportion within Mode</i>			χ^2 test	
		WEB (n=136)	RDD (n=115)	P&P (n=49)	WEB vs. RDD	WEB vs. P&P
Gender	Female	25%	34%	49%	2.40	9.62**
	Male	75%	66%	51%		
Age	18-24	16%	7%	0%	47.1**	36.1**
	25-34	45%	13%	16%		
	35-44	20%	31%	29%		
	45-54	15%	32%	49%		
	55-64	2%	12%	6%		
	65 or older	1%	4%	0%		
Race	White	85%	61%	59%	21.4**	36.0**
	Black	4%	18%	37%		
	Latino or Hispanic	6%	10%	2%		
	Asian or Pacific Islander	2%	4%	2%		
	Other	3%	7%	0%		
Education	Less than high school	5%	28%	2%	69.7**	38.7**
	High school grad	10%	27%	35%		
	Some college	15%	22%	39%		
	College degree	24%	16%	16%		
	Grad school	46%	6%	8%		
Employment	Employed full-time	60%	23%	22%	53.3**	30.0**
	Employed part-time	13%	11%	18%		
	Volunteer Work, full- or part-time	0%	0%	4%		
	Unemployed	19%	47%	51%		
	Student	4%	0%	2%		
	Retired	4%	19%	2%		
Marital Status	Single and never been married	23%	46%	63%	61.7**	43.9**
	Married	71%	23%	18%		
	Divorced	3%	20%	14%		
	Separated	1%	8%	0%		
	Widowed	2%	3%	4%		

**p<.001

parameters, thus estimates from the RDD sample could be used as benchmarks to evaluate the web samples.

Differences in Demographic Composition

A comparison of basic demographic variables revealed significant differences between the web samples and

observed demographics except gender, and there were significant differences between the web and P&P samples on all demographics. As shown in Table 1, it is apparently that the web sample contained much higher proportions of young, highly educated, married, White male respondents who had full time employment.

Among patients diagnosed with bipolar disorder, there were significant differences between the phone and web samples on age, education, employment, and marital status; and there were significant differences between the web and P&P samples on only age and education. Overall, compared to the differences observed among

Differences in Functionality/Disability

Difference between Modes

Respondents in all three modes (web, phone, P&P) answered five questions on functionality. Functionality

Table 2: Demographic Composition of Bipolar Patient Samples

		<i>Proportion within Mode</i>			χ^2 test	
		WEB (n=154)	RDD (n=301)	P&P (n=51)	WEB vs. RDD	WEB vs. P&P
Gender	Female	65%	55%	63%	3.75	.08
	Male	35%	45%	37%		
Age	18-24	17%	16%	2%	52.2**	9.93*
	25-34	30%	15%	33%		
	35-44	35%	19%	33%		
	45-54	12%	29%	20%		
	55-64	6%	15%	12%		
	65 or older	0%	7%	0%		
Race	White	84%	82%	84%	5.94	5.46
	Black	5%	6%	12%		
	Latino or Hispanic	5%	7%	4%		
	Asian or Pacific Islander	3%	1%	0%		
	Other	3%	4%	0%		
Education	Less than high school	3%	20%	6%	27.4**	11.3*
	High school grad	28%	25%	10%		
	Some college	43%	31%	45%		
	College degree	16%	17%	31%		
	Grad school	10%	7%	8%		
Employment	Employed full-time	40%	26%	41%	52.7**	7.19
	Employed part-time	12%	15%	24%		
	Volunteer Work, full- or part-time	1%	0%	0%		
	Unemployed	32%	43%	29%		
	Student	11%	0%	6%		
	Retired	5%	15%	0%		
Marital Status	Single and never been married	38%	34%	49%	17.7**	5.72
	Married	48%	35%	31%		
	Divorced	12%	24%	20%		
	Separated	1%	3%	0%		
	Widowed	0%	3%	0%		

**p<.01; *p<.05

schizophrenic patients, the differences shown in Table 2 much less dramatic.

was assessed by the Sheehan Disability Scale, which is scored simply by taking the unweighted sum of three

items: (a) “To what extent have emotional symptoms disrupted your work in the last month?” (b) “To what extent have emotional symptoms disrupted your social life in the last month?” (c) “To what extent have three items among bipolar patients. Whenever significant differences emerged, the web sample consistently yielded higher scores (indicating greater disability) than the other two samples. Differences were less dramatic among

Table 3: Survey Estimates of Disability

	<i>Means (standard errors)</i>			<i>t test</i>	
	WEB (n=136)	RDD (n=115)	P&P (n=49)	WEB vs. RDD	WEB vs. RDD
Schizophrenic Patient Samples					
Extent to which emotional symptoms disrupted work	7.1 (.29)	4.7 (.36)	4.0 (.67)	5.36**	4.33**
Extent to which emotional symptoms disrupted social life	6.9 (.26)	4.8 (.36)	4.6 (.50)	4.88**	4.06**
Extent to which emotional symptoms disrupted family life or home responsibilities	7.0 (.26)	4.0 (.34)	4.3 (.49)	7.04**	4.73**
Number of days symptoms cause you to miss school or work or leave you unable to carry out your normal daily responsibilities	2.6 (.20)	2.5 (.26)	1.7 (.34)	.16	2.11*
Number of days felt so impaired, that even though you went to school or work, your productivity was reduced	2.7 (.19)	2.5 (.27)	0.9 (.27)	.59	5.17**
Bipolar Patient Samples					
Extent to which emotional symptoms disrupted work	4.4 (.30)	4.3 (.22)	3.6 (.54)	.20	1.27
Extent to which emotional symptoms disrupted social life	5.7 (.23)	4.7 (.21)	3.9 (.44)	3.17*	3.49**
Extent to which emotional symptoms disrupted family life or home responsibilities	5.7 (.22)	4.5 (.21)	3.9 (.41)	4.04**	3.77**
Number of days symptoms cause you to miss school or work or leave you unable to carry out your normal daily responsibilities	1.4 (.15)	1.7 (.14)	0.8 (.23)	1.48	2.25*
Number of days felt so impaired, that even though you went to school or work, your productivity was reduced	1.8 (.16)	2.0 (.14)	1.3 (.27)	.95	1.48

**p<.01; *p<.05

emotional symptoms disrupted your family life or home responsibilities in the last month?”

T-tests of means between independent groups were conducted to assess if there were significant differences between the web vs. RDD samples, and between the web vs. P&P samples. As shown in Table 3, there were significant mode differences among schizophrenic patients on all three Sheehan items, and on two out of

bipolar patients, but the difference between web and P&P did reach statistical significance nonetheless, $t=2.12$, $p<.05$.

In addition to the Sheehan Disability Scale, two additional items were asked to assess respondents' level of functioning: (a) “On how many days in the last week did your symptoms cause you to miss school or work or leave you unable to carry out your normal daily

responsibilities?” (b) “On how many days in the last week did you feel so impaired, that even though you went to school or work, your productivity was reduced?” As shown in table 3, there was no significant difference between the web and RDD samples on these two measures, but significant differences emerged between the web and P&P samples such that the web samples reported more days lost due to their disorder.

Comparing Estimates against Available Benchmarks

The web samples have consistently produced higher estimates of disability than the other two modes, but it remains unclear which estimates were more accurate. To address this question, we sought out norms available from large scale epidemiological research. One large-scale study that administered the Sheehan Disability Scale on patients screened for bipolar disorder. In one study of bipolar patients, the initial screener was mailed to a representative sample of 127,800 respondents on the NFO panel, from which 3,059 were selected to participate in the main study. Among the selected group, 2,450 (80% response rate) completed the main survey, and 1,167 eventually screened positive for bipolar disorder. Responses from these 1,167 respondents provided norms on the three Sheehan Disability Scale items that could be

we computed the average discrepancy, it became apparent that the web estimates were most discrepant from the available norms. Nonetheless, an average discrepancy of 10% surprisingly low for a self-selected sample with coverage error and no validation of diagnosis.

Unfortunately, the same analyses could not be performed on the schizophrenic patient samples because no norms on the disability scales were available from large-scale epidemiological studies.

Demographics Account for Differences in Disability

Regressions were ran to explore whether the significant differences on the Sheehan items were at least partially attributable to the observed demographic differences. To do so, progressive step-wise OLS regressions were ran for each measure with the following predictors: In Model 1, the predictor is mode (i.e., phone/P&P vs. web) alone. In Model 2, predictors are mode and the main effects of demographic variables, entered as dummy variables (e.g., age = 18-24, age = 25-34, age = 35-44) instead of a single variable (e.g, age) because we cannot assume that the impact of the demographic variables on the DV are definitely linear. In Model 3, predictors are mode, the main effects of demographic variables, and two-way

Table 4: Percent of Respondents Reporting Scores of 8 or more on the Sheehan Disability Scale items

	Norms from Calabrese et al., 2003 (n=1,167)	WEB (n=154)	RDD (n=301)	P&P (n=51)
Extent to which emotional symptoms disrupted work	12%	14%	11%	15%
Extent to which emotional symptoms disrupted social life	14%	29%	28%	18%
Extent to which emotional symptoms disrupted family life or home responsibilities	19%	31%	23%	10%
Average Discrepancy from Norms	n/a	10%	6%	5%

used to assess the quality of estimates obtained from the three samples in this study.

As shown in Table 4, the RDD sample produced estimates that approximated the Calabrese et al. (2003) estimates on the first and third Sheehan scale items, but greater discrepancy can be observed on the second item measuring social life. The P&P sample produced close estimates on the first and second items, but greater discrepancy can be observed on the third item. The web sample produced a close estimate on only the first item, but not the other two items. In sum, none of the samples provided perfect matches on all three items. However, if

interaction effects between mode and demographics. If the significant effect of mode in Model 1 becomes non-significant in Models 2 and 3, this change would indicate that the mode difference is attributable to sample differences in demographics. Indeed, more often than not, analyses revealed that the significant differences between modes were eliminated after controlling for observed demographic differences. This trend suggests that most of the observed differences in disability were due to the sample differences in demographic composition.

Among patients diagnosed with schizophrenia, both mode differences (between web-RDD and web-P&P) on the

extent to which emotional symptoms disrupted work were eliminated after controlling for demographics. Some demographics emerged as significant predictors of differences in functionality. For web vs. RDD, demographics that emerged as significant predictors included the main effects of being unemployed ($\beta=.27$, $p<.001$) or divorced ($\beta=.35$, $p<.001$), that is, disability was higher among schizophrenic respondents who were unemployed or divorced. For web vs. P&P, demographics that emerged as significant predictors included the main effects of being Black ($\beta=.58$, $p<.05$) or having graduate school education ($\beta=1.41$, $p<.05$), that is, disability was higher among Black respondents and respondents who completed at least some graduate school.

For web vs. RDD among patients diagnosed with schizophrenia, the significant mode difference on extent to which emotional symptoms disrupted social life was removed after controlling for demographics. Demographics that emerged as significant predictors included the main effects of having a high school degree ($\beta=.24$, $p<.001$) or a college degree ($\beta=.29$, $p<.001$), that is, disability was higher among schizophrenic respondents who possessed these two levels of educational attainment. However, for web vs. P&P, the significant mode difference remained significant ($p<.05$) after controlling for demographics. In other words, the mode difference on this aspect of functionality is not due to demographic differences between the web and P&P samples.

Both mode differences among patients diagnosed with schizophrenia on the extent to which emotional symptoms disrupted family life or home responsibilities were eliminated after controlling for demographics. For web vs. RDD, the 2-way interaction between mode \times age=25-34 was a significant predictor ($\beta=.87$, $p<.05$), such that in the phone sample, schizophrenic respondents age 25-34 (mean=3.1) were less disabled than other schizophrenic respondents (mean=4.1); whereas in the web sample, schizophrenic respondents age 25-34 (mean=7.6) were more disabled than other schizophrenic respondents (mean=6.5). For web vs. P&P, demographics that emerged as significant predictors included the main effects of being retired ($\beta=.04$, $p<.05$) or having graduate school education ($\beta=.89$, $p<.001$), that is, disability was higher among retired respondents and respondents who completed at least some graduate school.

Among patients diagnosed with bipolar disorder, both mode differences (between web-RDD and web-P&P) on the extent to which emotional symptoms disrupted social life were eliminated after controlling for demographics. For web vs. RDD, demographics that emerged as significant predictors reflected a main effect of age, such that in both samples, bipolar respondents who were age

65 or older reported significantly less disability than younger respondents, all groups $p<.01$. For web vs. P&P, no particular demographic category emerged as significant predictors.

Similarly, both mode differences (between web-RDD and web-P&P) on the extent to which emotional symptoms disrupted family life were eliminated after controlling for demographics. For web vs. RDD, there was again a main effect of age, such that in both samples, bipolar respondents who were age 65 or older reported significantly less disability than younger respondents, all groups $p<.01$. For web vs. P&P, demographics that emerged as significant predictors included the main effects of being retired ($\beta=.17$, $p<.05$) or unemployed ($\beta=.30$, $p<.05$), that is, disability was higher among retired or unemployed respondents.

The same regressions were also ran on the two additional items measuring daily functioning, on which there were some significant differences between the web and P&P samples. Among patients, diagnosed with schizophrenic, the web-P&P mode difference on “number of days symptoms cause you to miss school or work or leave you unable to carry out your normal daily responsibilities” was eliminated after controlling for demographics. Demographics that emerged as significant predictors included the main effects of having high school education ($\beta=.18$, $p<.05$) or having graduate school education ($\beta=.34$, $p<.01$), that is, disability was higher among respondents who completed high school education only and respondents who completed at least some graduate school. The main effect of being single and never been married was also significant ($\beta=.50$, $p<.01$), such that disability was lower among single respondents than other respondents.

In contrast, the web-P&P mode difference on “Number of days symptoms cause you to miss school or work or leave you unable to carry out your normal daily responsibilities” remained significant among schizophrenic patients ($p<.05$) after controlling for demographics. In other words, the mode difference on this aspect of functionality could not be attributed to demographic differences between the web and P&P samples.

Among patients diagnosed with bipolar disorder, the web-P&P mode difference on “Number of days symptoms cause you to miss school or work or leave you unable to carry out your normal daily responsibilities” was eliminated after controlling for demographics. Demographics that emerged as significant predictors included the main effect of being Hispanic ($\beta=.66$,

$p < .05$), that is, disability was higher among Hispanic respondents than other respondents.

Multi-country Web Samples

The mode comparison reported thus far was based only on patient samples within the United States. We also had on hand additional data from the larger worldwide study that predominantly used web samples. Eleven countries were included in the worldwide study: the United States, the United Kingdom, France, Germany, Sweden, Greece, Italy, Spain, Brazil, South Korea, and Australia. In each of 11 countries, the survey was conducted not only on patient samples, but also caregiver samples. Hence, data from the worldwide study allowed us to compare

The key findings are presented in Table 5. The top half of the table displays the mean estimates of disability, while the bottom half of the table shows the proportion of respondents in each of the corresponding demographic category. The third and last column of Table 5 shows the difference between the estimate from the U.S. web sample compared to the estimate derived from the aggregated web sample from the other 10 countries. The trends are very obvious. Discrepancies between the U.S. estimate and the non-U.S. estimate were consistently and remarkably greater for the U.S. sample of schizophrenic patients. All other discrepancies among bipolar patients and caregivers of bipolar patients, as well as caregivers of schizophrenic patients were relatively small; indicating good convergent validity. This set of findings again suggests that the U.S. web sample of schizophrenic

Table 5: Estimates from US vs. Non-US Web Samples

	Schizophrenia			Bipolar Disorder		
	U.S.	Non-U.S.	Difference	U.S.	Non-U.S.	Difference
DISABILITY ESTIMATES						
<u>Caregiver Samples</u>						
Extent to which emotional symptoms disrupted work	5.0	4.2	0.8	3.9	4.3	0.4
Extent to which emotional symptoms disrupted social life	6.4	5.9	0.5	5.9	5.5	0.4
Extent to which emotional symptoms disrupted family life or home responsibilities	6.6	5.9	0.7	6.3	5.7	0.6
<u>Patient Samples</u>						
Extent to which emotional symptoms disrupted work	7.1	4.3	2.8	4.4	4.5	0.1
Extent to which emotional symptoms disrupted social life	6.9	5.2	1.7	5.7	5.5	0.2
Extent to which emotional symptoms disrupted family life or home responsibilities	7.0	5.0	2.0	5.7	5.4	0.3
KEY DEMOGRAPHICS						
<u>Caregiver Samples</u>						
Holds college degree or higher	20%	17%	3%	16%	28%	12%
Employed Full Time	21%	12%	9%	28%	18%	10%
Married	30%	24%	6%	43%	38%	5%
<u>Patient Samples</u>						
Holds college degree or higher	70%	29%	41%	27%	36%	9%
Employed Full Time	60%	24%	36%	39%	33%	6%
Married	71%	24%	47%	48%	34%	14%

estimates from the U.S. samples against estimates from other countries.

patients was a unique aberration.

Summary

Among schizophrenic patients, the web sample contained much higher proportions of young, married, highly educated, White male respondents with full time employment than the phone and in-person samples. Unlike the differences observed among schizophrenic patients, the demographic differences among bipolar patients were much less pronounced. Among schizophrenic patients, the web sample yielded significantly higher estimates of disability than the phone and in-person samples. Differences were again much less pronounced among bipolar patients. These significant differences in disability between the web samples vs. other modes were at least partially attributable to its demographic composition. That is, more often than not, significant differences between modes were eliminated after controlling for demographics.

Examination of web sample data across 11 eleven countries revealed that that the U.S. web sample of schizophrenic patients produced uniquely discrepant estimates, while all other web samples showed satisfactory convergent validity.

Action Taken

Based on the findings from this research, it was determined that the sample of bipolar patients was acceptable but the sample of schizophrenic patients was not. The highly unlikely estimates derived from this web sample of schizophrenic patients on education and employment, in particular, suggested that this web sample was either extremely unrepresentative of the full population of patients with this ailment, or that these respondents were intentionally misrepresenting their demographic profiles. Irregardless, it was difficult to place much confidence in any substantive survey estimates derived from this web sample.

Hence, we replaced the web sample of schizophrenic patients with an entirely new sample, which was recruited and surveyed using the same methodology described above for the paper-and-pencil (P&P) sample. In short, these respondents showed up in person at specific facilities to complete the questionnaire using paper and pencil, in the presence of a moderator. They brought proof of their diagnosis when they came into the offices to complete the surveys.

Compared to the original web sample, the new replacement sample provided estimates that were much closer to estimates from the phone (RDD) and paper-and-pencil (P&P) samples in terms of both disability measures and demographics. There were a few instances in which there was a statistically significant difference in disability

estimates between the new sample and the phone sample; but these mode differences were completely removed after controlling for sample differences in demographic composition.

Substantive findings from this research have been presented at international conferences in psychiatry (McIntyre, Stave, O'Brien, Jina, and Chang, 2007a, 2007b), and are also documented in a manuscript currently under review (McIntyre, O'Brien, Jina, Chang, and Stave, 2007).

Acknowledgements

The UNITE survey was funded by Pfizer, Inc., New York, New York, USA. The authors acknowledge the efforts of GMI, Mercer Island, Washington, USA for provision of the web panel samples; and Opinion Research Corporation, Princeton, NJ, USA for conduct of the survey and analysis of data.

References

- Calabrese, Hirschfeld, Reed, Davies, Frye, Keck, Lewis, McElroy, McNulty, & Wagner (2003). Impact of Bipolar Disorder on a U.S. Community Sample. *Journal of Clinical Psychiatry*, 64:4, 425-432.
- McIntyre, Roger S, Channing Stave, Corrine O'Brien, Anil S Jina, LinChiat Chang (2007a). UNITE (Understanding Needs, Interactions, Treatment, and Expectations): A Global Survey of Carers and Patients with Schizophrenia and Bipolar Disorder. Paper presented at the 15th European Congress of Psychiatry. Madrid, Spain.
- McIntyre, Roger S, Channing Stave, Corrine O'Brien, Anil S Jina, and LinChiat Chang (2007b). A Global Survey of Carers and Patients with Schizophrenia and Bipolar Disorder. Paper presented at the 2007 International Congress on Schizophrenia Research. Colorado Springs, CO.
- McIntyre, Roger S, Corrine O'Brien, Anil S Jina, LinChiat Chang, and Channing Stave (2007). UNITE Survey: Self-reported comorbid somatic disorders and weight gain in patients with severe mental illness and related interactions with medication prescribers. *Manuscript under review.*