

**Practical utility of a new propensity score weighting technique for estimates based on landline-only cases in a dual frame landline/mobile telephone US national alcohol survey**

**Thomas K. Greenfield, PhD; Yu Ye, MS; Katherine J. Karriker-Jaffe, PhD; and William C. Kerr, PhD**

**Alcohol Research Group, Emeryville, CA USA**

**American Public Health Association Annual Meeting**

**Boston, MA, November 2-6, 2013**



Supported by a National Alcohol Research Center grant from the U.S. National Institute on Alcohol Abuse and Alcoholism (P50 AA005595; T. K. Greenfield, PI)



**A L C O H O L  
R E S E A R C H  
G R O U P**

## **Presenter Disclosures**

Thomas K. Greenfield

No relationships to disclose

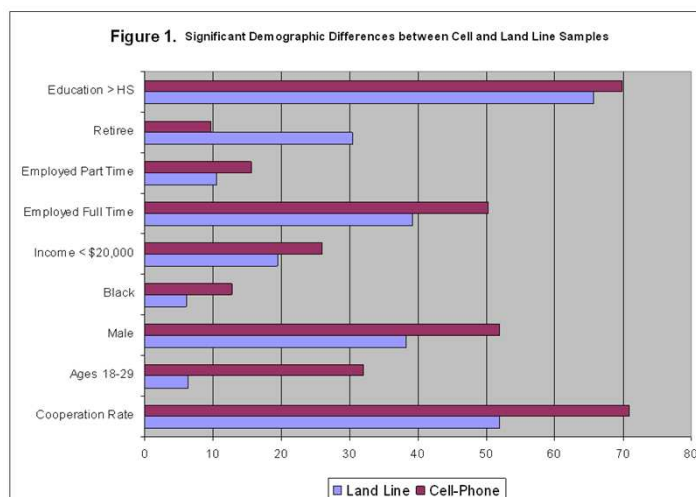
## Background

- ❖ The 2010 National Alcohol Survey (NAS) was a dual-frame Landline ( $n=5,382$ ) and Cell phone ( $n=1,012$ ) computer assisted telephone interview (CATI) survey, using random digit dial (RDD) sampling (97.5% of US households)
- ❖ But because the landline interviews could be longer (averaging 45 minutes) than the cell phone ones (averaging 30 minutes), many items could only be included on the landline telephone instrument.
- ❖ Therefore it was important to develop a method to improve the estimates for the Landline-only cases.
- ❖ Our earlier work found, as expected, strong demographic differences in the Landline and Cell-Phone Samples and the How-likely-to-be-Reached by cell phone or landline subgroups, and in their relationship to alcohol variables emphasizing the importance of supplementing RDD Landline telephone surveys with cell-phone samples.



3

## Differences between Cell and Landline Samples



Note: Unweighted Results; Source: Greenfield et al. (2011) RSA Annual Meeting [Abstract: ACER 35(6), 206A, 2011]



4

## Background

- ❖ Landline-only caseweights were developed, but do not completely solve the problem given that selected individuals with certain characteristics could be reached only by cell phones (as shown in the data)
- ❖ Increasingly some individuals cannot be reached by landlines because they ONLY have cell phones—e.g., younger people and racial/ethnic minorities. (Blumberg SJ, Luke JV (2011) *Wireless Substitution...NCHS*).
- ❖ However, heuristically, there may be some people much like cell-only individuals who DO have landlines and will were reached by a landline
- ❖ Therefore, we wanted to examine whether a propensity score (PS) weighting methods could be used to re-weight Landline cases to provide better estimates than census-based weights in the Landline-only sample.



5

## Background

- ❖ Approximately two-thirds (N=3,025) of the sample were asked questions to ascertain relative use of cell and landline phones.

Group:	Landline Only (n=644)	Landline Mostly (n=532)	Landline = Cell (n=911)	Cell Mostly (n=400)	Cell Only (n=538)	TOTAL (N=3,025)
<b>Reached By:</b>						
Cell Phone	N 0	70	179	132	538	919
Row %	0	7.6	19.5	14.4	39.0	100%
Landline	N 644	461	732	268	0	2,105
Row %	30.6	21.9	34.8	12.7	0	100%
Volume Mean (drinks / month)	11.7	13.5	14.7	23.9	30.5	F <sub>(4,3018)</sub> =16.8***



Source: Adapted from Greenfield et al. (2011) RSA Annual Meeting [Abstract: ACER 35(6), 206A, 2011]



6

## Aim & Research Question

**Aim:** To examine whether propensity score (PS) weighting methods could be used to re-weight landline cases to provide better estimates than census-based weights in the landline-only sample. **Approach:** PS weighting methods were used to reweight the landline sample cases to take account of the 'missing' mobile sample. In the dual-frame sample, a probability for each land-line individual being in the cell phone sample was estimated using sampling weighted logistic regression, with key predictors including demographics and drinking priors (age of drinking onset; parental drinking problems); New weights were constructed in the Landline-only sample reflecting the dual-frame weight and adding to this the additional PS-weight reflecting "cell-phone case propensity."

## Prior Related Research

The response propensity weighting approach was introduced for unit-nonresponse adjustment ([David et al., 1983](#)) after the development of the propensity score theory ([Rosenbaum and Rubin, 1983](#)). This method derives the probability of response based on a logit or probit regression model, the propensity score (PS). The response propensity weighting method has been applied in various surveys to adjust for sample non-respondents ([Rizzo et al., 1996](#); [Sommers et al., 2004](#); [Wun et al., 2007](#))

## Method: Predicting Variables

- Predictor measures, which are used to predict cell phone sample versus the landline sample in constructing the propensity score (PS):
  - gender, age, race/ethnicity, marital status,
  - socioeconomic status (education, family income, employment status),
  - religion (denomination and importance of religion in your life)
  - alcohol-related priors (age of first drinking, family/blood relative alcoholism).

## Method-Validation Measures

- ❖ The validating variables include 12-month drinking measures:
  - (a) current drinking (vs. abstaining),
  - (b) drink-ethanol adjusted volume from beverage-specific items,
  - (c) volume from combined-beverages graduated frequency series,
  - (d) heavy episodic drinking (number of 5+ days in a year).
- ❖ Two alcohol problem variables were also used:
  - (e) social consequences (2+ of 15 items)
  - (f) DSM-IV alcohol dependence (at least one symptom in each of 3 of 7 DSM-IV definitional domains).

## Initial Findings

- Before the PS weights were constructed, we first examined the differences between the cell phone and landline samples for the key predicting variables
- Those reached by Cell phone were more likely to be male, younger, blacks, single, lower income, full time employed or students. They were also less likely to be protestant or catholic, with religion less important in their life. Their age of onset for drinking was somewhat lower and they were more likely to have a blood relative with an alcohol problem.
- After the PS weighting the landline sample was very similar to the cell phone sample in terms of the predictor variable distributions, with none of the measures significantly different

## Further Findings

- With Census-base weighting the average absolute difference across all predictor variables was 5.1% for the Landline-only sample, whereas after PS weighting it was 1.4% for the Landline sample.
- This improved similarity is expected given these predictive variables were used to construct the PS weights.
- Next we examined whether the PS-adjusted weights help improve the landline sample for the alcohol-related validating measures

Comparing distributions of alcohol variables between landline/cell combined sample, landline sample, and landline sample using propensity score weighting (PSW)

	Landline/cell Combined	Landline	Landline w/ PSW	Abs dif <sup>1</sup> : Combined vs Landline	Abs dif <sup>1</sup> : Combined vs Landline w/ PSW
12-Month Drinker (%)	66.0	65.3	65.6	0.9%	0.5%
Drink-Ethanol Adjusted Volume (mean, in drinks)	275	253	272	7.9%	1.1%
Volume from Combined Drink GF (mean, in drinks)	220	210	220	4.6%	0.3%
Number of 5+ Days/Year (mean)	14.1	13.5	14.3	4.0%	1.2%
MEAN FOR DRINKING VARIABLES				<b>4.35%</b>	<b>0.78%</b>
2+ social consequence (%)	3.95	3.81	4.23	3.4%	7.1%
3+ DSM-IV dependence (%)	3.30	3.10	3.51	6.1%	6.4%
MEAN FOR PROBLEM VARIABLES				<b>4.75%</b>	<b>6.80%</b>

<sup>1</sup> Absolute difference = |combined sample value – value from landline sample or landline w/ PSW| / combined sample value

## Summary of Findings

- As shown in the table, for the six variables examined, improvement was found for four of the alcohol consumption variables (current drinking, volume from beverage specific questions, volume from the graduated frequency series, and number of 5+ days). The absolute difference from the 'gold standard' was 4.35% for the landline-only weighted sample and only 0.78% for the PS-weighted landline only sample.
- For the two drinking problem measures results of PS weighting were not as good. Compared to the 'gold standard' combined dual-frame sample estimates, the results for alcohol dependence were about equally different, but the discrepancies were in opposite directions; the PS weighting gave a worse estimate for 2+ consequences than the landline-only sample.

## Summary Conclusions

- Interestingly, the weighted estimates from the dual-frame sample for alcohol consumption and drinking problems were consistently higher compared to the landline-only sample, suggesting the landline-only sample weights may underestimate drinking.
- After applying the PS adjusted weight, this underestimation of drinking is hugely reduced, and closer to the 'gold standard' dual frame estimates.
- However, for drinking problems, there may be somewhat of an over-adjustment.

## Study limitations and last thoughts

- Although a number of our studies have used propensity scores to consider appropriate comparison groups in investigating exposure outcomes, this is our first excursion into propensity score weighting to take account of missing sample. Further work is needed to improve the technique.
- **Analytic challenges**
  - Selection of predictors – what is appropriate?
  - Alcohol consumption variables are better handled by the PS-weighting.
  - We need to understand why alcohol problems are not as well handled and whether we can improve the technique, possibly by using prior alcohol consumption variables when predicting alcohol consequences and dependence.
- **However, the first results appear promising and suggest the PS weighting method can be effective for Landline-only samples**
- **Note that we did not use as a predictor the variable whether respondents also use cell phones and how likely they are to be reached with them (available for 2/3 of the cases because added late)**



## Acknowledgements and Thanks

- ❖ NIAAA Center Grant P50 AA 005595
- ❖ Contact me at [tgreenfield@arg.org](mailto:tgreenfield@arg.org)
- ❖ THANK YOU!



17

## References Cited

- Blumberg SJ, Luke JV (2009) Wireless Substitution: Early Release of Estimates from the National Health Interview Survey, July-December 2008 [Accessed: 2012-06-14. Archived by WebCite® at <http://www.webcitation.org/68Q6oLoEC>], pp 13. National Center for Health Statistics, Hyattsville, MD.
- David M, Little RJA, Samuhel ME, Triest RK (1983) Nonrandom nonresponse models based on the propensity to respond, in Proceedings of the Business and Economic Statistics Section, pp 168-173. American Statistical Association.
- Kass G (1980) An exploratory technique for investigating large quantities of categorical data. *Appl Statistics* 29(2):119-127.
- Kish L (1965) Survey sampling. Wiley, New York.
- Rizzo L, Kalton G, Brick JM (1996) A comparison of some weighting adjustment methods for panel nonresponse. *Survey Methodology* 22:43-53.
- Robins JM, Hernán MÁ, Brumback B (2000) Marginal structural models and causal inference in epidemiology. *Epidemiology* 11(5):550-560.
- Rosenbaum PR, Rubin DB (1983) The central role of the propensity score in observational studies for causal effects. *Biometrika* 70:41-55.
- Sommers J, Riesz S, Kashihara D (2004) Response propensity weighting for the Medical Expenditure Panel Survey – Insurance Component (MEPS-IC), in Proceedings of the Survey Research Methods Section, pp 4410-4417. American Statistical Association.
- Wun L-M, Ezzati-Rice TM, Diaz-Tena N, Greenblatt J (2007) On modeling response propensity for dwelling unit (DU) level non-response adjustment in the Medical Expenditure Panel Survey (MEPS). *Stat Med* 26(8):1875-1884.



18