**Table II. Reporting of essential information to interpret survey quality, number, percentage, median and range (2004-2012) and rationale for reporting**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Percent of publications reporting information** | | **Values of reported information** | **Rational for reporting** |
| **Indicators** | **N (222)** | **%** | **Median (range)** |  |
| Year of survey | 219 | 99 | -- | Useful for any survey in order to know how current data are, to plan future surveys and to compare data from other surveys. |
| Eligibility criteria (minimum of behavior description) § | 222 | 97 | -- | Useful for any survey to determine the denominator being measured, to know measurement for the construction of the social network question needed for RDS analysis. Provides readers with possible criteria to use in different populations and settings; allows for comparison of data across countries. |
| Type of specimen collected for biological testing § | 193 | 87 | -- | Useful for any survey. Informs readers about the types of testing being conducted in different populations and settings. |
| Pre-survey research conducted | 88 | 40 | -- | Useful for any survey. Informs readers about the survey planning process, especially whether attempts were possibly made to assess the underlying network structure of the population. |
| Number of recruitment sites per survey area | 95 | 43 | 1 (1-6) | Especially useful in RDS as it alerts readers to the possible violation of the network being one complete component if participants at each site are not connected; informs readers about the possible clustering (or diffusion) of sample. |
| Interview method | 210 | 94 | -- | Useful for any survey. Provides information about level of confidentiality in the survey (i.e., ACASI may provide more confidentiality) and informs readers about the different types of methods used for interviewing hidden populations in RDS surveys. |
| Number of seeds at the start of the survey | 62 | 28 | 7.5 (1-48) | Specifically useful for RDS surveys. Informs readers about whether many seeds were added during data collection and the number of seeds in relation to the sample size and number of waves (too many seeds may result in too few waves needed to reduce seed dependence, adding too many seeds may be an indication that the population is not well networked); provides parameters for readers about seeds needed in different populations and settings. |
| Whether seeds were added during data collection | 17 | 8 | 5 (1-37) |
| Whether seeds failed during data collection | 14 | 6 | 4 (1-24) |
| Number of seeds by the end of the survey | 121 | 54 | 10 (1-156) |
| Amount or type of primary incentive (USD) | 186 | 84 | 3 (1.9-5) \* | Specifically useful for RDS surveys. Provides readers with parameters about amounts used in different populations and settings, provides an indication of potential bias during recruitment (if incentives are too high, more people may enroll who are not eligible). |
| Having no primary incentive | 6 | 3 | -- |
| Amount or type of secondary incentive (USD) | 177 | 80 | 10 (0.95-20) † |
| Having no secondary incentive | 18 | 8 | -- |
| Calculated target sample size | 89 | 40 | 300 (100-756) | Useful for any survey. Indicates if an original sample size was calculated and if that sample size was reached in order to ensure sufficient power and confidence for data analysis. Provides readers with parameters about sample sizes used in different populations and settings. Specific for RDS surveys: combining multiple survey sites is often a violation of the network being one network component. |
| Final sample size | 212 | 95 | 325 (100-1056)‡ |
| Final sample size for multiple cities combined | 28 | 12 | -- |
| Design effect used for sample size calculation § | 50 | 22 | 2 (1.3-3) | design effects, currently recommended to be at least 2 (16,27,28), are important for calculating a sufficient sample size to account for RDS not being a traditional random sample |
| Maximum number of waves | 95 | 43 | 9 (3-21) |  |
| Duration of data collection (in weeks) | 139 | 63 | 12 (2-124) | Informs readers of the time needed to gather samples of different sizes from different populations and settings; alerts readers of unusual recruitment lengths that may impact representativeness of the sample. |
| Maximum number of coupons distributed to each recruiter § | 163 | 73 | 3 (2-7) | Specifically useful for RDS surveys. The number of coupons used are normally three (7), but some surveys have used more. Analysis does not account for branching induced by the number of coupons provided to each participant so fewer coupons, when possible, is suggested to mimic a random walk process. |
| Whether equilibrium or convergence was assessed§ | 44 | 20 |  | Specifically useful for RDS surveys. Informs readers of seed dependence and is a diagnostic to assess bias. |
| Whether data were adjusted for network size | 157 | 70 | -- | Specifically useful for RDS surveys. Informs readers of the extent to which RDS was fully utilized, resulting in the ability to assess whether the survey may represent the network of the population from which the sample was gathered |
| Software used to adjust data § | 162 | 73 | -- | Specifically useful for RDS surveys. There are limited software packages available for analyzing RDS data. Analyzing RDS data in more popular, preexisting software (i.e., STATA, SPSS) may not eliminate RDS specific biases. |
| Citation for adjustment § | 59 | 26 | -- | Specifically useful for RDS surveys. Given the evolvement of the estimators for the analysis of RDS data, this is useful for providing information about the assumptions supporting the adjustments. |
| Heckathorn 1997, 2002 § | 19 | 32 | -- |
| Salganik & Heckathorn 2004 § | 28 | 47 | -- |
| Heckathorn 2007 § | 10 | 17 | -- |
| Volz & Heckathorn 2008 § | 7 | 12 | -- |
| Gile 2010 § | 4 | 7 | -- |
| Estimator used for adjustment § | 10 | 4 | -- |
| Whether seeds were discarded during analysis | 31 | 14 | -- | Some studies either did not collect data from seeds or did not include their data their analysis, which could likely result in the sample having addition seeds (analysis would assume wave 1 participants are seeds) thereby potentially impacting seed dependence and biasing the final estimate. |

\*Among those that reported a value (n=166). †Among those that reported a value (n=152). ‡Among those that reported a sample size for an individual city (n=185). § Not presented in Table I.

**Table III. Annual rate of change over time (2004 to 2012) based on robust regression of values and reported information.**

|  |  |  |
| --- | --- | --- |
| **Variable** | **Change (95% CI)** | **P-value** |
| Reporting having conducted pre-survey research | 0.04 (0.01, 0.08) | 0.01 |
| Reporting having adjusted RDS data | 0.03 (0.00 , 0.6) | 0.09 |
| Lower eligibility age value | -0.14 (-0.20, -0.07) | 0.00 |
| Number of seeds at start of survey | -0.14 (-0.43, 0.14) | 0.33 |
| Number of seeds at end of survey | -0.31 (-1.86, -0.66) | 0.00 |
| Survey durations | 0.15 (-0.35, 0.66) | 0.55 |
| Final sample sizes | -29.14 (-39.56, -8.73) | 0.00 |
| Sizes of estimated design effects | 0.07 (0.02, 0.12) | 0.01 |
| Length of longest recruitment chain | 0.29 (-0.39, 0.97) | 0.40 |