

# Abortion Patient Survey 2014 (APS 2014)

## Data Users' Guide and Codebook

Data collected by the Guttmacher Institute

Public-use dataset prepared by the Guttmacher Center for  
Population Research Innovation and Dissemination

### Contents

Introduction; about the data .....	1
Contact person .....	2
Available data formats .....	2
Weights and sample design variables .....	2
Key variables .....	2
Imputations .....	3
Codebook .....	4
Programming for recodes.....	5
Extended methodology .....	4
Bibliography of articles based on the dataset .....	8

### Introduction: About the data

This survey of abortion patients is the Guttmacher Institute's fifth in a series and uses a design and questionnaire similar to the four earlier studies, which were conducted in 1987, 1994–1995, 2000–2001 and 2008. One important way the 2014 APS differs from prior APS surveys is that the 2014 did not sample hospital facilities or include hospital patients within the sample. These individuals accounted for 4% of abortions in 2014 and their exclusion likely introduces little bias.

The data are from a nationally representative sample of women obtaining non-hospital abortions in 2014. A total of 8,380 abortion patients provided information about the demographic characteristics of age, race and ethnicity and educational attainment, as well as contraceptive use in the month they became pregnant, health insurance coverage during the time period of the abortion, how they were paying for abortion services, foreign-born status, and how long ago they made the appointment.

The final data file contains 169 variables. All the variables are numeric in type; there are no string variables.

Additional information about the questionnaires, sample design and other facts related to data collection can be found at the end of this document.

## **Contact person**

For inquiries about this dataset, please contact Project Investigator Rachel K. Jones, Ph.D., by email at [rjones@guttmacher.org](mailto:rjones@guttmacher.org) or via phone at (212) 248-1111 x2262.

## **Available data format**

The data are available as a **Stata** data file (.dta).

To obtain the dataset, complete the Guttmacher data request form available on the Data page of the Guttmacher Center's website, <http://www.guttmacher.org/popcenter>.

## **Weights and sample design variables**

The main weight variable is `weight3`.

In order to obtain accurate standard errors, most analyses should also account for the fact that the sample was stratified by using the `strata` variable (strata):

- Stata users: The complex-sample settings (`weight3` and `strata`) can be set the below command:

```
svyset [pweight=weight3], strata(strata)
```

See the extended methodology section below for more information about how the universe of clinics was stratified. Labels identifying strata were suppressed in order to deidentify the data.

## **Key variables**

The unique case identifier is `caseid`.

The unique facility number is `facid`.

Key demographic variables include `age` (age), `insurance` (insurance coverage) and `raceth` (race and ethnicity), `marstat` (marital status), `povcat3` (poverty level), and `gest` (gestation).

Zip code has been deleted in order to deidentify the data.

## **Contraceptive use variables**

Respondents were asked if they had stopped using a method of pregnancy prevention before they became pregnant (hadstop), the last method used (pill, condom, depepro, ring, implant, iud, withdraw, otrostop, otrobstop, neveruse), when they had stopped using the method (stopmo, stopyr, stilluse) and how long they had used it for (monthuse) (Questions 9-12).

These variables were used to construct `ieveruse`, `lastmethod`, `imethod`, `methfail`, `cpfail`, and `istopago` (as well as relevant flag variables). Additionally, the variables `rimplant` and `riud` are revised versions of the variables `iud` and `implant`; the revised variables exclude cases that initially appeared to be contraceptive failures as they were reported to have been used in the month of pregnancy, but information from other variables suggested they were not failures (e.g., respondent also reported pill use in the same month). See also: Jones, 2018 in the below bibliography.

Information about the coding used to create all of the contraceptive variables is available upon request.

## **Imputations**

Variables whose labels include the word “imputed” are ones where missing values were imputed. Variables with names ending in “flag” can be used along with the associated imputed variable (i.e., the one with the same the variable name) to generate the original (i.e., pre-imputation) variable.

## Extended methodology

### *Data collection*

This survey of abortion patients is the Guttmacher Institute's fifth in a series and uses a design and questionnaire similar to those for earlier studies conducted in 1987, 1994–1995, 2000–2001 and 2008. One important way the 2014 APS differs from prior APS surveys is that the 2014 survey did not sample hospital facilities and hospital patients are excluded. There individuals accounted for 4% of abortions in 2014 and their exclusion likely introduced little bias.

We developed a four-page questionnaire, collecting information about demographic items including age, race and ethnicity, insurance coverage and educational attainment, as well as contraceptive use in the month the respondent became pregnant. In addition, the 2014 APS assessed several new issues, including how long ago the patient had made the appointment and reason for choosing the facility where the abortion was obtained. All core demographic and contraceptive methods items were asked of all respondents. We explored alternate wording for three variables--current school enrollment (question 15), prior abortions (question 25) and pregnancy timing (question 34)--to see if they resulted in different response patterns. Each version was only asked of one-half of respondents (Module A and Module B). Within each facility, every other woman received Module A and every other woman received Module B. The questionnaires are included in a separate document.

The facilities in the survey were sampled from all clinics and physician's offices where abortions were performed in 2011, according to information from the Guttmacher Institute's 2011 Abortion Provider Census. The universe was stratified by 2011 caseload rounded to the nearest 10 (30–390 abortions, 400–1,990, 2,000–4,990, and 5,000 or more) and whether they were affiliated with any of several national organizations (e.g., the National Abortion Federation) for women's reproductive health; facilities were then listed by census region and state within each stratum. Facilities that reported fewer than 25 abortions in 2011 were not included because of the high likelihood that they would perform few or no abortions during the survey period. Their exclusion could cause little bias regarding the representativeness of women obtaining abortions because these facilities only accounted for 1% of all reported procedures in 2014. Next, every *n*th facility was sampled (*n* varied by stratum). Clinics with large caseloads were oversampled to obtain adequate representation of the variety of facilities in the sample.

Each facility was assigned a sampling period that was inversely proportional to its probability of being selected. Facilities were asked to administer the questionnaire to all women who obtained an abortion during a specified fielding period, which ranged from two weeks in the largest clinics to 12 weeks in the smallest facilities. If a facility declined to participate or did not obtain usable questionnaires from at least half of abortion patients, it was replaced by the next facility listed in its stratum, which in most cases was in the same state or in a neighboring state in the same region. Our goal was to recruit 113 facilities; the final sample was obtained from patients at 87 facilities (77% of the original goal). An additional 123 facilities were approached, but did not participate. Twenty of these facilities were found to no longer be providing services; 25 agreed to participate, but were unable to adhere to the study protocols or recruit a sufficient number of patients into the study; and 78 declined to participate. In all, 87 of the 190 active facilities (46%) approached agreed to participate.

The questionnaire, available in both English and Spanish, was distributed to women by facility staff. Participating facilities decided when during the patient's visit to distribute the questionnaire; in most cases, women completed it along with other paperwork while they waited for their procedure. The questionnaire included an introduction explaining the purpose of the survey and informing women that participation was voluntary and anonymous. The questionnaire and procedures were approved by the Guttmacher Institute's federally registered Institutional Review Board.

Participating facilities reported performing 11,024 abortions during the sampling period. Usable questionnaires were obtained from 8,380 patients, for a response rate of 76%. Facility staff supplied information about age, race, ethnicity, insurance coverage and method of payment for 1,066 of the patients who did not complete the questionnaire. (Reasons women did not complete the questionnaire included refusal to participate, failure of the clinic to distribute questionnaires and lack of time to complete the questionnaire.) No information was available for the remaining 1,578 women.

As in prior surveys, in order to correct for any bias produced by deviation from the original sampling plan and nonresponse, a three-stage weighting process was followed. First, individual weights were developed to adjust for the demographic characteristics of the 1,066 nonrespondents for whom the facility staff provided information. Second, facility-level weights adjusted for the other 1,578 nonrespondents for whom no demographic data were available. Third, stratum weights were constructed to correct for departures from the number of facilities to be sampled in each grouping by caseload and provider type. With the final weight adjusted to a mean of 1.0, the standard deviation is 0.24 and the range is 0.50 to 2.00.

Nonresponse on individual items was around 2% for most questions but ranged from 0.8% for age to 13% for family income. Missing information on key demographic variables was imputed on the basis of the responses of other women with similar characteristics using a "hot-deck" procedure. Specifically, we used cross-tabulations to identify the variables most strongly associated with each item requiring imputation. Respondents were sorted according to these variables in the order of the strength of the item's association with the variable to be imputed, so that similar cases were adjacent to one another in the file. A missing value was then replaced by the value of the preceding case in the file.

### *Data issues*

While many of the survey items were adopted directly from the previous patient surveys conducted by the Guttmacher Institute, several were revised to improve accuracy.

*Race and ethnicity.* To measure race and ethnicity, the 2014 Abortion Patient Survey adopted items from the 2013 Current Population Survey. Respondents were first asked "Are you Spanish, Hispanic or Latina?" and could answer yes or no (question 2). This was followed by "Please choose one or more races that you consider yourself to be." Six response categories were available: "American Indian or Alaska Native," "Asian," "Native Hawaiian or Pacific Islander," "black or African American," "white" and "other." As in our 2008 Abortion Patient Survey (but unlike in the Current Population Survey), the last response category provided a space to write in a specific race. The 2014 survey differed from the 2008 in that it allowed patients to choose more than one racial group; the prior survey asked patients to choose the racial group that best described them. To make the surveys comparable on this key characteristic, the current analysis

used a measure of race and ethnicity that was comparable to that of the 2008 survey; even though the earlier survey had encouraged respondents to provide only one race, 118 respondents indicated that they identified with multiple groups (e.g., checked more than one race). In keeping with prior surveys, we constructed a measure of race in which patients who indicated multiple races were typically classified as belonging to the least common of the racial groups checked off. Hence, respondents were classified as a specific racial group besides “other” when possible (e.g., a written response of “Chinese” was coded to “Asian”). Also in line with prior surveys, patients who checked off both “black” and one or more other racial groups were classified as black. In the combined measure used in this analysis, Hispanic ethnicity was given priority over any racial category. (Most commonly, 45% of Hispanic respondents indicated they were white, and 37% indicated “other” race.)

Slightly fewer than 5% of respondents identified with more than one race. When compared to the traditional measure, allowing for this option reduced the proportion of abortion patients who identified as black from 28% to 25%, and the proportions who were Asian and “other” were each reduced by about one percentage point. The more nuanced measure of race, which allows for more than one racial designation, will be used in subsequent analyses that do not require comparisons over time.

*Health insurance coverage.* The Affordable Care Act went into effect in January 2013, and introduced state-based marketplaces for individuals to shop for private health insurance. As such, we modified the health insurance item to include this as an answer category for respondents; specifically, patients were provided with the option of indicating that they obtained coverage through HealthCare.gov or a state-run health exchange (question 4). Of the 527 respondents who indicated they obtained coverage through HealthCare.gov or a state exchange, 252 also said they had coverage under a private plan or under Medicaid; these cases were coded to the more specific plan. However, we were unable to classify 275 respondents and they are coded; these individuals may have had private insurance or may have been covered by state health insurance (Medicaid).

An additional 149 respondents (1.7% of the sample) indicated coverage under multiple plans. Our measure of insurance allows for one type of coverage and priority was given to private insurance, followed by other, Medicaid and, finally, HealthCare.gov or the state health exchange. That is, if an individual indicated they had private insurance and some other type of coverage, they were coded to have private coverage.

More so than for characteristics such as race and age, insurance coverage is subject to reporting error; for example, some individuals may not know which type of health insurance coverage they have, especially if it was obtained through a parent, spouse, domestic partner or other family member. In the context of health care reform, these types of errors may be more pronounced.

*Income and poverty.* To construct our measure of family income, we asked abortion patients to report their total family income before taxes in the previous year, and the number of family members in their household at the time of the abortion (questions 21 and 22). This information was used to calculate three income categories of less than 100%, 100–199% and 200% or more

of the federal poverty level. We used these categories to refer to the patients who fall within them as poor, low-income and highest-income.

Both individual and family income levels are difficult to measure on surveys because these items often suffer from lower response rates than other types of questions. A higher level of nonresponse for this item (13%) may be the result of resistance to disclosing income, or to the fact that some patients (e.g., those living with parents and other adult family members) do not know their annual family income. However, as in 2008, the 2014 survey provided 12 annual income categories listed in \$5,000 increments (ranging from less than \$9,999 per year to \$75,000 or more per year), with weekly incomes given parenthetically to serve as a more tangible guide to assist with estimates. We do not believe that the accuracy of this measure changed over time.

*Religious Affiliation.* Our measure of religious affiliation was adopted from the National Survey of Family Growth (NSFG). Respondents were asked “What religion are you now, if any?” (question 17), and a follow-up item determined if they were fundamentalist (question 18). Following the NSFG, we asked about four categories of evangelism, but for purposes of this survey, we collapsed them into one category. Patients who selected “other” religion were asked to specify which religion, and 991 of the 1,239 eligible did so. In line with the NSFG, we coded patients who wrote in that they were Christian (no denomination given) as Protestant. Our measure of religious affiliation in the 2008 and 2014 surveys distinguishes between mainline Protestants, evangelical Protestants, Catholics, those affiliated with some other religion and those with no religious affiliation. (Individuals who indicated that they were evangelical but affiliated with Catholicism or some other religion were not included in our measure of evangelicals.)

*Sexual Orientation.* In recognizing that not all abortion patients have sex only with men or identify as straight or heterosexual, sexual orientation was measured for the first time in 2014. We adapted an item from the NSFG audio computer-assisted self-interview module that asked “Do you think of yourself as...” and provided three response categories: “heterosexual or straight,” “homosexual, gay or lesbian” or “bisexual” (question 32). In our modification of this question, we added a “something else” category, and allowed patients to write in a response. While 81 respondents, or 1%, chose this option, none of the write-in answers achieved enough responses to justify analysis as its own category. Nine percent of respondents did not answer this item, higher than for standard demographic items, likely because of its sensitive nature. Gender identity was not measured.

## Bibliography of articles using the dataset

“Sexual orientation and exposure to violence among U.S. abortion patients” Jones RJ, Jerman J, Charlton BM, *Obstetrics & Gynecology*, in press.

“Reported contraceptive use in the month of becoming pregnant among U.S. abortion patients in 2000 and 2014” Jones RK, *Contraception*, 2018, 97(4): 309-312, doi: 10.1016/j.contraception.2017.12.018

“Changes in population group abortion rates between 2008 and 2014 and lifetime incidence of abortion” Jones RJ, Jerman J, *American Journal of Public Health*, 2017, 107(12): 1904-1909, doi: 10.2105/AJPH.2017.304042 .

“Which abortion patients have had a prior abortion? Findings from the 2014 U.S. abortion patient survey” Jones RJ, Jerman J, Ingerick M, *Journal of Women's Health*, 2017, doi: <https://doi.org/10.1089/jwh.2017.6410>

“Characteristics and circumstances of US women who obtain very early and second-trimester abortions” Jones, RK and Jerman, J, PLOS ONE, 2017, doi:10.1371.

“Time to appointment and delays in accessing care among abortion patients.” Rachel K. Jones and Jenna Jerman. New York, Guttmacher Institute, 2016.  
<https://www.guttmacher.org/report/delays-in-accessing-care-among-us-abortion-patients>

“Characteristics of U.S. abortion patients in 2014 and changes since 2014.” Jenna Jerman, Rachel K. Jones & Tsuyoshi Onda. New York: Guttmacher Institute, 2016.  
<https://www.guttmacher.org/report/characteristics-us-abortion-patients-2014>