**Data Quality Statement: 2005-2021 Cancer Institute Tobacco Tracking Survey**

 Published date November 2021

**Report prepared for:**

Cancer Institute NSW
Level 9, Central Avenue
Australian Technology Park
Eveleigh  NSW  2015

**Report prepared by:**

Dr Paul Myers (Director, Quantitative Research Consulting)
Matthew Gidman-Rowse (Research Consultant)
Gillian Meikle (Project Manager)

The Social Research Centre
Level 5, 350 Queen Street
MELBOURNE VIC. 3000
Tel:  (613) 9236 8500
Fax: (613) 9602 5422
Email:  info@srcentre.com.au

Version:  1.0 – November 2021

**1.Introduction**

This data quality statement provides a summary of the context, methodology and data collection practices of the Cancer Institute Tobacco Tracking Survey (CITTS) between the years 2005 and 2021.

This data quality statement has been prepared by the Social Research Centre on behalf of the Cancer Institute NSW (the Institute). This report is based on the systematic review of documents relating to the methodology and data collection practices of the 2005–2021 CITTS. It was noted, however, that details about the conduct of the CITTS was not available for all years. As such, relevant members of the Institute were consulted to 'fill in the gaps' where required. Instances where information was unavailable for particular years have been noted throughout the report.

**2.Institutional Environment**

The Institute is a NSW Government agency charged with substantially improving cancer control through the NSW Cancer Plan (the Plan). The major goal of the Plan in the area of tobacco control is to reduce the incidence of cancer through reducing smoking prevalence in NSW.

Social marketing campaigns are one of many strategies undertaken to achieve this goal. Under the Plan, the Institute assumes primary responsibility for mass media campaigns in the area of tobacco control.  Since 2005, the Institute has implemented a large number of campaigns[[1]](https://www.cancer.nsw.gov.au/cms/getdoc/eff15ca2-be1c-43fd-9db0-c2d36eabe27b/citts-data-quality-statement.aspx?viewmode=3&showpanel=1&cmscontentchanged=false&lang=en-AU&langobjectlifetime=request" \l "_ftn1" \o ") with the aim of promoting cessation among smokers in NSW.  The target audience for the campaigns has typically been all NSW residents aged 18 years and over who currently smoke tobacco or have quit smoking in the past year (recent quitters).

The 2005-2021 CITTS was undertaken to monitor and evaluate the effectiveness of these campaigns.

The Institute is funded by the NSW State Government; is governed by the Cancer Institute NSW Board; and works with the NSW Ministry of Health, Local Health Districts, Specialty Health Networks and the pillars (Agency for Clinical Innovation, the Bureau of Health Information, Clinical Excellence Commission, Health Education Training Institute and NSW Kids and Families) to improve health outcomes across the State.

**3.Relevance**

**3.1.Research objectives**

The objective of the 2005–2021 CITTS is twofold: evaluate the effectiveness of various anti-tobacco campaigns in terms of overall campaign awareness and impact; and monitor smoking and quitting-related cognitions and behaviours among smokers (and recent quitters) in NSW.

**3.2.Survey overview**

The CITTS is a cross-sectional continuous tracking survey that commenced in April 2005. The survey has been conducted for up to 50 weeks a year with weekly quotas in place.

The 2005–2021 CITTS saw a number of methodological changes implemented in an effort to improve the relevance and accuracy of the resultant data. Key changes related to: sample frame; sample size; in-scope population definition; and approach to weighting.

The 2005-2012 CITTS was conducted via computer assisted telephone interviewing (CATI) using various landline sample frames. The 2013–2020 CITTS was conducted via a dual-frame sample frame – that is, a survey that utilised both landline and mobile phone numbers. From October 2020 the CITTS was conducted via a mobile only sample frame.

The in-scope population for the CITTS was initially smokers aged 18 years and over who are resident in three television markets: Metropolitan Sydney, Northern NSW and Southern NSW (which included Canberra). However, in 2007 the in-scope population was expanded to include recent quitters (i.e. those who had ceased smoking in the past 12 months) aged 18 years and over. The boundaries used to define the television markets for the 2005–2006 CITTS are not available. The boundaries used throughout 2007–2021 were obtained from the now defunct Regional TV Marketing.

The CITTS has been undertaken by a number of research agencies on behalf of the Institute over the years: BlueMoon Research and Planning (with data collection undertaken by Surveytalk) in 2005-March 2007; McNair Ingenuity Research in April 2007–2012; and the Social Research Centre in 2013–2021.

**3.3.Sample design and size**

**3.3.1.       Overview**

The sample design and size for the 2005–2021 CITTS was premised on weekly quotas for achieved sample size stratification of the weekly sample by NSW television market: Metropolitan Sydney, Northern NSW and Southern NSW. The implementation of these design criteria varied across the years, with the largest changes occurring with a change in contracted research agency. An overview of the weeks of data collection and total sample size by year is shown in Table 1.

**Table 1            Achieved sample size and weeks of data collection by year**

|  |  |  |
| --- | --- | --- |
| **Year** | **Weeks** | **Annual Sample Size** |
| 2005 | 34 | 1,543 |
| 2006 | 38 | 1,600 |
| 2007 | 50 | 2,604 |
| 2008 | 50 | 2,501 |
| 2009 | 50 | 2,500 |
| 2010 | 50 | 2,553 |
| 2011 | 50 | 2,500 |
| 2012 | 50 | 2,502 |
| 2013 | 50 | 2,522 |
| 2014 | 50 | 2,046 |
| 2015 | 50 | 2,041 |
| 2016 | 50 | 2005 |
| 2017 | 50 | 2004 |
| 2018 | 50 | 2003 |
| 2019 | 50 | 2000 |
| 2020 | 50 | 2000 |

**3.3.2. 2005-Mar 2007**

The sample design for the 2005-March 2007 CITTS consisted of disproportionate quotas by location, such that an equal number of interviews were conducted within each television market (Metropolitan Sydney, Northern NSW and Southern NSW); proportion to smoking population demographic quotas (according to age, sex and socio-economic status) within each television market.

Socio-economic status was based on the occupation of the main income earner in the household (not the respondent) and coded to a modified version of the 2006 Australian Standard Classification of Occupations (ANZCO). The source of smoker benchmarks is not available. The weekly sample size target was typically 40 interviews each week, although was increased on occasion.

**3.3.3. April 2007–2012**

The sample design for the 2007–2012 CITTS involved geographic stratification by television market in proportion to the population. The source of the benchmarks on which the sample design was based or the exact targets set are not available.

The sample size target for the 2007–2012 CITTS was 50 interviews per week. The only variations to this design were establishing a target of 280 interviews for Week 14 2007 and increasing the sample size target to 100 interviews for Week 50 2010.

**3.3.4. 2013–2020**

The 2013–2020 CITTS saw the introduction of a dual frame sample design – this involved allocating a portion of the overall sample to a separate mobile phone stratum, thereby enabling ‘mobile phone only’ respondents to be included in the survey.

The sample design for the 2013–2020 CITTS accounted for the following considerations:

* Establishing strata with hard quotas for landline and mobile phone sample frames. For Weeks 1–12 2013 the landline stratum was allocated 50 interviews per week from a weekly total of 70 interviews. This was implemented to maintain the existing time series, while evaluation of adopting a dual frame design was undertaken. For Weeks 12–25 2013 the weekly number of interviews was reduced to 50, with a 60:40 split between the landline and mobile sampling frames. The weekly number of interviews was then reduced to 40 from Week 26 2013, with a revised 50:50 split between the landline and mobile sampling frames.
* Stratifying the landline sample in proportion to population based on television market divisions: 60.6% Metropolitan Sydney, 21.3% Northern NSW and 18.1% Southern NSW.  Geographic based strata were not implemented for the mobile stratum.

Geographic strata were not put in place for the random-digit dialling (RDD) mobile stratum (as no a priori geographic information was available for mobile sample records).The sample size target for Week 50 2014 and Week 50 2015 was increased to 80 interviews.

**3.3.4. October 2020 onwards**

From 2020 October the CITTS saw a transition to a mobile only dual frame sample design where the landline stratum was replaced with a listed mobile stratum.

The revised sample design accounted for the following considerations:

* 50:50 split between the listed mobile and RDD mobile sampling frames.
* Stratifying the listed mobile sample in proportion to population based on television market divisions: Metropolitan Sydney (60.6%), Northern NSW (20.6%) and Southern NSW (18.9%).
* Geographic quotas were not put in place for the RDD mobile stratum. RDD mobile numbers identified as not being in NSW or ACT by the sample provider SamplePages (through matching postcode information where available with a small number of RDD mobile records) were pre-screened as ‘Out of scope’ and retained in the sample frame.

**3.4.Ethics approval and quality assurance**

The CITTS was originally reviewed and approved by the NSW Population and Health Services Research Ethics Committee in 2010 (HREC/10/CIPHS/13). Approval is in place until 30 June 2025. The research agencies contracted to undertake the CITTS adhered to various industry quality accordance practices and policies applicable at the time.

**4.Timeliness**

The CITTS was designed as a continuous tracking survey and therefore, data has essentially been continuously collected since 4 April 2005. Exceptions were:

* A four week fieldwork pause between 8 August and 4 September in 2005
* No data collection during January to March 2006, as anti-tobacco campaigns were not run during this period
* Two week data collection pause over Christmas and New Year period each year.

The frequency with which survey data was provided to the Institute by research agencies varied somewhat over the years.  The following provides further details.

* Between 2005 and March 2007 survey data was provided at the end of each calendar year.
* Between April 2007 and 2013 survey data was provided on a biannual basis following the completion of Week 25 and Week 50 of fieldwork.
* Between 2014 and 2021 final survey data was provided on a biannual basis to the Institute; however, interim data was also provided at the conclusion of certain campaigns to facilitate timely campaign evaluation and reporting. From 2019 onwards a live dashboard was implemented with select survey results visible weekly.

**5. Accuracy**

**5.1. Sample frame and sampling**

The sample frame for the CITTS was the Electronic White Pages (EWP); however, it moved to Random Digit Dialling (RDD) in 2007.  Since then, various RDD sample frames have been used. Table 2 outlines the nature and commercial provider of the sample frame by year.

**Table 2            Sample frame and provider by year**

|  |  |  |  |
| --- | --- | --- | --- |
| **Year** | **Sample frame** | **Telephone type** | **Sample provider** |
| 2005 | Electronic white pages | Landline | Not documented |
| 2006 | Electronic white pages | Landline | Not documented |
| 2007 | In-house RDD | Landline | In-house |
| 2008 | List assisted RDD | Landline | SamplePages |
| 2009 | List assisted RDD | Landline | SamplePages |
| 2010 | List assisted RDD | Landline | SamplePages |
| 2011 | List assisted RDD | Landline | SamplePages |
| 2012 | List assisted RDD | Landline | SamplePages |
| 2013 | Exchange based RDD | Landline and mobile | SampleWorx |
| 2014 | Exchange based RDD | Landline and mobile | SampleWorx |
| 2015 | Custom RDD | Landline and mobile | SamplePages |
| 2016 | Custom RDD | Landline and mobile | SamplePages |
| 2017 | Custom RDD | Landline and mobile | SamplePages |
| 2018 | Custom RDD | Landline and mobile | SamplePages |
| 2019 | Custom RDD | Landline and mobile | SamplePages |
| 2020 | Custom RDD | Landline and mobile | SamplePages |
| Oct 2020 – Jun 2021 | Custom RDD & Listed mobile | Mobile only | SamplePages |

**5.1.1. Electronic white pages: 2005–2006**

The sample frame for the 2005–2006 CITTS was the Electronic White Pages (EWP). Although not documented, it is likely the sample frame for the 2005–2006 CITTS was the 2004 EWP. Further, it is unknown which provider was used to access the EWP.

**5.1.2. In-house RDD: 2007**

In 2007, the CITTS adopting a Random Digit Dialling (RDD) sample frame. Details regarding the source and exact nature of the sample frame used for 2007 is unavailable. However, records indicate a form of 'list-assisted' or 'known blocks' sampling methodology was used to generate telephone numbers, which was most likely undertaken by the research agency of the time. The basic premise of the ‘list-assisted’ approach is to randomly generate telephone numbers from 'seed' numbers obtained from available phone number lists.

**5.1.3. List assisted RDD: 2008–2012**

The list-assisted RDD sample frame sourced from the commercial sample provider SamplePages and brokered by the Association for Market and Social Research Organisations (AMSRO). SamplePages withdrew its list-assisted RDD product in late 2009, instead replacing it with a 'listed' sample frame. Information on how this change was managed for the CITTS is unavailable.

**5.1.4. Exchange based RDD: 2013–2014**

The 'exchange based' (or ‘allocated numbers’) RDD sample frame sourced from the commercial sample provider SampleWorx. The building blocks for the sample frame were exchange prefixes published by the Australian Communications and Media Authority (ACMA). Numbers are then randomly generating numbers that fall within these ranges and tested with the exchange to verify whether the number is currently a valid or invalid number. Valid numbers are allocated to a pool of numbers from which sampling is undertaken. Sample frames for landline and mobile telephone numbers were generated in the same fashion.

**5.1.5. Custom RDD: 2015-2020**

The custom RDD sample frame sourced from the commercial sample provider SamplePages. The essence of the custom approach is that landline and mobile phone numbers were randomly generated from exchange prefixes published by ACMA and tested at the time of sample generation, rather than being drawn from a pre-existing pool of numbers. Landline and mobile telephone numbers were generated in the same fashion.

**5.1.6. RDD Mobile and Listed Mobile: 2020 Onwards**

The RDD mobile and listed mobile sample frame sourced from the commercial sample provider SamplePages. The generation of RDD mobile phone numbers is the same as detailed in section 5.1.5. The listed mobile numbers are selected at random from SamplePages’ list of mobile numbers, verified as belonging to NSW or ACT residents. These numbers are sourced from a composite phone database built by contributors from different organisations, including charities, telemarketing companies and other business entities. The list is updated monthly.

**5.2. Respondent selection**

The in-scope population for the 2005–2006 CITTS was daily, weekly, or less often than weekly smokers aged 16 years and over, resident in NSW. For 2007–2015, the in-scope population was NSW residents aged 18 years and over who currently smoke tobacco (daily, weekly, or less often than weekly) or have quit in the past year (recent quitters).

Various respondent selection process were used between 2005–2015:

* For 2005–2008 CITTS, households were initially screened for the number of smokers and recent quitters. Throughout 2005–2006 in cases where two or more in-scope persons were present in the household the “most recent birthday” method was used. In the 2007–2008 CITTS when more than one smoker or recent quitter resided in a household a random selection (via eldest, second eldest etc.) was undertaken.
* For 2009–2012 households initially screened for the number of people in the household regardless of smoking status. A random selection was then undertaken (via eldest, second eldest etc.), following which the selected respondent was screened for smoking status.
* For 2013–2020, different respondent selection processes were implemented for the landline and mobile sample frames. With the landline sample, the “next birthday” method was used to randomly select respondents from households where two or more in-scope persons were present. The phone answerer was the selected respondent for the mobile phone sample.
* From October 2020, the CITTS had a single selection process, where the phone answerer was the selected respondent for the mobile phone sample frames.

**5.3.Response rates**

Response rates were not available for 2005–2006. For the 2007–2012 CITTS, response rates were calculated according to the American Association of Public Opinion Research (AAPOR) standard. Table 3 shows response rates based to the total sample for the 2007–2015 CITTS.

**Table 3            AAPOR Response Rate 3 by year and sample frame**

|  |  |  |
| --- | --- | --- |
| **Year** | **Sample frame** | **RR3** |
| **2007** | Landline | 48.8% |
| **2008** | Landline | 52.1% |
| **2009** | Landline | 21.2% |
| **2010** | Landline | 22.6% |
| **2011** | Landline | 27.0% |
| **2012** | Landline | 25.2% |
| **2013** | Combined | 43.3% |
|   | Landline | 58.6% |
|   | Mobile | 36.3% |
| **2014** | Combined | 42.9% |
|   | Landline | 55.4% |
|   | Mobile | 38.9% |
| **2015** | Combined | 40.1% |
|   | Landline | 48.7% |
|   | Mobile | 36.6% |
| **2016** | Combined | 41.3% |
| **2017** | Combined | 38.3% |
| **2018** | Combined | 34.3% |
| **2019** | Combined | 29.7% |
|  | Landline | 24.3% |
|  | Mobile | 35.0% |
| **2020 (Jan – Oct)** | Combined | 24.9% |
|  | Landline | 18.6% |
|  | Mobile | 31.2% |
| **2020 (Oct – Dec)** | Combined | 20.4% |
|  | Listed Mobile | 15.9% |
|  | Mobile | 25.0% |
| **2021 (Jan – Jun)** | Combined | 22.5% |
|  | Listed Mobile | 18.0% |
|  | Mobile | 26.9% |

**5.4.Weighting**

The approach to the weighting of CITTS survey data was revised a number of times between 2005 and 2021. The revisions coincided with changes in research agencies and sample design.

**5.4.1. 2005–2006**

The weighting approach adopted for the 2005–2006 CITTS involved applying post-stratification weights to ensure the final sample matched smoker population benchmarks for location (Metropolitan Sydney, Northern and Southern NSW television markets), gender (male, female), age (16–to34, 35–to44, 55+) and socio-economic status (blue collar, white collar). The source of benchmarks are unknown.

**5.4.2. 2007–2012**

The weighting approach adopted for the 2007–2012 CITTS involved applying post-stratification cell weights to ensure the final sample matched general population (not smoker population) benchmarks for location (Metropolitan Sydney, Northern NSW, Southern NSW television markets), gender (male, female) and age (18–to24, 25–to39, 40–to54, 55+). Final weights were applied once each quarterly block of data collection had been completed and scaled so that they summed to the number of respondents in each cohort (i.e. proportional rather than target weights).

**3.4.3. 2013–2020**

The weighting approach adopted for the 2013–2020 CITTS involved a two-stage process.

* Applying a design weight to adjust for an individual’s chance of selection based on the number of in-scope sample members in a household, the number of landlines in the household used for private calls; and/or having a mobile phone.
* Applying a post stratification weight to ensure the final sample was weighted to relevant population benchmarks for age (18 to 24 years, 25 to 39 years, 40 to 54 years, 55+ years), gender (Female, Male), part of state (Metropolitan Sydney, Rest of state incl. ACT) and telephone status (Mobile only, Dual, Landline only).

Demographic benchmarks for smokers used for weighting were obtained from the 2014–15 National Health Survey. Benchmarks for telephony status were obtained from a combination of the 2014–15 National Health Survey and the 2013–15 Roy Morgan Single Source (RMSS) survey. From 2015 onwards, benchmarks for the distribution of characteristics are generally derived from figures published by the Australian Bureau of Statistics.

**3.4.4. October 2020 onwards**

The weighting approach adopted for a mobile only sample frame required one step only. This was to apply a post stratification weight to ensure the final sample was weighted to relevant population benchmarks for age (18 to 24 years, 25 to 39 years, 40 to 54 years, 55+ years), gender (Female, Male) and part of state (Metropolitan Sydney, Rest of state incl. ACT).

Final weights were applied once each six-month block of data collection had been completed and scaled so that they summed to the number of respondents in each cohort.

**5.5. Instrumentation**

The original CITTS questionnaire drafted in 2005 was modelled on that used to evaluate the National Tobacco Campaign. This included measures relating to smoking and quitting behaviour, unprompted and prompted campaign awareness and campaign impact or effects. In addition, a number of measures have been included from other sources, which are documented in the following sections, as well as ad hoc development of questions as required.

The average interview length for the 2007–2021 CITTS is provided in Table 27 below and as can be seen has varied slightly across the years.

**Table 4            Interview length (minutes)**

|  |
| --- |
| **Weekly range** |
|   | **Shortest** | **Longest** | **Average** |
| 2007 | n/a | n/a | 21.0 |
| 2008 | n/a | n/a | 23.4 |
| 2009 | n/a | n/a | 23.4 |
| 2010 | n/a | n/a | 21.0 |
| 2011 | n/a | n/a | 23.4 |
| 2012 | n/a | n/a | 21.6 |
| 2013 | 13.54 | 22.46 | 17.41 |
| 2014 | 12.10 | 20.40 | 16.25 |
| 2015 | 13.56 | 21.04 | 17.18 |
| 2016 | 13.53 | 22.38 | 18.19 |
| 2017 | 16.37 | 23.31 | 19.24 |
| 2018 | 15.25 | 21.52 | 18.40 |
| 2019 | 17.06 | 23.04 | 21.42 |
| 2020 | 17.55 | 22.54 | 20.41 |
| 2021 (Jan-Jun) | 17.17 | 22.08 | 21.02 |

**6.Coherence**

There have been a number of key changes to the design and conduct of the 2005–2021 CITTS. Key changes across years include: research agency, sample design, sample frame, in-scope definition, respondent selection process, weighting approach and questionnaire.

Direct comparison of results from the 2005–2021 CITTS has not been undertaken with other statistics products and was not a principle goal of the collection.

**7.Interpretability**

Information to assist in the interpretation of the 2005–2021 CITTS is provided in a number of forms: a Data Quality Statement and data dictionary supporting each modularised data file will be made available to assist in the interpretation.

**8.Accessibility**

Historically, the survey data collected via the CITTS has not been proactively released publicly. The CITTS data is publicly available by formal request.

The following products relating to the 2005–2021 CITTS will be available for public use.

* Data Access Policy.
* Data Quality Statement.
* Survey data for the 2005–2021 CITTS via 4 modularised data files: demographics; campaign recall (unprompted awareness); campaign recognition and diagnostics; and smoking and quitting.
* Data dictionary for each modularised data file.

[[1]](https://www.cancer.nsw.gov.au/cms/getdoc/eff15ca2-be1c-43fd-9db0-c2d36eabe27b/citts-data-quality-statement.aspx?viewmode=3&showpanel=1&cmscontentchanged=false&lang=en-AU&langobjectlifetime=request" \l "_ftnref1" \o ") Refer to the anti-tobacco campaign evaluation report for an overview of campaigns from 2005 to 2010 - [https://www.cancerinstitute.org.au/media/368182/Anti-tobacco-campaign-evaluation–2005–2010.pdf](https://www.cancer.nsw.gov.au/about-cancer/document-library/anti-tobacco-campaign-evaluation-2005-2010)

**Tags:**

[Reports](https://www.cancer.nsw.gov.au/category-search?c=350)