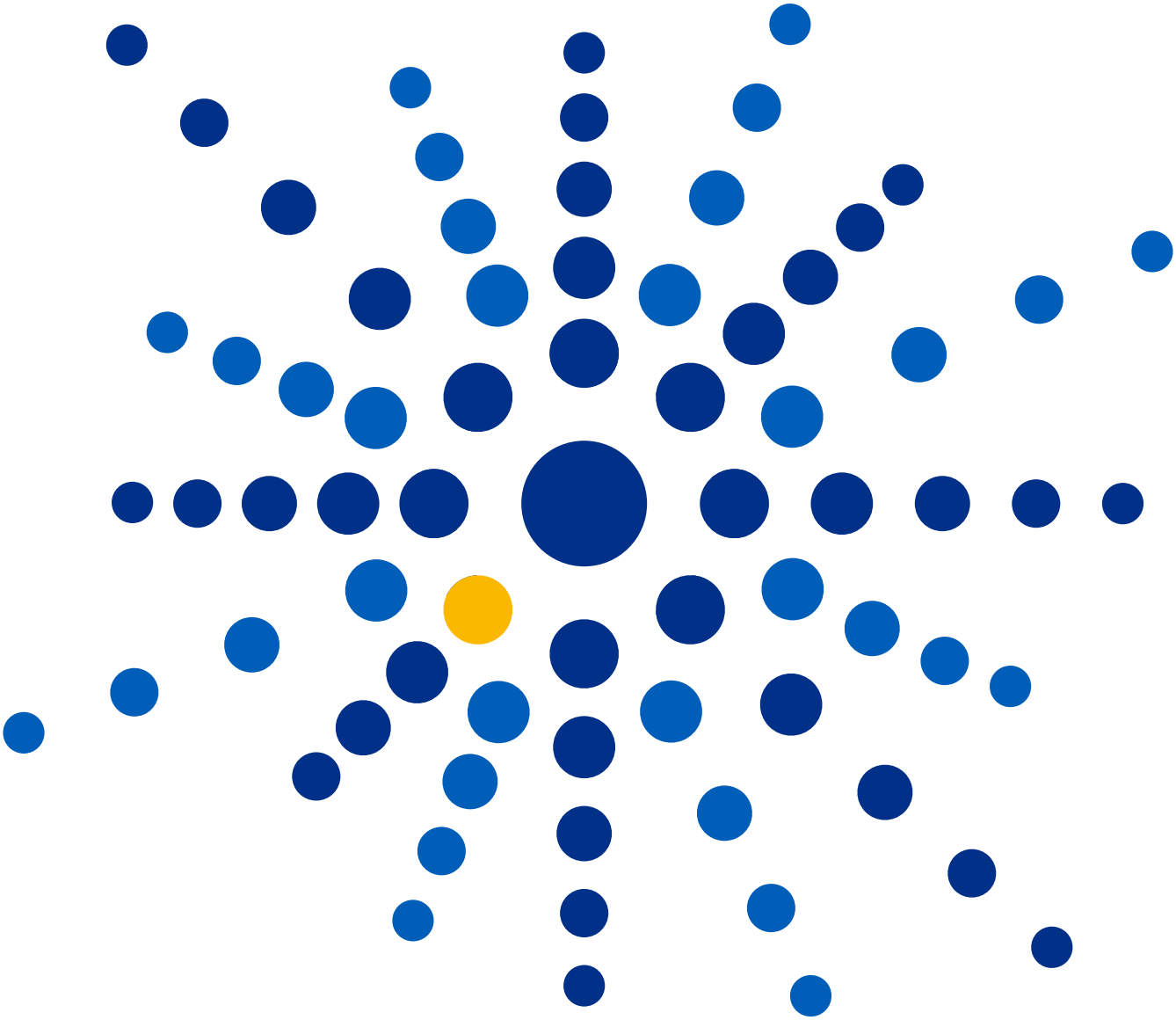

Methods

Christos Byron | Zoe Morgan | Sally Bridges | Marton Papp | Pablo Cabrera-Alvarez | Susan Purdon | Freya Tyrer | Jane Smith | Valdeep Gill | Terry Brugha | Sally McManus

Summary

Survey design summary

The Adult Psychiatric Morbidity Survey (APMS) 2014 is the fourth in a series of national mental health surveys. Each survey involved interviewing a large stratified probability sample of the general population, covering people living in private households. The full adult age range was covered, with the youngest participants aged 16 and the oldest over 100. The two-phase survey design involved an initial interview with the whole sample, followed up with a structured assessment carried out by clinically trained interviewers with a subset of participants. People were assessed or screened for a range of different types of mental disorder, from common conditions like depression and anxiety disorder through to rarer neurological and mental conditions such as psychotic disorder, attention-deficit/hyperactivity disorder (ADHD), and autism spectrum disorder (ASD). The long questionnaire also covered many aspects of people's lives that are linked to mental health, and this information can be used to profile the circumstances and inequalities experienced by people with mental disorders.

Aims and rationale for the survey

- To estimate the prevalence of a range of types of common and rare mental disorders in the population.
- Measure the gap between presence of each disorder and receipt of treatment.
- Produce trends in disorder and treatment through comparisons with previous surveys in the series.
- Enable the circumstances of people with different mental disorders to be compared with those of people without disorder.

Design strengths

- **By sampling from the general population rather than from lists of patients,** APMS data can be used to examine the 'treatment gap'. That is, the survey data can be used to explore what proportion of people with a condition are not in contact with services or in receipt of any treatment, or who are in receipt of inappropriate treatment.

- **The use of validated mental disorder screens and assessments** allows for identification of people with sub-threshold symptoms and those with an undiagnosed disorder.
- **Consistent methodology and coverage over time** allows for trends in a number of conditions to be monitored.
- **An advantage surveys like APMS have over routinely collected health data** is that for each participant a large amount of data on a range of topics is collected and relationships can be examined. In particular, the questionnaire covers detailed and current information about people's social and economic circumstances, information which does not tend to be collected in a consistent or comprehensive way in administrative datasets.
- **The use of a computer assisted self-completion module** to cover the most sensitive topics – such as suicide attempts, illegal behaviours, and experience of abuse and violence – means that the survey includes information that some participants may have never disclosed before.
- **At the end of the survey a question is asked about permission for follow-up.** The study therefore presents an opportunity for longitudinal data collection and a sampling frame that allows a random sample of people with very specific experiences, who may not otherwise have been identifiable, to be invited for further research.
- **The APMS dataset is being deposited at the UK Data Service** and is designed to be suitable for extensive further analysis. There is only scope for a small part of the data collected to be covered in this report.

Design limitations

- **The sampling frame covers only those living in private households,** and therefore those who were living in institutional settings such as large residential care homes, offender institutions, prisons, in temporary housing (such as hostels or bed and breakfasts) or sleeping rough, would not have had a chance to be selected. People living in such settings are likely to have worse mental health than those living in private households (Gill et al. 1996). However, the proportion of

the overall population not living in private households is so small that this would have little (or no significant) impact on the prevalence estimates for the disorders examined on APMS.

- **Some people selected for the survey were not able to take part in a long interview.** These include those with serious physical health conditions, who may feel unwell or be staying in hospital during the fieldwork period, and those whose mental capability may be impaired, for example due to cognitive decline as a result of dementia or injury, or because of a learning impairment. Where a selected participant could not take part due to a physical or mental health condition, some information about this was recorded by the interviewer on the doorstep. This information may be biased due to it having been collected often from another household member.

- **Some people selected for the survey could not be contacted or refused to take part.** The achieved response rate (57%) is in line with that of similar surveys (Barnes et al. 2010). A problem for all such studies is how to take account of those who do not take part, either because contact could not be established with the selected household or individual or because they refused to take part. The weighting (outlined in Section 14.7) addresses this to some extent.

- **The mental health assessments used are not as reliable as a clinical interview.** In a clinical interview, a trained psychologist or psychiatrist may take many sessions and much explorative questioning and clinical judgement to reach a diagnosis. In the context of a questionnaire administered by a lay interviewer, this is not possible. However, the assessments used have been validated and are among the best available for the purpose in hand.

- **Socially undesirable or stigmatised feelings and behaviours may be underreported.** While this is a risk for any study based on self-report data, the study goes some way to minimising this by collecting particularly sensitive information in a self-completion format.

- **As for all surveys, it should be acknowledged that prevalence rates are only estimates.** If everyone in the population had been assessed the rate found may be higher or lower than the survey estimate. Confidence intervals are given for key estimates in the methods chapter (Chapter 14). For low prevalence disorders, relatively few positive cases were identified. Particular attention should be given to uncertainty around these estimates and to any subgroup analysis based on these small samples. All comparisons made in the text have been tested and only statistically significant differences are described.

14.1 Introduction

The Adult Psychiatric Morbidity Survey (APMS) series began in 1993, and surveys have been conducted every seven years since. APMS 2014 is the fourth general population survey of adult mental health in the series. The first two were carried out by ONS in 1993 and 2000, and covered England, Scotland and Wales. The 2007 and 2014 surveys were carried out by NatCen Social Research, covered England only, and had no upper age limit to participation (which was 64 in 1993 and 74 in 2000). Like the preceding surveys, APMS 2014 consisted of two phases, with the second phase interview being conducted with a sub-sample of phase one participants by clinically trained interviewers coordinated by the University of Leicester.

The APMS series is part of a wider programme of surveys currently commissioned by NHS Digital, and funded by the Department of Health. Core topics are covered in every survey in the series, such as anxiety and depression, psychosis and substance use disorders. New topics in 2014 included screening for bipolar disorder and experience of childhood neglect.

This chapter provides a description of the survey methodology used on APMS 2014, including an outline of the:

- Sample design for the phase one and phase two interviews
- Topic coverage
- Piloting and questionnaire development

- Fieldwork procedures
- Survey response
- Weighting strategies
- Data analysis approach used in this report
- Quality assurance.

Further methodological detail is provided in the following appendices:

A Publications using data from the psychiatric morbidity survey series

B Assessment of psychiatric disorders

C Derived variables used in the main report

D Phase one questionnaire and phase two contents

E Fieldwork documents

14.2 Sample design

Overview of the sample design

The sample for APMS 2014 was designed to be representative of the population living in private households (that is, people not living in communal establishments or sleeping rough) in England. People living in communal or institutional establishments tend to be either aged 16 to 24 years (and living in higher education halls of residence) or aged 65 years or over (and living in a nursing or care home setting) (ONS 2015). Older people living in communal settings are likely to have worse mental health than older people living in private housing, and this should be borne in mind when considering the survey's account of the older population's mental health. Between the 2001 and 2011 censuses the proportion of young people recorded as living in communal establishments increased slightly and the proportion of older people in such settings decreased. However, overall, communal establishment residents represented less than 2% of all usual residents in England.

The survey adopted a multi-stage stratified probability sampling design. The sampling frame was the small user Postcode Address File (PAF) because this has excellent coverage of private households in England. The small user PAF consists of those Royal Mail delivery points which receive fewer than 50 items of mail each day. Therefore, most large institutions and businesses are excluded from the sample but some small businesses and institutions may receive fewer than 50 items each day and so be included. Once the interviewer had verified that an address does not contain a private household, such addresses were recorded as ineligible. The small proportion of households living at addresses not on the PAF (less than 3%) were not covered by the sample frame (ONS 2014).[1]

The stratified multi-stage random probability sample used for the phase one interview involved two stages of sample selection: the sampling of the primary sampling units (PSUs) followed by the sampling of addresses within the selected PSUs.

Wakefield local boost sample

In addition to the national sample, a sample for an additional local area boost was also drawn. The fieldwork involved the full phase one interview, but did not include a phase two assessment. The boost took place in Wakefield and was funded by a collaboration of the Wakefield Local Authority, NHS Wakefield CCG and South West Yorkshire Partnership NHS Foundation Trust. The Wakefield dataset will be deposited with the UK Data Service, with additional weighting variables that will allow for the datasets to be combined and analysed together as well as separately. This report focuses entirely on the national sample, which includes Wakefield only in proportion to its population.

Selection of primary sampling units (PSUs)

The PSUs were individual or groups of postcode sectors. A postal sector contains on average 2,550 delivery points. Small postal sectors were grouped with contiguous sectors so that each group contained at least 500 delivery points.

1 Addresses selected for all NatCen surveys in the last three years were excluded from the sampling frame. However, because they had been selected at random in the first place, this did not introduce selection bias. The benefit of this procedure is to reduce the burden of surveys on the public, which, it is hoped, will help to maintain response in the long term.

Before selection, the list of PSUs in the population was ordered (stratified) by a number of strata and a systematic random sample was selected from the ordered list. This ensures the different strata in the population are correctly represented and increases the precision of survey estimates.[2]

APMS 2014 used a sampling methodology that was consistent with previous surveys in the series, and very similar to that used in 2007. First, all PSUs in England were stratified by the 10 Strategic Health Authorities (SHA) as shown in Table 14.1. Within each SHA, PSUs were listed in increasing order of the proportion of people in non-manual occupations (according to the 2011 Census)[3] and cut-off points were drawn approximately one third and two thirds down the ordered list to create three roughly equal-sized groups. Within each of the 30 strata created (10x3), PSUs were listed in order of Census estimates of the percentage of households without a car and cut-off points were drawn to create three roughly equal-sized groups. Within each of the 90 strata created (30x3), PSUs were listed in order of the percentage of households owner-occupied. 682 PSUs were then systematically selected from the ordered list with probability proportional to the delivery point count of each PSU. Using the same stratification methodology, an additional sample of 16 PSUs was selected at a later stage to boost the size of the available achieved sample, thus bringing the total number of selected PSUs to 698. As stated above, this approach was comparable with that used for the other surveys in the series and is designed to produce a sample representative of the wider population, with biases in sample selection addressed through weighting. **Table 14.1**

Sampling addresses and households

In the second stage of sampling 22 delivery points were randomly selected within each of the selected PSUs. About half-way through fieldwork, progress was reviewed and it was decided that the sample should be reduced by removing

2 An estimate from a survey is precise if similar results are obtained with repeated surveys. One measure of precision is the standard error around an estimate.

3 The NS-SEC (National Statistics Socio-economic Classification) measure relating to household reference persons (the person in whose name the accommodation is owned or rented) does not easily lend itself to a manual/non-manual breakdown. Hence the social grade measure available for all *people* aged 16 and over in households was used, where non-manual was defined by social classes AB (higher and intermediate managerial/administrative/professional) and C1 (supervisory, clerical, junior managerial/administrative/professional).

a random selection of 636 addresses from the remaining two fieldwork quarters. Finally, three of the sampled addresses were not issued to interviewers due to problems with the addresses. Therefore the total sample of issued addresses was 14,717.[4]

Interviewers visited the addresses to identify private households with at least one resident aged 16 or over. When visited by an interviewer, 1,398 of the selected addresses were found not to contain private households. These addresses were thus ineligible, and were excluded from the survey sample. At eligible addresses found to contain more than one dwelling/household, interviewers used multi-dwelling/household selection grids to select one dwelling/household at random.

In summary, out of the 14,717 addresses in the original sample, 13,122 (89%) were found to contain at least one private household, 1,398 (9%) were non-residential addresses, and 197 (1%) were addresses of unknown eligibility.

Sampling one adult per household

One adult aged 16 years or over was randomly selected for interview in each eligible household. This was done in preference to interviewing all eligible adults because:

- It helped interviewers to conduct the interview in privacy and thereby obtain more reliable information.
- Individuals within households tend to be similar to each other and, where households differ markedly from each other, the resultant clustering can lead to an increase in standard errors around survey estimates. By selecting one person in each household this clustering effect was overcome.
- Given the length of the interview process, interviewing one household member helped to reduce the burden placed on each household.

4 Consisting of 698 PSUs each with 22 addresses, minus 636 deselected and 3 not issued.

Sampling for the phase two

The approach taken for selecting which phase one participants would be invited for a phase two assessment was based on that used in the 2007 survey, but amended to select on the basis of two disorders (psychosis and autism) not four (borderline and antisocial personality disorder were also covered at phase two in 2007). Further refinements to the sampling fractions, including introducing different sampling fractions for men and women, were possible due to the availability of data from the 2007 survey on the performance of the ASD screening items.

For each phase one participant, the probability of selection for a phase two assessment was calculated as the higher of two disorder-specific probabilities: psychosis probability and ASD probability. The probabilities were generated based on participants' responses to screening questions in the phase one questionnaire and whether they were male or female. These disorder-specific probabilities of selection to phase two were then corrected for in disorder specific weights, described in Section 14.7.

14.3 Topic coverage

APMS 2014 phase one interview

The table below summarises the topic coverage of the phase one interviews. The interview structure consisted of initial modules of questions administered by the interviewer, a self-completion section, and further interviewer administered modules. A few sections were asked only of particular age-groups, for example questions on cognitive decline were restricted to those aged over sixty years. This was done in part to minimise respondent burden. The full phase one questionnaire is reproduced in Appendix D and the documentation lodged with the UK Data Service describes each of the survey items.

APMS 2014 phase one interview content

	Age of participant		
CAPI interview: face to face interview [1]	**16–59**	**60–69**	**70+**
Details of household members and relationships	•	•	•
General health and activities of daily living	•	•	•
Caring responsibilities	•	•	•
Mental wellbeing (WEMWBS)[a]	•	•	•
Physical health conditions	•	•	•
Sensory impairment[a]	•	•	•
Learning impairment[a]	•	•	•
Mental illness diagnoses[a]	•	•	•
Treatment and service use	•	•	•
Common mental disorders	•	•	•
Suicidal behaviour and self-harm	•	•	•
Psychosis screening questionnaire	•	•	•
Attention-deficit/hyperactivity disorder	•	•	•
Work related stress	•	•	–
Tobacco	•	•	•
Alcohol – any drinking	•	•	•
CASI interview: self completion			
Alcohol (AUDIT, SADQ)	•	•	•
Drug use and dependence	•	•	•
Personality disorder	•	•	•
Social functioning (SRQ)	•	•	•
Bipolar disorder[a]	•	•	•

continued	Age of participant		
CASI interview: self completion	**16–59**	**60–69**	**70+**
ASD	●	●	●
Posttraumatic stress disorder	●	●	●
Military experience	●	●	●
Domestic violence and abuse	●	●	●
Child neglect[a]	●	●	●
Suicidal behaviour and self-harm	●	●	●
Discrimination	●	●	●
Sexual identity and behaviour[a]	●	●	–
Menopause[a]	●	–	–
CAPI interview: face to face interview [2]			
Cognitive and intellectual functioning:			
TICS-M	–	●	●
National Adult Reading Test (NART)	●	●	●
Animal naming test	–	●	●
Stressful life events (LTE)	●	●	●
Parenting	●	●	●
Social support networks (IMSR)	●	●	●
Religion	●	●	●
Social capital and participation	●	●	●
Socio-demographics	●	●	●
Consents (for data linkage and phase two contact)	●	●	●

[a] *These are new modules included in APMS for the first time in the 2014 survey.*

Mental disorders covered on APMS 2014

A summary of the measures used to assess or screen for each of the mental disorders included in APMS 2014 is listed below, with further technical detail in Appendix B.

Measures used to assess and screen for mental disorder

Condition	Diagnostic status	Classification system	Assessment tool	Survey phase	Reference period
Generalised anxiety disorder (GAD)	Present to diagnostic criteria	ICD-10	CIS-R (Lewis et al. 1992)	One	Past week
CMD not otherwise specified (NOS)	Present to diagnostic criteria	ICD-10	CIS-R	One	Past week
Obsessive and compulsive disorder (OCD)	Present to diagnostic criteria	ICD-10	CIS-R	One	Past week
Depressive episode	Present to diagnostic criteria	ICD-10	CIS-R	One	Past week
Panic disorder	Present to diagnostic criteria	ICD-10	CIS-R	One	Past week
Phobia	Present to diagnostic criteria	ICD-10	CIS-R	One	Past week
Alcohol use disorders	Screen positive	ICD-10	AUDIT (Saunders et al. 1993); SADQ (Stockwell et al. 1994)	One	Past six months
Drug dependence	Screen positive	DSM-IV	Based on Diagnostic Interview Schedule (Malgady et al. 1992)	One	Past year
Psychotic disorder	Present to diagnostic criteria	ICD-10	SCAN (WHO 1999)	One/ two	Past year

continued

Condition	Diagnostic status	Classification system	Assessment tool	Survey phase	Reference period
Any personality disorder	Screen positive	–	SAPAS (Moran et al. 2003)	One	Lifetime
Borderline personality disorder (BPD)	Present to diagnostic criteria	DSM-IV	Self-report SCID-II (First et al. 1997)	One	Lifetime
Antisocial personality disorder (ASPD)	Present to diagnostic criteria	DSM-IV	Self-report SCID-II	One	Lifetime
Posttraumatic stress disorder (PTSD)	Screen positive	DSM-IV	PTSD-Check List (Blanchard et al. 1996)	One	Past week
Attention-deficit/hyper-activity disorder (ADHD)	Screen positive	DSM-IV	Adult Self-Report Scale-v1.1 (WHO 2003)	One/ two	Past six months
Bipolar Disorder (BD)	Screen positive	DSM-IV	Mood Disorder Questionnaire (Hirschfeld et al. 2000)	One	Lifetime
Attempted suicide	Occurrence of behaviour	–	Self completion	One	Past year
Autism	Present to diagnostic criteria	DSM-IV	Autism Diagnostic Observation Schedule (ADOS: Lord et al. 2003)	One/ two	Lifetime

The phase two interview assessed psychotic disorder and autism. In addition, a further assessment of ADHD was introduced to the phase two interview in 2014. The approach taken to the phase two assessment of psychosis is described in Chapter 5, the phase two assessment of autism is described in Chapter 6. The phase two assessment of ADHD is not covered in this report but will be covered in subsequent publications.

Coverage of the 1993, 2000, 2007 and 2014 APMS interviews

The following table summarises the topic coverage of the 1993, 2000, 2007 and 2014 APMS phase one questionnaires. In 1993 the survey was administered by paper and pen, from 2000 a consistent computer assisted interviewing approach was used. The aim has been to have consistent core coverage, with additional modules covered in different years.

Summary of APMS coverage in 1993, 2000, 2007 and 2014

Face to face interview	1993	2000	2007	2014
General health	–	●	●	●
Activities of daily living	–	–	●	●
Caring responsibilities	–	–	●	●
Service use and medication	●[a]	●	●	●
Self-perceived height and weight	–	–	●	–
Common mental disorders	●	●	●	●
Suicidal behaviour and self-harm	●[b]	●	●	●
Psychosis screening questionnaire	●	●	●	●
Attention-deficit/hyperactivity disorder	–	–	●	●
Work related stress	–	–	●	●
Smoking	●	●	●	●
Drinking	●	●	●	●
Self completion				
Problem drinking	●[c]	●	●	●
Drug use	●	●	●	●
Personality disorder	–	●	●	●
Social functioning	–	–	●	●

continued				
Self completion				
Problem gambling	–	–	●	–
ASD	–	–	●	●
Posttraumatic stress disorder	–	–	●	●
Military experience	–	–	●	●
Bipolar disorder	–	–	–	●
Domestic violence, abuse and neglect	–	–	●	●
Suicidal behaviour and self-harm (repeated)	–	–	●	●
Eating disorder	–	–	●	–
Discrimination	–	–	●	●
Face to face interview				
Intellectual functioning:				
TICS-M	–	●	●	●
National Adult Reading Test (NART)	–	●	●	●
Animal naming test	–	●	●	●
Key life events	●	●	●	●
Social support networks	●	●	●	●
Religion	–	–	●	●
Social capital and participation	–	–	●	●
Socio-demographics	●	●	●	●

[a] *In APMS 1993 only participants who screened positive for CMD were asked about use of services and receipt of treatment.*

[b] *In APMS 1993 only participants with depression in the past week were asked about suicidal behaviour.*

[c] *APMS 1993 data on problem drinking is not compatible with that collected in 2000, 2007 and 2014.*

Key variations across the survey series

A primary purpose of the survey series is to assess change in the population prevalence of disorders over time. For this reason comparability with previous surveys was a priority and so both the questionnaire and the approach taken to its administration were largely the same. However, there have been some changes in coverage and method over time, and these are summarised below. They were made as a result of consultation with data users and potential data users.

Area

The 1993 and 2000 surveys covered England, Scotland and Wales, while the 2007 and 2014 surveys covered England only.

Age range

APMS 2007 and 2014 sampled adults aged 16 and over without an upper age limit. APMS 2000 included adults aged 16–74 and APMS 1993 covered adults aged 16–64.

New topics added

The following topics were included for the first time in the 2014 survey:

- Sensory impairment
- Previous diagnosis of mental illness and learning impairment
- Bipolar disorder
- Child neglect
- Menopause
- Sexual behaviour

Summary of amendments to existing modules

The full questionnaire was reviewed prior to launch in 2014. A detailed list of all questionnaire changes are included with the archived dataset, including information on the rationale for changes. In summary, amendments made to modules that were in the 2007 questionnaire include:

- **Mental wellbeing:** single item measures were replaced with the validated 14 item Warwick Edinburgh Mental Well-Being Scale (WEMWBS) (Tennant et al. 2007).
- **General health:** the SF12[5] was replaced with detailed questions about specific chronic conditions.
- **Caring responsibilities:** additional questions were asked about the nature of the relationship between the participant and the person they provide care for.
- **Medications and service use:** an extended list of medications and services were asked about, to reflect changes in prescribing practice and services available; new questions were added on requesting treatment.
- **Common mental disorder:** questions on social phobia were added (the mini Social Phobia Inventory, Weeks et al. 2007).
- **Work-related stress:** the module was extensively revised, including adding additional questions on bullying in the workplace.
- **Tobacco:** new questions were added on smoking cessation and e-cigarettes.
- **Personality disorder:** the addition of a screen for any personality disorder (SAPAS).
- **Suicidal behaviour and self-harm:** while some questions were retained in the face to face section of the interview, most were moved into the self-completion section.
- **Drug use:** new questions were added on use of ketamine and mephedrone.
- **PTSD:** the screening tool changed to the PTSD-Check List (PCL) for better comparability with other surveys.
- **Military experience:** additional questions on deployment were added.
- **Interpersonal violence and abuse:** additional questions about the assailant were added.

5 The 12-item Short-Form Survey (SF12) www.rand.org/health/surveys_tools/mos/12-item-short-form.html. This change was made partly due to steep increases in the license costs for use of this tool.

- **Key life events:** changes were made to the questions about key life events to make it clearer when the events had taken place.
- **Religion:** questions on spirituality were replaced with questions on belief.
- **Poverty:** questions were added on material deprivation.

Phase two sample selection

For the 2014 survey, data from the 2007 survey were available on which to model sampling fractions for autism. This allowed for the development of more precise and discriminating probabilities. The 2014 approach is outlined in Section 14.2.

14.4 Piloting and questionnaire development

Guidance and consultation

The APMS series is long-established, and the 2014 survey design is based on that used in previous surveys in the series. The survey development that did take place, to ensure that the survey meets current needs, drew on the expertise of a wide range of advisors and data users. These included:

- Project oversight and management from key managers at NHS Digital.
- A Steering Group comprised of representatives from the Department of Health, Public Health England, NHS England, Royal College of Psychiatrists, Improving Access to Psychological Therapies, and academic leads in psychiatric epidemiology (Professor Paul Bebbington) and economics (Lord Professor Richard Layard). This group was co-ordinated by NHS Digital.
- An APMS Academic Group, co-ordinated by the research team, and drawing on the expertise of leading academics from a range of universities and medical schools.
- A convened group of senior NatCen interviewers with practical experience of survey delivery in field.

Cognitive testing

Two waves of cognitive testing were carried out in 2013 and 2014.[6] Because one of the main questionnaire modules being developed addressed psychiatric diagnoses, participants with personal experience of mental illness were oversampled. The purpose of this stage of development work was to test the questions new to the 2014 survey and some questions from 2007 that were identified by data users and others as in need of revision. In addition, the cognitive testing explored alternative survey names and visual branding, and lead to the development of a study logo. Participants' interpretations of questions were explored, as well as their views on acceptability, language and terminology. Reports on the findings of the cognitive testing were submitted to NHS Digital.

Dress rehearsal

Following the cognitive testing, the questionnaire was refined in preparation for a full dress rehearsal. The dress rehearsal enabled testing of the flow, content and timings of the interview as a whole, and of individual modules, together with the operation of fieldwork procedures. The dress rehearsal included phase two interviews conducted by clinically trained interviewers co-ordinated by the University of Leicester. The phase two pilot sample included people both men and women of a range of ages. Again, a report on the dress rehearsal was submitted to NHS Digital.

14.5 Fieldwork procedures

Training and supervision of interviewers

Phase one interviewers

The NatCen interviewers selected to work on the first phase of the survey tended to be particularly experienced, and most had worked previously on other health-related surveys. They were fully briefed on the administration of the survey. Topics covered on the one-day survey-specific training included introducing the survey, questionnaire content, confidentiality and responding to participant distress.

6 For more details on cognitive testing see Collins D (2003) *Pretesting survey instruments: An overview of cognitive methods* in Quality of Life Research 12. Kluwer Academic Publishers.

Written instructions were provided for interviewers. As the fieldwork took place over the course of a year, refresher sessions were available for interviewers who took a break from the survey. Less experienced interviewers were accompanied by a project supervisor during the early stages of their fieldwork to ensure that the interviews were administered correctly. Routine supervision of 10% of interviewer work was subsequently carried out.

Phase two interviewers

The phase two interviewers were recruited and co-ordinated by the University of Leicester. They were all experienced in psychological research interviewing, and most had worked on APMS 2007. Phase two interviewers received an extensive, month-long induction and training programme, run by a senior research psychologist and a psychiatrist. They also received training sessions from NatCen on using computer assisted interviewing. Whilst in the field these interviewers received regular supervision sessions and technical support.

Advance letters

An advance letter was sent to each sampled address. This introduced the survey and stated that an interviewer would be calling to seek permission to interview. A sample advance letter is provided in Appendix E.

Making contact

At initial contact, the interviewer established the number of households at the address, and made any selection necessary (see Section 14.2). The interviewer randomly selected one adult per household, and then attempted to interview that person. As in previous waves in the series, the survey title used in the field was the 'National Study of Health and Wellbeing'. This was felt to be more readily understandable than 'psychiatric morbidity', an observation confirmed in the cognitive testing (see Section 14.4). Interviewers had various materials they could use on the doorstep and leave with participants, including a survey leaflet that introduced the study and provided details of a number that people could call (see Appendix E).

Collecting the data

The phase one and the phase two interviews each took about an hour and a half to complete on average, although some were shorter and others took as long as three hours. The phase one and phase two interviews both involved computer assisted interviewing (CAPI). In phase one, some information was collected by self-completion, also using the laptop. Despite the self-completion section being very long, 75% of participants completed this entirely alone. In 16% of cases the interviewer read out the self-completion and entered the participant's responses, and for 3% of cases the interviewer read out the questions but the participant entered their own responses. 6% of participants did not complete the self-completion section of the interview at all, this was mainly older participants.

At the end of the phase one interview, permission was sought for the participant's survey responses to be linked with other health datasets, including the NHS Central Register and Hospital Episode Statistics. 77% gave permission for data linkage. The documentation for this is included in Appendix E. Verbal permission was also sought for a University of Leicester interviewer to contact the participant again in order to explain the phase two interview, should they be selected: 78% agreed.

If the selected participant was not capable of undertaking the interview alone, for reasons of mental or physical incapacity, the option was available for additional information to be collected from another member of the family or a carer on their reasons for not being able to take part. In 2007 this took the form of a proxy interview, with data collected in the laptop. In 2014, information was collected on the doorstep.

Token of appreciation and helpline information

A high street voucher was given to all those who took part in a phase one interview as an appreciation for their time. In addition, those who were selected and took part in the phase two interview were given an additional high street voucher. All participants were also offered a list of helpline numbers that they could call. These included the numbers for organisations providing information about the various disorders covered in the survey as well as for those providing support to people in crisis. The helplines leaflet also emphasised contacting a GP for support and advice as a first step (see Appendix E).

14.6 Survey response

Response at phase one

Of the 14,717 addresses in the original sample, 13,122 (89%) were found to include at least one private household, 1,398 (9%) were non-residential addresses, and 203 (1%) were addresses of unknown eligibility. This left 13,313 cases in the sample set. Of these 3872 (29%) were refusals in field and 300 refusals direct to the office. 782 were coded as non-contacts and 813 were unproductive for another reason. 7,546 productive interviews were achieved, representing a 57% response rate. This included 18 partial interviews where the participant completed the treatment, service use and CIS-R modules, but did not reach the end of the interview.

Response rates of adults at phase one

	Number	Percentage
Potentially eligible households	**13,313**	
Field refusals	3,872	29%
Office refusals	300	2%
Non-contacts	782	6%
Other unable/unproductive	813	6%
Productive adults	**7,546**	**57%**
Full interviews	7,528	
Partial interviews	18	

Response at phase two

7,528 participants provided a full phase one interview. A probability of selection was calculated for each participant based on their answers to the phase one screening questions on psychosis and ASD as outlined in Section 14.2. Overall 78% of phase one participants agreed to be contacted about the phase two interview. After the application of the highest of the two disorder specific sampling fractions, 875 participants were issued for a phase two interview. Phase two interviews were conducted with 630 of these (72%), and there were 204 refusals and 41 non-contacts.

14.7 Weighting the data

Weighting the phase one data

The survey data were weighted to take account of selection probabilities and non-response, so that the results were representative of the household population aged 16 years and over. Weighting occurred in four steps.

First, address selection weights (wt1) were applied to take account of the differential selection probabilities of addresses (after the removal of 636 addresses from the originally drawn sample, see Section 14.2). For each of the 698 sampled PSUs, the weight was calculated as follows: wt1 = total addresses on PAF / (698 x number of sampled addresses per PSU). All addresses in the same PSU were assigned the same weight.

Second, to reduce household non-response bias, a household level weight was calculated from a logistic regression model using interviewer observation and area-level variables (collected from Census 2011 data) available for responding and non-responding households. The dependent variable was whether the household responded or not. The independent variables considered for inclusion in the model were the presence of any physical barriers for entry to the property (e.g. a locked common entrance or the presence of security staff), Government Office Region (GOR), Index of Multiple Deprivation 2010 (IMD 2010) quintiles,[7] population density (number of persons per hectare), percentage of households owner-occupied, and the percentage of adults in a non-manual occupation.

Not all the variables were retained for the final model: variables not significantly related to the propensity of households to respond were dropped from the analysis. The variables significantly associated with response were: GOR, whether there were entry barriers to the selected address, the percentage of households owner-occupied and population density. The model shows that the propensity for a household to respond was lower in Yorkshire and Humberside, East of England, and in inner and outer London (relative to the North East), higher for households

7 IMD 2010 is a measure of multiple deprivation at the small area level. www.gov.uk/government/statistics/english-indices-of-deprivation-2010

with no physical barriers for entry to the property, higher in areas where a relatively high percentage of households were owner-occupied and lower in areas with a relatively high population density.

The non-response weight (wt2) for each eligible household was calculated as the inverse of the probability of response estimated from the final model. The full model is given in Table 14.2. **Table 14.2**

Third, selection weights (wt3) were applied to take account of the different probabilities of selecting participants in different sized households.[8] The weight was equal to the number of adults (16+) in the household, the inverse of the probability of selection.[9]

The composite weight for selection and participation was calculated as the product of the weights from the previous stages: wt4 = wt1 x wt2 x wt3.

The final stage of the weighting was to adjust the composite weight (wt4) using calibration weighting.[10] Calibration takes an initial weight (in this case wt4) and adjusts (or calibrates) it to given control totals. The process generates a weight which produces survey estimates that exactly match the population for the specific characteristics (control totals) used in the adjustment. Calibration reduces any residual non-response bias and any impact of sampling and coverage error for the measures used in the adjustment. The population control totals used were the ONS 2014 mid-year population estimates for age-by-sex and region, shown in Tables 14.3 and 14.4. After calibration, the APMS 2014 weighted data matches the estimated population in terms of age-by sex and region as shown in Table 14.5. **Tables 14.3 to 14.5**

An additional weight was calculated for the combined APMS 2007 and 2014 datasets by re-calibrating the combined data to the ONS 2014 mid-year population

8 The selection of multiple dwellings and/or households was done as a paper exercise and was not collected in the CAPI. As a result, there was no information on selection at addresses with multiple dwelling units or at dwelling units with multiple households, so it was not possible to adjust for this in the weighting. Evidence from other large scale English-only household surveys (e.g. the Health Survey for England) show that only a very small percentage of addresses in England (under 1%) turn out to include multiple dwellings/households. Therefore, any bias from not adjusting for multiple dwellings/households should be negligible.

9 The selection weight wt3 was trimmed at 4 to avoid a small number of very high weights which would inflate the standard errors, reduce the precision of the survey estimates and cause the weighted sample to be less efficient.

10 The calibration weighting was carried out in STATA (StataCorp. 2013).

estimates for age-by sex and region. The combined weight is being used solely for analysis of the combined 2-year dataset.

Weighting the phase two data

Two weighting variables have been developed specifically for use when analysing outcomes derived from phase two data: presence of psychosis and presence of autism. These weights were designed to generate condition-specific datasets that are representative of the general population, and based on all the participants with relevant information.

The phase two participants have a set of survey weights different from those generated for phase one, with one set of weights being applicable for psychosis and a second set being applicable for autism. Participants get a phase two weight if they were eligible for phase two, were selected, and then responded.

For analysis of prevalence of disorders assessed at phase two (autism and psychosis), the weighted phase two participants are added to the set of phase one participants who were not eligible for phase two, the prevalence being assumed to be zero for the not eligible group. Those not eligible are given their phase one weights.

The phase two weights account for two factors:

1. Not all those eligible for phase two were selected with equal probability: all those screened in with a positive psychosis score were selected (although those selected in the final two months of fieldwork were subsequently excluded), as were all men with an Autism-Spectrum Quotient (AQ) score of eight or more and all women with an AQ score of 11 or more. But for men with an AQ score of between 4 and 7, and women with an AQ score of between 4 and 10, sub-sampling was used.

2. Some of the eligible phase one participants did not agree to be contacted for phase two during their phase one interview so were automatically excluded from the phase two selection. Others were selected for phase two but then declined to take part. These refusals introduce the possibility of phase two non-response bias. The phase two weights incorporate a non-response adjustment to ensure that those responding have a similar weighted profile to those eligible.

The phase two weights were calculated by modelling, via logistic regression, the probability of being selected and responding to phase two, conditional on being eligible for selection. The weight per phase two participant was then calculated as the inverse of the predicted probability from the model, multiplied by their phase one weight. The predicted probabilities simultaneously account for selection probabilities and for observable non-response biases.

The variables included in the model were: phase one psychosis and AQ scores; gender; marital status, ethnic group (four categories (Moran et al. 2003)); and age group. Other variables, such as employment status, qualification, and the index of multiple deprivation quintile group, were tested in the regression model but excluded because not significant (the implication being that there is no statistical evidence of non-response bias on these variables).

14.8 Data analysis and reporting

Introduction

APMS 2014 is a cross-sectional survey of the general population. While it allows for associations between mental disorder and personal characteristics and behaviour to be explored, it is important to emphasise that such associations cannot be assumed to imply causality. A list of the variables used in the analysis in this report is provided in Appendix C: all will be included in the archived dataset.

Weighted analysis and unweighted bases

As outlined in Section 14.7 above, all the data presented in the substantive chapters of this report are weighted to account for likelihood of selection and non-response. Bases are presented as unweighted to show the number of participants included, should weighted bases be required these can be generated from the archived dataset.

Testing for seasonal variation

The fieldwork for the psychiatric morbidity surveys conducted in 1993 and 2000 was conducted around March to August of their respective calendar years.

Fieldwork for APMS 2007 and 2014 was spread across a whole year, so that any seasonal variation in rates could be explored. This raised the issue of whether month of interview would need to be controlled for when examining trends in disorders assessed on the basis of symptoms in a recent reference period. To check on this, we looked at the rate of any common mental disorder and the rate of depression in terms of the month in which the interview took place. Adjusting for month of interview did not significantly affect rates of disorder and so is not included this report.

Age-standardisation

Rates of disorder in some analyses have been age-standardised in this report to allow for comparisons between groups after adjusting for the effects of any differences in their age distributions. When sub-groups are compared in respect of a variable on which age has an important influence, differences in age distributions between sub-groups are likely to affect the observed differences in the proportions of interest.

Most analyses in this report (sample size permitting) are presented separately for men and women, as well as for the total population. Age-standardisation was undertaken separately within each sex, expressing male data to the overall male population and female data to the overall female population. When comparing data for the two sexes, it should be noted that no age-standardisation has been undertaken to remove the effects of the sexes' different age distributions. It should also be noted that where data for all adults combined is presented as age-standardised, this has been produced in the way outlined above, with male data expressed to the age profile of the male population and female data expressed to the profile of the female population.

Age-standardisation was carried out using the direct standardisation method. The reference population was the Office for National Statistics' Census based mid-year 2014 population estimates for England.

Age-standardisation was not conducted for some analyses. These include analysis by household type. Our age-standardisation approach requires cases to be present in each 'cell'. Because some household type groups (e.g. 'one or more adults aged 65+') did not have cases in some age/sex combinations (e.g. men aged 16–24),

there was no rate in the cell to weight up to the population prevalence. Moreover, where there are very few cases in a cell, this tends to cause instability in the age-standardised rate generated. This is one of the reasons why observed rates are often presented alongside rates that have been age-standardised.

Standard analysis breaks

Most of the disorders covered in this report are analysed by a core set of breaks: age, sex, ethnic group, household type, employment status, benefit status, and region, described briefly below and are defined in more detail in the Glossary.

Ethnic group

Participants identified their ethnicity according to one of fifteen groups presented on a show card, including 'other – please state'. These groups are based on those used in the latest Census and are drawn from the ONS harmonised ethnic group questions for use on national surveys. The groups were subsumed under four headings: White; Black/Black British; Asian/Asian British; and those who reported their ethnic group as mixed, multiple or other. For some analyses by ethnic group the White group was further divided into 'White British' (which included those giving their ethnic group as White and English, Scottish, Welsh or from Northern Ireland) and White other. About 15% of the sample (1,131 participants) identified with an ethnic group other than White British. This is in line with the combined prevalence of these groups in the adult population resident in England. It should be noted that these small groups are highly heterogeneous, for example the 'Black' group could include both recent migrants from Somalia and Black people born in Britain to British parents. The results of analysis by ethnic group should therefore be treated with caution.

Household type

In APMS, basic information (age, sex and relationship status) was collected from the participant about all members of the household in which they lived. This enabled a variable to be derived that summarised the structure of the household in which people lived, particularly in relation to the number and ages of the other people lived with. This enables the circumstances of people living alone to be compared with those of people living with others, as well as identifying participants living with children. A 'small family' was defined as one or two adults living with

one or two children, and a 'large family' was defined as at least one adult living with three or more children.

Employment status

Detailed information was collected from participants on the nature of their employment status in the previous week. Participants were classified as either employed (including working in a family business); unemployed (and therefore looking and available for work); or economically inactive (including those who are unable to work due to disability or illness, students, retired, or looking after the home). The standard International Labour Organisation[11] definition was used, and is described more fully in the Glossary. Where this analysis break has been used, generally the base has been restricted to participants aged between 16 and 64.

Benefit status

Participants were asked, using a series of showcards, whether or not they were in receipt of each of a range of benefits. For the purposes of the analyses presented in this report, three variables were derived. One allows participants reporting current receipt of Employment and Support Allowance (ESA) to be compared with those who do not receive this benefit. ESA is a benefit available to people of working age but who are unable to work due to disability or impairment. While Incapacity Benefit (IB) was not actually available at the time of the interview, a few participants (11) reported receiving this, and they were included in the ESA group. 'Any out of work benefit' included those reporting ESA or IB, combined with those in receipt of Jobseeker's Allowance (JSA). Analyses using these two variables were restricted to participants aged 16–64. A third benefit group was used in the analysis in this report: those living in a household in receipt of Housing Benefit. This is a household level benefit and relates to support provided to help with rent costs.

Region

The former GORs were used for the analysis by region. The APMS sample is too small for analysis by geographical groupings below region.

11 www.ilo.org/global/about-the-ilo/lang--en/index.htm

Treatment and service use

When looking at treatment and service use, participants screening positive for each disorder were compared with those who did not. Because of the relatively low prevalence of many of the disorders assessed in APMS 2014, this generally meant that the base size for the group with the disorder was usually small. Age-standardising a small group can be problematic, for the reasons outlined in Section 14.8, and so the treatment and service use tables were not age-standardised in most of the chapters.

Sampling errors and design factors

The percentages quoted in the main report are estimates for the population based on the information from the sample of people who took part in this survey. All such survey estimates are subject to some degree of error. The confidence interval (CI) is calculated from the sampling error, which is a measure of how such a survey estimate would vary if it were calculated for many different samples. If the survey was repeated many times, such a 95% CI would contain the true value 95% of the time. For this survey, a multi-phase stratified design was used, rather than a simple random sample, and the sampling errors need to reflect this.

The effect of a complex sample design on estimates is quantified by the design factor (deft). It is the ratio of the standard error for a complex design to the standard error which would have resulted from a simple random sample. A deft of two, for example, indicates that the standard errors are twice as large as they would have been had the sample design been a simple random sample. The sampling errors, design effects and CI for key prevalence variables can be found in Tables 14.6 to 14.15. The calculations were carried out using the statistical package SPSS v21 (IBM Corp. 2012). **Tables 14.6 to 14.16**

Quality assurance

Quality assurance has been defined as any method or procedure for collecting, processing or analysing survey data that is aimed at maintaining or enhancing reliability or validity (Statistics Canada 1998). It was an ongoing process throughout APMS, from preparation and sampling through data collection and data analysis to report writing, as detailed in this chapter. NatCen has a quality management system

with sets of procedures that were followed throughout. The purpose of establishing standard procedures, as highlighted by the WHO in relation to its World Health Surveys, is to help ensure that (Üstun et al. 2005):

- Data collection is relevant and meaningful
- Data can be compared across surveys and between subgroups
- Practical implementation of the survey adheres to proper practice
- Errors in data collection are minimised
- Data-collection capability is improved over time.

Examples of quality control measures built into, or to check afterwards, the survey process included:

- The computer programme used by interviewers had in-built soft checks (which can be suppressed) and hard checks (which cannot be suppressed); these included querying uncommon or unlikely answers, and answers out of the acceptable range.
- For phase one interviewers, telephone checks were carried out with participants at 10% of productive households to ensure that the interview had been conducted in a proper manner.
- The phase two interview was less structured, and required clinical skill and assessment by a graduate psychologist. The work of these research psychologists was supervised by a senior research psychologist. The experienced trainer also accompanied all of the interviewers on at least one of their participant visits 3 months into fieldwork, to ensure that they were conducting the interview as per protocol and to validate the coding. If a further supervised visit was felt necessary, this was also carried out.
- An ADOS (ASD assessment) reliability day was carried out, where all phase two interviewers returned to Leicester for their ADOS interviewing to be validated. Furthermore, if a phase two interviewer was unsure about any rating during fieldwork, they made extensive notes and then contacted the field research manager to discuss.

14.9 Tables

Sampling

Table 14.1 Regional stratifier used and number of PSUs selected

Table 14.2 Final response model

Table 14.3 2014 mid-year household population estimates for adults in England, by age and sex

Table 14.4 2014 mid-year household population estimates for adults in England, by Government Office Region

Table 14.5 Weighted and unweighted sample distribution, by Government Office Region, age and sex

Standard errors and confidence intervals for key estimates

Table 14.6 True standard errors and 95% confidence intervals for CIS-R score and prevalence of common mental disorders (CMDs)

Table 14.7 True standard errors and 95% confidence intervals for treatment rate among people with a CMD

Table 14.8 True standard errors and 95% confidence intervals for major trauma and screen positive for posttraumatic stress disorder (PTSD)

Table 14.9 True standard errors and 95% confidence intervals for prevalence of psychotic disorder in past year

Table 14.10 True standard errors and 95% confidence intervals for autism

Table 14.11 True standard errors and 95% confidence intervals for screening positive for personality disorder

Table 14.12 True standard errors and 95% confidence intervals for number of attention-deficit/hyperactivity disorder characteristics present in the past six months

Table 14.13 True standard errors and 95% confidence intervals for prevalence of bipolar disorder

Table 14.14 True standard errors and 95% confidence intervals for prevalence of hazardous and harmful drinking in the past year

Table 14.15 True standard errors and 95% confidence intervals for prevalence of drug dependence

Table 14.16 True standard errors and 95% confidence intervals for prevalence and recency of suicidal thoughts, suicide attempts and self-harm

14.10 References

Barnes W, Bright G, Hewat C. (2010) *Making sense of Labour Force Survey response rates*. ONS.

Blanchard EB, Jones-Alexander J, Buckley TC, Forneris CA. Psychometric properties of the PTSD checklist (PCL). *Behaviour Research and Therapy,* 1996; 34: 669–673.

First MB, Gibbon M, Spitzer RL, William JBW and Benjamin L (1997) *Structured Clinical Interview for DSM – IV Axis II Personality Disorders,* American Psychiatric Press: Washington.

Gill B, Meltzer H, Hinds K and Petticrew M (1996) *OPCS Surveys of Psychiatric Morbidity in Great Britain, Report 7: Psychiatric morbidity among homeless people*, HMSO: London.

Hirschfeld RM, Williams JB, Spitzer RL, et al. Development and validation of a screening instrument for bipolar spectrum disorder: the Mood Disorder Questionnaire. *American Journal of Psychiatry*, 2000; 157(11): 1873–5.

IBM Corp. (2012). *IBM SPSS Statistics for Windows, Version 21.0.* Armonk, NY: IBM Corp.

Lewis G, Pelosi AJ, Araya R, Dunn G. Measuring psychiatric disorder in the community; a standardised assessment for use by lay interviewers. *Psychological Medicine*, 1992; 22: 465–486.

Lord C, Risi S, Lambrecht L, Cook EH Jr, Leventhal BL, DiLavore PC, Pickles A, Rutter M. The autism diagnostic observation schedule-generic: a standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders*, 2002; 30: 205–223.

Malgady RG, Rogler LH, Tryon WW. Issues of validity in the Diagnostic Interview Schedule. *Journal of Psychiatric Research*, 1992; 26: 59–67.

Moran P, Leese M, Lee T, Walters P, Thornicroft G, Mann A. Standardised Assessment of Personality – Abbreviated Scale (SAPAS): preliminary validation of a brief screen for personality disorder. *British Journal of Psychiatry,* 2003; 183: 228–32.

ONS (2014) *The Coverage of the Postcode Address File and Address Base for Sampling*.

ONS (2015) *2011 Census Analysis: What Does the 2011 Census Tell Us About People Living in Communal Establishments?* www.ons.gov.uk/ons/rel/census/2011-census-analysis/what-does-the-2011-census-tell-us-about-people-living-in-communal-establishments-/story-what-does-the-2011-census-tell-us-about-people-living-in-communal-establishments-.html

Saunders JB, Aasland OG, Babor TF, Dela Fuente JR, Grant M. Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO collaborative project on early detection of persons with harmful alcohol consumption, part II. *Addiction*, 1993; 88: 791–804.

StataCorp. (2013). *Stata Statistical Software: Release 13.* College Station, TX: StataCorp LP.

Statistics Canada (1998). *Quality Guidelines,* 3rd ed. Ottawa.

Stockwell T, Sitharan T, McGrath D, Lang. The measurement of alcohol dependence and impaired control in community samples. *Addiction*, 1994; 89: 167–174.

Tennant R, Hiller L, Fishwick R, et al. The Warwick-Edinburgh Mental Well-being Scale (WEMWBS): development and UK validation. *Health and Quality of Life Outcomes*, 2007; 5: 63.

Üstun TB, Chatterji S, Mechbal A, Murray CJL. (2005) *Quality assurance in surveys: standards, guidelines and procedures.* World Health Organisation. Geneva: Switzerland. http://unstats.un.org/unsd/hhsurveys/pdf/Chapter_10.pdf

Weeks J, Spokas M, Heimberg RG. Psychometric evaluation of the mini-social phobia inventory (Mini-SPIN) in a treatment-seeking sample. *Depression and Anxiety,* 2007; 24(6): 382–392.

World Health Organisation (1999) *SCAN Schedules for Clinical Assessment in Neuropsychiatry Version 2.1*, World Health Organisation: Geneva.

World Health Organisation (2003) *Adult ADHD Self-Report Scale-V1.1 (ASRS-V1.1) Screen. WHO Composite International Diagnostic Interview*. World Health Organization.

This chapter should be cited as:

Byron C, Morgan Z, Bridges S, Papp M, Cabrera-Alvarez P, Purdon S, Tyrer F, Smith J, Gill V, Brugha T, McManus S. (2016) 'Chapter 14: Methods' in McManus S, Bebbington P, Jenkins R, Brugha T. (eds) *Mental health and wellbeing in England: Adult Psychiatric Morbidity Survey 2014*. Leeds: NHS Digital.

Designed by Soapbox soapbox.co.uk

