

The Netherlands Mental Health Survey and Incidence Study-3 (NEMESIS-3): Objectives, methods and baseline characteristics of the sample

Margreet ten Have  | Marlous Tuithof | Saskia van Dorsselaer |
Frederiek Schouten | Ron de Graaf

Trimbos Institute, the Netherlands Institute of Mental Health and Addiction, Utrecht, The Netherlands

Correspondence

Margreet ten Have, the Netherlands Institute of Mental Health and Addiction, Da Costakade 45, 3521 VS Utrecht, The Netherlands.
Email: mhave@trimbos.nl

Funding information

Ministerie van Volksgezondheid, Welzijn en Sport

Abstract

Objectives: NEMESIS-3 (Netherlands Mental Health Survey and Incidence Study-3) is a psychiatric epidemiological cohort study of the Dutch general population that replicates and expands on two previous NEMESIS-studies conducted in 1996–1999 and 2007–2018 respectively. The main aims of NEMESIS-3 are to provide up-to-date information on the prevalence, incidence, course and consequences of mental disorders, their risk indicators, and to study the relevant time trends. This paper gives an overview of the objectives and methods of NEMESIS-3, especially of the recently completed first wave, and describes the sample characteristics.

Methods: NEMESIS-3 is based on a multistage, stratified random sampling of individuals aged 18–75 years. Face-to-face interviews were laptop computer-assisted and held at the respondent's home. A slightly modified Composite International Diagnostic Interview (CIDI) version 3.0 was used to assess both Diagnostic and Statistical Manual of Mental Disorders-IV (DSM-IV) and DSM-5 mental disorders. Two follow-up waves are planned three and six years after baseline.

Results: In the first wave, performed from November 2019 to March 2022, 6194 individuals were interviewed: 1576 respondents before and 4618 respondents during the COVID-19 pandemic. The average interview duration was 91 min and the response rate was 54.6%. The sample consisted of 50.4% women and had a mean age of 47.9 years. The sample was reasonable nationally representative, although some sociodemographic groups were somewhat underrepresented.

Conclusions: Despite the COVID-19 restrictions, we were able to build a large and comprehensive dataset of good quality, permitting us to investigate the latest trends in mental health status, various new topics related to mental health, and the extent to which the pandemic has had an effect on the population's mental health.

KEYWORDS

general population, longitudinal, psychiatric epidemiology, survey methodology

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2022 The Authors. International Journal of Methods in Psychiatric Research published by John Wiley & Sons Ltd.

1 | INTRODUCTION

In the last two and a half decades two nationally representative prospective surveys studied the mental health status of the Dutch general population: NEMESIS-1 and NEMESIS-2 (Bijl, van Zessen, et al., 1998; De Graaf et al., 2010a). NEMESIS-1 was conducted from 1996 to 1999 in three waves, and NEMESIS-2 from 2007 to 2018 in four waves. Some of the main findings are detailed below. We then describe the objectives, methods and sample characteristics of NEMESIS-3, the first wave of which was recently completed.

1.1 | Some previous NEMESIS-findings

Both NEMESIS-1 and NEMESIS-2 estimated that approximately 40% of the population of 18–64 years had an Axis-I disorder in their lifetime and about 20% in the 12 months before the baseline interview (Bijl et al., 1998; De Graaf et al., 2012a). From an international perspective, these prevalence rates are similar to those in other Western European countries and somewhat lower than those in the United States (Kessler et al., 2007).

The longitudinal design of the studies made it possible to study both incidence and course of mental disorders. In NEMESIS-2 similar incidence rates were found to those of the largest incidence study of Diagnostic and Statistical Manual of Mental Disorders-IV (DSM-IV) disorders worldwide (the National Epidemiologic Survey on Alcohol and Related Conditions; Grant et al., 2009), but the rates were lower than the rates of DSM-III-R disorders in NEMESIS-1 (De Graaf et al., 2013). These differences may reflect a true decrease in incidence but may also be due to differences in sample characteristics, survey design and the classification system used to assess mental disorders in both NEMESIS-studies (DSM-IV vs. DSM-III-R).

The course of mental disorders was examined in NEMESIS in various ways, for example, by studying episode duration, recurrence, chronicity, and remission of disorders (e.g., Scholten et al., 2021; Schopman et al., 2021; Spijker et al., 2002; Ten Have et al., 2017, 2018, 2021; Tuithof et al., 2013). Worldwide, such longitudinal data on disease course from the general population are still scarce, but are needed to inform patients, to facilitate therapists with treatment planning, and to inform policymakers on disease management.

Other key topics of the previous NEMESIS-studies were functioning with, and after remission of, a mental disorder (e.g., Bijl & Ravelli, 2000a; Bos et al., 2018; Buist-Bouwman et al., 2004; Ormel et al., 2004; Schopman et al., 2018), costs of mental disorders due to care use or work loss (De Graaf et al., 2012b; Smit et al., 2006) and mental health care use (e.g., Bijl & Ravelli, 2000b; De Graaf et al., 2010b; Ten Have et al., 2003, 2013a, 2013b).

In conclusion, NEMESIS-1 and NEMESIS-2 have generated a lot of knowledge about the mental health status of the Dutch population. NEMESIS-3 offers the opportunity to study the latest trends in mental health and service use for mental health problems, and to look at new topics related to mental health.

1.2 | Objectives of NEMESIS-3

NEMESIS-3 replicates and expands on NEMESIS-1 and closely resembles NEMESIS-2 in terms of study design and clinical instrument. Since there is speculation about an increase in the past decade in the prevalence of mental disorders (Wennekers et al., 2019), an important objective of NEMESIS-3 is to study time trends in the prevalence of mental disorders for the 12-year period between the baseline waves of NEMESIS-2 and NEMESIS-3.

To properly answer the question whether a larger proportion of the population in the Netherlands has poorer mental health, we cannot rely on foreign studies and the number of large, well-conducted psychiatric epidemiological population studies that have been repeated several times over the past decade is limited. Nor can we use registration figures, as a possible increase in mental health care use does not necessarily mean that a larger proportion of the population has become mentally ill (Ruths et al., 2021; Steffen et al., 2020).

Furthermore, as the fieldwork for the first wave of NEMESIS-3 took place before and during the COVID-19 pandemic, an additional objective now is to investigate the extent to which the pandemic has had an effect on the population's mental health. To date, only a few studies have addressed this question by comparing prevalence rates of mental disorders before and during the pandemic (Kessler et al., 2022). That is because most studies lack information on the pre-pandemic period, have designs with one-time data-collection early in the pandemic and are based on self-selected non-probability or convenience samples with important biases.

As we expanded the age range of respondents from 18 to 64 years in the previous NEMESIS-studies to 18–75 years, it became possible to study the mental health of those aged 65 and over as well, an age group that is often underrepresented in psychiatric-epidemiologic studies (Andreas et al., 2017).

In conclusion, the main aims of NEMESIS-3 are to provide up-to-date information on prevalence, incidence, course and consequences of mental disorders, their risk indicators, and to study the relevant time trends.

2 | METHOD

2.1 | Study design

For NEMESIS-3 a random sample of 148 of the 355 Dutch municipalities was drawn, stratified by four regions (north, east, south, west) and five levels of population density. The four largest cities, Amsterdam, Rotterdam, The Hague, and Utrecht, were included in advance. By means of the distribution of the number of inhabitants aged 18–75 over the 24 ($4 \times 5 + 4$) strata, the number of potential respondents was determined per municipality. Then, from the Dutch population register (Basisregistratie Personen), a random sample of individuals aged 18–75 was drawn. Individuals with insufficient command of the Dutch language and institutionalized individuals (i.e.,

those living in hostels, hospices, prisons) were excluded. Individuals who were temporarily living in institutions, however, could be interviewed later during the fieldwork when they returned home.

The study proposal, field procedures and information for respondents were submitted to the Medical Research Ethics Committee (METC Utrecht). Under the Dutch Medical Research Involving Human Subjects Act (WMO) (reference number WAG/mb/19/017126, May 15, 2019), NEMESIS-3 did not require official approval. The field procedures, information for respondents and informed consent form were approved by the local Ethical Review Committee. Respondents provided written informed consent to participate in the interview, after full written and verbal information about the study was given before and at the start of the interview.

Two more waves are planned after baseline, each three years after the previous wave. The second wave starts in January 2023. In order to keep participants engaged, various techniques are applied, such as sending Christmas and birthday cards, sending digital newsletters (7 during the first wave), and adding updates about the study to the website for respondents.

2.2 | Fieldwork

In September 2019, 10 pilot interviews were conducted, resulting in some minor changes in the questionnaires. The fieldwork for the baseline wave was performed from November 2019 to March 2022. A relatively long fieldwork period was chosen to have sufficient time to re-contact potential respondents. This period had to be extended owing to three lockdown periods during the COVID-19 pandemic.

To optimize response, the fieldwork comprised three phases. In the first phase, initial contact attempts were made. Potential respondents were sent an invitation letter and brochure explaining the study goals and procedures. During the COVID-19 pandemic, they also received a flyer on how the face-to-face interview could take place safely. Included was an incentive of €5 as a personal motivation for cooperation. In the experience of the fieldwork agency and other Dutch fieldwork agencies (Scherpenzeel & Toepoel, 2012) offering this amount in advance—together with the invitation letter and brochure—positively influences the response. The letter and brochure included a free phone number for additional information and referred to the study website which contained further information such as answers to frequently asked questions and details about the processing of personal data.

Shortly after the invitation letter was sent, interviewers contacted the selected individuals to seek permission for an interview and to make an appointment. Interviewers were encouraged to make contact in person. They were allowed to make contact by phone if after several efforts face-to-face contact was not achieved or if a potential respondent was 55 years or older. For 77% of all potential respondents a phone number was available. In return for the respondent's time and cooperation, an incentive of €20 was given after the interview. In this first phase, interviewers tried to complete up to 10 phone calls or visits on different days of the week and different

times of the day. To minimize a tendency to focus on easy-to-recruit people, in the first two phases of the fieldwork interviewers were given only small numbers of addresses.

In the second phase, "soft refusers", those unable time-wise or those who could not be reached in the first phase, were re-contacted. In principle, they were contacted in person. Refusers were approached by another interviewer. Before re-contact was made, a glossy postcard was sent. Five different postcards—for refusers (2 for different types of refusal), those unable time-wise, those not reached, and those who mentioned COVID-19 as a reason not to participate—were developed. The incentive was now €25, and later in this phase €40. In this phase, again up to 10 visits or phone calls were made.

Owing to the COVID-19 pandemic, we temporarily had to halt the fieldwork three times during the first two fieldwork phases: from mid-March to late August 2020; from mid-January to early April 2021; and from mid-December 2021 to mid-January 2022. Potential respondents already contacted received a letter to inform them that they would be contacted again after the lockdown and that interview appointments already made during the second and third lockdown could be converted into a video-interview. Only specially trained interviewers were allowed to conduct video-interviews and only with respondents who were considered suitable for this medium. A small pilot study showed that it was possible to perform such interviews and that good contact was possible.

In the third phase ("endgame"), a random selection of 33% of the remaining potential respondents was chosen for special recruitment efforts. The remaining 67% were not approached again. The same strategy was applied in NEMESIS-2 (De Graaf et al., 2010a), NCS-R (National Comorbidity Survey Replication; Kessler et al., 2004) and ESEMeD (European Study of the Epidemiology of Mental Disorders; Alonso et al., 2004) to use the budget in this phase as efficiently as possible. Again, a letter was sent to the randomly selected individuals. The incentive was now €50. Twenty-five interviewers with a high response rate in the previous phases were selected for this phase and were paid extra. In this last phase, up to four visits or phone calls were made. Hard refusers were not approached again.

2.3 | Interviewers

The fieldwork started with 102 professional interviewers headed by three supervisors from the fieldwork agency I&O Research. Selection criteria for interviewers were: experience with systematic face-to-face data collection, experience with sensitive topics, and ability to achieve a good response in other studies. During a 3-day course, interviewers were trained in administering the diagnostic instrument and the additional questionnaire and in motivating potential respondents to participate in the study. Interviewers had to perform a series of test interviews with semi-scripted responses before starting their fieldwork. During the fieldwork, the fieldwork agency had contact with the interviewers on a regular basis. 10 newsletters were sent to interviewers to keep them motivated and to inform them about the study's progress.

Shortly before the first restart of the fieldwork in September 2020, a half day meeting was organized for the interviewers to refresh field work procedures and parts of the diagnostic instrument and additional questionnaire, to discuss fieldwork problems, to exchange tips on improving response, to discuss COVID-19 measures during the interview and to explain the new added COVID-19 questions to the interview. Interviewers who had not yet completed many interviews were given an extra 1 day training. Those who were not able to attend one of these meetings were trained online. Shortly before the second restart of the fieldwork in April 2021, a 2-h online meeting was organized for interviewers with similar topics to those in the previous meeting.

For various reasons, some of the interviewers stopped working in the first year. Therefore, and because of the delay in the fieldwork due to COVID-19, in September 2020, 40 new interviewers were trained, again during a three-day course.

2.4 | Quality control

The NEMESIS-team and the fieldwork agency monitored the fieldwork over the entire data collection period. Every month, the sex and age of all newly interviewed respondents were checked to establish if the right persons had been interviewed. Moreover, comments that interviewers added to the questionnaire responses as well as the response pattern on certain key questions were reviewed. Per interviewer, several items were monitored, such as response rate, interview duration, number of affirmative answers to the diagnostic screener, and respondent's interview evaluation.

The diagnostic instrument starts with a screener comprising key questions on disorders. Respondents who reply in the affirmative to these questions are given the remainder of the disorder section. The answers to these key questions were monitored, because a "learning" effect can occur whereby an interviewer gets into the habit of avoiding parts of the interview by skipping screening questions, resulting in a decrease in prevalence rate of mental disorders over the fieldwork period (see Matschinger et al., 2005). This is also the reason why interviewers were paid separately per interview, interview duration and recruitment time. Payment per completed interview alone could result in interviews being rushed, a tendency to answer screening questions negatively, and a low response rate because interviewers tend to focus on easy-to-recruit respondents (De Graaf et al., 2010a). We found no evidence for a decrease in prevalence rate during the fieldwork period.

Furthermore, back-checks were carried out. Within 2 weeks after the interview 20% of respondents received a short questionnaire to check: whether an interview had taken place, whether a laptop and answer card booklet were used, the length of the interview; and to ask for respondent's interview evaluation. Interviewers were informed several times during the fieldwork period that all these quality control procedures were performed. This monitoring process gave us no reason to doubt the interview quality.

2.5 | Diagnostic instrument

Presence of lifetime and 12-month mental disorders were measured with a slightly adapted version of the Composite International Diagnostic Instrument (CIDI) 3.0 used in NEMESIS-2. The CIDI 3.0 was developed for use in the World Health Organization World Mental Health Survey Initiative (WHO-WMHS). The Dutch CIDI 3.0 was first used in ESEMeD, which was part of this initiative (Alonso et al., 2004; De Graaf et al., 2008). In NEMESIS-2 an improved version of this CIDI 3.0 was used (De Graaf et al., 2010a).

In NEMESIS-3, three types of change were made in the CIDI 3.0: some less relevant questions were removed, a few referrals or skips were adjusted, and some questions were added. These changes are detailed in Table 1. Before omitting any question or changing any referral or skip in the original CIDI 3.0 used in NEMESIS-2, we checked its effect on the prevalence rates within NEMESIS-2. These were only small and non-significant. The differences in 12-month and lifetime prevalence rates before and after we adapted the CIDI 3.0 changed for any mood disorder from 6.1% to 6.0% and from 20.2% to 20.4%, respectively; for any anxiety disorder it stayed at 10.1% and changed from 19.6% to 20.0%; and for any substance use disorder—where we have deliberately built in a slightly stricter threshold—it changed from 5.6% to 5.5% and from 19.1% to 17.5%. The questions we added to the diagnostic instrument were needed to determine DSM-5 disorders, as the original focus of the CIDI 3.0 is to assess DSM-IV disorders.

Most DSM-5 definitions of mental disorders were based on the information already available in the CIDI 3.0 and were reached by making small changes in the algorithms, for example, by omitting a symptom or criterion from the algorithms of a particular disorder (omitting the symptom "legal problems" to diagnose alcohol and drug use disorders and the criterion "bereavement" to diagnose major depressive disorder) or by adding a new symptom to the algorithms of a particular disorder (adding the symptom "craving" to assess alcohol and drug use disorders). The diagnosis "persistent depressive disorder", new in DSM-5, includes both the DSM-IV diagnostic categories of chronic major depression and dysthymia, which were already defined for NEMESIS-2. In the DSM-5, agoraphobia is diagnosed irrespective of the presence of panic disorder. If an individual's presentation meets criteria for panic disorder and agoraphobia, both diagnoses are assigned. In the DSM-IV, agoraphobia was viewed in the context of panic disorder. The criteria for assessing DSM-5 agoraphobia are now similar to those of other phobias. As such these diagnostic criteria have now also been used in the new algorithm of agoraphobia. The most changes were made in the algorithms to assess alcohol and drug use disorders (i.e., omitting the symptom "legal problems", adding the symptom "craving", adding the cluster criterion, and combining the number of abuse and dependency symptoms to diagnose these disorders).

In NEMESIS-3, the following disorders were assessed: mood disorders (major depressive disorder, persistent depressive disorder, bipolar disorder); anxiety disorders (panic disorder, agoraphobia,

TABLE 1 Adaptations made in the CIDI 3.0 to assess both DSM-IV and DSM-5 disorders

<p>Omitting questions</p>	<p>In NEMESIS-3, the following changes were made in the CIDI 3.0. First, as the CIDI is a time-consuming instrument, we shortened it slightly. We omitted questions which were not necessary to assess the presence of a disorder, questions that hardly affected prevalence rates, or questions that were not used for NEMESIS-2 publications. For example, in the CIDI 3.0 used in NEMESIS-2, organic exclusion rules were used to construct psychiatric diagnoses in order to ascertain that certain symptoms were not exclusively due to a somatic cause, an injury, or use of drugs, alcohol or medication. In NEMESIS-3, these questions on the organic cause were omitted, because hardly any respondent in NEMESIS-2 reported that the disorder was exclusively caused by this, and if so, the answer to the open question on the cause was usually not specific enough to exclude the disorder for an organic cause.</p>
<p>Changing referrals or skips</p>	<p>Second, another type of CIDI modification relates to referrals or skips in two disorder sections. The drug use disorder section was adapted in such a way that if a drug was never used more frequently than once a month, no more questions were asked about the symptoms of this specific drug use disorder. Expert consultation supported our notion that this modification most likely increased the validity of these disorders. In addition, we changed some skips in the depression section to keep respondents longer in this section if they did not meet the criteria of a full-blown major depression. This made it possible to ask respondents with subthreshold depression about the number of hours per day they had these depressive feelings, the burden these caused and the presence of other depressive symptoms. With this modification, subthreshold depression can be studied better than in NEMESIS-2 (Tuithof et al., 2018).</p>
<p>Adding questions</p>	<p>Third, another CIDI modification relates to the ability to assess both DSM-IV disorders—the original focus of the CIDI 3.0—and DSM-5 disorders. For example, in the alcohol use disorder section and the drug use disorder section we added questions to measure whether two or more symptoms occurred in the same 12-month period, as this cluster criterion was added in the DSM-5. For drug use disorders, the questions on clustering were assessed for a maximum of two drugs. If a respondent used more than two drugs including cannabis, the questions on clustering always related to cannabis use and the drug that caused the most problems. Because the CIDI 3.0 generates both DSM-IV and international classification of diseases, 10th revision (ICD-10) diagnoses, the new DSM-5 substance use disorder symptom “craving” was already included in the CIDI 3.0. Thus, here we did not have to add a symptom question to the CIDI, but made changes in the algorithms defining a disorder (see main text). To be able to assess ADHD according to DSM-5 criteria, the childhood symptom questions now refer to their presence prior to age 12 instead of age 7, in keeping with the changes from DSM-IV to DSM-5.</p>

social anxiety disorder [social phobia], specific phobia, generalized anxiety disorder); substance use disorders (alcohol and drug use disorders); attention-deficit hyperactivity disorder. In establishing diagnoses, hierarchy rules were used. These diagnoses are used if psychiatric comorbidity is not the subject of study. Suicidality was also measured with the CIDI. In our additional questionnaire, screens for other mental health problems were included (e.g., psychotic experiences, autistic traits, insomnia; see below).

Clinical calibration studies in various countries (Haro et al., 2006) found that the CIDI 3.0 assesses mood, anxiety and substance use disorders with generally good validity in comparison to blinded clinical reappraisal interviews with the Structured Clinical Interview for DSM-IV (SCID). Studies of earlier CIDI versions also concluded that the CIDI assesses common mental disorders with generally acceptable reliability and validity (Andrews & Peters, 1998; Wittchen, 1994). Although the criteria for common mental disorders according to DSM-IV and DSM-5 are quite similar, the validity and reliability of our modified CIDI 3.0 to assess DSM-5 diagnoses have not been investigated.

2.6 | Additional questionnaire

In addition to the CIDI, a further questionnaire was administered. Its design enables a comparison of key questions and scales in NEMESIS-2. The topics included in this questionnaire of the baseline wave of NEMESIS-3 are described in Table 2.

3 | RESULTS

3.1 | Response

Table 3 shows the results of the sample procedure and the response. As in the NCS-R (Kessler et al., 2004), ESEMeD (Alonso et al., 2004) and NEMESIS-2 (De Graaf et al., 2010a), the response rate calculation was based on weighted data because in the endgame phase of the hard-to-recruit respondents 33% of the unresolved cases from phase 1 and 2 were chosen for special recruitment efforts. The final weighted response rate was 54.6%. In phases 1, 2 and 3, the response was 41.8%, 16.3% and 13.1%, respectively. Of the 6194 respondents, 5312 were interviewed in phase 1, 674 in phase 2 and 208 in phase 3.

We examined to what extent respondents differed per phase of the fieldwork, thus by amount of recruitment efforts. Those who were interviewed in phase 2 were more often employed, lived without a partner, and lived in urban areas, than those who were interviewed in phase 1. They also were less often higher educated and less often 65 years or older. Compared to phase 1, interviewees in phase 3 less often had higher vocational or university education. No differences were found between phase 2 and 3 respondents. More importantly, respondents interviewed in the different phases of the fieldwork did not differ in prevalence of lifetime and 12-month mental disorder categories (mood, anxiety, substance use), after controlling for differences in sociodemographic characteristics between the groups and in being interviewed before or during the COVID-19 pandemic.

TABLE 2 Topics included in the additional questionnaire of the baseline wave of NEMESIS-3

Socio-demographics	Sex, age, educational attainment, Dutch or non-Dutch origin, religion, urbanicity of place of residence, living situation, number and age of children, employment situation and income of respondent and her/his partner.
Psychotic experiences	Questions on psychotic experiences (delusions and hallucinations) constituted an extension and improvement to those of the CIDI 1.1, and were used in NEMESIS-2 as well (Honings et al., 2016; Van Nierop et al., 2012).
Autistic traits	These are measured with the short version of the autism spectrum quotient (AQ). This encompasses the most discriminating 10 items of the full-length version of the AQ (Allison et al., 2012). The AQ has good discriminative validity and screening properties. The short version includes 10 statements where respondents have to indicate the degree to which they agree or disagree with the statement. The statements relate to subdomains that are characteristic for autism spectrum conditions, for example, attention to detail, attention switching, communication, imagination, and social interaction.
Insomnia	Insomnia was assessed with the Women's health initiative Insomnia Rating Scale (IRS; Levine et al., 2003; Ten Have et al., 2016), which consists of five questions concerning sleep in the past month. The questions address difficulty falling asleep, waking up during the night, early morning awakening, trouble getting back to sleep after waking up, and sleep quality. Validation studies on the IRS found a high test-retest reliability and strong associations with other actigraphy-derived sleep measures (Levine et al., 2003).
Tobacco use	Smoking history, frequency in previous 4 weeks, frequency of use of e-cigarettes (with nicotine) in previous 4 weeks were assessed.
Somatic health problems	With regard to 17 chronic somatic disorders, questions were asked on presence, treatment and burden in the previous 12 months, and age of onset, as was done in NEMESIS-2 (De Graaf et al., 2010a). Comparisons between self-reports of chronic physical disorders and medical records show moderate to good concordance (Baker et al., 2004; National Center for Health Statistics, 1994).
Physical activity	Physical activity was measured by number of physically active days, that is, performing at least moderate intensive activities for at least 30 min a day, in an average week, based on the international physical activity questionnaire (IPAQ) (Craig et al., 2003). Also, the number of hours per week engaging in physical exercise/sport was recorded.
Loneliness	Loneliness was assessed with the 11-item De Jong Gierveld loneliness scale (De Jong-Gierveld & Kamphuis, 1985), a commonly used and psychometrically sound measure of loneliness (De Jong-Gierveld & Van Tilburg, 1999).
Childhood adversities	Childhood adversities were measured with six negative life events before the age of 16 (such as death of parent and divorce of parents), and by items on abuse and neglect during one's youth (frequency of emotional neglect, psychological, physical and sexual abuse, and bullying-victimization), as was done in NEMESIS-2 (Ten Have et al., 2019).
Negative life events	The presence of 10 negative life events in the previous 12 months was measured, based on the brugha life events section (Brugha et al., 1985). Examples are: death of a relative or friend, divorce, and financial difficulties.
Functioning	General functioning was assessed by the Medical Outcomes Study (MOS) short-form health survey (SF-36) (Stewart et al., 1988; Ware & Sherbourne, 1992). "Days out of role" was measured as number of days lost from work or other normal activities (three questions based on the World Health Organization (WHO) disability assessment schedule (WHODAS) (Von Korff et al., 2008). Furthermore, three questions were asked about memory problems.
Emotional exhaustion	This was assessed among respondents with a paid job with the 5-item emotional exhaustion scale of the utrecht burnout scale (the Dutch version of the Maslach Burnout Inventory (MBI)—General survey (Schaufeli et al., 1996; Schaufeli & van Dierendonck, 2000).
Personal functioning, wellbeing	This was assessed with an adapted version of the 5-item brief INSPIRE (Williams et al., 2015; https://www.researchintorecovery.com/inspire%23s3). The brief INSPIRE measures 5 important domains of recovery: Connectedness, hope, identity, meaning, and purpose and empowerment. The original version asks how the care provider supports respondent's recovery (e.g., My care provider helps me to feel supported by other people), the adapted version asks how the respondent scores on important domains of recovery without reference to care provider's support (e.g., I feel supported by other people).
Service use	Service use was measured by means of the service use section of NEMESIS-1 and -2 (Bijl & Ravelli, 2000b; Ten Have et al., 2013a, 2013b). Service use was assessed with the question: "In the previous 12 months, have you attended. . . for emotional problems or alcohol or drugs problems of your own?" this question was asked in respect of the following care providers/services: General medical professionals (general practitioners; mental health nurses; company doctors; social work; home care or district nurses; physiotherapists or haptonomists; medical specialists or other professionals working within this care sector), mental health services (psychiatrists, psychologists, psychotherapists, part-time or full-time psychiatric treatment), or other types of care (alternative care providers, pastoral care, self-help groups or telephone help lines). If the respondent had attended a general medical or mental health care provider/service, information on the type of care, number of visits, and age of onset of treatment were inquired. Furthermore, medication use for emotional or substance use problems in the previous 12 months, duration and onset, and unmet need of care in the previous 12 months, were measured.

TABLE 2 (Continued)

Sexual orientation	Sexual orientation was asked with a question on attraction to women or/and men with 5 categories.
Femininity/masculinity	Self-ascribed femininity (for women) or masculinity (for men) was assessed with a question with 7 categories (based on Bockting et al., 2009).
Impact of COVID-19	Before the restart of the fieldwork in Sep 2020 we added questions on the impact of the COVID-19 pandemic: Having been ill or a close or loved one having been ill because of COVID-19 (including its severity), having had to work from home, (threatened) job loss, worry about one's own health or health of a close or loved one, worry about keeping one's work and about income; and changes in lifestyle behaviors (physical activity, alcohol use, smoking, sleep), well-being (worry, anxiety, depressed mood, being irritated or tense, feelings of loneliness), as well as changes in the relationship with partner and children during the COVID-19 pandemic compared to the period before.

TABLE 3 Response and non-response in NEMESIS-3

	N	Percentage
Initial sample ^a	15,067	
Gross sample	14,585	100
Deceased	93	0.6
Moved to unknown address, or unknown at the address	1244	8.5
Problems with Dutch language	650	4.5
Cognitive problems	41	0.3
Long term unable ^b	269	1.8
Invalid address found out during the fieldwork	155	1.1
Net sample of eligible persons	12,133	83.2
Net sample of eligible persons	12,133	100
Refusal ^{c,d}	4532	37.4
Temporarily unable ^{c,e}	416	3.4
No contact ^c	957	7.9
Unfinished interview, or appointment made after the fieldwork period ^f	34	0.3
Unweighted uncorrected response	6194	51.1
Unweighted corrected response ^g	6194	51.2
Weighted corrected response ^h	6194	54.6

^aOf the 15,067 addresses, 482 were not part of the sample for the following reasons: no residential address (208), secret address (270), moved abroad (4).

^bLong term unable: physically unable (111), mentally unable (158).

^cThe non-response categories are the reasons of the last contact attempt.

^dRefusal: refusal by the respondent (3,841), by someone else (691).

^eTemporarily unable: physically unable (15), mentally unable (10), other reason (391).

^fUnfinished interview (19), appointment made after the fieldwork period (11), not present at the interview appointment (4).

^gIn calculating this response, we assumed a slightly lower net sample of eligible persons. As there were 180 people who could not be contacted during the entire fieldwork period, we assume that a similar proportion of these people is not eligible due to language problems or other reasons ($100\% - 83.2\% = 16.8\%$; $180 \cdot 0.168 = 30$). Therefore, the net sample of eligible persons becomes: $12,133 - 30 = 12,103$.

^hThe response rate calculation was based on weighted data because the endgame phase undersampled hard-to-recruit respondents, as a random 67% of the unresolved cases in phase 1 and 2 were terminated and 33% were chosen for special recruitment efforts in the endgame phase.

Table 3 also shows the reasons for non-response at the last contact attempt. If we define refusal not on the basis of the last contact attempt but on ever having refused during the different contact attempts in the

different phases, refusal was more often seen (43.0%). Fear of contamination by COVID-19 during the interview was mentioned as the reason for refusal by 2.1% of the refusers after the first lockdown.

Partial non-response was negligible due to the computer-assisted face-to-face interview method.

3.2 | Sample characteristics

Table 4 shows the distribution of sociodemographic variables of the study sample (first column) and the total population according to Statistics Netherlands (1-1-2020; www.cbs.nl) (last column). The mean age of the sample is 47.9 years [standard deviation (SD) = 16.4] with 50.4% female; 41.5% with higher vocational/university education; 67.2% living with a partner; 67.5% with paid employment; and 17.3% of non-Dutch origin. Non-Dutch origin was defined as the

respondent or at least one parent not born in the Netherlands. Respondents reflected the Dutch population reasonably well. However, people of 18–34 years, higher secondary educated people, those not living with a partner, people living in bigger towns, and people of non-Dutch origin were somewhat underrepresented.

To facilitate generalization of the data to the Dutch general population, based on post-stratification, a weighting factor was constructed to correct for different response rates in the various sociodemographic groups, and for differences in the probability of selection in phase 3. The following sociodemographic characteristics of the population from 2020 obtained from Statistics Netherlands were used: sex, age, partner status (living with or without partner), educational level (three categories) and urbanicity (six categories).

	NEMESIS-3		Dutch population
	Unweighted	Weighted	
Sex			
Male	49.6	50.0	50.0
Female	50.4	50.0	50.0
Age at interview			
18–24	10.7	12.1	12.0
25–34	15.1	17.5	17.5
35–44	16.2	16.2	16.2
45–54	17.7	19.4	19.4
55–64	20.4	18.6	18.6
65–75	19.8	16.3	16.3
Education			
Primary education, lower secondary	22.1	23.2	23.2
Higher secondary	36.5	42.2	42.2
Higher vocational, university	41.5	34.6	34.6
Living situation			
With partner	67.2	62.9	63.0
Without partner	32.8	37.1	37.0
Employment situation			
Paid job	67.5	69.0	68.4
No paid job	32.5	31.0	31.6
Urbanicity			
Very high	22.6	25.9	25.9
High	29.4	30.4	30.4
Medium	16.1	15.1	15.1
Low	22.8	20.9	20.9
Very low	9.2	7.6	7.6
Dutch or non-Dutch origin			
Dutch	82.7	81.2	75.2
Non-Dutch	17.3	18.8	24.8

TABLE 4 Demographic characteristics of the NEMESIS-3 sample ($N = 6194$) and of the Dutch population according to Statistics Netherlands, in percentages

The weighting factor was constructed with the statistics programme R (R Core Team, 2021), using the “survey” package (Lumley, 2010). After weighting, the distribution of the sociodemographic characteristics of the study sample came very close to that of the Dutch population.

3.3 | Interview characteristics

In total, 1576 respondents (25.4%) were interviewed before and 4618 respondents (74.6%) during the COVID-19 pandemic.

Almost all interviews were held at the respondent's home. Although it was stated at recruitment that the interview should be held privately, in 14.0% of the interviews another person was present. In 68.7% of these cases ($n = 595$; 9.6%) this was for at least half of the time. A significantly lower prevalence of mental disorders was found among those interviewed with another person present for at least half of the time compared to the other respondents [any lifetime disorder: odds ratio (OR) = 0.76 (95% confidence interval CI = 0.64–0.91); any 12-month disorder: OR = 0.79 (95% CI = 0.63–0.99); controlled for demographic characteristics.

Most interviews were face-to-face (laptop computer-assisted). A total of 500 (8.1%) interviews were done via video calling. These cases were more often younger, higher educated, employed and living in urban areas. Adjusted for these demographic characteristics, they did not differ in prevalence rate of any 12-month (OR = 0.95; 95% CI = 0.76–1.19) and lifetime disorder (OR = 1.10; 95% CI = 0.91–1.33). This was also true for those 4618 respondents interviewed during the COVID-19 pandemic.

The average duration of the interview was 91 min: 43 min for the diagnostic instrument the Composite International Diagnostic Interview (CIDI) and 48 min for the additional questionnaire. Interview duration varied widely, mainly depending on the number of mental disorders a respondent reported having ever experienced. Video-interviews lasted on average 92 min.

After the interview, respondents were asked to evaluate it: 82.9% rated it positively, 16.4% neutrally and 0.8% negatively. Compared to those with a positive evaluation, those who were negative or neutral were more often female, living without a partner, unemployed, living in urban areas, more often had any lifetime or 12-month mental disorder, and had a longer interview duration. No differences were found for age, education and interview mode (face-to-face vs. video).

4 | DISCUSSION

With the fieldwork of the first wave of NEMESIS-3 we were able to build a large and comprehensive dataset of high quality. Because we adapted the CIDI 3.0 by adding questions, we are now able to assess both DSM-IV and DSM-5 diagnoses. This gives us the opportunity to calculate up-to-date prevalence rates as well as to examine trends in prevalence rates of DSM-IV mental disorders between NEMESIS-2

and NEMESIS-3. As the fieldwork of the first wave took place before and during the COVID-19 pandemic, an additional aim now is to investigate the extent to which the pandemic has had an effect on the population's mental health. To date, few studies have addressed this question by comparing prevalence rates of mental disorders before and during the pandemic (Kessler et al., 2022). As we expanded the age range of respondents from 18 to 64 years in the previous NEMESIS-studies to 18–75 years, it is now possible to study the mental health of those aged 65 and over as well, an age group that is often underrepresented in psychiatric epidemiology studies (Andreas et al., 2017). For each measurement wave, new topics will be included that we can relate to mental health.

In NEMESIS-3, a random sample of individuals was drawn. In the previous NEMESIS-studies, a random sample of addresses of private households was drawn and the last-birthday method was used to select the individual within a household. Particularly in NEMESIS-3 and to a lesser extent NEMESIS-2 the recruitment methods were intensive: visit in person as much as possible, several times in cases of no previous hard refusal, a long fieldwork period with different phases, interviewer briefings, the use of respondent letters, brochures and postcards, and respondents' incentives. For NEMESIS-3, this resulted in a response rate of 54.6%, which is lower than the first wave of NEMESIS-1 in 1996 (64.2%; Bijl, van Zessen, et al., 1998) and of NEMESIS-2 performed between 2007 and 2009 (65.1% De Graaf et al., 2010a).

This is in line with an international trend toward declining response rates in all types of surveys (De Leeuw et al., 2018), which is attributed to societal changes (e.g., smaller households, work/life balance, privacy concerns), technological innovations (e.g., mobile phones and surveys), decreasing trust in surveys and increasing survey burden (Beullens et al., 2018). For example, in the European Social Surveys (ESS) conducted between 2006 and 2018 the response rate fell from 59.8% to 49.6%.

A small part of the decline in response rate in NEMESIS-3 may be due to the fact that we did not use a last-birthday selection method that usually generates lower refusal rates (www.european-socialsurvey.org: round 9 ESS sampling guidelines), and may be the result of the COVID-19 pandemic. Although fear of contamination was not often given explicitly as reason for non-response, the interviewers and their supervisors felt that it was in fact more often the reason for refusal. Owing to the COVID-19 pandemic, the fieldwork had to be stopped three times. Although potential respondents who had already been approached were contacted later, we assume that this also had a negative effect on the response rate.

With similar recruitment efforts, the refusal (and non-contact) conversion in phase 2 and 3 was lower in NEMESIS-3 compared to NEMESIS-2 (response in phase 1-3: 42%, 16% and 13%; and 43%, 28% and 20%, respectively). It appears that nowadays people are less inclined to take part once they have refused before.

Although the sample was reasonably nationally representative, people of 18–34 years, higher secondary educated people, those not living with a partner, people living in bigger towns, and people of non-Dutch origin were somewhat underrepresented. The fact that males

were hardly underrepresented is remarkable because in NEMESIS-2 (De Graaf et al., 2010a) and NEMESIS-1 (Bijl, van Zessen, et al., 1998), the opposite was the case. This is probably because we made use of a random sample of individuals instead of households using the last-birthday method to select an individual within a household, which is less affected by selection bias (www.european-socialsurvey.org: round 9 ESS sampling guidelines). The relatively high response rate among subjects of 18–24 years old in NEMESIS-3 may also be attributed to this method. Moreover, possibly on account of the COVID-19 pandemic these younger age groups were more often at home than during the fieldwork periods of the former NEMESIS studies. In NEMESIS-2 this was the age category with the lowest response rate (De Graaf et al., 2010a).

We do not know in NEMESIS-3 whether non-respondents are more or less likely to have mental disorders than respondents. From previous NEMESIS-studies we know that respondents who no longer participated in multiple follow-up waves do not differ in psychopathology at baseline from the people who continued to participate, after controlling for sociodemographic differences between the groups (De Graaf et al., 2018). The finding that NEMESIS-3 respondents interviewed in the three fieldwork phases with increasing recruitment efforts did not differ in lifetime and 12-month prevalence of three main categories of common mental disorders—a result also found in NEMESIS-2—shows that hard-to-persuade or hard-to-reach people do not differ in psychopathology. Whether the non-respondents also do not differ from the respondents in the prevalence of the disorders assessed in NEMESIS-3 remains unknown.

Despite the COVID-19 pandemic and the inclusion of 65–75 year old respondents in NEMESIS-3, the percentage of those interviewed with another person present for at least half of the time was lower than in NEMESIS-2 (9.6% vs. 12.6%). But in contrast to NEMESIS-2 (De Graaf et al., 2010a), a significantly lower prevalence of any lifetime and 12-month mental disorder was found among these respondents. Underreporting of mental disorders among those interviewed with another person present is the most plausible reason for this finding since we adjusted the analyses for sociodemographic differences between the groups. We have no clear explanation why these results differ from our previous study. But it shows how important it is to train interviewers to explain to the respondent at different opportunities that the interview should be held in private.

Finally, the vast majority of the NEMESIS-3 respondents evaluated the interview positively (82.9%) and this was even much better than in NEMESIS-2 (67.9%). Those who were negative or neutral were more often female, living without a partner, unemployed, living in urban areas, more often had any lifetime or 12-month mental disorder, and had a longer interview duration. Most of these findings are similar to those from NEMESIS-1 and NEMESIS-2 (De Graaf et al., 2004, 2010a). The finding that the interview was generally perceived as positive is important, especially for ethics committees assessing the burden of respondents to participate in studies such as NEMESIS.

ACKNOWLEDGEMENTS

The Netherlands Mental Health Survey and Incidence Study-3 (NEMESIS-3) is being conducted by the Netherlands Institute of Mental Health and Addiction (Trimbos Institute) in Utrecht. Financial support has been received from the Ministry of Health, Welfare and Sport. The fieldwork for the first wave was carried out by I&O Research in Enschede, the Netherlands. We thank the advisory committee for their contribution to the study: Prof. Wim van den Brink (chair), Carolien van den Brink, PhD, Prof. Peter de Jonge, Prof. Dike van de Mheen, Prof. Brenda Penninx, Didi Rhebergen, PhD, and Ineke Stoop, PhD.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest that could be perceived as prejudicing the impartiality of the research reported.

DATA AVAILABILITY STATEMENT

The data on which this manuscript is based are not publicly available. However, data from NEMESIS-3 are available upon request. The Dutch ministry of health financed the data and the agreement is that these data can be used freely under certain restrictions and always under supervision of the Principal Investigator (PI) of the study. Thus, some access restrictions do apply to the data. The PI of the study is first author of this paper and can at all times be contacted to request data. At any time, researchers can contact the PI of NEMESIS-3 and submit a research plan, describing its background, research questions, variables to be used in the analyses, and an outline of the analyses. If a request for data sharing is approved, a written agreement will be signed stating that the data will only be used for addressing the agreed research questions described and not for other purposes.

ORCID

Margreet ten Have  <https://orcid.org/0000-0003-0086-2199>

REFERENCES

- Allison, C., Auyeung, B., & Baron-Cohen, S. (2012). Toward brief "red flags" for autism screening: The short autism spectrum quotient and the short quantitative checklist in 1,000 cases and 3,000 controls. *Journal of the American Academy of Child & Adolescent Psychiatry*, 51(2), 202–212. <https://doi.org/10.1016/j.jaac.2011.11.003>
- Alonso, J., Angermeyer, M., Bernert, S., Bruffaerts, R., Brugha, T. S., Bryson, H., de Girolamo, G., de Graaf, R., Demyttenaere, K., Gasquet, I., Haro, J. M., Katz, S. J., Kessler, R. C., Kovess, V., Lépine, J. P., Ormel, J., Polidori, G., Russo, L. J., Vilagut, G., ... Vollebergh, W. A. M. (2004). Sampling and methods of the European Study of the Epidemiology Of Mental Disorders (ESEMeD) project. *Acta Psychiatrica Scandinavica*, 109(Suppl. 420), 8–20. <https://doi.org/10.1111/j.1600-0047.2004.00327.x>
- Andreas, S., Schulz, H., Volkert, J., Dehoust, M., Sehner, S., Suling, A., Ausín, B., Canuto, A., Crawford, M., Da Ronch, C., Grassi, L., Hershkovitz, Y., Muñoz, M., Quirk, A., Rotenstein, O., Santos-Olmo, A. B., Shalev, A., Strehle, J., Weber, K., ... Härter, M. (2017). Prevalence of mental disorders in elderly people: The European MentDis_ICF65+ study. *British Journal of Psychiatry*, 210(2), 125–131. <https://doi.org/10.1192/bjp.bp.115.180463>
- Andrews, G., & Peters, L. (1998). The psychometric properties of the composite international diagnostic interview. *Social Psychiatry and*

- Psychiatric Epidemiology*, 33(2), 80–88. <https://doi.org/10.1007/s001270050026>
- Baker, M. M., Stabile, M., & Deri, C. (2004). What do self-reported, objective measures of health measure? *Journal of Human Resources*, 4(4), 1067–1093. <https://doi.org/10.3368/jhr.XXXIX.4.1067>
- Beullens, K., Loosveldt, G., Van den plas, C., & Stoop, I. (2018). Response rates in the European social survey: Increasing, decreasing, or a matter of fieldwork efforts? *Survey Methods: Insights from the Field*, 1–2. <https://doi.org/10.13094/SMIF-2018-00003>
- Bijl, R. V., & Ravelli, A. (2000a). Current and residual functional disability associated with psychopathology: Findings from The Netherlands Mental Health Survey and Incidence Study (NEMESIS). *Psychological Medicine*, 30(3), 657–668. <https://doi.org/10.1017/S0033291799001841>
- Bijl, R. V., & Ravelli, A. (2000b). Psychiatric morbidity, service use, and need for care in the general population: Results of The Netherlands Mental Health Survey and Incidence Study. *American Journal of Public Health*, 90, 602–607. <https://doi.org/10.2105/AJPH.90.4.602>
- Bijl, R. V., van Zessen, G., & Ravelli, A. (1998). Prevalence of psychiatric disorder in the general population: Results of The Netherlands Mental Health Survey and Incidence Study (NEMESIS). *Social Psychiatry and Psychiatric Epidemiology*, 33(12), 587–595. <https://doi.org/10.1007/s001270050098>
- Bijl, R. V., van Zessen, G., Ravelli, A., de Rijk, C., & Langendoen, Y. (1998). The Netherlands mental health survey and incidence study (NEMESIS): Objectives and design. *Social Psychiatry and Psychiatric Epidemiology*, 33(12), 581–586. <https://doi.org/10.1007/s001270050097>
- Bockting, W., Benner, A., & Coleman, E. (2009). Gay and bisexual identity development among female-to-male transsexuals in North America: Emergence of a transgender sexuality. *Archives of Sexual Behavior*, 38(5), 688–701. <https://doi.org/10.1007/s10508-009-9489-3>
- Bos, E. H., ten Have, M., van Dorsselaer, S., Jeronimus, B. F., de Graaf, R., & de Jonge, P. (2018). Functioning before and after a major depressive episode: Pre-existing vulnerability or scar? A prospective three-wave population-based study. *Psychological Medicine*, 48(13), 2264–2272. <https://doi.org/10.1017/S0033291717003798>
- Brugha, T., Bebbington, P., Tennant, C., & Hurry, J. (1985). The list of threatening experiences: A subset of 12 life event categories with considerable long-term contextual threat. *Psychological Medicine*, 15(1), 189–194. <https://doi.org/10.1017/S003329170002105X>
- Buist-Bouwman, M. A., Ormel, J., de Graaf, R., & Vollebergh, W. A. M. (2004). Functioning after a major depressive episode: Complete or incomplete recovery? *Journal of Affective Disorders*, 82, 363–371. <https://doi.org/10.1016/j.jad.2004.02.007>
- Craig, C. L., Marshall, A. L., Sjostrom, M., Bauman, A. E., Booth, M. L., Ainsworth, B. E., Pratt, M., Ekelund, U., Yngve, A., Sallis, J. F., & Oja, P. (2003). International physical activity questionnaire: 12-country reliability and validity. *Medicine & Science in Sports & Exercise*, 35(8), 1381–1395. <https://doi.org/10.1249/01.mss.0000078924.61453.fb>
- De Graaf, R., Ormel, J., ten Have, M., Burger, H., & Buist-Bouwman, M. (2008). Mental disorders and service use in The Netherlands. Results from the European study of the epidemiology of mental disorders (ESEMeD). In *The WHO World mental health surveys: Global perspectives on the epidemiology of mental disorders* (pp. 388–405). Cambridge University Press.
- De Graaf, R., Ten Have, M., Tuithof, M., & Van Dorsselaer, S. (2013). First-incidence of DSM-IV mood, anxiety and substance use disorders and its determinants: Results from The Netherlands mental health survey and incidence study-2. *Journal of Affective Disorders*, 149(1–3), 100–107. <https://doi.org/10.1016/j.jad.2013.01.009>
- De Graaf, R., ten Have, M., & van Dorsselaer, S. (2010b). *De psychische gezondheid van de Nederlandse bevolking. NEMESIS-2: Opzet en eerste resultaten*. Trimbos-instituut.
- De Graaf, R., ten Have, M., & van Dorsselaer, S. (2010a). The Netherlands mental health survey and incidence study-2 (NEMESIS-2): Design and methods. *International Journal of Methods in Psychiatric Research*, 193(3), 125–141. <https://doi.org/10.1002/mpr.317>
- De Graaf, R., Ten Have, M., Van Dorsselaer, S., Schoemaker, C., & Vollebergh, W. A. M. (2004). Negative and positive participant responses to the composite international diagnostic interview. Results of The Netherlands mental health survey and incidence study. *Social Psychiatry and Psychiatric Epidemiology*, 39(7), 521–527. <https://doi.org/10.1007/s00127-004-0773-1>
- De Graaf, R., ten Have, M., van Gool, C., & van Dorsselaer, S. (2012a). Prevalence of mental disorders, and trends from 1996 to 2009. Results from The Netherlands mental health survey and incidence study-2. *Social Psychiatry and Psychiatric Epidemiology*, 47(2), 203–213. <https://doi.org/10.1007/s00127-010-0334-8>
- De Graaf, R., Tuithof, M., van Dorsselaer, S., & ten Have, M. (2012b). Comparing the effects on work performance of mental and physical disorders. *Social Psychiatry and Psychiatric Epidemiology*, 47(11), 1873–1883. <https://doi.org/10.1007/s00127-012-0496-7>
- De Graaf, R., Van Dorsselaer, S., Tuithof, M., & Ten Have, M. (2018). Sociodemographic and psychiatric predictors of attrition in the third follow-up of The Netherlands Mental Health Survey and Incidence Study-2 (NEMESIS-2). Trimbos-instituut.
- De Jong-Gierveld, J., & Kamphuis, F. (1985). The development of a rasch-type loneliness scale. *Applied Psychological Measurement*, 9(3), 289–299. <https://doi.org/10.1177/014662168500900307>
- De Jong-Gierveld, J., & van Tilburg, T. (1999). Manual loneliness scale. (Updated from the printed version: 15-6-2017). https://home.fsw.vu.nl/tg.van.tilburg/manual_loneliness_scale_1999.html
- De Leeuw, E., Hox, J., & Luiten, A. (2018). International nonresponse trends across countries and years: An analysis of 36 years of labour force survey data. *Survey insights: Methods from the field*. <https://doi.org/10.13094/SMIF-2018-00008> Retrieved from <https://surveyinsights.org/?p=10452>
- Grant, B. F., Goldstein, R. B., Chou, S. P., Huang, B., Stinson, F. S., Dawson, D. A., Saha, T. D., Smith, S. M., Pulay, A. J., Pickering, R. P., Ruan, W. J., & Compton, W. M. (2009). Sociodemographic and psychopathological predictors of first incidence of DSM-IV substance use, mood and anxiety disorders: Results from the wave 2 national epidemiologic survey on alcohol and related conditions. *Molecular Psychiatry*, 14(11), 1051–1066. <https://doi.org/10.1038/mp.2008.41>
- Haro, J. M., Arbabzadeh-Bouchez, S., Brugha, T. S., De Girolamo, G., Guyer, M. E., Jin, R., Lepine, J. P., Mazzi, F., Reneses, B., Vilagut, G., Sampson, N. A., & Kessler, R. C. (2006). Concordance of the composite international diagnostic interview version 3.0 (CIDI 3.0) with standardized clinical assessments in the WHO World mental health surveys. *International Journal of Methods in Psychiatric Research*, 15(4), 167–180. <https://doi.org/10.1002/mpr.196>
- Honings, S., Drukker, M., van Nierop, M., van Winkel, R., Wittchen, H.-U., Lieb, R., ten Have, M., de Graaf, R., van Dorsselaer, S., & van Os, J. (2016). Psychotic experiences and incident suicidal ideation and behaviour: Disentangling the longitudinal associations from connected psychopathology. *Psychiatric Research*, 245, 267–275. <https://doi.org/10.1016/j.psychres.2016.08.002>
- Kessler, R. C., Angermeyer, M., Anthony, J. C., de Graaf, R., Demyttenaere, K., Gasquet, I., De Girolamo, G., Gluzman, S., Guereje, O., Haro, J. M., Kawakami, N., Karam, A., Levinson, D., Medina Mora, M. E., Oakley Browne, M. A., Posada-Villa, J., Stein, D. J., Tsang, C. H. A., Aguilar-Gaxiola, S., ... Ustun, T. B. (2007). Lifetime prevalence and age-of-onset distributions of mental disorders in the World health organization's World mental health surveys. *World Psychiatry*, 6, 168–176.
- Kessler, R. C., Berglund, P., Chiu, W. T., Demler, O., Heeringa, S., Hiripi, E., Jin, R., Pennell, B. E., Walters, E. E., Zaslavsky, A., & Zheng, H. (2004). The US National Comorbidity Survey Replication (NCS-R): Design

- and field procedures. *International Journal of Methods in Psychiatric Research*, 13(2), 69–92. <https://doi.org/10.1002/mpr.167>
- Kessler, R. C., Ruhm, C. J., Puac-Polanco, V., Hwang, I. H., Lee, S., Petukhova, M. V., Sampson, N. A., Ziobrowski, H. N., Zaslavsky, A. M., & Zubizarreta, J. R. (2022). Estimated prevalence of and factors associated with clinically significant anxiety and depression among US adults during the first year of the COVID-19 pandemic. *JAMA Network Open*, 5(6), e2217223. <https://doi.org/10.1001/jamanetworkopen.2022.17223>
- Levine, D. W., Lewis, M. A., Bowen, D. J., Kripke, D. F., Kaplan, R. M., Naughton, M. J., & Shumaker, S. A. (2003). Reliability and validity of women's health initiative insomnia rating scale. *Psychological Assessment*, 15(2), 137–148. <https://doi.org/10.1037/1040-3590.15.2.137>
- Lumley, T. (2010). *Complex surveys: A guide to analysis using R*. John Wiley and Sons.
- Matschinger, H., Bernert, S., & Angermeyer, M. C. (2005). An analysis of interviewer effects on screening questions in a computer assisted personal mental health interview. *Journal of Official Statistics*, 21, 657–674.
- National Center for Health Statistics. (1994). *Vital and health statistics. Evaluation of national health interview survey diagnostic reporting. Series 2: Data evaluation and methods research*. Department of Health and Human Services no.120.
- Ormel, J., Oldehinkel, A. J., Nolen, W. A., & Vollebergh, W. (2004). Psychosocial disability before, during, and after a major depressive episode. A 3-wave population-based study of state, scar and trait effects. *Archives of General Psychiatry*, 61(10), 387–392. <https://doi.org/10.1001/archpsyc.61.10.990>
- R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Retrieved from <http://www.R-project.org/>
- Ruths, S., Haukenes, I., Hetlevik, Ø., Smith-Sivertsen, T., Hjørleifsson, S., Hansen, A. B., Riiser, S., Meling, H. M., & Baste, V. (2021). Trends in treatment for patients with depression in general practice in Norway, 2009–2015: Nationwide registry-based cohort study (the Norwegian GP-DEP study). *BMC Health Services Research*, 21(1), 697. <https://doi.org/10.1186/s12913-021-06712-w>
- Schaufeli, W. B., Leiter, M. P., Maslach, C., & Jackson, S. E. (1996). Maslach burnout inventory-general survey (MBI-GS). In C. Maslach, S. E. Jackson, & M. P. Leiter (Eds.), *The Maslach burnout inventory manual*. Consulting Psychologists Press.
- Schaufeli, W. B., & van Dierendonck, D. (2000). *Utrechtse burn-out schaal: Handleiding*. Pearson Assessment and Information B.V.
- Scherpenzeel, A., & Toepoel, V. (2012). Recruiting a probability sample for an online panel: Effects of contact mode, incentives, and information. *Public Opinion Quarterly Advance*, 76(3), 470–490. <https://doi.org/10.1093/poq/nfs037>
- Scholten, W., Ten Have, M., Van Geel, C., Van Balkom, A., De Graaf, R., & Batelaan, N. (2021). Recurrence of anxiety disorders and its predictors in the general population. *Psychological Medicine*, 1–9. <https://doi.org/10.1017/S0033291721002877>
- Schopman, S., ten Have, M., van Dorsselaer, S., de Graaf, R., & Batelaan, N. (2018)., & MP Leiter Limited functioning after remission of an anxiety disorder as a trait effect versus a scar-effect: Results of a longitudinal general population study. *Journal of Clinical Psychiatry*, 79(3). <https://doi.org/10.4088/JCP.16m11256>
- Schopman, S. M. E., Ten Have, M., Van Balkom, A. J., De Graaf, R., & Batelaan, N. M. (2021). Course trajectories of anxiety disorders: Results from a 6-year follow-up in a general population study. *Australian and New Zealand Journal of Psychiatry*, 55(11), 1–9. <https://doi.org/10.1177/00048674211009625>
- Smit, H. F. E., Cuijpers, P., Oostenbrink, J., Batelaan, N. M., de Graaf, R., & Beekman, A. J. (2006). Costs of nine common mental disorders: Implications for curative and preventive psychiatry. *The Journal of Mental Health Policy and Economics*, 9(4), 193–200. <http://www.icmpe.org/test1/journal/journal.htm>
- Spijker, J., de Graaf, R., Bijl, R. V., Beekman, A. T. F., Ormel, J., & Nolen, W. A. (2002). Duration of major depressive episodes in the general population. Results of The Netherlands mental health survey and incidence study (NEMESIS). *British Journal of Psychiatry*, 181(3), 208–213. <https://doi.org/10.1192/bjp.181.3.208>
- Steffen, A., Thom, J., Jacobi, F., Holstiege, J., & Bätzing, J. (2020). Trends in prevalence of depression in Germany between 2009 and 2017 based on nationwide ambulatory claims data. *Journal of Affective Disorders*, 271, 239–247. <https://doi.org/10.1016/j.jad.2020.03.082>
- Stewart, A. L., Hayes, R. D., & Ware, J. E. (1988). The MOS short form general health survey: Reliability and Validity in a Patient Population. *Medical Care*, 26(7), 724–735. <https://doi.org/10.1097/00005650-198807000-00007>
- Ten Have, M., de Graaf, R., van Dorsselaer, S., & Beekman, A. (2013a). Lifetime treatment contact and delay in treatment seeking after first onset of a mental disorder. *Psychiatric Services*, 64(10), 981–989. <https://doi.org/10.1176/appi.ps.201200454>
- Ten Have, M., de Graaf, R., van Dorsselaer, S., Tuithof, M., Kleinjan, M., & Penninx, B. (2019). Childhood maltreatment, vulnerability characteristics and adult incident common mental disorders: 3-year longitudinal data among >10,000 adults in the general population. *Journal of Psychiatric Research*, 113, 199–207. <https://doi.org/10.1016/j.jpsychires.2019.03.029>
- Ten Have, M., de Graaf, R., van Dorsselaer, S., Tuithof, M., Kleinjan, M., & Penninx, B. W. J. H. (2018). Recurrence and chronicity of major depressive disorder and their risk indicators in a population cohort. *Acta Psychiatrica Scandinavica*, 137(6), 503–515. <https://doi.org/10.1111/acps.12874>
- Ten Have, M., Nuyen, J., Beekman, A., & De Graaf, R. (2013b). Common mental disorder severity and its association with treatment contact and treatment intensity for mental health problems. *Psychological Medicine*, 43(10), 2203–2213. <https://doi.org/10.1017/S0033291713000135>
- Ten Have, M., Oldehinkel, A., Vollebergh, W., & Ormel, J. (2003). Does educational background explain inequalities in care service use for mental health problems in the Dutch general population. *Acta Psychiatrica Scandinavica*, 107(3), 178–187. <https://doi.org/10.1034/j.1600-0447.2003.00074.x>
- Ten Have, M., Penninx, B. W. J. H., Tuithof, M., van Dorsselaer, S., Kleinjan, M., Spijker, J., & de Graaf, R. (2017). Duration of major and minor depressive episodes and associated risk indicators in a psychiatric epidemiological cohort study of the general population. *Acta Psychiatrica Scandinavica*, 136(3), 300–312. <https://doi.org/10.1111/acps.12753>
- Ten Have, M., Penninx, B. W. J. H., van Dorsselaer, S., Tuithof, M., Kleinjan, M., & de Graaf, R. (2016). Insomnia among current and remitted common mental disorders and the association with role functioning: Results from a general population study. *Sleep Medicine*, 25, 34–41. <https://doi.org/10.1016/j.sleep.2016.07.015>
- Ten Have, M., Tuithof, M., van Dorsselaer, S., Kleinjan, M., Penninx, B., Batelaan, N., & de Graaf, R. (2021). Duration of anxiety disorder and its associated risk indicators: Results of a longitudinal study of the general population. *Depression and Anxiety*, 38(3), 328–336. <https://doi.org/10.1002/da.23103>
- Tuithof, M., ten Have, M., van den Brink, W., Vollebergh, W., & de Graaf, R. (2013). Predicting persistency of DSM-5 alcohol use disorder and examining drinking patterns of recently remitted individuals: A prospective general population study. *Addiction*, 108(12), 2091–2099. <https://doi.org/10.1111/add.12309>
- Tuithof, M., ten Have, M., van Dorsselaer, S., Kleinjan, M., Beekman, A. T. F., & de Graaf, R. (2018). Course of subthreshold depression into a depressive disorder and its risk factors. *Journal of Affective Disorders*, 241, 206–215. <https://doi.org/10.1016/j.jad.2018.08.010>

- Van Nierop, M., Van Os, J., Gunther, N., Myin-Germeys, I., de Graaf, R., ten Have, M., van Dorsselaer, S., Bak, M., & van Winkel, R. (2012). Phenotypically continuous with clinical psychosis, discontinuous in need for care: Evidence for an extended psychosis phenotype. *Schizophrenia Bulletin*, 38(2), 231–238. <https://doi.org/10.1093/schbul/sbr129>
- Von Korff, M., Crane, P. K., Alonso, J., Vilagut, G., Angermeyer, M. C., Bruffaerts, R., De Girolamo, G., Gureje, O., de Graaf, R., Huang, Y., Iwata, N., Karam, E. G., Kovess, V., Lara, C., Levinson, D., Posada-Villa, J., Scott, K. M., & Ormel, J. (2008). Modified WHODAS-II provides valid measure of global disability but filter items increased skewness. *Journal of Clinical Epidemiology*, 61(11), 1132–1143. <https://doi.org/10.1016/j.jclinepi.2007.12.009>
- Ware, J. E., & Sherbourne, C. D. (1992). The RAND-36 short-form health status survey: 1: Conceptual framework and itemselection. *Medical Care*, 30(6), 473–481. <https://doi.org/10.1016/j.psychres.2007.05.013>
- Wennekers, A., Boelhouwer, J., van Campen, C., & Kullberg, J. (2019). De sociale staat van Nederland 2019. *Sociaal en Cultureel Planbureau*. ISBN 978 90 377 0923 0 2
- Williams, J., Leamy, M., Bird, V., Le Boutillier, C., Norton, S., Pesola, F., & Slade, M. (2015). Development and evaluation of the INSPIRE measure of staff support for personal recovery. *Social Psychiatry and Psychiatric Epidemiology*, 50(5), 777–786. <https://doi.org/10.1007/s00127-014-0983-0>
- Wittchen, H. U. (1994). Reliability and validity studies of the WHO-Composite international diagnostic interview (CIDI): A critical review. *Journal of Psychiatric Research*, 28(1), 57–84. [https://doi.org/10.1016/0022-3956\(94\)90036-1](https://doi.org/10.1016/0022-3956(94)90036-1)

How to cite this article: ten Have, M., Tuithof, M., van Dorsselaer, S., Schouten, F., & de Graaf, R. (2023). The Netherlands Mental Health Survey and Incidence Study-3 (NEMESIS-3): Objectives, methods and baseline characteristics of the sample. *International Journal of Methods in Psychiatric Research*, 32(1), e1942. <https://doi.org/10.1002/mpr.1942>