Improving accuracy of major depression age-of-onset reports in the US National Comorbidity Survey

BÄRBEL KNÄUPER, Department of Psychology, Free University of Berlin, Berlin, Germany
CHARLES F. CANNELL, NORBERT SCHWARZ, Institute for Social Research, University of Michigan, USA
MARTHA L. BRUCE, Cornell University, USA
RONALD C. KESSLER, Department of Health Care Policy, Harvard Medical School, USA

ABSTRACT This paper describes a series of questions designed to improve the accuracy of age-of-onset reports in the US National Comorbidity Survey (NCS) and empirically evaluates the impact of these questions on reports about age of onset of major depressive episodes. The logic underlying the series of question is traced to cognitive psychological research on autobiographical memory. Data are presented showing that the new question series yielded more substantively plausible age-of-onset reports than those obtained a decade earlier in the Epidemiologic Catchment Area (ECA) Study. The test–retest consistency of age-of-onset reports was also higher in the NCS than the ECA. Despite these improvements, considerable inconsistency in age-of-onset reports remains in the NCS test-retest data. The paper closes with a discussion of potentially promising future directions to improve retrospective age-of-onset reports in new psychiatric epidemiological surveys.

Key words: age of onset, major depression, accuracy, autobiographical memory, diagnostic interview, epidemiology surveys

All psychiatric diagnostic interviews use retrospective information of one sort or another. They differ, though, in the length of the recall period and in the relative importance of information about the present and the past. Clinical intake interviews, for example, are primarily concerned with current complaints, but they interpret these complaints in the context of retrospective reports about history. Community epidemiological risk factor surveys, in comparison, are usually concerned with lifetime prevalence, whereas needs assessment surveys are usually concerned with prevalence over a time interval used for administrative planning, such as the last six months or year.

In all of these cases, the clinician or researcher needs to develop strategies for eliciting information that is as accurate as possible about the past. Success in this task requires an understanding of how information is stored in memory and of how successful memory retrieval operates. It requires questioning strategies that use these understandings to help respondents maximize their information retrieval capabilities. And it requires an appreciation of the limits of autobiographical memory and a willingness to make use of partial information when complete information cannot be retrieved.

The current report discusses only one aspect of the information-retrieval process: the recollection of age of onset of psychiatric disorders in epidemiological surveys. Age-of-onset information is often collected in clinical interviews as part of the history-taking process and is used along with other retrospective information about course and family history to build up an initial understanding of the nature of the disorder being treated. Exact dating is generally not critical here, although it is useful to know the general life stage in which symptoms first occurred. In cases where there is an interest in establishing the importance of a triggering event, however, accurate dating of age of onset in relation to a presumed precipitant becomes more important. In cases where there is an interest in distinguishing primary and secondary comorbidity, furthermore, accurate comparative age-of-onset information is needed to distinguish between temporally primary and secondary disorders.

Precise age-of-onset information is more important in epidemiological surveys that focus on lifetime disorders. Retrospective age-of-onset reports are used in such surveys to estimate survival curves and synthetic cohort effects (such as Fombonne, 1994). Age-of-onset
information is also used to establish temporal priorities for the purposes of carrying out retrospective case-control risk factor analyses (for example, Kessler, Davis and Kendler, 1997). Unlike the clinical situation, where the therapist has an opportunity to build up an increasingly complex retrospective portrait of the patient's history over many sessions, the epidemiological researcher typically needs to collect retrospective information in the course of only one interview using fully structured interview methods. There are inherent limitations to the accuracy one can expect in situations of this sort. However, there are also opportunities to improve the accuracy of retrospective reports in fully structured epidemiological interviews. These opportunities are based on research about the ways in which memories are stored and successfully retrieved.

**Accuracy of age-of-onset reports in psychiatric epidemiological surveys**

We are aware of only two general population psychiatric epidemiological studies that assessed the accuracy of age-of-onset reports. Bromet, Dunn, Copnell, Dew and Schulberg (1986) studied the test–retest reliability of age-of-onset reports for major depression in a community sample of 391 women interviewed 18 months apart using the semi-structured Schedule for Affective Disorders and Schizophrenia – Lifetime Version. Among the respondents who reported lifetime major depression (53%) and generalized anxiety disorder (67%), Prusoff, Merikangas and Weissman (1988) found similar results, but also documented that consistency of age-of-onset reports decreases as a function of length of time between the test and retest interviews. Wittchen and colleagues also assessed the validity of the DIS age-of-onset reports in comparison to dates established on the basis of record checks and in-depth clinical interviews. Slightly more than 50% of the dates reported in the DIS were confirmed in this way.

**Memory storage and retrieval for age-of-onset**

Cognitive psychologists have done considerable research on the organization of information in memory and the strategies that are most successful in retrieving this information (Jabine, Straf, Tanur and Tourangeau, 1984; Schwarz and Sudman, 1994; Jobe and Mingay, 1991). This research shows that information about when an episodic experience first occurred is one of the least well-remembered aspects of experience (Rubin and Baddeley, 1989; Sudman, Bradburn and Schwarz, 1996).

Recollection of first experiences is often challenging in itself unless there is something distinct about the first occurrence. The situation is further complicated by the fact that adults generally have no direct recollection of experiences that occurred during the first five years of their lives (Wettler and Sweeney, 1986). This means that it is impossible for most people to accurately date their age at onset of experiences that happened in these early years of life. In addition, the longer the time lapse between the present and the experience in question, when the latter occurred after age of five years, the less likely it is that the experience will be recalled. Experiences that are recent, distinctive, salient, and unique are more likely to be remembered than experiences that are remote, typical, and regular (Brewer, 1986; Belli, 1988; Menon, 1994).

Assuming that a first experience can be recalled, recollection of the age when it occurred is even more
Accuracy of depression onset reports

The current approach

Methodological research carried out by Cannell and his associates (Cannell, 1985; Cannell, Oksenberg and Converse, 1977) has demonstrated that accurate retrospective recall can be maximized by modifying standard interviewing procedures. Their approach begins by making it clear to respondents that active and extensive memory search is needed. They then motivate memory search and facilitate the processes involved in accurate memory search. The necessity of doing this in asking about age-of-onset of psychiatric disorders was made clear to us in preliminary pilot interviews carried out in preparation for the US National Comorbidity Survey (NCS) (Kessler, McGonagle, Zhao, Nelson, Hughes, Eshleman, Wittchen and Kendler, 1994). These interviews demonstrated that many respondents failed to appreciate our desire that they should engage in active memory search and provide precise and accurate answers to the age-of-onset questions asked in our diagnostic interview. Based on this evidence, we set out to develop a revised series of questions about age of onset that would improve on the questions used in the DIS and the Composite International Diagnostic Interview (CIDI) (World Health Organization, 1990).

The DIS approach simply asks ‘How old were you the first time you had (SX)?’ The DIS training manual instructs interviewers to ‘probe for a more specific response’ when the respondent answers this question with a ‘don’t know’ response, but gives no guidance whatsoever on how this probing should be done other than to say that a respondent who fails to provide an age after probing should be asked whether or not the disorder began in the past 12 months. This instruction makes it possible for the interviewer to code the structured DIS age-of-onset question that is divided into the categories of more than a year ago, past year, past six months, past month, and past two weeks. As a practical matter, though, it does not lead to improved reporting accuracy because virtually none of the respondents in our experience who respond ‘don’t know’ to the initial question ended up reporting that his or her age of onset was in the past year. The CIDI approach is only slightly different from the DIS...
approach. Rather than asking about 'how old' the respondent was at the time of onset, the CIDI asks 'When was the first time you had (SX)?' The CIDI training manual instructs interviewers to probe 'don't know' responses with 'Could you give me your best guess?' and, if this is not productive, to ask whether or not the disorder began in the past 12 months.

Our pilot work for the NCS showed that these approaches do not provide clear instructions to respondents about the intent of the implied task (to carry out a serious and thorough memory search). Indeed, when we used standard cognitive probes to ask pilot test respondents how they interpreted the standard CIDI age-of-onset question, it became clear that most respondents interpreted the question as asking whether they could remember a first episode right now rather than whether they would engage in active memory search in an attempt to remember the first episode. This kind of misinterpretation of the intent of the question is bound to generate superficial responses.

In order to correct these problems, we developed a revised series of age-of-onset questions for the NCS. The sequence began with what is referred to in the survey methodology literature as a 'prequest', a question aimed at clarifying the nature of the information to be sought in subsequent questions. The prequest question was 'Can you remember your exact age the first time you had (SX)?' (emphasis in original). Our intent here was to make it clear to respondents that we were interested in a serious memory search and to convey to the respondent that accurate reporting in response to the next question was very important to us. In addition, the prequest allowed us to sort respondents into two groups who could versus could not recall their earliest episodes. This sorting was important for the task of facilitating memory recall, our next goal, because it allowed us to tailor subsequent questions depending on whether the first onset was clearly remembered or not.

In pilot work, we probed respondents who reported that they could clearly remember their age of onset as to how they were able to remember so clearly. Their responses typically involved a combination of a critical trigger event (for example, parental divorce leading to depression) or incident (for example, a first panic attack backstage before a fifth-grade class musical production) and a clear recollection of a life situation linked to an age (for example, being in Mrs Smith's fifth-grade class). Based on this result, the production phase of the NCS simply asked respondents who reported recollection of their exact age-of-onset to tell us that age: 'How old were you?'

For respondents who reported that they could not recall their exact age, in comparison, we asked a different follow-up question phrased in such a way as to make it clear that we wanted an estimate: 'about how old were you the first time you had (SX)?' (emphasis in original). This question served two purposes. First, it gave us an opportunity to acknowledge our acceptance of the respondent's inability to provide an exact response. Second, it provided us with approximate information that could be useful both to us in narrowing the range of uncertainty and to the respondent in focusing memory search in response to the next question. In keeping with these purposes, interviewers were instructed to accept a range response without probing.

This question was then followed by another question designed to provide an upper bound on our uncertainty concerning age-of-onset and to permit respondents to answer with as much accuracy as possible even though they could not recall their exact age of first onset. This question asked 'What is the earliest age you can clearly remember having (SX)?' (emphasis in original). This question is much less demanding than the original question about exact age of first onset. Yet it encourages active memory search, it focuses memory search on the life stage reported in the previous question, and the answer provides very useful, albeit partial, information that allows us to bound our uncertainty. It is much more useful to recognize that exact information is unattainable and to capture useful partial information in this way than to ask respondents to tell us more than they know.

Methods and procedures

The remainder of the report presents data about responses to the age-of-onset questions in the NCS applied to the diagnosis of major depressive episodes. The first set of results concerns the distribution of NCS age-of-onset responses in comparison to the distribution found by Simon and Von Korff (1992) for the standard major depression DIS age-of-onset question in the ECA data. The second set of results deal with test–retest reliability data of age-of-onset reports in the NCS clinical reappraisal study of depression in comparison to the two-wave test–retest consistency data in the ECA.

As described in more detail elsewhere (Kessler et al., 1994), the NCS was a face-to-face household survey of
Accuracy of depression onset reports

**Figure 1:** ECA – reported years since onset by age at interview.

**Figure 2:** NCS – reported years since onset by age at interview.
were reanalysed with somewhat finer age divisions than in the original Simon and Von Korff analysis for this purpose. Revised ECA data using the same more refined age breaks were kindly provided by Professor Gregory Simon. The report of the ECA analysis is confined to respondents 54 years or younger to make it comparable to the NCS study.

Results are reported in Figures 1 (ECA) and 2 (NCS), where we show the proportions of respondents in each survey with lifetime major depression who reported their first episodes as occurring in the age ranges 0–10, 11–20, 21–30, 31–40, and 41+ years prior to the interview as a function of age at interview. The results in Figure 1 show, as reported by Simon and Von Korff (1992), that onset reports in the ECA data cluster in the 10 years prior to the interview for all age groups, including the oldest. The results in Figure 2, in comparison, show that the peak onset ages reported are in early adulthood (roughly 20–30 years of age) in all the NCS cohorts. Only 17.7% of NCS respondents in the oldest sample cohort reported that their first episode occurred in the 10 years prior to the interview compared to 38.3% in the equivalent age group of the ECA. This comparison strongly suggests that the modified approach to asking about age of onset in the NCS yielded more substantively plausible respondent reports, with peak ages of onset in early adulthood and absence of reports that cluster in the 10 years prior to the interview. The most plausible interpretation of this finding is that the approach used in the NCS led respondents to conduct a more thorough memory search.

Consistency of age-of-onset reports in the NCS and ECA

The test–retest reliability of major depression age-of-onset reports in the NCS and ECA are presented in Table 1. Part I of the table shows absolute value differences in years between test and retest age-of-onset reports for the 20 NCS respondents who were included in the evaluation of the test–retest reliability of major depression. The results in the first column show that 65% of these NCS respondents reported ages of onset within one year of each other in the two interviews and that 80% reported ages of onset within five years of each other. The results in Part II of the table show that reporting consistency is highest for respondents who reported first onsets in childhood or late adulthood and lowest for those who reported first onsets in adolescence. This pattern of variation is substantively plausible, as childhood onsets are likely to be defined in terms of clear upper bound ages and late
adult onsets are likely to be easiest to recall because they are comparatively recent in a sample with a restricted upper age range. Onsets in adolescence, early adulthood, or middle adulthood, could be more difficult to recall accurately.

As shown in the second column of Table 1, aggregate results are markedly different in the subsample of 349 ECA respondents who reported a history of depression in the baseline ECA interview and were re-interviewed one year later. The most critical of these differences is that, whereas 100% of the NCS respondents who met criteria for depression in the baseline survey acknowledged being depressed and reported an age-of-onset in the re-interview, this was true of only 58.7% of baseline ECA respondents. As shown in the second column of the table, the agreement percentages for age-of-onset reports are substantially lower in the ECA than the NCS. This occurs both because of the great inconsistency between the NCS and the ECA in terms of admitting depression and because of lower consistency of age-of-onset reports than in the NCS in the subsample of ECA respondents who consistently reported being depressed. Only 31% of baseline ECA depressives had age-of-onset reports within one year of each other and 45.6% within five years of each other. As shown in Part II of the table, the distribution of reporting inconsistency across life stages in the ECA is dramatically different from that in the NCS.

Table 1: Consistency in major depression age of onset reports over time in the NCS and ECA

<table>
<thead>
<tr>
<th>I. Differences in age of onset reports between Time 1 and Time 2†</th>
<th>NCS</th>
<th>ECA</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-1 years</td>
<td>65.0%</td>
<td>31.0%</td>
</tr>
<tr>
<td>2-5 years</td>
<td>15.0</td>
<td>14.6</td>
</tr>
<tr>
<td>6-10 years</td>
<td>15.0</td>
<td>7.5</td>
</tr>
<tr>
<td>11 or more</td>
<td>5.0</td>
<td>5.7</td>
</tr>
<tr>
<td>No age of onset information</td>
<td>0.0</td>
<td>41.3</td>
</tr>
<tr>
<td>(n)</td>
<td>(20)</td>
<td>(349)</td>
</tr>
<tr>
<td>NCS v. ECA difference</td>
<td>χ²(4) = 15.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>p &lt; 0.01</td>
<td></td>
</tr>
</tbody>
</table>

II. Per cent consistency in life stage of onset reports between Time 1 and Time 2 |

<table>
<thead>
<tr>
<th>Life stage</th>
<th>NCS</th>
<th>ECA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Childhood (0-9 yrs)</td>
<td>100.0%</td>
<td>16.7%</td>
</tr>
<tr>
<td>Adolescence (10-19 yrs)</td>
<td>66.7</td>
<td>39.8</td>
</tr>
<tr>
<td>Early adulthood (20-29 yrs)</td>
<td>83.3</td>
<td>41.5</td>
</tr>
<tr>
<td>Middle adulthood (30-39 yrs)</td>
<td>80.0</td>
<td>45.2</td>
</tr>
<tr>
<td>Late adulthood (40-54 yrs)</td>
<td>100.0</td>
<td>27.3</td>
</tr>
<tr>
<td>Total</td>
<td>85.0</td>
<td>40.4</td>
</tr>
<tr>
<td>(n consistent / n total)</td>
<td>(17/20)</td>
<td>(141/349)</td>
</tr>
<tr>
<td>NCS v. ECA difference</td>
<td>χ²(1) = 15.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>p &lt; 0.01</td>
<td></td>
</tr>
</tbody>
</table>

† Meaningful comparison of the consistency of age of onset reports in the NCS and ECA required us to apply the same coding rules to inexact responses in both surveys. The following rules were used: (1) For ranges of two years (e.g. 18-19) the lower age was chosen; (2) for ranges of several years the mean age in the range was chosen; (3) ‘teenager’ was coded as 16 years, ‘early teens’ as 13 years, ‘late teens’ as 18 years, ‘childhood’ as 6 years; (4) ‘as long as I can remember’ was coded as an onset of 2 years.
Consistency was highest in the NCS for respondents who reported in the baseline interview that their age of onset was either in childhood or late adulthood (100% consistency), but these were the life stages with the lowest reporting accuracy in the ECA (16.7% and 27.3%, respectively). The extremely low consistency of reports about childhood onsets, which was not only 16.7% in the total ECA sample but only 25% in the subsample of ECA respondents who admitted depression in both interviews, is especially telling because it means that ECA respondents had a tendency to give up in trying to recover a memory of an early age of onset and simply reported a more recent and more easily recalled episode. The NCS respondents, in comparison, were dramatically more consistent in reporting onsets early in life because they were asked explicitly about whether they had a clear memory of their first episode and, if not, were asked to tell us about the rough age of onset.

Discussion

Absence of a gold standard to measure the true age of onset of depression makes it impossible to document exactly how much the accuracy of age-of-onset reports improved due to the new questions asked in the NCS. Three important indirect indicators nonetheless suggest that these questions did, in fact, lead to improvements rather than the reverse. The first is that the distribution of reported onset ages in the NCS is substantively more plausible than the distribution found a decade earlier in the ECA study. The second is that the test–retest consistency of age-of-onset reports is higher in the NCS than the ECA. The third is that the pattern of variation in test–retest consistency as a function of reported life stage is more substantively plausible in the NCS than the ECA.

It should be noted that we were able to achieve this improvement by making fairly modest changes to the age-of-onset questions asked in the CIDI. These changes aimed at improving the understanding of the question and at simplifying the recall task for those respondents who had difficulty remembering their exact age of onset. The results suggest that this strategy led respondents to conduct a more thorough memory search and successfully encouraged respondents to particularly recall episodes that occurred a long time ago. However, the improvement was incomplete, with only 65% of respondents in the depression test–retest study being able to report the same age-of-onset in a second interview. It is important to note that this lack of complete consistency was not due exclusively to respondents who reported in the baseline NCS that they could not remember their exact age of onset. A full 85% of respondents who met criteria for major depression in the NCS told us that they could remember their exact age of onset, even though only 65% reported the same age in the retest interview. This means that future improvements in recall accuracy will require us to go beyond simply distinguishing respondents who do and do not report vivid recall of their first episodes.

The literature on memory processes suggests a number of potentially useful strategies that could be used in future research to increase the accuracy of age-of-onset reports. It is noteworthy that consistency of reports increased substantially when the data were analysed at the level of life stages. One way to use these findings is to begin with the exact age prequest and to follow this with a question about life stage. Once life stage of onset is established, a series of contextual questions (for example, ‘Were you already in school when you had these problems for the first time?’) could be used to narrow the range of uncertainty in dating. A related possibility is to have respondents create their own cues by asking them to enter personal facts and events on a personal time line ranging from their birth to the time of the interview. These entries could then serve as visual aids and reference points for recalling the age of onset of the disorder. A number of cognitive psychologists have shown that the use of such references points improve the accuracy of recall of dates (Loftus and Marburger, 1983; Baddeley, 1990; Means and Loftus, 1991).

Some related attempts to obtain more comprehensive assessments of time-related psychiatric information already exist in dating life events in relation to episodes of major depression (Brown and Harris, 1978; Kessler and Wethington, 1991). A very promising approach recently developed along these lines is the Life Chart Interview (LCI) of Lyketsos, Nestadt, Cwi, Heithoff and Eaton (1994). Building on earlier work by Thornton, Freedman and Camburn (1982), the LCI uses personal landmark events and an interactive dating procedure to record information about the time sequence of psychopathology and life events on a life chart calendar. It is conceivable that future research, by combining a modification of the disorder-specific age-of-onset approach used in the NCS with the lifetime review of events used in the LCI, could succeed in substantially narrowing the range of uncertainty in age-of-onset reports of psychiatric disorders.
Accuracy of depression onset reports

Acknowledgements
The National Comorbidity Survey (NCS) is a collaborative epidemiological investigation of the prevalences, causes, and consequences of psychiatric morbidity and comorbidity in the United States supported by the National Institute of Mental Health (R01 MH46376, R01 MH49098, and R01 MH52861) with supplemental support from the National Institute of Drug Abuse (through a supplement to MH46376) and the W.T. Grant Foundation (90135190). Ronald C. Kessler, Principal Investigator. Preparation for this report was also supported by a Research Scientist Award to Kessler (K05 MH00507). Collaborating NCS sites and investigators are: The Addiction Research Foundation (Robin Room), Duke University Medical Center (Dan Blazer, Marvin Swartz), Harvard Medical School (Richard Frank, Ronald Kessler), Johns Hopkins University (James Anthony, William Eaton, Philip Leaf), the Max Planck Institute of Psychiatry (Hans-Ulrich Wittchen), the Medical College of Virginia (Kenneth Kendler), the University of Miami (R. Jay Turner), the University of Michigan (Lloyd Johnston, Roderick Little), New York University (Patrick Shrout), SUNY Stony Brook (Evelyn Bromet), and Washington University School of Medicine (Linda Cottler, Andrew Heath). A complete list of all NCS publications along with abstracts, study documentation, interview schedules, and the raw NCS public use data files can be obtained directly from the NCS Homepage by using the following URL: http://www.hcp.med.harvard.edu/ncs. We would like to thank Jamie Abelson, Bob Belli, Chris Nelson, Ulrich Schimack, and Hans-Ulrich Wittchen for helpful comments on an earlier draft of this paper and Greg Simon for providing the ECA data used in Table 1.

References
Loftus EF, Marburger W. Since the eruption of Mt St Helens, has anyone beaten you up? Memory and Cognition 1983; 11: 114–20.


Correspondence to Bärbel Knäuper at the Free University of Berlin, Psychologie (WE 10), Habelschwerdter Allee 45, D-14195 Berlin, Germany. Email: bknauper@zedat.fu-berlin.de. Alternatively address correspondence to Ronald C. Kessler, Harvard Medical School, Department of Health Care Policy, 180 Longwood Avenue, Boston, MA 02115-5899, USA. Email: kessler@hcp.med.harvard.edu