Studying Chinese Suicide with Proxy-Based Data: Reliability and Validity of the Methodology and Instruments in China

Jie Zhang, PH.D.[1], **Yeates Conwell, M.D.**[2], **William F. Wieczorek, PH.D.**[3], **Chao Jiang, M.D.**[4], **Shuhua Jia, M.D.**[4], and **Li Zhou, M.D.**[4]

[1]Department of Sociology, State University of New York College at Buffalo, 1300 Elmwood Avenue, Buffalo, New York 14222.

[2]Department of Psychiatry, University of Rochester, Rochester, New York.

[3]Center for Health and Social Research, State University of New York College at Buffalo, Buffalo, New York.

[4]Dalian Medical University, Dalian, Liaoning, China.

Abstract

This study examined the reliability and validity of the instruments as used in the psychological autopsy method in China. With data from 130 informants on 66 completed suicides and 130 informants on 66 normal community controls and 66 controls themselves, the validity was examined by comparing the responses of informants and the responses (gold standards) of the target participants in the control group. All the tested instruments were shown to be reliable, and proxy respondents were generally good judges of targets' suicidal intention, social support, depression, life events, personality traits, and mental disorders. Additionally, interrater reliabilities of the five interviewers were very good on selected scales. This study has laid a partial foundation for future psychological autopsy projects to be held in Chinese culture.

With a wide range of suicide rates reported for China as a result of different methods of sampling and adjustment, the mean annual suicide rate for China is 23 per 100,000 (Phillips et al., 2002). Suicide has been an alarming social and health problem in China, claiming approximately 300,000 lives each year (Brown, 1997; Murray and Lopez, 1996; Phillips et al., 2002; World Health Organization, 1988–1995), with more women than men, higher rates in the rural areas than in the urban, and greater risks for young population age 15 to 24 years than the younger and the middle-aged (Macleod, 1998; Phillips et al., 1999; Pritchard, 1996; Qin and Mortensen, 2001; World Health Organization, 1988–1995; Yip, 2001; Zhang, 1996; Zhang, 2000b). Researchers both inside and outside of China are trying to identify the factors that account for the high suicide rates among Chinese rural young women (Ji et al., 2001; Lester, 1994; Yip et al., 2000; Zhang, 2000a; Zhang et al., 2002a). To understand the causes and social environmental factors of completed suicide, psychological autopsy (PA), using data from proxy respondents and medical records or other sources, may be the best scientific method available to researchers (Clark and Horton-Deutsch, 1992).

Psychological autopsy, "a procedure for the reconstruction of suicidal death through interviews with survivors" (Beskow et al., 1990), has been used in the West for suicide studies in the past 4 decades. However, it is a relatively new concept in Chinese studies of suicide. Cheng

(1995) and Cheng et al. (2000) reported applications of the method with ethnic minorities in Taiwan, and the work of M. Phillips (personal communication, June 2001) is in progress. Other than these studies, there is a lack of published evidence that the technique can be applied in Chinese culture in mainland China, given the traditional cultural background that is so different from the Western culture in which the methodology has been developed.

Based on the literature on PA published in the West, we have conducted a study using the methodology in China. Methodological guidelines for determining postmortem psychiatric diagnoses and other risk factors have been offered (Brent et al., 1988b; Clark et al., 1992; Ebert, 1987; Hawton et al., 1998) and initiated in China but require empirical validation. A general report on applying PA method can be found in the study by Zhang et al. (2002b), which examines the quality of the instruments used in Chinese culture.

Research in psychopathology has generally supported use of the best estimate method in making diagnostic judgments (Leckman et al., 1982; Maziade et al., 1992; Nunes et al., 1998; Reich and Earls, 1987; Reich et al., 1993; Weissman et al., 1987), although validation research in studies of suicidal behavior is limited (Kelly and Mann, 1997; Velting et al., 1998). Conner et al. (2001a) filled the gap by comparing the psychiatric diagnoses from elderly suicide attempters directly and from their proxy respondents. Their results support the best estimate methodology using proxy data for making psychiatric diagnoses in research of suicidal behavior in that age group.

Although stressful life events have been studied in several investigations (Asgard, 1990; Boardman et al., 1999; Brent et al., 1993; Brugha and Cragg, 1990; Gould et al., 1996; Heikkinen et al., 1992; Vijayakumar and Rajkumar, 1999), there have been only two reports on the validity of proxy-based life events data in adults. In a study of medically serious suicide attempters younger than 25 years, Beautrais et al. (1997) investigated the concordance between subjects and proxy respondents on stressful life events preceding suicidal behavior. Although there was substantial variability of agreement based on correlation coefficients between self-report and proxy report of life events that precipitated the suicide attempts, all correlations were statistically significant (Beautrais et al., 1997). A more recent study by Conner et al. (2001b) with 80 psychiatric inpatients admitted after a suicide attempt indicated that the proxy respondents were able to offer valid judgments on suicide attempters' stressful live events.

The effect of social support on suicide has been examined in PA investigation (Brent et al., 1994; Heikkinen et al., 1994; Murphy et al., 1992) and supported by investigations showing an association of low social support with suicidal ideation (Mireault and de Man, 1996), depression (Paykel, 1994), psychological distress (Finch et al., 1999), and other negative health outcomes (Lepine and Bouchez, 1998; Mireault and de Man, 1996; Rook, 1994; Schwarzer and Leppin, 1989). Although research on the comparability of self and proxy-based measures of social support is limited (Bages et al., 1997; Berk, 1995; Emmons and Colby, 1995), two studies using older adult samples have shown that proxy–respondent agreement on measures of social support was significantly correlated (Antonucci and Israel, 1986; Epstein et al., 1989).

Suicide intention has been demonstrated to be correlated with subsequent completed suicide acts (Beck and Steer, 1989; Beck et al., 1989; Brent et al., 1988a). However, the validity of proxy respondent reports of remote history of suicidal behavior or intent to die associated with suicidal acts has not been adequately reported. Conner et al. (2001b) reported that proxies were good judges of past history of suicide attempts and level of suicidal intent.

This current study further examines these measurements for psychopathological and social factors used in psychological autopsies, but with Chinese samples. Validation of the scales in

Chinese culture will allow researchers to use the method and instruments in Chinese proxy-based data collection.

Materials and Methods

Data used in the current study are part of the database resulting from a PA study conducted in Dalian, a northeastern city in China. The Dalian area with its surrounding counties has a total population of more than five million, approximately two thirds of whom are rural residents. The study was approved by the Institutional Review Board at both the Chinese and the US institutions. The health administrations of Jinzhou County and two townships in Zhuanghe County were approached to collect the names of suicides reported in the past 6 months. From February to August 2001, this consecutive sampling yielded a total number of 66 suicides in Jinzhou County and the two townships in Zhuanghe County.

Sixty-six age-, gender-, and location-matched living people were selected as the community control group. For each suicide case, there were two informants who were either next of kin or the best friend of the suicide. For each normal control, two informants and the control served as the sources of information. The interviews with the living normal controls provided a basis for validity tests. The total number of interviews was 330.

For each target person, suicide or control, there were two interviews with informants. Each interview was entered into a database, and then the two interviews for the same target were combined into a new file with programmed computations. If the responses to one item were the same for the two informants' files, the agreed value was written to the new file. If there was a disagreement between the two interviews, the programmed computation usually took the value that indicated a symptom or more extreme response. The interview with the target person in the community normal control group was kept separately in the database for other analyses.

The protocol was a structured questionnaire designed for interviews and had the following major parts:

Beck's Suicide Intent Scale (SIS; Beck et al., 1974)

Paykel's Interview for Recent Life Events (IRLE; Paykel et al., 1971)

Duke Social Support Index (DSSI; Landerman et al., 1989)

Structured Clinical Interview for DSM-III-R (SCID; Spitzer et al., 1988)

Hamilton Depression Rating Scale (HAM-D; Hamilton, 1960; Williams, 1988)

Center for Epidemiological Studies—Depression (CES-D; Radloff, 1977)

General Social Survey Attitudes Towards Suicide (GSS; Davis and Smith, 1993)

Personality Measurement (NEO-FF-I; McCrae and Costa, 1997)

Some of the instruments had been translated into Chinese and/or validated with Chinese samples, such as the SCID (Gu and Chen, 1993), HAM-D (Zheng et al., 1988), CES-D (Lin, 1989), and NEO-FF-I (Yang et al., 1999), but none of them was tested in proxy data collection. Translation and back translation of the rest of the instruments were conducted multiple times to ensure the accuracy of each item in the interview. Recognizing that translations cannot always be perfect because of the differences in culture (Zhang and Thomas, 1994), the bilinguals on the research team met frequently for the disputed items in the translation until consensus was reached. Interviewers were trained professionals who were either professors or graduate students at Dalian Medical University. Before the data collection began, each of them received further training in interviewing people with the SCID protocol.

To study the validity of the measurements, we used the data from the control group. The community normal living controls provided the gold standard, and their informants' responses to the questions about them were proxies which are at the examinations for validity. Table 1 illustrates demographic descriptions of the targets (community normal living controls) and their informants.

One other way to estimate the validity of proxy data is to compare the responses from both suicide attempters and their informants (Conner et al., 2001a; Conner et al., 2001b; Velting et al., 1998). However, the method has at least two limitations. First, attempted and completed suicide have different prevalence, social demographic characteristics, and psychological correlates (Maris et al., 1992; McIntosh et al., 1992; Tiller et al., 1998), although attempted suicide is a potent risk factor for completed suicide (Goldstein et al., 1991). Therefore, the study of attempted suicide is not an ideal substitute for research on completed suicide. Second, proxy respondents of suicide attempters are more likely to be grief-stricken than others, and although some data are available (Brent et al., 1988b), little is known about the influence of affective symptoms on proxy respondent reports.

Results

A very important indication of the quality of interview data collection is the interrater reliability. Although all the interviewers were well-trained professionals and had been further trained in using SCID, there is still a need to examine how consistent they were in rating a same diagnosis or response. Data from five of the interviewers were analyzed as shown in Table 2. The total score was used in each of the scales, and a correlation matrix of the five raters was run for each of the scales. A total score was calculated by summing up the scores of all the items for a continuous scale for each case, or adding up the number of occurrences for a dichotomous variable such as SCID and Life Events. The five raters were most consistent on HAM-D, CES-D, and GSS, because all correlations were significant at the .01 level. Next were SCID total score and suicidal intention, whose correlations were significant at the .05 level. For the Life Events total score, only two correlations were not significant. The interrater reliabilities for the NEO total score and for the scales of Social Support were not as good as others.

Validity of the PA methodology and the instruments used in the data collection is at the core of the current study. We performed different tests on dichotomous and continuous data. First, we categorized all the diagnosed mental disorders into six types and grouped the 63 life events into six categories. The dichotomous data were then recoded as 0 if the diagnosis or life event did not exist and 1 if it did exist. For each category in either mental disorder or life events, proxy-based diagnoses were compared with target person-based diagnoses. The latter served as the gold standard.

With the formulas provided in Gordis (1996), four statistics were generated to assess the validity of responses by proxies. Sensitivity was calculated to assess the ability of proxy respondents to detect the presence of a psychiatric disorder or life event. Specificity was used to examine the ability of proxy respondents to identify the absence of a psychiatric disorder or life event correctly. Percent agreement referred to the proportion of cases in which proxy-based decisions (either positive or negative) coincided with target person-based responses. Kappa quantified the level of agreement between target person-based and proxy-based responses after accounting for the level of agreement expected by chance.

Results of analyses on the SCID and the Life Events scale are presented in Table 3 and Table 4, respectively. There were five categories for mental disorder as opposed to six originally grouped, because no participants in the control group were diagnosed for nonaffective

psychoses (schizophrenia). Similarly, there were five categories of life events instead of six as originally grouped, because no participants in the control group were scored as positive in any of the life event items categorized as law/legal.

For the validity test on mental disease diagnoses, sensitivity was perfect (1.00) for four of the five categories, which indicated total coincidence between the proxy responses and the information provided by the targets. Only anxiety had a sensitivity score of .50. Only two of 66 target persons identified themselves as positive on anxiety, but the proxy respondents endorsed only one. Accordingly, the percent agreement and Kappa for anxiety were affected. On the other hand, all the five categories had a perfect specificity score, indicating that the proxies' ability to identify those who did not have a mental disease went without errors in the sample. It should be noted that the base rates of these diagnoses were so low that most agreements were from the no disagreement categories, and that result contributed to the perfect sensitivity, specificity, Kappa, and percent agreement.

More variance was shown in the data for life events than for mental disorders. Sensitivity was lowest for the relationship life event but greater than .50, and highest for health/hospital. Accordingly, specificity and percent agreement were high across the life events categories. The most important statistics in the assessment of agreement may be the coefficient Kappa because it is a rigorous standard of diagnostic accuracy (Cohen, 1960). The Kappa values were evaluated based on the following benchmarks originally provided by Landis and Koch (1977) and reprinted in Dunn (1989): poor, less than .01; slight, .01 to .20; fair, .21 to .40; moderate, .41 to .60; and substantial, greater than .80. Because the lowest Kappa was .30 for the relationship life event, all the Kappa values were at least fair, indicating adequate agreements between the proxies and targets in their responses to the life events items.

Another way to examine the agreement between proxies and target persons themselves is to run correlations, especially on continuous scales. The results are shown in Table 5. The SCID and life events were further tested with the correlation statistics. Almost all of the agreements were highly correlated at the .01 level, and suicide intention had a correlation significant at the .05 level.

Further tests on the agreements between proxies and target selves were performed on selected continuous scales by comparing their means. A low *t*-value and high probability level indicate the closeness of the two means, which is hoped for to support a non-difference hypothesis. Table 6 shows that a significant difference occurred only between the first informants and targets on instrumental social support. Self-report on instrumental social support was significantly higher than those scores reported by the first proxy respondents. There was a tendency for targets to score higher than their proxies on the scales of SCID, HAM-D, and NEO-FF-I. In other words, proxies could have underestimated the diagnoses or traits, but the differences were not significant.

Discussion

Quality control through standardized procedures is the essential foundation of scientific studies. Although all the Western-developed instruments used in the protocol have been validated and shown to have good reliability and validity, the Chinese versions have not been systematically tested for reliability and validity with Chinese samples, except for the CES-D (Lin, 1989; Zhang and Norvilitis, 2002), HAM-D (Zheng et al., 1988), and NEO-FF-I (Yang et al., 1999). However, those scales were never tested in proxy data collection. Quality control is even more critical in PA studies because of the indirect characteristics of the information to be collected and the different types of relationships between the interviewees and the targets.

The quality tests should be exercised not only on the instruments translated into Chinese but also on the procedure of data collection.

In a PA study, reliability tests should answer two major questions: (1) is the recording or judgment of the information precise, and (2) is the same information obtained with consistency? Whereas the second question is about the instrument and informant, the first is about the interviewer and the research team. To examine the first question, we used interrater reliability tests with two interviewers recording the information from the same interviewee. The interrater reliability of the five interviewers varied across the scales (Table 2), with the best scores for HAM-D, CES-D, and GSS. The worst scores were for Perceived Social Support and Instrumental Social Support, which clearly indicated discrepancies between self-knowledge and other perception. Other worse scores were for the NEO total and Life Events total. The length of each of the two scales (60 items for the NEO and 63 for the Life Events) may have contributed to the low consistency scores. Another possible factor lowering the interrater reliability of the NEO total could have been the way the NEO questions were asked. This part of the interview may be less structured, allowing the interviewers more flexibility in recording a response on the five-category scale from "strongly disagree" to "strongly agree." Generally, the interrater reliabilities for the five interviewers were good and acceptable.

To study the second question on reliability, we could have exercised the test and retest technique on some scales with the same informants, but we performed split-half and internal consistency tests instead. Based on both tests, the lowest reliability scores were for the two informants of suicide on SIS. Other low scores were for the first informant of the normal control group on the Instrumental Social Support and the two informants of the suicide group on the NEO—openness. Although the reasons for the low reliability for certain scales and samples are to be investigated, all other scales were either excellent or acceptable on both split-half tests and internal consistency tests for the different samples. They are all adequate measurements in terms of reliability for use in Chinese suicide research.

Is information derived without input from the targets themselves valid for identifying psychiatric diagnoses and social factors? The accuracy of the information provided by various informants is related to our major concerns on the validity of data. To examine the validity of the suicide data collected through proxy respondents, information may be obtained from medical records of the deceased and compared with the interview data, given that such information is kept secret to the interviewer (Kelly and Mann, 1997). However, this method is almost useless in China, because the majority of the suicides did not have a hospital record and had never visited a psychiatrist. Thanks to the availability of the living control group in the study, we are able to compare the information from the informants (proxy respondents) with the answers provided by the targets themselves.

All the diagnostic categories except anxiety were perfect on all the four statistics: sensitivity, specificity, percent agreement, and Kappa (Table 3). In the control group of 66 targets, only two were identified positive on anxiety, and only one proxy respondent did not agree. Small sample and less variance (low base rates of the diagnoses) accounted for the extreme values on the statistics. Generally, the informants of the control group were able to offer valid estimates on the targets' mental disease diagnoses.

For proxy respondents' estimates on targets' life events, Kappa was fair only for the relationship life event (Table 4). All categories of life events were acceptable based on the Kappa, sensitivity, specificity, and percent agreement. The data support the validity of PA methods in estimating targets' life events by proxy respondents.

The significant correlation coefficients between the informants and targets themselves indicate comparatively accurate estimates of the proxy respondents on the reality of the selected scales

(Table 5). Another alternative to examine the validity of selected scales in the protocol is to compare the means of the answers from the proxy respondents and the targets (Table 6). A nonsignificant *p*-value (> .05) indicates that the proxy respondents were basically to the point, in line with the information provided by the targets themselves. The significant difference was found only in the instrumental social support scale between the first informant and the target. For personal assistance in daily life, the targets must have observed more and scored the items higher than an informant did. Other than this, the *t*-tests suggest that all selected scales could be estimated adequately by the informants.

In summary, in the PA studies in China, proxy respondents are generally good judges of targets' suicidal intention, social support, depression, life events, personality traits, and mental disorders. These instruments, with adequate reliability and validity scores, may be used in studying completed suicides when targets are not available for comparisons. The group of interviewers, with adequate interrater reliabilities, is a solid combination for further projects.

Finally, it is important to notice several limitations of the study and the PA method in the Chinese setting. First, this study is limited by its sample size. With a total of 66 normal controls, there is virtually not enough variance in mental disorder measures for the control group. Second, a good and accurate translation of the scales into Chinese is important, but more important is their performance in Chinese environments. Further research is needed on the cultural applicability of these scales to Chinese suicides. Third, using normal living controls to assess the validity of proxy responses to completed suicide creates another problem: lack of grief in the controls. Because the informants of completed suicides are most likely to be in grief, it is asymmetrical to generalize the responses from the nongrieving normal people on living controls. Therefore, future research may need to find ways to adjust this imbalance. Further, future study of this type may focus on certain age groups to generalize some more meaningful details for a specific age group.

Acknowledgments

References

Antonucci TC, Israel BA. Veridicality of social support: A comparison of principal and network members' responses. J Consult Clin Psychol 1986;54:432–437. [PubMed: 3745595]

Asgard U. A psychiatric study of suicide among urban Swedish women. Acta Pscyhiatr Scand 1990;82:115–124.

Bages N, Warwick-Evans L, Falger PRJ. Differences between informants about type A, anger, and social support and the relationship with blood pressure. Psychol Health 1997;12:453–465.

Beautrais AL, Joyce PR, Mulder RT. Precipitating factors and life events in serious suicide attempts among youths aged 13 through 24 years. J Acad Child Adolesc Psychiatry 1997;36:1543–1551.

Beck, AT.; Schuyler, D.; Herman, I. Development of suicide intent scales. In: Beck, AT.; Resnick, HCP.; Lettieri, D., editors. The prediction of suicide. Bowie, MD: Charles Press; 1974. p. 45-46.

Beck AT, Steer RA. Clinical predictors of eventual suicide: A 5-to 10-year prospective study of suicide attempters. J Affect Disord 1989;17:203–209. [PubMed: 2529288]

Beck AT, Steer RA, Trexler LD. Alcohol abuse and eventual suicide: A 5- to 10-year prospective study of alcohol-abusing suicide attempters. J Stud Alcohol 1989;50:202–209. [PubMed: 2724967]

Berk RA. Use of primary caretakers as proxies to measure the health care needs of patients with AIDS. Public Health Nurs 1995;12:109–116. [PubMed: 7739981]

Beskow J, Runeson B, Asgard U. Psychological autopsies: Methods and ethics. Suicide Life Threat Behav 1990;20:307–323. [PubMed: 2087767]

Boardman AP, Grimbaldeston AH, Handley C, Jones PW, Willmott S. The North Staffordshire suicide study: A case-control study of suicide in one health district. Psychol Med 1999;29:27–33. [PubMed: 10077291]

Brent DA, Perper JA, Goldstein CE, Kolko DJ, Allan MJ, Allman CJ, Zelenak JP. Risk factors in adolescent suicide: A comparison of adolescent suicide victims with suicidal inpatients. Arch Gen Psychiatry 1988a;45:581–588. [PubMed: 3377645]

Brent DA, Perper JA, Kolko DJ, Zelenak JP. The psychological autopsy: Methodological considerations for the study of adolescent suicide. J Am Acad Child Adolesc Psychiatry 1988b;27:362–366. [PubMed: 3379020]

Brent DA, Perper JA, Moritz G, Baugher M, Roth C, Balach L, Schweers J. Stressful life events, psychopathology, and adolescent suicide: A case control study. Suicide Life Threat Behav 1993;23:179–187. [PubMed: 8249030]

Brent DA, Perper JA, Moritz G, Schweers J, Balach L, Roth C. Familial risk factors for adolescent suicide: A case-control study. Acta Pscyhiatr Scand 1994;89:52–58.

Brown P. No way out. New Scientist 1997;153(2074):34.

Brugha TS, Cragg D. The List of Threatening Experiences: The reliability and validity of a brief life events questionnaire. Acta Pscyhiatr Scand 1990;82:77–81.

Cheng ATA. Mental illness and suicide. Br J Psychiatry 1995;170:441–446. [PubMed: 9307694]

Cheng ATA, Chen THH, Chen CC, Jenkins R. Psychosocial and psychiatric risk factors for suicide: Case-control psychological autopsy study. Br J Psychiatry 2000;177:360–365. [PubMed: 11116779]

Clark, DC.; Horton-Deutsch, SL. Assessment in absentia: The value of the psychological autopsy method for studying antecedents of suicide and predicting future suicides. In: Mans, RW.; Berman, AL.; Maltsberger, JT.; Yufit, RI., editors. Assessment and prediction of suicide. New York: Guilford Press; 1992. p. 144-182.

Cohen J. A coefficient of agreement for nominal scales. Psychol Meas 1960;20:37.

Conner KR, Duberstein PR, Conwell Y. The validity of proxy-based data in suicide research: A study of patients 50 years of age and older who attempted suicide, I: Psychiatric diagnosis. Acta Psychiatr Scand 2001a;104:204–209. [PubMed: 11531657]

Conner KR, Conwell Y, Duberstein PR. The validity of proxy-based data in suicide research: A study of patients 50 years of age and older who attempted suicide, II: Life events, social support and suicidal behavior. Acta Psychiatr Scand 2001b;104:452–457. [PubMed: 11782238]

Davis, JA.; Smith, TW. General Social Survey, 1972–1993. Chicago: National Opinion Research Center; 1993.

Dunn, G. Design and analysis of reliability studies: The statistical evaluation of measurement errors. New York: Oxford; 1989.

Ebert BW. Guide to conducting a psychological autopsy. Prof Psychol Res Pract 1987;18:52–56.

Emmons RA, Colby PM. Emotional conflict and well-being: Relation to perceived availability, daily utilization, and observer reports of social support. J Pers Soc Psychol 1995;68:947–959. [PubMed: 7776190]

Epstein AM, Hall JA, Tognetti J, Son LH, Conant L Jr. Using proxies to evaluate quality of life: Can they provide valid information about patients' health status and satisfaction with medical care? Med Care 1989;27:S91–S98. [PubMed: 2921889]

Finch JF, Okun MA, Pool GJ, Ruehlman LS. A comparison of the influence of conflictual and supportive social interactions on psychological distress. J Pers 1999;67:581–622. [PubMed: 10444852]

Goldstein RB, Black DW, Nasrallah A, Winokur G. The prediction of suicide. Arch Gen Psychiatry 1991;48:418–422. [PubMed: 2021294]

Gordis, L. Epidemiology. Philadelphia, PA: Saunders; 1996.

Gould MS, Fisher P, Parides M, Flory M, Shaffer D. Psychosocial risk factors of child and adolescent completed suicide. Arch Gen Psychiatry 1996;53:1155–1162. [PubMed: 8956682]

Gu, J.; Chen, Y. [Instruction manual for the Structured Clinical Interview for DSM-III-R]. Kaohsiung, Taiwan: Kaohsiung Medical College Press; 1993. (Chinese ed)

Hamilton MA. A rating scale for depression. J Neurol Neurosurg Psychiatry 1960;23:56–62. [PubMed: 14399272]

Hawton K, Appleby L, Platt S, Foster T, Cooper J, Malmberg A, Simkin S. The psychological autopsy approach to studying suicide: A review of methodological issues. J Affect Disord 1998;50:269–276. [PubMed: 9858086]

Heikkinen ME, Aro HM, Lönnqvist JK. Recent life events and their role in suicide as seen by their spouses. Acta Psychiatr Scand 1992;86:489–494. [PubMed: 1471544]

Heikkinen M, Aro H, Lonnqvist J. Recent life events, social support and suicide. Acta Psychiatr Scand 1994;377:65–72.

Ji J, Kleinman A, Becker AE. Suicide in contemporary China: A review of China's distinctive suicide demographics in their socio-cultural context. Harvard Rev Psychiatry 2001;9:1–12.

Kelly TM, Mann JJ. Validity of DSM-III-R diagnosis by psychological autopsy: A comparison with clinician ante-mortem diagnosis. Acta Pscyhiatr Scand 1997;94:337–343.

Landerman R, George LK, Campbell RT, Blazer DG. Alternative models of the stress buffering hypothesis. Am J Commun Psychol 1989;17:625–642.

Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159–174. [PubMed: 843571]

Leckman JF, Sholomskas D, Thompson D, Belanger A, Weissman MM. Best estimate of lifetime psychiatric diagnosis. Arch Gen Psychiatry 1982;39:879–883. [PubMed: 7103676]

Lepine JP, Bouchez S. Epidemiology of depression in the elderly. Int Clin Psychopharmacol 1998;13:S7–S12. [PubMed: 9817614]

Lester D. The epidemiology of suicide in Chinese populations in six regions of the world. Chinese J Mental Health 1994;7:25–36.

Lin N. Measuring depressive symptomatology in China. J Nerv Ment Dis 1989;117:121–131. [PubMed: 2918295]

Macleod L. The dying fields. Far East Econ Rev 1998;23:62–63.

Maziade MRMA, Fournier JP, Cliché D, Merette C, Caron C, Garneau Y, Montgrain N, Shriqui C, Dion C, Nicole L, et al. Reliability of best-estimate diagnosis in genetic linkage studies of major psychoses: Results from the Quebec pedigree studies. Am J Psychiatry 1992;149:1674–1686. [PubMed: 1443244]

McCrae RR, Costa PT Jr. Personality trait structure as a human universal. Am Psychol 1997;52:509–516. [PubMed: 9145021]

McIntosh, JL. Methods of suicide. In: Maris, RW.; Berman, AL.; Maltsberger, JT.; Yufit, RI., editors. Assessment and prediction of suicide. New York: Guilford; 1992. p. 381-397.

Mireault M, deMan AF. Suicidal ideation among the elderly: Personal variables, stress and social support. Soc Behav Pers 1996;24:385–392.

Murphy GE, Wetzel RD, Robins E, McEvory L. Multiple risk factors predict suicide in alcoholism. Arch Gen Psychiatry 1992;49:459–463. [PubMed: 1599370]

Murray, C.; Lopez, A. The global burden of disease. Geneva, Switzerland: World Health Organization; 1996.

Nunes EV, Weissman MM, Goldstein RB, McAvay G, Seracini AM, Verdelli H, Wickramarantne PJ. Psychopathology in children of parents with opiate dependence and/or major depression. J Am Acad Child Adolesc Psychiatry 1998;37:1142–1151. [PubMed: 9808925]

Paykel ES. Life events, social support and depression. Acta Psychiatr Scand 1994;377:50–58.

Paykel ES, Prusoff BA, Uhlenhuth EH. Scaling of life events. Arch Gen Psychiatry 1971;25:340–347. [PubMed: 5116988]

Phillips MR, Li X, Zhang Y. Suicide rates in China, 1995–99. Lancet 2002;359:835–840. [PubMed: 11897283]

Phillips MR, Liu H, Zhang Y. Suicide and social change in China. Culture Med Psychiatry 1999;23:25–50.

Pritchard C. Suicide in the People's Republic of China categorized by age and gender: Evidence of the influence of culture on suicide. Acta Psychiatr Scand 1996;93:362–367. [PubMed: 8792906]

Qin P, Mortensen PB. Specific characteristics of suicide in China. Acta Psychiatr Scand 2001;103:117–121. [PubMed: 11167314]

Radloff LS. The CES-D scale: A self-report depression scale for research in the general population. Appl Psychol Meas 1977;1:385–401.

Reich W, Earls F. Rules for making psychiatric diagnoses in children on the basis of multiple sources of information: Preliminary strategies. J Abnormal Child Psychol 1987;15:601–616.

Reich W, Earls F, Frankel O, Shayka JJ. Psychopathology in children of alcoholics. J Am Acad Child Adolesc Psychiatry 1993;32:995–1002. [PubMed: 8407776]

Rook KS. Assessing the health-related dimensions of older adults' social relationships. Ann Rev Gerontol Geriatr 1994;14:142–181.

Schwarzer R, Leppin A. Social support and health: A meta-analysis. Psychol Health 1989;3:1–15.

Spitzer RL, Williams JBW, Gibbon M, First AB. Instruction Manual for the Structured Clinical Interview for DSM-III-R (SCID, 6/1/88 Revision). Biometrics Research. 1988

Tiller J, Kupinski J, Burrows G, Mackenzie A, Hallenstein H, Johnston G. Completed and attempted youth suicide in Victoria. Stress Med 1998;14:249–254.

Velting DM, Shaffer D, Gould MS, Garfinkel R, Fisher P, Davies M. Parent-victim agreement in adolescent suicide research. J Am Acad Child Adolesc Psychiatry 1998;37:1161–1166. [PubMed: 9808927]

Vijayakumar L, Rajkumar S. Are risk factors for suicide universal? A case-control study in India. Acta Psychiatr Scand 1999;99:407–411. [PubMed: 10408261]

Weissman MM, Gammon GD, John K, Merikangas KR, Warner V, Prusoff BA, Sholomskas D. Children of depressed parents: Increased psychopathology and early onset major depressive episode. Arch Gen Psychiatry 1987;44:847–853. [PubMed: 3662741]

Williams JBW. A structured interview guide for the Hamilton Depression Rating Scale. Arch Gen Psychiatry 1988;45:742–747. [PubMed: 3395203]

World Health Organization. World health statistics annual. Geneva, Switzerland: WHO; 1988–1995.

Yang J, McCrae RR, Costa PT Jr, Dai X, Yao S, Cai T, Gao B. Cross-cultural personality assessment in psychiatric populations: The NEO-PI-R in the People's Republic of China. Psychol Assess 1999;11:359–368.

Yip PSF. An epidemiological profile of suicides in Beijing, China. Suicide Life Threat Behav 2001;31:62–70. [PubMed: 11326769]

Yip P, Callanan C, Yuen H. Urban/rural and gender differentials in suicide rates: East and west. J Affect Disord 2000;57:99–106. [PubMed: 10708821]

Zhang J. Suicide in Beijing, China, 1992–1993. Suicide Life Threat Behav 1996;26:175–180. [PubMed: 8840421]

Zhang J. Gender differences in athletic performance and their implications in gender ratios of suicide: A comparison between the USA and China. Omega J Death Dying 2000a;41:117–123.

Zhang J. Understanding Chinese suicide with a comparison of national data. Am Rev China Stud 2000b; 1:9–29.

Zhang J, Jia S, Wieczorek WF, Jiang C. An overview of suicide research in China. Arch Suicide Res 2002a;6:167–184.

Zhang J, Norvilitis JM. Measuring Chinese psychological well-being with Western developed instruments. J Pers Assess 2002;79:502–521.

Zhang J, Thomas DL. Modernization theory revisited: A cross-cultural study of adolescent conformity to significant others in mainland China, Taiwan, and the USA. Adolescence 1994;29:885–903. [PubMed: 7892799]

Zhang J, Wieczorek WF, Jiang C, Zhou L, Jia S, Sun Y, Jin S, Conwell Y. Studying suicide with psychological autopsy: social and cultural feasibilities of the methodology in China. Suicide Life Threat Behav 2002b;32:370–379. [PubMed: 12501962]

Zheng Y, Zhao J, Phillips M, Liu J, Cai M, Sun S, Huang M. Validity and reliability of the Chinese Hamilton Depression Rating Scale. Br J Psychiatry 1988;152:660–664. [PubMed: 3167442]

TABLE 1

Descriptions of Targets and Informants

Characteristics	Informant First (N = 66) N (%)	Informant Second (N = 66) N (%)	Target (N = 66) N (%)
Gender			
Male	13 (19.7)	22 (33.3)	48 (72.7)
Female	53 (80.3)	44 (66.7)	18 (27.3)
Age, mean	44.3	41.3	44.9
SD	12.2	13.6	16.7
Family size, Mean	3.2	3.3	3.2
SD	1.0	1.1	1.0
Family income, mean	10,735	12,630	11,902
SD	8240.2	10125.2	11709.0
Education, y	7.1	8.0	7.6
SD	3.4	3.7	3.6
Education level			
No school	3 (4.5)	3 (4.5)	4 (6.1)
Elementary	28 (42.4)	19 (28.8)	23 (34.8)
Middle school	25 (37.9)	26 (39.4)	26 (39.4)
High school	2 (3.0)	5 (7.6)	5 (7.6)
Two-year college	8 (12.1)	11 (16.7)	6 (9.1)
University	—	2 (3.0)	2 (3.0)
Marital status			
Single	3 (4.5)	7 (10.6)	9 (13.6)
Married, together	61 (92.4)	55 (83.3)	52 (78.8)
Married, separate	—	—	1 (1.5)
Remarried	—	1 (1.5)	1 (1.5)
Widowed	2 (3.0)	3 (4.5)	3 (4.5)
Residence card			
Urban	8 (12.1)	10 (15.2)	10 (15.2)
Rural	58 (87.9)	56 (84.8)	56 (84.8)
Phone at home			
Yes	41 (62.1)	42 (63.6)	40 (60.6)
No	25 (37.9)	24 (36.4)	26 (39.4)

TABLE 2

Interrater Reliabilities for the Five Interviewers

	No. Correlations $p \leq .01$ (2-tailed)	No. Correlations $p \leq .05$ (two-tailed)	No. Correlations NS
Suicidal Intention	7	3	0
Social Support—			
Interaction	4	5	1
Social Support—			
Perceived	1	2	7
Social Support—			
Instrument	6	0	4
HAM-D	10	0	0
CES-D	10	0	0
GSS	10	0	0
NEO total score	3	3	4
Life Events total score	6	2	2
SCID total score	4	6	0

TABLE 3

Comparison of Mental Disease Diagnoses Provided by Targets and Informants

Diagnostic Category	Kappa	Sensitivity	Specificity	% Agreement
Alcohol	1.00	1.00	1.00	100
Eating	1.00	1.00	1.00	100
Anxiety	.66	.50	1.00	98
Depression	1.00	1.00	1.00	100
Organic mood	1.00	1.00	1.00	100

TABLE 4

Comparison of Life Events Reported by Targets and Informants

Life Events Area	Kappa	Sensitivity	Specificity	% Agreement
Work/Business	.55	.81	.78	79%
Family/Home	.50	.82	.68	76%
Health/Hospital	.71	.91	.91	91%
Marriage/Romance	.42	.71	.86	85%
Relationship	.30	.57	.85	82%

TABLE 5

Correlations Between Targets and Informants on Responses to Selected Scales

Scale	Mean (SD) Informant	Mean (SD) Target Self	Spearman's rho
Suicidal Intention	1.12 (.54)	1.05 (.21)	.21[*]
Social Support—Interaction	9.80 (1.81)	8.67 (2.36)	.53[**]
Social Support—Perceived	19.39 (2.33)	17.61 (3.13)	.55[**]
Social Support—Instrument	12.18 (.61)	13.39 (2.33)	.30[**]
HAM-D	7.39 (10.07)	6.94 (7.56)	.57[**]
Life Events total score	2.63 (2.15)	1.79 (1.51)	.57[**]
NEO total score	196.50 (11.07)	197.41 (11.97)	.43[**]
SCID total score	1.03 (.35)	1.05 (.37)	.91[**]

[*] Correlation is significant at the .05 level (one-tailed).

[**] Correlation is significant at the .01 level (one-tailed).

TABLE 6

Comparing the Means of Selected Scales Between Targets and Informants on Responses

		Mean (SD)	t	p		Mean (SD)	t	p
Social Support—Interaction	Info#1	8.58 (2.01)	−0.24	0.81	Info#2	8.85 (2.33)	0.45	0.66
	Self	8.67 (2.36)			Self	8.67 (2.36)		
Social Support—Perceived	Info#1	17.77 (3.33)	0.30	0.77	Info#2	17.77 (3.38)	0.29	0.77
	Self	17.61 (3.13)			Self	17.61 (.13)		
Social Support—Instrument	Info#1	12.73 (1.25)	−2.05	0.04	Info#2	12.97 (1.85)	−1.16	0.25
	Self	13.39 (2.33)			Self	13.39 (2.33)		
HAM-D	Info#1	5.68 (9.48)	−0.84	0.40	Info#2	4.65 (7.39)	−1.76	0.08
	Self	6.94 (7.56)			Self	6.94 (7.56)		
Number of Life Events	Info#1	1.89 (1.50)	0.40	0.69	Info#2	1.80 (1.61)	0.06	0.96
	Self	1.79 (1.51)			Self	1.79 (1.51)		
NEO-FF-I	Info#1	196.5 (11.1)	−0.45	0.65	Info#2	196.5 (11.9)	−0.43	0.67
	Self	197.4 (12.0)			Self	197.4 (12.0)		
SCID	Info#1	1.03 (.35)	−0.24	0.81	Info#2	1.03 (.35)	−0.24	0.81
	Self	1.05 (.37)			Self	1.05 (.37)		