

# SIDAM – A Structured Interview for the diagnosis of Dementia of the Alzheimer type, Multi-infarct dementia and dementias of other aetiology according to ICD-10 and DSM-III-R

M. ZAUDIG,<sup>1</sup> J. MITTELHAMMER, W. HILLER, A. PAULS, C. THORA, A. MORINIGO  
AND W. MOMBOUR

*From the Max Planck Institute for Psychiatry, Psychiatric Out-patient Department, Munich, Germany*

**SYNOPSIS** The SIDAM – a new instrument for the symptomatic diagnosis and measurement of dementia according to DSM-III-R and ICD-10 – is described. It comprises a brief structured clinical interview, a range of cognitive tests (e.g. including the Mini-Mental State (Folstein *et al.* 1975)) which constitute a short neuropsychological battery and a section for clinical judgement and third party information. All items rely on DSM-III-R and ICD-10 algorithms. The SIDAM has a high overall test-retest reliability which equally holds true on the diagnostic, criterion and item level. It is a brief (average of 28 min), practical and easily scored diagnostic instrument, which reliably separates subjects with DSM-III-R and ICD-10 dementia from those without such a disorder. Good congruence was found between SIDAM diagnoses and corresponding ICD-9 expert diagnoses. Furthermore, the SIDAM-Score (SISCO) allows a detailed measurement of even low levels of cognitive impairment and provides quantification of severity grading of cognitive dysfunction.

## INTRODUCTION

Dementia is one of the major health problems of old people. Determining the presence of the disorder and its severity (including early stages) is central to any clinical, biological and epidemiological study as well as to the physician's everyday practice. Unfortunately, diagnostic accuracy related to dementias is still less well developed than for many other disorders (Wells, 1982; Henderson & Huppert, 1984; Cummings & Benson, 1986). Thus, the publication of DSM-III/DSM-III-R (American Psychiatric Association, 1980, 1987) was a major advance, since it provides diagnostic criteria for dementias and a better specification of criteria (Jorm & Henderson, 1985). In contrast to DSM-III, operational criteria for grading severity of the dementias have been incorporated into DSM-III-R (APA, 1987) and, in part, into the draft of the ICD-10 research diagnostic criterion section F0

(‘Organic, including symptomatic, mental disorders’) (World Health Organization, 1987, 1988). Classification systems like ICD-10 and DSM-III-R may therefore offer a more comprehensive and reliable diagnostic system for dementias.

### Comprehensive structured interviews

To enhance reliability of psychiatric diagnoses and to diminish variation among raters and ratings (Spitzer & Williams, 1985), structured or semi-structured diagnostic interviews can be applied such as the Diagnostic Interview Schedule (DIS) (Robins *et al.* 1981), the Present State Examination (PSE) (Wing *et al.* 1974) or the Structured Clinical Interview for DSM-III-R (SCID) (Spitzer *et al.* 1986). However, neither of these instruments has been designed for diagnosing and differentiating specific categories of dementia. At present, the Cambridge Mental Disorders of the Elderly Examination (CAMDEX) is considered to be a comprehensive and reliable diagnostic instrument (Roth *et al.* 1986) which includes all relevant information to

<sup>1</sup> Address for correspondence: Dr M. Zaudig, Max Planck Institute for Psychiatry, Kraepelinstrasse 10, D-8000 Munich 40, Germany.

make a diagnosis and differential diagnosis. Other similarly structured and equally comprehensive instruments for diagnosis of dementias are the Geriatric Mental State Examination (GMS) (Copeland *et al.* 1976) and the Comprehensive Assessment and Referral Evaluation (CARE) (Gurland *et al.* 1977) which includes the GMS. These instruments are very useful for differential diagnosis (e.g. depression *versus* dementia), but are also lengthy and therefore often difficult to apply to older people. Furthermore, no diagnostic algorithms for DSM-III-R and ICD-10 are included.

### Brief clinical tests

Due to difficulties in diagnosing elderly people with large and comprehensive interviews, a variety of simple psychometric scales were developed to allow the examiner more flexibility in evaluating the patient's cognitive performance. Currently used brief cognitive tests are, for example, the Mini-Mental State Examination (MMSE) (Folstein *et al.* 1975) or the Short Portable Mental Status Questionnaire (SPMSQ) (Pfeiffer, 1975). For grading the severity of dementias, rating scales like the Clinical Dementia Rating (CDR) (Hughes *et al.* 1982) and the Global Deterioration Scale (GDS) (Reisberg *et al.* 1982) have been found useful. Moreover, a variety of behavioral rating scales are in use, by which the subject is observed and rated based on daily life activities and abilities, e.g. the Dementia Scale (DS) (Blessed *et al.* 1968) or Behavioral Scale for Dementia (Haycox, 1984).

Two problems, however, arise from the application of such instruments: first, a clinical examination always requires a variety of additional information, but most of these scales evaluate problems in only one or two fields, e.g. memory or daily life abilities (thus they are not comprehensive enough to arrive at a more complex clinical evaluation). Secondly, they fail to arrive at a specific diagnosis according to a psychiatric classification system such as DSM-III-R or ICD-10.

### Brief diagnostic instruments

The authors have therefore developed a structured clinical diagnostic instrument aimed at an early and comprehensive evaluation of dementias using operationally defined diagnostic criteria of ICD-10 and DSM-III-R. It has been

named SIDAM: Structured Interview for the diagnosis of Dementias of the Alzheimer type and Multi-infarct dementia and dementias of other aetiology. In addition to cognitive testing (which is a central part of the SIDAM), the SIDAM includes information such as psychosocial and behavioural data as well as psychopathological data. We intended to construct a brief diagnostic instrument – not to replace any of the comprehensive instruments – which should be easy to administer and as brief as possible but still comprehensive enough to make at least a syndromal diagnosis of dementia. The SIDAM therefore should contain all necessary information to arrive at a comprehensive syndromal diagnosis of dementia which is measured by the severity and extent of cognitive impairment. For all other elements of a diagnostic process (differential diagnosis, physical state and laboratory tests) only diagnostic guidelines should be included according to the diagnostic algorithms, such as those set down in DSM-III-R and ICD-10.

A test-retest study was conducted to evaluate the reliability of the SIDAM. This method reflects a more demanding 'real life' test than the method of inter-rater reliability, where a patient is simultaneously rated by two interviewers. The advantage of a test-retest study is that it more closely reflects actual practice where independent assessments are the rule (Cicchetti & Prusoff, 1983; Burke, 1988; Williams, 1988).

More specifically, we will address the following questions: (1) What is the test-retest reliability of the SIDAM on the item, criterion and diagnostic level? (2) What is the test-retest reliability of the MMSE, EMMS, the Rosen Score and the SIDAM-Score (SISCO)? (3) How useful is the SIDAM as a brief diagnostic instrument?

### STRUCTURE OF THE SIDAM

In developing the SIDAM (Zaudig *et al.* 1989, 1990), algorithms for the syndrome of dementia and for diagnosing dementias of primary degenerative cause, of vascular origin and of a specific aetiology according to DSM-III-R and ICD-10 were incorporated as well as a broad range of cognitive functions (for the SIDAM the WHO ICD-10 research criteria (1987) were used; they are still in draft form. The final version will

**B. Intellectual/ Cognitive Abilities and Personality**

**B.1. Impairment of abstract thinking**

29. I will now give you *two terms*.  
Would you please tell me the difference  
between both terms?

a) RIVER - LAKE

Answer: .....

?  0  1

b) LADDER - STAIRS

Answer: .....

?  0  1

30. I shall give you a *proverb*. -  
What is the meaning of: "The apple does not fall  
far from the tree" ?

Answer: .....

?  0  1

**B.2. Impaired judgement**

31. I am going to ask you a *question*:  
"A workman fell from a ladder and broke his two legs.  
In order to get immediate medical treatment,  
he ran to the nearby hospital."  
- *Do you think this was right?*

Answer: .....

?  0  1

32. What *happens* on this picture ?  
(show Luria-figure on sheet 2).

Answer: .....

?  0  1

**ICD-10:**

Criterion A.2. (B.1/B.2.) is met  
if at least one question  
of items 29 to 32 was coded "0" :

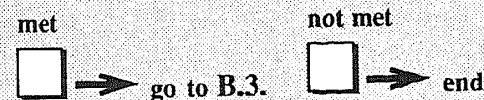


FIG. 1. Sample page from the SIDAM.

be published in about 1992). Problems in operationalizing criteria (DSM-III) in a formal manner had been discussed in detail by Jorm & Henderson (1985) and Robins (1989). Similar problems were found, when the SIDAM was constructed. Poor specification of diagnostic criteria for the syndrome of dementia was a major problem. Each of the DSM-III-R and ICD-10 criteria (each criterion includes certain cognitive functions) had to be translated into specific questions. If items of the MMSE/ Expanded Mini-Mental State (EMMS) were used, they had to be assigned to the specific DSM/ICD criteria. Another difficulty in constructing the SIDAM was to find cut-off points in deciding if a single criterion had been met or not. ICD-10 and DSM-III-R do not specify these cut-off points, apart from some minor exceptions in both systems. In general we defined cut-off points, intentionally, very broadly, to achieve high sensitivity. For example DSM-III-R criterion A (memory impairment) was 'translated' into 13 items (18 questions) for testing memory functions. At least one of the short-term memory questions *and* one of the long-term memory

questions have to be coded as error in order to consider criterion A (memory impairment) as fulfilled.

For the global concept of the syndrome of dementia and its different features we derived an overall score from the cognitive section, named SISCO (SIDAM SCORE). In addition, other measures of cognitive function were included, like the MMSE and EMMS. The SISCO consists of 40 items (55 questions), covers a broad range of cognitive functions and allows cut-offs for impairment of a given level of severity or mild cognitive dysfunction (Jorm & Henderson, 1985).

The SIDAM is constructed primarily to differentiate between non-demented and demented subjects according to the categories of DSM-III-R and ICD-10. But DSM-III-R and ICD-10 criteria provide no category for a diagnosis of 'cognitive change of other type', like ICD-9 (310.1) or 'mild dementia' (Henderson & Huppert, 1984). Therefore, we sought to define this group (with mild to marked cognitive impairment but not meeting DSM/ICD-10 criteria for dementia) in accordance to the

concepts of 'mild dementia' (Henderson & Huppert, 1984; Mowry & Burville, 1988) and 'limited dementia' (Gurland *et al.* 1982, 1983), which refers to memory impairment not interfering with the subject's ability to live independently (details are reported by Zaudig *et al.* 1990).

## DESCRIPTION OF THE SIDAM

The SIDAM is subdivided into the following main parts.

(a) A brief semistructured clinical overview with open-ended questions to the patient and/or informant; it covers for instance basic information on the subject's past and present medical and psychiatric history.

(b) A standardized performance test to evaluate cognitive performance using 40 items (one item may be composed of several questions). An example of this part of the SIDAM is shown in Fig. 1. Correct answers can be summed, resulting in the SISCO, with a range from 0 (worst cognitive impairment) to 55 (no cognitive impairment). The SISCO can be divided into 10 subscores (syndrome-scores) which may be used independently in order to evaluate performance in specific areas such as orientation, memory or judgement. In this way, a very specific profile of cognitive abilities can be obtained. The SISCO covers a broader range of cognitive functions than the MMSE and detects also mild degrees of cognitive impairment. Furthermore, all items of the MMSE (Folstein *et al.* 1975; Anthony *et al.* 1982) and the Expanded Mini-Mental State (EMMS) (Farmer & Helzer, 1990) are included in this SIDAM part. The MMSE is incorporated also into the DIS (Robins *et al.* 1981). Its administration and scoring were carried out according to the DIS-guidelines.

(c) A structured schedule for clinical judgement of dementia-related disabilities containing the Hachinski Score (Hachinski *et al.* 1975) and the Modified Ischemic Score (Rosen *et al.* 1980) into the diagnostic algorithm for evaluation of multi-infarct dementia. Third-party information regarding personality change and social performance must be considered in these SIDAM sections. This part of SIDAM also contains guidelines for differential diagnoses of psychiatric disorders, but does not by itself contain the necessary information to arrive at a specific

diagnosis like mood disorder; here, the clinician needs to use his own judgement or specific diagnostic instruments (e.g. the SCID interview).

(d) A severity grading for dementia (DSM-III-R/ICD-10).

(e) A listing of present and past medical disorders (sources of this kind of information should be physicians or medical records).

(f) A summary record for all scores contained in the SIDAM.

(g) An easily scoreable diagnostic algorithm (DSM-III-R: Criteria A-E; ICD-10: Criteria A<sub>1</sub>-D) for dementia.

Generally, third-party information in some SIDAM parts (*a*, *c* and *e* as listed above) is mandatory, but there may be some exclusions, e.g. a subject is clearly healthy and has no cognitive impairment.

## METHOD

### Subjects

For the reliability study a sample of 60 subjects (48 female and 12 male) with a mean age of 77 years (61-90 years: 7.33 years s.d.) was examined. Subjects were recruited mainly from a community home for the aged and from the outpatient department of the Max Planck Institute for Psychiatry, both in Munich. In both institutions there was a preponderance of cognitively impaired and demented subjects (which was intended by the investigators). Most subjects responded to a public announcement in the above-mentioned institutions; some known patients were also included. Thirty-five per cent were living alone in the community, 23% with their spouse or other family members, 42% were in the community home and a few of them in warden-controlled sheltered accommodation. Fifty-five per cent of the subjects were widowed, 24% still married, 13% were never married, 8% were divorced. Except for two patients, all had had at least eight years of school education.

All subjects met the following inclusion criteria: at least 60 years old, currently no acute psychiatric disorder (e.g. schizophrenia, alcohol dependence), no acute confusional or delirious state.

### Design

The pre-test consisted of four patient interviews with SIDAM, each of them conducted by one of

four psychiatrists and observed and scored by the other three. Each interview was followed by a general discussion of the patient's responses and the interviewer's scoring.

Each of the 60 patients participating in the study of test-retest reliability was interviewed twice by different interviewers. To diminish any bias due to individual interviewer style we used a balanced design: the pairing of interviewers was systematically rotated so that each of the six pairs (AB/AC/AD/BC/BD/CD) examined an equal number of patients and each member of a pair was the first interviewer; in an equal number of cases each psychiatrist conducted 30 interviews and each interviewer was a test-retest partner for each of his three colleagues across 10 subjects – five times as test interviewer and five times as retest interviewer. In order to reduce subject variance (due to possible changes in the patient's mental state) the time interval between test and retest examination was usually kept between 24 and 72 hours, with a mean of 26 hours. The interview setting was also standardized as far as possible; test and retest interviews were conducted in the same room at the same hour of the day and with only the patient and the interviewer present.

All interviewers were instructed not to discuss their own clinical diagnosis or SIDAM scoring (SIDAM-diagnosis) with the other team members until test and retest interviews had been finished. Once a week, after completion of the interviews, all results were discussed (including all additional information about the patients) by the four clinicians and final clinical consensus diagnoses were assigned according to ICD-9.

#### Characteristics of the interviewers

Four psychiatrists participated as interviewers. They ranged in clinical experience from senior psychiatrists (M. Z., J. M., A. M.) to a third-year resident (A. P.); A. M. is a Spanish psychiatrist with fluency in the German language.

#### Diagnostic procedure and psychopathological description of the sample

In all 120 interviews a SCID interview (Structured Clinical Interview for DSM-III-R, Spitzer *et al.* 1986; German version: Wittchen *et al.* 1987) was conducted with selective use of the section on mood disorders to exclude depressive syndromes, since SIDAM does not by

itself contain the necessary information to establish a specific differential diagnosis of mood disorder. Following this, the whole SIDAM was administered with a mean duration of 28 minutes (s.d. = 8 min); in calculating this figure the time taken to obtain the necessary information from a third party was included, but not the time to arrive at a differential diagnosis (SCID interview). After completion of the SIDAM each subject was rated by the IMPS (In-patient Multidimensional Psychiatric Scale, Lorr & Klett 1967; Mombour *et al.* 1973; Mombour, 1974; Hiller *et al.* 1986) and tested with the WAIS digit symbol test and the WAIS block design test (Wechsler-Adult-Intelligence Scale (WAIS), Wechsler, 1945, 1954, 1958). For reliable grading of severity of the dementia, we used the Clinical Dementia Rating (CDR) of Hughes *et al.* (1982) and the Global Deterioration Scale (GDS) of Reisberg *et al.* (1982). Furthermore, each patient was rated according to DSM-III-R Axis V (Global Assessment of Functioning (GAF)) (APA, 1987).

#### Statistical analyses and methods

Estimates of the SIDAM reliability are mainly based on the kappa ( $\kappa$ ) statistic (Cohen, 1960) where the observed overall agreement between two binary ratings ( $p_o$ ) is corrected for agreement expected by chance. The kappa statistic currently represents the standard method to assess diagnostic agreement in psychiatry (Fleiss, 1981; Shrout *et al.* 1987). Confidence intervals for kappa of 95% were shown additionally (Fleiss, 1981), values of above 1.0 were rounded to 1.0, since 1.0 represents maximum agreement.

We additionally employed Yule's  $Y$  (Yule, 1912). Unlike  $\kappa$ , this coefficient is independent of sample base rates, and it has been proposed by Spitznagel & Helzer (1985) as an alternative to  $\kappa$  in studies with low base rates.  $Y$  is closely related to  $\kappa$ , since it can be regarded as an approximation to maximum  $\kappa$  over all possible base rates (Spitznagel & Helzer, 1985). Whenever a single cell of the classification table became 0,  $Y$  values were calculated using the pseudo-Bayes estimation procedure (otherwise,  $Y$  would have reached the endpoint value of 1 despite incomplete congruence). Tests of significance were performed for  $\kappa$  (Bartko & Carpenter, 1976), one-tailed throughout (considering only positive deviations from 0), with a significance level of

5%. For a more detailed interpretation ( $\kappa$  values of 0.40 and above were regarded as acceptable (0.40–0.49), fair (0.50–0.59), good (0.60–0.69),  $\kappa$  values of 0.70 and above as excellent (Burke, 1986; Fleiss, 1981, p. 218). The same standards were applied to  $Y$  values. The reliability of quantitative measures (syndromes, scores) is expressed by correlation coefficients (Spearman rank correlation for ordinal-scaled data, Pearson correlation for interval-scaled data).

## RESULTS

### Group description and comparisons with regard to ICD-9 consensus diagnoses

According to the ICD-9 consensus diagnoses there were 17 normal (no cognitive impairment) subjects, 25 subjects with a 'cognitive change of other type' (ICD-9: 310.1) and two subjects with other psychiatric diagnoses. Sixteen subjects were diagnosed as suffering from dementia (ICD 9: 290.x, 294.x). With regard to the ICD-9 consensus diagnoses, the subjects with no cognitive impairment are characterized by a mean MMSE-score of 28.7 (range 24–30), the CDR median was 0 (which means no cognitive impairment) and the GDS median was 2. The GAF score ranged from 70 to 90 with a mean of 79.8 illustrating no more than slight psychosocial problems. In subjects with 'cognitive change of other type' (ICD-9: 310.1) the mean MMSE score was 26.3, the CDR median was 0.5 and the GDS median 3. As expected, a low mean MMSE score of 14.9 was found for dementia, along with a CDR median of 2 (moderate dementia) and a GDS median of 5 (early dementia).

### Test-retest reliability on diagnostic level

In Table 1 the concordance rates (percentage agreement,  $\kappa$ , and  $\gamma$ ) of the SIDAM diagnoses with regard to ICD-10 and DSM-III-R are summarized. On the diagnostic level the percentage agreement varied between 95 and 98%. Kappa values ranged from 0.64 to 0.95. SIDAM ICD-10 dementia (all dementias without subgrouping) showed the highest  $\kappa$  value with 0.95, respectively (Table 1).

For ICD-10 SIDAM-diagnoses for dementia due to cerebrovascular disease  $\kappa$  was 0.64. For all other specific diagnoses, excellent  $\kappa$  values

(0.73 and 0.85) were found (Table 1). Similar results could be demonstrated for SIDAM DSM-III-R diagnoses with a  $\kappa$  of 0.64 for multi-infarct dementia and a  $\kappa$  of 0.87 in all dementias without subgrouping (Table 1).

### Test-retest reliability on criterion level

Original ICD-10 and DSM-III-R criteria are represented by all the SIDAM-sections. In order to evaluate SIDAM ICD-10 criteria we generally found fair to excellent concordance rates with  $\kappa$  between 0.57 and 0.95 (Table 2). For ICD-10 criterion D the percentage agreement was 85%,  $\kappa = 0.69$  (Yule's  $\gamma = 0.74$ ). DSM-III-R criteria presented with excellent concordance rates from  $\kappa = 0.79$  to 0.87 (Table 2). For DSM-III-R criteria  $E_1$  and  $E_2$  we found 97 and 100% agreement and kappa values of 0.49 ( $\gamma = 1.0$ ) and  $\kappa = 1.0$  ( $\gamma = 1.0$ ) respectively.

### Test-retest reliability on item level

Agreement values were calculated for each of the 139 questions which are summarized to 52 items in the SIDAM. In 74% of the questions, significant  $\kappa$  values of 0.4 and more were obtained (in 28% more than 0.7). The overall proportion of agreement ranged from 55 to 100%. For some subsections of cognitive performance mean  $\kappa$  values were calculated: 0.70 for orientation (10 items); 0.44 for short-term memory with recall items included (14 items); 0.53 for abstract thinking (2 items); 0.50 for judgement (2 items); 0.69 for long-term memory (7 items). In the SIDAM-section of 'personality change' the mean  $\kappa$  value for personality change (3 items) was 0.52.

For most of the SIDAM MMSE-items (19 items)  $\kappa$  values higher than 0.40 were found with a mean  $\kappa$  value of 0.58. EMMS-items showed acceptable to excellent kappa values. According to the items of the modified ischemic score (Rosen *et al.* 1980), in general fair to excellent  $\kappa$ -values were found (0.50–0.83), with the exception of two items: 'stepwise deterioration' ( $\kappa = 0.25$ ) and 'somatic complaints' ( $\kappa = 0.30$ ). These low  $\kappa$  values (both insignificant) were due to information variance for both values (subjects gave different information to the interviewers) and criterion variance for the latter (vague definition of the item). The percentage agreement was 92 and 88%, respectively. The mean  $\kappa$  for the Rosen-score (8 items) was 0.64. DSM-III-R

Table 1. Concordance rates for SIDAM ICD-10 and DSM-III-R diagnoses

SIDAM diagnosis	Test	Retest		$\kappa^*$	$\kappa$ 95% confidence interval	Y	% Agreement
		-	+				
ICD-10		56	1	<b>0.85</b>	(0.56-1.0)	<b>0.85</b>	98
Dementia, not otherwise specified (F 03)		0	3				
Dementia of the Alzheimer type (F 00.x)		55	0	<b>0.73</b>	(0.38-1.0)	<b>0.85</b>	97
		2	3				
Vascular dementia (F 01.x)		54	2	<b>0.64</b>	(0.26-1.0)	<b>0.80</b>	95
		1	3				
All dementias (F 00.x-F 03)		47	0	<b>0.95</b>	(0.85-1.0)	<b>0.93</b>	98
		1	12				
DSM-III-R		51	1	<b>0.64</b>	(0.26-1.0)	<b>0.80</b>	95
Senile dementia, not otherwise specified (290.00)		2	3				
Primary degenerative dementia of the Alzheimer type, senile onset, uncomplicated (290.00/331.00)		53	1	<b>0.82</b>	(0.57-1.0)	<b>0.88</b>	97
		1	5				
Multi-infarct dementia uncomplicated (290.40)		54	2	<b>0.64</b>	(0.26-1.0)	<b>0.80</b>	95
		1	3				
All dementias (290.x)		43	1	<b>0.87</b>	(0.73-1.0)	<b>0.89</b>	95
		2	14				

\* All  $\kappa$  values are significant at the 5% level.

Table 2. Concordance rates for SIDAM diagnostic criteria (ICD-10 and DSM-III-R)

Diagnostic criteria	Test	Retest		$\kappa^*$	$\kappa$ 95% confidence interval	Y	% Agreement
		-	+				
ICD-10		0	4	<b>-0.05</b>	(-0.24-0.14)	<b>-0.02</b>	90
Criterion A <sub>1</sub>		2	54				
Criterion A <sub>2</sub>		33	5	<b>0.57</b>	(0.36-0.79)	<b>0.59</b>	80
		7	16				
Criterion C		44	2	<b>0.81</b>	(0.64-0.99)	<b>0.84</b>	93
		2	12				
Criterion A <sub>1</sub> + A <sub>2</sub>		47	0	<b>0.95</b>	(0.85-1.0)	<b>0.93</b>	98
		1	12				
Criterion B		47	0	<b>0.95</b>	(0.85-1.0)	<b>0.93</b>	98
		1	12				
DSM-III-R		21	1	<b>0.79</b>	(0.64-0.95)	<b>0.84</b>	90
Criterion A		5	33				
Criterion B		25	2	<b>0.83</b>	(0.69-0.97)	<b>0.84</b>	92
		3	30				
Criterion C		43	1	<b>0.87</b>	(0.73-1.0)	<b>0.89</b>	95
		2	14				
Criterion D		43	1	<b>0.82</b>	(0.66-0.99)	<b>0.86</b>	93
		3	13				

\* All  $\kappa$  values are significant at the 5% level except for ICD-10 criterion A<sub>1</sub>.

criteria for severity of dementia (mild, moderate, severe) were found to have a test-retest correlation of  $r = 0.67$  (rank correlation); a higher correlation was found for ICD-10 severity of memory impairment (mild, moderate, severe) with  $r = 0.80$  (rank correlation). However, in general the proportion of SIDAM items with low  $\kappa$  coefficients proved relatively small. The answers related to orientation, long-term mem-

ory, higher cortical function or judgement of personality change showed good or excellent agreement.

#### Test-retest reliability of SIDAM syndrome scores

In the test performance part of the SIDAM, 10 different cognitive syndromes are contained, each of which can be scored (syndrome scores)

Table 3. Concordance rates of a sample of SIDAM items

SIDAM item (◆ MMS item)	(Item number)	Test-Retest-	$\kappa^*$	$\kappa$ 95% confidence interval	Y	% Agreement
		combinations + + / - - / + - / - +				
Orientation	(01)	47/08/03/02	0.71	(0.47-0.95)	0.78	92
◆ Year						
◆ Season	(02)	41/10/04/02	0.70	(0.48-0.92)	0.76	90
◆ Date	(03)	33/17/06/04	0.64	(0.44-0.84)	0.66	83
◆ Day	(04)	44/09/02/05	0.65	(0.41-0.89)	0.73	88
◆ Month	(05)	51/06/01/02	0.77	(0.52-1.0)	0.85	95
Abstract thinking						
Difference river - lake	(29a)	47/04/03/06	0.39	(0.06-0.71)	0.53	85
Difference ladder - stairs	(29b)	44/06/04/06	0.44	(0.15-0.74)	0.54	83
Proverb	(30)	41/12/03/04	0.70	(0.49-0.90)	0.73	88
Judgement						
Story	(31)	44/06/02/08	0.45	(0.16-0.74)	0.60	83
Luria figure	(32)	47/06/06/02	0.57	(0.28-0.85)	0.75	88
Higher cortical functions						
Cube	(33)	16/34/03/07	0.64	(0.43-0.84)	0.67	83
3-D-figure	(34)	17/30/06/07	0.55	(0.33-0.76)	0.55	78
◆ Watch	(35a)	58/01/01/00	0.66	(0.03-1.0)	0.73	98
◆ Pencil	(35b)	58/01/00/01	0.66	(0.03-1.0)	0.73	98
Thumb on ear	(38)	49/06/04/01	0.66	(0.38-0.94)	0.79	92
Knuckles	(39)	44/10/01/05	0.71	(0.49-0.93)	0.81	90
Personality change	(41)	12/41/05/02	0.70	(0.49-0.90)	0.75	88
Altered personality						
Social behaviour	(42)	03/50/04/03	0.40	(0.03-0.76)	0.56	88
Impairment of social functioning						
Social and work performance	(44)	13/43/02/02	0.82	(0.65-0.99)	0.84	93
Memory						
◆ Recall apple	(16a)	23/25/07/05	0.60	(0.40-0.80)	0.60	80
◆ Recall table	(16b)	19/31/05/05	0.65	(0.46-0.85)	0.66	83
◆ Recall penny	(16c)	11/33/10/06	0.39	(0.14-0.64)	0.42	73
Recall design	(27)	14/30/08/08	0.43	(0.19-0.66)	0.44	73
Recall street	(28c)	14/30/15/01	0.46	(0.22-0.70)	0.68	73
World War II	(23)	34/21/03/02	0.83	(0.68-0.97)	0.83	92
Pope	(24)	55/03/02/00	0.73	(0.38-1.0)	0.79	97

\* All  $\kappa$  values are significant at the 5% level.

Table 4. Test-retest reliability of SIDAM syndrome scores (Spearman rank correlations)

Orientation (OR)	0.83
Immediate recall (RE)	0.47
Short-term memory (SM)	0.81
Long-term memory (LM)	0.86
Memory (ME)	0.89
Intellectual abilities (IN)	0.72
Verbal abilities/calculation (VC)	0.59
Constructional abilities (VS)	0.76
Aphasia/Apraxia (AA)	0.60
Higher cortical functions (HI)	0.77

independently. Test-retest reliability of all SIDAM syndrome scores was computed by means of Spearman rank correlations ( $r$ ). In general acceptable  $r$  values were found: 0.83 for 'orientation'; 0.72 for 'intellectual abilities';

0.77 for 'higher cortical functions'; and 0.89 for 'memory' (Table 4).

The 'memory syndrome' score ranges from 0 (severe memory impairment) to 20 (no impairment) and can be divided in three subscores ('immediate recall', 'short-term memory' and 'long-term memory'). Reliability of the 'immediate recall' syndrome was lowest with  $r = 0.47$ . Reliability was good for 'short-term memory' ( $r = 0.81$ ) and 'long-term memory' with  $r = 0.86$ .

Similarly the 'higher cortical function' syndrome is subdivided in three subsyndromes, each with different scores. Reliability of the three subscores was acceptable with  $r$  ranging from 0.59 ('verbal abilities/calculation') to 0.60 ('aphasia/apraxia') and 0.76 for 'constructional abilities'.

### Test-retest reliability of various scores included in the SIDAM

The SISCO is the overall score of the test performance part of the SIDAM with a maximum score of 55 (no cognitive impairment) and a minimum score of 0 (worst cognitive impairment). A high test-retest reliability of 0.97 (Pearson correlation coefficient) was found, for the MMSE 0.96 and for the EMMS 0.95 respectively. For the Rosen score (modified ischemic score), a reliability of  $r = 0.80$  (Spearman rank correlation) was calculated.

### Relationship between ICD-9 consensus diagnoses and SIDAM DSM-III-R and SIDAM ICD-10 diagnoses

Excellent concordance was demonstrated between ICD-9 consensus diagnosis and SIDAM DSM-III-R diagnoses of dementia with a  $\kappa$  of 0.92, and 0.86 with regard to SIDAM-ICD-10 diagnoses of dementia. In a more detailed analysis pertaining to subgroups (SDAT, MID)  $\kappa$  values between 0.64 and 0.95 were calculated. With respect to ICD-9 consensus diagnoses, in 44 cases no diagnosis of dementia was given. In only one such case (out of 44), the test-interviewer (but not the retest-interviewer) arrived at a SIDAM-DSM-III-R (but not at a SIDAM ICD-10) diagnosis of dementia; ICD-9 consensus diagnosis was 'cognitive change of other type' (ICD-9: 310.1). These results show that false positive diagnoses are very unlikely when administering the SIDAM.

## DISCUSSION

The SIDAM was conceived as a brief diagnostic instrument for the purpose of diagnosing dementia syndromes according to the criteria of DSM-III-R and the draft of the ICD-10 research criteria. The primary aim of this study was to describe and evaluate our instrument in a test-retest design.

For ICD-10 SIDAM-diagnoses,  $\kappa$  values between 0.64 and 0.95 have been demonstrated and similar results were found for DSM-III-R SIDAM-diagnoses with  $\kappa$  values between 0.64 (multi-infarct dementia) and 0.87 (all dementias). The relatively lower  $\kappa$  value for multi-infarct dementia was considerably influenced by the low base rate of this disorder in our study. A

higher  $\gamma$  value indicates that better  $\kappa$  values might be expected if a greater proportion of patients with this disorder could be included.

Diagnostic reliability of SIDAM DSM-III-R/ICD-10 diagnoses of dementia was found to be generally higher than that reported for the diagnosis of organic mental disorders (OMD), established without specific criteria (e.g. ICD-8/9): Spitzer & Fleiss (1974) obtained lower figures in a re-analysis of several reliability studies (most of them were inter-rater reliability studies) with a mean kappa for OMD of 0.77. In the DSM-III Field Trial (Spitzer *et al.* 1979) a  $\kappa$  value of 0.74 (inter-rater reliability) was found for the diagnosis of dementia (DSM-III). If this is compared with the results presented here, the present findings show that the use of a structured clinical instrument like the SIDAM can enhance the accuracy of diagnosing dementia. Using the GMS, Copeland *et al.* (1976) and Henderson *et al.* (1983) previously achieved a mean kappa value of 0.45 and an average phi value of 0.56 (test-retest interval: 11.7 days), respectively, for the test-retest reliability of the GMS.

Concordance rates on the criterion level were demonstrated to have excellent  $\kappa$  values for SIDAM DSM-III-R-criteria ( $\kappa = 0.79-0.87$ ) and SIDAM ICD-10 criteria as well with one exception: ICD-10 criterion A<sub>1</sub> (memory impairment) was found to have just a chance agreement with a  $\kappa$  of  $-0.05$ ! This result is due to our intentionally very broad definition of criterion A<sub>1</sub>; since it is not specified in ICD-10 how many questions are required to be scored as 'error' in order to meet A<sub>1</sub>, we defined one error (out of 18 questions) as sufficient for fulfilling criterion A<sub>1</sub>. Thus, every subject met criterion A<sub>1</sub> with six discrepancies between raters. For SIDAM DSM-III-R criterion A (memory impairment) at least two questions (out of 18 questions) scored as 'error' were required (one related to short-term memory *and* one related to long-term memory); here a much better concordance with a  $\kappa = 0.79$  was achieved. This comparison illustrates the need for stringent, not too sensitive cut-off points and a better operationalization of criteria, which is discussed in detail by Jorm & Henderson (1985), and Mowry & Burvill (1988).

Considering test-retest reliability on item level, we found in 74% significant kappa values of 0.4 and more, in 28% a kappa of more than

0.7. Some lower  $\kappa$  values were explained by symptom fluctuation and in a few cases by insufficient symptom description (criterion variance). These items will be reviewed and definitions and guide notes may be improved. In general, our results may indicate lower  $\kappa$  values for some of those items representing controlled information processing (Jorm, 1986), which requires the attentional resources of the individual (e.g. serial sevens), in contrast to items representing automatic processing (e.g. long-term memory).

It is entirely plausible that subjects with cognitive decline have problems with attention and concentration and therefore will demonstrate daily fluctuations in their performance. Some of the lower  $\kappa$  values may thereby be explained. The mean  $\kappa$  values obtained in the cognitive section are comparable to the results reported by Henderson *et al.* (1983) for GMS cognitive items e.g. the mean  $\kappa$  value of ten cognitive SIDAM items was 0.70, the average phi value for 11 cognitive items in the GMS was 0.60. A high mean kappa (8 items) of 0.64 was found for the modified ischemic score (Rosen *et al.* 1980); the Spearman rank correlation was 0.80; the SIDAM-MMSE proved to be reliable with a mean kappa of 0.58 (19 items = 30 questions) and a Pearson correlation coefficient of 0.96; Folstein *et al.* (1975) found a 24 h test-retest reliability of the MMSE of  $r = 0.83$  (Pearson coefficient) in patients thought to be cognitively stable. Since many of our patients demonstrated fluctuation of their cognitive performance our results are not *a priori* comparable.

The SIDAM syndrome scores and the SISCO allow a quantified assessment of cognitive state and specific cognitive functions. They demonstrated satisfying reliability. A Pearson  $r$  of 0.98 was computed for the SISCO. It was encouraging that diagnostic reliability of the SIDAM was high and almost no false positive diagnoses were found as compared to ICD-9 consensus diagnoses.

Using clinical judgement (ICD-9 consensus diagnoses) as a valid yardstick, high concordance rates were found with SIDAM DSM-III-R dementias ( $\kappa = 0.92$ ) and SIDAM ICD-10 dementias ( $\kappa = 0.86$ ); however, these results may be biased to some extent since the participating psychiatrists were engaged in testing and two

had also been responsible for construction of the SIDAM. This kind of bias has been discussed by Gurland (1981), Copeland *et al.* (1988) and Jorm *et al.* (1988). Therefore, we have used additional scales like the CDR and GDS, since predictive validity requires some criteria against which SIDAM can be judged; in general SIDAM diagnoses correlated well with these scales (Zaudig *et al.* 1990; Zaudig & Hiller, 1990). Finally, it has to be considered that the SIDAM as a brief diagnostic instrument cannot be expected to replace a comprehensive diagnostic instrument or a complete clinical evaluation since it does not include the necessary information to exclude other psychiatric disorders. Accurate diagnosis always depends on medical/psychiatric history, full mental status examination, physical status and laboratory tests.

## CONCLUSION

The SIDAM provides reliable quantification and diagnoses of cognitive impairment and dementia. Each section and syndrome of the SIDAM has been found to have an acceptably high level of test-retest reliability, which equally holds true on the diagnostic, criterion and item level. It reliably separates subjects with DSM-III-R/ICD-10 dementia from those without such a disorder. SIDAM diagnoses of dementia and cognitive impairment correlate well with clinical diagnoses and demonstrate high sensitivity and specificity (Zaudig *et al.* 1990). SIDAM has proved to be of value as a brief (28 min), practical and easily scored instrument in clinical routine and research, and seems very useful as a case-finding instrument for dementias. It was found to be acceptable to patients and interviewers. For most of the DSM-III-R and ICD-10-criteria, a set of specific questions was implemented and cut-off points were defined successfully. The SISCO and SIDAM syndrome scores were shown to be reliable and useful in quantitatively assessing cognitive status and specific functions as well as severity of cognitive impairment.

Normative data (study population of 150 subjects) of the SIDAM-score (SISCO) sensitivity, specificity and predictive validity are available and will be published soon (Zaudig *et al.* 1990; Zaudig & Hiller, 1990). It is intended to investigate the validity of the SIDAM, in-

cluding long-term follow-up studies and correlation with pathological and biochemical measures.

Thanks are due to Professor J. Helzer and Dr R. Fischbach (Washington University, St Louis, USA) for critical comments and helpful discussions on an earlier draft of the SIDAM.

The SIDAM is available in German, English, Spanish and several other languages (e.g. Japanese); enquiries should be addressed to M. Zaudig.

## REFERENCES

- American Psychiatric Association (1980). *Diagnostic and Statistical Manual of Mental Disorders, (third edn)*. American Psychiatric Association: Washington, DC.
- American Psychiatric Association (1987). *Diagnostic and Statistical Manual of Mental Disorders, (third edn) Revised*. American Psychiatric Association: Washington, DC.
- Anthony, J. C., Le Resche, L. A., Niaz, U., von Korff, N. R. & Folstein, M. F. (1982). Limits of the Mini-Mental State as a screening test for dementia and delirium among hospital patients. *Psychological Medicine* **12**, 397-408.
- Bartko, J. J. & Carpenter, W. T. (1976). On the methods and theory of reliability. *Journal of Nervous and Mental Diseases* **163**, 307-317.
- Blessed, G., Tomlinson, B. E. & Roth, M. (1968). The association between quantitative measures of dementia and of senile change in the cerebral grey matter of elderly subjects. *British Journal of Psychiatry* **114**, 797-811.
- Burke, J. C. (1986). Diagnostic categorization by the Diagnostic Interview Schedule (DIS): a comparison with other methods of assessment. In *Mental Disorders and the Community* (ed. J. Berrett and R. M. Rose), pp. 255-285. Guilford Press: New York.
- Burke, J. C. (1988). Diagnostic assessment. In *An Annotated Bibliography of DSM-III*, pp. 13-22. American Psychiatric Press: Washington, DC.
- Cicchetti, D. H. & Prusoff, B. A. (1983). Reliability of depression and associated clinical symptoms. *Archives of General Psychiatry* **40**, 987-990.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational Psychological Measurement* **20**, 37-46.
- Copeland, J. R. M., Kelleher, N. J., Kellett, J. N., Gourlay, A. J., Gurland, B. J., Fleiss, J. E. & Sharpe, L. (1976). A semistructured clinical interview for the assessment of diagnosis and mental state in the elderly: the Geriatric Mental State Schedule I. Development and reliability. *Psychological Medicine* **6**, 439-449.
- Copeland, J. R. M., Dewey, N. E., Henderson, A. S., Kay, D. W. K., Neal, C. D., Harrison, N. A. M., McWilliam, C., Forshaw, D. & Shiwach, R. (1988). The Geriatric Mental State (GMS) used in the community: replication studies on the computerized diagnoses AGE-CAT. *Psychological Medicine* **18**, 219-223.
- Cummings, J. L. & Benson, B. F. (1986). Dementia of the Alzheimer type. An inventory of diagnostic clinical features. *Journal of the American Geriatrics Society* **34**, 12-19.
- Farmer, A. & Helzer, J. A. (1990). Improving the sensitivity of the Mini-Mental State Examination. A test in a large sample of alcoholics. (In the press.)
- Fleiss, J. L. (1981). *Statistical Methods for Rates and Proportions*, edn. 2. John Wiley: New York.
- Folstein, N. F., Folstein, S. E. & McHugh, P. R. (1975). 'Mini-Mental State'. A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research* **12**, 189-198.
- Gurland, B. J. (1981). The borderlines of dementia: the influence of sociocultural characteristics on rates of dementia appearing in the senium. In *Clinical Aspects of Alzheimer's Disease and Senile Dementia (Aging, Vol. 15)* (ed. N. E. Miller and E. D. Cohen), pp. 61-84. Raven Press: New York.
- Gurland, B. J., Kuriansky, J., Sharpe, L., Simon, R., Stiller, P. & Birkett, P. (1977). The comprehensive assessment and referral evaluation (CARE). Rationale, development and reliability. *International Journal of Ageing and Human Development* **8**, 9-42.
- Gurland, B. J., Dean, L., Copeland, J., Gurland, B. & Golden, R. (1982). Criteria for diagnosis of dementia in the community elderly. *Gerontologist* **22**, 180-186.
- Gurland, B. J., Copeland, J., Kelleher, J., Kuriansky, J., Sharpe, L. & Lee Dean, L. (1983). *The Mind and Mood of Ageing: Mental Health Problems of the Elderly Community in New York and London*. Haworth Press: New York. Croom Helm: London.
- Hachinski, V. V., Iliff, L. D., Zihlka, E., Du Boulay, G. H., McAllister, V. L., Marshall, J., Ross Russell, R. W. & Symon, L. (1975). Cerebral blood flow in dementia. *Archives of Neurology* **32**, 632-637.
- Haycox, J. A. (1984). A behavioural scale for dementia. In *Biology and Treatment of Dementia in the Elderly* (ed. P. A. Shamoian), pp. 2-13. American Psychiatric Press: Washington DC.
- Henderson, A. S. & Huppert, F. A. (1984). The problem of mild dementia. *Psychological Medicine* **14**, 5-11.
- Henderson, A. S., Duncan-Jones, D. & Finlay-Jones, R. A. (1983). The reliability of the Geriatric Mental State Examination. *Acta Psychiatrica Scandinavica* **67**, 281-289.
- Hiller, W., v. Zerssen, D., Mombour, W. & Wittchen, H.-U. (1986). *Die Inpatient Multidimensional Psychiatric Scale (IMPS)*. Beltz: Weinheim.
- Hughes, C. P., Berg, L., Danziger, W. L., Coben, L. A. & Martin, R. L. (1982). A new clinical scale for the staging of dementia. *British Journal of Psychiatry* **140**, 566-572.
- Jorm, A. S. (1986). Controlled and automatic information processing in senile dementia: a review. *Psychological Medicine* **16**, 77-88.
- Jorm, A. S. & Henderson, A. S. (1985). Possible improvements to the diagnostic criteria for dementia in DSM-III. *British Journal of Psychiatry* **147**, 304-399.
- Jorm, A. S., Scott, R., Henderson, A. S. & Kay, D. W. K. (1988). Educational level and differences on the Mini-Mental State: the role of test bias. *Psychological Medicine* **18**, 727-731.
- Lorr, N. & Klett, C. J. (1967). *Manual for the Inpatient Multidimensional Psychiatric Scale* (rev.) Consulting Psychologists Press: Palo Alto, Ca.
- Mombour, W. (1974). Syndrome bei psychiatrischen Erkrankungen. eine vergleichende Untersuchung mit zwei Schätzskalen (IMPS und AMDP-Skala). *Archiv für Psychiatrie und Nervenkrankheiten* **219**, 331-350.
- Mombour, W., Gammel, G., v. Zerssen, D. & Heyse, H., (1973). Die Objektivierung psychiatrischer Syndrome durch multifaktorielle Analyse des psychopathologischen Befundes. *Nervenarzt* **44**, 352-358.
- Mowry, D. J. & Burvill, P. W. (1988). A study of mild dementia in the Community using a wide range of diagnostic criteria. *British Journal of Psychiatry* **153**, 328-334.
- Pfeiffer, E. (1975). A Short Portable Mental Status Questionnaire (SPMSQ) for the assessment of organic brain deficit in elderly patients. *Journal of the American Geriatrics Society* **23**, 133-441.
- Reisberg, B., Ferris, S. H., De Leon, M. J. & Crook, T. (1982). The Global Deterioration Scale (GDS) for assessment of primary degenerative dementia. *American Journal of Psychiatry* **139**, 1136-1139.
- Robins, L. N. (1989). Diagnostic grammar and assessment: translating criteria into questions. *Psychological Medicine* **19**, 57-68.
- Robins, L. N., Helzer, J. E., Croughan, J. & Ratcliffe, K. S. (1981). National Institute of Mental Health Diagnostic Interview Schedule: its history, characteristics and validity. *Archives of General Psychiatry* **38**, 381-389.
- Rosen, W. G., Terry, R. D., Fuld, P. A., Katzman, R. & Peck, A. (1980). Pathological verification of ischemic score in differentiation of dementias. *Annals of Neurology* **7**, 486-488.

- Roth, M., Tym, E., Mountjoy, C. Q., Huppert, F. A., Hendrie, H., Verma, S. & Goddard, R. (1986). CAMDEX. A standardized instrument for the diagnosis of mental disorder in the elderly with special reference to the early detection of dementia. *British Journal of Psychiatry* **149**, 698-709.
- Shrout, P. E., Spitzer, R. L. & Fleiss, J. L. (1987). Quantification of agreement in psychiatric diagnoses revisited. *Archives of General Psychiatry* **44**, 172-177.
- Spitzer, R. L. & Fleiss, J. L. (1974). A re-analysis of the reliability of psychiatric diagnoses. *British Journal of Psychiatry* **125**, 341-347.
- Spitzer, R. L. & Williams, J. B. W. (1985). Classification in Psychiatry. In *Comprehensive Textbook of Psychiatry*, edn. 4. (ed. H. Kaplan and B. Saddock), pp. 591-613. Williams and Wilkins: Baltimore.
- Spitzer, R. L., Forman, J. B. W. & Nee, J. (1979). DSM-III Field Trials. I. Initial interrater diagnostic reliability. *American Journal of Psychiatry* **136**, 815-817.
- Spitzer, R. L., Williams, J. B. W. & Gibbon, M. (1986). *Structured Clinical Interview for DSM-III-R (SCID)*. Biometrics Research Department, New York State Psychiatric Institute: New York.
- Spitznagel, E. L. & Helzer, J. E. (1985). A proposed solution to the base rate problem in the kappa-statistic. *Archives of General Psychiatry* **42**, 725-728.
- Wechsler, D. (1945). A standardized memory scale for clinical use. *Journal of Psychology* **19**, 87-95.
- Wechsler, D. (1954). Measurement and evaluation of intelligence of older persons in Old Age in the Modern World. *Report of the 3rd Congress, International Association of Gerontology*, pp. 275-278. Livingstone: Edinburgh.
- Wechsler, D. (1958). *The Measurement and Appraisal of Adult Intelligence*. Williams and Wilkins: Baltimore, MD.
- Wells, C. E. (1982). Refinements in the diagnosis of dementia. *American Journal of Psychiatry* **139**, 621-622.
- Williams, J. B. W. (1988). A structured interview guide for the Hamilton Depression Rating Scale. *Archives of General Psychiatry* **45**, 742-747.
- Wing, J. K., Cooper, J. E. & Sartorius N. (1974). *The Measurement and Classification of Psychiatric Symptoms*. Cambridge University Press: London.
- Wittchen, H.-U., Zaudig, M., Schramm, E., Spengler, P., Mombour, W., Klug, J. & Horn, R. (1987). *Strukturiertes klinisches Interview für DSM-III-R (SCID) (Testversion)*. Beltz: Weinheim.
- World Health Organization (1979) *ICD 9: Mental Disorders. Glossary and Guide to Their Classification in Accordance with Ninth Revision of the International Classification of Diseases*. WHO: Geneva.
- World Health Organization (1987). *ICD-10: Mental, Behavioural and Developmental Disorders: Clinical descriptions and diagnostic guidelines (1986 Draft for Field Trials); Research Diagnostic Criteria*. WHO: Geneva.
- World Health Organization (1988) *ICD-10: Mental, Behavioural and Developmental Disorders: Clinical descriptions and diagnostic guidelines (1987 Draft for Field Trials)*. WHO: Geneva.
- Yule, G. I. (1912). On the methods of measuring association between two attributes. *Journal of the Royal Statistical Society* **75**, 581-642.
- Zaudig, M. & Hiller, W. (1990). On diagnosing dementia. Comparison, evaluation and validity of ICD-10 and DSM-III-R criteria of dementia. (In the press.)
- Zaudig, M., Mittelhammer, J. & Hiller, W. (1989). *Structured Interview for the Diagnosis of Dementia of the Alzheimer Type, Multi-Infarct-Dementia and Dementia of other Etiology according to ICD-10 and DSM-III-R - SIDAM*. Logomed: Munich.
- Zaudig, M., Mittelhammer, J. & Hiller, W. (1990). *Strukturiertes Interview zur Diagnose der Demenz vom Alzheimer Type, der Multiinfarktdemenz und Demenzen anderer Ätiologie nach DSM-III-R und ICD-10. Manual*. Logomed: München.