

Interrater agreement for the schedule for affective disorders and schizophrenia epidemiological version for school-age children (K-SADS-E)

Concordância entre observadores da entrevista semi-estruturada para diagnóstico em psiquiatria da infância, versão epidemiológica (K-SADS-E)

Guilherme V Polanczyk, Mariana Eizirik, Victor Aranovich, Daniel Denardin, Tatiana L da Silva, Tatiana V da Conceição, Thiago G Pianca and Luis Augusto Rohde

ADHD Outpatient Clinic at Hospital de Clínicas de Porto Alegre, Department of Psychiatry, Federal University of Rio Grande do Sul. Porto Alegre, RS, Brazil

Abstract **Objective:** The main objective of this study was to assess the interrater agreement for the Schedule for Affective Disorders and Schizophrenia Epidemiological version for School-Age Children (K-SADS-E).

Methods: Four interviewers being trained with the K-SADS-E scored independently 29 videotaped interviews performed with psychiatric outpatients in the ADHD Outpatient Clinic at Hospital de Clínicas de Porto Alegre. Interrater agreement analysis was performed using the kappa coefficient (k).

Results: Kappa coefficients were .93 ($p < .001$) for affective disorders, .9 ($p < .001$) for anxiety disorders, .94 ($p < .001$) for attention-deficit/hyperactivity disorders and disruptive behavior disorders.

Conclusion: These findings suggest an excellent interrater agreement for the diagnosis of several mental disorders in childhood and adolescence by the Brazilian Portuguese version of the K-SADS-E.

Keywords K-SADS-E. Diagnosis. Diagnostic interview. Reliability. Agreement.

Resumo **Objetivos:** Avaliar a concordância entre observadores da Entrevista Semi-Estruturada para Diagnóstico em Psiquiatria da Infância, versão epidemiológica (K-SADS-E).

Métodos: Quatro observadores em fase final de treinamento na aplicação do instrumento K-SADS-E pontuaram independentemente 29 entrevistas, registradas em vídeo-tape, de pacientes ambulatoriais do Programa de Transtorno de Déficit de Atenção/Hiperatividade do Hospital de Clínicas de Porto Alegre. Os resultados foram analisados pelo coeficiente de kappa (k).

Resultados: Os coeficientes k foram 0,93 ($p < 0,001$) para transtornos do humor, 0,9 ($p < 0,001$) para transtornos de ansiedade e 0,94 ($p < 0,001$) para transtornos de déficit de atenção/hiperatividade e do comportamento disruptivo.

Conclusão: Os resultados demonstram uma excelente concordância entre observadores na formulação diagnóstica dos vários transtornos da infância e adolescência, por meio da versão em português da K-SADS-E.

Descritores K-SADS-E. Diagnóstico. Entrevista diagnóstica. Confiabilidade. Concordância.

Introduction

Standardized instruments have been increasingly used in the psychiatric assessment of children and adolescents in clinical and research settings. There are many instruments with different characteristics which vary according to their purposes such as diagnostic formulation, identification of symptoms, on the assessment of personality and emotional dynamics or cognitive development.

Diagnostic interviews have an important role for the establishment of judicious diagnoses and for the assessment of epidemiological patterns of psychopathologies. Several instru-

ments in the field of child and adolescence psychiatry have this profile, such as the Schedule for Affective Disorders and Schizophrenia – epidemiological version (K-SADS-E), the Diagnostic Interview for Children and Adolescents (DICA) and the National Institute of Mental Health Diagnostic Interview Schedule for Children (NIMH-DISC). The K-SADS-E, which has a Brazilian Portuguese translation currently used for clinical and researching purposes in our country, is the most used version of the K-SADS interview.^{1,2} The DICA is also a semi-structured interview, mainly aimed at clinical and epidemiological research, based on the criteria of the International Clas-

sification of Diseases 9th Edition, World Health Organization (ICD-9, WHO). It has a version adapted for the DSM-III-R criteria and a self-reported version, with computer-based analyses.³ Another widely used interview is the DISC IV, DISC's most recent version, an interview conceived to be administered by raters without clinical training, making its application more feasible in large-scale epidemiological studies.^{2,3}

Currently the K-SADS-E is widely used in research, having excellent psychometric properties reported by international studies.² As it is a semi-structured interview what allows the interviewer's clinical judgement to punctuate the questions, its performance greatly depends on the interviewers' training. However, we have not found studies in Brazil assessing the psychometric properties of this instrument and, for that reason, the main objective of this study was to assess the interrater agreement in the derivation of child and adolescence psychiatric diagnoses by means of the Brazilian Portuguese version of the K-SADS-E.

Method

Subjects

The K-SADS-E was consecutively applied to 29 patients who were called for initial ambulatorial assessment at the ADHD Outpatient Clinic at Hospital de Clínicas de Porto Alegre. The interviews were all videotaped. The parents of the patients answered to the questions after giving oral informed consent for the registration of the interviews, which was also videotaped, after receiving the explanations about the objectives, being assured the total secrecy regarding the collected information.

Instrument

The K-SADS-E is a semi-structured interview for children and adolescents aged 6 to 18 years which assesses current episodes and the severest episode in the past (lifetime) of mood disorders (major depression, dysthymia and mania), anxiety disorders (separation anxiety, panic disorder, agoraphobia, social phobia, simple phobia, generalized anxiety disorder, obsessive-compulsive disorder, post-traumatic stress disorder), eating disorders (anorexia and bulimia), attention-deficit/hyperactivity disorder (ADHD), conduct disorder and oppositional defiant disorder, substance abuse and dependence, elimination disorders (enuresis and encopresis), speech disorder, Tourette's disorder, psychotic disorders and pervasive disorders. The fourth reviewed version of the interview, used in our Clinic and in this study, is based on the diagnostic criteria of the DSM-III-R and DSM-IV and has essential questions for the diagnosis, which, if not met, allow the interviewer to go to the following disorder. The interviewer codifies the symptoms as present or absent and graduates the impairment caused by the disorders as mild, moderate or severe. Its administration lasts for 50 to 90 minutes and its informants are parents and children or adolescents.⁴

Assessments

The symptomatology of the 29 patients was informed by their

parents and were punctuated using the K-SADS-E by three trained interviewers, being recorded on videotape. Each interview was discussed in a clinical committee, conducted by a child and adolescents psychiatrist with large clinical experience (L.A.R.), being thus generated the final diagnoses.

Four observers, medical students in their final stage of training in the application of the K-SADS-E, individually and independently punctuated each interview analyzing the videotapes. The observers were blind to the result of the initial interview and to all information about patients.

The training process consisted in four phases: 1st) seminars about the structure and diagnostic criteria of the instrument, conducted by a child and adolescents psychiatrist (L.A.R.), 2nd) live observation of 5 interviews performed by trained observers, 3rd) live administration of the K-SADS-E interview in 10 patients with the presence of trained observers and 4th) punctuation of interviews recorded in videotape and agreement analysis.

Statistical analysis

We performed the concordance analysis of diagnoses generated by each of the four observers using the kappa coefficient (k). The adopted analysis strategy was the combination of the four observers in pairs and the calculation of the k coefficient for each of the six possible pairs for each disorder and for current and past diagnosis. We have considered, thus, as the final k coefficient the arithmetic mean of the six combinations for each disorder and for the current and past diagnosis.

P values $<.05$ were deemed indication of statistical significance.

We used the criteria proposed by Landis & Kock⁵ (1977) to interpret the values of k coefficients: excellent agreement, $k >.75$; good agreement, k from $.59$ to $.75$; medium agreement, k between $.40$ and $.58$; poor agreement, $k <.40$. Disorders present in less than 10% of the sample were not considered for the analysis, due to the instability of the k coefficient when the number of observations is small.⁶⁻⁸

Results

We interviewed 29 patients aging 10.3 ± 3.4 years, being 69% males. Each patient had in average 5.6 diagnoses, with a minimum of 1 and a maximum of 13 diagnoses. All parents of requested patients gave their consent to the videotaping of the interview.

The k coefficient for past diagnosis of major depression was $.91$ and $.93$ for current diagnosis. Aggregated mood disorders had a k coefficient of $.93$ for past and current diagnoses. Aggregated anxiety disorders had k of $.93$ for past diagnosis and $.90$ for current diagnosis. K coefficients for past diagnosis of agoraphobia were $.80$ and $.79$ for current diagnosis, whereas for generalized anxiety disorder was 1, both for current and past diagnosis. Aggregated disruptive disorders and ADHD had a k coefficient of $.94$ for past and current diagnoses. Table shows k coefficients for each specific disorder and aggregation of disorders, separated according to current and past diagnoses.

Table - Mean of kappa coefficients (K) between observers by disorder.

	Past Diagnoses		Current Diagnoses	
	N	k*	N	k*
Mood disorders	21	.93	15	.93
Major disorder	8	.91	5	.93
Dysthymia	10	.96	7	.90
Bipolar mood disorder	3	.85	3	.85
Psychosis	7	.82	7	.82
Anxiety disorders	57	.93	47	.90
Separation anxiety	10	.84	6	.82
Agoraphobia	15	.80	13	.79
Social phobia	7	.95	7	.95
Simple phobia	14	.88	12	.81
Generalized anxiety disorder	5	.1	5	.1
Obsessive-compulsive disorder	4	.80	4	.80
Attention-deficit/Hyperactivity and disruptive behavior disorder	40	.94	39	.94
ADHD	21	.94	21	.94
Oppositional defiant disorder	12	.91	11	.91
Conduct disorder	7	.93	7	.93
Speech disorder	8	.85	4	.85
Elimination disorder	11	.94	7	.96
Enuresis	8	.1	5	.1
Encopresis	3	.83	2	-
Pervasive disorder	5	.94	5	.94

N= absolute number of diagnoses of each disorder in the sample; k= Kappa coefficient; ADHD= Attention-deficit/Hyperactivity disorder;

*k coefficients with $P < .001$.

Discussion

The values of k coefficients found for affective, anxiety, attention and disruptive behavior, elimination and developmental disorders indicate that the Brazilian Portuguese version of the K-SADS-E is an interview with optimal interrater agreement in the psychiatric diagnoses of many child and adolescence disorders.

Variabilities occur very frequently in diagnostic processes, mainly related to disorders involving behaviors or emotions in which cultural and personal features have great influence. The utilization of objective diagnostic criteria would lessen this variability. However, Winokur⁹ reports that their use not necessarily produces diagnoses with consistently high reliability as different raters interpret them differently. Studies which assess the diagnostic agreement between raters, therefore, try to minimize the variabilities that occur in the diagnostic process in order to maximize the replicability of diagnoses and the discrimination of patients by different raters. As the interviewers in our Clinic are medical students, their adequate training and the assessment of their performance by means of properties such as the interrater agreement, is fundamental.

In psychiatric services in which the therapeutic decisions are based on the results of semi-structured interviews such as the K-SADS-E,¹⁰ optimal psychometric features of these instruments are essential. However, similarly to other authors,¹¹⁻¹³ we believe that the K-SADS-E be better clinically used as part of a battery of assessments involving self-reported instruments with acknowledged validity, a semi-structured diagnostic interview, instruments for the psychological and clinical assessments, integrating the information supplied by patients, their parents and teachers. The finding of a patient with thirteen psychiatric disorders indicates a limitation of semi-structured instruments which derive descriptive diagnoses. When we find patients with severe mental disorders, with diverse symptomatology and significant impairment in different aspects of their mental functioning (such as in the pervasive disorders or in severe mental retardation) the

instrument has positive values for several diagnoses, as there is no diagnostic hierarchy, such as occurred with the mentioned patient, who had a pervasive disorder.

A previous study which assessed the interrater agreement for the K-SADS-E, using audio material showed k coefficients of .73 for major depression, .65 for separation anxiety, .75 for phobic disorders, .51 for oppositional defiant disorder, .77 for ADHD and .68 for conduct disorder.⁴

The k values found were higher than those reported in the international literature, what probably was due to the extensive training process to which the observers are submitted in our Clinic, as recommended by the authors of the instrument when the interviewers are not psychiatrists or psychologists.¹⁴ Perhaps such an agreement would not occur if the application of the interview were performed by examiners in the usual context. We also think that the visualization of the interviews by means of videotapes could help in the adequate interpretation of the questions.

Our results should be interpreted in the context of some methodological limitations. The studied sample was small and it was not possible to analyze the k coefficient for some diagnoses that had frequencies lower than 10% in the sample. Only one strategy for reliability assessment was used in this study, although studies which use the test-retest method,^{11,15,16} considered as the most consistent method for reliability analysis,⁴ have reported similarly high values k coefficients. Our results cannot be generalized to samples of the general population, as existent data clearly indicate that the reliability regarding semi-structured interviews tend to be higher in clinical settings, with patients who are really ill, than in population samples in which children are healthier.^{16,17}

Conclusion

The results of this study show optimal interrater agreement in the diagnostic formulation of child and adolescence psychiatric disorders using the Brazilian Portuguese version of the K-SADS-E.

References

1. Mercadante MT, Asbarh F, Rosário MC, Ayres AM, Ferrari MC, Assumpção FB, Miguel EC. K-SADS, entrevista semi-estruturada para diagnóstico em psiquiatria da infância, versão epidemiológica. 1º ed. São Paulo: PROTOC – Hospital das Clínicas da FMUSP; 1995.
2. American Psychiatric Association. Task force for the handbook of psychiatric measures. Handbook of psychiatric measures. Washington (DC): American Psychiatric Association; 2000.
3. Duarte CS, Bordin IAS. Instrumentos de avaliação. Rev Bras Psiquiatr 2000;22(Supl 2):55-8.
4. Ambrosini PJ. Historical development and present status of the schedule for affective disorders and schizophrenia for school-age children (K-SADS). J Am Acad Child Adolesc Psychiatry 2000;39(1):49-58.
5. Landis J, Koch G. The measurement of observer agreement for categorical data. Biometrics 1977;33:159-74.
6. Del-Ben CM, Vilela JA, Crippa JA, Hallak JE, Labate CM, Zuardi AW. Confiabilidade da entrevista clínica estruturada para o DSM-IV – Versão clínica traduzida para o português. Rev Bras Psiquiatr 2001;23(5):156-9.
7. Grove WM, Andreasen NC, McDonald-Scott P, Keller MB, Shapiro RW. Reliability studies of psychiatric diagnosis: theory and practice. Arch Gen Psychiatry 1981;38:408-13.
8. Shrout PE, Spitzer RL, Fleiss JL. Quantification of agreement in psychiatric diagnosis revisited. Arch Gen Psychiatry 1987;44:172-7.
9. Winokur G, Zimmerman M, Cadoret R. 'Cause the Bible tells me so. Arch Gen Psychiatry 1988;45:683-4.
10. Ambrosini PJ, Metz C, Prabucki K, Lee JC. Videotape reliability of the third revised edition of the K-SADS. J Am Acad Child Adolesc Psychiatry 1989;28(5):723-8.
11. Kaufman J, Birmaher B, Brent D, Rao U, Flynn C, Moreci P, et al. Schedule for affective disorders and schizophrenia for school-age children – present and lifetime version (K-SADS-PL): initial reliability and validity data. J Am Acad Child Adolesc Psychiatry 1997;36(7):980-8.
12. Leckman J, Sholomskas D, Thompson D, Belanger A, Weissman M. Best estimate of lifetime psychiatric diagnosis: a methodological study. Arch Gen Psychiatry 1982;39:879-83.
13. Young JG, O'Brien JD, Gutterman EM, Cohen P. Research on the clinical interview. J Am Acad Child Adolesc Psychiatry 1987;26(5):613-20.
14. Endicott J, Spitzer RL. A diagnostic interview. The schedule for affective disorders and schizophrenia. Arch Gen Psychiatry 1978;35:837-44.
15. Chambers WJ, Puig-Antich J, Hirsch M, Paez P, Ambrosini PJ, Tabrizi MA, et al. The assessment of affective disorders in children and adolescents by semistructured interview. Arch Gen Psychiatry 1985;42:696-702.
16. Jensen P, Roper M, Fisher P, Piacentini J, Canino G, Richters J, et al. Test-retest reliability of the diagnostic interview schedule for children (DISC 2.1): parent, child and combinations algorithms. Arch Gen Psychiatry 1995;52:61-71.
17. Boyle MH, Offord DR, Racine Y, Sanford M, Szatmari P, Fleming JE, et al. Evaluation of the diagnostic interview for children and adolescent for use in general population samples. J Abnorm Psychol 1993;21:663-71.

Correspondence:

Luis Augusto Rohde
Serviço de Psiquiatria da Infância e Adolescência, Hospital de
Clínicas de Porto Alegre
Rua Ramiro Barcelos, 2350
90035-003 Porto Alegre, RS, Brazil
E-mail: lrohde@terra.com.br
