Psychological Medicine, 2005, **35**, 771–782. © 2004 Cambridge University Press
doi:10.1017/S0033291704003460 Printed in the United Kingdom

REVIEW ARTICLE

Tests of executive functions in first-degree relatives of schizophrenic patients: a meta-analysis

ANDREI SZÖKE[1,2]*, FRANCK SCHÜRHOFF[1,2], FLAVIE MATHIEU[2], ALEXANDRE MEARY[1,2], SERBAN IONESCU[3] AND MARION LEBOYER[1,2]

[1] *Service de Psychiatrie Adulte, Hôpital Albert Chenevier et Henri Mondor (Assistance Publique – Hôpitaux de Paris), 94000 Créteil, France*; [2] *Unité INSERM U 513, 'Neurobiologie et Psychiatrie', Hôpital Henri Mondor, 94000 Créteil, France*; [3] *Faculté de Psychologie, Université 'René Descartes', Paris 5, 75270 Paris, France*

ABSTRACT

Background. Executive dysfunctions in relatives of schizophrenic patients may be trait markers of genetic liability and thus help us to elucidate the aetiology of schizophrenia. As a large amount of data has been published, a synthesis through a meta-analysis was needed to demonstrate the existence of executive impairments in relatives of schizophrenic patients and to assess their magnitude.

Method. We conducted a meta-analysis of articles that compared performances of controls and relatives of schizophrenic patients on the four tests most frequently used to assess executive functions: the Wisconsin Card Sorting Test (WCST), the Trail Making Test (TMT), the Stroop Test and the Verbal Fluency (VF) Test. When needed and possible, published data were supplemented with information from the authors. After assessing the homogeneity of the data, effect sizes were estimated and publication bias was tested by use of funnel plots.

Results. Relatives of schizophrenic patients performed less well than controls on all executive tests analysed. Effect estimates were in the small to moderate range (from 0·26 to 0·49) for Stroop, WCST and TMT, but were greater for the fluency tests (0·65 for phonological and 0·87 for semantic VF).

Conclusion. Relatives of schizophrenic patients appear to have wide, although not severe, executive dysfunctions. As the sensitivity of the different tests for impairments in relatives is not the same, the choice of test and method used should be carefully assessed.

INTRODUCTION

Genetic epidemiology studies have provided strong evidence that genetic factors play an important role in the aetiology of schizophrenia. However, molecular genetics studies have failed to identify susceptibility genes reliably. Several authors (Tsuang *et al.* 1990; Leboyer *et al.* 1998) have argued that this situation is due to the clinical-phenotypical and aetiological-genetic complexity of schizophrenia. One of the ways proposed to overcome these difficulties (Gottesman & Shields, 1973; Leboyer *et al.* 1998; Freedman *et al.* 1999; Gottesman & Gould, 2003; Leboyer, 2003) is the use of endophenotypes, i.e. trait markers of genetic liability that are present not only in patients but also in subjects at genetic risk.

Cognitive deficits are putative endophenotypes because they have often been found in

* Address for correspondence: Dr Andrei Szöke, Service de Psychiatrie Adulte, Hôpital 'Albert Chenevier', 40 rue de Mesly, 94000 Créteil, France.
(Email: Andrei.szoke@ach.ap-hop-paris.fr)

Table 1. *Characteristics of studies selected for meta-analysis*

	Study (first author, year)	Diagnostic in probands[a]	Type of relatives[b]	Controls with psychotic first-degree relatives excluded	Exclusion criteria[c]		Differences in demographic variables[d]	WCST[e]		TMT[e]	VF[e]		Stroop[e]
					Relatives	Controls		PE[f]	C[g]		S[h]	L[i]	
1	Asarnow (2002)	S	P	No	1	1	Yes			**X**			
2	Chen (2000)	S	S	Yes	5	5	No	x	**X**		**X**		
3	Condray (1992)	S	S	Yes	1/2/5	5	No	x	x				
4	Dollfus (2002)	S	P	Yes	2	1	Yes	x	**X**		**X**	**X**	**X**
5	Egan (2001*a*)	S & SA	S	Yes	0	1	No		x	**X**	x	X	
6	Egan (2001*b*)	S & SA	S	Yes	0	1	No	x					
7	Faraone (2003)[j]	S	M	Yes	1	1	Yes			**X**[k]			x[k]
8	Franke (1992)	S	S/M	No	5	5	No	**X**	**X**				
9	Franke (1993)	S	S	No	5	5	No	**X**	x	**X**		**X**	
10	Goldberg (1995)	S & SA	T	No	2	5	No	x	x	x		x	x
11	Harris (1996)	S	P	No	0	0	No			X			
12	Keefe (1994)	S	M	Yes	1/2	4	Yes	x	x	**X**	**X**	**X**	
13	Keri (2001)	S	S	Yes	5	5	No	x	x			x	
14	Koren (1998)[j]	S & SA	M	Yes	1	1	No	**X**	x				
15	Krabbendam (2001)	S & SA	M	Yes	1	1	No				x		**X**[k]
16	Laurent (1999)	S	M	Yes	4	3	No	x	x	**X**	**X**	**X**	
17	Laurent (2000*a*)	S	M	Yes	3	5	No						x
18	Laurent (2000*b*)	S	M	Yes	3	3	No	x	x	**X**	**X**	**X**	
19	Mirsky (1995)	S	M	No	0/5	2	Yes		**X**				
20	Rybakowski (2002)	S	M	Yes	0/5[k]	5[k]	No	x	**X**	x			**X**
21	Saoud (2000)	S	S	Yes	4	4	Yes	x	**X**				
22	Scarone (1993)	S	S	No	5	5	Yes	x	x				
23	Stratta (1997)	S	M	Yes	2	2	Yes	x	x				
24	Wolf (2002)	S	C	No	1	0	Yes	**X**	**X**				
25	Yurgelun-Todd (1993)	S	S	No	1	5	Yes	x	x	x			

[a] S, schizophrenia; SA, schizoaffective disorder. [b] P, parents; S, siblings; C, children; T, twins; M, mixed. [c] Exclusion criteria [1, psychosis; 2, psychosis and schizotypal personality disorder (SPD); 3, all DSM Axis I diagnostics; 4, all DSM Axis I diagnostics and SPD (or cluster I personality disorders), 5, all DSM Axis I and II diagnostics]. [d] Demographic variables are age, gender, school grades or socio-economic status [Yes = difference present or unspecified; No = absent (or adjusted for)]. [e] Bold, capital letter ('**X**') for studies in which controls performed significantly better than relatives. [f] Perseverative errors. [g] Categories. [h] Semantic fluency. [i] Letter fluency. [j] Criteria from Faraone *et al.* (1995) but age not limited to 60 years. [k] Personal communication by the first author of the study.

schizophrenic patients (see Heinrichs & Zakzanis, 1998 and Johnson-Selfridge & Zalewski, 2001 for meta-analytical reviews), in subjects from high-risk samples which later developed psychotic disorders (Erlenmeyer-Kimling *et al.* 2000) and because their measures are reliable and stable over long periods of time (Rund, 1998). The strict demonstration of a specific deficit in schizophrenic patients is precluded by psychometric considerations (see Chapman & Chapman, 2001 for discussion). However, based on the extensive accumulated data (see e.g. Goldberg *et al.* 2003 for a synthesis on the subject), most authors consider that executive functions are, along with attention and memory, one of the most impaired cognitive domains. Although there is not a unique definition of executive functions current consensus regards them as processes whose primary purpose is to facilitate adaptation to novel situations by means of modulation and control of more fundamental or routine cognitive skills (Burgess, 1997).

For the reasons outlined above, many studies have been conducted to assess executive functions in relatives of schizophrenic patients and to determine whether executive dysfunctions can be used as endophenotypes. A large amount of data has accumulated with discrepant results (Table 1). This could be due to the small sample sizes which could lead to false-negative results (because of limited statistical power) or to false-positive results due to sampling variation (coupled with a publication bias towards positive results). To assess the existence and magnitude of executive impairments in relatives of

https://doi.org/10.1017/S0033291704003460 Published online by Cambridge University Press

schizophrenic patients, using all available data, and to help us understand the causes of differences in effect-size estimates (use of different tests of executive functions, differences in inclusion criteria, etc.) a synthesis through meta-analysis was needed.

Several authors (Schröder *et al.* 2002; Sitskoorn *et al.* 2003; Snitz *et al.* 2003) have recently published meta-analyses of studies that assessed cognitive functions in relatives of schizophrenic patients. However, these studies have been published only as abstracts meaning that little information is available regarding the complete articles, the meta-analytical procedures used and the neurocognitive tests analysed. As both article selection and meta-analytical procedures may influence results, it is difficult to judge the extent to which these findings can be generalized in the absence of this information. Furthermore, mean effect estimates are only available for wide cognitive domains (executive, attention, memory, etc.), which is a major limitation for two reasons. First, as there is no general agreement concerning the correspondence between specific neurocognitive tests and the cognitive domains explored, it is not clear which tests were used to evaluate differences between relatives and controls for different cognitive domains. Second, sensitivity may vary considerably even between tests thought to explore the same cognitive functions. Thus, global estimates do not offer clues about which tests are most sensitive to the impairments exhibited by the relatives of schizophrenic patients.

In this context, we decided to carry out a meta-analytical review of studies that assessed executive functions in relatives of schizophrenic patients, with detailed methodological considerations, and report the effect estimates for each test.

As pointed by Goldberg *et al.* (2003) executive functions involve use of the information rather than fundamental processing of information. As such, executive performance depends on 'lower processes' which provide input (perception, memory, etc.) or express the output (motor ability). This is reflected in 'task impurity' (Rabbit, 1997) and explains the existence of controversies on which tasks are (mainly) executive or not. As our main goal in this paper was to summarize current research rather than express a theoretical point of view, we decided to select the tests to be included in our meta-analysis empirically, as the tests most often used in articles dealing with executive function assessment.

As the results of different executive tasks usually show little correlation (Rabbit, 1997), we chose to analyse tasks separately and not to include articles that reported only composite scores from multiple tasks.

The main purpose of our study was to find out whether, taken as a whole, published studies of executive functions provide reliable evidence for impairments in relatives of schizophrenic patients.

Our study also created the opportunity to answer several related questions:

(*a*) What tests are the most commonly used in the medical literature to evaluate executive functions?

(*b*) Are the results of different studies comparing relatives of schizophrenic patients and controls homogeneous? If not, which study characteristics can explain the heterogeneity?

(*c*) Do the different tests of executive functions reveal differences of a similar magnitude between relatives and controls? If not, which tests are the most sensitive?

METHOD

Literature search

First, we performed a literature search to identify the cognitive tasks most often used in recent studies of executive functions. To do this, we analysed abstracts of medical articles concerned with the assessment of executive functions, published during the last 5 years. Articles were found by performing a literature search in Medline for:

(test OR **assessment)** AND **(executive)**
in title or abstract,
limited to: Human, Adult and Publication Date from 1998 to 2003.

After the selection of articles, available abstracts were searched for the tests used. The number of studies in which each test was used was calculated.

Once the most frequently used tests had been identified, we searched for articles relevant to

https://doi.org/10.1017/S0033291704003460 Published online by Cambridge University Press

our meta-analysis. To do this, we used three complementary strategies.

First, we searched the Medline database for:

(schizo* OR **psychotic)** AND **(relatives** OR **children** OR **parents** OR **sib*)** AND (**executive**)
limited to: Human, Adult and Publication Date from 1978 to 2003.

The '*' symbol stands for all possible endings (for example schizo* means that the database is searched for schizophrenia, schizophrenic, schizoaffective, schizophreniform, etc.). The last term (executive) was replaced successively by the names of the tests selected in the previous step and all their usual abbreviations (for example for the Trail Making Test, 'trail' and 'TMT' were also used). The year 1978 was chosen as the date limit because this is the date when the research diagnostic criteria (RDC) were published (Spitzer *et al*. 1978).

Second, to avoid omissions due to articles not yet indexed on Medline, a manual search was done, for the last 2 years, in the seven journals judged to be the most relevant for our meta-analysis (*Schizophrenia Research*, *Schizophrenia Bulletin*, *Archives of General Psychiatry*, *American Journal of Psychiatry*, *Psychological Medicine*, *British Journal of Psychiatry* and *Psychiatry Research*).

Third, all references in the articles previously identified were screened to find other relevant publications.

The literature search was performed at the beginning of September 2003 and did not, therefore, include articles published after August 2003.

Selection of articles included in meta-analysis

To be included in our meta-analysis, articles had to meet the following criteria:

(*a*) diagnosis of schizophrenia (or schizoaffective disorder) in patients according to the RDC, DSM-III (or a later version), ICD-9 or ICD-10;

(*b*) inclusion of a group of first-degree relatives;

(*c*) inclusion of a control group;

(*d*) age of subjects above 18 years;

(*e*) results of tests reported individually (i.e. not only composite scores);

(*f*) statistics convertible to effect size (i.e. mean and standard deviation, or *F* or *t* statistics);

(*g*) data for relatives and controls were independent from other published data; when several studies were done in non-independent populations, we included only the study with the largest sample size.

In addition, we excluded studies that only included psychotic relatives of schizophrenic patients.

When insufficient details were provided to allow us to decide whether to include or to exclude a study, the authors were contacted for further information.

Recorded variables

For each article included, we recorded the following variables:

(1) principal author's name, date of publication and journal name;

(2) tests used and, for tests for which more than one method exists, the method used;

(3) proposant's diagnosis (i.e. schizoaffective patients included or not);

(4) exclusion or not of controls that have psychotic first-degree relatives;

(5) type of relatives included in the relatives group (i.e. sibs, parents, children, twins or mixed sample);

(6) exclusion criteria, based on lifetime psychiatric diagnosis in the relatives or control groups;

(7) presence or absence of demographic differences between the two groups that could affect neuropsychological functions (age, gender, educational level or socio-economic status);

(8) results of tests (recorded as mean and standard deviation or *F* and *t* statistics).

When results were provided for several subsamples of relatives and/or controls, we included only the largest samples that complied with the inclusion/exclusion rules with the following exception: when we had the choice, we preferentially used samples of relatives that did not include psychotic subjects. The rationale to exclude psychotic relatives, when possible, is based on the aim of this meta-analysis which is to assess differences in executive performances between controls and unaffected relatives. From this point of view, inclusion of psychotic subjects in the group of relatives could artificially exaggerate differences between groups.

https://doi.org/10.1017/S0033291704003460 Published online by Cambridge University Press

When means and standard deviations were available only in subsamples, we also calculated these statistics for the total samples using classical procedures (Armitage *et al*. 2002).

To avoid errors in variables transcription, all variables were independently recorded by two of us and in case of disagreement data were re-analysed to reach agreement.

Meta-analytical procedures

For all executive function tests that we analysed, we used the following procedure:

(1) Effect-size estimates were calculated as Hedges' unbiased *g* (Hedges & Olkin, 1985). To facilitate comparison between effect sizes across studies and across different executive tests, positive effect sizes always reflect better performance in controls. If necessary, data were transformed for this purpose.

(2) Homogeneity of data was tested as described by Hedges (1994). Homogeneity of effect-size estimates is an essential prerequisite in order to combine them in a global estimate. The absence of statistical heterogeneity between studies means that, despite differences in study design, variation in effect-size estimates is not greater than expected by chance.

(3) In presence of heterogeneity, recorded study characteristics [variables (2)–(7)] that could explain the extra variance were used to identify homogeneous subgroups of studies. Homogeneity was then tested in subgroups using a one-factor, fixed-effect model (Hedges, 1994).

(4) Forest plots were used to summarize data. To facilitate finding solutions for the previous step, studies were represented in ascending order of their effect estimates.

(5) For homogeneous data, effect sizes were combined to obtain an estimate of the global effect size. The pooled estimator of the population effect size, across all studies included in the meta-analysis, was obtained by weighting each effect-size estimate against the inverse of its variance (Wang & Bushman, 1999, p. 155).

(6) Funnel plots were used to evaluate the existence of publication bias as described by Sterne & Egger (2001).

Statistical analyses were performed using the SAS® V8 package as described by Wang & Bushman (1999).

RESULTS

Selection of executive tests to be analysed

Our procedure identified 100 abstracts of studies assessing executive functions. The tests used were cited in 63 of these abstracts. Four out of the 20 tests cited were cited 15 or more times: Wisconsin Card Sorting Test (WCST) (41 times), fluency tests (27 times), Trail Making Test (TMT) (22) and Stroop Test (15 times). The 16 other tests were cited no more than five times each (details available on request).

According to these data, we decided to concentrate our data analysis on the four tests most often used.

Selection of articles to be included in our meta-analysis

Our search identified 58 articles referring to the four tests selected (list of references available on request). Only 25 of these articles met the inclusion criteria for our meta-analysis. The selection process is summarized by the flow diagram in Fig. 1.

Table 1 summarizes the characteristics of the studies retained for analysis.

Analysis of effect sizes

Wisconsin Card Sorting Test

We found 20 studies that compared WCST performance in relatives of schizophrenic patients and controls. As the most commonly reported measures were number of categories completed (reported in 19 studies) and perseverative errors (reported in all 20 studies), we analysed these two variables.

Perseverative errors. Analysis of homogeneity demonstrated that data from the 20 studies were heterogeneous ($p = 0{\cdot}009$). The first factor that could explain this heterogeneity is the use of different forms of the WCST. A one-factor model (fixed effects) was used to assess the ability of this variable to explain data heterogeneity. Four different forms of the WCST were used but only one of them – the classical method (Heaton, 1981) – was used in enough studies to allow a separate meta-analysis. For this reason, we grouped the studies according to the method used: 'classic' and 'other'. The 'other' category included studies that used the computerized version, Nelson's method or Milner's method.

https://doi.org/10.1017/S0033291704003460 Published online by Cambridge University Press

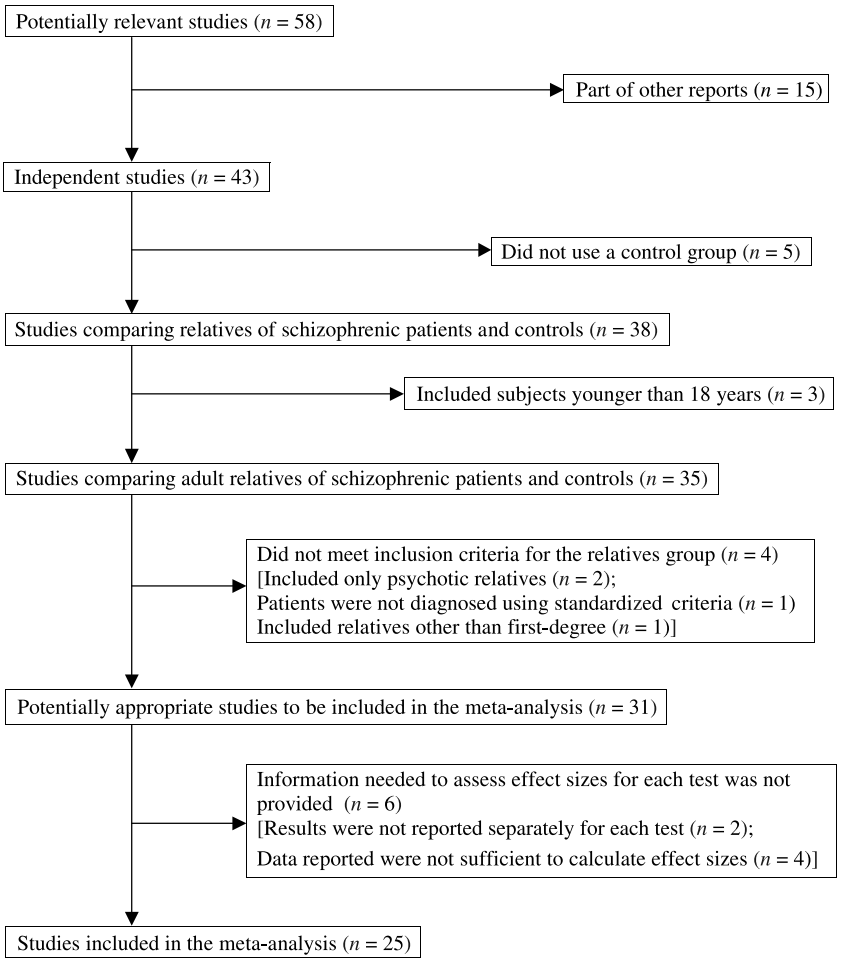

FIG. 1. Flow diagram of the process used to select the articles included in the meta-analysis.

Analysis showed that this classification system was relevant as it identified two homogeneous groups of studies ($p > 0{\cdot}05$), differing from each other ($p < 0{\cdot}05$). However, the homogeneity of the 'other' group could be artefact as the number of studies included was small ($n = 6$), and it also contains three different methods. For this reason, we restricted our analysis to 'classic' studies.

In this group of studies, relatives performed less well than controls: estimated effect 0·26 (95% CI 0·14–0·38).

Fig. 2 shows the effect estimates in the studies (and subgroups) selected for meta-analysis. Studies are classified as 'classic' or 'other' according to the method used and sorted by the size of the estimated effect.

Number of categories. The heterogeneity test revealed that data from the 19 studies were not homogeneous ($p = 0{\cdot}003$). Grouping the studies according to the method used revealed a significant difference between studies using the classic method and studies using other methods ($p = 0{\cdot}01$).

For the reasons outlined above, we included in our analysis only studies that used the classic method. Homogeneity testing revealed that these studies were not heterogeneous ($p = 0{\cdot}10$).

https://doi.org/10.1017/S0033291704003460 Published online by Cambridge University Press

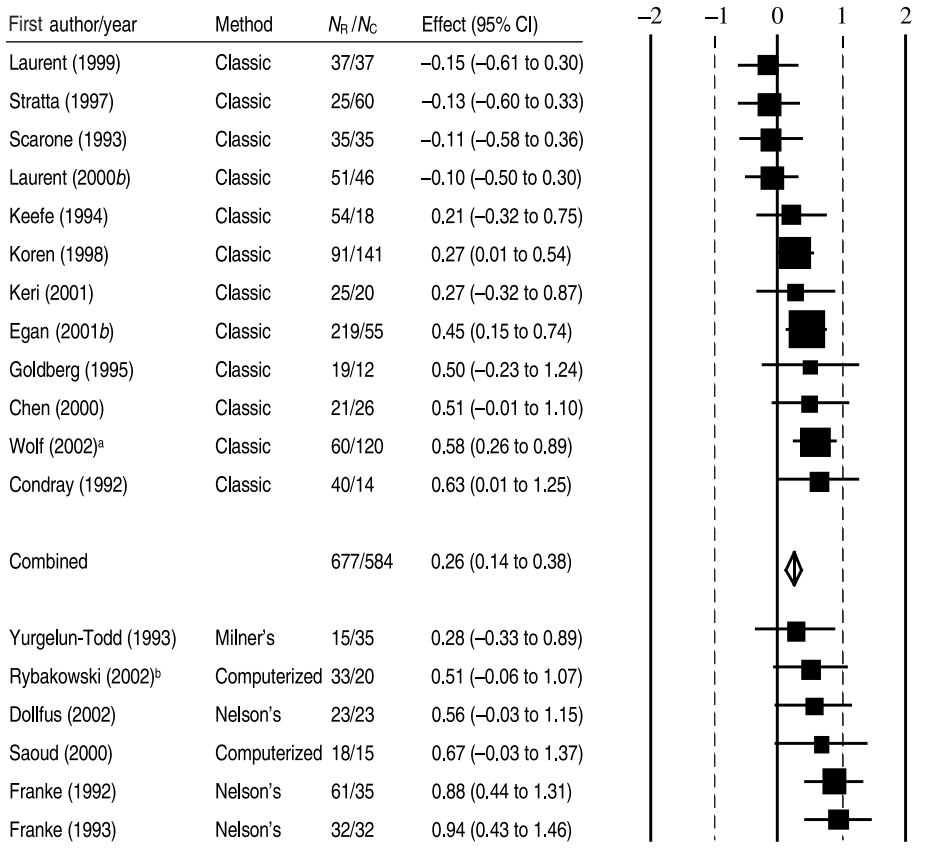

FIG. 2. WCST effect sizes of the primary studies included in our meta-analysis. Each study is represented by a square proportional to sample size and a horizontal line that corresponds to the confidence intervals. The centre of the square indicates the point estimate. N_R/N_C represents number of relatives and number of controls respectively. [a] Only relatives without psychotic disorders. [b] Only relatives without Axis I or II diagnosis.

The effect estimate was 0·31 (95% CI 0·21–0·42), with controls performing better than relatives of schizophrenic patients.

Stroop Test

Six studies provided data on performances on the interference (colour-word) trial of the Stroop Test. The data from these studies were not heterogeneous ($p = 0·15$). The effect estimate in the total sample was 0·38 (95% CI 0·21–0·55), with controls outperforming relatives.

Fluency tests

Phonological (*letter*) *fluency*. We found eight studies reporting results of letter fluency tests. The reported results were not heterogeneous ($p = 0·12$). The estimated effect for the total sample was 0·65 (95% CI 0·48–0·82).

Semantic (*category*) *fluency*. Seven studies provided data on category verbal fluency. Analysis demonstrated a high degree of heterogeneity ($p < 0·001$). To solve this problem, we plotted effect estimate and 95% confidence intervals in parallel with the characteristics of the studies (Fig. 3). As all these studies excluded controls having first-degree relatives with psychotic disorders, the corresponding column is not shown in Fig. 3.

Two studies gave clearly different effect-size estimates from the five others. The only characteristic that differentiates these two studies from the other studies was the inclusion

https://doi.org/10.1017/S0033291704003460 Published online by Cambridge University Press

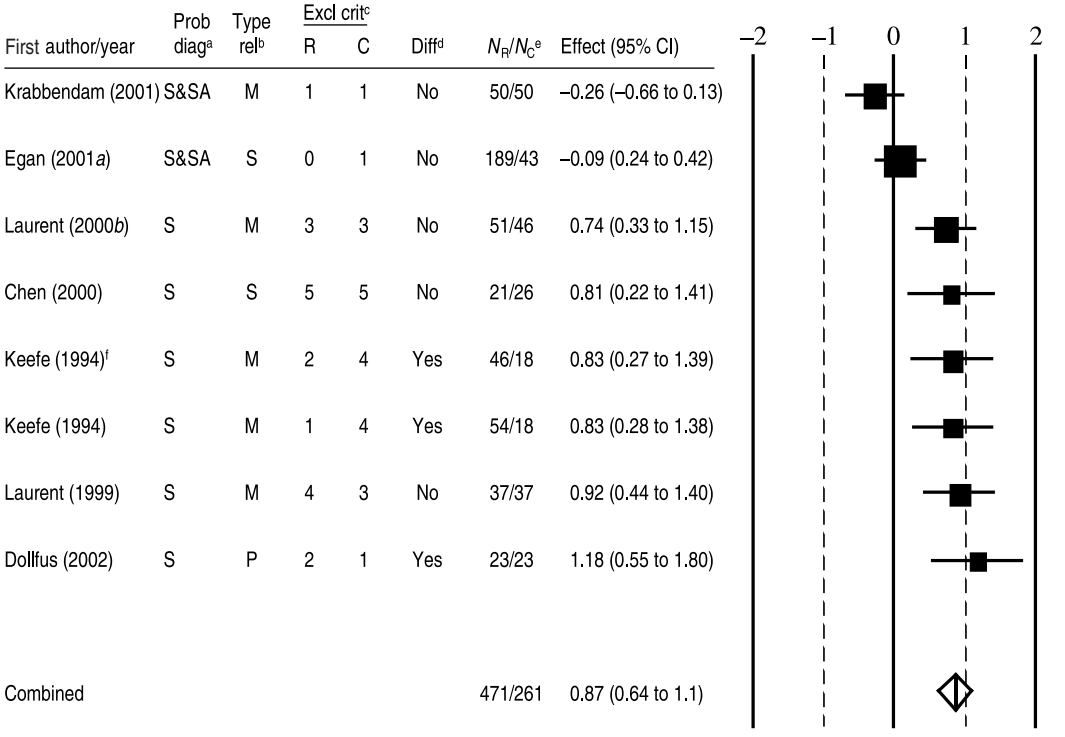

FIG. 3. Category Verbal Fluency effect sizes of the primary studies included in our meta-analysis. Each study is represented by a square proportional to sample size and a horizontal line that corresponds to the confidence intervals. The centre of the square indicates the point estimate. [a] Proband's diagnostic (S, only schizophrenic patients; S&SA, schizoaffective patients also included). [b] Type of relative sample (S, siblings; P, parents; M, mixed). [c] Exclusion criteria in relatives (R) and controls (C) [1, psychosis; 2, psychosis and SPD; 3, all DSM Axis I diagnostics; 4, all DSM Axis I diagnostics and SPD (or cluster I personality disorders); 5, all DSM Axis I and II diagnostics]. [d] Demographic variables are age, gender, school grades or socio-economic status [1, difference present or unspecified; 0, absent (or adjusted for)]. [e] N_R/N_C represents number of relatives and number of controls respectively. [f] Without relatives with SPD, not included in the meta-analysis.

of relatives of probands with either schizophrenia or schizoaffective disorder. When this characteristic was entered in a one-factor model (fixed effects), the two groups of studies were homogeneous ($p=0{\cdot}83$ for the 'only relatives of schizophrenic probands' group and respectively 0·18 the two studies including relatives of schizoaffective patients) but differed from one another ($p<0{\cdot}001$).

When only studies reporting data in relatives of schizophrenic subjects (186 relatives and 150 controls) were analysed, the estimated effect was 0·87 (95% CI 0·64–1·10), with relatives performing less well than controls.

Trail Making Test, part B

For the TMT, two variables are usually considered as measures of executive function: time to complete part B of the TMT and the difference between the times taken to complete parts B and A of the test. As the difference between the two parts of the TMT was rarely reported, we restricted our analysis to part B. Twelve studies reported data on part B of the TMT. Results from these studies were not heterogeneous ($p=0{\cdot}47$). The estimated effect for the combined studies was 0·49 (95% CI 0·37–0·62), with better performance for controls than for relatives.

Results from all meta-analyses are summarized in Table 2. The funnel plots did not suggest publication bias for any of the six measures of executive functions.

DISCUSSION

We quantitatively reviewed studies that compared the performances of relatives of schizo-

https://doi.org/10.1017/S0033291704003460 Published online by Cambridge University Press

Table 2. *Principal results of the meta-analysis of executive functions in relatives of schizophrenic probands*

Test	Subsample (for heterogeneous total sample)	Studies included	No. of relatives	No. of control subjects	Estimated effect (95% CI)
WCST perseverative errors	Classic method	12	677	584	0·26 (0·14–0·38)
WCST categories	Classic method	13	846	773	0·31 (0·21–0·42)
Stroop Test		6	294	269	0·38 (0·21–0·55)
Verbal Fluency – phonologic		8	431	233	0·65 (0·48–0·82)
Verbal Fluency – semantic	Relatives of schizophrenic patients	5	186	150	0·87 (0·64–1·10)
TMT, part B		11	672	507	0·49 (0·37–0·62)

phrenic patients and controls for the four tests of executive functions most used in the medical literature (WCST, TMT, VF and Stroop). Our results show that relatives of schizophrenic patients are impaired in all analysed measures of executive functions, but the estimated effect differs depending on the test used.

Our literature search identified 58 articles dealing with this subject. Only 25 of these articles met our inclusion criteria. The main reason for exclusion was inclusion of the same subjects in multiple reports (15 out of the 33 articles excluded). When necessary and possible, information in published articles was supplemented by contacting the authors.

Several of our methodological choices deserve discussion. First, rather than subjectively choosing the tests to be included in our meta-analysis, we decided to select the most frequently used tests of executive functions in the medical literature over the last 5 years. As we analysed only the abstracts and as one third of the abstracts contained no information concerning the tests used, our figures clearly do not reliably reflect the frequency of the tests used. However, our aim was not to determine the exact use of each test, but to obtain a simple, although rough, estimate of the most frequently used tests. The sizable gap between the fourth and the fifth most commonly used tests makes it improbable that other tests were used more frequently than the four we selected for our meta-analysis.

Second, we included only data from published studies, supplemented when necessary, with information from the authors concerning inclusion/exclusion criteria, data on individual tests (when only composite scores were reported) or statistics convertible to effect sizes (when only p values were available). It is a matter of debate whether unpublished data should, or should not, be used in meta-analyses. As data may not have been published because they show non-significant differences, not including them favour greater differences between the experimental (i.e. relatives) group and the control group. However, unpublished data may not have been published because the studies in question did not comply with scientific standards needed for publication, meaning that these data are less reliable. Inclusion of only information that complemented published studies allowed us to supplement information without including less reliable data. In addition, to test the existence of a publication bias, we used funnel plots.

For the four tests (and six measures) that we analysed, relatives of schizophrenic patients were more impaired than controls. The effect estimates were small to moderate (from 0·26 to 0·49) for three of the tests (Stroop, WCST, TMT) but were greater for the fluency tests (0·65 and 0·87). These results suggest that fluency tests are more sensitive to impairments in relatives of schizophrenic patients than the other three tests. However, for some of the tests (Stroop and VF) the total number of studies and subjects was small and there is some overlap in the 95% confidence intervals of the effect estimates for the different tests (Table 2) suggesting that more research is needed before firm conclusions could be drawn. The presence of deficits in all the tests analysed does not allow us to draw conclusions regarding the existence of a specific deficit in executive functions (for more

extensive discussion on the problem of specific *versus* generalized deficit see Chapman & Chapman, 2001). Two previous studies (Sitskoorn *et al.* 2003; Snitz *et al.* 2003) found that the effect sizes for tests of executive functions were in the small to moderate range (0·2–0·5). This is concordant with our results for WCST, Stroop and TMT – B, but not for fluency tests. Unfortunately, as these studies were published only as abstracts, there are insufficient details to allow us to explain these partially discordant findings.

Our study revealed that three measures, from two executive tasks (semantic fluency, WCST preservative errors and categories), were not homogeneous. For the WCST measures, the use of four different methods (i.e. classical, Nelson's, Milner's and computerized) is the most probable explanation for this heterogeneity. It is noteworthy that Nelson's method and the computerized version of the WCST seem to be more sensitive than the classical method at detecting impairments in relatives of schizophrenic patients. However, before a firm conclusion can be draw, more studies must be done using these two methods.

For semantic fluency, heterogeneity of studies is more problematic to explain. Entering proband diagnosis in a one-factor model solved this heterogeneity by excluding from the analysis two studies (Egan *et al.* 2001*a*; Krabendam *et al.* 2001) with particularly low effect estimates. However, several indices suggest that our meta-analytical findings on semantic fluency should be interpreted with caution. First, the two excluded studies were not only those that included relatives of schizoaffective probands, but also those with the smallest standard error and largest total samples, thus suggesting the existence of a potential publication bias. Second, the number of studies before, but especially after, exclusion of those studies was small. Finally, unless the proportion of relatives of schizoaffective probands was very high, this explanation seems fairly unlikely. More studies are needed to explore the degree of impairment in verbal semantic fluency among relatives of schizophrenic patients. In addition, it seems important for future studies to separate relatives of schizoaffective patients from relatives of schizophrenic patients.

Our study was restricted to the four most used tests in the medical literature and, as a matter of fact, the only tests for which sufficient studies exist to merit a meta-analysis. As more data will accumulate it will be interesting to compare results from these classical neurocognitive tests with results from newer tests designed to assess more specific executive processes: planning (e.g. the Tower of London; Shallice, 1982), initiation and suppression of response (the Hayling test; Burgess & Shallice, 1996), cognitive estimation (Shallice & Evans, 1978), etc.

As separate data for non-psychotic relatives were not available in some studies we also included in our analysis samples of relatives containing psychotic subjects. However, because there were only two such studies (Harris *et al.* 1996; Egan *et al.* 2001*a*), the percentage of psychotic relatives included was small, and their results were homogeneous with those of studies not including psychotic relatives, we consider that the global effect estimates were not significantly influenced by this issue.

It is somewhat surprising that results were not more heterogeneous in spite of major differences in study designs. In particular, it should be noted that differences in inclusion criteria for relatives (affected relatives included or not, type of relatives, etc.) or controls (taking into account positive family history or not) did not affect the homogeneity of results. This suggests that if these variables affect test results, the effect is not very large. However, an alternative explanation could be limited statistical power due to the small number of studies included in the meta-analysis.

In conclusion, our meta-analysis suggests that relatives of schizophrenic patients have widely, although not severely, impaired executive functions. As the sensitivity of the different tests for impairments in relatives is not the same, the choice of test and method used should be carefully assessed.

ACKNOWLEDGEMENTS

We thank Dr Szabolcs Keri, Dr Annie Laurent, Dr Mohamed Saoud and Dr Janusz Rybakowski for providing the information needed to check the inclusion criteria for their articles. We are grateful to Dr Steven Faraone and Dr Lydia Krabbendam for providing supplementary, unpublished data. This research was supported by grants from Délégation à la

Recherche Clinique de l'AP-HP (PHRC AOM 98152) and INSERM. A.S. and A.M. received grants from the Fondation pour la Recherche Médicale.

DECLARATION OF INTEREST

None.

REFERENCES

Armitage, P., Berry, G. & Matthews, J. N. S. (2002). *Statistical Methods in Medical Research* (4th edn). Blackwell Science: Oxford.

Asarnow, R. F., Nuechterlein, K. H., Subotnik, K. L., Fogelson, D. L., Torquato, R. D., Payne, D. L., Asamen, J., Mintz, J. & Guthrie, D. (2002). Neurocognitive impairments in nonpsychotic parents of children with schizophrenia and attention-deficit/hyperactivity disorder: the University of California, Los Angeles Family Study. *Archives of General Psychiatry* **59**, 1053–1060.

Burgess, P. W. (1997). Theory and methodology in executive function research. In *Methodology of Frontal and Executive Function* (ed. P. Rabbitt), pp. 81–116. Psychology Press: Hove.

Burgess, P. W. & Shallice, T. (1996). Response suppression, initiation and strategy use following frontal lobe lesions. *Neuropsychologia* **34**, 263–273.

Chapman, L. J. & Chapman, J. P. (2001). Commentary on two articles concerning generalized and specific cognitive deficits. *Journal of Abnormal Psychology* **110**, 31–39.

Chen, Y. L., Chen, Y. H. & Lieh-Mak, F. (2000). Semantic verbal fluency deficit as a familial trait marker in schizophrenia. *Psychiatry Research* **95**, 133–148.

Condray, R., Steinhauer, S. R. & Goldstein, G. (1992). Language comprehension in schizophrenics and their brothers. *Biological Psychiatry* **32**, 790–802.

Dollfus, S., Lombardo, C., Bénali, K., Halbecq, I., Abadie, P., Marié, R.-M. & Brazo P. (2002). Executive/attentional cognitive functions in schizophrenic patients and their parents: a preliminary study. *Schizophrenia Research* **53**, 93–99.

Egan, M. F., Goldberg, T. E., Gscheidle, T., Weirich, M., Rawlings, R., Hyde, T. M., Bigelow, L. & Weinberger, D. R. (2001*a*). Relative risk for cognitive impairments in siblings of patients with schizophrenia. *Biological Psychiatry* **50**, 98–107.

Egan, M. F., Goldberg, T. E., Kolachana, B. S., Callicott, J. H., Mazzanti, C. M., Straub, R. E., Goldman, D. & Weinberger, D. R. (2001*b*). Effect of COMT Val108/158 Met genotype on frontal lobe function and risk for schizophrenia. *Proceedings of the National Academy of Science USA* **98**, 6917–6922.

Erlenmeyer-Kimling, L., Rock, D., Roberts, S. A., Kestenbaum, C., Cornblatt, B., Adamo, U. H. & Gottesman, I. I. (2000). Attention, memory, and motor skills as childhood predictors of schizophrenia-related psychoses: the New York high-risk project. *American Journal of Psychiatry* **157**, 1416–1422.

Faraone, S. V., Seidman, L. J., Kremen, W. S., Pepple, J. R., Lyons, M. J. & Tsuang, M. T. (1995). Neuropsychological functioning among the nonpsychotic relatives of schizophrenic patients: a diagnostic efficiency analysis. *Journal of Abnormal Psychology* **104**, 286–304.

Franke, P., Maier, W., Hain, C. & Klingler, T. (1992). Wisconsin Card Sorting Test: an indicator of vulnerability to schizophrenia? *Schizophrenia Research* **6**, 243–249.

Franke, P., Maier, W., Hardt, J. & Hain, C. (1993). Cognitive functioning and anhedonia in subjects at risk for schizophrenia. *Schizophrenia Research* **10**, 77–84.

Freedman, R., Adler, L. E. & Leonard, S. (1999). Alternative phenotypes for the complex genetics of schizophrenia. *Biological Psychiatry* **45**, 551–558.

Goldberg, T. E., David, A. & Gold, J. M. (2003). Neurocognitive deficits in schizophrenia. In *Schizophrenia* (2nd edn) (ed. S. R. Hirsch and D. Weinberger), pp. 168–184. Blackwell Science: Oxford.

Goldberg, T. E., Torrey, E. F., Gold, J. M., Bigelow, L. B., Ragland, R. D., Taylor, E. & Weinberger, D. R. (1995). Genetic risk of neuropsychological impairment in schizophrenia: a study of monozygotic twins discordant and concordant for the disorder. *Schizophrenia Research* **17**, 77–84.

Gottesman, I. I. & Gould, T. D. (2003). The endophenotype concept in psychiatry: etymology and strategic intentions. *American Journal of Psychiatry* **160**, 636–645.

Gottesman, I. I. & Shields, J. (1973). Genetic theorizing and schizophrenia. *British Journal of Psychiatry* **122**, 15–30.

Harris, J. G., Adler, L. E., Young, D. A., Munro Cullum, C., Rilling, L. M., Cicerello, A., Intemann, P. M. & Freedman, R. (1996). Neuropsychological dysfunction in parents of schizophrenics. *Schizophrenia Research* **20**, 253–260.

Heaton, R. (1981). *The Wisconsin Card Sorting Test Manual*. Psychological Assessment Resources: Odessa, FL.

Hedges, L. V. (1994). Fixed effect models. In *The Handbook of Research Synthesis* (ed. H. Cooper and L. V. Hedges), pp. 285–299. Russell Sage Foundation: New York.

Hedges, L. V. & Olkin, I. (1985). *Statistical Methods for Meta-analysis*. Academic Press: New York.

Heinrichs, R. W. & Zakzanis, K. K. (1998). Neurocognitive deficit in schizophrenia: a quantitative review of evidence. *Neuropsychology* **12**, 426–445.

Johnson-Selfridge, M. & Zalewski, C. (2001). Moderator variables of executive functioning in schizophrenia: meta-analytic findings. *Schizophrenia Bulletin* **27**, 305–316.

Keefe, R. S., Silverman, J. M., Roitman, S. E., Harvey, P. D., Duncan, M. A., Alroy, D., Siever, L. J., Davis, K. L. & Mohs, R. C. (1994). Performance of nonpsychotic relatives of schizophrenic patients on cognitive tests. *Psychiatry Research* **53**, 1–12.

Keri, S., Kelemen, O., Benedek, G. & Janka, Z. (2001). Different trait markers for schizophrenia and bipolar disorder: a neurocognitive approach. *Psychological Medicine* **31**, 915–922.

Koren, D., Seidman, L. J., Harrison, R. H., Lyons, M. J., Kremen, W. S., Caplan, B., Goldstein, J. M., Faraone, S. V. & Tsuang, M. T. (1998). Factor structure of the Wisconsin Card Sorting Test: dimensions of deficit in schizophrenia. *Neuropsychology* **12**, 289–302.

Krabbendam, L., Marcelis, M., Delespaul, P., Jolles, J. & van Os, J. (2001). Single or multiple familial cognitive risk factors in schizophrenia? *American Journal of Medical Genetics* **105**, 183–188.

Laurent, A., Biloa-Tang, M., Bougerol, T., Duly, D., Anchisi, A. M., Bosson, J. L., Pellat, J., d'Amato, T. & Dalery, J. (2000*a*). Executive/attentional performance and measures of schizotypy in patients with schizophrenia and in their nonpsychotic first-degree relatives. *Schizophrenia Research* **46**, 269–283.

Laurent, A., d'Amato, T., Naegele, B., Murry, P., Baro, P., Foussard, N., Spitz, F. & Dalery, J. (2000*b*). Executive and amnestic functions of a group of first-degree relatives of schizophrenic patients [in French]. *Encéphale* **26**, 67–74.

Laurent, A., Moreaud, O., Bosson, J. L., Naegele, B., Boucharlat, J., Saoud, M., d'Amato, T. & Dalery, J. (1999). Neuropsychological functioning among non-psychotic siblings and parents of schizophrenic patients. *Psychiatry Research* **87**, 147–157.

Leboyer, M. (2003). Searching for alternative phenotypes in psychiatric genetics. *Methods in Molecular Medicine* **77**, 145–161.

Leboyer, M., Bellivier, F., Nosten-Bertrand, M., Jouvent, R., Pauls, D. & Mallet, J. (1998). Psychiatric genetics: search for phenotype. *Trends in Neurosciences* **21**, 102–105.

Mirsky, A. F., Yardley, S. L., Jones, B. P., Walsh, D. & Kendler, K. S. (1995). Analysis of the attention deficit in schizophrenia: a study of patients and their relatives in Ireland. *Journal of Psychiatric Research* **29**, 23–42.

Rabbit, P. (1997). Introduction: methodologies and models in the study of executive function. In *Methodology of Frontal and*

Executive Function (ed. P. Rabbitt), pp. 1–38. Psychology Press: Hove.

Rund, B. R. (1998). A review of longitudinal studies of cognitive functions in schizophrenia patients. *Schizophrenia Bulletin* **24**, 425–435.

Rybakowski, J. K. & Borkowska, A. (2002). Eye movement and neuropsychological studies in first-degree relatives of schizophrenic patients. *Schizophrenia Research* **54**, 105–110.

Saoud, M., d'Amato, T., Gutnecht, C., Triboulet, P., Bertaud, J. P., Marie-Cardine, M., Dalery, J. & Rochet, T. (2000). Neuropsychological deficits in siblings discordant for schizophrenia. *Schizophrenia Bulletin* **26**, 893–902.

Scarone, S., Abbruzzese, M. & Gambini, O. (1993). The Wisconsin Card Sorting Test discriminates schizophrenic patients and their siblings. *Schizophrenia Research* **10**, 103–107.

Schröder, C., Maier, W. & Wagner, M. (2002). A meta-analysis of neurocognitive deficits in relatives of patients with schizophrenia. *Schizophrenia Research* (Suppl.) **53**, 126.

Shallice, T. (1982). Specific impairments in planning. *Philosophical Transactions of the Royal Society of London B* **298**, 199–209.

Shallice, T. & Evans, M. E. (1978). The involvement of the frontal lobes in cognitive estimation. *Cortex* **14**, 294–303.

Sitskoorn, M., Aleman, A., Ebisch, S., Appels, M. & Kahn, R. S. (2003). Meta-analysis of neurocognitve deficits in relatives of patients with schizophrenia: evidence for a strong cognitive endophenotypic marker. *Schizophrenia Research* (Suppl.) **60**, 90.

Snitz, B. E., MacDonald, A. W. & Carter, C. S. (2003). Cognitive impairments in adult first-degree relatives of schizophrenia patients: a meta-analytic review. *Schizophrenia Research* (Suppl.) **60**, 158.

Spitzer, R. L., Endicott, J. & Robins, E. (1978). Research diagnostic criteria: rationale and reliability. *Archives of General Psychiatry* **35**, 773–782.

Sterne, J. A. & Egger, M. (2001). Funnel plots for detecting bias in meta-analysis: guidelines on choice of axis. *Journal of Clinical Epidemiology* **54**, 1046–1055.

Stratta, P., Daneluzzo, E., Mattei, P., Bustini, M., Casacchia, M. & Rossi, A. (1997). No deficit in Wisconsin Card Sorting Test performance of schizophrenic patients' first-degree relatives. *Schizophrenia Research* **26**, 147–151.

Tsuang, M. T., Lyons, M. J. & Faraone, S. V. (1990). Heterogeneity of schizophrenia: conceptual models and analytic strategies. *British Journal of Psychiatry* **156**, 17–26.

Wang, M. C. & Bushman, B. J. (1999). *Integrating Results through Meta-analytic Review Using SAS Software*. SAS Institute Inc.: Cary, NC.

Wolf, L. E., Cornblatt, B. A., Roberts, S. A., Maminski Shapiro, B. & Erlenmeyer-Kimling, L. (2002). Wisconsin Card Sorting deficits in offspring of schizophrenics in the New York high-risk project. *Schizophrenia Research* **57**, 173–182.

Yurgelun-Todd, D. A. & Kinney, D. K. (1993). Patterns of neuropsychological deficits that discriminate schizophrenic individuals from siblings and control subjects. *Journal of Neuropsychiatry and Clinical Neuroscience* **5**, 294–300.