Twin Research (2000) 3, 134–141

www.nature.com/tr

Zygosity diagnosis in young twins by parental report

MJH Rietveld[1], JC van der Valk[1,2], IL Bongers[1], TM Stroet[1], PE Slagboom[3] and DI Boomsma[1]

[1]Vrije Universiteit, Department of Biological Psychology, Amsterdam
[2]Erasmus University, Department of Child and Adolescent Psychiatry, Rotterdam
[3]TNO-PG, Gaubius Laboratory, Department of Vascular and Connective Tissue Research, Leiden, The Netherlands

This study reports on zygosity determination in twins of childhood age. Parents responded to questionnaire items dealing with twin similarity in physical characteristics and frequency of mistaking one twin for another by parents, relatives and strangers. The accuracy of zygosity diagnosis was evaluated across twins aged 6, 8, and 10 and across parents. In addition, it was examined whether the use of multiple raters and the use of longitudinal data lead to an improvement of zygosity assignment. Complete data on zygosity questions and on genetic markers or blood profiles were available for 618 twin pairs at the age of 6 years. The method used was predictive discriminant analyses. Agreement between zygosity assigned by the replies to the questions and zygosity determined by DNA markers/blood typing was around 93%. The accuracy of assignment remained constant across age and parents. Analyses of data provided by both parents and collected over multiple ages did not result in better prediction of zygosity. Details on the discriminant function are provided. Twin Research (2000) 3, 134–141.

Keywords: twin zygosity, childhood, questionnaire, review, discriminant analysis

Introduction

In 1927, Siemens[1] suggested that the diagnosis of zygosity in twins can take place by evaluating the degree of resemblance on genetically determined traits. Development of this method resulted in the frequent use of questionnaires, often including those criteria originally proposed by Siemens, for example.[2] Several studies have shown that the establishment of zygosity based on mailed questionnaires is of considerable accuracy, with around 95% correctly classified compared with blood or DNA typing. Studies on the diagnosis of zygosity by mailed questionnaires are summarised in Appendix 1.[3–23]

The purpose of this paper is twofold. First, the validity of zygosity classification across childhood is examined in a large sample. One might expect the physical dissimilarity between dizygotic twins to become more obvious as they grow up. If so, the accuracy of classification is likely to improve with increasing age of the participants. A few studies have reported on this issue by evaluating the precision of zygosity diagnosis between samples varying in age,[8,19,23] and by test–retest estimation.[9] With the exception of the study of Cohen et al,[9] the findings are suggestive of an increased precision in zygosity prediction for older participants. However, findings may have suffered from a lack of statistical power due to a relatively small number of co-operating twins and parents.

To our knowledge there are no studies investigating this issue in a longitudinal sample. Since the availability of longitudinal data of various birth cohorts is increasing in several twin registers,[24] the establishment of zygosity incorporating longitudinal data deserves our attention. The Netherlands Twin Register collects questionnaire data on zygosity items at multiple ages in the same children by parental report. By making use of this longitudinal dataset it is possible to examine whether analysing all available data collected at different ages increases the precision of classification or whether it is sufficient or possibly advisable to rely on information obtained at a specific age only. We are especially interested to determine if reliable classification of zygosity can take place as early as age 6.

The second objective is to investigate how to make optimal use of information provided by multiple carers. The majority of participating families registered with the Netherlands Twin Register returns two completed questionnaires, usually filled in by the mother and father of the twin pair. In other twin studies of young children, typically the mother is used as primary informant.[17] It is of interest to find out whether the precision of the establishment of zygosity can further improve if information provided by a second informant is included in the analyses.

The Netherlands Twin Register has access to complete data on bloodgroup typing or DNA polymorphism and zygosity questionnaires collected in a sample of 618 twin pairs at age 6. This large number

Correspondence: MJH Rietveld, Vrije Universiteit, Department of Biological Psychology, De Boelelaan 1111, 1081 HV Amsterdam, The Netherlands. Fax: +31 20 444 8832; Tel: +31 20 444 8812; E-mail: mjh.rietveld@psy.vu.nl
Received 28 February 2000; accepted 24 March 2000

of participants provides sufficient statistical power to investigate the above issues.

Materials and methods

Subjects

The Netherlands Twin Register (NTR) is a population-based register, which contains 40%–50% of all multiple births after 1986.[25] As part of a current longitudinal study on the development of behaviour problems, two questionnaires are sent to the registered parents or primary carers at multiple points in time with an average interval of 2 years. The present study used information by parental report on twin similarity and twin confusion at three ages in childhood, for cohorts born between 1986 and 1991. On the first occasion of data collection, around the sixth birthday of the twins (mean = 6.36 years, SD = 0.95), information on zygosity by report of the father was not requested. At the second and third assessment, age 8 (mean = 7.90 years, SD = 0.50) and age 10 (mean = 10.27 years, SD = 0.40) respectively, both parents provided information on zygosity items. For this study, only pairs of same sex with DNA/blood zygosity data were included in the analyses (n = 691 pairs). Twin pairs with missing items on the parental zygosity questions were excluded. Table 1 reports on the numbers of same sex twin pairs with complete data on the zygosity items and DNA/blood typing at each age.

Complete longitudinal data were available from 253 mothers (age 6, 8, and 10), and from 224 fathers (age 8 and 10). Data from both raters were collected in 316 twin pairs at age 8, and in 257 twin pairs at age 10. The sample participating in this study was predominantly of Caucasian origin, with around 2% classified into other ethnic groups.

Zygosity questionnaire

The questionnaire used in the present study asked for information regarding similarity of the children and experiences of mistaking one for another (Appendix 2). When the twins were aged 6, parents provided information on eight items. In addition, a question concerning knowledge of zygosity classification based on DNA/blood testing was included. This item was used to identify those families with knowledge of zygosity prior to completing the questions. Two more items were added to the zygosity questionnaire at the second and third measurement occasion.

Table 1 Number of twin pairs participating in the present study

	Age 6 Mother	Age 8 Mother	Age 8 Father	Age 10 Mother	Age 10 Father
Questionnaire and DNA/blood data	618	394	335	324	279
MZ	388	243	210	200	163
DZ	230	151	125	124	116

Genotyping and blood polymorphism

A total of 691 same sex twin pairs participated in DNA/blood testing; 62% donated blood samples for analyses of blood grouping profiles and 38% provided a mouth swab sample for DNA isolation. Zygosity determination was performed using eight highly polymorphic di-, tri- and tetranucleotide genetic markers. The zygosity testing included a multiplex PCR of markers D2S125, D8S1130, D1S1609, D5S816 and a second multiplex reaction of markers 15 ActC, D21S1437, D7S2846, and D10S1423. These two multiplex PCR reactions were performed essentially by the protocol provided in the website of the Marshfield Institute (http://www.marshmed.org/genetics/). For the purpose of zygosity determination based on blood grouping profiles, red cells were typed with test sera for the following red cell blood group antigens: AB, CcDEe, MNSs, P_1, Kk, Kp^aKp^b, Fy^aFy^b, Jk^aJk^b, Lu^aLu^b. More details on the collection and treatment of these blood samples are given by Van Dijk et al.[26]

Statistical procedures

All parents of twins with DNA/blood data were informed about the zygosity results. Since the employment of DNA/blood testing varied across age, two groups of families could be distinguished. One group of parents with knowledge of the DNA/blood test results before completion of the questionnaire, and one group of parents whose twin pair had not yet participated in the DNA/blood testing. Since prior knowledge of the test results may affect responses to the zygosity questions, it was established first whether the two groups of parents differed in their item response pattern. If so, generalisation of the application of the statistical function to samples for which no information on biological indices is available is seriously hampered. The tests were performed on each item separately by employment of χ^2 tests.

Predictive discriminant analysis was used for classifying subjects into MZ and DZ groups.[27,28] In the present study, the discriminant analysis generated a linear function of the weighted sum of the questionnaire items with the weightings chosen, such that the distinction between MZ and DZ twins was optimal. The estimated success of classification or hit rate is the proportion of correctly classified observations in the sample. It is sometimes argued that this hit rate is optimistically biased since the classification rule is derived from and applied to the same sample. This bias can be avoided in two ways, either through use of large samples or through application of an external classification analysis. In this study, both routes are taken. As a criterion for sample size, it is proposed that the minimum of observations in the smallest group should be at least five times the number of questionnaire items. As can be seen in Table 1, this requirement was easily met by each individual dataset. The leave-one-out procedure was chosen as the preferred external analysis. This method omits an observation, recalculates the classification rule from the remaining observations, classifies the deleted observation, and repeats these steps for each observation in the sample. The number of deleted observations correctly classified are counted and reported as cross-validated hit rates. Considering the proportion of same-sex MZ and DZ twins in the population, equal prior probabilities of group membership were used. To define the underlying construct that the discriminant function represents, inspection of the correlations between the discriminant function and each of the questionnaire variables was performed. The discriminant function and descriptive statistics were calculated using Statistical Package for Social Sciences/ Windows 9.0.

Results

At age 6, out of 618 pairs with DNA/blood data, 411 mothers knew the result of zygosity testing and 199 mothers had not yet received a request for DNA/ blood testing for their twins. Eight mothers had not answered the question. The ratio MZ : DZ was equal in both groups and data were pooled across zygosities to examine mothers' responses between groups. A difference in response pattern was observed for 1 item only, 'do strangers have difficulty telling them apart?' (χ^2 = 5.17 (1), P = 0.02). A positive answer was given by 65% of those mothers who were ignorant of zygosity, compared with 75% among mothers with knowledge of the DNA/blood test result. Overall, the two groups did not seem to differ allowing the discriminant function to be applied to both groups simultaneously.

A summary of the results of the first series of discriminant analyses is given in Table 2. Each analysis indicated a very accurate hit rate. Between 91.6% and 94.2% of all twin pairs were assigned the correct zygosity by the discriminant function. The precision of classification was not equally distributed across zygosities. Irrespective of age, correct classification for MZ twins was estimated around 97%, whereas around 88% of DZ twins were identified correctly.

Table 2 Classification results by use of discriminant function analyses

		Age 6 Mother	Age 8 Mother	Age 8 Father	Age 10 Mother	Age 10 Father
Correctly classified	MZ	96.6%	95.1%	97.1%	97.5%	96.9%
	DZ	90.0%	86.8%	85.6%	88.7%	89.7%
Cross-validated	Total	94.2%	91.6%	91.9%	92.6%	93.9%

Next, twin pairs with longitudinal questionnaire data were considered. The analysis of data collected at age 6, 8, and 10 by report from the mother resulted in a hit rate of 93.7%. Analysis of fathers' reports collected at the twins' age of 8 and 10 yielded a correct classification of 94.2%. Finally, data from mother and father were analysed jointly. At age 8, 93.4% of all twin pairs were classified correctly. A hit rate of 93.8% was obtained at age 10.

The above cross-validated hit rates indicated a minimal difference in the precision of assignment across the use of various datasets. The use of multiple raters and longitudinal data did not lead to an increased precision of zygosity prediction. Because the majority of twin studies are performed within cross-sectional designs, we believe it is of much practical use to report upon the discriminant function coefficients resulting from the first series of analyses. These parameter values together with the associated classification scores are given in Appendix 3. For interpreting the discriminant function, we have listed the correlations between each function and each questionnaire item in Table 3.

Across age and parent, the majority of the correlations ranged from 0.50 to 0.80. Identification of those questionnaire items that show the largest overlap with the function helps to determine the underlying construct that the discriminant function represents. The zygosity questionnaire was developed along two dimensions, similarity of physical characteristics and confusion of identity. At either age and for either parent, the most informative correlations were not clustered in a sense that the function could easily be defined along one of these dimensions. Closer inspection revealed a few interesting details. With the exception of item 1 (facial appearance) and item 2 (hair colour), a relatively large degree of overlap was observed between mothers and fathers within age 8 and age 10 of the twins. Looking at the ranking of the items, parents evaluated the questions in the same general manner. When the percentage of correctly classified twins was taken into consideration, this indicated that parents are interchangeable in assessing identity and fraternity in their children. Another interesting finding was the very small correlation found for item 5 ('peas in a pod'). In contrast to numerous other studies, for example,

Table 3 Correlations between discriminant function and individual questionnaire items

Item	Age 6 Mother	Ranking	Age 8 Mother	Ranking	Age 8 Father	Ranking	Age 10 Mother	Ranking	Age 10 Father	Ranking
Facial appearance	0.72	1	0.67	3	0.72	2	0.62	6	0.66	3
Hair colour	0.67	3	0.70	2	0.67	4	0.71	2	0.58	6
Face colour	0.66	4	0.63	6	0.65	6	0.68	5	0.63	5
Eye colour	0.52	6	0.53	7	0.50	7	0.51	7	0.52	7
'Two peas'	0.47	7	0.46	8	0.43	8	0.39	8	0.40	8
Mother/father	0.32	8	0.27	9	0.28	9	0.24	9	0.28	9
Family members	0.68	2	0.64	4	0.66	5	0.70	3	0.63	4
Strangers	0.62	5	0.64	5	0.71	3	0.75	1	0.82	1
Photograph			0.15	10	0.15	10	0.12	10	0.23	10
Hair structure			0.76	1	0.75	1	0.70	4	0.68	2

Magnus et al,[16] this item was of minor importance in defining the discriminant function. Even smaller correlations were observed for item 6 (confusion by mother or father) and item 9 (tell twins apart in photograph). The association among these three items seems obvious given that these questions rely on parental impression of global similarity and parental confusion of twins' identities. Apparently, parents themselves did not have difficulties in telling who is who.

Discussion

The primary focus of this study was to evaluate the accuracy of zygosity determination in young children. As young as age 6, the precision in zygosity prediction was high, with 94% agreement between zygosity assigned by the parental replies to the questionnaire items and zygosity determined by blood typing or analyses of genetic markers. It was found that the accuracy of classification remained stable across childhood. The suggestion that determination improves with increasing age due to more obvious dissimilarities in dizygotic twin pairs was not confirmed. It was also found that mothers and fathers were equally effective in diagnosing their children.

Although the questionnaire items allow an accurate determination of zygosity, the accuracy resulting from the discriminant analyses was not equally distributed in monozygotic and dizygotic pairs. At each age and for both parents, a bias towards classification as monozygotic twins took place. This may have resulted either from a tendency by parents to overestimate similarities in their twin children or from a lack of sensitivity of these questions to detect fraternity. The former case seems less plausible, considering assessment of parental replies to a question that deals with their personal opinion of the twins' zygosity. This item is included in a questionnaire sent to parents shortly after registration with the NTR (before the twins' first birthday). Correct in 80% of the cases, parents misclassified true MZ twins more than four times as often as true DZ. This result may reflect either the fact that parents are misinformed by physicians or the parents' wish for fraternity, or a combination of both. A preference towards labelling a twin as dizygotic is commonly found both by use of parental report, as in Cohen et al[9] and self report.[29]

The sample used in the analyses was mainly Caucasian. This may imply that the use of the zygosity questionnaire and the application of the discriminant functions do not generalise to groups of non-Caucasian ethnic origin.

Concluding, the use of the zygosity questions and the employment of discriminant analysis as multivariate tool for classification seem of value in determining zygosity in young twins.

Acknowledgements

This work was financially supported by USF grant 96/22. The work of Jolande van der Valk was made possible by grant M165 from the Sophia Foundation for Medical Research.

References

1 Siemens HW. The diagnosis of identity in twins. J Hered 1927; 18: 201–209.
2 Goldsmith HH. A zygosity questionnaire for young twins: A research note. Behav Genet 1991; 21: 257–269.
3 Cederlöf R, Friberg L, Jonsson E, Kaij L. Studies on similarity diagnosis in twins with the aid of mailed questionnaires. Acta Genet Stat Med 1961; 11: 338–362.
4 Nichols RC, Bilbro WC. The diagnosis of twin zygosity. Acta Genet Stat Med 1966; 16: 265–275.
5 Jablon S, Neel JV, Gershowitz H, Atkinson GF. The NAS-NRC twin panel: Methods of construction of the panel, zygosity diagnosis, and proposed use. Am J Hum Genet 1967; 19: 133–161.
6 Hauge M, Harvald B, Fischer M, Gotlieb-Jensen K, Juel-Nielsen N, Raebild I, Shapiro R, Videbech T. The Danish Twin Register. Acta Genet Med Gemellol (Roma) 1968; 17: 315–332.
7 Schoenfeldt LF. A comparison of two analytic procedures for estimating twin zygosity. Hum Hered 1969; 19: 343–353.

8 Cohen DJ, Dibble E, Grawe JM, Pollin W. Separating identical from fraternal twins. Arch Gen Psychiatry 1973; 29: 465–469.
9 Cohen DJ, Dibble E, Grawe JM, Pollin W. Reliably separating identical from fraternal twins. Arch Gen Psychiatry 1975; 32: 1371–1375.
10 Martin NG, Martin PG. The inheritance of scholastic abilities in a sample of twins. I. Ascertainment of the sample and diagnosis of zygosity. Ann Hum Genet 1975; 39: 213–218.
11 Kasriel J, Eaves L. The zygosity of twins: Further evidence on the agreement between diagnosis by blood groups and written questionnaires. J Biosoc Sci 1976; 8: 263–266.
12 Sarna S, Kaprio J, Sistonen P, Koskenvuo M. Diagnosis of twin zygosity by mailed questionnaire. Hum Hered 1978; 28: 241–254.
13 Torgersen S. The determination of twin zygosity by means of a mailed questionnaire. Acta Genet Med Gemellol (Roma) 1979; 28: 225–236.
14 King M-C, Friedman GD, Lattanzio D, Rodgers G, Lewis AM, Dupuy ME, Williams H. Diagnosis of twin zygosity by self-assessment and by genetic analysis. Acta Genet Med Gemellol (Roma) 1980; 29: 121–126.
15 Sarna S, Kaprio J. Use of multiple logistic analysis in twin zygosity diagnosis. Hum Hered 1980; 30: 71–80.
16 Magnus P, Berg K, Nance WE. Predicting zygosity in Norwegian twin pairs born 1915-1960. Clin Genet 1983; 24: 103–112.
17 Bønnelykke B, Hauge M, Holm N, Kristoffersen K, Gurtler H. Evaluation of zygosity diagnosis in twin pairs below age seven by means of a mailed questionnaire. Acta Genet Med Gemellol (Roma) 1989; 38: 305–313.
18 Eisen S, Neuman R, Goldberg J, Rice J, True W. Determining zygosity in the Vietnam Era Twin Registry: an approach using questionnaires. Clin Genet 1989; 35: 423–432.
19 Ooki S, Yamada K, Asaka A, Hayakawa K. Zygosity diagnosis of twins by questionnaire. Acta Genet Med Gemellol (Roma) 1990; 39: 109–115.
20 Ooki S, Yamada K, Asaka A. Zygosity diagnosis of twins by questionnaire for twins' mothers. Acta Genet Med Gemellol (Roma) 1993; 42: 17–22.
21 Spitz E, Moutier R, Reed T, Busnel MC, Marchaland C, Roubertoux PL, Carlier M. Comparative diagnosis of twin zygosity by SSLP variant analysis, questionnaire, and dermatoglyphic analysis. Behav Genet 1996; 26: 55–63.
22 Charlemaine C, Duyme M, Aubin J-T, Guis F, Marquiset N, de Pirieux I, Strub N, Brossard Y, Jarry G, Le Groupe Romulus, Frydman R, Pons JC. Twin zygosity diagnosis by mailed questionnaire below age twelve months. Acta Genet Med Gemellol (Roma) 1997; 46: 147–156.
23 Chen WJ, Chang H-W, Wu M-Z, Lin CCH, Chang C, Chiu Y-N, Soong W-T. Diagnosis of zygosity by questionnaire and polymerase chain reaction in young twins. Behav Genet 1999; 29: 115–124.
24 Boomsma DI. Twin registers in Europe: an overview. Twin Research 1998; 1: 34–51.
25 Boomsma DI, Orlebeke JF, Van Baal GCM. The Dutch twin register: growth data on weight and height. Behav Genet 1992; 22: 247–251.
26 Van Dijk BA, Boomsma DI, De Man AJM. Blood group chimerism in human multiple births is not rare. Am J of Med Genet 1996; 61: 264–268.
27 Huberty CJ. Applied Discriminant Analysis. John Wiley & Sons: New York, 1994.
28 Panel on Discriminant Analysis, Classification, and Clustering. Discriminant analysis and clustering. Stat Sci 1989; 4: 34–69.
29 Kendler KS, Neale MC, Kessler RC, Heath AC, Eaves LJ. A test of the equal-environment assumption in twin studies of psychiatric illness. Behav Genet 1993; 23: 21–27.

Appendix 1 Summary of studies on zygosity determination by written questionnaire

Study	Subjects	Mailed questionnaire	Method of classification[1]	Results
Cederlöf, Friberg, Jonsson, Kaij, 1961[3]	200 pairs, age 35–75	1 similarity item,[2] 1 multivariate,[3] confusion item; completed by both twins	decision rules	98% of MZ correct; 92% of DZ correct; 10% of total sample left unclassified
Nichols, Bilbro, 1966[4]	123 pairs, high school juniors	5 similarity items and 1 multivariate confusion item; completed by both twins	decision rules; intuitive decision was made in case the previous method left cases unclassified (7%)	93% of total sample correct
Jablon, Neel, Gershowitz, Atkinson, 1967[5]	232 pairs, age 30–45	A short description of 'identical' and 'non-identical' was given by the investigators, followed by one single item that dealt with twins' own opinion; completed by both twins (complete agreement within pair) or individual twins	Evaluation of zygosity diagnosis was performed on one item only: the joint opinion of a pair, and the opinion of the individual twin	No difference in accuracy, between individual twins and pairs. 89% of MZ correct, 97% of DZ correct
Hauge, Harvald, Fischer, Gotlieb-Jensen, Juel-Nielsen, Raebild, Shapiro, Videbech, 1968[6]	335 pairs, adults	Not clearly specified: multiple similarity items as well as 1 multivariate confusion item; completed by one twin or both twins, or by relatives	decision rules	97% of total sample correct
Schoenfeldt 1969[7]	124 pairs, sample is identical to Nichols, Bilbro, 1966	Identical to Nichols and Bilbro, (1966)	decision rules based on one single score calculated from scores of both twins; discriminant analyses on same single score	decision rules: 92% of total sample correct (cross-validated 79%); discriminant: 88% of total sample correct (cross-validated 88%)
Cohen, Dibble, Grawe, Pollin, 1973[8]	Two samples: (1) 120 pairs, mean age 9.4 (2) 35 pairs mean age 4.2	7 similarity items and 1 multivariate confusion item, completed by the mother. Samples differed in age and in knowledge of zygosity by the mother	discriminant analyses; cutting point on summed raw scores	No difference in response pattern between groups varying in age and informed mothers. Groups were pooled; discriminant: 98% of total sample correct; cutting point: 93% of MZ correct and 73% of DZ correct, with the remaining left unclassified
Cohen, Dibble, Grawe, Pollin, 1975[9]	275 pairs, age 1–6	Identical to Cohen et al, (1973), completed by the mother	discriminant analyses; cutting point on summed raw score; principal component factor analysis	hit rate is estimated at 90%
Martin, Martin, 1975[10]	47 pairs, age 15	A description of 'identical' and 'non-identical' was given by the investigator, followed by one single item that dealt with the twins' own opinion; their joint answer had to be confirmed by the parents	Since parents and twins all had to agree on the zygosity of the pair, evaluation of zygosity diagnosis was performed on one item only	100% of total sample correct
Kasriel, Eaves, 1976[11]	178 pairs, adults	1 similarity item and 1 univariate[3] confusion item: completed by both twins	decision rules	96% of total sample correct
Sarna, Kaprio, Sistonen, Koskenvuo, 1978[12]	104 pairs, age 20–69	1 similarity item and 1 univariate confusion item; completed by both twins	deterministic decision tree	93% of total sample correct with 7% unclassified
Torgersen, 1979[13]	215 pairs, age 18–67	1 similarity item and 1 multivariate confusion item; completed by both twins	cutting point on single summed raw score composed of scores of both twins; discriminant analyses on same summed raw score; decision tree	cutting point: 95% of total sample correct; discriminant: 94% of MZ correct, 96% of DZ correct; decision tree: 96% of total sample correct

Continued on next page

Appendix 1 continued from previous page

Study	Subjects	Mailed questionnaire	Method of classification[1]	Results
King, Friedman, Lattanzio, Rodgers, Lewis, Dupuy, Williams, 1980[14]	173 pairs, adults	1 similarity item that dealt with twins' own opinion; completed by both twins	Evaluation of zygosity diagnosis was performed on one item only	83% of MZ correct, 97% of DZ correct
Sarna, Kaprio, 1980.[15] This study is a follow-up of Sarna et al, 1978.[12]	Two samples: (1) 52 pairs previously left unclassified (2) 104 pairs	Identical to Sarna et al, (1978), completed by both twins	logistic regression, with (1) 0.50 and (2) 0.70 limit for a posteriori probability discriminant analyses	logistic regression: (1) all cases classified with 75% correct of total sample, cross-validated, (2) 100% correct of total sample with 53% left unclassified, cross-validated; discriminant: identical results
Magnus, Berg, Nance, 1983[16]	207 pairs, age 33–61	Originally[4] composed of 13 similarity items, 1 multivariate confusion item, and 1 item reflecting twins' own opinion: completed by one twin or both twins	discriminant analyses applied to 2 groups: (1) data from one twin only, (2) data from both twins. Intrapair means of scores was used in case both twins responded	(1) 96% of total sample correct, cross-validated, (2) 98% of total sample correct, cross-validated
Bønnelykke, Hauge, Holm, Kristoffersen, Gurtler, 1989[17]	125 pairs, age 0.5–6.5	4 similarity items and 1 univariate confusion item; completed by the mother	decision rules	91% of total sample correct, 4% misclassified, and 5% left unclassified
Eisen, Neuman, Goldberg, Rice, True, 1989[18]	4774 male pairs with insufficient blood typing data, adults	Identical to Magnus, (1983), completed by both twins	discriminant analyses as employed by Magnus (1983); 3 types of logistic regression including race-specific analysis	By combining the various methods, 9% of MZ twins were classified incorrectly. Variation in discriminating questions was observed for race
Ooki, Yamada, Asaka, Hayakawa, 1990[19]	Two samples: (1) 189 pairs age 12–16; (2) 93 pairs age 52–77	Identical to Torgersen, (1979), completed by both twins	cutting point on single summed raw score composed of scores of both twins; discriminant analyses on same summed raw score	cutting point: (1) 92% of MZ correct, 88% of DZ correct, (2) 100% of MZ correct, 77% of DZ correct; discriminant: (1) 92% of total sample correct, cross-validated in older sample resulted in 95% correct, (2) 94% of total sample correct, cross-validated in younger sample resulted in 67% correct
Ooki, Yamada, Asaka, 1993[20]	74 pairs, high-school age	Identical to Torgersen, (1979), completed by both twins and by the mother	cutting point on single summed raw score composed of (1) scores of both twins, and of (2) scores by mother	(1) 98% of MZ correct, 77% of DZ correct, (2) 93% of MZ correct, 92% of DZ correct
Spitz, Moutier, Reed, Busnel, Marchaland, Roubertoux, Carlier, 1996[21]	79 pairs, age 8–12.5	Adapted from Goldsmith, (1991), originally composed of 18 items, completed by one parent	cutting point on mean score obtained by summing raw scores and dividing by number of items answered; logistic regression	cutting point: 97% of total sample correct; logistic regression: 92% of total sample correct

Continued on next page

Appendix 1 continued from previous page

Study	Subjects	Mailed questionnaire	Method of classification[1]	Results
Charlemaine, Duyme, Aubin, Guis, Marquiset, De Pirieux, Strub, Brossard, Jarry, Le Group Romulus, Frydman, Pons, 1997[22]	76 pairs, age < 1	Adapted from Bønnelyke et al, (1989), originally composed of 26 items; completed by one parent or both parents together	decision rules, various approaches; cutting point on summed raw score	decision rules: ranging from 87% to 99% of total sample correct; cutting point: 96% of total sample correct
Chen, Chang, Wu, Lin, Chang, Chiu, Soong, 1999[23]	Two samples: (1) 105 pairs age 12–16, (2) 47 pairs age 2–12	Adapted from Cohen et al, (1975), Goldsmith, (1991), and culture-specific items. Originally composed of 20 (parental report) and 27 (self report) items; completed by (1) both parents and both twins, (2) one parent	logistic regression; cutting point on 3-item profiles for (1) only	logistic regression: (1) 97% of total sample correct by parental report, 96% of total sample correct by self report; (2) 93% of total sample correct; cutting point: (1) identical to logistic regression

[1]Each study compares the assignment of zygosity based on questionnaire to the classification obtained through blood polymorphism or DNA markers, or a combination of both. [2]The question 'are twins alike as two peas in a pod?' is considered a similarity item. [3]Univariate versus multivariate: this reflects the number of sub-questions that deal with confusion of twin identity. Univariate: the occurrence of twin confusion is limited to one type of person, for instance 'strangers'. Multivariate: the occurrence of twin confusion by multiple types of persons, like 'parents', 'family members', 'teachers', etc. [4]'Originally' implies that the final analyses were performed on a reduced number of items.

Appendix 2 Translation of zygosity questionnaire, sent to parents when twins reach the age of 6

How much are the twins alike with respect to:				
1	Facial appearance	not	somewhat	exactly
2	Hair colour	not	somewhat	exactly
3	Face colour	not	somewhat	exactly
4	Eye colour	not	somewhat	exactly
5	Are they as alike as two peas in a pod?	no	yes	
6	Does the mother or father mistake one for the other?	no	yes	
7	Do other family members mistake one for the other?	no	yes	
8	Do strangers have difficulty telling them apart?	no	yes	
At age 8 and 10 of the twins, two more questions are added				
9	Do you have difficulty in correctly identifying each twin on new photographs?	no	yes	
10	Do the twins have the same hair structure?	not	somewhat	exactly

Appendix 3 Unstandardised canonical discriminant function coefficients, constants and classification score to construct the classification rule

Item	Age 6 Mother	Age 8 Mother	Age 8 Father	Age 10 Mother	Age 10 Father
Facial appearance	0.618128	0.424786	0.546325	0.166356	0.522894
Hair colour	0.431205	0.562038	0.385539	0.465518	0.176443
Face colour	0.521933	0.059957	0.156256	0.170350	0.218696
Eye colour	0.252118	0.242795	0.271036	0.192224	0.119514
Two peas	0.349174	0.329923	0.190973	0.086300	0.165164
Mother/father	0.025022	0.086795	–0.10002	0.061590	–0.00264
Family members	1.098133	0.343303	0.638154	0.825344	0.452154
Strangers	0.358312	0.432926	0.568857	1.054857	1.688902
Photograph		–0.10844	–0.03261	–0.07711	–0.26824
Hair structure		0.778413	0.601257	0.611719	0.459194
Constant	–7.30262	–6.58742	–6.76956	–6.92407	–6.68708

Items are rated 0, 1, or 2 on a three-point scale. Dichotomous items are rated 0 or 1. By multiplying each coefficient with the item score and summing these products with the constant, a zygosity score is obtained for each individual pair. This zygosity score is compared with the classification score that is generated by the discriminant function analysis. In this study, the classification score is 0.4 for each individual dataset. Pairs whose zygosity score is greater than 0.4 are assigned the label monozygotic, pairs with scores below this classification score are considered dizygotic.