Definitive Zygosity Scores in the Peas in the Pod Questionnaire is a Sensitive and Accurate Assessment of the Zygosity of Adult Twins

Zakariya A. Jarrar,[^] Kirsten J. Ward,[^] Massimo Mangino, Lynn F. Cherkas, Raj Gill, Irina Gillham-Nasenya, Darioush Yarand, Deborah Hart, Christopher J. Hammond,[*] and Tim D. Spector[*]
Department of Twin Research and Genetic Epidemiology, King's College London, London, UK

Twin researchers face the challenge of accurately determining the zygosity of twins for research. As part of the annual questionnaire between 1999 and 2006, 8,307 twins from the TwinsUK registry were asked to complete five questions (independently from their co-twin) to ascertain their self-perceived zygosity during childhood on up to five separate occasions. This questionnaire is known as the 'peas in the pod' question-naire (PPQ), but there is little evidence of its validation. Answers were scored and classified as monozygotic (MZ), dizygotic (DZ), or unknown zygosity (UZ) and were compared with 4,484 twins with genotyping data who had not been selected for zygosity. Of these, 3,859 individuals (46.5% of those who had a zygosity from PPQ) had zygosity classified by both the PPQ and genotyping. Of the 708 individual twins whose answers meant that they were consistently classed as MZ in the PPQ, 683 (96.5%) were MZ within the genotype data. Of the 945 individual twins consistently classed as DZ within questionnaire, 936 (99.0%) were DZ in the genotype data. Where both twins scored MZ consistently across multiple questionnaires, 99.6% were MZ on genotyping, 99.7% were DZ on genotyping if both twins consistently scored DZ. However, for the initial questionnaire, 88.6% of those scoring as MZ were genotypically MZ and 98.7% DZ. For twin pairs where both scored UZ, 94.7% were DZ. Using the PPQ on a single occasion provided a definitive classification of whether the twin was MZ or DZ with an overall accuracy of 86.9%, increasing to 97.9% when there was a consistent classification of zygosity across multiple questionnaires. This study has shown that the PPQ questionnaire is an excellent proxy indicator of zygosity in the absence of genotyping information.

■ **Keywords:** twins, twin studies, zygosity, questionnaire, genome-wide association study, next-generation sequencing, genotyping techniques, gene expression analysis

A common misunderstanding, until relatively recently, was that the presence of two placentae at birth must mean that twins are non-identical or dizygotic (DZ), while one placenta meant that the twins are identical or monozygotic (MZ; Ooki et al., 2004). Opposite-sex twins are DZ, but there are issues arising when using this methodology to determine the zygosity of same-sex twins. This meant that between 20% and 25% of MZ twins with two separate placentae were misclassified as DZ at birth, with approximately 9% of DZ twins with placental fusing wrongly classed as MZ (Ooki et al., 2004). As a result, a number of twin pairs have grown to adulthood understanding that they are identical or non-identical when they are not. This misclassification of zygosity at birth has implications on not only the medical treatment of twins during gestation, after their birth, and in later life, but also has personal implications on the twins who may grow up questioning their zygosity.

Within the scientific community, an essential component of any twin registry is the knowledge of the zygosity of the twins registered. There are a number of different methods of determining the zygosity in adult twins. These include self-report methods (e.g., asking the twins whether they are identical or non-identical), asking twins questions on how similar they are or were during childhood and growing up (Bønnelykke et al., 1988; Cederlöf et al.,

1961; Magnus et al., 1983; Ooki et al., 1989; Sarna et al., 1978; Song et al., 2010), or asking the parents to report on the similarity of twins (Peeters et al., 1998).

Determination of zygosity by self-report raises concerns that misclassification may have occurred at birth by an authoritative medical professional. More reliable methods include testing by blood group (Cederlöf et al., 1961; Kasriel & Eaves, 1976; Lykken, 1978) or using genetic tests (Song et al., 2010) but these may not necessarily be fast or cost-effective. Song et al. (2010) used 16 short tandem repeat genetic markers and demonstrated that using questions about similarities during childhood provided a sensitivity of 98.8% for MZ twin pairs and 88.9% for DZ twin pairs.

Genotypic data is the 'gold standard' for determining the zygosity of twins, but as this can be costly and time consuming when testing a large-scale study at one time, alternative methods of accessing zygosity are necessary, and as a result, latent class analysis has been used with good success (Eaves et al., 1993; Heath et al., 2003).

The TwinsUK Adult Twin Registry started in 1992 (Moayyeri, Hammond, Hart et al., 2013, Moayyeri, Hammond, Valdes et al., 2013; Spector & Williams, 2006) and, as with all twin registries, has faced the challenge of determining the zygosity of the twins. The aims of this study were to compare the accuracy of zygosity determined by a 'peas in the pod' similarity questionnaire (the PPQ) for both single twins and twin pairs with the zygosity determined by genotype data, the 'gold standard', and to examine the consistency of responses when the PPQ is administered annually compared to initial self-report and PPQ.

Materials and Method

Using a short similarity questionnaire from the Australian Twin Registry as a basis (unpublished and known as the Peas in the Pod Questionnaire or PPQ), the TwinsUK registry has adapted this to be more specific to the cohort. Between 1999 and 2006, the TwinsUK registry asked same-sex twins (aged 18–89 years of age) the PPQ (see Supplementary Material 1) as part of their annual questionnaire. The accompanying instructions asked for the questionnaire to be completed by each twin separately to ensure that the twins were not influenced in the answers they gave by their co-twin. Written, informed consent was obtained from participants of the study and all procedures contributing to this work comply with the ethical standards in the Helsinki Declaration of 1975, as revised in 2008.

Self-Reported Zygosity

Upon registration with the TwinsUK registry, each twin was asked to report what they believed their zygosity to be. A total of 13,291 twins had provided a self-reported zygosity: age 18 to 84, mean age of 58; 10,796 (83%) were female, 2,226 were male, and 269 had no gender assigned to them.

Peas in the Pod Questionnaire (PPQ)

The PPQ is a five-item questionnaire on the degree of similarity between twins (Supplementary Material 1). It asks four questions based on whether at school-age people at school, parents, close-friends, or strangers had difficulty telling twins apart (0 = yes, 1 = don't know, 2 = no), and a fifth question about whether the twins would be described during childhood as alike as two peas in a pod (0 points), as alike as ordinary siblings (2 points), or they didn't know (1 point). Scores between 0 and 4 were classed as MZ, scores between 8 and 10 were classed as DZ, and anything in between was scored as unknown zygosity (UZ), with each twin's PPQ scored separately.

A total of 8,307 twins answered a PPQ zygosity between 1999 and 2006: age 18 to 87, mean age of 51; 7,287 (88%) were female, and 1,020 were male.

Initial analysis looked at the zygosity provided from the first questionnaire, and then the individual twins were matched to their co-twin. The second stage of analysis looked at all data collected from the PPQ between 1999 and 2006. The PPQ was scored separately for each questionnaire on five separate occasions so that the scores for each questionnaire, and therefore the zygosity of the twin for each questionnaire were independent of the scores and zygosity from the other PPQs.

To create an overall zygosity over time, individuals who consistently had the same zygosity for each of the questionnaires were scored as that zygosity (consistent answers of MZ meant an overall zygosity of MZ, consistent answers of DZ meant an overall zygosity of DZ, and consistent answers of UZ meant an overall zygosity of UZ), and individuals whose zygosity was inconsistent over time — for example, MZ in one questionnaire and DZ in another — were scored as UZ.

The overall zygosity on an individual and paired basis was determined from all of the five PPQs. If both of the twins had the same overall zygosity, they were scored as this zygosity (i.e., if both had an overall zygosity of DZ, then the pair was scored as DZ) and if the twins disagreed within the pair, this was noted so that it would be possible to see which zygosity was correct after comparison with the zygosity determined via genotyping.

To determine the weight of each question within the PPQ, the final stage was to look at each individual question within the PPQ and score the answer to each question as MZ if the answer scored 0, UZ if the answer scored 1, and DZ if the answer scored 2.

To assess the sensitivity of the PPQ, a standard true positive rate calculation was used and to assess the specificity of the PPQ, a standard true negative rate calculation was used.

Genotype Data

As MZ twins share 100% of their segregating genes, only one MZ twin of a pair was routinely genotyped for

genome-wide association study (GWAS), whereas both members of a DZ or UZ pair were genotyped.

TwinsUK samples were genotyped with the Infinium 317K and 610K assays (Illumina, San Diego, USA) at two different centers, namely, the Wellcome Trust Sanger Institute and the Center for Inherited Diseases Research (USA), respectively. The normalized intensity data were pooled and the genotypes were called on the basis of the Illumina algorithm. No calls were assigned if the most likely call was less than a posterior probability of 0.95. Validation of pooling was done by visual inspection of 100 random, shared single-nucleotide polymorphisms (SNPs) for overt batch effects; none were observed. We excluded SNPs that had a call rate <97% (SNPs with minor allele frequency (MAF) ≥5%) or <99% (for 1% ≤ MAF <5%), Hardy–Weinberg p values $<10^{-6}$ and MAFs <1%. We also removed subjects where genotyping failed for >2% of SNPs. The overall genotyping efficiency of the genome-wide association (GWA) was 98.7%.

A total of 4,484 twins had zygosity determined via genotyping: age 20 to 90, mean age of 61; 4,136 (92%) were female, and 348 were male.

We computed identity-by-descent (IBD) estimates for all available pair of twins ($n = 4,484$ individuals) using the Plink (Purcell et al., 2007) option on a set of 9,357 SNPs (not in linkage disequilibrium, with a minor allele frequency > 20%, and overlapping among all Illumina platforms available). For ambiguous cases, IBDs were subsequently recalculated by using all the SNPs available. We defined as MZ those twin pairs with a p-hat value < 0.9 and DZ the twin pairs with a p-hat value ranging between 0.4 and 0.6.

The Results From the PPQ

The PPQ has been used by the TwinsUK registry since the start of the registry in 1992, but for the purpose of this study, we concentrated on the data that were available over a 7-year period of time (between 1999 and 2006), where the PPQ was asked within the annual questionnaire on five separate occasions. It was possible to determine the zygosity for 8,307 individuals who had answered the PPQ on at least one occasion.

Results

Since the start of the TwinsUK registry in 1992, 13,291 twins have provided a self-reported zygosity (including 6,644 complete twin pairs). Of these, 6,129 (46.1%) reported that they were MZ, 6,359 (47.8%) reported that they were DZ, and 803 (6.1%) reported that they were UZ.

Zygosity From Individual Twins

A total of 8,307 individual twins answered the PPQ. Looking at the first PPQ answered by each twin separately, 4,038 (48.6%) individuals had scores that indicated that they were MZ, 1,645 (19.8%) provided answers that indicated that

they were DZ, and 2,624 (31.5%) were graded as UZ as they scored 5–7 on the PPQ (see Figure 1).

Examining repeated PPQs, 3,562 (42.9%) of the twins consistently scored as MZ and 1,536 (18.5%) of the twins consistently were rated as DZ on PPQ scores. The remaining 3,209 (38.6%) of twins' answers did not always result in the same grade (MZ, DZ, and UZ) over time ($n = 3,062$, 36.9%) or were scored consistently as UZ within the PPQ ($n = 147$, 1.8%) (see Figure 1).

Zygosity Within Twin Pairs

From the 8,307 individuals who had answered the PPQ, there were 3,697 complete twin pairs (7,394 individuals). Of these, 1,387 pairs (37.5%), both scored that they were MZ, 480 pairs (13.0%) both scored that they were DZ, and 1,150 (31.1%) both scored that they were UZ. The remaining 680 (18.4%) pairs could not agree on their zygosity. Of these, 10 (1.5%) had one twin scoring as MZ and the other twin scoring as DZ, 379 (55.7%) where one twin scored as MZ and the other twin as UZ, and 291 (42.8%) where one twin scored as DZ and the other twin scored as UZ.

Comparison of Zygosity From the PPQ and Genotyping Data

Of the 4,484 individuals from the TwinsUK registry with zygosity determined by genotype data, 1,106 (24.7%) were MZ and 3,378 (75.3%) were DZ. When the zygosity obtained from the PPQ was matched with the 'true' zygosity from the genotyping data, there were 3,859 twins with zygosity for both methodologies, which included 1,806 complete twin pairs (see Figure 1).

Zygosity From Individual Twins

The zygosity from the first PPQ was matched with the genotyping data, 943 twins answered that they were MZ in the PPQ, of whom 735 (77.9%) were MZ in the genotyping data; 1,101 answered that they were DZ, of whom 1,071 (97.3%) were DZ in the genotyping data. From the 1,811 twins who were UZ from their first PPQ, 1,665 (91.9%) were DZ from the genotyping data (see Figure 2). Using the zygosity obtained from the first PPQ proved to have 96.1% sensitivity and 83.7% specificity.

Comparing the overall zygosity from the PPQ and genotype data, there were 3,859 individuals who had zygosity determined by both methodologies (providing coverage of 46.5% of the PPQ zygosity results; see Table 1). Using the overall zygosity obtained from all of the PPQs proved to have 98.6% sensitivity and 97.4% specificity.

Seven hundred and eight individual twins consistently indicated that they were MZ in the PPQ, of whom 683 (96.5%) were MZ in the genotype data. Nine hundred and forty-five twins consistently indicated that they were DZ in the PPQ, of whom 936 individual twins (99.0%) were DZ in the genotype data. From the 2,206 twins who were classed
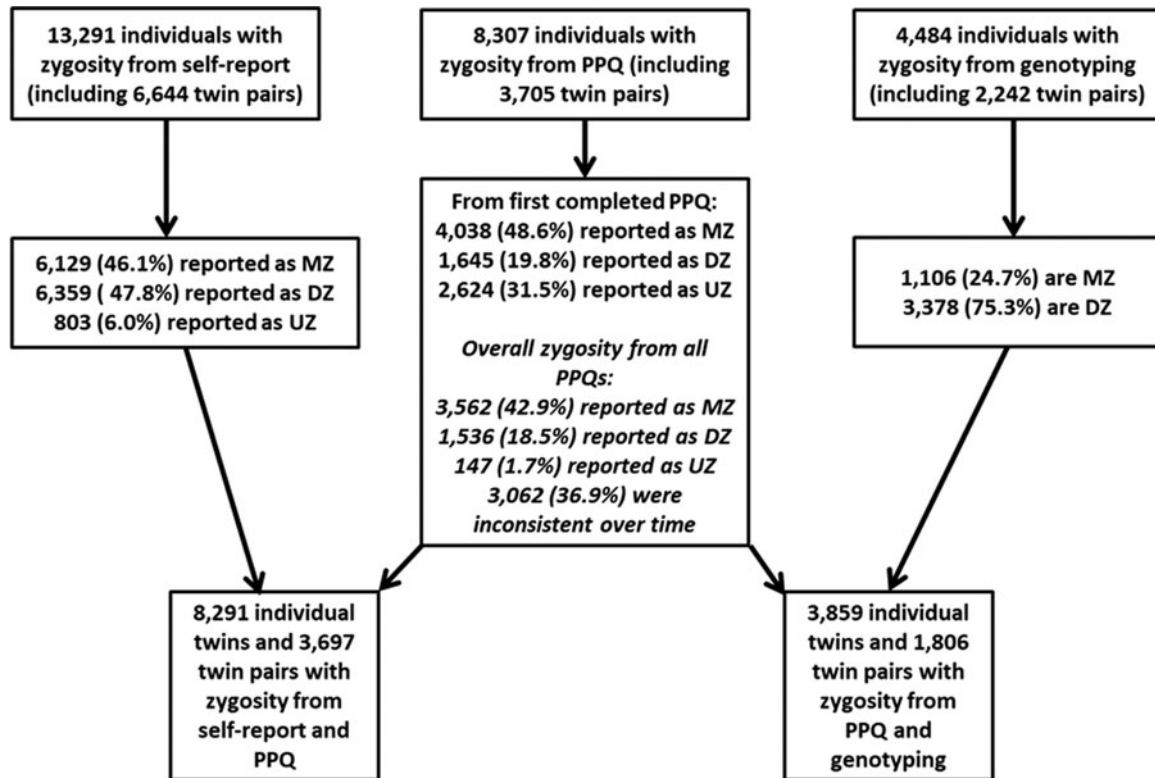
FIGURE 1
Flow diagram of zygosity data available from self-report, PPQ, and genotyping.
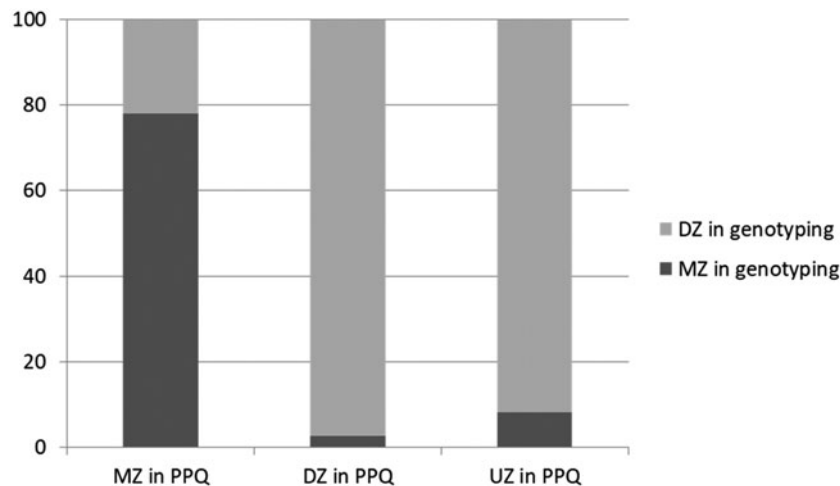

FIGURE 2
Percentage accuracy of the zygosity from the first PPQ with the zygosity from the genotyping data for individual twins.

as UZ from the PPQ, 1,987 (90.1%) were DZ in the genotype data.

The first time that the PPQ is completed by an individual twin and compared to the genotyping data, there is 77.9% accuracy if the score is MZ, 97.3% accuracy if the score is DZ, and when the score indicates a UZ, there is a 91.9% chance that they are in fact a DZ twin. This accuracy in-

creases to 88.6% for MZ and 98.7% for DZ when both twins' scores indicate that they are MZ or DZ, respectively.

Zygosity Within Twin Pairs

Taking the answer from both twins from the first PPQ that they had answered showed that there were 343 pairs who both agreed that they were MZ, of whom 304 (88.6%) were
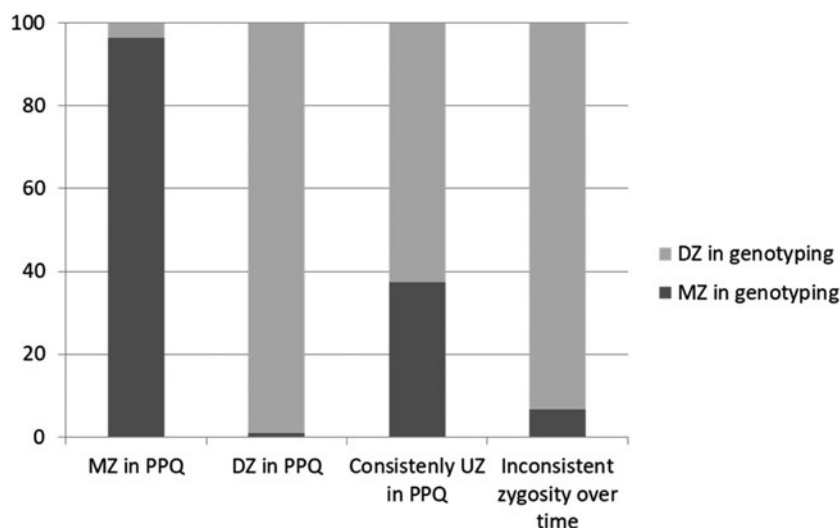
FIGURE 3

Percentage accuracy of the overall zygosity from all PPQs with the zygosity from the genotyping data for genotyping data for twin pairs.

TABLE 1

Comparison of Zygosity Determined by First PPQ and Genotype Data for Twin Pairs

First PPQ answer	Number of pairs in PPQ	Number of pairs who are MZ in GWAS (%)	Number of pairs who are DZ in GWAS (%)
Both answer monozygotic (MZ)	343	304 (88.6%)	39 (11.4%)
Both answer dizygotic (DZ)	390	5 (1.3%)	385 (98.7%)
Both answer unknown zygosity (UZ)	703	37 (5.3%)	666 (94.7%)
One answers MZ and one answers DZ	36	10 (27.8%)	26 (72.2%)
One answers MZ and one answers UZ	148	60 (40.5%)	88 (59.5%)
One answers DZ and one answers UZ	186	4 (2.2%)	182 (97.8%)

MZ in the genotype data (see Table 1). From the 390 pairs who both answered that they were DZ, 385 (98.7%) were DZ in the genotyping data (see Figure 3).

For twin pairs where it was not possible to determine the zygosity, 703 had both scored as UZ, and 666 (94.7%) of these were DZ from the genotype data and 370 did not agree on their zygosity, of whom 296 (80%) were DZ in the genotype data.

Looking at the overall PPQ scores from all questionnaires, there were 274 twin pairs where both scored as MZ consistently, of whom 273 (99.6%) were MZ in the genotype data. From the 312 twin pairs where both twins consistently scored as DZ within the PPQ, 311 (99.7%) were DZ in the genotype data. There were 901 twin pairs where it was not possible to ascertain the zygosity of either twin from the PPQ (both twins scored UZ, 5–7 points). From these UZ pairs, 837 (92.9%) of the twins were DZ in the genotype data (see Table 2).

There were a number of twin pairs where only one twin had scored as a consistent zygosity within the PPQ over time. When one twin consistently scored as MZ in the PPQ and the other twin as UZ, comparison with the genotype data showed that the MZ twin score was correct on 87.8% occasions. When one twin consistently scored as DZ in the PPQ over time and the other twin as UZ, comparison with the genotype data showed that the DZ twin score was correct on 99.1% occasions.

Weighting of the Individual Questions Within the PPQ

The answer for each question within the questionnaire was scored as MZ, DZ, or UZ and compared with the result on zygosity obtained from the genotyping.

For determining both MZ and DZ twins (see Figure 4), the most accurate question was question (e): 'In childhood, which of the following would best describe you and your twin?' This had an accuracy of 92.5% for MZ twins and an accuracy of 97.2% for DZ twins when compared to the genotyping data. The least accurate question for MZ twins was question (a): 'At school, did people have trouble telling you apart?', with an accuracy of 48.5%, and for DZ twins, question (b): 'Were your parents able to tell you apart?', with an accuracy of 80.3%.

These results suggest that it is possible to ascertain the zygosity of both MZ and DZ twins using just question (e): 'In childhood, which of the following would best describe you and your twin?', with the possible answers being 'As alike as two peas in a pod', 'Ordinary sibling likeness (Like

TABLE 2

Comparison of Overall Zygosity Determined by Multiple PPQs and Genotype Data for Twin Pairs

| Zygosity from PPQ	Number of pairs in PPQ	Number of pairs who are MZ in genotype data	Number of pairs who are DZ in genotype data
Both answer monozygotic (MZ)	274	273 (99.6%)	1 (0.4%)
Both answer dizygotic (DZ)	312	1 (0.3%)	311 (99.7%)
Both answer unknown zygosity (UZ)	901	64 (7.1%)	837 (92.9%)
One answers MZ and one answers DZ	5	1 (20.0%)	4 (80.0%)
One answers MZ and one answers UZ	90	79 (87.8%)	11 (12.2%)
One answers DZ and one answers UZ	224	2 (0.9%)	222 (99.1%)
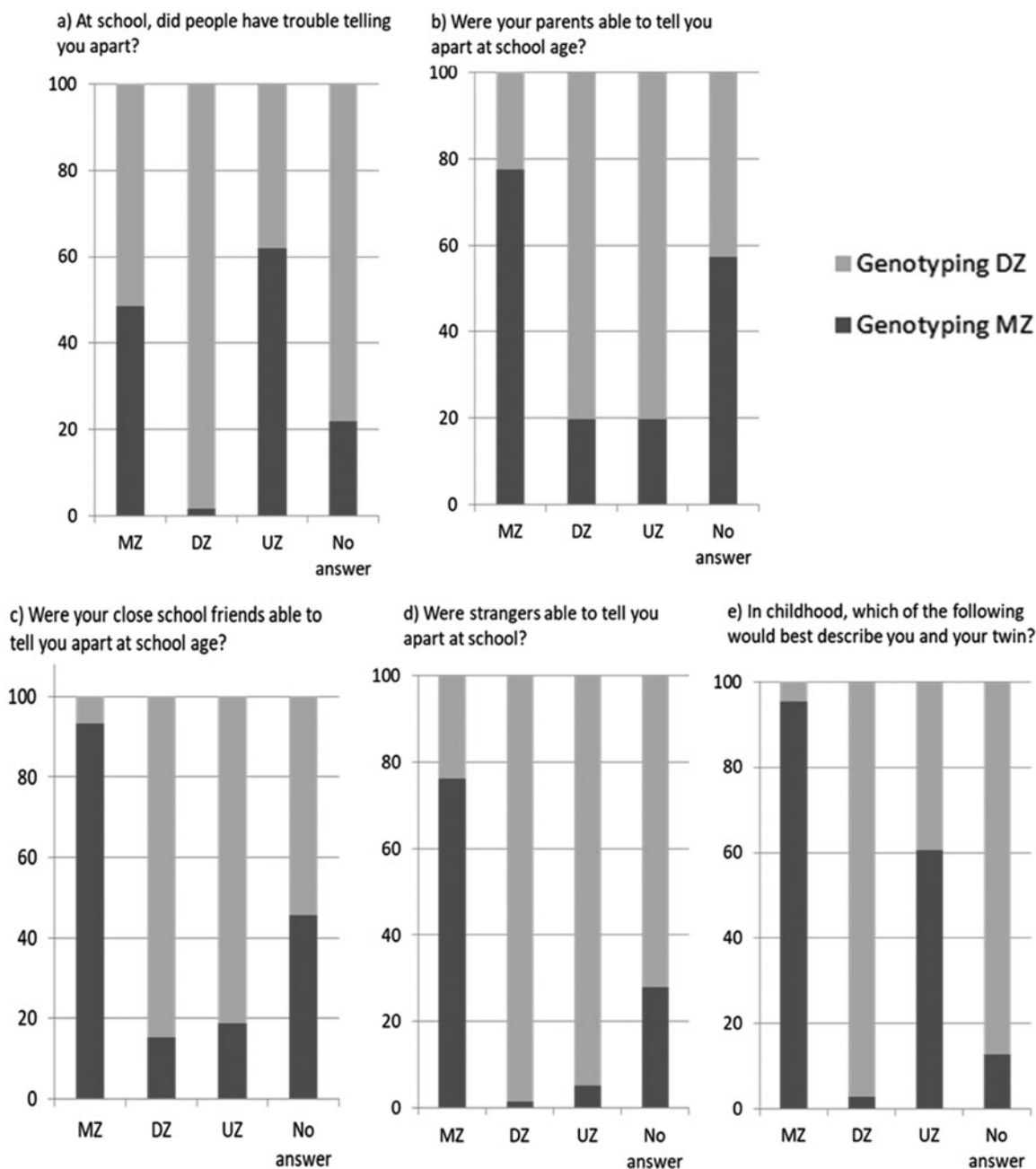
FIGURE 4

Percentage agreement between zygosity from PPQ and genotyping data for each question.
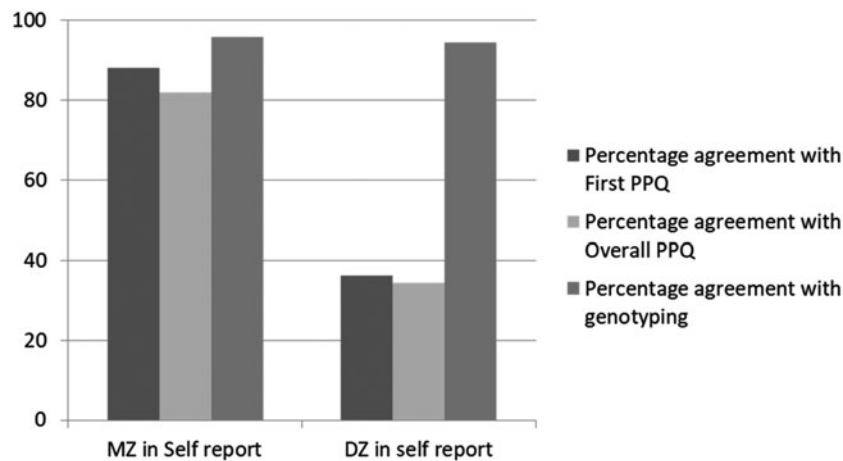
FIGURE 5
Percentage agreement between zygosity from self-report, first time answering the PPQ, overall PPQ, and genotyping data.

sisters or brothers)', and 'I don't know' (see Supplementary Material 1).

Comparison Between Self-Reported Zygosity, Zygosity From PPQ and Genotyping Data

Eighty-eight percent of the twins who self-reported that they were MZ at registration were also MZ within the first PPQ that they answered (see Figure 5). Eighty-two percent of the self-reported MZs remained MZ for the overall zygosity, and 95.9% of those who self-reported as MZ remained as MZ within the genotyping data. However, reporting that they were DZ at registration did not appear to be as consistently accurate, as 36.3% were DZ within the first PPQ that they answered, and 34.3% were DZ with the overall zygosity from all of the PPQs that they answered. However, 94.4% of those who had self-reported as DZ were also DZ within the genotyping data.

Discussion

It may well be that the historical misclassification of zygosity according to number of placentae explains the discrepancies between the self-report zygosity and genotyping results, particularly in our cohort of twins born before human genomic testing was practical. Ninety-five percent of individuals who self-reported as MZ and 94.4% as DZ had their zygosity confirmed by genotyping. In addition to self-report data, the TwinsUK registry has used the PPQ as an initial indicator of zygosity, and in the main now has the luxury of genetic techniques to confirm the zygosity of a twin pair.

Traditionally, genetic determination of zygosity has been seen as costly, particularly in large-scale research studies. However, with the reduction in costs of genotyping in recent years, it is not expensive when compared to the pregnancy and childbirth costs of twin pregnancies. We would

certainly advocate routine genotypic testing of twins at birth to allow families to have definitive zygosity ascertained.

Interestingly, although the PPQ was designed to improve classification of zygosity from self-report, the score obtained from the first time that the PPQ is answered appears to be less predictive of true zygosity — only 77.9% who were scored from PPQ as MZ were confirmed as MZ from the genotyping data, compared to 97.3% genetically confirmed as DZ from those who were scored as DZ.

Indecision about zygosity in questionnaire studies has previously been suggestive that twins are likely to be DZ (Kasriel and Eaves, 1976). Our results are similar, which show that 92.9% of twins whose answers scored as UZ on multiple questionnaires were found to be DZ. This percentage is higher than a previous study where both twins answered as UZ (Song et al., 2010) and could be explained by the fact that participants in our study were asked to complete multiple questionnaires, which may have increased the accuracy of the results. Eighty-eight percent of twin pairs who answered differently to each other were found to be DZ. As the questionnaire was asked on several occasions, participants in our study were instructed not to confer with their fellow twin about their perceived zygosity so as not to influence their answers, and we did not ask them to come to any pair-wise decisions on zygosity.

There is variation in the literature about the accuracy of questionnaire-based zygosity. In agreement with our study, some studies confirm DZ more frequently than MZ on questionnaire (Peeters et al., 1998). However, other studies confirm MZ more than DZ (Cederlöf et al., 1961). While questionnaires used in different studies contain similar components, there is no standardized questionnaire used between studies, which may account for some of the differences. Also, older studies used techniques, such as blood group, to determine zygosity, which may be less accurate than modern genotyping technologies, which have

near 100% accuracy (Chen et al., 2010). As discussed in the introduction, it may be that our particular twin population, on average born in the 1950s, was misinformed about zygosity due to misinformation from midwives and/or doctors, based on numbers of placentae. Finally, there may be a bias because often only one of the pairs who were consistently MZ on PPQ was genotyped, reducing the numbers of MZ pairs who would likely have been confirmed by genotyping.

The single question, 'In childhood, which of the following would best describe you and your twin?', where the answer was 'As alike as two peas in a pod' or 'Ordinary sibling likeness (like sisters or brothers)' was more predictive of the 'true' zygosity of the twins compared to the overall score obtained from all of the questions from the PPQ. The accuracy of the PPQ in our study was found to be 92.5% and 97.2% for MZ and DZ twins, respectively. This is very similar to results from Peeters' et al. (1998) study, which showed an accuracy of 92.8% and 97.1% for MZ and DZ twins, respectively. The highly reproducible results give weight to the validity of the PPQ in accurately diagnosing zygosity.

This study has demonstrated that consistency of the response to the same set of questions administered over a period of several years demonstrated even stronger predictability and that this response represents the 'true' zygosity. When both of the twins in the pair consistently agreed that they were MZ or DZ, the responses were their 'true' zygosity in 99.6% and 99.7% of twin pairs, respectively. Ninety-five percent of individuals who were categorized as UZ across longitudinal PPQs, or whose category changed across time, were DZ when genotyped. There were a number of limitations and biases present within this study. The study may not be generalizable to other twin populations as there is a high proportion of females within the TwinsUK cohort and a relatively older age of mean 61 years compared to other cohorts. The TwinsUK Cohort (Moayeri, Hammond, Hart et al., 2013) is predominantly female (~80%) as historically the initial focus of the study was osteoporosis, and so middle-aged female twin pairs were recruited since 1992; despite subsequent inclusion of men, like many twin cohorts, there is a female volunteer bias.

The TwinsUK Cohorts were born in the United Kingdom and so represent the cultural norms of their society and knowledge about twinning, and may be biased in that they are volunteers in a research cohort.

A further bias might occur through the fact that the questionnaire was asked on five separate occasions. Comparisons between the first time answering the questionnaire and the overall result after answering multiple questionnaires have shown that repeating the PPQ is more accurate than asking on a single occasion; however, there is the risk that the twins 'learnt' their true zygosity, and therefore the answers for questionnaires completed after this subsequently are a more accurate representation of their 'true' zygosity.

We have been able to compare the zygosity determined by the PPQ questionnaire with the zygosity determined via (GWAS) genotyping in a large number of twins ($n = 3,859$), but a possible limitation is due to the fact that only one twin was sent for genotyping from twin pairs who were thought to be MZ and both twins were sent for genotyping from twin pairs who were thought to be DZ. This means that there are fewer MZ twin pairs with genotyping data, resulting in a greater number of DZ pairs than MZ pairs available for comparison with the PPQ zygosity data. Despite this, we still had genotypic data on over 340 MZ twin pairs who self-reported as MZ.

Using data from adult twins in the TwinsUK cohort, we have validated the PPQ as an excellent proxy indicator of zygosity. In particular, if an initial PPQ from both twins scores as DZ, they are 98.7% likely to be DZ on genotyping (and 99.1% likely if one scores DZ and the other UZ on multiple PPQ testing). While only 88.6% of pairs where both initially scored MZ were truly MZ, this improved to 99.6% of pairs where they consistently scored MZ across multiple PPQs. The single 'alike as two peas in the pod' question was most discriminatory (92% and 97% accurate for MZ and DZ individuals, respectively). We would recommend that twin registries could use the PPQ as a quick and relatively inexpensive way of determining the zygosity of twins at registration. It may be unnecessary to genotype pairs where both twins score as DZ on the PPQ (or one as DZ and one as UZ), and similarly where both twins have MZ scores over serial questionnaires. However, depending on budget and time pressures, genotyping may be required to identify the true zygosity where there are other twin combinations (such as one twin scoring as MZ, or both UZ).

Acknowledgments

Research Centre based at Guy's, and St. Thomas' NHS Foundation Trust in partnership with King's College London (Grant code WT081878MA).

Disclosure of Interests

None.

Details of Ethical Approval

The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008. Written, informed consent was obtained from participants of the study.

Supplementary material

To view supplementary material for this article, please visit https://doi.org/10.1017/thg.2018.9

References

Bønnelykke, B., Hauge, M., Holm, N., Kristoffersen, K., & Gurtler, H. (1988). Evaluation of zygosity diagnosis in twin pairs below age seven by means of a mailed questionnaire. *Acta Geneticae Medicae et Gemellologiae*, *38*, 305–313.

Cederlöf, R., Friberg, L., Jonsson, E., & Kaij, L. (1961). Studies on similarity diagnosis in twins with the aid of mailed questionnaires. *Human Heredity*, *11*, 338–362.

Chen, J., Li, X., Chen, Z., Yang, X., Zhang, J., Duan, Q., & Ge, X. (2010). Optimization of zygosity determination by questionnaire and DNA genotyping in Chinese adolescent twins. *Twin Research and Human Genetics, 13*, 194–200.

Eaves, L. J., Silberg, J. L., Hewitt, J. K., Rutter, M., Meyer, J. M., Neale, M. C., & Pickles, A. (1993). Analyzing twin resemblance in multisymptom data: Genetic applications of a latent class model for symptoms of conduct disorder in juvenile boys. *Behavior Genetics*, *23*, 5–19.

Heath, A. C., Nyholt, D. R., Neuman, R., Madden, P. A., Bucholz, K. K., Todd, R. D., … Martin, N. G. (2003). Zygosity diagnosis in the absence of genotypic data: An approach using latent class analysis. *Twin Research*, *6*, 22–26.

Kasriel, J., & Eaves, L. (1976). The zygosity of twins: Further evidence on the agreement between diagnosis by blood groups and written questionnaires. *Journal of Biosocial Science*, *8*, 263–266.

Lykken, D. T. (1978). The diagnosis of zygosity in twins. *Behavior Genetics*, *8*, 437–473.

Magnus, P., Berg, K., & Nance, W. E. (1983). Predicting zygosity in Norwegian twin pairs born 1915–1960. *Clinical Genetics*, *24*, 103–112.

Moayyeri, A., Hammond, C. J., Hart, D. J., & Spector, T. D. (2013). The UK adult twin registry (TwinsUK Resource). *Twin Research and Human Genetics*, *16*, 144–149.

Moayyeri, A., Hammond, C. J., Valdes, A. M., & Spector, T. D. (2013). Cohort profile: TwinsUK and healthy ageing twin study. *International Journal of Epidemiology*, *42*, 76–85.

Ooki, S., Yamada, K., Asaka, A., & Hayakawa, K. (1989). Zygosity diagnosis of twins by questionnaire. *Acta Geneticae Medicae et Gemellologiae*, *39*, 109–115.

Ooki, S., Yokoyama, Y., & Asaka, A. (2004). Zygosity misclassification of twins at birth in Japan. *Twin Research*, *7*, 228–232.

Peeters, H., Van Gestel, S., Vlietinck, R., Derom, C., & Derom, R. (1998). Validation of a telephone zygosity questionnaire in twins of known zygosity. *Behavior Genetics, 28*, 159–163.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. Z., Bender, D., … Sham, P. C. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*, *81*, 559–575.

Sarna, S., Kaprio, J., Sistonen, P., & Koskenvuo, M. (1978). Diagnosis of twin zygosity by mailed questionnaire. *Human Heredity*, *28*, 241–254.

Song, Y. M., Lee, D., Lee, M. K., Lee, K., Lee, H. J., Hong, E. J., … Sung, J. (2010). Validity of the zygosity questionnaire and characteristics of zygosity-misdiagnosed twin pairs in the Health Twin Study of Korea. *Twin Research and Human Genetics*, *13*, 223–230.

Spector, T. D., & Williams, F. M. (2006). The UK Adult Twin Registry (TwinsUK). *Twin Research and Human Genetics*, *9*, 899–906.