Assessing the influence of consanguinity on congenital heart disease

Alan H. Bittles

Centre for Comparative Genomics, Murdoch University, and School of Medical Sciences, Edith Cowan University, Perth, Australia

ABSTRACT

Numerous articles have been published linking consanguineous marriage to an elevated prevalence of congenital heart disease, with ventricular septal defects and atrial septal defects the most commonly cited disorders. While initially persuasive, on closer examination many of these studies have fundamental shortcomings in their design and in the recruitment of study subjects and controls. Improved matching of cases and controls, to include recognition of the long-established community boundaries within which most marriages are contracted, and the assessment of consanguinity within specific levels and types of marital union would improve and help to focus the study outcomes. At the same time, major discrepancies between studies in their reported prevalence and types of congenital heart disease suggest an urgent need for greater standardization in the classification and reporting of these disorders.

Keywords: Congenital heart defects, consanguinity, endogamy, population stratification

INTRODUCTION

Marital unions between close biological kin are common in many populations, especially in South, Central and West Asia, and North and sub-Saharan Africa, and in migrant communities from those regions now resident in Europe, North America and Oceania.[1] From a medical genetics perspective, all marriages between couples related as second cousins or closer are regarded as consanguineous (derived from the Latin *con sanguineus*; i.e. sharing the same blood), and using this definition, it has been estimated that at least 10.4% of the current world population of 7.0 billion persons are consanguineous, with first cousin marriages especially popular.[1]

Despite the many investigations that have been conducted into the relationship between consanguinity and congenital heart disease (CHD), the precise nature and significance of the association remains unclear. At the most basic level it would be reasonable to assume that if a robust relationship could be demonstrated, it would be indicative of a causative mutation or the faulty expression of one or a number of otherwise rare recessive genes present in the parents in heterozygous form and adversely affecting embryonic and fetal development when homozygous. However, many of the more common CHDs appear to be genetically heterogeneous, whether diagnosed as isolated anomalies or accompanied by other heart defects.[2] If common predisposing genes are involved in the expression of CHD then, by definition, the likelihood of an association with consanguinity would be substantially reduced.[1]

The epidemiological data often are ambiguous, and although in many instances the results are suggestive of an association between consanguinity and CHD, they frequently fail to attain statistical significance. Several explanations can be advanced for this shortcoming, relating respectively to the diagnostic methods and disease criteria that are employed, the mode of recruitment of study subjects and the numbers and types of cases involved, the demographic and genetic structures of individual sub-populations and communities, and the methods of consanguinity assessment used, with the latter two factors especially relevant in the context of the Indian subcontinent.

Address for correspondence: Prof. Alan H. Bittles, Centre for Comparative Genomics, Murdoch University, South Street, Perth, WA 6150, Australia.
E-mail: abittles@ccg.murdoch.edu.au

DIAGNOSTIC METHODS AND CRITERIA FOR CHD IDENTIFICATION

In most studies, CHD has been diagnosed by echocardiography and/or cardiac catheterization, with a birth prevalence of 4–8/1000 typically cited and approximately 90% of CHD reported as being multifactorial in origin.[3] However, in a review by the American Heart Association Congenital Defects Committee endorsed by the American Academy of Pediatrics, an overall CHD incidence of 50/1000 births was described as conservative, on the grounds that a number of commonly occurring defects are routinely excluded from prevalence estimates. For example, bicuspid aortic valve, which affects 10–20/1000 people in the general population and may be associated with considerable morbidity and mortality in later life.[4] Further, in studies on Israeli neonates, the incidence of ventricular septal defects (VSD) alone was 53/1000.[5] The scale of the discrepancy in total CHD incidence between the generally accepted level of 4–8/1000 and 50+/1000 live births therefore raises significant concerns, and merits detailed attention and explanation.

As a first step it is important to determine exactly which types of heart defect are under investigation, the methods of ascertainment employed, the defects recorded and reported, and of equal importance which types of CHD were not identified or identifiable by the study techniques. With just a few exceptions,[5,6] a consistent positive association has been reported between consanguinity and VSD and atrial septal defects (ASD), but both positive and negative associations between consanguinity and patent ductus arteriosus, atrioventricular septal defect, pulmonary atresia, tetralogy of Fallot and other CHDs have been recorded in different populations.[7-10]

What is presently unclear is whether these inter-population differences are real and indicative of population-specific patterns of CHD, which in turn could be genetic and/or environmental in origin. Or if they also reflect variable diagnostic capacities, expertise and experience, the case definitions employed in the different study centers which have contributed to the literature, and whether ICD codes for malformations of the cardiac system were routinely applied, and disorders were subtyped in terms of embryological timing.

SELECTION AND COMPOSITION OF THE STUDY SAMPLE

A second area of concern relates to the composition of the study populations included in CHD investigations. While neonates only were recruited in a Lebanese study that reported a significant association between consanguinity

and CHD,[9] many other widely cited studies variously recruited symptomatic infants, including some with chromosomal anomalies such as Down syndrome, children of differing ages, and even adults in the third and fourth decades of life.[6-8,10-16] Likewise, while some studies included control for the potential effects of sociodemographic and lifestyle variables, such as maternal age and education, socioeconomic status, and smoking, and clinical conditions including pre-eclampsia, maternal diabetes, and pregnancy-induced hypertension, other studies did not.

Age-independent case selection makes rigorous inter-study comparisons well-nigh impossible, as defects that would, or at least could, have been diagnosed shortly after birth may have spontaneously corrected early in life. For example, in the Israeli study in which no association was found between consanguinity and the incidence of VSD, the defects detected were asymptomatic and 89% closed spontaneously within the first 10 months of life.[5] This outcome was confirmed in an associated study on preterm neonates which reported 87.5% spontaneous closure of VSD by age 6–11 months.[17]

Many of the more seriously affected cases of CHD diagnosed in developing countries die in childhood,[18] and unless assessed in early infancy they would not be included in prevalence statistics. By comparison, adults who are not diagnosed with CHD until their 20s and 30s but are included in prevalence statistics presumably do not have an especially life-threatening condition. Their inclusion may thus convey an erroneous impression as to the nature and clinical severity of the underlying cardiac lesions in the population. With these factors in mind, the unqualified use of mixed-age study groups in the investigation of CHD and the calculation of prevalence rates is questionable. More specifically, in the present context, is it appropriate to use the results obtained with such heterogeneous patient groups to assess the influence of consanguinity on CHD prevalence and clinical outcomes?

POPULATION STRATIFICATION IN CASE–CONTROL AND ASSOCIATION STUDIES

To ensure scientific and clinical relevance, case–control and association studies rely on the close matching of affected persons with other individuals who do not exhibit symptoms and have no family history of the disorder under investigation. As previously noted, for this reason cases and controls are routinely compared and matched in terms of age and gender, and with respect to variables such as residence and socioeconomic status. While this type of matching might be sufficient when largely environmental disorders are under investigation, when a genetic predisposition or cause is suspected it is of obvious importance that matching in terms of

inheritance also needs to be ensured.

This requirement is especially important in South Asia due to the very marked patterns and levels of population stratification. For example, in India, there are numerous co-resident ethnicities and religions, some 50,000–60,000 caste and non-caste communities, and up to 30% uncle–niece and first cousin marriages reported in the southern states of Andhra Pradesh, Karnataka and Tamil Nadu, and in southern Maharashtra.[19-22] Similarly, in Pakistan, the population is subdivided into ethnic communities and *biraderi* (literally translated as "brotherhoods"), that is, traditional social and occupational groupings inherited through the male lineage, and with first cousin marriage rates of 40–50+% consistently reported in regional and national surveys.[23-25] In both countries, the long history of marriage within strict community boundaries means that significant genetic differentiation predictably would have occurred, with some disease mutations unique to specific communities. Ongoing studies have suggested that on average each human carries 50–100 genetic variants that have been implicated in inherited disorders.[26] Ample opportunities would therefore have existed for major inter-community genetic differentiation during the estimated 3000 years since the initial establishment of the caste system.

It is the routine failure to control for population genetic subdivisions brought about by community endogamy that is of greatest significance and concern in association studies, including the investigation of CHD. Besides the widely recognized geographical, ethnic and religious divisions, the genetic effects of caste and sub-caste endogamy have been clearly demonstrated in India,[22,27,28] and an equivalent role is played by *biraderi* subdivisions in the Muslim populations of India and Pakistan.[29-31] Yet, in the selection of cases and controls, little attempt appears to have been made to control for these differences, despite the clear role of community-specific founder mutations in many genetic disorders,[1] for example, β-thalassemia in the populations of the subcontinent.[32,33]

This problematic situation was demonstrated in a bibliographic exercise based on information compiled from PubMed, using the terms association study, case–control study and clinical trial as keywords and matching of these terms with caste or *biraderi* and consanguinity.[34] Of the 948 association studies published on Indian populations, only 4.2% mentioned caste, while for 2030 case–control studies the comparable figure was 1.9%, and for 1381 clinical trials it was 0.5%. The data for Pakistan were essentially the same, with just 6.6% of association studies, 1.9% of case–control studies and 0.6% of clinical trials indicating caste or *biraderi* affiliations. The combined data therefore suggest non-recognition or inadvertent avoidance of the central roles played by caste and *biraderi* membership in determining the genetic structure of South Asian populations.

Failure to account for the probability of inter-community genetic differences means that the results obtained with case–control and gene association studies may at least in part represent flawed recruitment protocols and thus inappropriate investigative design, which could explain the notable lack of reproducibility in many gene association studies. However, another factor to be considered is that because of the highly endogamous nature of caste and *biraderi* lineages, even children described as "non-consanguineous" because their parents were unrelated prior to marriage may in fact be homozygous at a significant proportion of their gene loci.

This phenomenon relates to the deeper ancestral heritage of humans and has been clearly demonstrated in studies in Asian, European and North American populations.[35-38] Even families known not to have contracted consanguineous marriages for at least 5–10 generations exhibit numerous regions of the genome in which there are long uninterrupted runs of homozygosity that are indicative of close kin unions in much earlier generations.[36,37]

ASSESSING THE INFLUENCE OF CONSANGUINITY IN CASE–CONTROL AND ASSOCIATION STUDIES

The situation with regard to non-recognition of the possible influence of intra-familial marriage in case–control and association studies is actually more extreme, with in India consanguinity mentioned in just 1.0%, 0.2% and <0.1% of association studies, case–control studies, and clinical trials, respectively, possibly due to the strict avoidance of unions between biological relatives in many communities. But the fact that only 4.6% of association studies, 4.0% of case–control studies and no clinical trials conducted in Pakistan discussed the possible genetic effects of consanguinity in a country where 40–50+% of all marriages are between first cousins can be highly misleading, especially for clinicians and researchers who may be unfamiliar with the high prevalence of this form of marriage and its consequences in terms of increased genomic homozygosity.[34]

Even when consanguinity is included as a variable, as has been the case in many studies of CHD, the evaluation very frequently is reduced to a simple "consanguineous" versus "non-consanguineous" dichotomy. In an uncle–niece or a double first cousin marriage, the couple is assumed to have inherited 1/4 of their genes from a common ancestor, whereas in first cousin unions the assumption is that the couple has inherited 1/8 of their genes from a common ancestor, and for a second cousin

couple the comparable proportion is 1/32. This means that on average the progeny of an uncle–niece or a double first cousin marriage will be expected to have inherited identical gene copies at 1/8 of all their loci, defined as a coefficient of inbreeding $F = 0.125$. It follows that for first cousin progeny, $F = 0.0625$, that is, 1/16 loci predictably are homozygous, whereas for second cousins, $F = 0.0156$, that is, 1/64 of loci are homozygous.[39]

It is unclear and unhelpful why in investigating the impact of consanguinity on the prevalence and types of CHD, researchers very frequently have opted to coalesce three quite distinct levels of consanguinity ($F = 0.0156$, 0.0625 and 0.125), representing an eightfold average difference in genomic homozygosity, into a single ambiguous and ill-defined "consanguineous" category. Failure to recruit adequate numbers of cases in each specific category of consanguineous union is not a sufficient excuse for the adoption of this practice in order to attain statistical significance.

It is also important to recognize that the preferred types of consanguineous marriages vary between communities with, for example, uncle–niece marriages commonly contracted in Dravidian Hindu communities in South India, whereas first cousins are by far the most common form of consanguineous marriage in Muslim populations.[40] Also, since in many communities consanguinity is a long-established tradition, the effect of uninterrupted intra-familial marriages contracted in successive generations would have resulted in much higher cumulative coefficients of inbreeding than would be predicted on the basis of consanguineous marriage in a parental generation only.[41]

Perhaps because of these complex and variable factors, it was not unexpected that a study conducted on consanguineous progeny in Bangalore, the capital of Karnataka where uncle–niece and first cousin marriages remain popular, was unable to identify a single gene of major effect in a clinically heterogeneous cohort of patients with CHD.[42] As suggested by the authors, the statistical power of follow-up studies could be significantly increased by investigating consanguineous families in which multiple children have been diagnosed with the same form of CHD. However, a complementary approach which would enable more genetically meaningful comparisons to be drawn would be to match cases and controls according to their community of origin, rather than simply on the basis of parental consanguinity.[1,42]

DISCUSSION

Obviously, there must be some practical limitation in the degree to which cases and controls can be matched, and control for potential genetic variables is currently much more intractable than for environmental or sociodemographic factors. At the same time, it should be incumbent on authors to avoid claims for their findings that are based on unproven and/or questionable study designs, particularly when major variables of genetic significance, such as caste/*biraderi* affiliation and the specific levels of consanguinity of individual cases and controls, have often been effectively ignored or inadequately controlled.

In India, the selection of cases and controls matched solely on the basis of personal religion is inadequate because most members of the Muslim, Christian and Buddhist communities are the descendants of converts from Hinduism in the recent or relatively recent past, and so they would be expected to share significant portions of their genomes with sectors of the majority Hindu community.[43] Given the approximately 120 generations since the establishment of the caste system, the restricted numbers of founder individuals in each caste, the effects of subsequent caste sub-divisions, and the action of genetic drift, the opportunity for genetic differentiation and the accumulation of caste-specific mutations would have been equivalently large.

From a global perspective, more appropriate subject recruitment based on social and demographic realities should lead to improved investigative outcomes for CHD and other common disorders. This conclusion applies not only to South Asia but also to the Middle East, Africa and many other regions of the world where clan and tribal affiliations should be integral factors in the recruitment of cases and controls. At the same time, to facilitate credible inter-population comparisons, the standardization of diagnostic methods and particularly disease classificatory systems needs to be urgently addressed, together with improved statistical analysis and reporting of findings. Once these basic changes have been effected, it should be possible to more accurately assess the impact of consanguineous marriage on the health of populations in general, and more specifically on the prevalence and types of CHD at family and community levels.

REFERENCES

1. Bittles AH, Black ML. Consanguinity, human evolution, and complex diseases. Proc Natl Acad Sci U S A 2010;107 Suppl 1:1779-86.

2. Benson DW, Sharkey A, Fatkin D, Lang P, Basson CT, McDonough B, *et al.* Reduced penetrance, variable expressivity, and genetic heterogeneity of familial atrial septal defects. Circulation 1998;97:2043-8.

3. Christianson A, Howson CP, Modell B. March of Dimes Global Report on Birth Defects. White Plains, New York: March of Dimes Birth Defects Foundation; 2006.

4. Pierpont ME, Basson CT, Benson DW Jr, Gelb BD, Giglia TM,

Goldmuntz E, *et al.* Genetic basis for congenital heart defects: current knowledge: a scientific statement from the American Heart Association Congenital Cardiac Defects Committee, Council on Cardiovascular Disease in the Young: endorsed by the American Academy of Pediatrics. Circulation 2007;115:3015-38.

5. Roguin N, Du ZD, Barak M, Nasser N, Hershkowitz S, Milgram E. High prevalence of muscular septal defect in neonates. J Am Coll Cardiol 1995;26:1545-8.

6. Roodpeyma S, Kamali Z, Afshar F, Naraghi S. Risk factors in congenital heart disease. Clin Pediatr (Phila) 2002;41:653-8.

7. Becker SM, Al Halees Z, Molina C, Paterson RM. Consanguinity and congenital heart disease in Saudi Arabia. Am J Med Genet 2001;99:8-13.

8. Nabulsi MM, Tamim H, Sabbagh M, Obeid MY, Yunis KA, Bitar FF. Parental consanguinity and congenital heart malformations in a developing country. Am J Med Genet A 2003;116A:342-7.

9. Yunis K, Mumtaz G, Bitar F, Chamseddine F, Kassar M, Rashkidi J, *et al.* Consanguineous marriage and congenital heart defects: a case-control study in the neonatal period. Am J Med Genet Part A 2006;140:1524-30.

10. Chehab G, Chedid P, Saliba Z, Bouvagnet P. Congenital cardiac disease and inbreeding: specific defects escape higher risk due to parental consanguinity. Cardiol Young 2007;17:414-22.

11. Gev D, Roguin N, Freundlich E. Consanguinity and congenital heart disease in the rural Arab population in northern Israel. Hum Hered 1986;36:213-7.

12. Badaruddoza, Afzal M, Akhtaruzzaman. Inbreeding and congenital heart diseases in a north Indian population. Clin Genet 1994;45:288-91.

13. Hassan I, Haleem AA, Bhutta ZA. Profile and risk factors for congenital heart disease. J Pak Med Assoc 1997;47:78-81.

14. Gnanalingam MG, Gnanalingam KK, Singh A. Congenital heart disease and parental consanguinity in South India. Acta Paediatr 1999;88:473-4.

15. Subramanyan R, Joy J, Venugopalan P, Sapru A, al Khusaiby SM. Incidence and spectrum of congenital heart disease in Oman. Ann Trop Paediatr 2000;20,337-41.

16. Ramegowda S, Ramachandra NB. Parental consanguinity increases congenital heart diseases in South India. Ann Hum Biol 2006;33:519-28.

17. Du ZD, Roguin N, Barak M, Bihari SG, Ben-Elisha M. High prevalence of muscular venticular septal defect in preterm neonates. Am J Cardiol 1996;78:1183-5.

18. Bitar FF, Baltaji N, Dbaibo G, Abed el-Jawad M, Yunis KA, Obeid M. Congenital heart disease at a tertiary care center in Lebanon. Middle East J Anesthesiol 1999;15:159-64.

19. Bittles AH, Mason WM, Greene J, Rao NA. Reproductive behavior and health in consanguineous marriages. Science 1991;252:789-94.

20. NFHS. National Family Health Survey, India 1992-93. Bombay: International Institute for Population Sciences; 1995.

21. Gadgil M, Joshi MV, Manoharan S, Patil S, Prasad UV. Peopling of India. In: Balasubramanian D, Appaji Rao N, editors. The Indian Human Heritage. Hyderabad: Universities Press; 1998. p. 100-29.

22. Bittles AH. Endogamy, consanguinity and community genetics. J Genet 2002;81:91-8.

23. Ahmed T, Ali SM, Aliaga A, Arnold F, Ayub M, Bhatti MH, *et al.* Pakistan Demographic and Health Survey 1990/91. Islamabad and Columbia, MD: Pakistan National Institute of Population Studies and Macro International; 1992.

24. Bittles AH, Grant JC, Shami SA. Consanguinity as a determinant of reproductive behaviour and mortality in Pakistan. Int J Epidemiol 1993;22:463-7.

25. Hussain R, Bittles AH. The prevalence and demographic characteristics of consanguineous marriages in Pakistan. J Biosoc Sci 1998;30:261-75.

26. 1000 Genomes Project Consortium. A map of human genome variation from population-scale sequencing. Nature 2010;467:1061-73.

27. Indian Genome Variation Consortium. Genetic landscape of the people of India: a canvas for disease gene exploration. J Genet 2008;87:3-20.

28. Reich D, Thangaraj K, Patterson N, Price AL, Singh L. Reconstructing Indian population history. Nature 2009;461:489-94.

29. Shami SA, Grant JC, Bittles AH. Consanguineous marriage within social/occupational class boundaries in Pakistan. J Biosoc Sci 1994;26:91-6.

30. Wang W, Sullivan SG, Ahmed S, Chandler D, Zhivotovsky LA, Bittles AH. A genome-based study of consanguinity in three co-resident endogamous Pakistan communities. Ann Hum Genet 2000;64:41-9.

31. Qamar R, Ayub Q, Mohyuddin A, Helgason A, Mazhar K, Mansoor A, *et al.* Y-chromosomal DNA variation in Pakistan. Am J Hum Genet 2002;70:1107-24.

32. Sinha S, Black ML, Agarwal S, Colah R, Das R, Ryan K, *et al.* Profiling β-thalassaemia mutations in India at state and regional levels: implications for genetic education, screening and counseling programmes. Hugo J 2009;3:51-62.

33. Black ML, Sinha S, Agarwal S, Colah R, Das R, Bellgard M, *et al.* A descriptive profile of β-thalassaemia mutations in India, Pakistan and Sri Lanka. J Commun Genet 2010;1:49-57.

34. Bittles AH. Population stratification and genetic association studies in South Asia. J Mol Genet Med 2005;1:43-8.

35. Gibson J, Morton NE, Collins A. Extended tracts of homozygosity in outbred human populations. Hum Mol Genet 2006;15:789-95.

36. McQuillan R, Leutenegger AL, Abdel-Rahman R, Franklin CS, Pericic M, Barac-Lauc L, *et al.* Runs of homozygosity in European populations. Am J Hum Genet 2008;83:359-72.

37. Kirin M, McQuillan R, Franklin CS, Campbell H, McKeigue PM, Wilson JF. Genomic runs of homozygosity record population history and consanguinity. PLoS One 2010;5:e13996.

38. Leutenegger AL, Sahbatou M, Gazal S, Cann H, Génin E. Consanguinity around the world: what do the genomic data of the HGDP-CEPH diversity panel tell us? Eur J Hum Genet 2011;19:583-7.

39. Bittles AH. Consanguinity, genetic drift and genetic diseases in populations with reduced numbers of founders. In: Speicher M, Antonarakis SE, Motulsky AG, editors. Human Genetics - Principles and Approaches. 4th ed. Heidelberg: Springer; 2009. p. 507-28.

40. Bittles AH. A community genetics perspective on consanguineous marriage. Commun Genet 2008;11:324-30.

41. Woods CG, Cox J, Springell K, Hampshire DJ, Mohamed MD, McKibbin M, et al. Quantification of homozygosity in consanguineous individuals with autosomal recessive disease. Am J Hum Genet 2006;78:889-96.

42. McGregor TL, Misri A, Bartlett J, Orabona G, Friedman RD, Sexton D, et al. Consanguinity mapping of congenital heart disease in a South Indian population. PLoS One 2010;5:e10286.

43. Bittles AH. Congenital heart disease and consanguineous marriage in South India. Ann Hum Biol 2007;34:682-3.

Author Help: Online submission of the manuscripts

Articles can be submitted online from http://www.journalonweb.com. For online submission, the articles should be prepared in two files (first page file and article file). Images should be submitted separately.

1) **First Page File:**
   Prepare the title page, covering letter, acknowledgement etc. using a word processor program. All information related to your identity should be included here. Use text/rtf/doc/pdf files. Do not zip the files.

2) **Article File:**
   The main text of the article, beginning with the Abstract to References (including tables) should be in this file. Do not include any information (such as acknowledgement, your names in page headers etc.) in this file. Use text/rtf/doc/pdf files. Do not zip the files. Limit the file size to 1 MB. Do not incorporate images in the file. If file size is large, graphs can be submitted separately as images, without their being incorporated in the article file. This will reduce the size of the file.

3) **Images:**
   Submit good quality color images. Each image should be less than 4 MB in size. The size of the image can be reduced by decreasing the actual height and width of the images (keep up to about 6 inches and up to about 1800 x 1200 pixels). JPEG is the most suitable file format. The image quality should be good enough to judge the scientific value of the image. For the purpose of printing, always retain a good quality, high resolution image. This high resolution image should be sent to the editorial office at the time of sending a revised article.

4) **Legends:**
   Legends for the figures/images should be included at the end of the article file.