



Published in final edited form as:

J Med Genet. 2008 August ; 45(8): 498–499. doi:10.1136/jmg.2008.059055.

The fragile X prevalence paradox

Paul J Hagerman [Professor]

Department of Biochemistry and Molecular Medicine, University of California, Davis, School of Medicine, 4303, Tupper Hall, One Shields Ave, Davis, California 95616, USA

Although fragile X syndrome (FXS; OMIM 300624) is generally regarded as the most common inherited form of cognitive impairment,^{1–3} there is little consensus as to its prevalence in the general population or to sex-specific differences in prevalence. Estimates of FXS prevalence (~1/4000–1/8000) that are based on population projections from cohorts of children with special education needs (SEN) generally underestimate the extent of clinical involvement (for a comprehensive summary, see Song *et al*⁴), as many individuals affected by the behavioural, emotional and/or learning disabilities of FXS have IQs in the normal or borderline range.^{5,6} The latter may not be included in cohorts that use cognitive impairment as an inclusion criterion, a problem that is particularly marked for girls, with the majority having IQs within the normal range.⁷

A second difficulty with population studies is the tendency to conflate disease prevalence (projections from SEN cohorts) with allele/carrier frequencies within the general population, as prevalence estimates will only approach carrier frequencies for complete penetrance. Thus, seemingly paradoxical results among studies of prevalence may reflect the effects of selection bias, conflation and differing defining criteria for FXS across studies. Aspects of this “paradox” are exemplified by two important screening studies of Israeli women. Using the same cut-off point (55 CGG repeats) for premutation (PM) carriers, Toledano-Alhadeef *et al*⁸ found a higher frequency of carriers (127/14 334=1/113) than did Pessoa *et al*⁹ (62/9459=1/152), despite the fact that Pessoa *et al* reported a higher frequency for full mutation (FM) alleles (4 FM carriers; 1/2365) than did Toledano-Alhadeef *et al* (3 FM carriers; 1/4778). Although the numbers of FM alleles in the two studies are too small to attach significance to the difference in frequencies, the trend is consistent with the exclusion of individuals with a family history of learning difficulties by Toledano-Alhadeef *et al*⁸ whereas such individuals were not excluded by Pessoa *et al*.⁹ Samples approaching 50 000 would be needed to establish significance of the difference between the observed frequencies. Interestingly, there was no significant difference between the two studies for the allele distributions within the PM ranges. Indeed, Toledano-Alhadeef *et al*⁸ observed a greater percentage of PM alleles of >70 CGG repeats than did Pessoa *et al*.⁹

A generally unexploited feature of the genetics of FXS is that there is a defined (albeit uncharacterised) relationship between the frequencies of PM alleles (which are relatively easy to determine in an unbiased manner) and the frequencies of FM alleles in a given population. Thus, it is possible to use the frequencies of PM alleles to define expectations for the number of FM alleles. As noted above and in Song *et al*,⁴ the best population estimates of PM carrier frequencies have been obtained through large-scale screening of normal pregnant or preconception women (eg 62 of 9459 subjects;⁹ 127 of 14 334 subjects,⁸). Although these frequencies are themselves subject to a number of caveats, including exclusion of individuals with known family history of cognitive impairment and possible regional founder effects, they

Correspondence to: Professor P J Hagerman, Department of Biochemistry and Molecular Medicine, University of California, Davis, School of Medicine, 4303, Tupper Hall, One Shields Ave, Davis, California 95616, USA; pjhagerman@ucdavis.edu.

Competing interests: None.

can be used as a starting point for providing expectations for the frequency of male and female individuals who harbour FM alleles and of male individuals with PM alleles. The other advantage of using the frequency estimates for female (PM) carriers is that there are published estimates for the CGG repeat-dependence of both the probability (p_i ; i = number of CGG repeats) of PM to FM transmission¹⁰ and the relative frequency (f_i) of PM alleles.¹¹ These two quantities can be used to estimate an aggregate probability, S , of the transmission (loss) of a PM allele to a FM allele. S , in turn, can be used to estimate the remaining frequencies expected within the same population. S is obtained by summing the product of the fraction (f_i) of pre-mutation alleles of i CGG repeats and the probability (p_i) of a PM to FM transition for the same allele size (i); thus,

$$S = \sum_{i \geq 55} (f_i \times p_i) = 0.107.$$

Strictly speaking, these are lower-bound estimates for FM allele frequency, as the less common FM to FM transmissions are not counted.

Relative frequency (f_i) was estimated by a cubic fit to the allele data in table 3 of Jacquemont *et al.*,¹¹ correcting for the range of CGG repeat sizes in each grouping, yielding

$$f_i = 1.2261 - 0.0396(i) + 0.0004284(i)^2 - 1.5498e - 6(i)^3$$

where i is the number of CGG repeats ($55 \leq i \leq 100$). The corresponding transition probability was estimated by fitting the data in table 1 of Nolin *et al.*,¹⁰ yielding

$$p_i = 5.9475 - 0.2515(i) + 0.0033524(i)^2 - 1.3397e - 5(i)^3.$$

Both cubic fits were truncated to 100 CGG repeats, as the product function is essentially zero beyond that range.

Using an aggregate value (189/239,793; 1/126) for the frequency of PM carriers from the studies of Pessoa *et al.*⁹ and Toledano-Alhadeff *et al.*⁸ the expected (average) frequency for FM males and females is $1/126 \times 0.5 \times S = 1/2355$ and the expected frequency for PM males is $1/126 \times 0.5 \times (1-S) = 1/282$. Remarkably, the predicted value for females with the FM allele is very close to the observed value (1/2365) from the study of Pessoa *et al.*⁹ which did include women with a history in the extended family of learning difficulty or developmental problems. Clearly, given the uncertainties inherent in the analysis, as well as the use of the aggregate allele frequency, such agreement must be considered fortuitous. The lower value (1/4778) observed by Toledano-Alhadeff *et al.*⁸ probably reflects the results of excluding any family history of learning difficulty. An important prediction of the higher frequency estimate (1/282) for PM alleles in males is that as many as $\sim 1/3000$ males aged >50 years in the general population may have the carrier-specific neurodegenerative disorder, fragile X-associated tremor/ataxia syndrome (FXTAS^{12,13}).

Implicit in the foregoing analysis is the assumption that, whereas the allele frequencies will probably depend on a specific population being studied, the transmission fraction, S , is likely to be a more general function of the genetic mechanisms underlying CGG-repeat expansion. A test of this assumption (constancy of S across populations where absolute carrier frequencies may vary widely) is important to our understanding of the factors involved in dynamic repeat

instability. In this regard, a second pair of screening studies in eastern Canada¹⁴ found lower PM allele frequencies for both female ($41/10\ 624 = 1/259$)¹⁴ and male ($13/10\ 572 = 1/813$) subjects.¹⁵ However, the allele frequency for males ($1/580$) predicted from the female allele frequency and S (0.107), is within the 95% confidence limits of the observed frequency for males. These results suggest that founder effects may contribute to the approximately two-fold difference in absolute allele frequencies between the Canadian and Israeli populations.

It is hoped that the current comments will help to frame the discussion of both prevalence and frequency estimates for the fragile X family of disorders and to provide further impetus for larger scale screening of unbiased populations (eg newborn screening). To help frame this discussion, several points should be considered:

1. As noted above, an as yet untested prediction of the dynamic instability of the *FMRI* gene is that, whereas population founder effects will result in large variation in allele frequencies, the transmission fraction, S , should be relatively constant, reflecting underlying genetic mechanisms for expansion.
2. The frequency of PM alleles in a given population is likely to represent the most robust measure of variation (eg, founder effects) across populations, and estimates of this quantity should be the starting point for discussions of corresponding frequencies in males and in individuals (males and females) harbouring FM alleles.
3. The transmission model, based on the observation that PM to FM transmission is strictly matrilineal, predicts equal numbers of FM alleles in males and females. Screens for FM alleles in at-risk populations that yield substantial frequency differences for males and females are likely to possess significant selection bias.
4. As a corollary to (3), prevalence values for fragile X syndrome that are substantially lower than the corresponding frequency estimates for FM alleles should always raise the possibility that some criteria for clinical involvement (eg lowered IQ) are too restrictive. FXS also includes individuals within the normal/borderline IQ range who have learning deficits and emotional and/or behavioural difficulties. Starting from the vantage point of allele status, as obtained through newborn screening, a much better foundation would be laid for defining the true nature and variation of the phenotypic spectrum of FXS.
5. Direct, unbiased estimates of the frequency of FM alleles in a given population will require the screening of samples of at least 50 000 individuals. Such efforts will require new tools for high-throughput genotyping that are capable of direct detection of FM alleles in both males and females.¹⁶

In summary, despite numerous estimates of the frequencies of FM (*FMRI*) alleles, or prevalence of FXS, direct (general population) frequency estimates are still lacking, particularly for males. In lieu of such studies, perhaps the best current estimate for the frequency of the FM in females is $\sim 1/2500$, based on direct screening⁹ and by projection from the frequency of PM alleles (see above). This frequency estimate should also apply to male carriers of FM alleles, a prediction that remains to be tested by direct population screening. This prediction is consistent with the lower-bound estimate of $\sim 1/3600$ ⁷ from a male SEN cohort within the USA. It is therefore recommended that the oft-quoted $1/4000$ – $1/6000$ figures for FXS prevalence be abandoned in favour of an approximate frequency of $\sim 1/2500$ for individuals (male and female) with the FM allele. This frequency should be only slightly higher than the prevalence of FXS if the full spectrum of involvement is considered. Of course, these numbers are expected to display geographical variation due to founder effects.

Acknowledgments

I thank the National Fragile X Foundation for their support. This work was also supported by the National Center for Research Resources (UL1 RR024922) and the National Institute of Aging (RL1 AG032119). I also thank Drs F Tassone, R Hagerman and D Nguyen for their helpful comments and discussion regarding this work.

References

1. Sherman, S. Epidemiology. In: Hagerman, RJ.; Hagerman, PJ., editors. Fragile X syndrome: diagnosis, treatment and research. Vol. 3rd. Baltimore: The Johns Hopkins University Press; 2002.
2. Hagerman RJ. Lessons from fragile X regarding neurobiology, autism, and neurodegeneration. *J Dev Behav Pediatr* 2006;27:63–74. [PubMed: 16511373]
3. Hagerman RJ, Rivera S, Hagerman PJ. The fragile X family of disorders: a model for autism and targeted treatments. *Curr Ped Rev.* 2008in press
4. Song FJ, Barton P, Sleightholme V, Yao GL, Fry-Smith A. Screening for fragile X syndrome: a literature review and modeling study. *Health Technol Assess* 2003;7:1–106.
5. Bennetto, L.; Pennington, BF. Neuropsychology. In: Hagerman, RJ.; Hagerman, PJ., editors. Fragile X syndrome: diagnosis, treatment and research. Vol. 3rd. Baltimore: The Johns Hopkins University Press; 2002. p. 206–48.
6. de Vries BB, Wiegers AM, Smits AP, Mohkamsing S, Duivenvoorden HJ, Fryns JP, Curfs LM, Halley DJ, Oostra BA, van den Ouweland AM, Niermeijer MF. Mental status of females with an FMR1 gene full mutation. *Am J Hum Genet* 1996;58:1025–32. [PubMed: 8651263]
7. Crawford DC, Meadows KL, Newman JL, Taft LF, Scott E, Leslie M, Shubek L, Holmgreen P, Yeargin-Allsopp M, Boyle C, Sherman SL. Prevalence of the fragile X syndrome in African-Americans. *Am J Med Genet* 2002;110:226–33. [PubMed: 12116230]
8. Toledano-Alhadeif H, Basel-Vanagaite L, Magal N, Davidov B, Ehrlich S, Drasinover V, Taub E, Halpern GJ, Ginott N, Shohat M. Fragile-X carrier screening and the prevalence of premutation and full-mutation carriers in Israel. *Am J Hum Genet* 2001;69:351–60. [PubMed: 11443541]
9. Pessa R, Berkenstadt M, Cuckle H, Gak E, Peleg L, Frydman M, Barkai G. Screening for fragile X syndrome in women of reproductive age. *Prenat Diagn* 2000;20:611–14. [PubMed: 10951469]
10. Nolin SL, Brown WT, Glicksman A, Houck GE Jr, Gargano AD, Sullivan A, Biancalana V, Brondum-Nielsen K, Hjalgrim H, Holinski-Feder E, Kooy F, Longshore J, Macpherson J, Mandel JL, Matthijs G, Rousseau F, Steinbach P, Vaisanen ML, von Koskull H, Sherman SL. Expansion of the fragile X CGG repeat in females with premutation or intermediate alleles. *Am J Hum Genet* 2003;72:454–64. [PubMed: 12529854]
11. Jacquemont S, Leehey MA, Hagerman RJ, Beckett LA, Hagerman PJ. Size bias of fragile X premutation alleles in late-onset movement disorders. *J Med Genet* 2006;43:804–09. [PubMed: 16723388]
12. Amiri K, Hagerman RJ, Hagerman PJ. Fragile X-associated tremor/ataxia syndrome: an aging face of the fragile X gene. *Arch Neurol* 2008;65:19–25. [PubMed: 18195136]
13. Berry-Kravis E, Abrams L, Coffey SM, Hall DA, Greco C, Gane LW, Grigsby J, Bourgeois JA, Finucane B, Jacquemont S, Brunberg JA, Zhang L, Lin J, Tassone F, Hagerman PJ, Hagerman RJ, Leehey MA. Fragile X-associated tremor/ataxia syndrome: clinical features, genetics, and testing guidelines. *Mov Disord* 2007;22:2018–30. [PubMed: 17618523]quiz 140
14. Rousseau F, Rouillard P, Morel ML, Khandjian EW, Morgan K. Prevalence of carriers of premutation-size alleles of the *FMR1* gene and implications for the population genetics of the fragile X syndrome. *Am J Hum Genet* 1995;57:1006–18. [PubMed: 7485149]
15. Dombrowski C, Levesque S, Morel ML, Rouillard P, Morgan K, Rousseau F. Premutation and intermediate-size *FMR1* alleles in 10572 males from the general population: loss of an AGG interruption is a late event in the generation of fragile X syndrome alleles. *Hum Mol Genet* 2002;11:371–8. [PubMed: 11854169]
16. Tassone F, Pan R, Amiri K, Taylor AK, Hagerman PJ. A rapid polymerase chain reaction-based screening method for identification of all expanded alleles of the fragile X (*FMR1*) gene in newborn and high-risk populations. *J Mol Diagn* 2008;10:43–9. [PubMed: 18165273]