Longitudinal Analysis Casts Doubt on the Presence of Genetic Anticipation in Heritable Pulmonary Arterial Hypertension

Emma K. Larkin¹, John H. Newman², Eric D. Austin³, Anna R. Hemnes², Lisa Wheeler², Ivan M. Robbins², James D. West², John A. Phillips III³, Rizwan Hamid³, and James E. Loyd²

¹Vanderbilt University Institute for Medicine and Public Health and ²Division of Allergy, Pulmonary, and Critical Care Medicine, Department of Medicine, and the ³Division of Pediatric Pulmonary Medicine and Genetic Medicine, Department of Pediatrics, Vanderbilt University School of Medicine, Nashville, Tennessee

Rationale: Analysis of the age of onset in heritable pulmonary arterial hypertension (HPAH) has led to the hypothesis that genetic anticipation causes younger age of onset and death in subsequent generations. With accrual of pedigree data over multiple decades, we retested this hypothesis using analyses that eliminate the truncation of data that exists with shorter duration of follow-up.

Objectives: To analyze the pedigrees of families with mutations in bone morphogenetic protein receptor type 2 (BMPR2), afflicted in two or more generations with HPAH, eliminating time truncation bias by including families for whom we have at least 57 years of data.

Methods: We analyzed 355 individuals with BMPR2 mutations from 53 families in the Vanderbilt Pulmonary Hypertension Registry. We compared age at diagnosis or death in affected individuals ($n = 249$) by generation within families with multigenerational disease. We performed linear mixed effects models and we limited time-truncation bias by restricting date of birth to before 1955. This allowed for 57 years of follow-up (1955–2012) for mutation carriers to develop disease. We also conducted Kaplan-Meier analysis to include currently unaffected mutation carriers ($n = 106$).

Measurements and Main Results: Differences in age at diagnosis by generation were found in a biased analysis that included all birth years to the present, but this finding was eliminated when the 57-year observation limit was imposed. By Kaplan-Meier analysis, inclusion of currently unaffected mutation carriers strengthens the observation that bias of ascertainment exists when recent generations are included.

Conclusions: Genetic anticipation is likely an artifact of incomplete time of observation of kindreds with HPAH due to BMPR2 mutations.

Keywords: hereditary; pulmonary hypertension; genetics

Pulmonary arterial hypertension (PAH) is a rare, lethal disease of the pulmonary vasculature. Heritable pulmonary arterial hypertension (HPAH) is observed in families in about 6% of most large PAH registries, the majority of which (80%) are in the gene, *bone morphogenetic protein receptor type 2* (*BMPR2*)

AT A GLANCE COMMENTARY

Scientific Knowledge on the Subject

It is currently believed that genetic anticipation is a feature of heritable pulmonary hypertension related to mutations in BMPR2. This means earlier age of diagnosis in subsequent generations.

What This Study Adds to the Field

This study casts doubt on the current understanding of the disease and makes a search for molecular causes of genetic anticipation less compelling. Also, it informs families that the risk of earlier disease is not passed on to the next generations.

(1–3). In 1995, we reported that genetic anticipation in HPAH appeared likely based on comparisons of age at death between two to three affected generations within families (4). The possibility of genetic anticipation had been suggested in earlier reports of familial primary pulmonary hypertension, now called HPAH (5, 6). Other pulmonary hypertension registries have made similar observations (7, 8). Genetic anticipation has a known molecular basis in several diseases, but a biological cause has been elusive in HPAH (9). Genetic anticipation is the observation of progressively earlier onset and increased severity in successive generations in affected families (10). Genetic anticipation has been found to be artifactual in some illnesses, the result of (1) a bias of incomplete time of follow-up to allow apparently unaffected individuals to develop disease, and (2) a time bias inherent in the age comparison of an affected parent–offspring pair (10–12). However, discovery of the meiotic expansion of repeating segments of trinucleotides and other DNA expansions as the basis for genetic anticipation in multiple neurological diseases, including Huntington disease, fragile X syndrome, and spinocerebellar ataxia, has led to reconsideration of the possibility of genetic anticipation in other diseases with familial clustering (13, 14). These diseases of interest include Crohn disease, rheumatoid arthritis, psychiatric disorders, and HPAH (15–18). Given that about 80% of heritable PAH is associated with mutations in the gene *BMPR2*, we previously considered the possibility of trinucleotide repeats in *BMPR2* in our affected families. We evaluated the gene for repeats but have not identified such regions to date, including a search of the promoter region. In 1995, we did not yet have any HPAH gene identified, and that analysis was limited to more recent generations because of the newness of the registry.

HPAH is an autosomal dominant disease with reduced penetrance (19). It is estimated that about 20% of mutation carriers

(Received in original form May 16, 2012; accepted in final form August 19, 2012)

Supported by National Heart, Lung, and Blood Institute grant 5 K08 HL093363 (A.R.H.); National Institutes of Health grants K23 HL0987431 (E.D.A.), 1R01HL102020-01 (R.H.), and R01 HL 095797-01A2 (J.D.W.); and National Center for Research Resources/ National Institutes of Health grant UL1 RR024975 (J.H.N.).

Author Contributions: E.K.L.: conceived and performed statistical analysis; J.H.N., J.E.L., J.A.P., and L.W. developed the Registry and published initial analysis; A.R.H. and I.M.R. contributed patients; E.D.A., responsible for quality of database; R.H. and J.D.W. contributed to intellectual development of the topic.

Correspondence and requests for reprints should be addressed to John H. Newman, M.D., T 1219 Vanderbilt Medical Center North, Nashville, TN 37232-2650. E-mail: john.newman@vanderbilt.edu

Am J Respir Crit Care Med Vol 186, Iss. 9, pp 892–896, Nov 1, 2012

Copyright © 2012 by the American Thoracic Society

Originally Published in Press as DOI: 10.1164/rccm.201205-0886OC on August 23, 2012

Internet address: www.atsjournals.org

develop clinical pulmonary hypertension, although women are more often affected than men, at a ratio of about 3:1 (19). Family members have been diagnosed at ages as young as 12 months and as late as 67 years old (4, 6). The mean age at diagnosis in our family registry from all mutations is 35.8 ± 14 years (1 standard deviation of the mean) of age (4). The cumulative incidence of disease is approximately 90% by age 55 years, so inclusion of persons followed for at least 57 years (1955–2012) should capture the majority of new cases (4).

We now report a new analysis, comparing affected members in multiple generations from the families whose members were born before 1955. The interval between 1955 and the present, 2012, allows for 57 years of observation for expression of disease in mutation carriers who might have been unaffected at the initial comparisons. Using simple truncation analysis (not evaluating younger generations) to compare age at diagnosis in the generations who have lived at least 57 years, we have found that decreased age at diagnosis in subsequent generations is eliminated. Our results suggest that prolonged observation was needed to remove the bias of ascertainment in this disease of incomplete penetrance and wide range of age of disease onset. We also performed the more biased previous method of analysis by including more recent generations and found similar results to prior report.

METHODS

The Vanderbilt Pulmonary Hypertension Registry has been continuously active since its inception in 1980 and has data on several thousand total family members from more than 200 families with two or more cases of PAH. All data were obtained with Vanderbilt University Institutional Review Board approval #9401 and with informed consent. Of those with BMPR2 mutations, we have collected date of birth and age at diagnosis or death on 249 affected and 106 unaffected mutation carriers from 53 pedigrees. Data were obtained from family records, interviews, autopsies, and medical records (4, 20). We have date of birth and age at diagnosis or death in every decade from 1900 to 2012, plus 26 members (22 unaffected and 3 affected) born before 1900. For 30 individuals who died of pulmonary hypertension with missing age of diagnosis, age of death was used. For 16 individuals, we used age of transplant as the age of death. Age at death was usually within 2 to 3 years of age of diagnosis before the development of effective therapies in the 1990s and so is very close in time for generational and statistical analysis. In addition, the younger generations have a large number of patients who are alive, and thus date of death is not applicable, whereas date of diagnosis is known. Thirty-eight of the 53 total families in this analysis are shared in both the before and after 1955 date of birth cutoff. The overlap of families reduces the likelihood that the findings are an artifact of studying different cohorts or unique family effects. We made comparisons by generation, starting with the first generation identified with disease, whatever date that occurred. We excluded those families with Alk-1 or endoglin mutations, Caveolin 1 mutations, and those families (about 15% of our total registry) for whom we have not found the disease-related mutation (s). Thus, all families studied have well-identified BMPR2 mutations that segregate faithfully with disease. We restricted analyses to the first four generations due to limited sample size in generations five or greater.

Statistical Analysis

To account for familial correlations, we used a linear mixed effects model to compare the mean age of diagnosis across generations with a trend test. To study the possible impact of truncation bias in previous publications, we limited to generations to those affected who have at least 57 years of follow-up (born before 1955), which is the age at which more than 90% of all cases are diagnosed, and well beyond the average age of onset of HPAH. This eliminates individuals from later generations that include only affected children diagnosed at young ages without allowing for other children/siblings to develop disease.

These affected-only analyses do not consider the distribution of unaffected carriers in onset of disease. Because unaffected carriers may develop the disease in the future, we conducted Kaplan-Meier survival analysis by the number of generations since the index case (21). We censored unaffected carriers at the age they were last known to be alive or deceased by an unrelated cause of death. This Kaplan-Meier analysis does not consider familial correlations or within-family pairings of parents and offspring; however, this approach includes all known unaffected carriers as well.

To get an estimate of penetrance in our families, we analyzed all affected sibships in our database. In these sibships, about 50% of members should inherit the BMPR2 mutation. The calculated penetrance would thus be: number affected/(total in sibship $\times$ 0.5). We then compared penetrance in males and females and in aggregate. We included members up to date of birth 1965 to avoid splitting too many sibships, which span decades.

RESULTS

The HPAH registry contains 53 families in which the BMPR2 mutation has been identified. Of the 2,497 individuals who are/were at risk for disease, the mutation status is known in 722 individuals (29%), either through mutation testing, a positive diagnosis of PAH, or obligate carrier status. Of the 722 with known mutation status, 428 (59%) are positive. We analyzed 355 who are in the first four generations beginning with information about affection status. Table 1 shows the characteristics of the individuals in the registry who have tested positive for BMPR2 mutations and are included in this paper. The mean age of diagnosis or death is 34.6 ± 14.9 years. Of the 106 unaffected mutation carriers, 54% ($n = 57$) have died, compared with 85% ($n = 211$) of the 249 affected individuals, with 4 individuals lost to follow-up and censored at the date of diagnosis. Fifty-six unaffected individuals died of causes unrelated to HPAH.

The age at diagnosis or death is shown in Figure 1 for 115 affected family members by generation born before 1955. A test of linear trend for age of diagnosis or death is not significant ($P = 0.64$). We determined that the mean age of diagnosis or death is 43.6 years (SE, 2.5) for generation 1; 38.6 years (SE, 2.5) for generation 2; 38.6 years (SE, 3.7) for generation 3; and 46.9 years (SE, 6.1) for generation 4, with an overall P value of 0.27 for any group differences.

Figure 2 shows the biased analysis of all 249 affected members from 53 families born up to 2012, so that time bias was not minimized. Inclusion of younger generations shows a statistical decrease in age at diagnosis with succeeding generations, supporting genetic anticipation ($P < 0.0001$). The time truncation bias can also be shown by Spearman correlation

TABLE 1. CHARACTERISTICS OF FAMILY PEDIGREES WITH BMPR2 MUTATIONS

Characteristic	Value (N = 355)
Median pedigree size in BMPR2 (interquartile range)	4 (3–8)
n (%) in each generation	
First	93 (26)
Second	138 (39)
Third	81 (23)
Fourth	43 (12)
Female, n (%)	226 (64)
Affected, n (%)	249 (70)
Age at diagnosis of affecteds, yr (N = 249)	
Mean $\pm$ SD	34.9 ± 14.9
Range	0–74
Age at death or age of last contact in unaffecteds, yr (N = 106)	
Mean $\pm$ SD	66.7 ± 15.3
Range	20–98

Figure 1. Age at diagnosis of 115 patients from 30 families with two to four generations of members affected with heritable pulmonary arterial hypertension, born in or before 1955. The Tukey plots show median (bar within box), mean (bold star), interquartile range (boxes), the 10th and 90th percentile (whiskers), and outliers (beyond whiskers). There are no differences among generations 1-2-3-4 in any combination.

analysis of age at diagnosis of death and generation number, by sequentially restricting date of birth to require increasing length of follow-up. Including all births, the correlation is -0.23 ($P = 0.0002$). The magnitude of the correlation falls

to -0.18 ($P = 0.005$) before 1975; -0.06 ($P = 0.55$) before 1955; and -0.04 ($P = 0.76$) before 1935. These affected-only analyses include data on all generations, even if disease skipped a generation.

Figure 2. Age at diagnosis of 249 patients from 53 families with two to four generations of members with heritable pulmonary arterial hypertension born in or before 2011. Inclusion of recent generations without sufficient time for all at-risk members to express disease results in statistical lowering of the age at diagnosis ($P < 0.0001$), especially in the recent two generations.

Figure 3 is a Kaplan-Meier plot of age at diagnosis including all 249 affected members, plus 106 mutation carriers who do not have pulmonary hypertension and are censored by age at the latest time of contact or unrelated cause of death. Data that include younger family members require age at diagnosis, as many are still living. It is likely that some currently asymptomatic family members will become affected as they age in the future and would further influence the results against anticipation. Figure 3 includes the more recent biased sample plus all mutation carriers at risk for disease censored by age at last contact. The addition of this group eliminates statistical evidence for decreasing age at diagnosis in successive generations driven by the fourth generation ($P = 0.02$).

There was no difference in analysis based on age of onset versus age at death (data not shown). We performed the same analysis for the 44 non-BMP2 familial patients in our database and found statistical evidence for differences in the naive analysis, which disappeared with appropriate time truncation (data not shown).

To roughly assess penetrance, we counted 1,683 total siblings at risk from affected sibships, of whom 842 were likely to carry a mutation (50%). There were 232 total affected, including 177 females and 59 males. The overall penetrance was about 27%, the female to male ratio of affecteds is 3:1, the female penetrance is about 42%, and the male penetrance is about 14%. The overall 27% is slightly higher but similar to the current estimate of 20% in the literature (19).

DISCUSSION

Genetic anticipation is an old concept that has been verified in a small number of specific neurological diseases that are due to trinucleotide repeats and other expansions in the heritable gene, which usually causes dysfunction of the protein product (13). These expansions can enlarge during meiosis and cause an earlier age of onset in successive generations (13). However, the original observations that led to the theory of genetic anticipation were based on parent-child pairs in asylums, without regard to evaluation of siblings of either pair, and the theory was found to be flawed. Penrose listed several classes of bias of ascertaining disease that leads to the confusion (12). The first is initial discovery of a family by ascertainment of an offspring, which must lead to a time bias because the parent had to survive

long enough to procreate. The second is the inherent age bias in any parent-child pair, whomever is the first case discovered. The third is failure to follow sibs long enough to discover age at disease onset of every potential case. BMP2-related HPAH is a special example of the difficulty in ascertaining disease. Although it is autosomal dominant, only an estimated 20% of mutation carriers develop clinical pulmonary hypertension, and disease may manifest at any age (4, 6, 19, 20). Thus, decades must elapse before a generation of offspring can be assumed to be spared pulmonary hypertension. Finally, because most family members at risk have not been genotyped, it is unclear who is actually at risk and thus needs long term follow-up.

Our earlier analysis was based on ascertainment of two or more generations of affected family members, and in a paired analysis, the differences were highly significant (4, 6). We have collected data since the inception of the Vanderbilt Pulmonary Hypertension Registry in 1980 and now have pedigree data spanning back to 1900 in some families and as far forward as the present. Importantly, using mostly the same families (38 families have affected members before and after 1955), but censoring recent generations wherein insufficient time has elapsed, the previously observed generational differences disappear. Even the choice of 1955 as the cutoff birth age for analysis means that some mutation carriers could still potentially manifest disease at an older age, greater than 57 years. The addition of such carriers who manifest disease at a late age would further diminish the bias of ascertainment.

The cumulative incidence of disease is approximately 90% by 57 years of age; thus, we are likely capturing the majority of new cases (4). Using 57 years of follow-up eliminates evidence of the residual truncation bias, whereby the maximum difference between the first generation and last generation is 3.6 years. Using age of death alone as an outcome confirmed the results about lack of anticipation (data not shown).

Genetic anticipation is being sought in multiple diseases in which family clusters suggest a genetic link (11, 12, 15–18, 22). Recent analysis of Crohn registries has used methods similar to our new approach to HPAH, and the results cast doubt on the phenomenon of genetic anticipation in most diseases without trinucleotide repeat expansions (11). There is new evidence that diseases of telomere shortening may display genetic anticipation as another mechanism in pulmonary fibrosis, as shortening

Figure 3. A Kaplan-Meier plot of disease-free status of all 355 patients in the 53-family cohort, including known mutation carriers who are disease free and therefore censored (open circles) at the most recent age of query. Adding the censored members reverses the statistically significant differences by generation noted in the time biased analysis in Figure 2. FPAH = familial pulmonary arterial hypertension.

worsens in subsequent generations (23). Further research, time of observation, and appropriate analysis will reveal for these diseases whether genetic anticipation is a biological phenomenon or a statistical artifact. We have no information on whether earlier or easier detection has affected the apparent incidence of hereditary PAH relative to idiopathic PAH over the last decades. The overall percentage of hereditary pulmonary hypertension was similar in the recent French registry to the original National Institutes of Health registry (1, 2). This analysis does not eliminate the possibility in lead time biases in diagnosis because younger generations benefit from the knowledge that an older generation had a very rare disease.

Given that we do not know the mutation status in ~70% of at-risk individuals, we cannot make firm comments on penetrance in BMPR2-related PAH. Older unaffected family members with unknown mutation status in a pedigree remain at some risk regardless of age. Continued screening or surveillance for the onset of symptoms would be recommended for family members regardless of age. Our analysis of affected sibships does suggest that the overall penetrance of HPAH with BMPR2 mutations is about 27%, slightly higher than current estimates. Because the sex ratio is 3:1 affected females to males, the penetrance estimates should be adjusted to about 42% for females, 14% for males.

In summary, enough time must elapse to allow siblings to develop genetic disease to ensure that all affected cases are captured to allow for appropriate comparisons of age at diagnosis or death. With extended analysis, it appears that the phenomenon of genetic anticipation may not exist in BMPR2-related heritable pulmonary hypertension.

Author disclosures are available with the text of this article at www.atsjournals.org.

References

- Rich S, Dantzker DR, Ayres SM, Bergofsky ER, Brundage BH, Deter KM, Fishman AP, Goldring RM, Groves BM. Primary pulmonary hypertension: a national prospective study. *Ann Intern Med* 1987;107:206–213.
- Humbert M, Sitbon O, Chaouat A, Bertocchi M, Habib G, Gressin V, Yaici A, Weitzenblum E, Cordier JF, Chabot F, *et al*. Pulmonary arterial hypertension in France: results from a national registry. *Am J Respir Crit Care Med* 2006;173:1023–1030.
- Cogan JD, Vnencak-Jones CL, Phillips JA III, Lane KB, Wheeler LA, Robbins IM, Garrison G, Hedges LK, Loyd JE. Gross BMPR2 gene rearrangements constitute a new cause for primary pulmonary hypertension. *Genet Med* 2005;3:169–174.
- Loyd JE, Butler MG, Foroud TM, Conneally PM, Phillips JA III, Newman JH. Genetic anticipation and abnormal gender ratio at birth in familial primary pulmonary hypertension. *Am J Respir Crit Care Med* 1995;152:93–97.
- Melmon KL, Braunwald E. Familial pulmonary hypertension. *N Engl J Med* 1963;269:770–775.
- Loyd JE, Primm RK, Newman JH. Familial primary pulmonary hypertension: clinical patterns. *Am Rev Respir Dis* 1984;129:194–197.
- Sztrymf B, Yaici A, Girard B, Humbert M. Genes and pulmonary arterial hypertension. *Respiration* 2007;74:123–132.
- Sztrymf B, Yaici A, Sitbon O, Simmonneau G, Humbert M. Idiopathic pulmonary hypertension: what did we learn from genes? *Sarcoidosis Vasc Diffuse Lung Dis* 2005;26:352S.
- Loyd JE, Slovis B, Phillips JA, Butler MG, Foroud TM, Conneally PM, Newman JH. The presence of genetic anticipation suggests that the molecular basis of familial primary pulmonary hypertension may be tri-nucleotide repeat expansion. *Chest* 1997;111:82S–83S.
- McInnis MG. Invited Editorial. Anticipation: an old idea in new genes. *Am J Hum Genet* 1996;52:973–979.
- Picco JF, Goodman S, Reed J, Bayless TM. Methodologic pitfalls in the determination of genetic anticipation: the case of Crohn disease. *Ann Intern Med* 2001;134:1124–1129.
- Penrose LS. The problem of anticipation in pedigrees of dystrophin myotonia. *Ann Eugen* 1948;14:125–132.
- Ashley CT, Warren ST. Trinucleotide repeat expansion and human disease. *Annu Rev Genet* 1995;29:703–728.
- Schalling M, Hudson TJ, Buetow KH. Direct detection of novel expanded tri-nucleotide repeats in the human genome. *Nat Genet* 1993;4:135–139.
- Boonstra PS, Gruber SB, Raymond VM, Huang SC, Timshel S, Nilbert M, Mukherjee B. A review of statistical methods for testing genetic anticipation: looking for an answer in Lynch syndrome. *Genet Epidemiol* 2010;34:756–768.
- Bassett AS, Honer WG. Evidence for anticipation in schizophrenia. *Am J Hum Genet* 1994;54:864–870.
- McDermott E, Khan MA, Deighton C. Further evidence for genetic anticipation in familial rheumatoid arthritis. *Ann Rheum Dis* 1996;55:475–477.
- Fresko I, Soy M, Hamuryudan V, Yurdakul S, Yavuz S, Tümer Z. Genetic anticipation in Behçet's syndrome. *Ann Rheum Dis* 1998;57:45–48.
- Machado RD, Eickelberg O, Elliott CG, Geraci MW, Hanaoka M, Loyd JE, Newman JH, Phillips JA 3rd, Soubrier F, Trembath RC, *et al*. Genetics and genomics of pulmonary arterial hypertension. *J Am Coll Cardiol* 2009;54:S32–42.
- Newman JH, Wheeler L, Lane KB, Loyd E, Gaddipati R, Phillips JA III, Loyd JE. Mutation in the gene for bone morphogenetic protein receptor II as a cause of primary pulmonary hypertension in a large kindred. *N Engl J Med* 2001;345:319–324.
- Kaplan EI, Meier P. Nonparametric estimation of incomplete observations. *J Am Stat Assoc* 1958;53:457–481.
- McInnis MG, McMahon FJ, Chase GA, Simpson SG, Ross CA, DePaulo JR. Anticipation in bipolar affective disorder. *Am J Hum Genet* 1993;53:385–390.
- Armanios M. Telomerase and idiopathic pulmonary fibrosis. *Mutat Res* 2011;730:52–58.