Twin Research and Human Genetics
Volume 17 ▮ Number 6 ▮ pp. 505–511 ▮ © The Author(s) 2014 ▮ doi:10.1017/thg.2014.59

Improving the Power to Detect Risk Variants for Allergic Disease by Defining Case-Control Status Based on Both Asthma and Hay Fever

Manuel A. R. Ferreira
QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia

Asthma and hay fever are likely to share hundreds if not thousands of genetic risk variants. Despite this, the extent to which the power to identify shared risk variants could be improved by considering information from both diseases when designing or analyzing genetic studies has not been studied in detail. Simulations were performed to quantify the power to detect an association between case-control status and a bi-allelic risk variant shared between asthma and hay fever across a range of disease and genetic models, as well as different ascertainment and analytical strategies. For a fixed sample size, when designing a new genome-wide association study (GWAS), selecting for genotyping cases with both asthma and hay fever (A+H+), and controls with neither disease (A-H-) was the study design that provided the greatest power to identify a shared risk variant. On the other hand, when analyzing an existing GWAS, power was greatest across a wide range of scenarios, when cases were defined as individuals who suffered from either disease (A+ or H+) and controls as those who suffered from neither (A-H-). Bivariate analysis of asthma and hay fever provided comparable but slightly decreased power. In conclusion, new GWAS can be designed and existing GWAS reanalyzed more efficiently to identify risk variants for allergic disease by using ascertainment or analytical strategies that consider both asthma and hay fever information.

■ **Keywords:** genetic correlation, multivariate, pleiotropy, allergy, gene

Asthma and hay fever are two common, comorbid allergic diseases (Guerra et al., 2002; Leynaert et al., 1999) that have been estimated to share 50–90% of their genetic make-up (Duffy et al., 1990; Thomsen et al., 2006). Because both are highly polygenic, potentially hundreds or thousands of genetic risk variants are likely to be shared between the two diseases. Therefore, it is important to determine whether studies seeking to identify genetic risk factors for these diseases can be designed or analyzed more efficiently by incorporating case-control information from both asthma and hay fever.

A recent GWAS explored the possibility that genetic risk factors shared between asthma and hay fever can be identified with greater efficiency if the analysis was restricted to cases who suffer from both diseases and to controls who suffer from neither (Ferreira et al., 2014). Using a relatively modest sample size, 6,685 cases and 14,091 controls, a total of 11 independent common risk variants were identified at the genome-wide significance level. Importantly, all 11 variants had a significant but weaker association with asthma and hay fever when both diseases were considered as separate phenotypes. In fact, for three variants, the association did not reach the genome-wide significance level for risks of asthma or hay fever, despite the larger number of samples (up to 13,000 more) included in the analyses. These results provided empirical evidence that defining the case-control status of genotyped individuals based on information for both asthma and hay fever, instead of asthma alone, could have a significant impact on the power to identify risk variants that are shared between the two diseases.

In that study, however, no attempt was made to identify the disease and genetic models for which an improvement in power could be expected by excluding from the analysis individuals with one disease but not the other.

RECEIVED 13 August 2014; ACCEPTED 19 August 2014. First published online 9 October 2014.

ADDRESS FOR CORRESPONDENCE: Manuel A. R. Ferreira, PhD, QIMR Berghofer Medical Research Institute, Locked Bag 2000, Royal Brisbane Hospital, Herston QLD 4029, Australia. E-mail: manuel.ferreira@qimr.edu.au

https://doi.org/10.1017/thg.2014.59 Published online by Cambridge University Press

Therefore, the main goal of the present study was to quantify the power of the analytical approach used by Ferreira et al. (2014) across a range of models, and to compare it against the power obtained with alternative analytical approaches. Specifically, a series of simulations were performed to address the two following questions: first, when designing a GWAS, can the power to detect a risk variant shared between asthma and hay fever be improved over alternative ascertainment strategies by genotyping as cases individuals who have both asthma and hay fever, and as controls those who have neither disease? Second, when analyzing an existing GWAS, can the power to detect variants shared between asthma and hay fever be improved over alternative analytical strategies by defining case-control status based on both asthma and hay fever information? Addressing these questions may help design and analyze more efficiently GWAS for allergic disease.

Methods

Power of a New GWAS: Impact of Case-Control Definition Used for Sample Ascertainment

Simulations were performed to estimate the power to detect a genetic association between case-control status and a bi-allelic variant that influences the risk of both asthma and hay fever, using a fixed sample size ($N = 5{,}000$). Power was compared between three study designs that ascertain samples for genotyping based on three different case-control definitions:

> **Study design 1:** Cases are defined as individuals who suffer from asthma (A+), and controls as individuals who are asthma free (A-). This design ('A+ vs. A-') ignores hay fever status and is commonly used by published GWAS of asthma.
>
> **Study design 2:** Cases are defined as individuals who suffer from both asthma and hay fever (A+H+), and controls as those individuals who suffer from neither disease (A-H-). This study design is referred to as 'A+H+ versus A-H-'.
>
> **Study design 3:** Cases are defined as individuals who suffer from either asthma or hay fever (A+ or H+), and controls as those individuals who suffer from neither disease (A-H-). Explicitly, the cases include three groups of individuals: those with asthma but not hay fever (A+H-); those without asthma but with hay fever (A-H+); and those with both asthma and hay fever (A+H+). This study design is referred to as 'A+ or H+ versus A-H-'.

The power provided by each study design was assessed using simulated data, generated as follows. First, genotype data for a causal single nucleotide polymorphism (SNP) with 20% risk allele frequency were simulated in up to 4 million people. For each person, disease status was then simulated for asthma and for hay fever, assuming (1) a liability threshold model; (2) a heritability of 60% for both asthma and hay fever liabilities; (3) the causal SNP explained 0.1% of the variation in asthma liability; and (4) an environmental correlation between asthma and hay fever of 0.3.

Data were simulated for 192 (16 × 3 × 4) different models, obtained by setting (1) the population prevalence at 5%, 10%, 15% or 20% for asthma, and 15%, 25%, 35% or 45% for hay fever (4 × 4 = 16 joint scenarios); (2) the genetic correlation between the two disease liabilities at 0.3, 0.6 or 0.9; and (3) the proportion of variation in hay fever liability explained by the SNP (i.e., SNP heritability) at 0% (no effect on hay fever), 0.05% (half that for asthma), 0.1% (same as for asthma), or 0.2% (twice that for asthma).

After genotype, asthma and hay fever data were simulated for a given model in a population of up to 4 million individuals, we created three datasets ($N = 5{,}000$, 30% cases) for analysis as follows:

- Case-control dataset for study design 1: selective ascertainment of 1,500 individuals with asthma (A+) and 3,500 individuals without asthma (A-).
- Case-control dataset for study design 2: selective ascertainment of 1,500 individuals with asthma and hay fever (A+H+) and 3,500 individuals without asthma and without hay fever (A-H-).
- Case-control dataset for study design 3: selective ascertainment of 1,500 individuals with asthma or hay fever (A+ or H+) and 3,500 individuals without asthma and without hay fever (A-H-).

For each simulated dataset, the association between the causal SNP and case-control status was estimated by logistic regression and the observed asymptotic *p*-value retained. This procedure was repeated 500 times to obtain an estimate of power ($\alpha = 0.05$) for a given study design and model considered. Specifically, the power to detect a significant association was calculated as the proportion of 500 replicate datasets that had an association *p* value < .05.

Two additional association analyses were performed in the datasets generated for study design 1 (A+ vs. A-). In this study design, although samples were ascertained for genotyping based on asthma status alone, hay fever case-control status was also simulated and so was available for analysis — this mirrors real-life examples of asthma GWAS. Including this additional information may improve power to identify a significant association with a SNP that influences the risk of both asthma and hay fever. Therefore, a bivariate measure of association between the SNP and both asthma and hay fever was obtained using two methods. First by applying a formal multivariate test of association (Ferreira & Purcell, 2009) and second by performing two univariate analyses (asthma and hay fever separately), retaining the smallest *p*-value and correcting this for multiple testing through 200 permutations. Power for these two analyses was estimated as described above.

https://doi.org/10.1017/thg.2014.59 Published online by Cambridge University Press

Power of an Existing GWAS: Impact of Case-Control Definition Used to Select Samples for Analysis

A related but distinct question concerns the analysis of studies with existing SNP data. In this case, we used the simulation procedure described above to test if the power to detect variants that influence asthma risk and are shared with hay fever could be improved by defining case-control status of previously genotyped individuals based on both asthma and hay fever status, instead of asthma alone, as performed by Ferreira et al. (2014). Genotype, asthma and hay fever data were simulated using the same procedure described above for the same 192 models in a population of up to 4 million individuals. We then created two datasets with 5,000 individuals as follows:

- Asthma case-control dataset: selective ascertainment of 1,500 individuals with asthma (A+) and 3,500 individuals without asthma (A-). This corresponds to the same dataset generated for study design 1 above.
- Cross-sectional dataset: random ascertainment of 5,000 individuals, that is, without consideration for either asthma or hay fever status.

For both simulated datasets, three case-control association analyses were performed, defining case and control status respectively as: (1) A+ individuals versus A- individuals; (2) A+H+ individuals versus A-H- individuals, that is, individuals with one disease but not the other (A+H- and A-H+) were set to missing and so excluded from the analysis; and (3) A+ or H+ individuals (including A+H-, A-H+ and A+H+) versus A-H- individuals. Note that analyses (1) and (3) are based on a sample size of $N = 5{,}000$, whereas analysis (2) is based on a sample size of $N < 5{,}000$, with N varying across the 192 models tested (range for asthma case-control dataset: 3,000 to 4,058; range for cross-sectional dataset: 2,848 to 4,329). Two additional association analyses were performed in each dataset, as above: (4) bivariate analysis of asthma and hay fever status; and (5) separate univariate analyses of asthma (A+ vs. A-) and hay fever (H+ vs. H-), retaining the lowest p-value and correcting this for multiple testing through permutations. Analyses (4) and (5) were based on a sample size of $N = 5{,}000$. Data simulation and analysis was repeated 500 times to obtain an estimate of power for a given study design and model considered, as described above.

Results

Power of a New GWAS: Impact of Case-Control Definition Used for Sample Ascertainment

Simulations were performed to address the following question: when designing a GWAS, can the power to detect a variant that influences both asthma and hay fever risks be improved over alternative ascertainment strategies by genotyping as cases people who have both asthma and hay fever, and as controls people who have neither disease? The alternative study designs considered were (1) selecting cases and controls based on asthma status alone; and (2) selecting as cases, individuals who suffer from asthma or hay fever, and as controls, those who suffer from neither disease.

For a fixed sample size ($N = 5{,}000$, including 30% cases), a disease prevalence of 10% for asthma and 25% for hay fever, a SNP heritability for asthma of 0.1%, and assuming that both diseases share 60% of their genetic risk factors, selecting for genotyping A+H+ cases and A-H- controls (study design 2, green circles in Figure 1A) instead of A+ cases and A- controls (study design 1, blue diamonds in Figure 1A) considerably increased the power to detect an association with a risk SNP shared between the two diseases. Intuitively, the power gain was greatest when the SNP explained a larger fraction of the heritability of hay fever. On the other hand, there was a significant drop in power if individuals were selected for genotyping as cases if they suffered from either asthma or hay fever (including A+H-, A-H+ and A+H+), and as controls if they suffered from neither disease (study design 3, A+ or H+ vs. A-H-, yellow squares in Figure 1A). The exception to this was when the SNP heritability was higher for hay fever than for asthma. Similarly, bivariate association analyses of asthma and hay fever in the dataset with A+ cases and A- controls originally selected for genotyping (red up-triangles and orange down-triangles in Figure 1A) only provided improved power compared to the A+ versus A- univariate analysis when the hay fever SNP heritability was similar to, or higher than, that simulated for asthma (i.e., ≥0.1%). Comparable results were obtained for models with different disease prevalences and genetic correlations (supplementary Figures 1–3), as well as frequency of cases selected for genotyping (not shown). In general, the improvement in power obtained with study design 2 over study design 1 decreased with increasing genetic correlation and decreasing asthma prevalence.

To illustrate why study design 2 (A+H+ vs. A-H-) was more effective in improving power when compared to study design 3 (A+ or H+ vs. A-H-), the frequency of the risk allele was determined in each of the four disease groups: A+H+, A+H-, A-H+ and A-H-. When the SNP increased the risk of asthma but not hay fever (i.e., hay fever SNP heritability was 0%), there were two noteworthy observations: first, the frequency of the risk allele was lower in A+H+ when compared to A+H- asthma cases, and higher in the A-H- when compared to the A-H+ asthma controls. Both of these differences increased with increasing genetic correlation (supplementary Figure 4). As a result, study design 2 (A+H+ vs. A-H-) was less powerful than study design 1 (A+ vs. A-). Second, the risk allele frequency in the A-H+ group was the lowest of all four groups. Therefore, because in study design 3, most cases were ascertained from the A−H+ group (66% of all cases, compared to 20% from A+H+ and 14% from A+H- groups, Figure 1C), power was significantly reduced.

https://doi.org/10.1017/thg.2014.59 Published online by Cambridge University Press

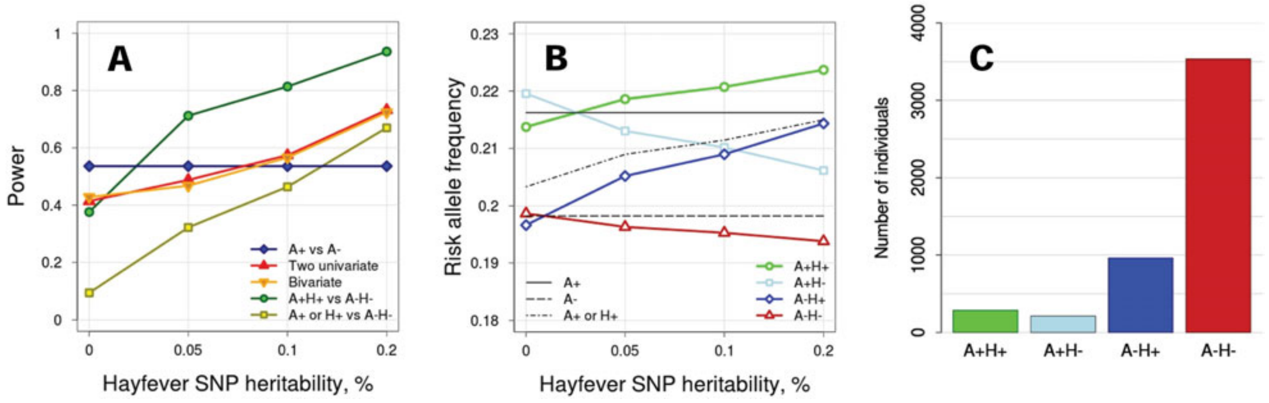

FIGURE 1

(Colour online) Impact of ascertainment strategy on the power to detect a risk variant shared between asthma and hay fever. A: Power according to the study design used to ascertain samples for genotyping. B: Frequency of the SNP risk allele in the overall population in subgroups of individuals defined by asthma and/or hay fever status. C: Number of individuals in each of the four subgroups defined by asthma and hay fever status, when 5,000 individuals were randomly ascertained from the overall population, assuming a population prevalence of 15% for asthma and 25% for hay fever, a genetic correlation of 0.6 and an environmental correlation of 0.3.

As the causal SNP explained an increasing proportion of hay fever heritability, the frequency of the risk allele steadily increased in the A+H+ group and decreased in the A-H- group (Figure 1B). As such, because study design 2 sampled cases exclusively from the A+H+ group and controls from the A-H- group, it was always the most powerful study design. On the other hand, study design 3 ascertained cases from two additional groups: in one of these (A+H-) the risk allele was progressively depleted with increasing hay fever SNP heritability, while in the other (A-H+) the opposite occurred (Figure 1B). Overall, this resulted in a risk allele frequency in the A+ or H+ case group that was always lower than that observed in the A+H+ case group and so study design 3 was always less powerful than study design 2.

Power of an Existing GWAS: Impact of Case-Control Definition Used for Analysis

A related but distinct question concerns the analysis of studies with existing genome-wide SNP data. In this case, the aim was to test if the power to detect variants associated with both asthma and hay fever risks could be improved over alternative analytical strategies by defining cases and controls based on disease status of both asthma and hay fever. The alternative analytical strategies considered were: (1) univariate analysis of asthma; (2) univariate analyses of asthma and hay fever separately, with correction for multiple testing; and (3) bivariate analysis of asthma and hay fever.

First, data were simulated under the assumptions summarized above for an asthma case-control study with 5,000 genotyped individuals, with selective ascertainment of 1,500 asthma cases (A+) and 3,500 asthma-free controls (A-); hay fever information was available for these 5,000 individuals but was not used to select samples for genotyping. When comparing A+ cases against A- controls, this dataset provided 54% power to identify a significant association ($\alpha = 0.05$) with a SNP that explained 0.1% of asthma heritability, irrespective of the effect of the same SNP on hay fever risk (blue diamonds, Figure 2A).

Next, cases were reclassified as those suffering from both asthma and hay fever (A+H+) and controls as those suffering from neither disease (A-H-); as a result, sample size dropped to 860 cases and 2,750 controls, a 43% and 21% reduction in sample size, respectively. With this case-control classification, power dropped to 26% when the SNP had no effect on hay fever (green circles, Figure 2A). However, power recovered to 52%, 63% and 74%, when the SNP hay fever heritability increased to 0.05%, 0.1% and 0.2%, respectively. These results demonstrate that when a SNP is a risk factor for the disease used for sample ascertainment and is also, to the same or greater extent, a risk factor for a genetically correlated disease, power can be improved by excluding from analysis genotyped individuals who suffer from one disease but not the other. The improvement in power was greatest when hay fever had a higher prevalence, but was largely unaffected by the degree of genetic correlation between the two diseases (supplementary Figures 5–7).

Power was also estimated using a different phenotype reclassification of the asthma case-control dataset of 5,000 genotyped individuals, in this case considering as cases individuals suffering from either asthma or hay fever (A+ or H+, $N = 2{,}250$) and as controls those suffering from neither disease (A-H-, $N = 2{,}750$); as such, the full sample size was retained in this analysis. For a prevalence of 10% for asthma and 25% for hay fever, results with the A+ or H+ versus A-H- classification (yellow squares, Figure 2A) were very similar to those obtained with the A+H+

https://doi.org/10.1017/thg.2014.59 Published online by Cambridge University Press

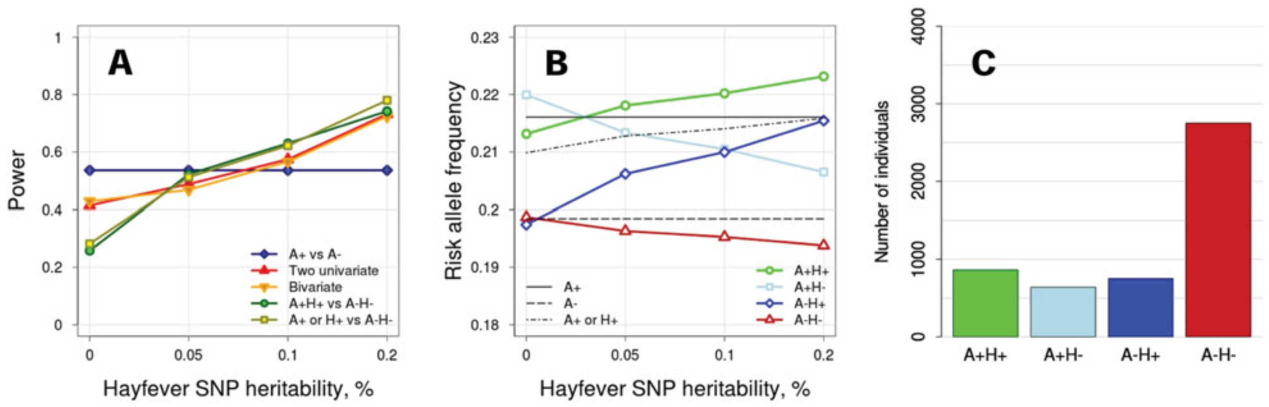

FIGURE 2

(Colour online) Impact of analytical strategy on the power to detect a risk variant shared between asthma and hay fever in an existing case-control GWAS of asthma. A: Power according to the phenotype classification used to define case-control status. B: Frequency of the SNP risk allele in a case-control GWAS of asthma (30% cases) in subgroups of individuals defined by asthma and/or hay fever status. C: Number of individuals in each of the four subgroups defined by asthma and hay fever status, when 1,500 asthma cases (A+) and 3,500 asthma-free controls (A-) were ascertained from the overall population, assuming a population prevalence of 15% for asthma and 25% for hay fever, a genetic correlation of 0.6 and an environmental correlation of 0.3.

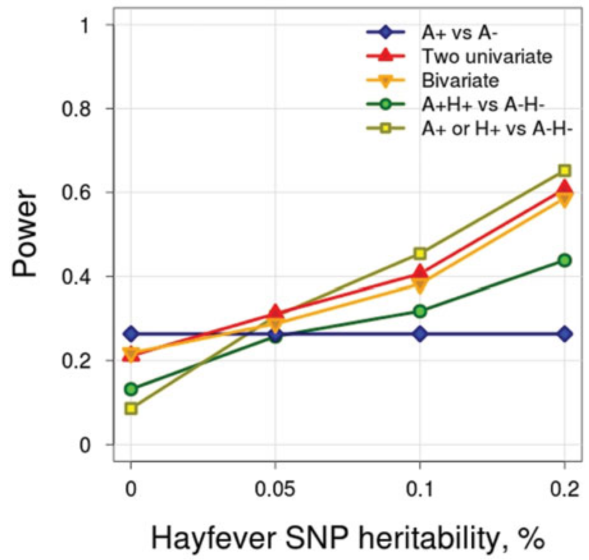

FIGURE 3

(Colour online) Impact of analytical strategy on the power to detect a risk variant shared between asthma and hay fever in an existing cross-sectional GWAS. The figure shows the power according to the phenotype classification used to define case-control status.

versus A-H- classification, suggesting that the lower risk allele frequency observed in the A+ or H+ group when compared to the A+H+ group (Figure 2B) was offset by the larger case sample size included for analysis (2,250 vs. 860, Figure 2C). On the other hand, for higher hay fever prevalences (35% or 45%) the A+H+ versus A-H- classification was the most powerful approach, while for lower hay fever prevalences (15%) A+ or H+ versus A-H- provided greatest power (supplementary Figures 5–7).

When the SNP influenced the risk of both diseases, bivariate analysis of asthma and hay fever provided similar, if slightly lower, power when compared to the A+H+ versus A-H- and A+ or H+ versus A-H- association analyses (Figure 2A).

Lastly, the same phenotype reclassification was also applied to a simulated cross-sectional study with 5,000 genotyped individuals; in this case, asthma and hay fever information was available for all 5,000 individuals but neither was used to select samples for genotyping. Figure 1C illustrates the average sample size for each of the four-disease groups in such a cross-sectional study, for a model with a disease prevalence of 10% for asthma and 25% for hay fever, a genetic correlation of 0.6 and an environment correlation of 0.3. The frequency of the risk allele in each of these four disease groups is illustrated in Figure 1B.

Reclassifying case and control status in the cross-sectional study resulted in the largest improvement in power to detect a risk locus shared between the two diseases when cases were defined as those suffering from either asthma or hay fever, and controls as those suffering from neither disease (A+ or H+ vs. A-H-, yellow squares, Figure 3). This approach provided considerable power gains over the A+ versus A- and the A+H+ versus A-H- classifications when the SNP effect for hay fever was similar to, or larger than, that for asthma. Similar results were obtained with different genetic correlations and disease prevalences (supplementary Figures 8–10). However, when the SNP affected asthma but not hay fever, this approach provided the lowest power. Bivariate analysis of asthma and hay fever provided an optimal balance between minimal drop in power when the SNP was not shared between the two loci, and significant improvement in power when the SNP was a shared risk factor.

https://doi.org/10.1017/thg.2014.59 Published online by Cambridge University Press

Discussion

Results from the simulations performed in this study demonstrate that when designing a new GWAS, the power to identify risk variants shared between asthma and hay fever can be substantially improved over alternative study designs by selecting for genotyping cases with both diseases and controls who have neither (A+H+ vs. A-H-). This result is intuitive: when the risk SNP influences the risk of both asthma and hay fever, the risk allele frequency is highest in the A+H+ group and lowest in the A-H- group, and so selectively ascertaining individuals from these two groups maximizes the allele frequency difference between cases and controls. On the other hand, genotyping cases with either disease and controls who have neither (A+ or H+ vs. A-H-) was only found to be an efficient approach to identify shared genetic risk factors, when the hay fever SNP heritability was larger than that of asthma. This was largely because in most models tested, the inclusion of a substantial number of A-H+ individuals in the case group resulted in a risk allele frequency in cases that was low compared to study designs that specifically excluded these individuals from the case group (i.e., A+ vs. A-, or A+H+ vs. A-H-). Therefore, new GWAS aiming to identify variants that influence allergic disease should consider adopting a study design that selectively ascertains A+H+ cases and A-H- controls for genotyping. The caveat of this approach is that the dataset generated cannot be used to estimate the risk of the identified variants on asthma and hay fever individually.

A related but distinct question regards the analysis of existing GWAS for which information is available on both asthma and hay fever status. Results from the simulations performed suggest that when the effect of the variant on hay fever liability is similar to or larger than that for asthma, analytical strategies that take into account both asthma and hay fever status can be more powerful than those that consider asthma status alone. Which analytical strategy was more powerful varied with the study design used to ascertain samples for genotyping (case-control or cross-sectional) and disease prevalence. In general, for existing case-control GWAS of asthma, power was greatest when individuals with one disease but not the other were excluded from analysis (A+H+ vs. A-H- classification), as performed by Ferreira et al. (2014). In this case, bivariate analysis of asthma and hay fever provided comparable if slightly lower power, as did the A+ or H+ versus A-H- approach. In contrast, for existing cross-sectional GWAS with both asthma and hay fever information available for analysis, the most powerful approach was to use the A+ or H+ versus A-H- classification, followed by bivariate analysis of asthma and hay fever. The A+H+ versus A-H- provided comparable power to these two approaches only when the prevalence of both diseases was high.

Therefore, collectively these results suggest that the A+ or H+ versus A-H- classification, or bivariate analyses of asthma and hay fever, may outperform the A+H+ versus A-H- approach used by Ferreira et al. (2014), across a wide range of scenarios and so should be considered in future reanalyses of existing GWAS of allergic disease. Results from bivariate analyses are less straightforward to meta-analyze and interpret, and so comparing cases who suffer from either disease against controls who suffer from neither is likely to provide a more practical analytical approach for consortium-driven studies. This approach has been recently used to identify genetic-risk factors shared between bipolar disorder and schizophrenia (Ruderfer et al., 2013), as well as between Crohn's disease and ulcerative colitis (Jostins et al., 2012).

In conclusion, results from this study provide further support to the notion that new GWAS can be designed — and existing GWAS reanalyzed — more efficiently to identify novel risk variants shared between asthma and hay fever by using ascertainment or analytical strategies that consider both asthma and hay fever information.

Acknowledgments

I would like to thank Peter Visscher, Nick Martin, John Hopper, David Evans and Melanie Bahlo for discussions and suggestions that helped improve this work. I am supported by a Career Development Fellowship from the National Health and Medical Research Council of Australia.

Supplementary Material

To view supplementary material for this article, please visit http://dx.doi.org/10.1017/thg.2014.59.

References

Duffy, D. L., Martin, N. G., Battistutta, D., Hopper, J. L., & Mathews, J. D. (1990). Genetics of asthma and hay fever in Australian twins. *American Review of Respiratory Disease*, *142*, 1351–1358.

Ferreira, M. A., & Purcell, S. M. (2009). A multivariate test of association. *Bioinformatics*, *25*, 132–133.

Ferreira, M. A., Matheson, M. C., Tang, C. S., Granell, R., Ang, W., Hui, J., . . . The Australian Asthma Genetics Consortium Collaborators. (2014). Genome-wide association analysis identifies 11 risk variants associated with the asthma with hay fever phenotype. *Journal of Allergy and Clinical Immunology*, *133*, 1564–1571.

Guerra, S., Sherrill, D. L., Martinez, F. D., & Barbee, R. A. (2002). Rhinitis as an independent risk factor for adult-onset asthma. *Journal of Allergy and Clinical Immunology*, *109*, 419–425.

Jostins, L., Ripke, S., Weersma, R. K., Duerr, R. H., McGovern, D. P., Hui, K. Y., . . . Cho,, J. H. (2012). Host–microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature*, *491*, 119–124.

Leynaert, B., Bousquet, J., Neukirch, C., Liard, R., & Neukirch, F. (1999). Perennial rhinitis: An independent

https://doi.org/10.1017/thg.2014.59 Published online by Cambridge University Press

risk factor for asthma in nonatopic subjects: Results from the European Community Respiratory Health Survey. *Journal of Allergy and Clinical Immunology*, *104*, 301–304.

Ruderfer, D. M., Fanous, A. H., Ripke, S., McQuillin, A., Amdur, R. L., Schizophrenia Working Group of the Psychiatric Genomics Consortium, . . . Kendler, K. S. (2013). Polygenic dissection of diagnosis and clinical dimensions of bipolar disorder and schizophrenia. *Molecular Psychiatry*. Advance online publication.

Thomsen, S. F., Ulrik, C. S., Kyvik, K. O., Ferreira, M. A., & Backer, V. (2006). Multivariate genetic analysis of atopy phenotypes in a selected sample of twins. *Clinical & Experimental Allergy*, *36*, 1382–1390.

https://doi.org/10.1017/thg.2014.59 Published online by Cambridge University Press