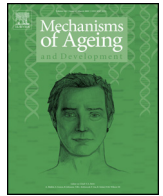




Contents lists available at [ScienceDirect](http://www.sciencedirect.com)

Mechanisms of Ageing and Development

journal homepage: www.elsevier.com/locate/mechagedev



Limitations and risks of meta-analyses of longevity studies

Paola Sebastiani^{a,*}, Harold Bae^b, Anastasia Gurinovich^c, Mette Soerensen^d,
Annibale Puca^{e,f}, Thomas T. Perls^g

^a Department of Biostatistics, Boston University School of Public Health, 801 Massachusetts Avenue, Boston, MA 02118, United States

^b College of Public Health and Human Sciences, Oregon State University, 151 Milam Hall, Corvallis, OR, United States

^c Bioinformatics Program, Boston University, 44 Cummington Mall, Boston, MA 02215, United States

^d The Danish Aging Research Center, Epidemiology, Biostatistics and Biodemography, Department of Public Health, University of Southern Denmark, J.B. Winsloews Vej 9 B, St., 5000 Odense C, Denmark

^e IRCCS MultiMedica, 20138 Milan (Mi), Italy

^f University of Salerno, Department of Medicine and Surgery, 84081 Baronissi, SA, Italy

^g Geriatrics Section, Department of Medicine, Boston University School of Medicine & Boston Medical Center, 88 E Newton St., Boston, MA 02118, United States

ARTICLE INFO

Article history:

Received 5 October 2016

Received in revised form 21 January 2017

Accepted 23 January 2017

Available online xxx

Keywords:

Human longevity

Meta-analysis

Odds-ratios

ABSTRACT

Searching for genetic determinants of human longevity has been challenged by the rarity of data sets with large numbers of individuals who have reached extreme old age, inconsistent definitions of the phenotype, and the difficulty of defining appropriate controls. Meta-analysis – a statistical method to summarize results from different studies – has become a common tool in genetic epidemiology to accrue large sample sizes for powerful genetic association studies. In conducting a meta-analysis of studies of human longevity however, particular attention must be made to the definition of cases and controls (including their health status) and on the effect of possible confounders such as sex and ethnicity upon the genetic effect to be estimated. We will show examples of how a meta-analysis can inflate the false negative rates of genetic association studies or it can bias estimates of the association between a genetic variant and extreme longevity.

© 2017 Elsevier B.V. All rights reserved.

1. Introduction

The hunt for genetic and non-genetic factors that promote human longevity continues to fascinate and engage aging-researchers. While substantial progress has been made in the epidemiology of extreme human longevity, particularly regarding evidence supporting Fries's "compression of morbidity" hypothesis in oldest centenarians (Andersen et al., 2012; Fries, 1980; Sebastiani et al., 2013a; Ismail et al., 2016), the discovery of genetic factors that promote longevity and extreme longevity has been challenged by the rarity of the phenotype, the need for large samples to reach an extreme level of statistical significance in genome-wide association studies and also, in the case of association studies, the lack of clarity in the definition of both cases and controls. These challenges are related and often work against each other. For example, we have

shown that the heritability of longevity expressed as sibling relative risk increases with more and more extreme definitions of longevity (Sebastiani et al., 2015), and while using an extreme definition of longevity should result in a more heritable phenotype, accruing a sufficiently large sample of very old individuals is very difficult. It has taken the New England Centenarian Study more than 20 years to accrue almost 200 supercentenarians, those who have survived to age 110 years and older (Sebastiani and Perls, 2012).

In order to boost their statistical power, some studies relax their definition of longevity (Erikson et al., 2016), or the definition of controls, and more recently meta-analysis has emerged as a way to increase statistical power by aggregating results from many smaller studies. While the method of meta-analysis has many important properties and can be useful to increase evidence in support of or against a hypothesis, it can often produce misleading results. Here we discuss some of the major challenges of meta-analysis of longevity studies.

* Corresponding author.

E-mail addresses: sebas@bu.edu (P. Sebastiani), Harold.Bae@oregonstate.edu (H. Bae), agurinov@bu.edu (A. Gurinovich), msoerensen@health.sdu.dk (M. Soerensen), apuca@unisa.it (A. Puca), thperls@bu.edu (T.T. Perls).

<http://dx.doi.org/10.1016/j.mad.2017.01.008>

0047-6374/© 2017 Elsevier B.V. All rights reserved.

2. Meta-analyses of genome-wide association studies of longevity

2.1. Definition of genetic effect

Genetic association studies of longevity typically use a case-control study design, in which cases are individuals who have reached some defined old age and controls are often a random sample from the population with the assumption that the phenotype is so rare, it is assumed demographically unlikely that the control will eventually survive to the age of interest. The genetic effect of an allele g that can be estimated with data collected using this study design is the odds ratio for extreme longevity, comparing carriers and non-carriers of the g allele. This odds ratio (OR) is defined as

$$\theta = \frac{p(EL|g)/(1 - p(EL|g))}{p(EL|\bar{g})/(1 - p(EL|\bar{g}))} \quad (1)$$

where EL denotes “extreme longevity”, $p(EL|g)$ is the probability of achieving extreme longevity in carriers of the g allele, and $p(EL|\bar{g})$ is the probability of achieving extreme longevity in non-carriers of the g allele. This odds ratio is equivalent to

$$\theta = \frac{p(g|EL)/(1 - p(g|EL))}{p(g|AL)/(1 - p(g|AL))} \quad (2)$$

where $p(g|EL)$ is the prevalence of the g allele in cases, AL denotes “average longevity”, and $p(g|AL)$ is the prevalence of the g allele in controls. The parameter θ in Eq. (2) is “exposure-odds” that is the estimable quantity in a case-control study design and it can be converted into the parameter of interest in Eq. (1), i.e. the “disease odds”, or in this case, “EL odds” by using Bayes’ theorem (Jewell, 2003).

Many factors can affect the magnitude and statistical significance of the genetic effect θ , particularly definitions of the case (EL), of the control (AL), genetic confounders such as ethnicity that may affect the prevalence of the g allele in cases and controls, and of course the sample size. Regarding controls, they should ideally be matched by birth year to avoid unmeasurable confounding due to secular effects. However, even the longest running longitudinal studies have not been able to adhere to such an inclusion criterion for controls. Thus, most studies settle for controls who have not reached a certain age. As discussed later, this definition can also be problematic.

Meta-analysis of results from different studies has become a standard procedure in genetic association studies to remedy the limited sample sizes of individual studies. The underlying assumption of this approach is that “more is always better” and aggregating results from different studies will strengthen the results and increase the statistical significance of the true positive associations. Using real data, we show however that this approach can lead to more false negatives, not fewer.

2.2. Meta-analysis: non-ignorable assumptions

A meta-analysis is a statistical method to summarize results from different studies that has become extremely common in genetic epidemiology (Evangelou and Ioannidis, 2013). A meta-analysis of genetic effects estimated through case-control studies will typically receive as input the estimated odds ratios and standard errors from each study and aggregate the results using some form of weighted average (fixed-effects meta-analysis). Weights defined as the inverse of the standard errors are used in the “inverse-variance weighting” method of meta-analysis that appears to be the most common approach in genetic epidemiology (Evangelou and Ioannidis, 2013). An alternative weighting system is used in Mantel Haenszel (MH) meta-analysis that results in a more robust estimate and standard error when some of the

studies are small (Mantel and Haenszel, 1959) or when the genetic variant is rare. The key assumption underlying a fixed-effects meta-analysis is that the different studies estimate the same population parameter and differences in the study-specific effect sizes are the result of sampling variability (Borenstein et al., 2010). Heterogeneity among studies can be accounted for by using a random-effects meta-analysis, which essentially assumes a hierarchical model to describe the within-study variability and the variability between effects estimated from different studies. Although a random-effects meta-analysis accommodates more variability, it still makes the assumption that there is a well-defined “population parameter” (odds ratio) to be estimated. If heterogeneity of the studies originates from study-specific genetic effects then a meta-analysis can be inconclusive or produce misleading results.

2.3. Published meta-analyses of genome-wide association studies of longevity

A comprehensive review by Broer and van Duijn (2015) lists just 5 genome wide association studies of longevity (Newman et al., 2010; Deelen et al., 2011; Nebel et al., 2011; Malovini et al., 2011; Walter et al., 2011), and one of extreme longevity (Sebastiani, 2012) published up to 2012. The article by Newman et al. (Newman et al., 2010) reported a meta-analysis of 4 genome-wide association studies from the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium. The meta-analysis included 1836 individuals who survived to age 90 and older (cases) and 1955 individuals who died at ages between 50 and 80 years (controls). Despite what was a large sample size for a study of longevity, this meta-analysis did not identify any genome-wide significant associations, not even for variants in the well replicated *APOE* locus on chromosome 19 that has achieved genome-wide significance levels with much smaller studies of centenarians (Sebastiani, 2012). Additional meta-analyses that expanded upon findings from the CHARGE consortium were reported by Deelen et al. (2014) and Broer et al. (2015). The meta-analysis in Deelen et al. (2014) included 14 different studies and increased the overall sample size of the discovery cohort to 7729 long lived individuals and 16,121 controls. However, the large sample size was reached by using a “relaxed definition” of longevity as survival to age 85 and older, and only SNPs in the *APOE* locus reached genome-wide significance in the discovery cohort. The meta-analysis in Broer et al. (2015) used data from new studies that joined the CHARGE consortium to synthesize the results of genome-wide association studies of longevity defined as surviving beyond age 90. This analysis also identified only SNPs in the *APOE* locus with genome-wide significance.

The lack of novel genetic findings linked to human longevity from these meta-analyses has been exclaimed as evidence of a weak genetic contribution to the phenotype of extreme human longevity and has fed the myth, based upon twin studies of survival to the mid octogenarian years, that the heritability of survival to extreme old age is approximately 25% (Sebastiani and Perls, 2012; Broer et al., 2015; Newman and Murabito, 2013). However, increasing the sample size through a meta-analysis does not necessarily increase the statistical power if including a large number of heterogeneous studies decreases the signal to noise ratio. We argue that an inconsistent definition of the phenotype across studies is a possible source of heterogeneity of the study-specific effects that may reduce the usefulness of a meta-analysis. In fact, we and others have shown that the sibling relative risk of extreme longevity defined as living beyond a threshold age increases as the threshold age increases (Sebastiani et al., 2015; Perls et al., 2002; Perls, 1998), and we have advocated for more standard definitions of extreme longevity by choosing the threshold age based, for example, on percentile survival from some standard life tables. A consistent and specific definition of extreme longevity should result in a

متن کامل مقاله

دریافت فوری ←

ISIArticles

مرجع مقالات تخصصی ایران

- ✓ امکان دانلود نسخه تمام متن مقالات انگلیسی
- ✓ امکان دانلود نسخه ترجمه شده مقالات
- ✓ پذیرش سفارش ترجمه تخصصی
- ✓ امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
- ✓ امکان دانلود رایگان ۲ صفحه اول هر مقاله
- ✓ امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
- ✓ دانلود فوری مقاله پس از پرداخت آنلاین
- ✓ پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات