Computers in Biology and Medicine 43 (2013) 1293–1299

Contents lists available at ScienceDirect

Computers in Biology and Medicine

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

Examining the stability of genetic risk effect as evidence accumulates in the context of meta-analysis

Elias Zintzaras [a,b,*]

[a] Department of Biomathematics, University of Thessaly School of Medicine, Larissa, Greece
[b] The Institute for Clinical Research and Health Policy Studies, Tufts Medical Center, Tufts University School of Medicine, Boston, MA, USA

ARTICLE INFO

Keywords:
Meta-analysis
Cumulative meta-analysis
Genetic association
Risk
Stability
Monte Carlo test

ABSTRACT

The recursive cumulative meta-analysis (RCM) of genetic association studies explores the relative change of the cumulative risk effect (e.g. *OR*) in time, indicating the stability of risk effect as evidence accumulates. However, the stability in risk effect is currently evaluated empirically with a graphical approach. A Monte Carlo permutation test for examining the instability in RCM is proposed. The statistic used is a function of the difference between the observed change in risk effect and the expected change, and is expressed (stepwise) cumulatively from the last published GAS to the first one. The permutation method is based on the individual studies and the number of studies in each time step. The test was demonstrated using data from two large scale meta-analyses of GAS. The performance of the test was also explored by simulating data from meta-analyses with different settings in terms of heterogeneity and significance. Significance instability was detected when wide oscillations in risk effect were presented and vice versa. The proposed test for assessing stability may provide the framework for claiming or denying the existence of an association as evidence accumulates.

1. Introduction

Genetic association studies (GAS) test the association between genetic variants and disease [1,2]. However, GAS may lack reproducibility (i.e. later studies often fail to confirm the positive results of initial studies) and they frequently generate controversial or inconclusive results [1]. Meta-analysis is a robust tool to investigate this type of discrepant results.

In addition to conventional meta-analysis, cumulative meta-analysis (CM) and recursive cumulative meta-analysis (RCM) provide a framework to investigate the trend and stability of risk effect as evidence accumulates, respectively [3–5]. In CM, studies are ordered by a covariate (e.g., publication year) and then, the pooled risk effect (e.g. odds ratio, *OR*) is obtained when a new study is published, i.e. at each information step. In RCM, the relative change in cumulative risk effect in each information step is calculated [1]. In addition to updating the genetic effects, CM provides a measure of how much the genetic effect changes as evidence accumulates. Furthermore, RCM predicts major changes in risk effect that may occur in the future [1,4]. Consequently, CM and RCM are methods to explore heterogeneity in risk effect for a genetic model over time. If oscillations in risk effect remain in time, then more information is required to draw safe conclusion on the magnitude of the risk effect [1,4]. Therefore, it is important to examine the results of CM and RCM prior to conclusions of pertinent gene–disease associations. However, stability in RCM is currently evaluated empirically with a graphical approach [4]. A formal statistical test for examining instability of risk effect, and consequently trend of association, as evidence accumulates, does not exist.

Hereby, a Monte Carlo permutation test to examine instability in RCM is proposed. The statistic used is a function of the difference between the observed change in risk effect and the expected change, and is expressed (stepwise) cumulatively from the last published GAS to the first one [6]. The test was illustrated using data from two meta-analyses [9,10] and simulated data from meta-analyses with different settings in terms of heterogeneity and significance.

2. Methods

2.1. Meta-analysis

The integral parts of a meta-analysis are the estimation of a summary metric and the exploration of heterogeneity between studies. In a meta-analysis for a dichotomous outcome, the association (or risk effect) for a specific genetic model is indicated

* Correspondence address: Department of Biomathematics, University of Thessaly School of Medicine, 2Panepistimiou St, Biopolis, Larissa 41110, Greece. Tel.: +1 30 2410 565063.
E-mail address: ezintzaras@tuftsmedicalcenter.org

0010-4825/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.compbiomed.2013.04.017

as pooled odds ratio ($\theta=OR$) with the corresponding 95% confidence interval (CI). The heterogeneity between studies is tested using the Q-statistic [1]. If $P<0.10$ then heterogeneity is considered to be statistically significant. The pooled *OR* is conventionally estimated using random effects (RE) model [1]. Random effects modeling assumes a genuine diversity in the results of various studies, and it incorporates to the calculations between study variances. When there is a lack of heterogeneity the *RE* model coincides with the fixed effects model [1]. In the present study, only the *RE OR* was used for all analyses since it is more conservative.

In CM, studies are ordered by a covariate (e.g. publication year) and then, the pooled *OR* is updated at each information step [1,3–4]. In RCM, the relative change in cumulative pooled *OR* in each information step [cumulative *OR* in next year (y+1)/ cumulative *OR* in current year (y)] is calculated [1,3]. Thus, CM and RCM indicate the trend in estimated risk effect and the stability in risk effect, respectively. When stability in risk effect is reached, then the relative change in cumulative *OR* will approach one [4].

2.2. Test statistic

In order to test instability in RCM, the following metric was calculated for each information step (i.e. the change from the current step to the next step):

$$D_\ell = \sqrt{w_\ell^*[\theta_\ell - E(\theta_\ell)]^2},\ \ell = 1,\ldots s \text{ steps},$$

where $w_\ell = n_\ell/n_s$ is the weighting factor for step i, n_ℓ is the cumulative number of subjects in step l, θ_ℓ is the change of the effect size (i.e. $\theta_\ell = OR_{\ell+1}/OR_\ell$ where *OR* denotes the cumulative *OR*) and E is the expected value of θ_ℓ which is set to $E(\theta_\ell)=1$.

In stability, it is expected that the effect size θ does not change from the current year to the next year, and thus, for step ℓ the change of the effect size is $E(\theta_\ell)=1$. In a study involving k years, the number of total step changes is $s=k-1$. Then, instability was tested using the following statistics:

$$S_r = \sum_{l=(s+1)-r}^{s} D_l,\ \forall\ \ell = 1,\ldots,s \text{ steps}$$

Thus, there are several test statistics S_r, one for each r. The summation runs from $\ell=(s+1)-r$ to $\ell=s$, that is, it goes backwards in time, covering always the last r (not the first) steps (or years). The P-value of S_r (P_{S_r}) is determined using the Monte Carlo permutation test. Then, instability is reached at step r when $P_{S_{r'}} \geq 0.05\ \forall\ r' \leq r-1$ and $P_{S_r} < 0.05$. Thus, the period covered by the steps 1 to $r-1$ are characterized by stability.

2.3. Monte Carlo permutation test

In order to assess the statistical significance of the metrics, a Monte Carlo method was used. Let us assume that N studies are included in the meta-analysis and ν_i is the number of studies involved in each step i so that $N(N-1)=\Sigma^k \nu_i$. The permutation method is based on two independent elements: the N studies and the number of studies ν_i. Then, in a run, the effect sizes θ_i (N studies in total) are randomly permuted so that in step i corresponds ν_i studies. At the same time, the number of studies involved in each step across all steps is also randomly permuted. This two stage random permutation procedure ensures complete randomness and keeps the structure of the meta-analysis unchanged. After each run, the value of the S_r test-statistic is computed from the permuted data, called $S_{r'}$. The procedure is repeated for 1,000,000 runs and a null distribution of the statistic is constructed.

The null hypothesis is that the observed changes in θ across steps are unrelated to the expected value which was set equal to one (i.e. no change, or stability). Then, each permutation of observed changes in *OR* relative to no changes is equally likely. The significance of observed S_r is assessed empirically against the null distribution of $S_{r'}$. The significance level $P_{S_r} = P(S_{r'} \geq S_r)$, is estimated as the proportion of randomized test statistics less than or equal to the observed test-statistic value. A significantly small value of the observed S_r statistic value relative to the distribution of randomized values is evidence against the null hypothesis of stability (i.e. existence of instability). The test was implemented using Compaq Visual Fortran90 with IMSL library [9].

3. Results

The test is now illustrated using data from two meta-analyses, alcohol dehydrogenase 2 (*ADH2*) **1/*2* variant vs. alcoholism [7] and angiotensin-converting enzyme (*ACE*) gene insertion/deletion (*I/D*) variant vs. coronary artery disease (CAD) [8], and simulated data from meta-analyses with different settings in terms of heterogeneity and significance.

3.1. ADH2 *1/*2 variant and alcoholism

The meta-analysis involved 33 studies that investigated the association between the *ADH2 *1/*2* variant and alcoholism (Fig. 1). The genetic model of allele contrast was adopted for the analysis [1]. Details on the analyzed studies and meta-analysis results are provided in Zintzaras et al. [9]. The studies were published from 1990 to 2004, providing 15 distinct information steps (years). The allele contrast *ADH2 *1* vs. **2* showed significant and large heterogeneity among studies ($P<0.01$) and significant association: RE $OR=1.79$ (95% CI: 1.47, 2.17). In CM, the *OR* showed an upward trend in the period 1990–1996, and then it remained fairly constant (Fig. 2). In RCM, instability in the *OR* change (θ_ℓ) appeared before year step 1996, whereas in the period 1996–2004 the *OR* was stable (Fig. 3). Thus, from the first published study (1992) till 1996, there are oscillations indicating bias or heterogeneity, and therefore, valid conclusions regarding association can not be made. Bias or heterogeneity can be due to a number of causes, related to study design, study quality (including departures from Hardy–Weinberg equilibrium), sample size, gender and ethnic or environmental heterogeneity between populations.

Using the Monte Carlo permutation test, the S_r statistic was non-significant (P_{S_r}-values ranged from 0.83 to 0.94) for the last eight information steps (2004/2003 to 1997/1996) and then, in the 9th step (1996/1995), the significant level P_{S_9} was significant ($P_{S_9} < 0.01$). The P_{S_r}-values up to the first significant step are shown in Fig. 4, and the null distribution for S_9 statistic is depicted in Fig. 5. Thus, the test verified the previous visual evaluation. There is indication that *OR* reached stability after year 1996. Given that the association remained significant after this year, it is reasonable to conclude that the accumulated evidence is sufficient for claiming association. Therefore, it may not be necessary to replicate the findings of the association further in future studies.

Of course, the above conclusion does not take into account different sources of heterogeneity, such as ethnic diversity (mainly Whites and East Asians) and the results should be interpreted with caution. Then, a more detailed analysis should consider a stability testing for each ethnicity separately (e.g. in Whites stability reached after 1992 whereas in East Asians, after 1994). In addition, the *OR* was calculated for the allele contrast since it is a model-free approach; alternatively, the generalized odds ratio can be used [10]. However, other genetic models (such as dominant, recessive, additive or co-dominant) can be applied, only, if there is a prior knowledge of the mode of inheritance [11]; nevertheless, the same model must be used for all individual studies.