Authors: Carol J. Etzel and Tracy J. Costello
Department of Epidemiology, The University of Texas M.D. Anderson Cancer Center, Houston, Texas 77030

3.6 COMBINING INFORMATION ACROSS GENOME-WIDE SCANS

3.6.1 Introduction

With the formation of international consortia to investigate complex disorders and a variety of cancers, meta-analysis is quickly becoming a valuable tool to combine linkage results and narrow chromosomal regions of interest. The presumed etiology of a complex disease is a combination of effects from multiple genes and the environment. The possibility of identifying some of these genes, which most likely have small effects, from one, independent study using traditional linkage analysis methods, is small. Instead, pooling raw data across independent studies (i.e. a mega-analysis) or pooling linkage results across independent studies (i.e. a meta-analysis) may be the best means to identify these numerous genes with typically small effects. Although among-study heterogeneity that may include differing marker maps, marker informativity, sample sizes, phenotype definition, ascertainment schemes, and linkage tests can be problematic for a meta-analysis; methods have been proposed to handle such problems and will be discussed here.

The basis of meta-analytic methods in genetic linkage is derived from pooling methods that have been available in the field of statistics for over 75 years. Such noble statisticians as Fisher (1925), Tippett (1931), and Pearson (1933) provide the earliest references to meta-analysis. These early methods were based on testing a consensus or omnibus null hypothesis (i.e., all null hypotheses from the individual studies are true) by combining the p-values from each of the individual studies. These methods are nonparametric in the sense that they do not rely on any distributional assumptions regarding the data in the individual studies; however, it is assumed that each of the studies are testing a common (and combinable) null hypothesis. Folks (1984) provides an excellent review of these early meta-analytic methods and the reader is encouraged to refer to Folks for a more detailed description of each of these early methods.

Meta-analysis for genome-wide scans has roots in methods developed to complete a meta-analysis for individual markers. Such methods involved pooling p-values (using Fisher's method, 1925) or genetic effects or estimates of proportion of alleles shared identity by descent (*ibd*) among relative pairs (Li and Rao 1996; Gu et al. 1998). In this chapter, we will review recent applications and extensions of meta-analytic methods for combining information across independent genome scans. We will also provide the reader with strategies to choose the method to best suit his/her meta-analytic needs. The reader will have a better understanding of the need for and application of meta-analysis in the field of genetic linkage involving genome scans.

3.6.2 Meta-Analytic Methods for Genome Scans

In this section, we review meta-analytic methods that have been proposed and applied to genome-wide scan studies. Our coverage of such methods may not be exhaustive as we have tried to focus on such methods where power and type I error have been evaluated or methods (due to their ease of application) that have been widely used.

3.6.2.1 Meta-analytic methods based on p-values and tests of significance

As mentioned in the Introduction, general applications of meta-analysis have been developed from methods based on combining p-values. The method proposed by Fisher (1925) has been widely used in genetic linkage and many extensions have been developed for meta-analyses involving genome-wide scans. Suppose that we wish to complete a meta-analysis on *k* studies. Each study *k* has m number of markers. Let M_{st} denote the t^{th} marker from study *s*, for $s = 1,\ldots k$ and $t = 1,\ldots,m$. Further define p_{st} as the p-value for marker M_{st} that provides evidence for linkage at marker M_{st} . We are not assuming that each study used the same sampling scheme or linkage test; however the studies must be testing the same null hypothesis of no linkage. Using Fisher's (1925) method, we can define

$$X_t^2 = -2\sum_{s=1}^{k} \ln\left(p_{st}\right) \tag{1}$$

as the combined evidence for linkage at marker $M_{.t}$ across all studies. We can further define the p-value associated with X_t^2 as

$$P_t = \Pr\left(\chi_{2k}^2 > X_t^2\right) \tag{2}$$

where χ_{2k}^2 is distributed as a chi-square variate with *2k* degrees of freedom. The power and type I error of this method was evaluated by Guerra et al. (1999) where a per marker alpha level of 0.1% was used to account for genome-wide testing. They concluded that although Fisher's method is applicable for genome scans, the power to detect linkage using this method is was not equivalent to that achieved by pooling raw data.

One of the caveats to using this method to complete a genome-wide meta-analysis is that an investigator is not guaranteed that all of the studies included in a meta-analysis will have used the exact same marker map. Or if the investigator is relying on published data, it is not guaranteed that results of all linkage studies are published, or of those that have been published, that results for all markers involved in a particular study will be readily available. Instead only information on local minimum p-values may reach publication. Therefore,

the straightforward application of Fisher's method may not be feasible. Alternatives to Fisher's method have been proposed (informally and formally) in order to apply this meta-analytic method across whole regions of the human genome instead of single loci. One such informal application was proposed by Allison and Heo (1998) to combine data from several studies that used different tests for linkage and different markers to detect linkage within the Human *OB* region. Their technique involved obtaining a single p-value within the *OB* region from each of five published studies that investigated linkage to body mass index using different testing procedures for different sets of markers. Fisher's method was then used to combine the p-values across the five studies. They concluded that meta-analysis is a vital statistical tool that highlights the importance of published literature in the absence of available raw data and increases the power to detect genes influencing complex traits. They note that their approach illustrates that one can conduct a meta-analysis over multiple linkage studies investigating a single phenotype despite what they describe as "worst case conditions." However, we argue that the situations that Allison and Heo describe are realistic of early linkage publications and worst case conditions are those in which no meta-analysis can be performed.

Badner and Gershon (2002b) formally considered a similar modification of Fisher's method so that meta-analysis can be performed for regions across the human genome instead of one marker at a time. In their paper, they defined equation (2) as the Multiple Scan Probability (MSP) with p_{st}^{*} substituting for p_{st}, where p_{st}^{*} is defined as the minimum observed p-value for study *s* over a specified linkage region *t* corrected for the size of the linkage region. Their correction factor was based on the Feingold et al. (1993) estimate of the probability of a p-value being observed in a specified region size, namely

$$p_{st}^{*} = Cp_{st} + 2\lambda GZ\left(p_{st}\right)\varphi\left(\Phi^{-1}\left(p_{st}\right)\right)V\left[\Phi^{-1}\left(p_{st}\right)\sqrt{4\lambda\Delta}\right] \quad (3)$$

where p_{st} is the observed p-value from study *s* over region *t*, C is the number of chromosomes, λ is the rate of crossovers per Morgan (which varies based on the linkage method employed and family structure), G is the size of region *t* in Morgans, $\Phi^{-1}(\bullet)$ is the standard normal inverse function, $\varphi(\bullet)$ is the normal density function, Δ is the average distance in Morgans between adjacent markers and the function *V* is a discreteness correction factor for Δ. Feingold et al. (1993) show that $V\left(x\right) \approx \exp\left(-0.583x\right)$, for $x < 2$. Under certain conditions, they also show that equation (3) is equivalent to the Lander and Kruglyak (1995) p-value correction factor. Badner and Gershon (2002b) show via simulation that the type I error rate for this modification is at least as low as for any single genome scan study and that power to detect linkage using this method is equivalent to that of pooling raw data. This method has been applied

to studies involving autism (Badner and Gershon 2002b) and bipolar disorder and schizophrenia (Badner and Gershon 2002a).

Another caveat to applying Fisher's method to genome-wide scans is that many widely used linkage tests are one-sided (i.e., LOD scores have a lower bound of 0) whereas the distributional assumptions for Fisher's original method assume that the p-values were derived from two-sided tests. Province (2001) suggested an extension of Fisher's general method to adjust for the potential bias of combining linkage results from such one-sided tests. Citing the one-to-one correspondence between LOD scores and p-values (Ott 1991)

$$p_{st} = 1 - \Phi\left[sign(LOD_{st})\sqrt{2\ln(10)\left|LOD\right|} \right], \qquad (4)$$

where $\Phi(\bullet)$ is the standard normal distribution function, Province recommended that LOD scores equal to zero should be assigned a p-value equal to $\frac{1}{2\ln(2)} \approx 0.72$ instead of equal to 0.50 as given by equation (4) or equal to 1.0 as suggested by maximum-likelihood theory. By doing so, the resulting test statistic obtained from Fisher's method using p-values extracted from published or derived LOD scores would roughly follow the assumed chi-square distribution with the appropriate number of degrees of freedom (2 times the number of studies) under the null of no linkage. This extension of Fisher's method has been applied to genome scan studies involved in the National Heart, Lung and Blood Institute Family Blood Pressure Program looking for obesity-related genes (Wu et al. 2002), hypertension-related genes (Province et al. 2003) and diabetes (An et al. 2005).

The Fisher p-value method and its subsequent extensions do not necessarily account for among-study heterogeneity with one of the most obvious differences being sample size and hence admittedly are subject to potential biases from not accounting for such differences among studies. Although decision criteria could be developed such that only studies that are most homogeneous (with respect to sample size or pedigree selection) be included in a meta-analysis, this may exclude too many studies with viable linkage information and hence limit the sample size for the meta-analysis (see discussion below). William Rice (1990) suggested a reparameterization of Fisher's method such that the evidence for linkage from each study can be weighted by the corresponding study's sample size. In doing so, he suggested that the p-value, p_{st}, be transformed into a standard normal variate

$$z_{st} = \Phi^{-1}\left(p_{st}\right)$$

where $\Phi^{-1}(\bullet)$ is the standard normal inverse function. A weighted average of the z-values at marker *t* (or region *t* if applying this reparameterization to the Badner and Gershon extension) can be calculated

$$z_{\bullet t} = \frac{\sum_{s=1}^{k} N_s z_{st}}{\sum_{s=1}^{k} N_s}$$

where N_s is the sample size (number of pedigrees, number of sib-pairs, etc.) for study *s*.

Under the omnibus null hypothesis of no linkage, $\frac{z_{\bullet t}}{\sqrt{Var(z_{\bullet t})}}$ follows a standard normal distribution where

$$Var\left(z_{\bullet t}\right) = \frac{\sum_{s=1}^{k} N_s^2}{\left(\sum_{s=1}^{k} N_s\right)^2}.$$

Other novel meta-analytic methods for genome scans that use p-values or other outcomes of significance tests involving linkage which are not extensions of Fisher's method have been proposed specifically for genome-scan meta-analysis. One such widely used method, the Genome Search Meta-analysis Method (GSMA), developed by Wise et al. (1999) is based on a non-parametric ranking of p-values or LOD scores within specified genetic regions (or bins). Suppose that we have split the chromosomes into *m* bins. For each genome-scan study *s*, ($s = 1, \ldots k$ =number of total studies), the most significant linkage result (whether it be p-value, LOD score or another linkage test statistic) within each bin *t* ($t = 1, \ldots, m$) is identified. The bins are then ranked within the studies where the most significant bin receives the highest rank. The ranks for each bin are then summed across the studies, such that

$$V_t = \sum_{s=1}^{m} R(X_{st}) \qquad (5)$$

where X_{st} is the most significant linkage result for bin *t* of study *s*, and $R(\bullet)$ is the ranking function. As with Fisher's method, there are no assumptions that each study used the same sampling scheme or linkage test, or that each genome scan use the same set of markers. Additionally, however, they showed through simulation that the GSMA is useful when studies use different ascertainment schemes, marker maps, or statistical methods to detect linkage. Wise et al. (1999) derived the null distribution of V_t given in (5) and Koziol et al. (2004)

refined the derivation of the null distribution using probability generating functions and provided approximations to the GSMA null distribution.

Wise (2001) further proposed an extension of the GSMA method such that candidate region studies can be included in the meta-analysis with genome-wide studies. In this extension, a simulation procedure is developed to assign ranks to the candidate regions where the ranks reflect the expected ranks under the null hypothesis of no linkage for a genome-wide study. By assigning the ranks to the candidate regions in this manner, Wise concludes that the false positive rate is not inflated due to the higher marker density of candidate region studies.

Babron et al. (2003) updated the GMSA method by first replacing the rank V_t in equation (5) with the average rank of bin *t* and the ranks of its two flanking bins, defined as V_{-t} and V_{+t} in order to adjust for arbitrary bin construction. Second, they defined a weighting scheme for the ranks such that the rank of study *s* in bin *t*, namely X_{st} in (5), is weighted by the number of pedigrees in study *s* in order to account for differing information content across studies. Although Babron et al. (2003) suggested weights to account for differing information content, a formal test for heterogeneity among the studies for the GSMA method was not introduced until 2005. Zintzaras and Ioannidis (2005b) propose three weighted metrics to measure among-study heterogeneity for the GSMA method: 1. sum of the weighted squared mean rank deviations, 2. sum of the weighted absolute mean rank deviations and 3. weighted sum of the distinct absolute rank differences. Furthermore, Zintzaras and Ioannidis (2005a) have developed a software program HEGESMA to perform the GSMA meta-analysis (unweighted or weighted as specified by the user) as well as provide the user with heterogeneity results.

In their original paper, Wise et al. (2001) suggested a bin width of 30 cM, but recently, Marazita et al. (2004) proposed repeating the GSMA with variable bin-length starting points in order to determine minimum regions of maximum significance (MRMS). The resulting bin-shifting method identifies narrower regions of positive findings compared to the original GSMA which then leads to narrower regions to be followed-up with fine-scale mapping.

Since its original publication, the GSMA has been the most widely used meta-analytic method for genome scans, specifically due to its ease of use and invariance to whether the studies are from one-sided or two-sided tests or if only the most significant results have been reported. A number of investigators have applied the GSMA method to a variety of complex diseases, multiple sclerosis and other autoimmune diseases (Wise, Lanchbury, and Lewis 1999; Fisher, Lanchbury, and Lewis 2003; Sagoo et al. 2004), inflammatory bowel disease (Williams et al. 2002; van Heel et al. 2004), asthma (Wise 2001), celiac disease (Babron et al. 2003), schizophrenia and bipolar disorders (Levinson et al. 2003; Lewis et al. 2003; Segurado et al. 2003), obesity (Johnson et al. 2005), diabetes (Demenais et al. 2003), coronary heart disease (Chiodini and Lewis 2003) and

hypertension (Liu, Zhao, and Chase 2004; Koivukoski et al. 2004) to name a few.

3.6.2.2 Meta-analytic methods based on effect sizes

A meta-analysis based on combining the results from significance tests can be limited or misleading, especially in cases where the concordance or discordance of significant linkage between two studies may not reflect the existence of true linkage, but rather may be based on the amount of heterogeneity between multiple studies. Although adjustments for heterogeneity have been proposed for these methods, combining effect sizes may be a better approach as many of these methods are based on random effects models that naturally allow the user to adjust for among-study heterogeneity.

Loesgen et al. (2001) developed a meta-analytic test that computes a weighted average estimate of score statistics

$$Z_{MA_t} = \frac{\sum_{s=1}^{k} w_{st} Z_{st}}{\sqrt{\sum_{s=1}^{k} w_{st}^2}} \qquad (6)$$

where Z_{st} is the NPL score statistic and w_{st} is the assigned weight from study *s* at position *t*. They proposed several weighting schemes such as sample size, information content and an exponential function based on marker distance. Dempfle and Loesgen (2004) compared the power of the method proposed by Loesgen et al. (2001) to Fisher's method, the GSMA and other p-value based meta-analytic methods. They showed that meta-analysis performed using weighted effect sizes had more power to detect linkage than the p-value methods with nominal increases in false positive rates. Further, they found that their method based on effect sizes was more robust and consistent across simulation aspects compared to the p-value based methods.

Etzel and Guerra (2002) developed a meta-analysis technique to combine Haseman-Elston test statistics across studies that have distinct marker maps. For this method they suppose that $\hat{\beta}_{st}$, the Haseman-Elston (1972) slope estimate, and S_{st}^2, the corresponding variance estimate of $\hat{\beta}_{st}$ for the marker *t* of study *s* are available for each of *k* studies. They further define $\{L_q, q=1,\ldots,v\}$ as the set of analysis points such that L_1 and L_t are at each endpoint of a chromosome segment, respectively, and the distance between any two adjacent points L_i and L_{i+1} is constant and equal to L/t where L is the length of the chromosome segment. For each analysis point, they calculate the statistics $\hat{\beta}_{stq}$ and S_{stq}^2 utilizing markers within D cM of L_q, where

$$\hat{\beta}_{stq} = \frac{\hat{\beta}_{st}}{\left[1-2\theta_{stq}\right]^2} \qquad \text{and} \qquad S^2_{stq} = \frac{S^2_{st}}{\left[1-2\theta_{stq}\right]^4}.$$

The value θ_{stq} is the recombination fraction between marker *t* of study *s* and analysis point L_q as estimated using a general mapping function, for example, Kosambi. Next, they calculate the weighted least-squares estimate $\tilde{\beta}_q$ at L_q,

$$\tilde{\beta}_q = \frac{\sum_{s=1}^{k}\sum_{t=1}^{n_{sq}} w_{st}\hat{\beta}_{stq}}{\sum_{s=1}^{k}\sum_{t=1}^{n_{sq}} w_{st}} \quad \text{and} \quad w_{st} = \frac{1}{\sigma^2_B + S^2_{stq}}$$

where *k* is the number of studies and n_{iq} is the number of markers within D cM of L_q for study *s* and σ^2_B is between-study variance. The estimator $\hat{\sigma}^2_{B_q}$ for σ^2_B at L_q is

$$\hat{\sigma}^2_{B_q} = \frac{1}{\sum_{s=1}^{k} n_{sq} - 1}\sum_{s=1}^{k}\sum_{t=1}^{n_{sq}}\left[\hat{\beta}_{stq} - \bar{\beta}_{..q}\right]^2 - \frac{1}{\sum_{s=1}^{k} n_{sq}}\sum_{s=1}^{k}\sum_{t=1}^{n_{sq}} S^2_{stq},$$

where $\bar{\beta}_{..q}$ is the average of the $\hat{\beta}_{stq}$ that are within D cM of L_q. The variance of $\tilde{\beta}_q$ is $\frac{1}{\sum_{s=1}^{k}\sum_{t=1}^{n_{sq}} w_{st}}$. The analysis point $L_{q'}$ such that $t_{q'} = \frac{\tilde{\beta}_{q'}}{\sqrt{\text{Var}\left[\tilde{\beta}_{q'}\right]}}$ is minimum and significant at a specified level is the point estimate of location of the QTL. Likewise, the estimate of genetic variance is given by $\hat{\sigma}^2_g = \frac{\tilde{\beta}_{q'}}{-2}$. Etzel and Guerra (2002) further describe a bootstrapping procedure to construct confidence intervals for location of the putative QTL and genetic variance. Through simulation, they show that the empirical power using this procedure remained high even when power at the individual study level was low. This procedure was used to assess linkage of immunoglobulin E (IgE), an asthma related quantitative trait, using the nine data sets provided by the Genetic Analysis Workshop 12 and found suggestive linkage for two regions on chromosome 4 and one region on chromosome 11.

The method proposed by Loesgen et al. (2001) assumes that all studies use the same marker map but different linkage tests and the method proposed by Etzel and Guerra allows for differing marker maps among the studies involved; however, the Etzel and Guerra method is limited by the fact all studies must use the same linkage test. In 2005, Etzel et al. (GAW14) proposed a meta-analytic procedure that combines the methods of Loesgen et al. (2001) and Etzel and Guerra (2002) and results in a more flexible procedure to combine effect sizes across linkage studies that perform different linkage tests on different marker maps. The resulting Meta-Analysis for Genome Studies (MAGS) method is based on a weighted average of effect sizes that are obtained through the reported linkage summary statistics. Suppose that we wish to complete a meta-analysis on *k* studies. Each study *k* has m_k number of markers. It is not assumed that the studies have the same number of markers, $m_i \neq m_k, i \neq j$, nor it is assumed that the studies have the same marker maps. For a specified chromosome, let M_{st} denote the t^{th} marker from study *s*, for $s = 1, \ldots k$ and $t = 1, \ldots, m_k$. Define $\{L_q, q = 1, \ldots, l\}$ as the set of analysis points such that the L_q are equally spaced across the chromosome. For each set of M_{st} on a chromosome, let Z_{st} be the associated score statistic. As noted by Dempfle and Loesgen (2004), Z_{st} can be the NPL score statistic as most standard multipoint linkage analysis software packages includes the calculation of such statistics. However, Z_{st} can also be derived from other linkage related statistics, such as an HLOD score or even a p-value with the correct transformation (see Appendix A). For each analysis point L_q, calculate the weighted normal variate:

$$Z_{MA_q} = \frac{\sum_{s=1}^{k}\sum_{t=1}^{m_k} I_{q\{M_{st}\}} w_{stq} Z_{st}}{\sqrt{\sum_{s=1}^{k}\sum_{t=1}^{m_k} I_{q\{M_{st}\}} w_{stq}^2}},$$

where w_{stq} is the weight given to marker M_{st}. The indicator function $I_{q\{M_{st}\}}$ is defined as 1 if marker M_{st} is within a set distance D *cM* from analysis point L_q and 0 otherwise. The weight w_{stq} for marker M_{st} can be a function of study sample size, information content at that marker, and/or distance (recombination fraction, θ_{stq}) between marker M_{st} and analysis point L_q, say

$$w_{stq} = f\left(n_s\right) g\left(I_{q\{M_{st}\}}\right) h\left(\theta_{stq}\right).$$

The p-value for each analysis location then be compared to a set level to determine areas with combined evidence for linkage. NOTE: If all studies use the same marker map, then the combined set of markers can replace the analysis points L_q and the expression for Z_{MA_t} simplifies to the statistic proposed by Dempfle and Loesgen (2004). Etzel et al. (2005) applied this procedure to the simulated data from the Genetic Analysis Workshop 14 and correctly identified the disease loci on chromosomes 1, 3 and 5; however, found low evidence of linkage to the disease modifier genes on chromosomes 2 and 10.

3.6.3 Choosing a method to best suit your analytic needs

Data can be obtained from published sources, open-source websites or through consortia group agreements. At times, the researcher may be limited in choosing a preferred meta-analytic method due to the type of data available for a meta-analysis: complete data on all studies through a consortium; data obtained by contacting corresponding authors from published articles; data from published reports; or some combination of these three. However, the researcher who is able to obtain the data of his/her choosing should then select the meta-analysis method based on the most robust methodology for identifying linkage within each individual study. Below, we propose some scenarios that reflect reasonable situations in which a meta-analysis would be performed and provide advice regarding the type of meta-analytic method to use.

3.6.3.1 Scenario 1: Raw data available on all studies.

This scenario could arise when the researcher is a member of a data consortium whereby members of the consortium freely share all data from their individual studies. For a meta-analysis, this is the most ideal situation since the researcher is relatively free to reanalyze the data (separately from each study) using a preferred linkage method and then combine the resulting linkage outcome using any one of the above mentioned meta-analysis methods. In order fully account for between-study heterogeneity, the researcher should choose one of the meta-analysis methods that allows for such an adjustment (Dempfle and Loesgen (2004), Etzel et al. (2005) or Zintzaras and Ioannidis (2005b)). Even if the marker maps are different among the studies in the consortium, the researcher could develop a simple scheme to align the marker maps in order to perform the meta-analysis. The researcher even has the option to not perform a meta-analysis, but to complete a mega-analysis instead, such that the raw data from each of the studies are combined into one common database. Some notable examples of this approach were applied to multiple sclerosis (The Transatlantic Multiple Sclerosis Genetics Cooperative 2001; GAMES and The Transatlantic Multiple Sclerosis Genetics Cooperative 2003), celiac disease

(Babron et al. 2003), asthma (Iyengar, Jacobs, and Palmer 2001), diabetes (Demenais et al. 2003) and obesity related phenotypes (Heo et al. 2002). A master marker map can be established by using a marker location database. If there are any missing values, one could consider imputation as in Heo et al. (2002). The combined data is then analyzed using a standard linkage method. It has been shown (Guerra et al. 1999), that a mega-analysis may have more power to detect linkage than compared to a meta-analysis; however, one should consider the different types of heterogeneity that may be inherent in each of the different studies. This heterogeneity may adversely confound or overshadow the results from a mega-analysis and may arise from differing study designs (linkage results on extended pedigrees may not combine well with linkage results from sib-pairs, discordant pairs or parent-offspring triads), varying ethnic/racial groups across study populations (different genes acting in different populations) and varying sample sizes.

3.6.3.2 Scenario 2: All studies use similar linkage tests and similar marker maps

This scenario could also arise when the researcher is a member of a data consortium whereby the members individually analyze their own data using a common linkage method and freely share linkage results instead of raw data. Likewise, this scenario could occur when the researcher personally contacts corresponding authors from published studies and requested complete linkage analysis results from their data. If these data are obtained from corresponding authors, or extracted from the literature, the researcher should collect the most detailed information possible: i.e., score statistics instead of p-values, marker information content, recruitment criteria and sample schemes. For this scenario, we once again recommend that the researcher choose a meta-analysis method that is flexible enough to account for between-study heterogeneity: (Dempfle and Loesgen (2004) or Etzel et al. (2005) if score statistics are available or Zintzaras and Ioannidis (2005b) if only p-values are provided.

3.6.3.3 Scenario 3: All studies used similar linkage tests but with different marker maps.

This scenario is similar to scenario 2 except for the commonality of the marker maps between the studies and likewise, this scenario could occur for the same reasons as scenario 2. The added complexity of differing marker maps will not hinder a meta-analysis over the individual studies, as long as the researcher uses a method that is flexible in this respect. Once again, we advise that the researcher request as detailed linkage information as possible and apply a meta-analysis based on the effect size method proposed by Etzel et al. (2005) if score statistics are available or the GSMA modification proposed by Zintzaras and Ioannidis (2005b) if only p-values are provided.

3.6.3.4 Scenario 4: P-values or LOD scores from different linkage tests and different marker maps from published data are available from all studies

In this scenario, it is assumed that the researcher is basing the meta-analysis on summary linkage results (p-values or LOD scores) that are available from published articles with no follow-up information obtained from the corresponding authors. Although the availability of data in this scenario may seem limited and can vary greatly depending on the disease of interest, manuscript type and journal of publication, many meta-analyses are based on such data (Allison and Heo (1998) for instance). For this case, the GSMA method developed by Wise et al. (1999) would be the best method to employ as long as the available data allow. If possible, the researcher could also employ any of the modifications to the GSMA method if s/he has ample auxiliary information to do so. In cases where application of the GSMA method is not possible (such as the scenario posed by Allison and Heo), then application of Fisher's method is still viable.

3.6.4 Discussion

Herein, we review current meta-analytic techniques for the combination of linkage data across studies in order to arrive at a consensus for linkage to a complex disease. We also propose several scenarios to help guide the researcher in their choice of which meta-analytic technique to employ. However, we caution that meta-analysis is more than just a method one can use to combine data together. Although the choice of method is important, the researcher must also keep in mind that the application of a method is just a small part of a complete meta-analysis. Just as study design and participant recruitment is important at the beginning of any linkage study, a researcher who is about to embark on a meta-analysis should also develop a study design and participant study plan which includes a literature review plan, as well as study inclusion/exclusion criteria. The researcher must also gather as much information on original studies as possible, which may include contacting corresponding authors. If raw data are provided, the researcher needs to decide how to treat missing data. The researcher may have ample data to complete a meta-analysis; however, roadblocks to complete the meta-analysis may exist. Most of these roadblocks include differences among the studies with respect to: marker maps or denseness of maps, family structure, environmental factors, population substructure, distinct genetic etiology/different pathways within the disease of interest, marker informativity, sample sizes, ascertainment schemes, phenotype definitions and/or linkage tests. Additional challenges include publication bias and time-lag bias. Although we presented meta-analytic methods that can handle some of these problems, no one single meta-analysis method exists that can handle all such problems. Therefore, a researcher must be willing to accept the limitations of his/her own meta-analysis.

Two topics that we have not discussed in detail within this chapter involve determining an appropriate significance level for a meta-analysis performed on genome scans and the effect of publication bias (only positive linkage results published). The topic of genome-wide significance levels for individual studies remains in controversy and to fully detail the debate with respect to a meta-analysis would be a lengthy chapter in itself. Instead, we leave it to the researcher to consider an appropriate significance level, but advise the researcher to look to Morton (1955), Lander and Kruglyak (1995), Feingold et al. (1993), Sawcer et al. (1997), Rao (1998), Rao and Gu (2001), and Levinson et al. (2003) to gain more insights into the determination of an appropriate significance level.

Publication bias in a meta-analysis may become a factor when the results of the study impact the probability that it will be published in the literature. In this event, if the published literature was biased in favor of statistically significant results, you would find a relative lack of studies reporting negative evidence for linkage and you could incorrectly conclude a region to be more significantly involved in the disease in question than it really is. Iyengar and Greenhouse (1988) present two procedures to handle this potential bias by estimating what they term the 'fail safe sample size.' They first describe the procedure presented by Rosenthal (1979) which determines the minimum number of unpublished studies with null results required to reverse the conclusion of the meta-analysis over the published studies and note that Rosenthal (1984) provides some ad hoc guidelines for interpretation. Iyengar and Greenhouse (1988) extend the approach described by Rosenthal (1979) and present a second procedure based on selection models that uses a maximum likelihood approach to model the reporting process by weighting the results in the meta-analysis. They note that by using the MLE approach, you can examine how changing your assumptions about the selection model change the parameter estimates and inference of the meta-analysis.

3.6.5 References

Allison, D.B. and Heo, M. (1998). Meta-analysis of linkage data under worst-case conditions: a demonstration using the human OB region. Genetics, 148:859-865.

An, P. et al. (2005). Genome-wide linkage scans for fasting glucose, insulin, and insulin resistance in the National Heart, Lung, and Blood Institute Family Blood Pressure Program: evidence of linkages to chromosome 7q36 and 19q13 from meta-analysis. Diabetes, 54:909-914.

Babron, M.C. et al. (2003). Meta and pooled analysis of European coeliac disease data. European Journal of Human Genetics, 11:828-834.

Badner, J.A. and Gershon, E.S. (2002a). Meta-analysis of whole-genome linkage scans of bipolar disorder and schizophrenia. Molecular Psychiatry, 7:405-411.

Badner, J.A. and Gershon, E.S. (2002b). Regional meta-analysis of published data supports linkage of autism with markers on chromosome 7. Molecular Psychiatry, 7:56-66.

Chiodini, B.D. and Lewis, C.M. (2003). Meta-analysis of 4 coronary heart disease genome-wide linkage studies confirms a susceptibility locus on chromosome 3q. Arteriosclerosis, Thrombosis, and Vascular Biology, 23:1863-1868.

Demenais, F. et al. (2003). A meta-analysis of four European genome screens (GIFT Consortium) shows evidence for a novel region on chromosome 17p11.2-q22 linked to type 2 diabetes. Human Molecular Genetics, 12:1865-1873.

Dempfle, A. and Loesgen, S. (2004). Meta-analysis of linkage studies for complex diseases: an overview of methods and a simulation study. Annals of Human Genetics, 68:69-83.

Etzel, C.J. and Guerra, R. (2002). Meta-analysis of genetic-linkage analysis of quantitative-trait loci. American Journal of Human Genetics, 71:56-65.

Etzel, C.J., Liu, M., and Costello, T.J. (2005). An Updated Meta-Analysis Approach for Genetic Linkage. BiomedCentral Genetics (in press).

Feingold, E., Brown, P.O., and Siegmund, D. (1993). Gaussian models for genetic linkage analysis using complete high-resolution maps of identity by descent. American Journal of Human Genetics, 53:234-251.

Fisher, R.A. (1925). Statistical methods for research workers. Oliver & Loyd, London.

Fisher, S.A., Lanchbury, J.S., and Lewis, C.M. (2003). Meta-analysis of four rheumatoid arthritis genome-wide linkage studies: confirmation of a susceptibility locus on chromosome 16. Arthritis and Rheumatism, 48:1200-1206.

Folks, L.J. (1984). Combination of Independent Tests. In Krishnaiah PR, Sen PK (eds) Handbook of Statistics. Elsevier Science Publishers, 6:56-70.

GAMES and The Transatlantic Multiple Sclerosis Genetics Cooperative (2003). A meta-analysis of whole genome linkage screens in multiple sclerosis. Journal of Neuroimmunology, 143:39-46.

Gu, C. et al. (1998). Meta-analysis methodology for combining non-parametric sibpair linkage results: genetic homogeneity and identical markers. Genetic Epidemiology, 15:609-626.

Guerra, R. et al. (1999). Meta-analysis by combining p-values: simulated linkage studies. Genetic Epidemiology, 17 Suppl 1:S605-S609.

Haseman, J.K. and Elston, R.C. (1972). The investigation of linkage between a quantitative trait and a marker locus. Behavior Genetics, 2:3-19.

Heo, M. et al. (2002). A meta-analytic investigation of linkage and association of common leptin receptor (LEPR) polymorphisms with body mass index and waist circumference. International Journal of Obesity and Related Metabolic Disorders, 26:640-646.

Iyengar, S. and Greenhouse, J. (1988). Selection Models and the File Drawer Problem. Statistical Science, 3:109-117.

Iyengar, S.K., Jacobs, K.B., and Palmer, L.J. (2001). Improved evidence for linkage on 6p and 5p with retrospective pooling of data from three asthma genome screens. Genetic Epidemiology, 21 Suppl 1:S130-S135.

Johnson, L. et al. (2005). Meta-analysis of five genome-wide linkage studies for body mass index reveals significant evidence for linkage to chromosome 8p. International Journal of Obesity and Related Metabolic Disorders, 29:413-419.

Koivukoski, L. et al. (2004). Meta-analysis of genome-wide scans for hypertension and blood pressure in Caucasians shows evidence of susceptibility regions on chromosomes 2 and 3. Human Molecular Genetics, 13:2325-2332.

Koziol, J.A. and Feng, A.C. (2004). A note on the genome scan meta-analysis statistic. Annals of Human Genetics, 68:376-380.

Lander, E. and Kruglyak, L. (1995). Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. Nature Genetics, 11:241-247.

Levinson, D.F. et al. (2003). Genome scan meta-analysis of schizophrenia and bipolar disorder, part I: Methods and power analysis. American Journal of Human Genetics, 73:17-33.

Lewis, C.M. et al. (2003). Genome scan meta-analysis of schizophrenia and bipolar disorder, part II: Schizophrenia. American Journal of Human Genetics, 73:34-48.

Li, Z. and Rao, D.C. (1996). Random effects model for meta-analysis of multiple quantitative sibpair linkage studies. Genetic Epidemiology, 13:377-383.

Liu, W., Zhao, W., and Chase, G.A. (2004). Genome scan meta-analysis for hypertension. American Journal of Hypertension, 17:1100-1106.

Loesgen, S. et al. (2001). Weighting schemes in pooled linkage analysis. Genetic Epidemiology, 21 Suppl 1:S142-S147.

Marazita, M.L. et al. (2004). Meta-analysis of 13 genome scans reveals multiple cleft lip/palate genes with novel loci on 9q21 and 2q32-35. American Journal of Human Genetics, 75:161-173.

Morton, N.E. (1955). Sequential tests for the detection of linkage. American Journal of Human Genetics, 7:277-318.

Ott, J. (1991). Analysis of Human Genetic Linkage. Johns Hopkins University Press, Baltimore.

Pearson, K. (1933). On a method of determining whether a sample of size n supposed to have been drawn from a parent population having a known probability integral has probably been drawn at random. Biometrika, 25:379-410.

Province, M.A. (2001). The significance of not finding a gene. American Journal of Human Genetics, 69:660-663.

Province, M.A. et al. (2003). A meta-analysis of genome-wide linkage scans for hypertension: the National Heart, Lung and Blood Institute Family Blood Pressure Program. American Journal of Hypertension, 16:144-147.

Rao, D.C. (1998). CAT scans, PET scans, and genomic scans. Genetic Epidemiology, 15:1-18.

Rao, D.C. and Gu, C. (2001). False positives and false negatives in genome scans. Advances in Genetics, 42:487-498.

Rice, W.R. (1990). A consensus combined p-value test and the family-wide significance of component tests. Biometrics, 46:303-308.

Rosenthal, R. (1979). The "file drawer problem" and tolerance for null results. Psychological Bulletin, 86:638-641.

Rosenthal, R. (1984). Meta-Analytic Procedures for Social Research. Sage Press, Beverly Hills, CA.

Sagoo, G.S. et al. (2004). Meta-analysis of genome-wide studies of psoriasis susceptibility reveals linkage to chromosomes 6p21 and 4q28-q31 in Caucasian and Chinese Hans population. Journal of Investigative Dermatology, 122:1401-1405.

Sawcer, S. et al. (1997). Empirical genomewide significance levels established by whole genome simulations. Genetic Epidemiology, 14:223-229.

Segurado, R. et al. (2003). Genome scan meta-analysis of schizophrenia and bipolar disorder, part III: Bipolar disorder. American Journal of Human Genetics, 73:49-62.

The Transatlantic Multiple Sclerosis Genetics Cooperative (2001). A meta-analysis of genomic screens in multiple sclerosis. Multiple Sclerosis, 7:3-11.

Tippett, L.H.C. (1931). The Methods of Statistics. 1st ed. Williams & Norgate, London.

van Heel, D.A. et al. (2004). Inflammatory bowel disease susceptibility loci defined by genome scan meta-analysis of 1952 affected relative pairs. Human Molecular Genetics, 13:763-770.

Williams, C.N. et al. (2002). Using a genome-wide scan and meta-analysis to identify a novel IBD locus and confirm previously identified IBD loci. Inflammatory Bowel Diseases, 8:375-381.

Wise, L.H. (2001). Inclusion of candidate region studies in meta-analysis using the genome screen meta-analysis method: application to asthma data. Genetic Epidemiology, 21 Suppl 1:S160-S165.

Wise, L.H., Lanchbury, J.S., and Lewis, C.M. (1999). Meta-analysis of genome searches. Annals of Human Genetics, 63:263-272.

Wu, X. et al. (2002). A combined analysis of genomewide linkage scans for body mass index from the National Heart, Lung, and Blood Institute Family Blood Pressure Program. American Journal of Human Genetics, 70:1247-1256.

Zintzaras, E. and Ioannidis, J.P. (2005a). HEGESMA: genome search meta-analysis and heterogeneity testing. Bioinformatics, (in press).

Zintzaras, E. and Ioannidis, J.P. (2005b). Heterogeneity testing in meta-analysis of genome searches. Genetic Epidemiology, 28:123-137.

Appendix A
Example transformation of a linkage summary to a score statistic

1. Transform an HLOD to Chi-square variate: $X_{st} = 4.6 * HLOD_{st}$
2. Obtain P-value for each chi-square variate [Faraway, J.J. (1993). Distribution of the admixture test for the detection of linkage under heterogeneity. Genetic Epidemiology 10:75-83.]

   $$p_{st} = .5 * \left[1 - \Pr^2 \left(\chi_1^2 < X_{st} \right) \right]$$

2. Transform the resulting P-value to a normal variate by the inverse of the normal distribution: $Z_{st} = \Phi^{-1} \left(p_{st} \right)$