DNA polymorphisms and renal disease: a critical appraisal of studies presented at the annual ERA/EDTA and ASN conferences in 2004 and 2005

Adrian Mondry1,2,*, Marie Loh1,*, Kevin Ben Laurence1 and Nelson Low1

1Bioinformatics Institute, Agency for Science, Technology and Research, Singapore and 2Department of Surgery, Yong Loo Lin School of Medicine, National University of Singapore, Singapore

Abstract
Countless studies try to associate single DNA polymorphisms with disease, while there is growing evidence that many of these studies are of flawed design. Based on the Journal of the American Society of Nephrology (JASN) requirements for gene-disease association study quality, the abstracts presented at the two major international nephrology conferences in 2004 and 2005 organized by the European Renal Association/European Dialysis and Transplantation Association (ERA/EDTA) and American Society of Nephrology (ASN) are analysed to show how this problem affects nephrology. Over time, average sample numbers have increased, as have the numbers of abstracts compliant with the JASN requirements. This indicates a potential beneficial effect of the published stricter guidelines on study quality. Alternative options include pre-registration of studies in dedicated databases, secondary assessment of association studies through meta-analysis and participation in network approaches, such as the Human Genome Epidemiology Network (HuGE Net) and the Renal Genome Network.

Keywords: gene association studies; renal disease; renal genome network; sample size; the human genome epidemiology network

DNA polymorphisms and their association with disease

Knowledge and correct interpretation of the human genome sequence is the key to more profound insight into pathophysiological processes, and maybe to the development of new and better therapies. Over the past decades, many attempts have been made to associate DNA polymorphisms with disease. Current understanding of the knowledge that can be gained from genomic studies indicates that it is futile to try and establish a causal relationship between isolated individual DNA polymorphisms and complex diseases. In consequence, the American Society of Nephrology (ASN) bars studies from the review process that do not comply with a strict set of requirements providing additional evidence supporting the claim for gene-disease association (Table 1) [1]. Thus of the 30 listed in the Journal Citation Index, the Journal of the American Society of Nephrology (JASN) is the only nephrology journal that makes a formal requirement of the suggestions advocated by Nature Genetics as early as 1999 [2]. Yet, despite the increased emphasis on quality considerations, studies on mutations and single nucleotide polymorphisms remain popular [3–5]. One public database, KGDB (Human Kidney Gene Database) [6], provides a statistic that categorizes 317 of 401 (79%) references as dealing with either mutations or polymorphisms in kidney disease.

Non-compliance with demands for better study quality causes problems on several levels. Firstly, research costs time and money. Production of data that are essentially useless is a waste of resources. This affects not only the researchers themselves, but also the grant reviewers, and later on the review committee members at scientific conferences and the scientific periodicals in which the researchers hope to publish their work. Secondly, once published, the data will remain in circulation even though they may convey little information and even if the knowledge generated has to be retracted later on [7]. Thirdly, incorrect prognosis may be made in clinical applications, leading to inappropriate therapeutic recommendations or genetic counselling. In such circumstances, inappropriately designed studies may be considered as breaching ethical standards.

*The authors wish it to be known that, in their opinion, the first two authors contributed equally to this work.
**Association studies presented at the leading conferences in nephrology in 2004 and 2005**

The problem is seen across all medical disciplines. The research area of nephrology is of particular interest to monitor, because it adopted the conceptual and technical approaches of modern molecular biology somewhat belatedly [8]. This may have been due to the stunning, and still ongoing, successes of renal anatomists and physiologists in the ‘pre-molecular era’. In view of this great legacy, however, strictest adherence to the highest standards of research ethics should be paramount.

In order to see how the field of nephrology addresses this issue, we analysed the abstracts of all relevant studies presented at the two major annual conferences in the field.

Electronic abstract books of the 41st and 42nd annual conferences of the European Renal Association/European Dialysis and Transplantation Association (ERA/EDTA) and the 2004 and 2005 Renal Week of the ASN were searched for the terms ‘gene*’ and ‘polymorphism*’. The retrieved abstracts were then individually assessed in triplicate to select only original gene-disease association studies on human subjects. Only those studies that allowed identification of the genes and polymorphisms were included in the numerical description. Each selected abstract was then individually assessed to check whether it provided information that would allow the study to be considered for review in *JASN*. For this purpose, one of the secondary requirements listed in the *JASN* instructions to authors [1] had to be met. Compliance with the primary requirements was assessed only in those abstracts that met the secondary requirements. Primary and secondary requirements are summarized in Table 1. The numbers of retrieved abstracts of polymorphisms studied, and their compliance with the *JASN* author guidelines are summarized in Table 2. Genes coding for proteins of the renin angiotensin system (RAS) or involved in inflammatory processes were studied most frequently. Notable is the low sample size of average study cohorts in 2004, which was 245 (±289) for cases and controls combined. Excluding six studies as outliers or extreme values, the average study size dropped to 185 (±136) subjects, cases and controls combined. Thirty-one studies examined <100 subjects altogether. Only 17/107 (16%) studies met the ASN secondary criteria for acceptance to the review process. Ten of these 17 examined polymorphisms that were on the list of most frequently studied polymorphisms.

In the year 2005, average study cohorts increased to 354 (±147) for cases and controls combined. Excluding six studies as outliers or extreme values, however, the average study size dropped to 276 (±116) subjects, cases and controls combined. Thirteen studies even had total sample sizes <100. There were 38/77 (49%) studies which met the ASN secondary criteria for acceptance to the review process. Eight of these looked at genes on the most frequent list (Figure 1) in 2005. The increase in compliance with the ASN criteria was more obvious in the abstracts from the ASN than from the ERA/EDTA conference (Table 2).

### Gene-association studies—the illusion of an explanation?

Since the late 18th century, the idea of heredity has been virtually ubiquitous in discussions on the origins of disease [9]. Not surprisingly, the concept has both changed and prospered with the growing knowledge about the genomic basis of inheritance [10]. Today there is a general understanding that knowledge and

---

**Table 1.** Quality criteria required for a study to be considered for review by the *JASN*

<table>
<thead>
<tr>
<th>Primary requirements</th>
<th>Secondary requirements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Results must make sense biologically</td>
<td>Data showing an effect of polymorphism on protein function or gene expression</td>
</tr>
<tr>
<td>Sample size must be adequate</td>
<td>Confirmation of the association using a family-based method</td>
</tr>
<tr>
<td>Adjustment of <em>P</em>-value for multiple comparisons</td>
<td>Replication of the association in an independent sample</td>
</tr>
<tr>
<td></td>
<td>Measurement of and correction for population stratification in the samples</td>
</tr>
<tr>
<td></td>
<td>Special consideration to the analysis identifying the ‘risk haplotype’ associated with renal disease</td>
</tr>
</tbody>
</table>

All the primary requirements must be met, as well as at least one secondary requirement.

**Table 2.** Descriptive numerical breakdown of the abstracts, studies, polymorphisms and genes by conference in 2004 and 2005

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstracts retrieved</td>
<td>10937</td>
<td>5427</td>
<td>5510</td>
<td>4168</td>
<td>4408</td>
<td>1259</td>
<td>1102</td>
</tr>
<tr>
<td>Studies</td>
<td>184</td>
<td>107</td>
<td>77</td>
<td>40</td>
<td>46</td>
<td>67</td>
<td>31</td>
</tr>
<tr>
<td>Polymorphisms</td>
<td>205</td>
<td>119</td>
<td>109</td>
<td>52</td>
<td>71</td>
<td>82</td>
<td>46</td>
</tr>
<tr>
<td>Genes</td>
<td>92</td>
<td>59</td>
<td>57</td>
<td>34</td>
<td>39</td>
<td>41</td>
<td>28</td>
</tr>
<tr>
<td>Studies that met <em>JASN</em> criteria</td>
<td>55 (30%)</td>
<td>17 (16%)</td>
<td>38 (49%)</td>
<td>7 (17%)</td>
<td>27 (59%)</td>
<td>10 (15%)</td>
<td>11 (36%)</td>
</tr>
<tr>
<td>Published</td>
<td>15 (8%)</td>
<td>11 (10%)</td>
<td>4 (5%)</td>
<td>4 (10%)</td>
<td>2 (4%)</td>
<td>7 (10%)</td>
<td>2 (7%)</td>
</tr>
<tr>
<td>Published that met <em>JASN</em> criteria</td>
<td>5 (3%)</td>
<td>3 (3%)</td>
<td>2 (3%)</td>
<td>1 (3%)</td>
<td>2 (4%)</td>
<td>2 (3%)</td>
<td>0 (0%)</td>
</tr>
</tbody>
</table>

Associations reported in duplicate were counted only once.
correct interpretation of the human genome is the key to a more profound insight into pathophysiological processes and may lead to the development of new and better therapies.

In consequence, no current biomedical conference seems complete that fails to host presentations which try to demonstrate an association between a particular genotype and disease, and it seems that studies on mutations and single nucleotide polymorphisms continue to be popular [3–5]. Looking at the field of nephrology, however, the numbers presented here indicate that the peak of popularity for this type of study may have passed. The analysis of conference abstracts described above shows that gene-disease association studies made up $\approx 2$ and $1.4\%$ of all abstracts in 2004 and 2005, respectively. Less than a fifth of these studies met at least one of the secondary requirements in 2004, but this number increased greatly to almost half in 2005. The total number of conference abstracts increased from 2004 to 2005 by 1.5\%, but the number of gene-disease association studies decreased by 28\%.

One argument in favour of an effect of the stricter editorial guidelines, such as those imposed by JASN on study quality, may be the observation that there is a marked increase in studies meeting the secondary requirements in the ASN abstracts (from 17 to 59\%), while the abstracts at ERA/EDTA still show a lower compliance rate (15\% in 2004, 36\% in 2005).

The assessment of compliance with the primary requirements is much harder. Adjustment for $P$-value is certainly a methodological detail that has no place in an abstract, because an abstract must be concise and comprehensive, but cannot ever be all-inclusive. And indeed, the formal limitations and the known deficiency of the abstract as an information source may be considered a methodological weakness of the present study. Abstracts of medical research are deficient to a certain degree even under the most favourable circumstances. A comparison of 44 abstracts from six leading medical journals with their full-text articles showed deficient data reporting in up to 68\%, depending on the journal in which the article was published [11]. In view of this, one might hope that the fuller data disclosure in peer-reviewed articles would allow better assessment. Yet again, only a fraction of the studies assessed here were published within a year of the conferences. By June 2005,
only 10 gene-disease association studies published in peer-reviewed journals could be identified in PubMed that had been presented at the 2004 ERA/EDTA and ASN annual conventions. This amounts to as little as 10% of all relevant abstracts from 2004 at a time lag that largely exceeds the average publication lag of 262 days [12] in the major nephrology journals. The small number of published studies may be an indirect indicator that the more extensive peer review process that a full manuscript has to pass may be a more effective filter against low quality. Of these 10 articles, however, only three (3%) met the JASN secondary requirements. And more precise analysis of one of these [13], co-signed by some of the authors of the present study, shows that, that particular study might be considered as meeting the primary requirements of adequate sample size only if it is considered as a pilot study in the specific population under observation [14].

This points to the most obvious flaw of all studies in question: inadequate sample size. If no minor allele frequencies for a candidate polymorphism in a specific disease and population setting are known, a minimum of 239 samples for cases and controls each is required in order to achieve a statistical power of 80% at the 5% significance level [14]. In the analysed abstracts, however, the average sample sizes for cases and controls combined were 245 (±289) in 2004 and 354 (±147) in 2005.

If the few studies of extreme sample size, defined as exceeding one SD of the sample sizes documented in all studies, are excluded as outliers or extreme values, the average combined sample sizes drop to as little as 185 (±136) in 2004 and 276 (±116) in 2005. In other words: if all studies were to be considered pilot studies (which they are not), the average sample size is only about 38–58% of the required minimum. This would amount to a power of only 41–57% at the 5% significance level, or 48% for the 2 years combined. These values are based on the general assumption that average minor allele frequencies of single nucleotide polymorphism (SNP) in the population are around 7% [15] and the proportion of patients carrying at least one copy of a susceptibility SNP allele is typically close to 15% [14,16].

The power of a study is analogous to the sensitivity of a diagnostic test, and any nephrologist involved in gene-disease association studies should ask himself if he were willing to employ a test of such low sensitivity on one of his patients. The answer is hopefully ‘no’—and accordingly, it should be ‘no’ with regards to the planning, implementation and presentation of gene-disease association studies, too. The complexity of designing, identifying and understanding appropriate studies is likely to surpass the capacities of anyone but trained specialists in the field. Even experienced statisticians seem to have difficulties in correctly interpreting the significance of results obtained in case-control settings [17].

There are possibilities for improving study quality at several levels.

Learned associations and editors of scientific periodicals can make use of the pressure to publish by rigorously implementing meaningful guidelines. This would make studies of inadequate design unpublishable, and could thus reduce the attraction of conducting fast but useless research. The decrease in the number of abstracts on gene-disease association studies from 2004 to 2005 may reflect an effect of JASN’s decision in 2003, to adopt such stricter guidelines. Alternatively, learned societies might establish and maintain databases that list ongoing studies, in analogy to the databases that have become standard in clinical trials [18]. In order to make better use of the already published data, despite the often too small sample numbers, meta-analysis seems to provide a good means to create knowledge [5].

Most importantly, however, individual researchers need to realize that in most cases, adequate sample sizes cannot be recruited in single centres. The solution to this problem is the establishment of research networks. In order to allow scientists to share resources and expertise, a ‘network of networks’, the HuGE Network has been set up by the Human Genome Epidemiology Network [19]. More specifically looking at the field of nephrology, the Renal Genome Network (ReGeNet, www.regenet.org) was established in 2003, which promotes this networking approach between academic research groups, the ERA/EDTA Registry and commercial enterprises.

If the research community does not apply rigid quality criteria, the science of heredity remains what it was 250 years ago: ‘The illusion of an explanation’ [9].

Acknowledgements. The study was financed by an intramural research grant from Singapore’s Biomedical Research Council. A sub-analysis of the 2004 data only was presented at the 42nd ERA/EDTA meeting in 2005. That presentation was supported by an ERA/EDTA conference award and travel grant.

AM dedicates this article to Professor Berend Grabezee in grateful recognition of the invaluable clinical education he has conveyed.

Conflict of interest statement. None declared.

(See related article by Brenchley et al. Nephrol Dial Transplant 2006; 21: 2681–2683.)

References


Received for publication: 13.12.05
Accepted in revised form: 15.5.06