Genome-Wide Association Studies (GWAS)

A Checklist of Methodological and Conceptual Requirements

Emanated by a DFG Round Table of Experts
21 November 2008

Background

Modern high-throughput genotyping has provided scientists with a systematic means to study the heritable basis of both qualitative and quantitative human phenotypes at a genome-wide level (so called "genome-wide association studies", GWAS). However, these technologies are still comparatively expensive, and the cost currently incurred at reasonable sample sizes may well exceed the usual limits of public funding schemes. Therefore, scientists applying for the financial support of GWAS can be requested to comply with a minimum set of criteria regarding the conceptual and methodical quality of their projects.

The following checklist is intended for double use. First, such as other established guidelines (e.g. Good Epidemiological Practice - GEP, see further reading) it shall serve as a specific guideline for scientists who are planning a GWAS in humans and who are seeking funding for such a project by the Deutsche Forschungsgemeinschaft (DFG). The checklist can assist them in designing and planning their study and in preparing a grant application in addition to the general funding guidelines of the DFG, which are not to be replaced. Second, the checklist can help reviewers of grant applications to assess the quality of a GWAS project in standardized fashion. Furthermore, as such a list gives the study sections ("Fachkollegien") a quick overview, and comparisons among GWAS applications between different study sections might be simplified. To allow a review by international experts, these applications should be submitted in English language.

Checklist

General Aspects of the Project

1. The goal and rationale of the study is clear-defined.
2. The scientific or practical medical interest in the phenotype under study warrants investment of the financial resources requested.
3. The applicants have convincingly demonstrated that high-throughput genotyping of the suggested scope is necessary to achieve the goal of the study[1].
4. The novelty of the anticipated results is sufficiently high to warrant investment of the requested financial resources.
5. The applicants have devised a sensible and viable strategy for the replication and follow-up of any positive results of their project[2].
6. The applicants have devised a sensible and viable concept for the handling of data and samples after the project has terminated.

[1] For example, the use of existing, so-called "historical" genotypes may often represent a scientifically valid means to avoid additional genotyping.

[2] Ideally, the replication and, to some extent, the follow-up of positive results should be an integral part of the original project plan. For reasons of quality control, it may often be sensible, to use different genotyping technologies at different stages of the project.

7. The project is set in a clear context of other existing and relevant networks, consortia, platforms or funding streams and the generation of potential synergies is well described.

Phenotyping and Field Work

8. The phenotype under study is well described and can be determined unequivocally[3].
9. The sensibility of undertaking a GWAS for the phenotype under study has been demonstrated unequivocally, either by a formal proof of heritability[4] or by alternative scientific reasoning.
10. The applicants have made a convincing case for their capability to obtain all data and samples necessary to achieve the goal of the study or for their integration into collaborative efforts with that capability.
11. All field work and phenotyping, including the retrieval of environmental and life-style information, will be carried out by sufficiently qualified personnel.

Genotyping

12. The choice of genotyping technology is appropriate and has been sufficiently justified.
13. At any given stage of the project, samples will be analysed with the same (or highly comparable) technologies and following the same (or highly comparable) experimental protocols, independent of phenotype.
14. All data acquisition, transfer and storage procedures have been described in sufficient detail and comply with current technical standards and applicable legal regulations[5].
15. Appropriate mechanisms of quality control will be implemented at all stages of the project, including pre-analytical sample handling, sample processing and storage as well as data acquisition, management and analysis.
16. The collection of sufficient amounts of high-quality DNA will be completed before genotyping starts.
17. All experimental work including pre-analytical sample handling will be carried out by sufficiently qualified personnel.

Data Analysis

18. The statistical methods of data analysis including data pre-processing and cleaning, are appropriate, well described and their choice has been sufficiently justified[6].
19. Sample size and power considerations have been based upon realistic assumptions about the effect sizes, modes of inheritance, and genotype frequencies involved.[7]
20. Measures will be taken to avoid confounding by population stratification or to adjust for such confounding during data analysis.
21. All computing and data storage capacities necessary to achieve the goal of the study will be available at the time when the project starts.
22. All data handling will be carried out by sufficiently qualified personnel.

---

[3] Consistent phenotyping can be ensured, for example, by the use of standardised, validated and internationally accepted diagnostic criteria.

[4] Heritability may have been demonstrated, for example, by adoption studies, segregation analysis, or classical heritability estimates from close relatives. It is important, however, that the phenotype under study is either identical or highly comparable to the phenotype for which heritability has been demonstrated.

[5] It should be noted that the legal requirements of a genetic epidemiological study usually include the involvement of an institutional ethics committee.

[6] In their description of the statistical methodology, applicants should be explicit about the criteria used for declaring a positive result statistically significant.

[7] In cases where no prior information about the respective parameters is available, experience from previous GWAS may be used.

23. The biometrical and statistical analysis will be carried out by sufficiently qualified personnel.

Administration

24. Informed consent that covers all activities connected to the GWAS has been obtained, or will be obtained, from all participants.
25. All activities necessary to achieve the goal of the study, including external data and sample acquisition and transfer, have been secured by legally binding agreements.

Further Reading

Hoffmann W, Latza U, Terschüren C. Leitlinien und Empfehlungen zur Sicherung von Guter Epidemiologischer Praxis (GEP) - überarbeitete Fassung nach Evaluation. Gesundheitswesen 2005;67:217-225

Leitlinien und Empfehlungen zur Sicherung von Guter Epidemiologischer Praxis (GEP) - Langversion
Ergänzung der Leitlinien für Gute Epidemiologische Praxis (GEP) durch spezifische Ausführungsbestimmungen für einzelne Fachgebiete innerhalb der Epidemiologie. Deutsche Gesellschaft für Epidemiologie (DGEpi) in Zusammenarbeit mit der Deutschen Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie (GMDS), Deutschen Gesellschaft für Sozialmedizin und Prävention (DGSMP), Deutschen Region der Internationalen Biometrischen Gesellschaft (DR-IBS). Mit Änderungen nach Evaluation April 2004. Mit Ergänzung durch Ausführungsbestimmung zur Guten Praxis Sekundärdaten Analyse (GPS) März 2008. (http://www.gmds.de/publikationen/stellungnahmen.php).

McCarthy MI, Abecasis GR, Cardon LR, Goldstein DB, Little J, Ioannidis JP, Hirschhorn JN (2008) Genome-wide association studies for complex traits: consensus, uncertainty and challenges. Nat Rev Genet 9: 356-369.