Strengthening the reporting of genetic risk prediction studies: the GRIPS statement

A Cecile J W Janssens,[1] John P A Ioannidis,[2 3 4 5 6] Cornelia M van Duijn,[1] Julian Little,[7] Muin J Khoury,[8] for the GRIPS Group[¶]

[1]Department of Epidemiology, Erasmus University Medical Center, Rotterdam, Netherlands
[2]Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, Ioannina, Greece
[3]Biomedical Research Institute, Foundation for Research and Technology, Ioannina
[4]Department of Medicine, Tufts University School of Medicine, Boston, Massachusetts, USA
[5]Center for Genetic Epidemiology and Modeling and Tufts CTSI, Institute for Clinical Research and Health Policy Studies, Tufts Medical Center, Boston
[6]Stanford Prevention Research Center, Stanford University School of Medicine, Stanford, California, USA
[7]Department of Epidemiology and Community Medicine, University of Ottawa, Ottawa, Ontario, Canada
[8]Office of Public Health Genomics, Centers for Disease Control and Prevention, Atlanta, Georgia, USA

Correspondence to:
A C J W Janssens
a.janssens@erasmusmc.nl

Cite this as: *BMJ* 2011;342:d631
doi: 10.1136/bmj.d631

The number of known genetic markers of risk is increasing but the interpretation of their clinical effect is hampered by poor reporting of prediction studies. These guidelines from the GRIPS group aim to ensure transparent reporting of prediction studies

The recent successes of genome-wide association studies and the promises of whole genome sequencing fuel interest in the translation of this new wave of basic genetic knowledge to healthcare practice. Knowledge about genetic risk factors may be used to target diagnostic, preventive, and therapeutic interventions for complex disorders based on a person's genetic risk or to complement existing risk models based on classic non-genetic factors such as the Framingham risk score for cardiovascular disease. Implementation of genetic risk prediction in healthcare requires a series of studies that encompass all phases of translational research,[1 2] starting with a comprehensive evaluation of genetic risk prediction.

With increasing numbers of discovered genetic markers that can be used in future genetic risk prediction studies, it is crucial to enhance the quality of the reporting of these studies, since valid interpretation could be compromised by the lack of reporting of key information. Information that is often missing includes details in the description of how the study was designed and conducted (eg, how genetic variants were selected and coded, how risk models or genetic risk scores were constructed, and how risk categories were chosen), or how the results should be interpreted. An appropriate assessment of the study's strengths and weaknesses is not possible without this information. There is ample evidence that prediction research often suffers from poor design and bias, and these may also have an impact on the results of the studies and on models of disease outcomes based on these studies.[3-5] Although most prognostic studies published to date claim significant results,[6 7] very few translate to clinically useful applications. Just as for observational epidemiological studies,[8] poor reporting complicates the use of the specific study for research, clinical, or public health purposes and hampers the synthesis of evidence across studies.

Reporting guidelines have been published for various research designs,[9] and these contain many items that are also relevant to genetic risk prediction studies. In particular, the guidelines for genetic association studies (STREGA) have relevant items on the assessment of genetic variants, and the guidelines for observational studies (STROBE) have relevant items about the reporting of study design. The guidelines for diagnostic studies (STARD) and those for tumour marker prognostic studies (REMARK) include relevant items about test evaluation; the REMARK guidelines also have relevant items about risk prediction.[5 10-12] However, none of these guidelines are fully suited to genetic risk prediction studies, an emerging field of investigation with specific methodological issues that need to be addressed, such as the handling of large numbers of genetic variants (from 10s to 10 000s) and flexibility in handling such large numbers in analyses. We organised a two day workshop with an international group of risk prediction researchers, epidemiologists, geneticists, methodologists, statisticians, and journal editors to develop recommendations for the reporting of genetic risk prediction studies (GRIPS).

SUMMARY POINTS

The rapid and continuing progress in gene discovery for complex diseases is fuelling interest in the potential application of genetic risk models for clinical and public health practice

The number of studies assessing the predictive ability is steadily increasing, but the quality and completeness of reporting varies

A multidisciplinary workshop sponsored by the Human Genome Epidemiology Network developed a checklist of 25 items recommended for strengthening the reporting of genetic risk prediction studies (GRIPS), building on the principles established by prior reporting guidelines

These recommendations aim to enhance the transparency of study reporting and thereby improve the synthesis and application of information from multiple studies that might differ in design, conduct, or analysis

Genetic risk prediction studies

Genetic risk prediction studies typically develop or validate models that predict the risk of disease, but they are also being investigated for use in predicting prognostic outcome, treatment response, or treatment related harms. Risk prediction models are statistical algorithms, which may be simple genetic risk scores (for example, risk allele counts), may be based on regression analyses (for example, weighted risk scores or predicted risks), or may be based on more complex analytical approaches such as support vector machine learning or classification trees. The risk models may be based on genetic variants only or include both genetic and non-genetic risk factors.[13]

Reporting recommendations for evaluations of risk prediction models that include genetic variants

Report section	Item No	Recommendation
Title and abstract		
	1	(*a*) Identify the article as a study of risk prediction using genetic factors
		(*b*) Use recommended keywords in the abstract: genetic or genomic, risk, prediction
Introduction		
Background and rationale	2	Explain the scientific background and rationale for the prediction study
Objectives	3	Specify the study objectives and state the specific model(s) that is/are investigated. State if the study concerns the development of the model(s), a validation effort, or both
Methods		
Study design and setting	4*	Specify the key elements of the study design and describe the setting, locations, and relevant dates, including periods of recruitment, follow-up, and data collection
Participants	5*	Describe eligibility criteria for participants, and sources and methods of selection of participants
Variables: definition	6*	Clearly define all participant characteristics, risk factors and outcomes. Clearly define genetic variants using a widely used nomenclature system
Variables: assessment	7*	(*a*) Describe sources of data and details of methods of assessment (measurement) for each variable
		(*b*) Give a detailed description of genotyping and other laboratory methods
Variables: coding	8	(*a*) Describe how genetic variants were handled in the analyses
		(*b*) Explain how other quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen, and why
Analysis: risk model construction	9	Specify the procedure and data used for the derivation of the risk model. Specify which candidate variables were initially examined or considered for inclusion in models. Include details of any variable selection procedures and other model building issues. Specify the horizon of risk prediction (eg, 5 year risk)
Analysis: validation	10	Specify the procedure and data used for the validation of the risk model
Analysis: missing data	11	Specify how missing data were handled
Analysis: statistical methods	12	Specify all measures used for the evaluation of the risk model including, but not limited to, measures of model fit and predictive ability
Analysis: other	13	Describe all subgroups, interactions, and exploratory analyses that were examined
Results		
Participants	14*	Report the numbers of individuals at each stage of the study. Give reasons for non-participation at each stage. Report the number of participants not genotyped and reasons why they were not genotyped
Descriptives: population	15*	Report demographic and clinical characteristics of the study population, including risk factors used in the risk modelling
Descriptives: model estimates	16	Report unadjusted associations between the variables in the risk model(s) and the outcome. Report adjusted estimates and their precision from the full risk model(s) for each variable
Risk distributions	17*	Report distributions of predicted risks and/or risk scores
Assessment	18	Report measures of model fit and predictive ability, and any other performance measures, if pertinent.
Validation	19	Report any validation of the risk model(s)
Other analyses	20	Present results of any subgroup, interaction, or exploratory analyses, whenever pertinent
Discussion		
Limitations	21	Discuss limitations and assumptions of the study, particularly those concerning study design, selection of participants, and measurements and analyses, and discuss their impact on the results of the study.
Interpretation	22	Give an overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence
Generalisability	23	Discuss the generalisability and, if pertinent, the healthcare relevance of the study results
Other		
Supplementary information	24	State whether databases for the analysed data, risk models, and/or protocols are or will become publicly available and if so, how they can be accessed
Funding	25	Give the source of funding and the role of the funders for the present study. State whether there are any conflicts of interest

* Marked items should be reported for every population in the study.

Aims and use of the GRIPS statement

The 25 items of the GRIPS statement are intended to maximise the transparency, quality, and completeness of reporting on research methodology and findings in a particular study. It is important to emphasise that these recommendations are guidelines only for how to report research and do not prescribe how to perform genetic risk prediction studies. The guidelines do not support or oppose the choice of any particular study design or method—for example, the guidelines recommend that the study population should be described but do not specify which population is preferred in a particular study.

The intended audience for the reporting guidelines is broad and includes epidemiologists, geneticists, statisticians, clinician scientists, and laboratory based investigators who undertake genetic risk prediction studies, as well as journal editors and reviewers who have to appraise the design, conduct, and analysis of such studies. In addition, it includes "users" of such studies who wish to understand the basic premise, design, and limitations of genetic prediction studies in order to interpret the results for their potential application in healthcare. These guidelines are also intended to ensure that essential data from future genetic risk prediction studies are presented in standardised form, which will facilitate information synthesis as part of systematic reviews and meta-analyses.

Items presented in the checklist are relevant for a wide array of risk prediction studies, because GRIPS focuses on the main aspects of the design and analysis of risk prediction studies. GRIPS does not address randomised

trials that may be performed to test risk models, nor does it specifically address decision analyses, cost effectiveness analyses, assessment of healthcare needs, or assessment of barriers to healthcare implementation.[14] Once the performance of a risk model has been established, these next steps toward implementation require further evaluation.[10 15] For the reporting of these studies, which go beyond the assessment of genetic risk models as such, additional requirements apply. However, proper documentation of genetic predictive research according to GRIPS might facilitate the translation of research findings into clinical and public health practice.

Development of the GRIPS statement

The GRIPS statement was developed by a multidisciplinary panel of 25 risk prediction researchers, epidemiologists, geneticists, methodologists, statisticians, and journal editors, seven of whom were also part of the STREGA initiative.[11] They attended a two day meeting in Atlanta, Georgia, USA, in December 2009 that was sponsored by the US Centers for Disease Control and Prevention on behalf of the Human Genome Epidemiology Network (HuGENet).[16] Participants discussed a draft version of the guidelines that was prepared and distributed before the meeting. This draft version was developed on the basis of existing reporting guidelines—namely, STREGA,[11] REMARK,[5] and STARD.[12] These were selected out of all available guidelines (see www.equator-network.org) because of their focus on observational study designs and genetic factors (STREGA), prediction models (REMARK), and test evaluation (REMARK and STARD).

During the meeting, methodological issues pertinent to risk prediction studies were addressed in presentations. Workshop participants were asked to change, combine, or delete proposed items and add additional items if necessary. Participants had extensive post-meeting electronic correspondence. To harmonise our recommendations for genetic risk prediction studies with previous guidelines, we chose the same wording for the items wherever possible. Finally, we tried to create consistency with previous guidelines for the evaluation of risk prediction studies of cardiovascular diseases and cancer.[2 17] The table gives the final version of the checklist.

GRIPS explanation and elaboration article

To accompany this GRIPS statement, we have written an explanation and elaboration document,[18] modelled on those developed for other reporting guidelines.[19-22] The document illustrates each item with at least one published example that we consider transparent in reporting, explains the rationale for its inclusion in the checklist, and presents details of the items that need to be addressed to ensure transparent reporting. The explanation and elaboration document was produced after the meeting. The document was prepared by a small subgroup and shared with all workshop participants for additional revisions and final approval.

Concluding remarks and future directions

High quality reporting reveals the strengths and weaknesses of empirical studies, facilitates the interpretation of the scientific and healthcare relevance of the results—especially within the framework of systematic reviews and meta-analyses—and helps build a solid evidence base for moving genomic discoveries into applications in healthcare practice. The GRIPS guidelines were developed to improve the transparency, quality, and completeness of the reporting of genetic risk prediction studies. As outlined in the introduction, GRIPS does not prescribe how studies should be designed, conducted, or analysed, and therefore the guidelines should not be used to assess the quality of empirical studies.[23] The guidelines should be used only to check whether all essential items are adequately reported.

Finally, the methodology for designing and assessing genetic risk prediction models is still developing. For example, newer measures of reclassification were first introduced in 2007,[24] and several alternative reclassification measures have been proposed.[25] Which measures to apply and when to use measures of reclassification are still subject to ongoing evaluation and discussion.[26] Furthermore, alternative strategies for constructing risk models other than simple regression analyses are being explored, and these may add increased complexity to the reporting. In formulating the items of the GRIPS statement, these methodological advances were anticipated. It is for this reason that the GRIPS statement recommends how a study should be reported and not how a study should be conducted or analysed. Therefore, methodological and analytical developments will not immediately impact the validity and relevance of the items, but the GRIPS statement will be updated when this is warranted by essential new developments in the construction and evaluation of genetic risk models.

Members of the GRIPS group are: A Cecile J W Janssens, John P A Ioannidis, Sara Bedrosian, Paolo Boffetta, Siobhan M Dolan, Nicole Dowling, Isabel Fortier, Andrew N Freedman, Jeremy M Grimshaw, Jeffrey Gulcher, Marta Gwinn, Mark A Hlatky, Holly Janes, Peter Kraft, Stephanie Melillo, Christopher J O'Donnell, Michael J Pencina, David Ransohoff, Sheri D Schully, Daniela Seminara, Deborah M Winn, Caroline F Wright, Cornelia M van Duijn, Julian Little, and Muin J Khoury.

Funding: The workshop was sponsored by the Centers for Disease Control and Prevention on behalf of the Human Genome Epidemiology Network. Opinions in this paper do not necessarily represent the official position or policies of the authors' employers, and the funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: All authors have completed the unified competing interest form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) and declare support from the Human Genome Epidemiology Network; ACWJ is supported by grants from the Erasmus University Medical Centre Rotterdam, the Centre for Medical Systems Biology in the framework of the Netherlands Genomics Initiative (NGI) and the VIDI grant of the Netherlands Organisation for Scientific Research (NWO) and Tufts CTSI is supported by the National Institutes of Health/National Center for Research Resources (UL1 RR025752); they have no other relationships or activities that could appear to have influenced the submitted work.

In order to encourage dissemination of the GRIPS statement, this article will also be published by *Annals of Internal Medicine*, *BMJ*, *Circulation: Cardiovascular Genetics*, *European Journal of Clinical Investigation*, *European Journal of Epidemiology*, *European Journal of Human Genetics*, *Genetics in Medicine*, *Genome Medicine*, and *Journal of Clinical Epidemiology*. A detailed explanation and elaboration document is published as supporting information.[18]

Provenance and peer review: Not commissioned; externally peer reviewed.

bmj.com Previous articles in this series

- Statistics Notes: Correlation in restricted ranges of data (*BMJ* 2011;342:d556)
- Interpretation of random effects meta-analyses (*BMJ* 2011;342:d549)
- Strategy for intention to treat analysis in randomised trials with missing outcome data (*BMJ* 2011;342:d40)
- Three techniques for integrating data in mixed methods studies (*BMJ* 2010;341:c4587)
- Assessing equity in systematic reviews: realising the recommendations of the Commission on Social Determinants of Health (*BMJ* 2010;341:c4739)

1 Khoury MJ, Gwinn M, Yoon PW, Dowling N, Moore CA, Bradley L, et al. The continuum of translation research in genomic medicine: how can we accelerate the appropriate integration of human genome discoveries into health care and disease prevention? *Genet Med* 2007;9:665-74.
2 Hlatky MA, Greenland P, Arnett DK, Ballantyne CM, Criqui MH, Elkind MS, et al. Criteria for evaluation of novel markers of cardiovascular risk: a scientific statement from the American Heart Association. *Circulation* 2009;119:2408-16.
3 Kyzas PA, Denaxa-Kyza D, Ioannidis JP. Quality of reporting of cancer prognostic marker studies: association with reported prognostic effect. *J Natl Cancer Inst* 2007;99:236-43.
4 Kyzas PA, Loizou KT, Ioannidis JP. Selective reporting biases in cancer prognostic factor studies. *J Natl Cancer Inst* 2005;97:1043-55.
5 McShane LM, Altman DG, Sauerbrei W, Taube SE, Gion M, Clark GM, et al. Reporting recommendations for tumor marker prognostic studies (REMARK). *Nat Clin Pract Oncol* 2005;2:416-22.
6 Kyzas PA, Denaxa-Kyza D, Ioannidis JP. Almost all articles on cancer prognostic markers report statistically significant results. *Eur J Cancer* 2007;43:2559-79.
7 Tzoulaki I, Liberopoulos G, Ioannidis JP. Assessment of claims of improved prediction beyond the Framingham risk score. *JAMA* 2009;302:2345-2352.
8 Von Elm E, Egger M. The scandal of poor epidemiological research. *BMJ* 2004;329:868-9.
9 Simera I, Moher D, Hoey J, Schulz KF, Altman DG. A catalogue of reporting guidelines for health research. *Eur J Clin Invest* 2010;40:35-53.
10 Von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP, et al. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *PLoS Med* 2007;4:e296.
11 Little J, Higgins JP, Ioannidis JP, Moher D, Gagnon F, von Elm E, et al. Strengthening the reporting of genetic association studies (STREGA): an extension of the STROBE statement. *PLoS Med* 2009;6:e22.
12 Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *BMJ* 2003;326:41-4.
13 Janssens ACJW, van Duijn CM. Genome-based prediction of common diseases: methodological considerations for future research. *Genome Med* 2009;1:20.
14 Khoury MJ, Gwinn M, Ioannidis JP. The emergence of translational epidemiology: from scientific discovery to population health impact. *Am J Epidemiol* 2010;172:517-24.
15 Moons KG, Altman DG, Vergouwe Y, Royston P. Prognosis and prognostic research: application and impact of prognostic models in clinical practice. *BMJ* 2009;338:b606.
16 Khoury MJ, Dorman JS. The Human Genome Epidemiology Network. *Am J Epidemiol* 1998;148:1-3.
17 Freedman AN, Seminara D, Gail MH, Hartge P, Colditz GA, Ballard-Barbash R, et al. Cancer risk prediction models: a workshop on development, evaluation, and application. *J Natl Cancer Inst* 2005;97:715-23.
18 Janssens ACJW, Ioannidis JPA, Bedrosian S, Boffetta P, Dolan SM, Dowling N, et al. Strengthening the reporting of genetic risk prediction studies (GRIPS): explanation and elaboration. *PLoS Med* 2011; doi:10.1371/journal.pmed.1000420.
19 Altman DG, Schulz KF, Moher D, Egger M, Davidoff F, Elbourne D, et al. The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Ann Intern Med* 2001;134:663-94.
20 Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, et al. The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration. *Ann Intern Med* 2003;138:w1-12.
21 Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JPA, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *BMJ* 2009;339:b2700.
22 Vandenbroucke JP, von Elm E, Altman DG, Gotzsche PC, Mulrow CD, Pocock SJ, et al. Strengthening the reporting of observational studies in epidemiology (STROBE): explanation and elaboration. *PLoS Med* 2007;4:e297.
23 Vandenbroucke JP. STREGA, STROBE, STARD, SQUIRE, MOOSE, PRISMA, GNOSIS, TREND, ORION, COREQ, QUOROM, REMARK... and CONSORT: for whom does the guideline toll? *J Clin Epidemiol* 2009;62:594-6.
24 Cook NR. Use and misuse of the receiver operating characteristic curve in risk prediction. *Circulation* 2007;115:928-35.
25 Pencina MJ, D'Agostino RB Sr, D'Agostino RB Jr, Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Stat Med* 2008;27:157-72.
26 Janssens ACJW, Khoury MJ. Assessment of improved prediction beyond traditional risk factors: when does a difference make a difference? *Circ: Cardiovasc Genet* 2010;3:3-5.

Lost in translation

I recently had the misfortune of being admitted to hospital with a minor medical ailment that required an overnight stay. I spent time on two wards with groups of women, all more than double my age, who had a wide variety of surgical and medical conditions. When the time came for the consultant ward round it was impossible for me not to hear every word of each consultation as curtains offer very little in the way of privacy. I even knew that one lady had cancer before she did because I heard the team discussing what they were going to say to her before they came into the bay.

What interested me was listening to these same ladies explaining to their relatives later in the day what the doctor had said. Not one of the other patients told their relatives the correct information. How was it that I knew what was wrong with them, including the ongoing plan, and had clearly heard the doctors telling them, and yet the patients seemed not to have heard the same information?

One of the plans had been quite muddled, so I can understand why the patient was a bit confused, but the others had all been clear to me. Listening to my fellow roommates revealed some interesting things. One woman was deaf in one ear and couldn't hear anything the doctors had been saying—but she didn't like to tell them, and so relied on her next door neighbour to fill her in. Another patient was so terrified when the doctors came round in a big group that she hadn't been able to take anything in, and another had severe dementia and confabulated much of her history to her relatives. What became clear was that the patients did not understand many of the medical terms the doctors used. Things that seemed simple to me and explanations that gave me a clear picture of what was going on left the other patients confused because they did not understand the language of medicine. Pneumonia, CT scan, saturations, ultrasound, UTI, and arrhythmia were all terms that left patients confused.

It took a lot of willpower for me not to try to be "the doctor" and accept that I was just another patient. However, I hope I have learnt some valuable lessons from my brief hospital stay that I can take back to my everyday practice.

My patients are people. They may be scared and intimidated, and taking the time to talk to them and explaining what is happening to them is crucial. They may not admit when they don't understand something, but avoiding medical jargon, allowing time for questions, and making sure they can hear me are good first steps. It is always worth asking patients what they understand about the situation so far, as it gives you a good starting point to begin your conversation.

However, one of my biggest lessons came when I found myself pulling on a pair of red cat pyjamas. Never let your husband pack your overnight bag.

Jennifer Graves respiratory specialist registrar, Southmead Hospital, Bristol jennifercrow21@hotmail.com

Cite this as: *BMJ* 2010;340:c414