Am. J. Hum. Genet. 63:678–681, 1998

HUMAN GENETICS '98: ETHICAL ISSUES IN GENETICS
Pedigrees—Publish? or Perish the Thought?

Peter H. Byers and John Ashkenas

The American Journal of Human Genetics

Pedigrees are the lifeblood of research in human genetics. They provide a shorthand description of the genetic and social relationships among relatives, and they allow modes of transmission of heritable traits or patterns of inheritance of haplotype to be assessed at a glance. For this reason, they are preferable to more poetic but cumbersome descriptions of family structure, such as the following early attempt: "Adam lived an hundred and thirty years, and begat a son in his own likeness, and after his image; and called his name Seth: And the days of Adam after he had begotten Seth were eight hundred years: and he begat sons and daughters" (Gen. 5:3–4, King James Version). At another extreme are tables of LOD scores, which display aggregate family data in a completely anonymous form but which lack the visual immediacy of the pedigree.

Although published pedigrees no longer contain overt identifying information, concern persists that they carry sufficient information to permit individuals who so wish to recognize some or all subjects depicted, to the subjects' possible detriment. In a discussion of patient anonymity, the International Committee of Medical Journal Editors (1995) responded to these concerns with the declaration that "identifying information should not be published in written descriptions, *photographs,* or *pedigrees* [emphasis added] unless the information is essential for scientific purposes and the patient (or parent or guardian) gives written informed consent for publication."

Although this statement implicitly equates photographic and pedigree data as risks to subject anonymity, there are compelling reasons to doubt their equivalence. A photograph per se is an identifying document. Hence, the practice of masking facial features in published clinical photographs, although not always sufficient to preserve anonymity, is certainly laudable in general. It is conceivable but unlikely that such masking will interfere with the future uses of the published data: a photograph that was not intended to display a facial dysmorphism is unlikely to do so effectively, especially if the anomaly is subtle enough that the initial investigators failed to notice it. Pedigrees, in contrast, are seldom, if ever, sufficient to identify either an individual or a family, but they carry a wealth of information that may prove useful in ways the researcher cannot anticipate.

Recently, Botkin and colleagues (1998) surveyed investigators and journal editors, asking about the extent to which potentially identifying information had been disguised prior to publication, the manner in which the information contained in the pedigrees was communicated to the subjects, and the extent to which alterations were described to the readers. The masking of pedigree data was acknowledged by both authors and journal editors and was permitted or, occasionally, encouraged by the latter. Few journals offered instructions for how to inform readers about modifications present in the published diagrams. Interestingly, the journals surveyed (including this one) handled the information in pedigrees and in photographs quite differently in that they did not require individual release for the publication of pedigrees. We think that, despite the Medical Journal Editors' policy stated above, photographic and pedigree data do not pose equivalent threats to confidentiality. The crucial differences between the two kinds of data may justify the different handling they appear to receive.

About half the research articles published in *The American Journal of Human Genetics* contain pedigree diagrams, which vary in degree of detail, depending on the heuristic purposes of the authors. For example, whereas large multigenerational families might be shown to display the fine mapping of a locus, all data regarding paternal lines of descent are routinely removed from pedigrees that are intended to display mitochondrial inheritance. Similarly, truncated nuclear families may suffice to illustrate the Mendelian transmission of a well-defined clinical condition or a mutant allele. Evidently, authors choose, in the interest of clarity of presentation, to omit

Received July 17, 1998; accepted for publication July 17, 1998; electronically published August 21, 1998.

Address for correspondence and reprints: Dr. Peter H. Byers, *The American Journal of Human Genetics,* University of Washington, Box 357470, Seattle, WA 98195-7470. E-mail: AJHG@u.washington.edu

This article represents the opinion of the authors and does not represent an official position of the ASHG.

0002-9297/98/6303-0003$02.00

information that might have been available to them. In some instances, it is clear that these pedigrees also contain alterations, the most common of which is masked gender. Although it is usually stated that masking was done for the purpose of anonymity, we have not always been able to discern the reason for masking. For instance, when pedigrees have been both truncated and masked for gender, the truncation alone often seems to be sufficient to keep an outsider who lacks specific additional information from identifying the subjects. In addition, we have not always been confident that the pedigree depicts the entire and correct family structure.

Dilemmas of Masking

In rare instances, we have been asked by contributors whether it was permissible to mask individual identity in pedigrees or, more rarely, to try to mask the entire pedigree so as not to allow identification of the family. These requests can have unusual ramifications if, for instance, a mating that sets the phase for a known altered allele involves a biological father who is not the acknowledged social father. If the remaining matings are sufficiently informative, the alleles or haplotypes that arise from the problematic mating might be omitted from a published figure without substantial loss of information. The reasons for these alterations should be disclosed to the editor but need not necessarily be acknowledged in print; in less sensitive cases, the reader should be made aware that published pedigrees are masked or altered.

Although alteration of pedigrees may enhance anonymity for the subjects, it is not without pitfalls. It is one of the few instances in which publication of falsified data is knowingly condoned. Important genetic phenomena, including anticipation and its biological basis, the maternal inheritance of mitochondrial disorders, and parent-of-origin effects in imprinted genes, were all recognized, in part, from analysis of pedigrees. It would be difficult to document such exceptions to Mendelian inheritance without appropriate illustrative diagrams—namely, unmasked and unaltered pedigrees.

It would be impossible to use previously published data with any confidence if we did not know that the data represented in pedigrees were correct. Ideally, if modified pedigrees must be presented, then the raw data should be available to other investigators, as long as subject anonymity can be assured. The practical issues of maintaining such a database over many years would be daunting if the task fell to the journals; ultimately it remains the responsibility of the investigator (or, in some cases, the institution that sponsors the research) to assure that the primary, unaltered data are available for retrospective analysis. With the advent of electronic "communities" of researchers who make their data available for each other has come a new middle ground between published and unpublished material. Participants in international mapping consortia would surely risk misleading their colleagues if they withheld or distorted data from their postings to the communal Internet site. The risk to subject privacy is relatively low when unaltered data are submitted to such protected web sites, and we think that, in general, primary data should be shared in this format without masking or truncation.

Real and Imagined Risks from Publication of Pedigrees

The possible abuses of pedigree data fall into several categories. First, pedigree publication could harm the subjects financially if they allow insurers to identify and then to raise the premiums of—or deny coverage to—the individuals in the study. In most instances, the risk of such disclosure seems remote. As we suggest in a fanciful "Wanted" poster in figure 1, published pedigrees contain too little information to allow unique identification in the absence of additional knowledge. Admittedly, as insurers accumulate more information on their customers, it may be possible to scan scientific literature for "matches" to the pedigrees they may have on file, but there is little evidence that agencies have sought or would have the capacity to apply such aggressive screening of potential clients. Here again, the differences between pedigree data and clinical photographs stand out. Although we are not aware of any attempts by insurers to identify and exclude individuals from coverage on the basis of published clinical photographs, such an abuse would clearly be possible and could be applied in a piecemeal fashion with no great investment. Any attempt to misuse pedigrees in this fashion, by law enforcement agencies, insurers, or others, would require enormous and systematic efforts, to achieve any useful results. Rather than altering scientific practice to anticipate this potential concern, we suggest that legal remedies could be devised if and when the problem arises.

A second concern is the inadvertent transmission of data concerning affected status to individuals within a pedigree or to friends, colleagues, or strangers. Some genetic disorders have a very public face, in which the separation of affected from unaffected individuals is done by laypeople and genetic information is necessary only to identify underlying causes. In these instances, the fortuitous identification of an individual does not change the fact of affected status but may make public the etiology and, perhaps, the natural history of the disorder. Publishing a pedigree is unlikely to cause adverse economic effects, but individuals may rightly feel that their

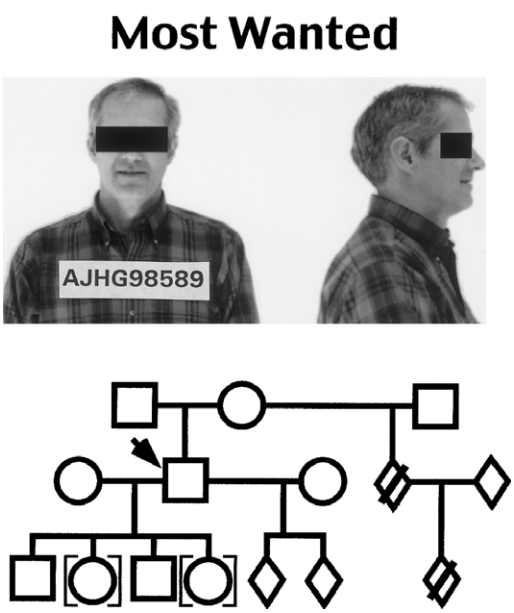

Figure 1 Two kinds of masking. Personal information about a subject's identity and family relations could be of interest to law enforcement agencies, insurance providers, family members, and acquaintances. The obligation of clinicians and geneticists to protect their subjects from the publication of sensitive information may conflict with their own desires to advance their fields and their careers. The need to balance the interests of subjects in genetic studies with those of the scientific community is used to justify the masking of data in pedigrees, usually by removing information about the sex of some or all of the individuals depicted. How does this practice differ from the masking of the eyes in published clinical pictures? Which of these kinds of data pose a threat to subjects' privacy and well-being? Because of crucial differences in the power of photographs and pedigrees to yield identifying information, we judge that the need to mask pedigrees is far less compelling.

privacy has been invaded. It is not clear, however, that this invasion is any greater than occurs when new scientific information about the condition is published without pedigrees. The most pressing issues relate to the disclosure of carrier or presymptomatic status. Publication of pedigrees that disclose these states to previously uninformed subjects could have adverse effects on the subjects depicted—especially if medical care and counseling are not provided to those individuals. The greatest anxiety arises when the disorders in question are stigmatized in our society, as with malignancy, mental illness, or progressive intellectual impairment. Botkin and colleagues (1998) could not identify instances in which they were convinced that publication of pedigrees had caused significant harm of this kind.

It may be useful to distinguish between publishing a pedigree in a clinical study and doing so in a research project, which must proceed in the absence of crucial information. In the former cases, demonstration of the inheritance of a known mutant allele often confirms that a subject is at high risk of a disease, and the investigator and physician consult in advance with family members and explain the consequences of allele detection. In the research setting, however, permission to perform the study is often granted with the proviso that data will not be provided to individuals or families. Subjects are informed of the vagaries introduced by recombination, by incomplete penetrance, or by possible laboratory error. For this reason, once haplotype data become available, researchers are often caught uncomfortably between the institutional policies under which they operate and the desires of subjects to learn potentially useful information about their own health.

Interests of Investigators and Research Subjects

Although the interests of the investigator and the subjects are generally compatible, it is clear that they will not always be wholly coincident. Obligate carrier or affected individuals who have not participated directly in the study raise particularly thorny issues about publication of pedigrees, because their consent is, by definition, lacking. Equally vexing to the investigator (and sometimes to the family) is the individual who participates in the first phase of the study and then withdraws, after his or her genotype is determined, and specifically requests that none of those data be used in the study. If this individual is key to the interpretation of the data, then other approaches must be taken, and the pedigree data may not be publishable.

Concern about the disclosure of genetic status by publication of a pedigree is legitimate and will require additional fine tuning of research protocols, to be sure that no information that could be construed to be identifying is contained within published pedigrees. It would be damaging to the field of genetics if pedigrees were allocated the same identifying status as are photographs of individuals. Explicit information about the use of pedigrees in published material should be included in consent documents. Furthermore, researchers must be prepared to satisfy editors and reviewers that submitted pedigrees will not allow the naive individual to identify the subjects in the diagram. Journals should expect authors to document that the study was approved by the appropriate institutional review boards and should not have to receive written consent for every individual in a pedigree. Such a requirement would impose a far-too-arduous responsibility on the investigator, the subjects, the reviewers, and the editors. Particularly worrisome would be any two-tiered consent system in which consent for participation in the study occurred at one time and separate consent to publish was obtained at a later time.

The issues surrounding pedigrees are not always clear.

The questions of *what* is identifying and *to whom* it is identified need to be resolved, in each instance, by open discussion among families, clinicians, investigators, and members of institutional review boards. In general, it is our impression that there is no substitute for thoughtful discussion between investigator and subjects at the beginning of the research process.

References

Botkin JR, McMahon WM, Smith KR, Nash JE (1998) Privacy and confidentiality in the publication of pedigrees: a survey of investigators and biomedical journals. JAMA 279: 1808–1812

International Committee of Medical Journal Editors (1995) Protection of patients' rights to privacy. BMJ 311:1272