A dozen practical comments about genetics research.
Elizabeth L. Hohmann MD
January 5, 2006

1. Definition of Genetics Research. Any correlation of phenotype and genotype is genetic research. All genetic research protocols should contain a statement about the non-medical (i.e. social and psychological) risks of genetic research. Avoid the "soft" terms "biomarkers, blood markers or polymorphisms;" call a gene a gene, using simple lay definitions. There is argument as to whether RNA expression studies are considered genetic research. Generally this does not have all the same implication is genotype/phenotype correlations, though one can argue that unique expression "signatures" are genetic data unique to individuals and inheritable in some situations. It is up to the investigator to propose whether they consider expression studies "genetics" requiring the discussion of the non-medical risks of genetics in consent forms, depending upon what is proposed in their study.

2. Anonymous vs. identifiable (coded) samples. Clearly differentiate between totally anonymous samples (cannot ever be linked back to an individual, by anyone) and coded samples (carry a code which allow *someone, usually the investigator who collected the sample,* to link the sample back to a specific individual). Do not use the term "anonymous" for coded samples. Do not equate "identifiers" with names or SSN's. Codes serve as "identifiers" as well. Avoid use of the ambiguous phrase "your identity will be unknown to the researchers using the samples" as this can refer to either truly anonymous or coded samples.

3. The value of coded samples retaining a link to subjects. Samples are generally most valuable to investigators if they are coded and *can* be linked back to specific individuals who continue to be followed at our institutions, for many reasons. However, use of such samples is usually somewhat more restricted because these links are maintained. There is most latitude allowed for testing of completely anonymous samples (i.e. de-identified samples, ones which *cannot* be traced back to the subject by *anyone*). Such samples may remain linked to important anonymous clinical data: (i.e. 30 year old man with schizophrenia unresponsive to drug X; 45 year old woman with hyperlipidemia etc). In general, non-identifiable samples may be used and saved for multiple purposes, even those unrelated to the research for which they were collected, as long as they are indeed de-identified. It is possible to save identifiable samples and then de-identify them before other unrelated uses are proposed (submit a secondary use request to the IRB), if this is done carefully and completely. Usually this involves taking out an aliquot and generating a new unlinked code, now associated with only limited important phenotype data but no identifiers or pre-existing codes. Note that this only "works" for larger populations/collections; you may not be truly de-identifying samples if you have 5 patients with rare disease X and retain enough clinical clues to identify them individually.

4. Subjects have "jurisdiction" over use of identifiable samples. Individuals should not lose control of coded DNA samples, i.e. you may not use coded samples for unrelated genetic testing if a subject has not given consent for that. Such testing would require either 1) IRB approval AND re-consenting subjects or 2) IRB-approval AND truly de-identifying samples so that it is *impossible* to link the data back to an individual. Once a sample collection is formally and fully anonymized, it may be possible for future uses to be considered " research, but not human subjects research," a determination that requires an

IRB action. It may be possible for the IRB to consider a waiver of consent for a specific defined secondary use of a sample previously collected for one specific purpose. Secondary uses of DNA means studies ***unrelated*** to your project and not defined in your protocol or consent form. However, the IRB will generally not grant waivers of consent for broad unspecified research uses (i.e. we generally do not allow banking of tissues for broad, future unspecified uses under a waiver; people need to consent to that formally). See the next point for more on this.

5. Related research uses. You may store coded DNA and use it for future, as yet unknown genetic tests *related to your research interest*, as long as this is stated in the consent form. This may be stated in a general way, i.e. "genes related to asthma, inflammation, and obstructive pulmonary disease" or "genes related to Alzheimer's disease and memory disorders." It would not be acceptable to test identifiable samples collected under those consent statements for breast cancer, or addiction research, however. *It is also possible* to propose that subjects' samples be used for general, as yet unspecified research purposes. Previously, some IRBs believed that it was not possible for people to consent to unspecified future uses, but over the past decade this has become more commonplace, and acceptable to both subjects and IRB's. Generally, consent forms requesting wide uses of samples for many different but unspecified uses require more detail on the labeling, duration of storage, confidentiality protections, a description of who controls the samples and how investigators wishing to use the samples "apply" and get them. Subjects need to know they are providing their samples for wide, general uses. This is typically a "tissue bank protocol" or "sample bank." Such a protocol will not describe all the specific possible uses, but will describe how the samples are stored, distributed, controlled etc, and how confidentiality is protected.

6. Sharing samples with collaborators. You may send either coded samples or anonymous samples to other investigators or industrial sponsors, as long as these uses are defined, and specified in the consent form and protocol. Uses of coded samples even if the sponsor does not have a key should be specified and related/limited to the disease under study. If samples will be not be shared with any outside investigators, that should be stated in the consent form, as that will be reassuring to subjects. If you do choose to share samples for collaborative research, there must be a statement noting that collaborators will not know the subject's identity, but one should not state or imply that such coded samples are "anonymous" unless that is really true (see #1 and 2 above). Clarify that the keys to coded samples will never be released to sponsors or collaborators (i.e. identifiable samples will remain in custody of the investigator to whom they were provided). Unrelated genetic testing of coded samples by collaborators or industrial sponsors is not acceptable (see point 6), and would usually be in violation of clinical trials agreements and/or materials transfer agreements. A few sponsors are routinely collecting additional *anonymous* (see above) samples for wide unspecified uses in pharmacogenomics. This is acceptable if truly anonymous and detailed in protocols and consent forms. Such uses are typically optional and described in a separate consent form.

7. Storage of DNA samples – site/duration. Subjects should know how long their samples will be stored, and where. If there is an option to change one's mind and remove samples from further testing, state how that would occur in the consent form. (Usually via submission of written request to PI.) If samples are stored in the truly anonymous fashion this isn't possible. You should explain this to subjects, with a comment to the effect that once the sample is provided, it will not be possible to withdraw it, because cannot be linked to a specific subject.

8. Optional components of studies. If there are portions of the study which are *optional*, that should be stated in the consent form, and subjects allowed to "pick and choose" on the form. For example they may wish to have their DNA used only for genes related to breast cancer research, but not have their DNA stored for other purposes, shared with collaborators, saved long term, or have immortalized cell lines made etc. *It is up to you as investigator to decide how much choice you wish to provide your subjects, balancing scientific needs with patient autonomy, willingness to participate, and "the hassle factor" related to picking and choosing*, *and tracking who picked which option!* It may be good for subjects to have some options, as people will have varying degrees of comfort with genetic testing, but this obligates the investigators to track choices and samples and be sure that subjects' wishes are fulfilled. Some subjects may be overwhelmed by choices. If various activities are *not* optional (e.g. you want to make immortalized cell lines from everybody's blood and store these indefinitely in an identifiable form, and you don't want in your study any subjects who can't agree to this), that must be clearly stated so such subjects may choose not to participate. In general, we recommended that investigators not routinely propose returning to subjects for "re-consenting," because of logistical difficulties related to maintaining contact information, possible progression of illness, and other factors. In general we prefer that investigators provide some reasonable envelope of work that subjects can understand and be comfortable with, for example "genes related to allergy, asthma, and inflammation" or "genes related to diseases of the heart and blood vessels." This informs subjects generally and provides investigators with reasonable scientific latitude without needing to re-contact subjects.

Generally, when studies involve treatment, especially new treatments not otherwise available, the IRB will require that genetics are optional – we don't make genetics a condition for treatment.

9. Immortalized cell lines. Carefully consider whether you wish to have immortalized cell lines retain codes and identifiers. Generally, inexhaustible DNA stores are considered a greater confidentiality risk than finite samples which will eventually be "used up." Proposals requesting indentifiable lines will undergo greater scrutiny by the IRB. If you are making immortalized cell lines *exclusively for non-genetic uses* (e.g. immunology studies, or tissue culture experiments), you should describe those uses in the consent form. In this setting you don't need to refer to genetics and genetic testing risks, but you should address other issues related to confidentiality, storage, etc, and again it is recommended that these lines not be identifiable long term unless essential to the science.

10. Return of information to subjects. When standard medical tests are done and clinically relevant information is obtained, the IRB is generally interested in seeing that such information makes its way back to the subject and/or their physician of choice in a medically appropriate way. For example it would not be acceptable for a healthy volunteer to enroll in a study, be found to have a blood pressure of 180/100, and not be told about this because the study provided an option to not receive medical information. There are known risks of hypertension and standard treatments. The universe of healthy volunteer is large and one can avoid enrolling subjects who don't wish to take the risk of having their blood pressure taken, and have hypertension discovered. The IRB would expect that these subjects would receive appropriate referrals for care, if hypertension is discovered. Investigators should practice good medicine.

This approach cannot be directly applied to genetics, but some extrapolations can be made. First, the diseases may be rare, and therefore one may be limited overall by the number of

subjects, probands and family members affected or at risk, for study. Second, risks may not be clearly known, and there may not be standard interventions (or any interventions). We rely on investigators to provide appropriate background information to support the approach they wish to take in providing genetic information to subjects. This will be dependent upon the type and certainty of the information obtained by the research, and status of medical care for the condition under study. Some subjects will not wish to know they have a "disease susceptibility gene" when no clear treatments are available, for example, and that wish should be respected. There may be some studies where immediately applicable genetic information may be obtained, and it would be unethical to withhold such information from subjects (similar to finding a pulmonary nodule on a screening chest x-ray). If standard clinical information, which indicates that standard clinical activities be undertaken, it is usually advised that return of this information to subjects and doctors be discussed in consent forms, or that subjects at least be given the option to obtain this information, in a medically appropriate setting with the appropriate counseling or referrals. Subjects should know in what time frame and how this type of information will come back to them and their physicians. Information must not be provided in a vacuum - usually a genetic counselor, a physician investigator, and/or a personal physician is involved. It is the responsibility of the investigators to ensure that subjects and their physicians receive enough information to know what to do next, particularly if the data is outside the average clinician's knowledge base. That said, investigators seem to more frequently commit "errors of optimism" in offering to return genetic information to subjects when it is highly unlikely that clinically relevant, validated information will be obtained. Most genotype data currently collected on polygenic disorders is still being correlated with clinical phenotypes, penetrance, and treatments, and will **not** be clinically relevant, and is best not returned to subjects or treating physicians. Do not state that you will return research data to subjects if it's not yet "ready for prime time" clinically. In such instances it is possible to tell subjects that if clinically relevant tests or treatments are developed they would be able to access these treatments and tests through their own physicians in the future. Truly standard clinical tests to be used for clinical decision-making should be performed in CLIA-certified laboratories, or repeated in such laboratories. The IRB appreciates that some clinically useful genetic information may be obtained by research laboratories which are not CLIA-certified.

11. Re-contact of subjects for more sample or other studies. It is often useful to include an option for subjects to decide if they may be re-contacted for additional samples or research studies, if you believe this will be useful. This makes life easier for all and ensures that only subjects willing to be re-contacted are approached. See above under point 8 for more on "re-consenting" issues.

12. Use of specimens by commercial entities. A standard "use of specimens" clause relating to commercial applications such as development of diagnostic testing, or treatments, is usually relevant. This statement should respectfully clarify that investigators and the hospitals may benefit if applications are commercialized; subjects will not share in any financial gains or royalties derived from such products developed based upon their genetic information. Typically, such advances are the result of study of many samples, and not one specific individual's sample. It is recommended that the standard, IRB-approved statement be used for this purpose if at all possible. Industrial sponsors may provide template language with less tactful statements which may be offensive to some subjects. It is preferred that you use the term "providing samples" rather than "donating samples." Subjects may be reassured by a statement that they are not giving up any of their legal rights by signing a consent form, but that is not a required statement.

You are referred to the extensive guidelines for genetic research at the Partner's HRC website (guidance documents): http://healthcare.partners.org/phsirb/guidance.htm and may also contact Dr. Hohmann with specific questions by email: ehohmann@partners.org