



**Association of
American Medical Colleges**
2450 N Street, N.W., Washington, D.C. 20037-1127
T 202 828 0460 F 202 862 6161
www.aamc.org

Darrell G. Kirch, M.D.
President

August 2, 2006

Reed V. Tuckson, M.D.
Chair, Secretary's Advisory Committee
on Genetics, Health, and Society
National Institutes of Health
Bethesda, MD
c/o Amita Mehrotra
mehrotraa@od.nih.gov

Re: Request for Public Comment on "Policy Issues Associated with Undertaking a Large U.S. Population Cohort Study on Genes, Environment, and Disease." 71 FR 34146

Dear Dr. Tuckson:

It is my pleasure to comment on behalf of the Association of American Medical Colleges on the draft report of the Secretary's Advisory Committee on Genetics, Health, and Society (SACGHS), as announced in the *Federal Register* on June 13. The AAMC is a nonprofit organization representing all 125 accredited U.S. medical schools, nearly 400 major teaching hospitals, and 94 academic and scientific societies, representing 109,000 faculty members.

The objective of the report is to identify the extensive policy considerations attendant to a study of the genetic and environmental associations with health and disease in a large cohort of the U.S. population, and to identify steps to be taken and expertise to be consulted for resolving these considerations and determining whether the U.S. Government should proceed with such an ambitious project. The committee and staff should be commended for their understanding and sensitivity reflected in this comprehensive draft, and its consistent themes of inclusion, consultation, and cooperation among the many prospective stakeholders and participants. The AAMC has several comments for revising this draft prior to submission to the Secretary.

First, in the section beginning line 1139 and elsewhere, the draft correctly underscores the need to consider the effects of the proposed study's estimated cost (perhaps \$3 billion over ten years, line 1160) on other areas of science, including other biomedical and public health research, and particularly on hypothesis-driven, investigator-initiated science typically funded through the R01 mechanism. In any decade, a proposal of this

type and magnitude raises legitimate questions about the role of large, directed research projects (“big science”) in agencies like NIH with a diverse portfolio dedicated primarily to investigator-initiated research. As noted in the document, the Human Genome Project demonstrated that the NIH’s well-established and respected system for assessing research and health priorities can accommodate diverse project modalities as long as they are expertly designed and managed.

The draft further notes:

Some members of the scientific community have expressed concern about the impact that such a large allocation could have during flat funding periods and argue that a large project should be undertaken only if funded through sources that do not compromise investigator-initiated projects (1163-6).

This statement by some members of the scientific community does not fully capture the extent of the apprehension in academic research about major new initiatives, nor is it realistic in imposing the condition that (presumably all) “investigator-initiated projects” must be protected from likely cost impacts. While the current period is one of flat “nominal” funding, the NIH budget is declining in inflation-adjusted terms at about 3% a year. In fact, there are severe funding constraints across all Public Health Service (PHS) programs and all Federal discretionary programs with the exception of Defense and Homeland Security. The fiscal predicament follows a period in which, concomitant with the rapid expansion of the NIH budget, academic institutions invested substantial funds to hire new faculty and staff, finance renovation and construction of facilities, and make other long-term commitments, consistent with the national goal to double the nation’s capacity for biomedical research. These institutions now endeavor to sustain themselves in substantial part from an NIH budget declining in purchasing power and with less money for new projects. The strain is especially severe for young investigators beginning their careers, and could well result in the loss of a generation that would be antithetical to the nation’s policy goals.

The AAMC recommends that the final report make clear that the resource implications for the proposed study may be far more acute in the current and foreseeable funding environment than currently stated. The report should also note the need for development of specific milestones and other metrics of performance so that the Department of Health and Human Services and the research community can determine that resources are being effectively used and that the study’s prospective benefits are realized.

Other issues

The draft provides that the Secretary may wish to establish “a highly collaborative model of project leadership and management” across multiple HHS agencies. Models for such management ought to be identified if they exist. Moreover, highly collaborative

management can be dysfunctional unless clear lines of authority, responsibility, and accountability are established by the leadership.

The report correctly indicates that the feasibility of the cohort study could hinge on intellectual property and data access arrangements. We believe that recommendation 7 (beginning line 1283) is inadequate without the following revision (in italics):

To ensure that the public benefits from such discoveries, the Secretary should require that there be clear intellectual property policies, *consistent with established NIH and PHS policies, norms, and guidelines on intellectual property, and sharing of data and research resources*, in place for discoveries made using the data and samples collected through the project.

The proposed cohort study, because of its profile and significance, could well influence standard setting or threshold requirements for IP and access rights in other such studies, much as the International Human Genome Project's Bermuda Rules have become a model for follow-on genomic and proteomic research.¹ IP policies for the cohort study should not be constructed from a blank slate, but should build upon the significant policies already established at the NIH.

The draft report (beginning line 1848) also notes that there is currently no comprehensive legislation protecting Americans from genetic discrimination in the workplace. For the government to move ahead with a highly publicized, wide-ranging data collection effort without first legislating comprehensive protection against employer or health insurer misuse of information could undermine public confidence in this research and be a further barrier to recruitment. SACGHS should restate its prior support for H.R.1227, "anti-genetic discrimination" legislation now in the House of Representatives, which passed in the Senate and which the AAMC strongly supports.

Finally, in the overview of a "hypothetical large U.S. population cohort study" (beginning line 752), the document considers that the population would be recruited primarily door-to-door over a four-year period (line 784). This appears to be premature, as multiple approaches could be incorporated to augment door-to-door recruitment. The document also notes, as the final characteristic of the hypothetical study, that:

Informatics and data management needs would include 1) data capture, entry, and editing; 2) database design and management; and 3) analysis. (lines 795-6)

This sole description of the informatics infrastructure necessary to support the cohort study is woefully insufficient in addressing the policy issues that support of such infrastructure would require. These points in particular were included among the many

¹ See for example, Recommendation 2 in the report of the National Research Council, *Reaping the Benefits of Genomic and Proteomic Research: Intellectual Property Rights, Innovation, and Public Health* (Washington, DC: National Academies Press, 2006), 138.

Reed V. Tuckson, M.D.

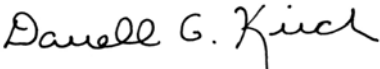
August 2, 2006

Page 4

thoughtful comments shared with us by colleagues at Vanderbilt University Medical Center, which are enclosed here as an attachment, and which AAMC endorses.

In summary, the AAMC hopes that this draft document and related deliberations remind the nation of the importance of population-based research to complete the promise of the Human Genome Project to improve public health. Many of the issues identified in the document underscore barriers to such research in general, and we are grateful to see them brought to the Secretary's attention.

Sincerely,

A handwritten signature in cursive script that reads "Darrell G. Kirch". The signature is written in black ink and is positioned above the printed name.

Darrell G. Kirch, M.D.

ATTACHMENT

Regarding: Comment on “Policy Issues Associated with Undertaking a Large U.S. Population Cohort Project on Genes, Environment, and Disease”

Vanderbilt Medical Center is pleased to provide comment on this important topic. Please see our specific comments below, along with the Line number where the issue is referenced.

Lines 236-237: ...the “Secretary may wish to establish a highly collaborative model of project leadership and management in multiple HHS agencies...”

Comment: We agree that the model should be highly collaborative. However, given that we are unaware of a successful current example of such a unique collaboration in a healthcare context, we would propose that an evaluation of the feasibility of the governance proposal (legality, practicality, and accountability) be a critical step. In addition, given the considerable uncertainties (management structure, and otherwise), it is important that any undertaking of this nature be held to prospectively defined milestones and include systematic go/no-go decisions.

Line 401: “Currently, a number of countries have begun such national population research projects.”

Comment: Fortunately, the fact that other programs exist throughout the world might provide a baseline with which to make estimations (although, of course, some of the parameters would not be applicable or relevant). We propose that the program consideration warrants a review of the cost-effectiveness of various external programs. It will be critical to develop metrics for success for any national program in the U.S. which take into account the investment, and explores a relationship between cost and impact (both the timing and quality of incremental research outcomes).

Lines 512-513: “Are there obstacles that would make the undertaking of such a project especially difficult in the United States compared to other countries?”

Comment: Perhaps this statement is unnecessary (since the answer is undoubtedly ‘yes’ e.g. lack of a national healthcare system, lack of uniform medical records, political environment, etc.) and the issue at hand seems, instead, to be “Do the obstacles inherent in the United States make the proposed program infeasible from a cost or other perspective?”

Lines 542-544: “...it requires a relatively large investment of public resources and, as such, warrants scrutiny of and deliberation about its relative value to science, society, and the country; and...”

Comment: The issues surrounding this are well characterized in the summary, however impossible to address specifically without knowing *how large* of an investment. Ideally, project leaders need to model and estimate under which scenarios the program might provide sufficient return on investment. In addition to understanding the absolute dollar amount, an added policy consideration is the context of the expenditure. For example, if

the project costs \$1 billion and is expected to have a direct impact on healthcare costs, such as drastically reducing adverse drug reactions (estimated at over \$130 billion annually) it could be a tolerable investment. Therefore, a follow-up policy consideration should be: “Will the program proposal include a plan for solicited requests for utilization of the resource which can directly impact societal healthcare expenses?”

Lines 569-572: “Due to the breadth of such an initiative, an interdisciplinary approach that includes geneticists, epidemiologists, toxicologists, social and behavioral scientists, public health experts, biostatisticians, information technologists, health providers, ethicists, community representatives, and others is needed.”

Comment: While people are defined by their expertise in this context, clearly the potential stakeholders, individually, will not all lead to an equal chance of success. We would propose that the named individuals to be involved should also be a subject of external public deliberation.

Lines 649-651: “The study population also must be large enough to detect simultaneous multiple variables that interact and cause disease, such as gene-environment or gene-gene interactions.”

Comment: In this case, the policy issues do not seem to be appropriately focused, because it is unknown exactly how large of a study population is needed. The challenge to address this would be developing statistical models (making informed assumptions for effect size, disease prevalence, allele frequency, desired power, epistasis/confounding, etc.) which can be used in consideration of the benefits and costs of the program.

Lines 666-668: “In comparison, prospective cohort studies of genes and environment enroll individuals prior to disease onset and prospectively collect environmental and biobehavioral marker data, allowing for the examination of contributing non-genetic and genetic factors in disease.”

Comment: A potential benefit not listed in the summary document is that a prospective cohort model, with no screening for diseases, would allow researchers to find a particular genetic mutation and then examine all phenotypes commonly associated with it, thereby potentially uncovering a greater number of mutations implicated across various, seemingly unrelated, diseases.

Lines 675-677: “But even more challenging is to collect indices of variation in biobehavioral reactivity that might produce cleaner intermediate phenotypic markers of early disease vulnerability.”

Comment: The importance that more and better characterized biomarkers can have in the research and drug development process should be elevated in the proposal for this program; we propose that ascertainment of biomarkers that correlate with drug responses (not solely susceptibility to disease) be among the core set of project deliverables to be deliberated.

Lines 747-748: “Because they are considered to be research tools and research infrastructure, they have primarily been funded by governments and non-profit organizations...”

Comment: We note that simply being funded by government does not require the project to be executed by government. For example, there are well-functioning, collaborative networks of academic medical centers on a number of subjects, coordinated by a highly involved program officer (for example, at an NIH center or institute); this is a concept that could be leveraged.

Line 784: “The project population would be recruited primarily door-to-door over a four-year period.”

Comment: The mechanisms of recruitment should be modeled against 1) enrollment of representative populations, or populations unable to be reached in other ways, 2) overall costs, and 3) effectiveness and efficiency. As an example, the next U.S. census (which also includes a door-to-door component along with considerable databasing of collected information) is expected to cost \$9.3 billion according to a GAO analysis of U.S. Census Bureau figures. The program leaders should consider employing a variety of recruitment methods that optimize the composition of the resulting population, as well as the costs associated with enrollment, instead of selecting a single method. In addition, an exclusively door-to-door method seems a bit arduous, since unlike the census, the major focus in a large scale population is not comprehensively determining where everybody lives.

Lines 790-792: “Data collection at entry would include the widest breadth of phenotypes and environmental factors needed to predict outcomes, balanced by cost and participant burden.”

Comment: We note that some participants will be able to contribute a considerable amount of time, while others will not be able to contribute as much. One policy consideration would be to consider a dynamic model that accommodates variability in time availability of individual participants.

Lines 795-796: “Informatics and data management needs would include 1) data capture, entry, and editing; 2) database design and management; and 3) analysis.”

Comment: Perhaps the biggest underestimation in the entire document, this area should be further examined and deliberated. We note that it is one major source of failure (or cause of limited success) of biobanks internationally.

Line 807: “...methods of public engagement will need to be considered...”

Comment: We suggest that a conjoint or ‘trade-off’ analysis be considered. That is, one quantitative method of measuring public reaction should include a study of the subsets of possible combinations of project components (commercialization assumptions, privacy components, time and inconvenience requirements, etc.) so that the relative order of *importance* each component has in the respondent’s decision (to participate in, or to approve of taxpayer funding for, etc.) can be determined.

Line 836: “...the significance of research findings resulting from analyses of the data/specimens.”

Comment: It should not be assumed that research findings will necessarily result from the resource (for example, it may require research to be stimulated by an accompanying set of solicitations).

Lines 853-854: “...for example, through representative democracy and the advisory committee system used by several funding agencies to assist in setting priorities and making funding decisions.”

Comment: One critical issue would be public input into the diseases studied on the database. Since we know the nature of the particular study can impact willingness to enroll, it also might impact support for the project. Prospectively collecting public input on disease areas to be studied in the context of this resource might be considered as one element of public engagement.

Lines 862-863: “Slightly more than half of the American public has a favorable attitude toward using genetic information to personalize and optimize health.”

Comment: In this context, we would propose that general favorability is less important than a) willingness to participate themselves, or b) whether they would commit tax dollars to fund the program.

Lines 866-869: “The survey data also suggest that even those who hold a favorable attitude toward using genetic information for health purposes have significant concerns about the privacy of their genetic information, including its storage for research purposes, and about the potential for its misuse.”

Comment: The assumptions about the implementation of the genetic anti-discrimination laws (which passed unanimously in the Senate but has stalled in committee in the House of Representatives) need to be clear in all public research efforts, as they certainly will impact response.

Lines 870-871: “In one recent survey, only one-quarter to one-third of the public agreed that the government should create a national database of DNA information to advance health.”

Comment: These are daunting data which perhaps suggest an insurmountable flaw in the design of the program, provided they are statistically representative. We would suggest that the program officials make these data publicly available for external peer review. Note, these data are especially critical because they directly impact the costs per patient enrolled – if contact with 5 people is required to obtain a single accepted enrollment, the costs are likely too onerous to manage.

Lines 954-956: “For example, does the United States have a unique genetic population or environment that is not found in other countries that are currently sponsoring studies?”

Comment: This seems like a reasonable assumption as countries in aggregate are certainly different from each other (different demographics, macroeconomic factors, healthcare systems, diets, environments, etc.). But perhaps the issue could be refined to

be: “Are there comparable subsets of populations resembling those in the United States anywhere else among existing (and accessible) biobanks?”

Line 957: “What is its value and cost?”

Comment: The value is unknown and might continue to be years into implementation: however, we suggest the costs need to be modeled before any further steps are taken.

Lines 959-961: “...will it be possible to sustain public and political support for such an investment, especially since such support will need to be reaffirmed annually as part of the federal budget process?”

Comment: This is an important consideration. We suggest that the policy discussion should include extensive use of deliverables and milestones. The procurement of funding should follow from meeting or not meeting these milestones and deliverables.

Line 962: “What would be the effect of funding the project on other areas of research or programs?”

Comment: The assumptions underlying this issue are of utmost importance to many individuals involved, either directly or indirectly. The effects must be reasonably known as the program is deliberated.

Line 966: “What should be the role of the private sector?”

Comment: We would like to add: “What would be the role of academic medical centers, which are notably the most trusted among various entities conducting human subject research?”

Lines 989-991: “Is the large cohort approach the only or most effective way to advance understanding of the interactions among genetics, environments, behaviors, their interactions, and common disease?”

Comment: A large cohort is not necessarily the same thing as a large, government-sponsored and managed cohort; therefore, we propose separating the two concepts in policy proposals.

Lines 1107-1109: “Some respondents noted that it may be wise to use existing cohorts for the purposes of the larger project in order to refine and pilot the processes of standardizing exposure, phenotype determination, and subject recruitment methods.”

Comment: A more thorough summation of the existing programs (both current data and resources as well as those that could be readily enhanced) should be included in the program overview proposal as it will aid in deliberations.

Lines 1171-1173: “Others have asked, because of the high costs of such a project, whether it is appropriate for the federal government to be the sole sponsor. Does the government have the necessary infrastructure to carry out this type of effort, or should it rely on the private investment of funds and technical resources to complete some of the work?”

Comment: We would propose that there should be a consideration of soliciting competitive bids for the project.

Lines 1298-1299: “Will the lack of uniform methods for collecting, storing, and centralizing genetic, behavioral, social, physiological, and clinical health information make a project of this scale difficult to implement?”

Comment: In any circumstance the project of this scale would be difficult to implement. We would suggest focusing this more on the feasibility of finding solutions to the challenges of managing data collected in a non-standard format.

Lines 1355-1357: “There are several issues to consider in ensuring that the subject population is representative of the entire population, not just along gender, racial, or ethnic lines, but also along socioeconomic strata.”

Comment: We suggest adding to the list of issues to consider in the recruitment of participants: “health status”. People with certain conditions may be more (or less) likely to enroll, so the impact of selection bias on representativeness needs to be considered.

Line 1391: “...requiring the development of new technologies.”

Comment: Related to collecting reliable environmental exposure data, we would propose making this a deliverable that determines whether certain outcomes can be measured. The GEI initiative may or may not be successful in generating the development of these sensors.

Lines 1409-1411: “The collection and dissemination of personal medical information in a large population project will be complicated in the United States, given that currently no universal electronic system for storing medical records exists.”

Comment: Given that there is no universal electronic system for storing medical records, it will be extremely difficult to collate and consolidate existing health information on participants. All data would then be prospectively collected purely for the sake of research and not derived from clinical care, which adds to the cost.

Line 1420: “...benefit returned to the participants...”

Comment: This statement does not seem to be correct, since it is not clear that there will be a benefit to individual participants.

Lines 1432-3: “...issues to be considered in approaching, educating, and enrolling...”

Comment: Surely among the most pressing issues felt by participants will be the potential for discrimination. The status of the anti-discrimination bill is of utmost importance and that should be resolved before this takes shape. It is paramount since the proposed entity potentially managing this program also manages federal health insurance programs such as Medicare and Medicaid. Individuals among these populations might disproportionately question whether the resource will be used to deny their insurance.

Line 1453: “...will the project provide healthcare to uninsured patients?”

Comment: This must be further clarified before it can become a policy issue.

Line 1468: “...will study results emanating from the project be returned to participants?”

Comment: The policy issue should also explore how they should be returned.

Line 1526: ‘...may not always be known at the time of collection ...’

Comment: It would be more accurate to say the research aims of a particular project would likely *never* be completely known at the time of collection.

Lines 1590-1591: “if DNA sequence information from these identified samples is matched to DNA sequence data in the database.”

Comment: The program needs to consider in particular individuals who have information and samples stored in nationwide databases such as CODIS and the military DNA ID system.

Lines 1719-1720: “Will such a project widen the gap between what can be diagnosed (or predicted) and what can be treated (or prevented)?”

Comment: Part of the policy discussion on this topic should include a proposal for a targeted, unified series of successive funding opportunities to propel translation of the research findings (to avoid widening this gap).