Development of a Classification Methodology for Genetic Tests:

Conclusions and Recommendations of the Secretary's Advisory Committee on Genetic Testing

September 2001

Secretary's Advisory Committee on Genetic Testing
National Institutes of Health
6705 Rockledge Drive, Suite 750
Bethesda, Maryland 20892
http://www4.od.nih.gov/oba/sacgt.htm

TABLE OF CONTENTS

SACGT Roster …..……..…….i

Introduction ……..……1

Review of SACGT's Initial Oversight Recommendations from July 2000 ……...…..……1

Test Classification Methodology – First Proposal ……...……2

Concerns and Second Thoughts Emerge about First Proposal ……...…….4

Test Classification Methodology – Revised Proposal ..……..…..4

Public Comments and a Reconsideration of the Classification Concept ……...……..6

Conclusion ……...……....7

Appendices

Appendix A: Request for Public Comment on a Proposed Classification Methodology for Determining Level of Review for Genetic Tests (Federal Register, Vol. 65, No. 236, December 7, 2001)

Appendix B: Summary of Public Comments on SACGT's Revised Proposal for a Methodology to Classify Genetic Tests

Secretary's Advisory Committee on Genetic Testing
ROSTER

Chair

Edward R.B. McCabe, M.D., Ph.D.
Professor and Executive Chair
Department of Pediatrics
University of California, Los Angeles
Physician-in-Chief
Mattel Children's Hospital
10833 Le Conte Avenue, 22-412 MDCC
Los Angeles, CA 90095

Members

Patricia A. Barr
407 Cold Spring Road
North Bennington, VT 05257

Kate C. Beardsley, J.D.
Partner
Buc & Beardsley
919 Eighteenth Street, N.W.
Washington, DC 20006

Ann Happ Boldt, M.S.
Certified Genetic Counselor
Department of Medical and Molecular Genetics
Indiana University School of Medicine
975 West Walnut Street
Indianapolis, IN 46202

Joann Boughman, Ph.D.
Executive Vice President
American Society of Human Genetics
9650 Rockville Pike
Bethesda, MD 20814

Wylie Burke, M.D., Ph.D.
Chair
Department of Medical History and Ethics
University of Washington - Box 357120
1959 NE Pacific, Room A204
Seattle, WA 98195

Patricia Charache, M.D.
Professor, Pathology, Medicine, and Oncology
Program Director
Quality Assurance and Outcomes Assessment
Department of Pathology
Johns Hopkins Medical Institutions
600 North Wolfe Street, Carnegie 469
Baltimore, MD 21287

Mary E. Davidson, M.S.W.
Executive Director
Genetic Alliance
4301 Connecticut Avenue, NW, Suite 404
Washington, DC 20008

Elliott D. Hillback, Jr.
Senior Vice President
Corporate Affairs
Genzyme Corporation
One Kendall Square
Cambridge, MA 02139

Barbara A. Koenig, Ph.D.
Executive Director
Stanford Center for Biomedical Ethics
Stanford University
701A Welch Road, Suite 1105
Palo Alto, CA 94304

Judith A. Lewis, Ph.D., R.N.
Associate Professor
Maternal Child Nursing
Director of Information Technology
School of Nursing
Virginia Commonwealth University
1220 East Broad Street
Richmond, VA 23298

Victor B. Penchaszadeh, M.D., M.S.PH.
Professor of Pediatrics
Albert Einstein College of Medicine
Chief, Division of Medical Genetics
Department of Pediatrics
Beth Israel Medical Center
First Avenue at 16th Street
New York, NY 10003

Reed V. Tuckson, M.D.
Senior Vice President
Consumer Health and Medical Care Advancement
UnitedHealth Group
MN 008-T910
9900 Bren Road East
Minnetonka, MN 55343

Ex Officio Members

Agency for Healthcare Research and Quality

John M. Eisenberg, M.D.
Administrator

Alternate:
David Lanier, M.D.
Deputy Director
Center for Primary Care Research

Centers for Disease Control and Prevention

Jeffrey P. Koplan, M.D., M.P.H.
Director

Alternate:
Muin Khoury, M.D., Ph.D.
Director
Office of Genetics and Disease Prevention

Food and Drug Administration

Bernard Schwetz, D.V.M., Ph.D.
Acting Principal Deputy

Alternate:
David W. Feigal, Jr., M.D., M.P.H.
Director
Center for Devices and Radiological Health.

Centers for Medicare and Medicaid Services

Thomas A. Scully
Administrator

Alternate:
Jeffrey L. Kang, M.D., M.P.H.
Director
Office of Clinical Standards and Quality

Health Resources and Services Administration

Elizabeth James Duke, Ph.D.
Acting Administrator

Alternate:
Michele Lloyd-Puryear, M.D., Ph.D.
Chief, Genetic Services Branch
Maternal and Child Health Bureau

National Institutes of Health

Ruth Kirschstein, M.D.
Principal Deputy Director

Alternate:
Francis Collins, M.D., Ph.D.
Director, National Human Genome Research Institute

Office of Human Research Protections

E. Greg Koski, Ph.D., M.D.
Director

Alternate:
Susan Zullo, Ph.D.
Senior Advisor to the Director

Executive Secretary

Sarah Carr
Executive Secretary
Secretary's Advisory Committee on Genetic Testing
Office of Recombinant DNA Activities
Office of Science Policy
National Institutes of Health
6705 Rockledge Drive, Suite 750
Bethesda, Maryland 20892

Development of a Classification Methodology for Genetic Tests: Conclusions and Recommendations of SACGT

Introduction

In July 2000, the Secretary's Advisory Committee on Genetic Testing (SACGT) submitted a report to the Secretary of Health and Human Services outlining the Committee's findings and recommendations regarding the oversight of genetic tests in the United States. The report--*Enhancing the Oversight of Genetic Tests: Recommendations of the Secretary's Advisory Committee on Genetic Testing*--responded to a specific request from the Assistant Secretary for Health to address five oversight questions. One of the questions related to whether an algorithm or methodology could be developed to classify tests for oversight purposes according to the level of scrutiny warranted by a test. In the July report, SACGT outlined criteria on which a methodology might be based and indicated that further study and analysis were needed to complete a fully developed methodology for classifying tests. This report describes SACGT's subsequent efforts to develop a classification methodology for genetic tests and explains why, in the final analysis, SACGT came to question the feasibility and utility of such a methodology.

Review of SACGT's Initial Oversight Recommendations from July 2000

In its July 2000 oversight report, SACGT reviewed the adequacy of current oversight of genetic tests and concluded that the level was inadequate for ensuring the safety, accuracy, and clinical validity of genetic tests. This conclusion was based on the rapidly evolving nature of genetic tests, their anticipated widespread use, and concerns expressed by the public about their potential for misuse or misinterpretation. SACGT offered 26 recommendations to enhance the current system of oversight for genetic testing. One key recommendation was that the Food and Drug Administration (FDA) should be involved in the review of all new genetic tests regardless of how they are formulated and provided (i.e., a kit versus laboratory service). SACGT further recommended that, given the growing number of genetic tests and the speed with which they are being developed, the agency's review process should be innovative and flexible in order to minimize the time and cost of review without compromising the quality of the assessment of test validity. To implement this expanded oversight of genetic tests, SACGT recommended that FDA be provided with sufficient resources and consider employing the deemed status mechanism used in other regulatory programs such as the Clinical Laboratory Improvement Amendments.

The question of whether a mechanism could be developed to assign tests to different categories was a critical part of the oversight deliberations. SACGT initially considered a classification scheme to be "an essential initial step in the process of test evaluation." In the oversight report, the Committee stated the following:

> Determining the level of review required of a particular genetic test will be crucial to ensuring that a test receives the appropriate level of review based on the characteristics of the test and its target disease or condition, the intended use of the test, and the potential for improved medical outcome. Because further work is needed to develop the criteria and the methodology to be used in classifying tests by the level of scrutiny required, a SACGT working group, augmented by representatives of relevant federal agencies, professional organizations, and the public and private sectors, will immediately begin to develop a proposed algorithm for the classification of genetic tests . . . It is recommended that these criteria and methodology be used in the classification of genetic tests.

Test Classification Methodology—First Proposal

In August 2000, SACGT convened a Working Group on Genetic Test Classification to assist in the development of a framework for classifying genetic tests. The Working Group was chaired by SACGT member Dr. Wylie Burke and composed of SACGT members, SACGT *ex officio* members, and ad hoc experts representing relevant professional and private sector organizations. The Working Group met on August 3, 2000 with the goal of producing a test classification schema for assessing the level of review warranted by a genetic test.

The Working Group developed a classification methodology based on four criteria that would designate genetic tests into one of two levels of review, Scrutiny Level I (SL I) and Scrutiny Level II (SL II). The four criteria were test volume; whether the test would be used for population-based testing; whether the test is diagnostic or predictive; and a set of three questions related to the availability of an intervention, the predictive value of the test, and potential for medical or social harms associated with the test. A diagram displaying how the criteria would be applied to determine the recommended level of review is shown in Figure 1.

The Working Group selected *test volume* as a pragmatic first criterion, based on the public health strategy of focusing resources on matters likely to affect the greatest number of people. This choice reflected the difficulties in data collection for rare diseases and the limited financial incentives available to cover research and development costs. However, this was not to suggest that tests intended for a small market would not warrant scrutiny but rather that they may *a priori* warrant SL I review. Furthermore, low volume tests that were found to raise significant medical or social concerns during the initial review could be elevated to SL II review. SACGT recognized that further discussion of test volume was needed to determine the threshold of high versus low volume.

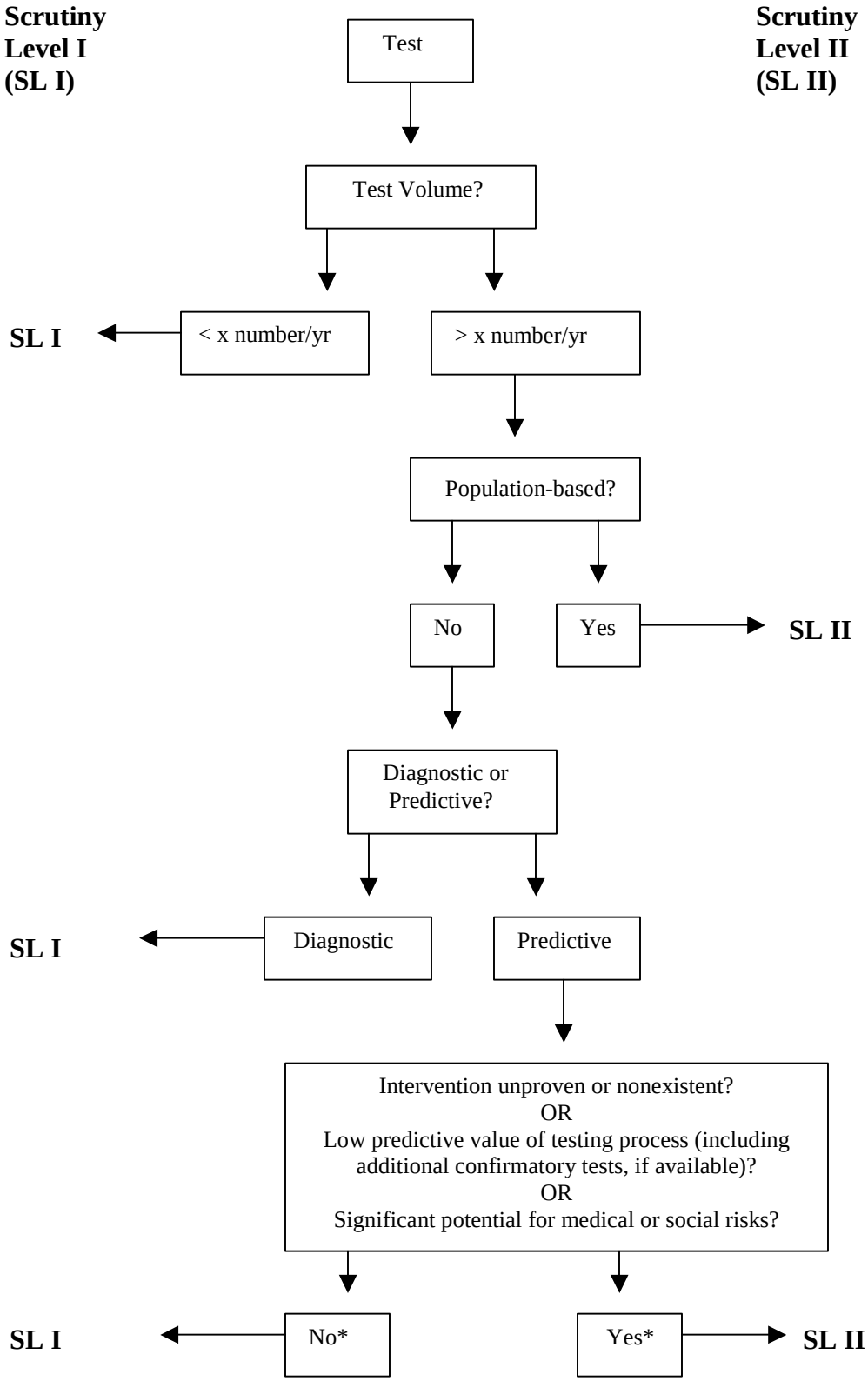
The second criterion was whether the test is to be used for *population-based screening*. Population-based screening is testing of groups or populations of currently healthy people rather than individuals or families. The Working Group recommended that tests to be used for population-based screening should undergo SL II because of the greater number of people who would be exposed to the potential risks of the tests.

If a high-volume test is not used on a population basis, the next consideration was to determine whether the test was to be used for *predictive or diagnostic* purposes. If a test is intended to identify or confirm the diagnosis of an affected individual, the Working Group suggested it receive an SL I review. If the test is intended to be used predictively to determine the probability that a healthy individual with or without a family history of a certain disease will develop that disease, the Working Group suggested that the test may warrant a higher level of scrutiny because of their greater potential for medical and social harms.

The Working Group suggested that three additional criteria needed to be applied to determine whether or not a predictive test would undergo SL I or SL II: 1) are interventions unproven or nonexistent for the condition being tested; 2) is the predictive value of the test low and are additional confirmatory tests not available; and 3) is there a significant potential for medical or social risks. If any one of the questions were answered affirmatively, the test would fall into an SL II review category.

On August 4, 2000, Dr. Burke presented the Working Group proposal to SACGT. After discussion and deliberation, SACGT concurred with the proposed methodology and agreed to prepare an addendum to the July 2000 oversight report recommending that the proposed methodology be considered a framework upon which FDA could build as it proceeded with the development of a program of review of genetic tests.

Figure 1. Test Classification Schema Formulated by SACGT on August 4, 2000



* An answer of "Yes" to any of these questions moves the evaluation to Scrutiny Level II. For the test to receive Scrutiny Level I, the answer must be "No" to all three questions.

Concerns and Second Thoughts Emerge about First Proposal

Following the August SACGT meeting, further discussion of the proposed methodology occurred in other venues, including a meeting of the Centers for Disease Control and Prevention (CDC) Genetic Laboratory Forum, a group convened by CDC's Division of Laboratory Systems to review the proposed model and to pilot-test it using several genetic tests. The pilot tests revealed a number of problems with the proposed methodology. For example, the criterion of test volume was seen as problematic since some low volume tests, which according to the proposed methodology would warrant SL I, might have heightened ethical and social implications that would warrant an SL II review. In addition, some tests may begin as low volume, but become high volume as test use widens, raising the issue of whether the test would need to be reviewed again. Other concerns raised by Forum participants related to the definition of population screening and the risks of off-label use, i.e., a diagnostic test is used for predictive or population-based purposes.

Test Classification Methodology—Revised Proposal

In response to such concerns, SACGT decided to reconsider the proposed classification methodology at its November 2-3, 2000 meeting. SACGT agreed that determining an appropriate threshold for low volume might not be possible and that other criterion, such as test purpose and social risk, might at times be more significant than volume. In addition, the reliability and validity of a test was important regardless of whether it was a low volume or high volume test.

Another concern related to tests for rare diseases. Members acknowledged that rare disease tests can have many purposes and uses but also raise social risks. For instance, newborn screening of all infants for phenylketonuria, a rare disorder, may necessitate greater scrutiny as a high-volume, population- based test. Overall, it was agreed that a better definition of rare disease and further discussion and delineation of social risks were needed.

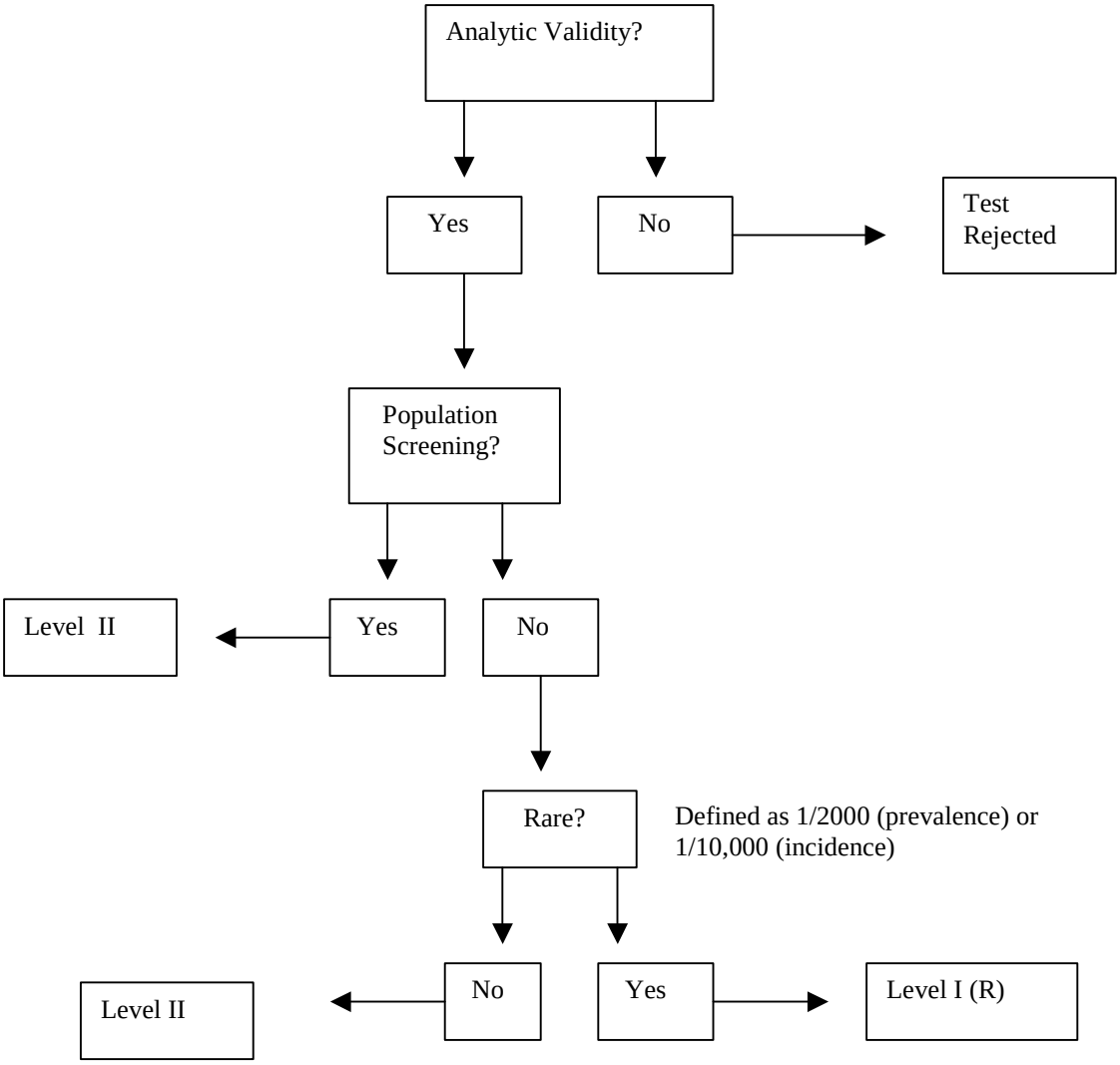
Questions also surfaced about the classification of prenatal tests. The proposed classification methodology implicitly categorized prenatal tests as diagnostic tests, thereby warranting SL I review. Members pointed out that, given the possible consequences of false results, the associated social issues, and the need for a high predictive value, prenatal tests may actually warrant SL II review.

The issue of off-label use was also discussed. Since FDA review would consider only the claim of intended use declared by the manufacturer for a particular test, the proposed classification methodology could allow a manufacturer to submit a test for SL I review, but use it for purposes that may have warranted SL II. For example, a test approved for a diagnostic claim could be used off-label for predictive or other purposes.

After further discussion and deliberation, SACGT concluded that the proposed methodology should be modified. While maintaining the two levels of review (Scrutiny Level I and Scrutiny Level II) and the criterion of population screening proposed in August, SACGT changed the other criteria for determining the level of review to analytical validity and frequency of disease. A diagram displaying how the revised criteria would be applied to determine level of review is shown in Figure 2.

Analytical validity was added as the first criterion in recognition of its fundamental importance to ensuring the quality and safety of genetic tests. Analytical validity is the ability of a test to measure or detect the analyte it is intended to measure or detect, and if the test were found to lack analytical validity, the test would not be reviewed further. The second criterion, *population screening*, was defined as a test intended for use in a group of individuals (>1000) who might be identified on the basis of a shared characteristic (e.g., ethno-cultural group, geographical location, gender, age, reproductive status,

Figure 2. Revised Test Classification Schema Formulated by SACGT on November 3, 2000



behavior, physical traits, or occupation) and may have a higher disease risk than the general population. Tests used for population screening would receive an SL II review.

The new third criterion, *frequency of disease*, divided tests by whether they tested for a common or rare disease. Disease frequency was adopted as a criterion because of the complexity of common diseases and greater challenge of demonstrating test accuracy and validity. The threshold for a rare disease was defined as either prevalence (less than one in 2,000 individuals) or incidence (less than one in 10,000 individuals). Thus, if a test were not to be used for population screening, the third criterion would be applied. A test to detect a rare disease would generally receive an SL I review, and tests for diseases greater than the prevalence or incidence threshold would receive an SL II review.

Once the revised methodology was agreed upon, SACGT decided to gather additional input from public and professional organizations to assess the merits and feasibility of the revised methodology. A request for public comments was published in the *Federal Register*, Vol. 65, No. 236, December 7, 2000 (attached at Appendix A) and individuals who commented on the recommendations in the oversight report were sent letters requesting comment on the revised methodology. A request for comments was also posted on the SACGT website.

Comments were requested on the rationale and feasibility of the proposed test classification methodology as well as on the following specific questions:

1. Is the number of review levels appropriate? Should there be more than two levels? Should all genetic tests receive the same level of review?
2. Are the criteria of analytical validity, population screening, and frequency of disease appropriate for determining the proper review level? Should other criteria, such as the intended use of a genetic test (e.g., diagnostic, predictive, carrier, prenatal, etc.) or clinical utility, be considered in the classification of tests? If so, how should they be incorporated into the methodology?
3. Are the proposed definitions for population and rare diseases appropriate?
4. SACGT has not proposed a specific threshold or minimum standard for analytical validity. Should a threshold for analytical validity be defined? If so, what should the standards be?
5. What characteristics of a rare disease test would raise the level of review from Level I to Level II?

Public Comments and a Reconsideration of the Classification Concept

At its February 2001 meeting, the Committee reviewed and discussed 34 comments submitted by individuals and organizations, including patient advocacy groups, academic organizations, professional societies, and industry (a summary of the public comments is attached at Appendix B). The comments raised a number of concerns regarding the feasibility of the classification methodology and the extent to which the proposed criteria would succeed in addressing the aspects of genetic testing that warrant a higher level of review.

Regarding the overall classification scheme and level of review, some of the comments suggested that certain types of tests are not consonant with the current classification methodology. Some of the comments proposed alternative classification schemes, while others urged the Committee to return to its initial schema.

Comments were also received on the appropriateness and definitions of the three proposed classification criteria. Regarding the criterion of frequency of disease, a number of comments were supportive of the proposed prevalence and incidence cut-offs between rare and common diseases. However, others pointed out that when prevalence is high but penetrance of a particular gene is low, it would be helpful to know

whether one gene was acting alone or in concert with other factors to cause disease. It is not the fact that a disease is common, but that the mutation may have less than high penetrance that raises concern.

Some commenters suggested other criteria that should be used to classify tests warranting a higher level of review. Suggestions included the following: test sensitivity and specificity; genetic heterogeneity; penetrance; low predictive value; potential for social stigma; predictive tests; tests for behavioral disorders; pharmacogenetic testing; complexity of test; difficulty of test interpretation; burden of disease; pattern of inheritance; late onset disorders; availability of proven treatments or prevention; clinical utility; prenatal testing; disease incidence or progression; availability and strength of confirmatory procedures; and the reliability of clinical corroboration.

Commenters also responded to SACGT's request for comments on what criteria would raise tests for rare diseases from SL I to SL II. Many of the criteria listed above were suggested, as well as others, including testing of healthy individuals, prenatal testing, commercial attractiveness of a test, carrier screening of ill-defined populations, risk of adverse effects, population screening, risky medical interventions, implications for family members, absence of medical intervention, burden of disease, and complexity of tests, including interpretation of results.

Conclusion

After consideration of the public comments and additional discussion, the Committee concluded that fundamental, irresolvable questions had been raised about the feasibility of categorizing tests for oversight purposes based on a limited set of elements in a simple, linear fashion. Thus, the Committee decided that further efforts to develop a classification methodology for genetic tests should be curtailed for the present. SACGT's decision to defer further work was also based on significant progress made by FDA to develop an innovative regulatory process for genetic tests. At its February 2001 meeting, SACGT was briefed by FDA on the agency's plans for pre-market review of genetic tests. They presented a pre-market review template developed through roundtable meetings with professional organizations and the private sector that would incorporate many of the parameters that were raised as SACGT deliberated on the test classification methodology. The Committee asked FDA to continue its work on this template and to evaluate how effective the template approach would be in identifying tests that warrant an increased level of review.

At its May 2001 meeting, SACGT was briefed in more detail about FDA's progress on the review template. SACGT members emphasized that the template would need to recognize features of genetic tests that may be unique, including social implications, such as the potential for stigmatization and the importance of informed consent, and to consider these issues during the review process. The Committee noted that because the template includes the purpose of the test, the condition which it detects, and what information it is intended to produce, it should be sensitive enough to allow the agency or another body to determine whether the test poses risks of social harm. After further discussion, SACGT reaffirmed its earlier decision to forego further development of a classification methodology for genetic tests. The Committee noted that the work of developing a classification methodology would likely be applicable to and useful for other issues that SACGT may address in the future.

Appendix A

Request for Public Comment on a Proposed Classification Methodology for Determining Level of Review for Genetic Tests (*Federal Register*, Vol. 65, No. 236, December 7, 2001)

Appendix B

Summary of Public Comments on SACGT's Revised Proposal for a Methodology to Classify Genetic Tests