National Research Council 1996

Executive Summary[1]

INTRODUCTION

Nearly a decade has passed since DNA typing methods were first used in criminal investigations and trials. Law enforcement agencies have committed substantial resources to the technology; prosecutors, defense counsel, and judges have struggled with the terminology and ideas of molecular biology, genetics, and statistics. In 1992, a broad-ranging report released by the National Research Council attempted to explain the basics of the relevant science and technology, to offer suggestions for improving forensic DNA testing and its use in law enforcement, and to quiet the controversy that had followed the introduction of DNA profiling in court. Yet, the report did not eliminate all controversy. Indeed, in propounding what the committee regarded as a moderate position—the *ceiling principle* and the *interim ceiling principle*—the report itself became the target of criticism from scientists and lawyers on both sides of the debate on DNA evidence in the courts. Moreover, some of the statements in the 1992 report have been misinterpreted or misapplied in the courts.

This committee was formed to update and clarify discussion of the principles of *population genetics* and statistics as they apply to DNA evidence. Thus, this second report is much narrower than the 1992 report. Issues such as confidentiality and security, storage of samples for future use, the desirability and legality of data banks on convicted felons, and international exchange of information are not in our charge. Rather, this report deals mainly with the computation of probabilities used to evaluate the implications of DNA test results that incriminate suspects. It focuses on situations where the *DNA profile* of a suspect (or sometimes a victim) apparently matches that of the *evidence DNA*. (We use the phrase "evidence DNA" to refer to the sample of biological material, such as blood or semen, usually taken from the crime scene or from the victim.) The central question that the report addresses is this: What information can a forensic scientist, population geneticist, or statistician provide to assist a judge or jury in drawing inferences from the finding of a match?

To answer this question, the committee reviewed the scientific literature and the legal cases and commentary on DNA profiling, and it investigated the various criticisms that have been voiced about population data, statistics, and laboratory error. Much has been learned since the last report. The technology for DNA profiling and the methods for estimating frequencies and related statistics have progressed to the point where the reliability and validity of properly collected and analyzed DNA data should not be in doubt. The new recommendations presented here should pave the way to more effective use of DNA evidence.

This report describes both the science behind DNA profiling and the data on the frequency of profiles in human populations, and it recommends procedures for providing various statistics that may be useful in the courtroom. The procedures are based on population genetics and statistics, and they render the ceiling principle and the interim ceiling principle unnecessary.

This executive summary outlines the structure and contents of the full report, and it gives the recommendations together with abbreviated explanations of the reasons behind them. This summary does not constitute a complete exposition, and it is no substitute for a careful reading of the chapters that follow. As the report will reveal, the committee agrees with many recommendations of the 1992 report but disagrees with others. Since the committee has not attempted to review all the statements and recommendations in the 1992 report, the lack of discussion of any statement should not be interpreted as either endorsing or rejecting that statement.

CONTENTS OF THE REPORT

Overview. The report begins with an extended summary of the chapters that make up the full report. This overview describes the essentials of the subject with a minimum of jargon, statistics, and technical details, and it includes a numerical example that illustrates how the procedures that are discussed and recommended would apply in a typical case. The main report offers fuller explanations, details, and justifications.

Chapter 1. The first chapter describes the 1992 NRC report, the changes since that report, the uses and validity of DNA typing, differences between DNA typing in criminal cases and in civil paternity litigation, reasons for the seemingly contradictory probability estimates that different experts sometimes present in court, and the committee's approach to the issue of "population structure."

Chapter 2. The second chapter describes the genetic and molecular basis of DNA typing. It introduces the fundamental concepts of genetics, and it surveys the genetic systems and the technologies used in DNA profiling.

Chapter 3. The third chapter concerns laboratory performance. Although our focus is on the statistics that can be used to characterize the significance or implications of a match between two DNA samples, these statistics do not float in a vacuum. They relate to specific claims or hypotheses about the origin of the DNA samples. If DNA from an evidence sample and DNA from a suspect share a profile that has a low frequency in the population, this suggests that the samples came from the same person; the lower the frequency, the stronger the evidence. But the possibility remains that the match is only apparent—that an error has occurred and the profiles differ from what the laboratory has reported. Chapter 3 describes ways that errors can arise and how their occurrence might be minimized. It contains recommendations regarding quality control and assurance, laboratory accreditation, proficiency tests, and confirmatory testing.

Chapter 4. Much of the controversy about the forensic use of DNA has involved population genetics. Chapter 4 explains the generally applicable principles, then considers the implications of the fact that the population of the United States includes different groups and subgroups with different mixes of genes. The chapter develops and illustrates procedures for taking this fact into account in computing *random-match probabilities* for an incriminating DNA profile in a population or a subgroup of a population.

Chapter 5. The fifth chapter considers how the estimated frequency of an incriminating DNA profile relates to conclusions about the source of the DNA in the evidence sample. It discusses how the frequencies are interpreted as probabilities and related quantities, the degree of uncertainty in such estimates, and the type of calculations that might indicate that a profile is unique. It concludes that the abundance of data in different ethnic groups within the major races and the methods outlined in Chapters 4 and 5 imply that the 1992 report's suggested ceiling principle and interim ceiling principle are unnecessary. In addition, it makes recommendations to help assure the accuracy of estimates for what are known as *VNTR profiles* and to handle the special situation in which the suspect was identified as a result of a search through a database of DNA profiles of known offenders.

Chapter 6. The sixth and final chapter discusses the legal implications of the conclusions and recommendations. It describes the most important legal rules that affect the use of DNA evidence, identifies the questions of scientific fact that have been disputed in court, reviews case law on the admissibility of DNA evidence, and explains how the conclusions and recommendations might be used in applying and developing the law. The report makes no recommendations on matters of legal policy, but it does suggest that the formulation of such policy might be assisted by behavioral research into the various ways that DNA test results can be presented in the courtroom.

Appendices. A glossary of scientific terms and a list of the literature cited are provided at the end of the report.

RECOMMENDATIONS

Major conclusions and recommendations are given at the end of the chapter in which the subject is discussed. For convenience, the report also lists them as a group at the end of the overview. This executive summary lists the recommendations only and gives some of the reasoning behind them.

Recommendations to Improve Laboratory Performance

Recommendation 3.1. Laboratories should adhere to high quality standards (such as those defined by TWGDAM and the DNA Advisory Board) and make every effort to be accredited for DNA work (by such organizations as ASCLD-LAB).

Recommendation 3.2. Laboratories should participate regularly in proficiency tests, and the results should be available for court proceedings.

Recommendation 3.3. Whenever feasible, forensic samples should be divided into two or more parts at the earliest practicable stage and the unused parts retained to permit additional tests. The used and saved portions should be stored and handled separately. Any additional tests should be performed independently of the first by personnel not involved in the first test and preferably in a different laboratory.

Comment. The committee offers these recommendations to improve laboratory performance rather than to try to estimate the probability that a particular laboratory makes a mistake by reporting that two DNA profiles match when in fact they do not match. Auditing and proficiency testing cannot be expected to give a meaningful estimate of the probability that a particular laboratory has made such an error in a specific case. An unrealistically large number of proficiency tests would be needed to estimate accurately even an historical error rate. For such reasons, proficiency test results should not be combined with the estimated frequency of an incriminating profile to yield the probability that a laboratory would report that DNA from a person selected at random contains the incriminating profile. No amount of effort and improved technology can reduce the error rate to zero, and the best protection a wrongly implicated, innocent person has is the opportunity for an independent retest.

Recommendations for Estimating Random-Match Probabilities

Recommendation 4.1. In general, the calculation of a profile frequency should be made with the product rule. If the race of the person who left the evidence-sample DNA is known, the database for the person's race should be used; if the race is not known, calculations for all the racial groups to which possible suspects belong should be made. For systems such as VNTRs, in which a heterozygous locus can be mistaken for a homozygous one, if an upper bound on the frequency of the genotype at an apparently homozygous locus (single band) is desired, then twice the allele (bin) frequency, 2p, should be used instead of p^2. For systems in which exact genotypes can be

determined, $p^2 + p(1 - p)Ø$ should be used for the frequency at such a locus instead of p^2. A conservative value of Ø for the US population is 0.01; for some small, isolated populations, a value of 0.03 may be more appropriate. For both kinds of systems, $2p_ip_j$ should be used for heterozygotes.

Comment. The formulas referred to and the terminology used in this recommendation are explained in the overview and in Chapter 4. The product rule, which gives the profile frequency in a population as a product of coefficients and *allele frequencies,* rests on the assumption that a population can be treated as a single, randomly mating unit. When there are partially isolated sub groups in a population, the situation is more complex; then a suitably altered model leads to slightly different estimates of the quantities that are multiplied together in the formula for the frequency of the profile in the population.

In most cases, there is no special reason to think that the source of the evidence DNA is a member of a particular ethnic subgroup within a broad racial category, and the product rule is adequate for estimating the frequency of DNA profiles. For example, if DNA is recovered from semen in a case in which a woman hitchhiker on an interstate highway has been raped by a white man, the product rule with the 2p rule can be used with VNTR data from a sample of whites to estimate the frequency of the profile among white males.[2] If the race of the rapist were in doubt, the product rule could still be used and the results given for data on whites, blacks, Hispanics, and east Asians.

Recommendation 4.2. If the particular subpopulation from which the evidence sample came is known, the allele frequencies for the specific sub-group should be used as described in Recommendation 4.1. If allele frequencies for the subgroup are not available, although data for the full population are, then the calculations should use the population-structure equations 4.10 for each locus, and the resulting values should then be multiplied.

Comment. This recommendation deals with the case in which the person who is the source of the evidence DNA is known to belong to a particular subgroup of a racial category. For example, if the hitchhiker was not on an interstate highway but in the midst of, say, a small village in New England and we had good reason to believe that the rapist was an inhabitant of the village, the product rule could still be used (as described in Recommendation 4.1) if there is a reasonably large database on the villagers.

If specific data on the villagers are lacking, a more complex model could be used to estimate the random-match probability for the incriminating profile on the basis of data on the major population group (whites) that includes the villagers. The equations referred to in the second sentence of Recommendation 4.2 are derived from this model.

Recommendation 4.3. If the person who contributed the evidence sample is from a group or tribe for which no adequate database exists, data from several other groups or tribes thought to be closely related to it should be used. The profile frequency should be calculated as described in Recommendation 4.1 for each group or tribe.

Comment. This recommendation deals with the case in which the person who is the source of the evidence DNA is known to belong to a particular subgroup of a racial category but there are no DNA data on either the subgroup or the population to which the subgroup belongs. It would apply, for example, if a person on an isolated Indian reservation in the Southwest, had been assaulted by a member of the tribe, and there were no data on DNA profiles of the tribe. In that case, the recommendation calls for use of the product rule (as described in Recommendation 4.1) with several other closely related tribes for which adequate databases exist.

Recommendation 4.4. If the possible contributors of the evidence sample include relatives of the suspect, DNA profiles of those relatives should be obtained. If these profiles cannot be obtained, the probability of finding the evidence profile in those relatives should be calculated with Formulae 4.8 or 4.9.

Comment. This recommendation deals with cases in which there is reason to believe that particular relatives of the suspect committed the crime. For example, if the hitchhiker described in the comment to Recommendation 4.2 had accepted a ride in a car containing two brothers and was raped by one of them, but there is doubt as to which one, both should be tested. If one brother cannot be located for testing and the other's DNA matches the evidence DNA, then the probability that a brother of the tested man also would possess the incriminating profile should be computed.

Recommendations on Interpreting the Results of Database Searches, on Binning, and on Establishing the Uniqueness of Profiles

Recommendation 5.1. When the suspect is found by a search of DNA databases, the random-match probability should be multiplied by N, the number of persons in the database.

Comment. Recommendations 4.1-4.3 specify the calculation of the random-match probability for an incriminating DNA profile in a relevant population (or subpopulation). When the defendant has been identified as a suspect from information that is unrelated to the DNA profile, the random-match probability is one statistic that helps to indicate the significance of a match. If the random-match probability is very low, it is unlikely that the samples match just because the defendant, though not the source of the evidence sample, coincidentally happens to share that very rare profile.

But when the defendant has been identified by a search through a large database of DNA profiles rather than by non-DNA evidence, the relevance of the random-match probability is less obvious. There are different ways to take the search process into account. Recommendation 5.1 proposes multiplying the random-match probability (P) by the number of people in the database (N). If the person who left the evidence DNA was not in the database of felons, then the probability that at least one of the profiles in the database would also match the incriminating profile cannot exceed NP.

Recommendation 5.2. If floating bins are used to calculate the random-match probabilities, each bin should coincide with the corresponding match window. If fixed bins are employed, then the fixed bin that has the largest frequency among those overlapped by the match window should be used.

This recommendation applies to the computation of a random-match probability when all or part of the profile involves VNTRs, which are fragments of DNA that are separated in the laboratory according to their lengths. Because the lengths of VNTRs cannot be measured exactly, an *uncertainty window* surrounds each measured VNTR, and two VNTRs are said to *match* when their uncertainty windows overlap. To calculate the frequency of matching VNTR profiles, one must find the proportion of VNTRs that fall within a *match window* around each VNTR in the incriminating profile. *Floating bins* do this exactly, whereas *fixed bins* do this approximately. Although the floating-bin procedure is statistically preferable, certain forms of the fixed-bin procedure usually lead to conservative approximations to the floating bin result.

Recommendation 5.3. Research into the identification and validation of more and better marker systems for forensic analysis should continue with a view to making each profile unique.

Comment. If a sufficient set of DNA characteristics is measured, the resulting DNA profiles can be expected to be unique in all populations. (Only identical twins would share such a profile.) Of course, it is impossible to establish uniqueness by profiling everyone in the world, but theory and experience suggests that this uniqueness is attainable in forensic typing. Indeed, some scientists would argue that the existing panoply of characteristics is already sufficient to permit unique identification in many cases. For example, it has been suggested that a probability much less than the reciprocal of the world population is a good indication of uniqueness. The committee has not attempted to define a specific probability that corresponds to uniqueness, but the report outlines a framework for considering the issue in terms of probabilities, and it urges that research into new and cumulatively more powerful systems continue until a clear consensus emerges that DNA profiles, like dermal fingerprints, are unique.

Recommendation for Research on Juror Comprehension

Recommendation 6.1. Behavioral research should be carried out to identify any conditions that might cause a trier of fact to misinterpret evidence on DNA profiling and to assess how well various ways of presenting expert testimony on DNA can reduce any such misunderstandings.

Comment. Scientifically valid testimony about matching DNA can take many forms. The conceivable alternatives include statements of the *posterior probability* that the defendant is the source of the evidence DNA, qualitative characterizations of this probability, computations of the *likelihood ratio* for the hypothesis that the defendant is the source, qualitative statements of this measure of the strength of the evidence, the currently dominant estimates of profile frequencies or random-match probabilities, and unadorned reports of a match. Courts or legislatures must decide which of these alternatives best meets the needs of the criminal justice system. At present, policymakers must speculate about the ability of jurors to understand the significance of a match as a function of the method of presentation. Solid, empirical research into the extent to which the different methods advance juror understanding is needed.

NOTES

[1]Abbreviations, symbols, and technical terms are defined in the list of abbreviations (p 212) and the glossary (p 214). The underlying concepts are explained in the overview and in appropriate chapters in the body of the report.

[2]The 2p rule involves replacing the quantity p^2 for a *single-banded VNTR locus* with the much larger quantity 2p in the product rule. This substitution accounts for cases in which one VNTR band from a heterozygote is not detected, and the person is mistakenly classified as a homozygote. The substitution also ensures that the estimate of the profile frequency will be larger than an estimate from a more precise formula that accounts for population structure explicitly. The technology for *PCR-based systems,* however, does not have these problems, and the 2p rule is inappropriate for these systems. Therefore, Recommendation 4.1 calls for using $p^2 + p(1-p)Ø$ (rather than 2p) in place of p^2 for such systems.