Seroanalysis and Expert Opinion in Practice

Prof. Dr. K. Hummel, Institute for Blood Group Serology,
POB 880, D-7800 Freiburg (FRG)

In 1980 an opinion on paternity in our Institute was faced with the question of whether a Puerto-Rican with a Spanish name - of whom there was no photograph - was a Spanish, Indio or Negro descent, or a mixture of these. The child's German mother had named the Puerto-Rican as the child's father. The pigmentation of the child's eyes, hair and skin was darker than it is common among, say, German children. Notwithstanding this it was impossible to draw any definite conclusion about the father's race from the child's appearance. This case encouraged the development of an appropriate technique - now known as seroanalysis - to solve such cases and was the first to which it was applied.

In the case concerned I started with two pairs of hypotheses:

I a) null hypothesis: the man is of Spanish descent (= X),
  b) counter hypothesis: the man is of Indio descent (= Y_1):

II a) null hypothesis: the man is of Spanish descent (= X),
  b) counter hypothesis: the man is of Negro descent (= Y_2).

This takes into account the presence on Puerto Rico of pure Spaniards, pure Indios (though seldom) and pure Negroes.

If one multiplies the phenotype frequencies of a person in different genetic systems in Spaniards, Indios and Negroes one obtains the corresponding serotype frequencies among Spaniards, Indios and Negroes. Taking a <u>neutral</u> prior probability, which means assuming that the man is equally likely to be Spanish, Indio or Negro, the <u>Bayes</u> function gives the probability that in blood terms the man is a <u>Spaniard</u>: $W_{Sp}(1) = f(X)/(f(X)+f(Y_1))$.

The probability that in blood terms he is Indio is given by $W_{Ind} = 1-W_{Sp}(1)$.

Taking as alternative hypothesis Negro descent, the probability for Spanish descent is $W_{Sp}(2) = f(X)/(f(X)+f(Y_2))$,
and for Negro descent $W_{Neg} = 1-W_{Sp}(2)$.

If in 100 cases with comparable circumstances one were to assume in each case <u>Indio</u> (and not Spanish) descent, the expectation of error would be $W_{Sp}(1)$, and if one always assumes <u>Negro</u> descent the expectation of error would be $W_{Sp}(2)$. The evaluation of the case in question gave values above 99,99% for $W_{Sp}(1)$ and $W_{Sp}(2)$.

Thus, it was "practically proven" that the man in question was of pure Spanish descent. Therefore, Bayes' Theorem enables one to allocate a person to a specific, uniformly treated racial circle, and this with a defined degree of probability.

Advances in Forensic Haemogenetics 3
Edited by H. F. Polesky and W. R. Mayr

In the case described, Spaniards, Indios and Negroes were treated as "uniform" in terms of blood. With reference to their respective superordinate major races - Caucasian, Amerindian and Negroid - these populations can be treated as separate entities. The major races are not completely homogeneous but are so removed from one another that in relation to one another they are "genetically uniform".

Populations combining different portions of major races may also be regarded as "uniform", provided that the mix has persisted for a very long time - at least many hundreds of years - and has not been disturbed by immigration during the period. Hence, it is possible in seroanalysis to treat Chinese, Japanese, Koreans, Vietnamese, South Sea Islanders, Australian Aborigines and other peoples of the Far East and Southeast Asia as "uniform" in terms of their blood, although they are all mixed in varying degrees of Pacific and Mongoloid races. - In other mixed populations the particular mix has not yet persisted long enough, or has been disturbed by fresh immigration, so it is impossible for one to speak of genuine panmixia. This holds, for instance, for North American Negroes and Latin American populations. Whereas the so treated "uniform" peoples of East and Southeast Asia have each become a "mixed people", the former are still a "people mix": mixed-bloods of every conceivable mixture live alongside pure representatives of the major races.

The greater the differences in serotype frequencies between populations, the greater the degree of certainty with which a person can be allocated to a specific, uniformly treated population. Among Caucasians there is, for example, a not inconsiderable difference between the frequencies of Central Europeans like Germans and those of Indians, Iranians, Afghans, North Africans or Lapps. The accuracy with which a person can be allocated to the mentioned fringe Caucasian populations is correspondingly high. There is far less certainty of accurately allocating a person to the Italian, Turkish, Greek or any other typical European population.

So far, the seroanalytical allocation we have been dealing with is a typical decision technique in statistics that gives the expectation of error for a categorical decision in favour of either the null or the counter hypothesis. For a W value of, e.g., 99% a specific allocation is "highly probable"; for W = 99,73% and above a person's membership of a specific ethnic group is "practically proven". Specific allocations are impossible for W values below 99% (down to 1%). But they can still be useful indicators, provided that they are based not on a uniform, panmixed population, but on a "people mix". Then percentages can be related to the blood portions, above all if the blood portions refer to different major races. For example, if, using Indio, Spanish and Negro frequencies, one denotes the frequency of an urban Mexican's phenotype as $f(X_i)$, $f(X_s)$, $f(X_n)$. The respective probabilities in favour of the three races are given by the three formulas:

$$W_i = \frac{f(X_i)}{f(X_i)+f(X_s)+f(X_n)},$$

$$W_s = \frac{f(X_s)}{f(X_i)+f(X_s)+f(X_n)},$$

$$W_n = \frac{f(X_n)}{f(X_i)+f(X_s)+f(X_n)}.$$

If the calculated values for a certain person are, say, W_i = 70% (Indio percent), W_s = 20% (Spanish percent), W_n = 10% (Negro percent), the person cannot be allocated to any one of the three racial groups; rather, the person belongs simultaneously to all the three - and this in accordance with the established blood pertions 70% Indio blood, 20% Spanish blood and 10% Negro blood. If one applies seroanalysis in this way (and only now can one strictly speak of analysis), one is no longer working with hypotheses and probability values in terms of decision theory, but in the field of descriptive measurement. Although the sum of the values obtained is 1 (or 100%), the values themselves are no longer probabilities of accuracy or error.

In connection with the biostatistical evaluation of blood group findings to calculate e.g. a figure for the probability of fatherhood, a seroanalytical allocation tells us which ethnic allele frequencies must be used for the person in question. Seroanalytical mix-values tell us in which proportions frequencies from diverse populations must be mixed to do justice to the person in question.

This differentiation between two seroanalytical procedures may cause some uneasiness about finding a common theoretical foundation for the seroanalytical method. But there is a relativily simple way to combine them. If in a case of allocation the probability of the alternative hypothesis is negligible, and one allocates a person categorically to a specific population, this simply means that one deprives the alternative hypothesis of all value by arbitrarily reducing its probability to zero. This must regarded pragmatically, given a high probability in favour of the null hypothesis, one decides - for reasons of economy - to disregard the alternative hypothesis. By analogy, in the case of an exclusion from paternity one treats the hypothesis "non-paternity" as 1 and the hypothesis "paternity" as zero, notwithstanding the fact that in reality 1 (= 100%) and zero in natural science and statistics will never be reached.

Thus, in terms of approach, seroanalysis is not a technique of deciding for or against. Rather, the method always provides percentage portions for frequencies. However, those frequencies which lie below certain limit values are subsequently disregarded; such low frequencies are either unrealistic or at least devoid of practical relevance.