1982

Will Blood Tell Genetic Markers in Criminal Cases?

Randolph N. Jonakait
New York Law School, farrah.nagrampa@nyls.edu

Recommended Citation
Emory Law Journal, Vol. 31, Issue 4 (Fall 1982), pp. 833-912

WILL BLOOD TELL? GENETIC MARKERS IN CRIMINAL CASES

by
Randolph N. Jonakait*

I. INTRODUCTION

Forensic science has recently proclaimed a revolution in its ability to differentiate one person's blood from that of another.[1] Scientists who study blood groupings believe that each person's blood is unique: that a person's blood is as individual as that person's fingerprints.[2] Although scientists concede that such individualization of blood is not now possible,[3] many feel that it will be only a

---

* Associate Professor of Law, Pace University Law School; A.B., Princeton University, 1967; J.D., University of Chicago Law School, 1970; LL.M., New York University Law School, 1971. The author gratefully acknowledges the editorial assistance of Dr. G. Miller Jonakait and the research assistance of Linda R. Tacoma.

[1] "In criminal matters, the evidential value associated with bloodstains remained very limited until about ten to fifteen years ago. Since then forensic blood grouping has gone through a period of revolution . . . . " Baird, *The Individuality of Blood and Bloodstains*, 11 J. CAN. SOC'Y FORENSIC SCI. 83, 121 (1978).

[2] *See, e.g.*, Grunbaum, *Potential and Limitations of Forensic Blood Analysis*, in HANDBOOK FOR FORENSIC INDIVIDUALIZATION OF HUMAN BLOOD AND BLOODSTAINS 1, 2 (B. Grunbaum ed. 1981). Grunbaum states that "A simile has been drawn between fingerprints and blood as forensic evidence because both are unique to the individual and both are unchanging throughout life." *See also* Diamond, *The Story of our Blood Groups*, in BLOOD, PURE AND ELOQUENT 691 (M. Wintrobe ed. 1980). Diamond states that:

> From this modest start of four types, the individual characteristics of blood cells have been so greatly expanded that it is now highly improbable that any two people except identical twins would have the same combination of red cell surface markers. In other words, every person on earth is unique in his or her combination of blood groups—a fact recognized only in the past thirty years.

Id. As this indicates, blood grouping evidence will never have precisely the same value as fingerprints, for although identical twins do have the same blood groupings, *see* Terasaki, *Resolution by HLA Testing of 1000 Paternity Cases Not Excluded by ABO Testing*, 16 J. FAM. L. 543, 549 (1978), even identical twins have different fingerprints, *see* Baxter, *Grouping of Blood Stains: Present and Future Trends*, 12 CAL. W.L. REV. 284, 284-85 (1976).

[3] A noted British forensic serologist has concluded: "It is not yet possible to individualize blood in the same way as one can a fingerprint, but this is because of a lack of knowledge of techniques and not because of the nature of blood." B. CULLIFORD, THE EXAMINATION AND TYPING OF BLOODSTAINS IN THE CRIME LABORATORY 15 (1971). An American forensic serologist states, however, that "[a]bsolute individualization of blood, while theoretically possible, is not a practical goal for any laboratory." Grunbaum, *supra* note 2, at 2. *See also* Baird, *supra* note 1, at 88. Two other scientists have argued that the present trends in blood-typing research will not lead to the desired end of individualizing bloodstains. They contend

matter of time until that goal is achieved.[4] Right now, they contend, recent discoveries have provided significant advances in the ability to tell one blood sample from another so that it is almost possible to determine whether a particular specimen of blood came from a specific person.[5] Already, the scientific advances produce powerful testimony in criminal trials.

The effectiveness of scientific testimony at a criminal trial is illustrated by what transpires once a bloodstain is found at a murder scene. The blood is typed along with the blood of the victim. The two samples do not match. Circumstantial evidence indicates that the blood was shed by the murderer. A suspect is arrested. A sample of his blood is taken and typed. At the subsequent trial, a forensic scientist testifies that the blood samples were grouped not only by traditional typing tests, but also by newer procedures which are able to determine many more factors in the blood than the older methods. He states that even though the tests are of recent origin, they are used by many crime laboratories and are well-accepted in his field. The scientist further testifies that studies have been made to determine how often each particular blood-type occurs in the general population, and that therefore, he can calculate the frequency with which a particular combination of blood factors appear. He concludes that the bloodstain at the murder scene matches the defendant's blood, and that only one in a thousand people have blood like this. This powerful evidence helps persuade the jury to convict the defendant.

---

that current research "lacks the potential for attainment of what is probably the primary goal of forensic serology—the 'fingerprinting' or individualization of human bloodstains. This is so because typing detects only qualitative differences . . . ." Sweet & Elvins, *Studies by Crossed Electroimmunodiffusion on the Individuality and Sexual Origin of Bloodstains*, 21 J. FORENSIC SCI. 498, 505 (1976). Since no one is unique in these types, they conclude, "[c]learly, if individualization of bloodstains is to become a reality a technique must detect not only differences in kind but also in quantity." *Id.*

[4] In 1976, a forensic scientist stated that "[t]he evolution of forensic blood grouping is likely to reach a point in the next decade where the goal of identifying the individual is achieved." Baxter, *supra* note 2, at 286. Culliford suggests that perhaps in the future, blood-typing of all convicted criminals will be done and kept on record just as fingerprints are filed. B. CULLIFORD, *supra* note 3, at 17.

[5] *See* Baird, *supra* note 1, at 121. Baird states that "we are now facing a situation wherein it is almost possible to characterize human bloodstains to the extent that they can be as individualistic as fingerprints." *Id.*

This narrative is a composite of several real cases,[6] and the forensic evidence is not purely hypothetical. Increasingly, the forensic serologists' belief in a blood-typing revolution is reflected in criminal litigation. More and more cases are being reported in which expert testimony concerning the new blood-typing procedures has been admitted.[7] Since crime laboratories have only recently started to use these tests, such testimony should increase dramatically in the next few years, and will become nearly as common, and just as persuasive, as fingerprint testimony. A tremendous new tool will be in the prosecution's hands.

The use of these new procedures has, however, received little scrutiny and few cases show any challenge to the recently developed procedures for typing blood.[8] Perhaps that is because blood-typing has long been considered acceptable by our courts.[9] This new revolutionary evidence, however, is not the same as previous blood-typing evidence. The new procedures are unrelated to the long-accepted ones and their use should be raising important issues. Are these procedures reliable? Do they perform as their partisans proclaim? If this new evidence is frequently wrong or misleading, the prosecution may have added a powerful weapon to its arsenal, but one that will work injustices. These questions have not yet been addressed.

The purpose of this article is to examine these procedures in detail, to decide what the proper legal test for the determination of the blood-typing procedure's admissibility in criminal cases should

---

[6] See *infra* note 8 and text accompanying notes 78-122 for a discussion of the cases from which this narrative is drawn.

[7] Beginning in 1978, there are reported cases in which defendants did not challenge the use of new blood typing procedures. *See, e.g.*, People v. Stephens, 81 Cal. App. 3d 744, 146 Cal. Rptr. 748 (2d Dist. 1978); State v. Hampton, 294 N.C. 242, 239 S.E.2d 835 (1978).

[8] In some cases, for example, even though the new blood grouping tests were introduced by the prosecution, the reported decision does not indicate that the defendant even challenged the evidence. *See* People v. Stephens, 81 Cal. App. 3d 744, 146 Cal. Rptr. 748 (2d Dist. 1978); State v. Anderson, 308 N.W.2d 42 (Iowa 1981); State v. Hanson, 587 S.W.2d 895 (Mo. Ct. App. 1979); State v. Hampton, 294 N.C. 242, 239 S.E.2d 835 (1978). In other cases, although the statistics generated by the blood grouping testimony were litigated, the serological tests themselves were not. *See* State v. Rolls, 389 A.2d 824 (Me. 1978); State v. Carlson, 267 N.W.2d 170 (Minn. 1978); State v. Fulton, 299 N.C. 491, 263 S.E.2d 608 (1980).

[9] Annot., 2 A.L.R.4th 500 (1980).

be, and to present information about the reliability of the procedures. As shall be demonstrated, the new tests have not yet been proven reliable, and therefore, should not yet be admitted into criminal cases.

A. *The ABO Groups and Agglutination*

Everyone's blood is not the same. In papers published in 1900 and 1901, Vienna-born Dr. Karl Landsteiner presented the results of his pioneering research which showed "that the [blood] sera of some of his colleagues clumped together, or agglutinated, the red cells of others."[10] He had in fact discovered the first three blood groups, later called A, B, and O.

The clumping described by Landsteiner was caused by the reaction between antigens, or synonymously, agglutinogens, in the red blood cells of one sample and the antibodies, or agglutinins, in the serum of the other blood sample.[11] The reaction between the various elements of the blood can be explained simply. Red blood cells contain either the A antigen (type A blood),[12] the B antigen (type B), both A and B antigens (type AB), or neither of the antigens (type O). The blood serum may contain antibody anti-A, anti-B,

---

[10] Baird, *supra* note 1, at 87. Baird's article presents a good discussion of the history of blood-typing. *See also* Diamond, *supra* note 2, *passim* (historical development of the agglutination tests).

Agglutination of blood cells is a clumping of the red and white cells that are normally suspended throughout the fluid, or serum portion of whole blood. This clumping of blood is a reaction to the introduction of foreign antibodies into the blood. *See* A. MOENSSENS & F. INBAU, SCIENTIFIC EVIDENCE IN CRIMINAL CASES §§ 6.06, 6.07 (1978); STEDMAN'S MEDICAL DICTIONARY 1276 (23d ed. 1976) (definition of "serum").

[11] A group of serologists stated that the classical and most useful definitions are: "[1] An antigen is any substance which, when introduced parenterally [that is, not through the intestines] into an individual who himself lacks that substance, stimulates the production of an antibody, and, when mixed with the antibody, reacts specifically with it in some observable way; [2] [a]n antibody is a substance which appears in the plasma or body fluids as a result of stimulation by an antigen and which reacts specifically with that antigen in some observable way." K. BOORMAN, B. DODD, & P. LINCOLN, BLOOD GROUP SEROLOGY 3 (1977) [hereinafter cited as BLOOD GROUP]. *See also* Terasaki, *supra* note 2, at 545 (giving definitions of "antigen" and "antibody" similar to those provided in BLOOD GROUP, *supra*).

[12] In 1911, antigen A was found to consist of two major subgroups, denominated A_1 and A_2. *See* Diamond, *supra* note 2, at 694. However, "only rarely are the reactions between A_1 and A_2 so sharp and distinct that they are useful in legal situations." 2 AM. JUR. PROOF OF FACTS, *Blood Types* 277 (Supp. 1981).

both antibodies, or neither. Antigen A reacts with antibody anti-A resulting in the agglutination of the blood. Therefore, "[i]f a person has A antigen in his red blood cells he cannot have an anti-A antibody in his serum, for this would agglutinate (clump) his own cells."[13]

Blood can be readily typed in this ABO system simply by introducing into the untyped blood serum known to contain either the anti-A or anti-B antibody. "A blood will be agglutinated by anti-A serum, B blood will be agglutinated by anti-B serum, AB blood will be agglutinated by both anti-A and anti-B serum, and O blood will not be agglutinated by either anti-A or anti-B serum."[14]

Landsteiner's discoveries, however, did not end with the ABO groups. He also was instrumental in the discovery of the M, N, and P blood groups, and of the rhesus (Rh) factor in 1940.[15] Since then, a host of other antigen-antibody blood groups have been isolated.[16] All of these factors have in common that they are detected by the traditional, agglutination tests. Each factor "depends upon an antigen-antibody reaction visualized by agglutination of whole red blood cells. No other methodology has yet been developed to identify these blood groups in routine practice."[17]

Landsteiner also discovered that he could find the blood-type of a dried bloodstain, and recognized that this technique could be valuable in criminal investigations.[18] Despite Landsteiner's well documented research on the M, N, and P groups, however, modern crime laboratories seldom test for any antigen group besides the ABO factors.[19] ABO testing, however, can only produce limited re-

---

[13] A. MOENSSENS & F. INBAU, *supra* note 10, at § 6.11.

[14] *Id.*

[15] *See* Baird, *supra* note 1, at 88.

[16] *See generally* R. RACE & R. SANGER, BLOOD GROUPS IN MAN (6th ed. 1975).

[17] Zajac, *Bloodstain Phenotyping in Crime Laboratory Casework*, in HANDBOOK FOR FORENSIC INDIVIDUALIZATION OF HUMAN BLOOD AND BLOODSTAINS 160, 163 (B. Grunbaum ed. 1981). All of the blood groups are detected by serologic tests, therefore, "[a]ny test that detects antigens by using antisera (antibodies) is called a *serologic* test." Terasaki, *supra* note 2, at 545.

[18] *See* Baird, *supra* note 1, at 88; Diamond, *supra* note 2, at 692.

[19] In the crime laboratory, bloodstain evidence is most frequently typed for the ABO group system, and often there is no determination of the other antigens. . . . Analysts are deterred from making antigen determinations other than ABO for

sults because a great many people fall into each of the ABO groups. Consequently, ABO typing by itself has little evidentiary value in a criminal case.[20]

B. *Electrophoresis and Polymorphic Enzymes and Proteins*

Until recently, almost all the information about the blood-types came from antigen-antibody reactions. In 1975, an authoritative textbook in this field stated, "The great bulk of the earlier blood group knowledge was gathered from results of simple agglutination tests."[21] Blood, however, contains many genetically controlled substances other than the red cell antigens. All of these blood components are controlled by chromosomes and genes. Normally, every person has twenty-three pairs of chromosomes; half of each pair comes from each parent. Chromosomes are made up of genes. "Each characteristic of an individual is determined by two genes, or alleles, located at corresponding locations on the paired chromosomes. The pair of genes together constitute what is known as the

---

> several reasons . . . . The determinations call for complicated and lengthy procedures requiring numerous manipulations. A large amount of sample is needed compared with the ABO and enzyme systems. Sensitive, specific, reliable and inexpensive antisera for forensic use are not always available for the other antigens.

Zajac, *supra* note 17, at 163. Another forensic scientist has concluded, for reasons similar to those stated by Zajac, that "it is therefore likely that as further advances are made in the typing of blood stains, the serological methods will play a less dominant role." Baxter, *supra* note 2, at 296. *But cf.* B. CULLIFORD, *supra* note 3, at 105 (contending that in time, almost all of the antigen system will be detectable in dried blood and that the antigen Gm system is potentially the most useful genetic marker for forensic purposes).

[20] A major problem in relying on the traditional ABO typing in the context of criminal investigations is that because the categories by blood-type are so broad, such testing does not sufficiently narrow the class of persons who could be suspects.

> For example, a suspect in a homicide may have had a bloodstain on his shirtcuff which investigators believed could have come from the victim. How significant would the following findings be in associating the suspect with the victim? Suppose the suspect was of Group O while the victim was of Group A and the bloodstain on the suspect's cuff was of Group A. We know the stain could not have come from the suspect but on this determination alone it could have come from the victim. On the other hand, it is realized that about forty-two percent of our population is of Group A. Therefore, the finding of group A blood on the cuff cannot in itself carry much weight.

Baird, *supra* note 1, at 103.

[21] R. RACE & R. SANGER, *supra* note 16, at 3.

genotype."[22]

Some genes, however, exert a stronger influence than others. "The 'more powerful' gene is called the dominant; the 'less influential' is called the recessive."[23] When a dominant is paired with a recessive gene, the dominant one controls the observable characteristics, or phenotype. In the ABO system, the A and B genes are dominant, while the O gene is recessive. Thus, if a person has the genetic pair AO, he has the genotype AO. Since, however, the A gene is dominant while the O is recessive, the A controls, and that person's phenotype is type A blood. He therefore has the same phenotype as the person with the genotype AA.[24]

In 1949, scientists discovered that a particular substance in the blood other than an antigen could appear in genetically distinct variations.[25] Scientists have since learned that blood components exist in genetically controlled types. Such substances are said to be *polymorphic,* which means "having different molecular forms but the same biochemical function."[26] Blood thus contains phosphoglucomutase (PGM), which takes different forms. One person may have PGM "1," while another has PGM "2-1." In each person, however, PGM has the same physiologic function—to "catalize a reaction in the metabolism of sugars in the body."[27] PGM,

---

[22] Baird, *supra* note 1, at 102. *See also* Terasaki, *supra* note 2, at 545 (discussion of basic genetic terms).

[23] 2 AM. JUR. PROOF OF FACTS, *Blood Types* 611 (1959).

[24] Genetic control of the blood groups is important in criminal cases because it assures that a person's blood-type always stays the same, except for shortly after a transfusion. *See* A. MOENSSENS & F. INBAU, *supra* note 10, at § 6.11 ("A person's blood group remains constant throughout life notwithstanding age, disease, or medication."). For a further discussion of the implications of a transfusion on blood-types, see *infra* note 57. *See also* Grunbaum, *supra* note 2, at 3 ("Since the genetic factors are permanent and unchanging throughout life, they are of primary importance.").

The laws of inheritance are not used in the attempt to individualize a blood sample although they are essential in paternity resolutions. *See Joint AMA-ABA Guidelines: Present Status of Serologic Testing in Problems of Disputed Parentage,* 10 FAM. L.Q. 247 (1976) [hereinafter cited as *Joint AMA-ABA Guidelines*].

[25] *See* A. MOURANT, A. KOPEC, & K. DOMANIEWSKA-SOBCZAK, THE DISTRIBUTION OF THE HUMAN BLOOD GROUPS AND OTHER POLYMORPHISMS 25 (2d ed. 1976) [hereinafter cited as DISTRIBUTION OF HUMAN BLOOD GROUPS].

[26] Grunbaum, *supra* note 2, at 3.

[27] Baird, *supra* note 1, at 109.

while existing in different molecular forms, always serves the same purpose, and therefore is polymorphic.

While many polymorphisms have already been discovered, it is likely that more will still be found.[28] Polymorphic enzymes and proteins,[29] however, are not typed by the traditional agglutination tests. Instead, electrophoresis is used.[30] Electrophoresis is a

> physical method for the separation of biologically important proteins through the use of electric current. Proteins are very complex molecules which assume positive, negative, or neutral charges, depending on the solution in which they are placed. When these charged molecules are placed on an appropriate medium and subjected to an electrical field, they will migrate toward the pole of the opposite charge. Blood proteins vary in size, shape, density, and charge; consequently they vary in electrophoretic mobility. Therefore, after electrophoresis, they are separated into distinct bands on the supporting medium.[31]

---

[28] The joint ABA-AMA committee listed sixty-three different blood systems. *See Joint AMA-ABA Guidelines, supra* note 24, at 253-56, Table 1. *See also* DISTRIBUTION OF HUMAN BLOOD GROUPS, *supra* note 25, at 26. Besides PGM, other polymorphic enzymes or proteins commonly include adenylate kinaise (AK), adensine deaminase (ADA), esterace D (EsD), erythrocyte acid phosphotase (EAP), and haptoglobin (Hp). See *infra* text accompanying note 39 for a further discussion of these polymorphic enzymes which are used as genetic markers.

[29] Even though "all immediate gene products are now known to be proteins," the polymorphisms are usually classified as the red cell antigens or the polymorphic enzymes and proteins. DISTRIBUTION OF HUMAN BLOOD GROUPS, *supra* note 25, at 25. *See* Grunbaum, *supra* note 2, at 3. Enzymes are one specific kind of protein that catalyze biochemical reactions, or in other words, help to convert one substance to another. *See* DISTRIBUTION OF HUMAN BLOOD GROUPS, *supra* note 25, at 25; Baird, *supra* note 1, at 109. The genetic variants of an enzyme are isoenzymes or isozymes. *See* DISTRIBUTION OF HUMAN BLOOD TYPES, *supra* note 25, at 25; Baird, *supra* note 1, at 103. Most of the polymorphic systems detected in forensic work are enzymes. The terms "blood group" or "blood-type" are usually limited to the antigen systems. The term "genetic marker" refers to any polymorphism whether detected by agglutination, electrophoresis, or some other means. *See* Baird, *supra* note 1, at 105; Grunbaum, *supra* note 2, at 3.

[30] Some scientists are attempting to perfect even newer methods with the expectation that these refined techniques should give greater phenotyping ability than standard electrophoresis. *See* Burdett, *Isoelectric Focusing in Agarose: Phosphoglucomatase (PGM Locus 1) Typing*, 26 J. FORENSIC SCI. 405 (1981). Nothing indicates that forensic laboratories are using anything other than the agglutination and electrophoretic tests in actual casework.

[31] Grunbaum, *supra* note 2, at 3. *See also* A. MOENSSENS & F. INBAU, *supra* note 10, at § 6.11; Baird, *supra* note 1, at 109-12.

Subsequent to the separation, dyes that unite with one specific protein are applied and the intensity of the resulting stain is measured.

C. *Genetic Markers in Dried Blood*

The forensic serologist faces a problem which other scientists studying blood do not. The scientist concerned with transfusions, genetic diseases, blood disorders, or paternity resolutions is not concerned with dried blood. The forensic scientist, however, will most often be confronted with blood in the form of a bloodstain —blood that has been shed by someone and aged.[32] Dried blood obviously is not the same as fresh blood. The antigen-antibody tests as described above do not work with dried blood. Rather, "direct agglutination tests are not possible for the detection of antigens in blood stains since the red cells have been destroyed . . . ."[33] Other methods, however, have been devised to find the traditional blood groups in a bloodstain. These methods also depend on the antigen-antibody reaction.

> [H]owever, the technique is somewhat more complicated and requires more expertise. [The methods] rely on the ability of the antigen present in the bloodstain material to absorb [i.e., picks up] its specific anti-serum. . . . [I]n the ABO system, if the bloodstain contains A, anti-A serum will be absorbed; if the stain contains antigen B, anti-B serum will be absorbed. . . . The blood group can then be determined by eluting (liberating) the absorbed serum, if any, and testing the same against known A, B, and O red blood cells. . . .[34]

---

[32] One group of forensic scientists reported that "more than 90% of serological clue materials are found on textiles," and thus more than 90% must be dried blood. *See* Denault, Takimoto, Kwan, & Pallos, *Detectability of Selected Genetic Markers in Dried Blood on Aging*, 25 J. FORENSIC SCI. 479, 480 (1980) [hereinafter cited as Denault Study].

[33] BLOOD GROUP, *supra* note 11, at 405. *See also* B. CULLIFORD, *supra* note 3, at 67; Baird, *supra* note 1, at 84.

[34] A. MOENSSENS & F. INBAU, *supra* note 10, at § 6.11. More than one technique is used to detect antigens in dried blood. The one described in the text is called absorption-elution. "Of the several methods available for the determination of antigens in dried bloodstains, absorption-elution is the most sensitive and the most widely employed. . . . [A]bsorption-elution has proved to be markedly more sensitive than the absorption-inhibition method; it has also been reported to be more successful than mixed agglutination for certain antigens." Denault Study, *supra* note 32, at 481. *See also* Baird, *supra* note 1, at 101-02; Baxter, *supra*

Although the technique for testing dried blood may be refined further, the last major advance in the typing of antigen groups in bloodstains came in 1960.[35] The typing of genetic markers in dried blood through the use of electrophoresis, however, did not begin in earnest until 1968.[36] The modern electrophoretic techniques with dried blood are almost identical to the ones used with fresh blood: "There is no difference in methodology from fresh blood to dried blood except in preparation of the samples and in the extra care that must be exercised in interpretation of results when degraded material is studied."[37]

As will be discussed below, however, blood begins to deteriorate the moment it leaves the body, and thus, dried blood differs from fresh. Not all genetic markers detected in fresh blood have been successfully found in bloodstains; "experience has shown that there are differences in the degree of detectability and persistence of the antigen, protein and enzyme systems."[38] Consequently, a time lag has existed between the discovery of a genetic marker in fresh blood with that marker's identification in dried blood. Therefore, while the discovery of the enzyme and protein genetic markers is recent, detection of them in dried blood is even more recent. For example, of the markers now widely typed in forensic laboratories, PGM was first identified in dried blood in 1969; adenylate kinase (AK) in 1966; adenosine deaminase (ADA) in 1971; esterase D (esD) in 1973; and erythrocyte acid phosphatase (EAP) in

---

note 2, at 286.

[35] The advances in the typing of antigen groups in bloodstains have come about in a somewhat hit or miss fashion. *See* Baird, *supra* note 1, at 101 ("The classical methods of grouping dried bloodstains have evolved over the years essentially by trial and error . . . ."). The first breakthrough came about in 1932, but "it was not until 1960, when Kind introduced the absorption-elution technique, that success was achieved with small stains. Since 1960 many workers have modified the original method of Kind . . . . " Baxter, *supra* note 2, at 286.

[36] *See* Baxter, *supra* note 2, at 286.

[37] Grunbaum, *Procedures for Phenotyping of Genetically Controlled Enzyme and Protein Systems*, in HANDBOOK FOR FORENSIC INDIVIDUALIZATION OF HUMAN BLOOD AND BLOODSTAINS 51, 103 (B. Grunbaum ed. 1981). The difference in sample preparation is that the stain must first be extracted in distilled water. *Id.* at 55. *See also* Zajac, *supra* note 17, at 161.

[38] Zajac, *supra* note 17, at 160.

1976.[39]

Consequently, many of the procedures used to identify genetic markers in dried blood have just been developed. Many of us who are untrained in scientific procedures are deeply impressed by recent advances in technology, and want to take advantage of scientific knowledge as soon as it becomes available. The law, however, has a duty to ask and to answer an important question before these tests become admissible as evidence: Should these new scientific tests be admitted into criminal cases?

II. The Legal Standard for Admitting the Genetic Marker Tests

Whether the blood tests are admissible can only be determined after first deciding what legal standard should govern their admissibility. This section of the paper will examine the traditional norm for assessing the admissibility of new scientific evidence. It will show that the reasons that the traditional standard ordinarily guarantees that only reliable scientific evidence will be admitted do not apply to new blood procedures. Instead of using the traditional legal test, the new procedures should be admitted into criminal trials only after comprehensive scientific experiments have proven that the genetic marker evidence is reliable.

A. *The* Frye *Standard*

In 1923, the Court of Appeals for the District of Columbia decided *Frye v. United States*.[40] In rejecting the admissibility of polygraph evidence, the court held that:

> Just when a scientific principle or discovery crosses the line between the experimental and demonstrable stages is difficult to define. Somewhere in this twilight zone the evidential force of the principle must be recognized, and while the courts will go a long way in admitting expert testimony deduced from a well-recognized scientific principle or discovery, the thing from which the deduction is made must be sufficiently estab-

---

[39] *See* Blood Group, *supra* note 11, at 407; Baird, *supra* note 1, at 113, 115.
[40] 293 F. 1013 (D.C. Cir. 1923).

> lished to have gained general acceptance in the particular field in which it belongs.[41]

The precise status of this test for determining the admissibility of novel scientific evidence is not entirely clear. Some jurisdictions have accepted it; some have rejected it; others have modified it; and still others have stated that they were applying it even though their decisions implicitly ignored it. It is clear, however, that this 1923 language has shaped the debate over the proper standards for admitting new scientific techniques,[42] and that "the *Frye* test has dominated the admissibility of scientific evidence for more than half a century."[43]

As the mixed reactions should indicate, the *Frye* requirements have their critics. Indeed, the attacks have been labelled "scathing."[44] The test has been characterized as vague and therefore difficult to apply.[45] One court has stated that the *Frye* standard "is usually construed as necessitating a survey and categorization of the subjective views of a number of scientists . . . but a determination of reliability cannot rest solely on a process of 'counting (scientific) noses.' "[46] The strongest criticism, however, is that the test keeps reliable evidence from the jury. Although a new scientific procedure might be proven reliable, the *Frye* standard suppresses such evidence until the scientific community has taken the time to pass judgment on the reliability of the procedure. The critics contend that this delay between proven reliability and scientific consensus denies the court from ruling on valuable evidence.[47] In

---

[41] *Id.* at 1014.

[42] For an excellent discussion of the current status of the *Frye* test, see Giannelli, *The Admissibility of Novel Scientific Evidence:* Frye v. United States, *a Half-Century Later,* 80 COLUM. L. REV. 1197, 1228 (1980).

[43] *Id.* at 1205.

[44] *Id.* at 1206.

[45] *Id.* at 1223. *See also* United States v. Williams, 583 F.2d 1194, 1198 (2d Cir. 1978), in which the Second Circuit states that the "[d]ifficulty in applying the *Frye* test has led a number of courts to its implicit modification."

[46] 583 F.2d at 1198. The *Williams* court stated that "[i]n testing for admissibility of a particular type of scientific evidence, whatever the scientific 'voting' pattern may be, the courts cannot in any event surrender to scientists the responsibility for determining the reliability of that evidence." *Id.*

[47] *See* Giannelli, *supra* note 42, at 1223.

other words, even though "every useful new development must have its first day in court,"[48] the *Frye* test delays that day longer than is necessary.

Defenders of *Frye* admit that the test erects a difficult standard that delays the admission of scientific evidence:

> This obviously sets forth a standard that is neither common to criminal litigation nor easily applied in the individual case. Equally obvious, the *Frye* standard retards somewhat the admission of proof based on new methods of scientific investigation by requiring that they attain sufficient currency and status to gain the general acceptance of the relevant scientific community.[49]

A conservative approach is warranted, defenders maintain, "[s]ince scientific proof may in some instances assume a posture of mystic infallibility in the eyes of a jury of laymen. . . . "[50] The chief advantage of the *Frye* test is that reliability of the new proof virtually is guaranteed since "the requirement of general acceptance in the scientific community assures that those most qualified to assess the general validity of a scientific method will have the determinative voice."[51]

While these arguments may be representative of the debate over the admissibility of scientific evidence, both sides overlook an important point. Both proponents and detractors of *Frye* agree that a stringent legal standard assures that a scientific test is reliable before it is admitted. One side believes *Frye* delays admissibility until too long after reliability has been proven, while the other believes caution is necessary to prevent lay jurors from giving technical evidence undue weight. As this author shall demonstrate, however, the *Frye* standard is insufficient to assure reliability in the case of the genetic marker tests. Whether these forensic procedures are reliable or not, *Frye* admits them. Since reliable evidence is the goal and since *Frye* does not serve that purpose here, a different

---

[48] *See* United States v. Stifel, 433 F.2d 431, 438 (6th Cir. 1970).

[49] United States v. Addison, 498 F.2d 741, 743 (D.C. Cir. 1974).

[50] *Id.* at 744.

[51] *Id.* at 743-44. *See also* People v. Kelly, 17 Cal. 3d 24, 31, 549 P.2d 1240, 1244-45, 130 Cal. Rptr. 144, 148 (1976); Giannelli, *supra* note 42, at 1205-07.

admissibility standard is needed for these procedures.

B. *Why* Frye *Normally Works*

Frye rests on the notion that science is a never-ending construct. A scientific discovery does not stand alone; instead, each new piece of knowledge is a block upon which others will build. Before that new building block is put into service, however, it is tested in a disinterested way to make certain that it is sound. If somehow an unsound piece slips through, it will be discovered as others try to use it.

Because this is the nature of science, a new scientific procedure will normally go through a three-step process: first, the procedure is developed, then other scientists will test the results, and finally, if the procedure is sound, general use of it will follow. Only at the end of this process is the reliability of the procedure clearly established. Suppose, for example, that some types of blindness are caused by tiny blood vessels in the eyes wearing out and bursting. This results in enough bleeding in the eyeballs to cause pressure that results in the deteriorating eyesight. A scientist believes that this condition could be arrested by the sealing of those capillaries. Working on animals, he develops a laser surgery technique that he believes can stop the bleeding without doing damage to the rest of the eyeball. He publishes his results.

Other eye specialists reading this report may be impressed by it, agree with the theory behind it, and understand how the new procedure should alleviate the condition. They would not yet, however, accept the procedure as reliable. Instead, verification through the use of controlled studies would next be required. Two sets of patients with the condition would be needed, one set to receive the new treatment, and one control set. This would be done for two reasons: first, to see if the procedure worked the cure as claimed; and second, to see if any limitations, such as side-effects, could be discovered. Such dangers, of course, should be identified before the technique goes into general practice.

If the new procedure passed the controlled studies, it might be regarded as reliable, but the best guarantee of reliability is still to come. The ultimate test would be the actual use of the technique

in general practice. Any limitations that were not unearthed in the studies would surface as the procedure was used under various circumstances. It might eventually be learned that while the surgery at first seems to be successful with diabetics, after a year or so the blood vessels burst again causing greater problems than before. Certainly, only when a procedure has gone through these three steps—development of the methodology, verification of the claimed results, and actual employment of the new technique—can the community of concerned scientists know both the procedure's reliability and its limitations.

Of course, not all scientific discoveries have the immediate clinical impact of this posited eye surgery. One can expect that any procedure affecting the health or safety of humans would undergo rigorous verification before it were used. That is not to say, however, that other discoveries would not have to mount similar steps to test their reliability and limitations. A group of scientists, for example, may be concerned with trying to learn how the nervous system develops in a fetus. At present, basic research is being conducted in laboratory animals in an attempt to learn about the biochemicals that relay "messages" from one nerve to another. Suppose one scientist believes he has perfected a more efficient way to detect minute amounts of one such chemical, for example, norepinephrine, in the tissue of rats. This new assay may not have any immediate human or clinical applications, but if it is truly the advance the developer hopes for, other scientists working in his field would also want to use it. These other scientists, however, may insist on verification of his work before they use the new procedure. Though human health may not be an immediate concern, the time, effort, and integrity of the scientists would be at stake. Finally, the ultimate check on the reliability and the limitations on the novel assay would become apparent as the technique was used by other scientists. This check would arise naturally because the detection of the chemical is not an end in itself. Instead, the scientists would view the chemical detection merely as one step in furthering the understanding of the developing nervous system. If for some reason the new assay delivered authentic looking results that were actually spurious, this would become known as scientists expanded their knowledge by unknowingly relying upon the inexact

procedure.

For example, a scientist may be studying the effect of nicotine ingestion by pregnant mothers on the developing fetus. In this basic research, nicotine is given to pregnant rats, and tissues normally known to contain norepinephrine are examined in the fetal rats. As a matter of convenience, in the past this scientist freeze-dried the tissues when obtained and stored them until a large batch had been collected. The scientist usually found it more economical to run the assays on many tissues at once than on each when taken separately. With the old procedure, this method produced no problems. Assume that the new assay, however, performs perfectly on fresh tissues, but for some reason does not work on those that have been frozen, and this limitation is not yet known. The scientist using this new procedure may think that he has discovered that the nicotine ingestion in the mother inhibits the production of the biochemical in the fetus. This presumably would be an important finding, and that scientist or others would want to build on this knowledge. Other experiments would follow. Similar work would perhaps be done, and other nervous system chemicals would be examined. The scientist might try to see if it mattered at what point in the pregnancy term the mother ingested the nicotine, or if the ingestion of lesser amounts of nicotine did not prevent the production of norepinephrine, or at what point in its development the fetus stopped producing that chemical. Further experiments likely would occur, and sooner or later someone working in the field might discover that the fetus actually had been producing the substance and that the originally reported results were wrong because the new assay did not work under these circumstances.

The ultimate guarantor of reliability in both these examples, then, is that the new test is put into practice and this practice eventually shows whether the procedure is unreliable or has limitations. The test creates incentives to check the reliability and limitations of the procedure before it becomes widespread. When a procedure has become so widely used in a field that any flaws in the procedure would have become known, the courts can conclude that not only has the new test been accepted as reliable by the

relevant scientific community, but that the test also is reliable.

The basic notion behind the *Frye* test, then, is a good one. Once a procedure is sufficiently established to have gained general acceptance in the particular field in which it belongs, it presumably has gone through an extended period of use and testing within the scientific community and is reliable. The question that naturally arises, however, is whether the typing of blood by the new forensic procedures is subject to the same dynamics as other new scientific tests?

C. *The* Frye *Test Does Not Guarantee Reliability of the Genetic Tests*

Unlike most scientific procedures, those used to detect genetic markers in dried blood will not, through repeated use, assure the validity and reliability of the procedures.[52] In contrast to the errors of other scientific investigators, the errors of the forensic serologist will not invariably become known. Unlike the use of most scientific tests, the detection of genetic markers in the forensic laboratory is the ending point; these results are not employed by scientists in ways that would invariably reveal limitations of the procedure.

Clinical practice does not need the tests of the criminal investigator since the procedures are not used in the treatment of patients. In fact, the role that many of the polymorphic enzymes play in the body is still unknown.[53] Even when the function has been discovered, the clinician still may not be interested in the details provided by forensic tests. Thus, in treating a person with a metabolic problem, it is conceivable that a doctor may want to know whether PGM is present in that person in normal amounts. While that doctor might be concerned with the *levels* of PGM, however,

---

[52] In scientific terminology, "validity" and "reliability" have different definitions. " 'Validity' refers to the ability of a test procedure to measure what it is supposed to measure—its accuracy. 'Reliability' refers to whether the same results are obtained in each instance in which the test is performed—its consistency. Validity includes reliability, but the converse is not necessarily true." Giannelli, *supra* note 42, at 1201 n.20.

[53] For example, although it has been discovered that PGM and AK both act as catalysts in the metabolism of sugars, the precise function of EAP is not known. *See* DISTRIBUTION OF HUMAN BLOOD GROUPS, *supra* note 25, at 38, 41; Baird, *supra* note 1, at 109, 113.

he does not care which of the polymorphic *forms* that enzyme takes. All the variants of PGM perform the same function in the body.[54] PGM "2" does the same as PGM "2-1." Consequently, the doctor might run tests to see how much PGM is present; he will not have any reason to find out what kind of PGM it is.[55]

Although clinical practice is not concerned with the polymorphic form a blood constituent may take in an individual, there is concern with how the blood of one person may react when put into the body of another—when there is a transfusion. Transfusion reactions, however, are only caused by the antigen-antibody systems.[56] The substances detected by electrophoresis do not have any adverse transfusion consequences. It does not matter if blood with PGM "1" is put into a patient with PGM "2-1."[57] It does matter, of course, if a person with type "B" blood gets a transfusion of type "A" blood because agglutination and possible death would result. Consequently, if an unreliable test for typing the ABO groups were used, a blood bank would quickly find out. Thus, the reliability of the blood bank's test is assured by its use.

The blood bank's procedures, however, are not those used by the forensic scientist.[58] It cannot be stressed too much that the forensic laboratory will deal with dried blood—blood that has aged in

---

[54] For a discussion of the polymorphic characteristics of blood, see *supra* text accompanying notes 25-27.

[55] Even if the doctor had run tests to determine which form that enzyme took, he would not know from his clinical work whether the tests to classify the genetic markers gave valid results since there would be no consequences for his patient if, for example, PGM "2-1" were incorrectly classified as "2."

[56] *See* Baxter, *supra* note 2, at 284 & n.2.

[57] Important forensic consequences might occur from such a transfusion. Culliford notes that a victim of an assault who has PGM "1" may get a transfusion of blood with PGM "2." He goes on to state that "this will not harm the patient but could make a blood sample taken as a control from the victim unreliable for forensic purposes." B. CULLIFORD, *supra* note 3, at 38. In other words, the classification of genetic markers of someone who has recently received a transfusion may be wrong. Culliford states that the forensic laboratory should know if such a transfusion occurred, and if so, should try instead to type a pre-transfusion sample of blood from the victim or wait four to six weeks to classify the victim's genetic markers. *Id.*

[58] The blood bank normally only tests for two systems: "The routine blood grouping tests in the normal clinical work are limited to the four A-B-O groups and the Rho blood factor. . . . [Other antigen tests] are not included in the ordinary tests for matching blood for transfusions." 2 AM. JUR. PROOF OF FACTS, *Blood Types* 279 (Supp. 1981).

ways not controlled by the serologist. The blood bank, however, does not use the procedures of the forensic scientist because it has no interest in identifying blood groups in dried blood. Those concerned with transfusions have no stake in trying to identify the type of blood found dried on a defendant's sweater. Rather, the blood bank merely types a liquid substance, either fresh blood or liquid blood preserved under known conditions.

The forensic scientist stands alone in the attempt to classify genetic markers in dried blood,[59] and thus, those tests have not been validated by use or research in other scientific areas. Furthermore, the forensic scientist does not put his results to a scientific use that would invariably show the limitations and liabilities, if any, of the procedures. When the forensic laboratory classifies a blood sample's PGM into one of its genetic variants, this classification is not merely a research step upon which that forensic scientist or others will continue to build. The scientist's job, conversely, ends with that categorization rather than putting these findings to any further use. Consequently, if for some unknown reason the procedures to detect genetic markers in dried and aged blood falsely detect PGM "1" as "2-1," this incorrect result will not necessarily be discovered by any subsequent scientific use or practice.

For purposes of the *Frye* test, this means the detection of genetic markers in dried blood is different from other scientific procedures. Acceptance by the relevant scientific community normally means that a procedure has been employed in such a way that the procedure's reliability would become known. This is not true for the forensic detection of genetic markers because these procedures are not used in ways that would reveal any limitations. The reliability of these forensic tests can only be shown from controlled experiments. Therefore, the *Frye* standard, which relies on the general acceptance of a test to affirm that test's reliability, does not serve its purpose. Mere widespread use of the forensic procedures proves nothing about their trustworthiness. Furthermore, an additional compelling argument why *Frye* should not be applied to the

[59] Genetics is the other branch of science that studies polymorphisms. The geneticist, however, does not have an interest in classifying genetic markers in dried blood. Instead, he, like the blood bank, works with fresh or preserved liquid blood.

genetic marker test derives from the unique nature of the field that uses these tests.

D. *Problems Concerning the Genetic Marker Field*

Before results of a test can be admitted into evidence, *Frye*, of course, requires that the test be generally acceptable in the relevant scientific field.[60] What then, is "the field" for these genetic marker procedures? The only field in which the tests for genetic markers in dried blood have been generally accepted and used happens to be a group with special characteristics. As we have seen, although several areas of science or medicine may be interested in genetic markers, only one attempts to detect them in dried blood.

The relevant scientific field does not contain all of the sciences concerned with blood. If it did, the tests could not be considered generally acceptable because only a small percentage of that group has any interest whatsoever in the detection of genetic markers.

The field does not even encompass the entire subset of scientists who study genetic markers—a subset that contains blood bankers, geneticists, and doctors interested in such areas as metabolism or transfusions. The forensic tests are not established throughout this field, either, because, as we have seen,[61] this group is not concerned with the detection of genetic markers *in dried, aged blood.*

Rather, the tests for the detection of genetic markers in dried blood can possibly be accepted in this field only if that field is defined as a subset of the subset — those who study the detection of genetic markers in dried blood.[62] That field becomes a quite special

---

[60] *See* Giannelli, *supra* note 42, at 1205.

[61] See *supra* text accompanying notes 53-59 for a discussion of the differing emphasis that criminal investigators, doctors, and blood bankers place on various blood testing procedures.

[62] In the leading case of People v. Williams, 164 Cal. App. 2d 858, 331 P.2d 251 (1958), the court defined the field in a similar way to allow the admissibility of the Nalline test, a procedure for the detection of narcotics in a person. It was conceded "that the medical profession generally is unfamiliar with the use of Nalline and therefore it cannot be truthfully said that the Nalline test has met with general acceptance by the medical profession as a whole . . . . " *Id.* at 862, 331 P.2d at 253. The court, however, went on to admit the test by concluding that "[i]t has been generally accepted by those who would be expected to be

one, consisting only of forensic scientists, or more precisely, the fraction of forensic scientists concerned with blood.

This specialized field has two important characteristics that ought to give a court pause before it accepts this group's certification that one of its forensic procedures is reliable. First, it is a small field. Not many scientists fall into the category of forensic serology.[63] An implicit foundation of the *Frye* standard is that a court can be assured of a scientific procedure's reliability if a large number of scientists accept the procedure. "If the 'specialized field' is too narrow, the consensus judgment mandated by *Frye* becomes illusory; the judgment of the scientific community becomes, in reality, the opinion of a few experts."[64]

Second, the field of forensic serology is almost entirely composed of law enforcement professionals. The only people trying to detect genetic markers in dried blood are forensic scientists, and these scientists are almost exclusively either employees of law enforcement agencies or are funded by such governmental bodies.[65] One could hardly classify this group as impartial on the question of the reliability of the techniques it uses. Indeed, many, if not most, have a stake in convincing a court that their tests are reliable since the significance of their work and careers would be greatly devalued if the evidence were barred from criminal trials.

A corollary is that the evidence about reliability and acceptability will only be presented by one of the adversaries at the trial. The characteristics of the community interested in the detection of

---

familiar with its use. In this age of specialization more should not be required." *Id.* at 862, 331 P.2d at 254. For a discussion of the identification of the appropriate field using the *Frye* standard, see Giannelli, *supra* note 42, at 1208-10. *See also* United States v. Williams, 583 F.2d 1194 (2d Cir. 1978), in which the court, in considering whether it was proper to admit spectographic voice-identification evidence, stated that "[s]election of the 'relevant scientific community,' appears to influence the result" of determining whether such evidence has been accepted as reliable. *Id.* at 1198.

[63] Although people may classify themselves as forensic serologists, there is no certification process for a forensic serologist. Letter from Dr. B. W. Grunbaum to R. Jonakait (July 18, 1981).

[64] Giannelli, *supra* note 42, at 1209-10.

[65] Much of the research into the detection of genetic markers was funded by the Law Enforcement Assistance Administration. See *infra* note 135 for a discussion of the LEAA practice of not funding verification studies.

genetic markers in dried blood indicates that the defense may have a difficult time finding a qualified person to analyze evidence from the defendant's viewpoint. Justifying the use of the *Frye* standard, one court observed that "the *Frye* test protects prosecution and defense alike by assuring that a minimal reserve of experts exists who can critically examine the validity of a scientific determination in a particular case."[66] This reserve of experts, however, does not exist here. Instead, the composition of this scientific field indicates that normally the evidence will be presented and analyzed only by someone from a small group of people, all of whom have a prosecution perspective.

This does not mean, of course, that this group's conclusion as to the reliability of its procedures would necessarily be incorrect or biased. A court, however, should be aware of forensic serology's assertion that the tests are reliable and valid is really the conclusion of the few and the partisan. Under such circumstances, it is only fair to insist that the courts carefully analyze the claim of reliability. The history of the paraffin test affords a valuable illustration and warning.

E. *Lessons from the Paraffin Test*

In 1933, Mexican Theodora Gonzalez announced a procedure to determine whether a person had recently fired a gun.[67] "The theory behind the test was that the results established the presence of particles of nitrates or nitrites, deposited on the hand by the bases of a discharged cartridge."[68] The popular name for the test—the paraffin test—evolved because hot paraffin was applied to the hands. It was believed that this substance caused "the pores of the skin to open up, [mixed] with the oil in the pores, and [caused] the pores to exude the gunpower residues imbedded therein."[69] Enough paraffin was applied until "a paraffin glove [was] finally molded, which, when cooled, [was] peeled gently from the

---

[66] United States v. Addison, 498 F.2d 741, 744 (D.C. Cir. 1974).

[67] A. MOENSSENS & F. INBAU, *supra* note 10, at § 4.12.

[68] *Id.*

[69] Conrad, *Evidential Implications of the Dermal Nitrate Test for Gunpowder Residues*, 44 MARQ. L. REV. 500, 503-04 (1961).

hand. . . . An organic reagent, diphenylamine . . . [was] added to the paraffin glove drop by drop. A positive reaction indicating the presence of gunpowder residues [was] indicated by the appearance of dark blue pin-point specks on the inner surface of the case."[70]

After its introduction, "the paraffin test was adopted quickly by law enforcement agencies," with the first reported case upholding the admissibility of this test decided in 1936.[71] Although reported cases on the procedure were nonexistent for the next generation, "the procedure was in general use in courts on a trial level."[72]

Experts were probably testifying at trial that the paraffin test was reliable and was generally accepted in its scientific field. For the same reasons presented above, however, hindsight shows us that the courts should have questioned these assertions. The parallels between the paraffin test and the genetic marker tests are close. First, like the dried blood procedures, scientists did not use the results of the paraffin test in such a way that any inherent unreliability would inevitably be discovered. No one besides forensic scientists would have been interested in this procedure. In actuality, for the forensic scientist, the tests were ends in themselves. Like dried blood testing, the detection of gunpowder residues was not a block upon which other research was built—research that would have exposed the limitations of the test. Thus, the paraffin test, like the dried blood procedures, never faced this strongest guarantor of reliability.

Second, as with the dried blood tests, generally the experts all would have been forensic scientists since they would be the only ones familiar with the tests. This meant that the evidence about reliability again would come only from a small number of people all with bias toward one side of the question. As with the forensic tests for genetic markers, reliability could truly be established only if careful experimentation were done to discover the test's limita-

---

[70] *Id.* at 502.

[71] *See* Giannelli, *supra* note 42, at 1224. The first case upholding the admissibility of the paraffin test was Commonwealth v. Westwood, 324 Pa. 289, 188 A. 304 (1936) (paraffin test admitted to show gunpowder residue on defendant's hand). See Conrad, *supra* note 69, at 507-09, for a discussion of *Westwood*.

[72] Conrad, *supra* note 69, at 509.

tions. Such a study was finally undertaken, although it was not until 1967 that the first comprehensive study was published.[73] This research showed that the paraffin test was unreliable because it gave positive results for many people who had never fired a gun.[74] Even though it is now recognized that this test is not valid because of the false positives,[75] "evidence based on the test was admitted in trials throughout the 1960's."[76]

Because of the lapse of time until the fallibility of the test was discovered, jurors, for over thirty years, were told that they could base their verdicts on what turned out to be unreliable evidence.[77] This evidence was admitted because courts accepted the assertions of forensic scientists about the tests. The scientists made those assertions even though comprehensive reliability studies had not been done. When such research was produced, the forensic scientists' confidence in their procedures was shown to be egregiously misplaced. This history should give pause to any court faced with the admissibility of the procedures claimed to detect genetic markers in dried blood. As will be discussed, the blood tests now stand where the paraffin tests stood before the comprehensive reliability studies were done.

Frye, with its premium on general acceptance in a scientific field, does not guarantee reliability of the genetic marker procedures that are used solely by forensic scientists and not used to

---

[73] See Giannelli, *supra* note 42, at 1224-25, for a discussion of the first thorough evaluation of the paraffin test.

[74] Research discovered that many people who never fired a gun but whose profession, occupation, or happenstance brought them in contact with nitrates can be expected to yield positive reactions to the test. Among them are photographers, engravers, match workers, farmer[s] and gardener[s] handling fertilizers, etc. Other substances which may be expected to yield positive tests include bleaching agents, chemical[s], cosmetics, explosives, certain types of foodstuffs, tobacco, and urine.

A. MOENSSENS & F. INBAU, *supra* note 10, at § 4.12. In contrast, another scientist contends that the test is sufficiently reliable to be admitted into evidence. *See* Conrad, *supra* note 69, at 504-06.

[75] *See* Stone, *Evidence of Firearms Discharge Residues,* 33 BAYLOR L. REV. 285, 286 (1981).

[76] Giannelli, *supra* note 42, at 1225.

[77] Since evidence as to whether a defendant fired a gun might be crucial evidence in many murder trials, this unreliable evidence might also have helped lead to executions.

further a research scheme or in any practical way that would reveal their limitations. The reliability of such tests could only be assured if comprehensive experimentation were done to study that reliability. Therefore, before these forensic techniques are admitted into criminal trials, courts ought to require that such comprehensive reliability studies be presented.

III. The Burden of Proof on Reliability

Courts that have admitted the new blood tests have not required that the tests themselves be established as scientifically reliable. Instead, these courts have placed upon the defendants the burden of proving the tests unreliable—something extremely difficult for defendants to establish. This portion of the paper will advocate that in order to guarantee the admission of trustworthy evidence, the burden of proving the tests reliable should be placed on the prosecution.

A. *The Courts' Approaches*

Assertions that the procedures for the detection of genetic markers in dried blood are reliable or accepted by the scientific community are less valuable than such assertions about other scientific techniques. Unlike other procedures, these tests for genetic markers do not have their reliability examined through their constant use in a practical way or in any continuing research. Rather, the assertions about the worth of the procedures come from a small group of interested people. The history of the paraffin test indicates that concerns over reliability are not merely abstract concerns, but instead are factors which should trigger a close look at the fallibility of the tests before they are admitted into criminal cases. Unfortunately, the courts, in the few cases that have actually litigated the admissibility of the techniques,[78] have not recognized these considerations, much less the importance of them. A look at several cases illustrates this.

---

[78] See *supra* note 8 for a list of cases in which the blood grouping tests were introduced by the prosecution.

In *Robinson v. State*,[79] the defendant was convicted of assault with intent to murder. The victim was severely cut on the throat and face. The defendant, when arrested approximately an hour after the attack, had blood on his clothes. At trial, the court permitted Jean Hostetler, a forensic chemist from the Montgomery County Police Department,[80] to testify that she analyzed the victim's blood as well as the blood on defendant's trousers using traditional ABO tests and electrophoretic procedures.[81] The blood samples matched each other.

The *Robinson* appellate court stated that "the only significant objection" to the admission of the genetic marker evidence was that

> Ms. Hostetler's testimony only established that the use of electrophoresis and enzyme comparison for the purpose of classifying blood samples is an accepted practice in the field of forensic chemistry. She did not and could not, according to appellant, demonstrate that this technique is generally accepted in the broader scientific community.[82]

This objection did not matter. The court stated that Maryland's interpretations of the *Frye* rule did not require "that a scientific technique be deemed legally unreliable simply because that technique has obtained general acceptance in only one branch of science."[83] The court concluded further that the testimony gave "an adequate foundation for . . . the trial court's ruling that the technique is accepted as reliable within the field of forensic

---

[79] 47 Md. App. 558, 425 A.2d 211 (1981).

[80] *Id.* at 573, 425 A.2d at 219.

[81] The opinion does not indicate which genetic markers were classified by electrophoresis.

[82] 47 Md. App. at 575, 425 A.2d at 220. The court noted that on cross-examination, the state's witness stated

> that the electrophoretic technique at issue here is not typically used by hospital or other non-police labs. However, she explained that the reason for this is that enzyme structure does not affect blood compatibility for the purposes of transfusions and the like; therefore, aside from use in forensic and research labs, there is simply no need for the precision that this technique affords.

Id.

[83] *Id.* at 576, 425 A.2d at 220.

chemistry."[84]

Since the court seemed to accept the conclusion that these tests were used almost exclusively in police laboratories,[85] the court was effectively stating that if the scientists employed by the law enforcement agencies testify that the test is reliable, then the evidence should be admitted.[86] Not considered by the court was the quite special nature of such a field and whether that field's unique attributes might affect its conclusions about a scientific procedure.

In addition to finding forensic chemistry to be the relevant field, the *Robinson* court found two bases for the conclusion that that field accepted the tests as reliable. The affirmative assertions of the witness formed one support:

> As to the reliability of the described electrophoretic technique, Ms. Hostetler testified that it was developed in the late 60's and that it is now an accepted practice in the field of forensic chemistry. She testified further that she personally knew that electrophoresis was being utilized by the Montgomery County Police, the Baltimore Police, the Maryland State Police and the Federal Bureau of Investigation . . . .[87]

The lack of contradictory evidence by the defense was the second basis for the court's conclusion: "Ms. Hostetler's testimony was uncontroverted. No expert witness was called by the defense,

---

[84] *Id.* at 576, 425 A.2d at 221.

[85] *See supra* note 82.

[86] Compare the *Robinson* court's conclusion about the relevant field with People v. Barbara, 400 Mich. 352, 255 N.W.2d 171 (1977), in which the court rejected the admission of lie detector evidence. The Michigan Supreme Court in *People v. Barbara* analyzed other court decisions that found general acceptance of the lie detector. The court then concluded that:

> [I]n order to find such acceptance, these courts adopted what was in fact a logical fallacy.
>
> Under the present state of the art the general acceptance of the polygraph among psychologists and physiologists cannot be demonstrated, because such acceptance does not exist. Therefore, these courts, in order to find general acceptance, found it amongst polygraphers. Once finding general acceptance, the courts then found they did not have to rely on scientific testimony, but were able to rely on the testimony of polygraphers to establish reliability of the device.
>
> Such reasoning is circular . . . .

Id. at 390, 255 N.W.2d at 187.

[87] 47 Md. App. at 575, 425 A.2d at 220.

nor was any other evidence adduced to suggest that electrophoresis is regarded as either experimental or controversial."[88] The court never analyzed whether these two purported bases in fact tended to show that the tests were reliable. If it had, the court should have seen that this foundation was flawed.

While the witness stated that the tests were accepted by forensic chemists, her opinion gave no explanation of why it was so accepted. Significantly, nothing was cited to show that the tests were reliable. No reference was given to any of the relevant literature. The statement that the procedures were accepted by the forensic scientists was the equivalent of stating that many forensic laboratories use the procedures.[89]

Widespread use, however, does not necessarily equal reliability. Many forensic labs used the paraffin test, but that did not mean that the procedure was accurate. Such verification requires controlled experiments, and controlled experiments require time and money. Few forensic laboratories are in business to do such research. Most exist to give answers to law enforcement agencies about specific pieces of physical evidence.

A forensic procedure becomes widespread not because all the people using it have made independent evaluations about reliability. Instead, a few research scientists develop the procedure. If the new technique appears to work, then the methods are taught to others. Since few labs can afford to employ only highly trained scientists, often those learning the procedures are not scientists,

---

[88] *Id.* at 576, 425 A.2d at 220.

[89] The lack of evidence on these points might have resulted because the witness was not a scientist, but merely a technician. One commentator has concluded that technicians should be limited in their testimony:

> [A] technician's testimony should never suffice to establish the validity of a novel technique: "[T]he technician merely follows prescribed routines, and is not expected to understand their underlying fundamentals. He knows how, but not why." Kirk, *The Interrelationship of Law and Science,* 13 BUFFALO L. REV. 393, 394 (1964). Because it is critical to know the 'why,' or, as in the case of empirical validation, the implications of not knowing the 'why,' the views of scientists are essential. Moreover, a technician would not be qualified to testify about the general acceptability of a technique because presumably only a scientist would be sufficiently conversant with the views held by those in the relevant field.

Giannelli, *supra* note 42, at 1214-15 (footnotes omitted).

but technicians. Thus, most of those who use the new tests have not verified the test's reliability, and few would have the training to conduct such research in any event. The users of the procedure, then, trust that the procedure is reliable, not because they have verified that fact but because the developers of the procedure say that it is reliable. Widespread use of electrophoretic tests in forensic labs thus does not indicate anything more about reliability than that a handful of people have attested to their reliability.

The *Robinson* court also stressed that the defense presented no experts of its own or other evidence to show that electrophoresis "is regarded as either experimental or controversial" in justifying the procedures' admission into the criminal trial.[90] This is shoddy reasoning. As one commentator has stated about different, but also controversial scientific evidence, "Incredibly, several courts have cited the absence of opposing experts to support their decision to admit voiceprints, apparently inferring reliability from a lack of opposition. This inference is unwarranted."[91] The inference is wrong for several reasons. A defendant may not have the means to present such a rebuttal. Defendants often are poor and may not have access to laboratories and experts. The prosecution, however, does not face this handicap. States and most large metropolitan areas have government-operated forensic laboratories. Federal laboratories likewise provide services to local and state law enforcement agencies.[92] The FBI lab will do forensic tests for local agencies without charge and even provide for the court appearance of the expert.[93]

Even if the defense has money to spend on experts, it will encounter difficulties in finding a person willing to probe the evidence from the defense's viewpoint. The defense cannot just go to the local forensic serologist to discuss the issues. Very few people are familiar with this evidence, and, as discussed above, almost all are connected with law enforcement agencies. Under these circumstances, the lack of opposing evidence at trial may not indicate re-

---

[90] 47 Md. App. at 576, 425 A.2d at 220.
[91] Giannelli, *supra* note 42, at 1243.
[92] *Id.* at 1244.
[93] *Id.*

liability, but rather only the narrowness and one-sided composition of the field.[94]

The *Robinson* approach implicitly states that once the prosecution comes forward with some evidence of reliability, however feeble, then the burden of proof shifts to the defendant. In *Robinson* the prosecution's production burden was met merely by introducing evidence that the genetic marker tests are widely used in the forensic field. Since the defendant produced no evidence, it was reasoned that the tests should be admitted. Most important in this approach is that the prosecution has to present no evidence that the tests are reliable other than the fact that a small band of people say they are.

Reasoning like that in *Robinson* also held sway in *People v. Young*,[95] even though the Michigan court professed to be applying

---

[94] Of course, even if the defense does manage to have access to experts, the adversary system can only work properly if the defendant is given ample notice and discovery of the prosecution's test. *See* Giannelli, *supra* note 42, at 1240; Annot., *supra* note 9, at 509. A corollary duty should require the prosecution to preserve the evidence so that the defense can retest it. *See* Giannelli, *supra* note 42, at 1243. Thus, in People v. Nation, 26 Cal. 3d 169, 604 P.2d 1051, 161 Cal. Rptr. 299 (1980), the California Supreme Court held that the prosecution has the duty of preserving a semen sample recovered in a sexual assualt case so that the defense could later examine the sample for genetic markers. In many criminal cases, such preservation of forensic evidence, may, however, present a problem. All that might be recovered are mere flecks of dried blood. Nothing of the evidence may be left after the prosecution runs its tests. This problem is compounded if the electrophoresis has been run with a starch gel base since no permanent record is left of these results. The only way of preserving the starch gel results is through photographs. *See* Grunbaum, *supra* note 37, at 54. Fairness to a defendant thus should require that if the prosecution's electrophoretic tests left no evidence for an independent analysis, then the prosecution should produce clear and unambiguous photographs of its results so that an independent expert can at least examine those pictures to see if he would reach the same results as the prosecution's expert.

[95] 106 Mich. App. 323, 308 N.W.2d 194 (1981). *Robinson*, *Young*, and State v. Washington, 229 Kan. 47, 622 P.2d 986 (1981), which is discussed at *infra* notes 110-22, are the only reported cases that give any detailed reasoning supporting the admission of the tests for the detection of genetic markers in blood. In People v. Bush, 103 Ill. App. 3d 5, 430 N.E.2d 514 (5th Dist. 1982), the defendant challenged the use of the electrophoretic procedures, but did not make the basis of his challenge clear to the appellate court. The court merely concluded that, "[c]onsidering the qualifications of the expert, the fact that she was supervised by senior personnel, and the description of the tests used, we find that proper foundation was laid for the admission of the evidence." *Id.* at 13-14, 430 N.E.2d at 521. In Jenkins v. State, 156 Ga. App. 387, 274 S.E.2d 618 (1980), the defendant contended that the testimony concerning the genetic marker test should have been stricken, "basing his motion on a contention that the reliability of the procedure and the acceptance of the procedure by the scien-

an evidentiary standard different from the one used in the Maryland case. Young was convicted of felony murder. Part of the evidence against him was that an electrophoretic "analysis of defendant's blood revealed properties which corresponded positively to those of bloodstains found at the [murder] scene."[96] This evidence was presented by Mark Stolorow, "a forensic serologist working for the Michigan State Police . . . . [who] had been active in the research and development of electrophoresis as a technique for comparing blood samples."[97]

Prior to *Young*, the Michigan Supreme Court had put its own gloss on the *Frye* standard in *People v. Barbara*,[98] in which the court held that lie detector evidence was not admissible. "The *Barbara* court further indicated . . . that the scientific technique used must be accepted in the particular field in which it belongs and that such acceptance can be established only by testimony of disinterested and impartial experts . . . . "[99] The importance of the disinterested witness requirement was evident in *People v. Tobey*,[100] a case in which the Michigan Supreme Court held that voiceprint evidence should not have been admitted during the criminal trial. The court pointed to the backgrounds of the two witnesses presented to establish admissibility and concluded, "Neither Nash nor Tosi, whose reputations and careers have been built on their voiceprint work, can be said to be impartial or disinterested."[101]

Stolorow similarly would not seem to be disinterested in the de-

---

tific community had not been shown." *Id.* at 388, 274 S.E.2d at 619. The *Jenkins* court disagreed and simply concluded that:

> "The opinion of an expert on any question of science is always admissible." Breland v. State, 134 Ga. App. 259 (214 S.E.2d 186). We do not find that the fact that the procedure is relatively new requires that the testimony be excluded. The witness explained the procedure, thereby giving the facts on which his opinion was based. The question, then, is not of admissibility but of the weight to be given the evidence by the jury.

Id. at 388, 274 S.E.2d at 619.

[96] 106 Mich. App. at 325, 308 N.W.2d at 195.

[97] *Id.* at 325, 308 N.W.2d at 195.

[98] 400 Mich. 352, 255 N.W.2d 171 (1977).

[99] People v. Young, 106 Mich. App. 323, 327, 308 N.W.2d 194, 196 (1981).

[100] 401 Mich. 141, 257 N.W.2d 537 (1977).

[101] *Id.* at 146, 257 N.W.2d at 539.

tection of genetic markers in dried blood. The *Young* court, however, distinguished the Michigan precedents: "[B]oth *Tobey* and *Barbara* involved scientific devices the accuracy of which was hotly disputed. Thus, we do not view the *Tobey* case as extending the . . . *Frye* rule to apply to all cases in which scientific evidence is sought to be admitted."[102] The court thus concluded "that because, in the present case, defendant offered no evidence that the electrophoresis technique used to compare the blood samples was scientifically inaccurate and because defendant has not convinced us that the accuracy of electrophoresis is seriously disputed the trial court did not abuse its discretion in admitting Mark Stolorow's testimony. . . . "[103]

The impact of the *Young* decision is that the burden of proof on the issue of reliability again was given to the defendant, because the prosecution's evidence that shifted that burden of proof was the widespread employment of electrophoresis. Here, however, the witness did not concede a use limited just to forensic science. Instead, after explaining the principles of the procedures,

> Stolorow testified that the technique of electrophoresis is simple and dates back many years in the field of biology. He further testified that only recently had the technique been adapted for use with very small samples of bloodstains, but that the techniques developed were, at the time of trial, employed routinely in case work. . . . He further stated that electrophoresis is used in Red Cross clinics, in hospitals, at universities, in medical schools, and wherever studies in genetics are performed, as well as in crime laboratories, and that he personally knew of half a dozen people in Michigan, not associated with crime laboratories, who performed electrophoresis analysis on a routine basis.[104]

The court's summary of the testimony indicates that the judges were misled. The court appears to have equated use of electrophoresis in the places where genetic studies are done with the tests performed by Stolorow. The forensic serologist had grouped the ge-

---

[102] People v. Young, 106 Mich. App. 323, 328, 308 N.W.2d 194, 196 (1981).
[103] *Id.* at 329, 308 N.W.2d at 197.
[104] *Id.* at 326, 308 N.W.2d at 195-96.

netic markers in a bloodstain, not just in fresh blood. The court, however, never understood that there might be importance in that distinction. Instead, if the court had been interested in determining whether the procedure presented in court was reliable, it should have asked some questions after stating "that only recently had the technique been adapted for use with very small samples of bloodstains, but that the techniques developed were, at the time of trial, employed routinely in case work."[105]

If, for instance, the procedures for the bloodstains are really the same as for the other electrophoretic procedures, which are "simple and date back many years,"[106] why is it that the techniques for dried blood have only recently been developed? What is the "casework" in which the dried blood tests are routinely used? Do the Red Cross clinics, hospitals, universities, medical schools or other places doing genetic research use the recently developed techniques for aged blood? If not, why should it be believed that the experiences of others working with liquid, fresh blood can be extrapolated to the tests here with dried blood?

Such questions would have led the court to realize that differences exist between the tests performed by the forensic expert and other applications of electrophoresis. The court should have realized that the new procedures for the detection of the markers in dried blood are only routinely used in police case work. Thus, the use of electrophoresis by the Red Cross clinics and others was irrelevant to the determination of the admissibility question before the court.[107] Reaching this conclusion, the court then should have asked the most important question: Besides the evidence that the tests are used by forensic serology, what other evidence is there that the procedures for the grouping of genetic markers in dried blood are reliable?

---

[105] *Id.* at 326, 308 N.W.2d at 195.

[106] *Id.*

[107] The court certainly should have realized that just because a scientific test is reliable in some areas that does not necessarily mean it is reliable in all its applications. *See, e.g.*, United States v. Brown, 557 F.2d 541, 557 (6th Cir. 1977), holding that while ion microprobic analysis had gained acceptance in the field of mass spectrometry, no showing had been made that technique had been accepted as reliable for comparing hair samples. *See also* Giannelli, *supra* note 42, at 1213.

The *Young* case illustrates that unless the defendant can shoulder the burden of showing that the tests are unreliable, the fact that the dried blood procedures are widely used in forensic work guarantees their admissibility. Interestingly, the *Young* court concluded that this burden was not met because the "defendant offered no evidence that the electrophoresis technique used to compare the blood samples was scientifically inaccurate and because defendant has not convinced us that the accuracy of electrophoresis is seriously disputed . . . ."[108] Apparently, the defense was not aware of the evidence of any such dispute, although the prosecution's own expert clearly was. Nothing in the opinion, however, indicates that the prosecution ever brought this evidence to the court's attention.[109]

Stolorow's knowledge is clear because of his role in the case of *State v. Washington*.[110] Washington was convicted of first degree murder and rape. Evidence indicated that the assailant had bled at the crime scene. At trial, Eileen Burnau, "criminalist" for the Kansas Bureau of Investigation, testified that she had analyzed the victim's and the defendant's blood as well as the blood samples found at the scene for ABO type and six polymorphic enzyme systems.

---

[108] People v. Young, 106 Mich. App. 323, 329, 308 N.W.2d 194, 197 (1981).

[109] Giannelli has argued: "Because the proponent has the burden of proof on the general acceptance issue, the proponent should be responsible for informing the trial court of opposing views in the literature . . . ." Giannelli, *supra* note 42, at 1218. While this may be a noble aspiration, one might question whether this will happen, at least when the only available experts may be connected with law enforcement agents whose careers will be enhanced if the tests are found to be admissible in court. *Young* certainly indicates that it is unreasonable to expect the experts to volunteer opposing evidence on their own. Specific questions probing this possibility would seem to be necessary at least. Of course, only if there are other experts available who are familiar with scientific literature will it be possible to check out the accuracy of a prosecution witness' assertion that the scientific literature contains nothing to indicate unreliability. For instance, in a trial with which I am familiar, a well qualified forensic serologist testified about the genetic marker tests he had done on a bloodstain found on the defendant's clothes. On cross-examination he was specifically asked whether the drying of blood made it more difficult to detect genetic markers. The forensic serologist just flatly stated, "No." The defense attorney, who did not appear to be knowledgeable about these blood tests and presented no expert testimony of his own, presented nothing to contradict this incredible assertion. For a discussion of how blood changes outside the body, see *supra* text accompanying note 33 and *infra* text accompanying notes 136-40.

[110] 229 Kan. 47, 622 P.2d 986 (1981).

The blood of the victim did not match that of the defendant's, although some of the blood at the scene did match the defendant's blood.

The defense challenged this evidence by calling Dr. Benjamin W. Grunbaum, who, the court said, "possesses an impressive string of credentials, including bachelors, masters, and doctoral degrees in biochemistry, and a masters in criminology with a specialty in criminalistic identification," and had been employed as a research biochemist at the University of California at Berkeley for 27 years, specializing in analytical biochemistry and microanalysis, which includes the examination of body fluids.[111] At trial Grunbaum testified that the tests performed were not reliable:

> Specifically, Dr. Grunbaum complained that blood, once outside the body, is always deteriorating. He explained that the degradation was particularly acute in the EAP and GLO enzymes [two of the enzymes typed in the blood samples]. Because of the length of time and the fact that the blood samples were dry, Dr. Grunbaum testified that the test of the samples in this case could not be reliable.[112]

The genetic markers were detected by a method of electrophoresis called the Multi-System analysis. According to Burnau, "The Multi-System analysis is based on the same principles as the other methods of enzyme analysis, but merely combines the testing."[113] Grunbaum likewise challenged the reliability of this system, noting that the University of California originally had been involved with its development, but due to its poor results had withdrawn from the project. Grunbaum eventually challenged the final results as having been falsified. Evidence from the Law Enforcement Assistance Administration (LEAA) admitted at the *Washington* trial indicated that it had not found support for this claim. The LEAA concluded that the allegations of test result falsification were unconfirmed, but still decided not to publish the report. "Grunbaum also testified that, apart from use in crime laboratories, the Multi-System analysis was not accepted within the scientific

---

[111] *Id.* at 50, 622 P.2d at 989.

[112] *Id.*

[113] *Id.* at 58, 622 P.2d at 994.

community."[114]

In rebuttal, the state called Mark Stolorow, who testified that he, too, had been part of the project that developed the analysis. He said that the workers had split into two groups. Grunbaum's group did the electrophoresis with a cellulose acetate base while the other group used a starch gel base, the one used in the Multi-System Analysis. Stolorow testified that although the entire research group did EAP readings on both cellulose acetate and starch gel mediums, "Grunbaum's readings were so inaccurate and erratic that they were deleted from the report."[115]

In contradiction to Grunbaum, the rebuttal witness also went on to state that the EAP enzyme was less likely to deteriorate in dried than in liquid form. Although the *Washington* decision did not state upon what information Stolorow based this conclusion, the court observed that his analysis "correspond[ed] to the State's introduction of a booklet called *The Detectability of Selected Genetic Markers and Dry Blood Upon Aging* published by the National Institute of Law Enforcement & Criminal Justice . . . . "[116] Finally, Stolorow asserted that the Multi-System Analysis was reliable and was used in over one hundred crime laboratories throughout the United States and Canada, and that Burnau stated that the FBI was using it in its laboratory.

In contrast to Stolorow's statements, the court noted that it could find no reported cases concerning the admission of this evidence. The court concluded that the admissibility should be measured by the *Frye* standard and held that:

> We have concluded from the record that there was sufficient evidence of reliability and acceptance by the scientific community to justify admission of the expert testimony in this case. . . . We are impressed by the testimony that the Multi-System analysis is reliable and generally accepted in the scientific field as illustrated by its present use in over 100 criminal laboratories in this country and that the FBI re-

---

[114] *Id.* at 51, 622 P.2d at 990.
[115] *Id.* at 51-52, 622 P.2d at 990.
[116] *Id.* at 52, 622 P.2d at 990.

> search laboratory uses the Multi-System analysis routinely and approves it. It should be noted that the basis for Dr. Grunbaum's opinion that the Multi-System analysis was unreliable was two-fold (1) that his testing had not produced one accurate result, and (2) that the EAP enzyme is subject to rapid deterioration, thus making blood groupings based on EAP analysis unreliable. This testimony was countered by testimony that, in testing, Dr. Grumbaum used a different medium than is used in the [other] method. . . . As to EAP enzyme deterioration, the State produced two separate pieces of evidence showing the contrary—that EAP did not rapidly deteriorate, once dried.[117]

Washington is the only opinion discovered in which the court made some effort to delve into the actual reliability of the dried blood genetic marker tests, as well as the only case found in which the defense presented its own expert. Even so, the approach is still strikingly similar to that in *Robinson* and *Young;* therefore, so is the result. Once the prosecution establishes that the tests are widely used by forensic scientists, the tests will be admitted unless the defendant can convince the court that the procedures are unreliable. Since no defendant has carried this burden, the tests have always been admitted.

A hypothetical will clarify the point. Assume that the same evidence about reliability is produced as was presented in *Washington,* but that the evidence also showed that only the Kansas Bureau of Investigation, and no other lab, used these procedures. The court could not then so glibly pronounce that sufficient evidence of reliability had been produced. Rather, a battle of the experts would likely ensue. Under such circumstances, the evidence still would consist of one expert with "an impressive string of credentials" maintaining that the procedures are unreliable and giving reasons for that conclusion. On the other hand, another expert would explain why the developers of the procedure rejected Grunbaum's complaints. Based on such competing allegations, it is hard to see how the court could determine who was right.

Certainly, that the developer's report was not published, for

---

[117] *Id.* at 55, 622 P.2d at 992-93.

whatever reason, had more significance than the court recognized. Whether or not the decision to withhold publication indicated anything about the validity of the work, it certainly reflected on the general acceptance of the study. If the data had not been disseminated to the scientific community so that an independent examination and analysis of the findings could be made by any interested scientist, then the study could not be generally accepted by scientists.[118]

The defense in *Washington* also produced evidence that EAP deteriorated as it aged. The court rejected this contention, noting that the state produced two separate pieces of evidence "showing the contrary—that EAP did not rapidly deteriorate, once dried."[119] The court never analyzed the data produced by the two sides to determine if either of the differing conclusions was supportable or if one of the conclusions was clearly wrong. The court merely rejected the defense's argument by summing up the amount of evidence. If it really were making its decision this way, the court knocked a hole in the proposition that the quality of evidence rather than the quantity should decide an issue.

In its tabulation, the *Washington* court should have been more interested that only three pieces of evidence were brought to bear on the question than in the fact that one side outnumbered the other.[120] The court should have wondered whether reliability had

---

[118] This lack of publication should have led to other questions. How is it that, without the final report, so many labs have used the procedures? Have the one-hundred laboratories each made independent assessments of the procedure's validity? If so, based on what information? Certainly it cannot be the report because it has not been distributed. A plausible conjecture is that a small group of scientists have taught the others the procedures and told them that the tests are reliable. Under such circumstances, since the FBI lab teaches the techniques, *id.* at 52, 622 P.2d at 990-91, the widespread use of the test would indicate only that a few people, who happend to be good salesmen, claimed to have had a scientific basis for their conclusion about reliability.

[119] *Id.* at 55, 622 P.2d at 993.

[120] In any event, the court's counting of evidence may have been wrong. The court did not cite the sources for either Grunbaum's or Stolorow's EAP opinions. Grunbaum's probably came from his own work. See *infra* text accompanying note 236. Conversely, I have found no articles that Stolorow has published concerning EAP deterioration. The Multi-Systems Analysis, however, was developed jointly by Aerospace Corporation and Beckman Industries, and Stolorow worked for this Beckman/Aerospace project. *See* State v. Washington, 229 Kan. 47, 51, 662 P.2d 986, 990 (1981). The booklet cited by the court as contra-

been established if this small amount of research was truly the only information on the subject. The presentations in *Washington* indicate that the evidence on EAP deterioration is minimal, but contradictory. The fair conclusion from such scanty evidence should be that although the defense has not proven the tests unreliable because of the degradation of the dried blood, neither has the prosecution established that the test is reliable. Instead, the evidence indicates that reliability remains an open question.

Since the question of reliability was not resolved by the evidence, the party who had the burden of proof on that issue should lose. The defendant lost, and therefore, he must have had that burden. The reason the burden was placed on him is the same as in the other cases—because many forensic labs use the tests. As the court acknowledged, "We are impressed by the testimony that the Multi-System analysis is reliable and generally accepted in the scientific field as illustrated by its present use in over 100 criminal laboratories in this country and that the FBI research laboratory uses the Multi-System analysis routinely and approves it."[121]

In *Washington,* as in *Robinson* and *Young,* once the prosecution presented evidence that the tests are widely used by the law enforcement laboratories,[122] the burden shifted to the defendant to

---

dicting Grunbaum, which also was published as an article in the *Journal of Forensic Science,* was written by employees of the Aerospace Corporation. *See* Denault Study, *supra* note 32, at 479. Since no independent study by Stolorow has been found on the disputed point, it is not unreasonable to assume that he was relying on this other Aerospace research in giving his opinion. If so, it was not two pieces of evidence contradicting Grunbaum that were given, but rather only one piece under two different guises.

[121] 229 Kan. at 55, 622 P.2d at 992.

[122] The defendant raised an issue about the field of forensic science. The court characterized the argument in the following manner: "The defendant futher suggests that there is a difference between the field of forensic science and the field of biochemistry, which renders a forensic expert incompetent to testify on the reliability of blood analysis tests, in contradiction of the testimony of a biochemist." 229 Kan. at 55-56, 622 P.2d at 993. This argument apparently contended that forensic scientists do not have the training, qualifications, and experience to make judgments about reliability. The court rejected this contention:

> The area of forensic science is obviously made up of many different types of experts, as evidenced by Stolorow's qualifications in chemistry and forensic chemistry. Stolorow, as a forensic chemist, and Dr. Grunbaum, as a biochemist, both specialize in microanalysis and serology. Defendant's position that the only acceptance of the Multi-System analysis is among law enforcement officers un-

show that the genetic markers tests in dried blood were unreliable. Even though the defendant produced evidence of unreliability, that evidence was inconclusive. Therefore, the defendant did not meet the burden implicitly placed on him by the court, and the evidence was admitted.

B. *The Proper Burden of Proof*

The courts uniformly have indicated that the dried blood genetic marker tests are admissible unless the defendant carries the burden of showing that the tests are unreliable.[123] This approach is improper. Evidence is probative only if it is trustworthy. As one commentator observed, "For evidence to contribute to the truth-determining function of a trial, it must be reliable."[124] This maxim is equally applicable to scientific procedures: "The probative value of scientific evidence, however, is connected inextricably to its reliability; if the technique is not reliable, evidence derived from the technique is not relevant."[125] Reliability of the dried blood genetic

---

trained in the chemical analysis of blood and other body fluids, is not supported by the testimony, and we find this argument to be without merit.

Id. at 56, 622 P.2d at 993. This last sentence is not clear. The court could be stating that acceptance of this technique is not limited to law enforcement officers. Such a conclusion is not supported by the rest of the opinion, which indicates that the test is only used in forensic labs. More likely, the court's sentence is rejecting the contention that the law enforcement officers are incompetent to do genetic marker analyses. Of course, even if some forensic scientists have qualifications as acceptable as Stolorow's, that does not mean that people performing the test at all one hundred labs do. More important, however, is that the court never recognized the crucial importance of the fact that this technique is confined to forensic work; that is, even if done by competent people, the results are not used in any way that shows whether or not those results are reliable.

[123] Some approaches to scientific evidence that view the *Frye* standard as too stringent suggest that traditional evidentiary rules of relevancy should apply. *See* Giannelli, *supra* note 42, at 1232-45. Even if this were the approach adopted by a court presented with the introduction of the forensic genetic marker tests, the same considerations analyzed here would control because, "in effect, the relevancy approach places the burden on the party opposing admissibility. . . . Instead of the prosecution carrying a substantial burden of establishing the reliability of a novel scientific technique, the defendant must shoulder the burden of establishing unreliability." *Id.* at 1246.

[124] *Id.* at 1200.

[125] *Id.* at 1235. *See also* United States v. Brown, 557 F.2d 541 (6th Cir. 1977), in which the court equated the *Frye* standard "with a showing that the scientific principles and procedures on which expert testimony is based are reliable and sufficiently accurate." *Id.* at 556.

marker tests should be the key concern in determining their admissibility. The only way the reliability of the dried blood techniques can be established is through thorough validation studies. Just as universal use of the paraffin tests did not assure reliability, it should be apparent that widespread use of the dried blood tests is no guarantor of dependability.

Since use by forensic scientists does not guarantee reliability, the only way anyone can determine if the tests work is if controlled experiments are conducted seeking to verify reliability under the conditions that will be met in actual casework. It might be argued that this has already been done. A scientist, for instance, may develop a procedure that detects and groups a polymorphic genetic marker. He then may modify this technique to use it on blood which he has dried overnight on a swatch of cotton sheeting. After repeating his work many times, he may announce that he can detect this genetic marker in dried blood. Does this establish the reliability of his procedure? Perhaps, but only for blood dried on a cotton sheet overnight under laboratory conditions. It does not establish that the technique is reliable for blood dried for one week, or two weeks, or months. It does not establish that the test works for blood dried on denim, or leather, or polyester. It does not show that it is reliable for blood dried at high humidity, or perhaps very low humidity, or at high or low temperatures. Perhaps it matters if the blood dried under intense light or no light; perhaps it matters if the cloth was contaminated with substances in addition to the blood. Such conditions all could be encountered in bloodstains important to a criminal case. More important, no one may ever know under what conditions blood taken from such a place as the defendant's clothes has aged and dried.[126] Only when experiments have been done exploring the many permutations of these possible aging conditions can the analysis of the bloodstain in a criminal case be certified as reliable.

Also, it is important to ascertain not merely whether the genetic marker can still be detected after being dried in a certain manner,

---

[126] *See* A. MOENNSENS & F. INBAU, *supra* note 10, at § 6.12 ("Exact determination of blood age is impossible.").

but whether the reading of the detected group has changed. This requires experiments that first reliably separate the fresh blood samples into groups according to different genetic markers, and then determine that subsequent readings taken at various stages of drying show the same genetic markers in the separate groups. Thus, the mere conclusion that the genetic marker was always found in blood dried for ten days is meaningless. Not only must it be found, it must be found in the same group as it was originally detected. If for some reason after drying, the blood group the genetic marker is found in appears to be different from the one in which it was originally found, then the test is unreliable.

If the burden of proving reliability is on the defendant, he would have to present the results of the verification experiments to the court. If sufficient studies did not exist, he would have to have them done or lose on this issue. Of course, it is not likely that a defendant could carry such a burden. Such scientific scrutiny takes a continuing commitment of time and money from trained experts. Laboratories and equipment are necessary. Time is needed to publish the results and disseminate them so that other scientists can analyze the methods, data, and conclusions as well as to see if the work can be replicated. It is nonsensical to expect a defendant either to perform such work or have it performed. For practical purposes, all a defendant can do to prove unreliability is to survey the published literature to see what verification studies have been done. If those studies are deemed insufficient or inconclusive by the court, the defendant will be unable to meet his burden. If the existing research has not been thorough, the disputed tests may be unreliable. Even so, the procedures would still be admitted.[127]

---

[127] For the same reason, the evidence should not be admitted with questions about reliability left to the jurors in their assessment of the weight to be given to the scientific procedure. This would presuppose an adversary system in which each side can bring all the information before the jury. Here, however, the defendant is at the mercy of the existing scientific literature. Even if the tests are unreliable, he will not be able to present that evidence to the jurors unless the appropriate studies have already been published. If the research has not yet been done, the defendant has no way of producing evidence of unreliability. Thus, the material for a determination about dependability cannot be brought before the jury and, consequently, that body cannot be expected to give accurate weight to the scientific evidence.

Instead, since reliability of the dried blood genetic marker tests should be the concern of the court, the court must place the burden of proving reliability upon the prosecution. If there are insufficient studies now, change will occur only when the prosecution has the incentive to do more. Only the prosecution, through its forensic labs and its ability to fund research, can assure the continuing commitments that valid work is done. Law enforcers, of course, have a stake in having the genetic marker tests admitted into court, but the current legal situation is such that they can get them admitted without having to establish reliability. Therefore, the research to establish reliability is at present an unnecessary waste of time and money for the forensic laboratories. The only way prosecutors will have the incentive to do the verification experiments and the only way to assure that reliable evidence is being introduced at criminal trials is for the courts to insist that the prosecution do more than merely present evidence that the dried blood genetic marker tests are used in many forensic labs. Instead, the courts must also impose upon the prosecution the burden of establishing that the tests are actually reliable.

IV. The Reliability of Genetic Marker Testing

The crucial concern for admitting the new blood procedures into criminal cases should be whether they have been proven reliable. A survey of the scientific literature indicates, however, that science has done little to verify the reliability of the tests. Instead, the scientists have concentrated primarily on perfecting the techniques to identify the markers with little concern shown for proving the reliability of these techniques under practical conditions. Conversely, as this portion of the paper will show, the little research which has been conducted regarding the reliability issue indicates not only that the tests have not been proven reliable, but also that the research has repeatedly demonstrated that genetic marker tests are unreliable.

The research concerning the detection of genetic markers in dried blood contains few comprehensive studies on how various aging factors affect the genetic markers. Instead, the researchers have concentrated on making the tests practical. Genetic marker testing

in dried blood will be of little value to law enforcement if the procedures are cumbersome, complex, and can only be done by highly trained scientists, because police laboratories are often staffed primarily by technicians. A major goal of the research thus has been the "[d]evelopment of a simple, rapid, reliable and economical technology for blood and bloodstain analysis for forensic uses. . . . "[128] Almost all of the research has concerned the improvement of measuring techniques. Noting the lack of comprehensive research, one author, who has surveyed the literature on the detection of genetic markers in dried blood, has observed that, "Forensic serological research has indeed been limited, and it has been concerned primarily with methodology development."[129]

---

[128] Grunbaum, *Foreword* to HANDBOOK FOR FORENSIC INDIVIDUALIZATION OF HUMAN BLOOD AND BLOODSTAINS v (B. Grunbaum ed. 1981). Other objectives were also set to make genetic marker testing practical for the forensic scientist. Some listed goals of Grunbaum's work were the "establishment of reliable data on a number of genetic markers in human blood, on a statistically valid sample representative of the four major ethnic groups in the California population," and "[t]he communication of project accomplishment from the academic research laboratory so that the information may be put to practical use for the benefit of civil and criminal justice." *Id.* Valid studies of the frequency with which genetic markers appear is essential for forensic purposes. "The usefulness of these determinations in the criminal justice system is dependent upon a knowledge of the frequency of occurrence of these genetic factors in a general population or in a specified subpopulation." Grunbaum, *supra* note 2, at 3. It would mean little if the genetic markers in a bloodstain matched a defendant's or crime victim's blood if half the population had those same markers. See *supra* note 20, for an illustration of what might happen if 50% of the population had the same typing as the defendant. The forensic scientist, consequently, must know how frequently particular markers appear in the population. Similarly, the research has little practical value unless it is successfully disseminated to the practicing forensic serologists. The goal of this research is to do the work necessary so that genetic marker testing can go into effect in the police laboratories. Nowhere, however, is the goal of comprehensively checking reliability under all practical conditions mentioned.

[129] Zajac, *supra* note 17, at 160. It does not follow that just because most of the research has concerned methodology that the methods are now foolproof. The resulting product of electrophoresis is not objective, but rather one that requires interpretation. Therefore, "the degree of skill and information of the analyst who 'reads' the pattern on the electrophoretogram" is important. Grunbaum, *supra* note 37, at 54. Grunbaum has concluded: "As a safeguard against possible error, it is advisable for determinations on a single sample to be made by two analysts working independently. If that is not possible, independent readings of all electrophoretograms by two analysts may serve to confirm the readings." *Id.* at 103. *See also* Denault Study, supra note 32, at 493. The authors therein stated that: "In the actual practice of serological evidence analysis, however, the importance of conducting tests at least in duplicate and sometimes in triplicate, and with proper controls, cannot be overemphasized." *Id.* Grunbaum has also stated that even though improved methodology has eliminated many problems, "blood analysis will never be an error-free

The production of apparently workable techniques, however, is only the first step towards reliability. Once techniques are perfected, controlled experiments seeking to verify dependability must be performed. These experiments should test the procedure's limitations and search to find out when, if ever, the methods seem to work but in fact give wrong results. Since the forensic genetic marker tests will not be used in a way that will reveal any errors, such studies will be the only way to verify reliability. If reliability is to be proven, this research ought to be especially thorough and comprehensive. Unfortunately, almost no such research has been done.

A. *Persistence Research; Warning Signals About Reliability*

Once the methodology was developed so that it became practical for law enforcement use, research virtually ceased. However, although not designed to probe directly the areas of potential unreliability, some off-shoots of the methodology research have bearing on the issue. For example, after some genetic markers were discovered in dried blood, follow-up work was done to discover how long the markers persist in the bloodstains. This information is important to the practicing forensic serologist who will not want to spend the time, effort, and money doing tests if it is certain that he will not get meaningful results. One such study noted that, "The purpose of this detectability study is to provide the practicing criminalist with adequate data to permit the selection of those blood genetic marker systems for serological analysis that have a high probability of yielding significant results."[130]

Persistence research, however, has produced conflicting conclusions. A 1980 review of the scientific literature stated that: "Detectability of antigens and enzymes in dried blood ranges from a few days to years, depending on the conditions of specimen exposure and on the particular genetic marker system being tested. Even for a single system there is little agreement among investigators on time limits for variant detectability."[131] These researchers

---

'cookbook' technology." Grunbaum, *supra* note 37, at 52.

[130] Denault Study, *supra* note 32, at 479.

[131] *Id.* at 480.

also noted that the current investigation was initiated "because of the large discrepancies in available data and because of the paucity of information on the effects of materials and humidity on the survival of genetic markers in bloodstains."[132] They further concluded that, "This is one of the few comprehensive detectability studies that takes into account, for a large number of genetic marker systems, the factors of bloodstain history, that is, the material on which the blood was deposited and the humidity at which it was stored."[133]

Despite efforts to make their study detailed, however, these scientists recognized that their data was extremely limited and that most of the research in this area has yet to be done:

> The information in this report will be useful to the forensic serologist, who must decide for each bloodstain problem the genetic markers to test. However, emphasis must be placed on the limitations of this study. *It is intended as a starting point for future research.* It covers a limited number of genetic markers and marker systems, aging conditions, analytical techniques, and investigative skills. Moreover, the tests were conducted on clean specimens free of impurities. It is realized that in actual practice serological evidence preserved under known and constant conditions is rare, and the specimens may be contaminated with impurities such as perspiration, urine, soil, and bacteria. These factors limit the application of the results of the study.[134]

In spite of this declaration, once funding for this specific study stopped, these scientists discontinued their research in this area.[135] Furthermore, the scientific literature does not indicate that anyone

---

[132] *Id.*

[133] *Id.* at 479-80.

[134] *Id.* at 496 (emphasis added).

[135] In a letter dated July 21, 1981, from Genevieve Denault to this author, Denault stated that the study had been done under contract to the Law Enforcement Assistance Administration (LEAA). When that contract ended, the scientists shifted their research to areas concerning the Air Force, including propellant stability and radiation effects in space. The LEAA funding for the genetic marker research was consistent with its practices in other areas by not funding verification studies. As Giannelli has concluded: "The LEAA can be criticized for failing to establish such an evaluative process as part of its procedures." Giannelli, *supra* note 42, at 1232.

else is conducting similar comprehensive research into the persistence of genetic markers in dried blood.

As limited as it has been, persistence research is crucial because it contains the seeds of reliability studies. Of course, the trustworthiness of the evidence presented in court does not depend on how long a scientist can detect a genetic marker in dried blood. If after a certain time the scientist gets no reading, then no evidence would be presented at a trial. Instead, reliability depends not only on whether a phenotype can be detected in the bloodstain, but also on whether the phenotype reading in the stain is the same as that found in fresh blood from the same person. Reliability is not established from the mere fact that a genetic marker system can be detected in aged blood; instead, the crucial question is whether the marker detected in aged blood is a reliable indication of that found in fresh blood from the same person.

The limited data available on this issue is startling and disturbing. As the following discussion will show, research has established that the forensic serologist testing aged blood will be able to "read" the genetic markers in a sample, but the readings may be different from those in fresh blood because of changes in the blood. In other words, *the blood typing will be wrong.* Not only will it be wrong, but the forensic serologist will have no way of knowing that a "false positive" has occurred. To date, almost nothing is known about why, when, or how often such alterations occur.

Blood changes quickly outside the body. "Blood, as an extremely perishable biological substance, alters from the moment that it leaves the body, or, if drawn from a cadaver, from the moment of death. The alteration may be fast or slow, depending on environmental conditions and possible exposure to chemical contamination."[136] The changes can be of different sorts. Sometimes the deterioration of the blood means that no genetic markers are observed; however, "[t]entative experimentation has indicated that

---

[136] Grunbaum, *supra* note 2, at 2. *See also* BLOOD GROUP, *supra* note 11, at 408; Denault Study, *supra* note 32, at 494; Rothwell, *The Effect of Storage upon the Activity of Phophoglucomotase and Adenylate Kinase Enzymes in Blood Samples and Bloodstains*, 10 MED. SCI. & L. 230 (1970); Zajac, *supra* note 17, at 1605.

these markers may change in a manner which may give false readings rather than no reading at all."[137] Thus, reliability is not proven by experiments that indicate a genetic marker is present in dried blood.

> While experiments with laboratory-prepared bloodstains indicate that blood group factors persist in aged samples, there is no guarantee that these markers will remain unchanged in evidential material. Phenotypes in some contaminated, dried, or putrified blood may be in an altered and degraded state *which gives an unambiguous but false reading.*[138]

Laboratory work is of limited value because it has seldom approached field conditions:

> Most of the experimental work has been done using dried bloodstains prepared in the laboratory under "ideal" conditions and with blood samples which usually contain an anticoagulant and/or preservative. Consequently, numerous reports cite "no problems" or "no mistypings" in "blind trials". However, recent publications are reporting definite alterations in apparent phenotypes in bloodstains and degraded samples for some of the enzyme and protein systems.[139]

No technique now exists to change the unambiguous, but false reading to the correct one, and no reason exists to believe that such procedures will ever be developed: "To date, we have no evidence that these changes can be reversed. *Improvements and changes in electrophoretic methodology will be to no avail if the determiner enzyme/protein has altered through aging and deterioration.*"[140]

---

[137] Zajac, *supra* note 17, at 160. *See also* BLOOD GROUP, *supra* note 11, at 408; B. CULLIFORD, *supra* note 3, at 142.

[138] Zajac, *supra* note 17, at 163 (emphasis added).

[139] *Id.* at 167.

[140] Grunbaum, *supra* note 37, at 103 (emphasis in original). Although an analogy between fingerprints and genetic markers may sound appealing, the differences in the detectability of any alterations constitute an important reason why the comparison fails. "After a fingerprint or partial print has been found and photographed, the expert can be quite certain of what he has to work with and he may proceed with an individualization based on visual examination of morphological details. If the fingerprint pattern is in any way physically altered prior to photographing, this alteration is apparent." Grunbaum, *supra* note 2, at 2.

So far, scientists have hardly studied this crucial area of the false positives.[141] Not surprisingly, therefore, "to date little is known regarding many of the enzyme/protein systems in terms of stability, degradation, and reliability of phenotyping stains and samples of questionable history."[142] The safest conclusion is that much research is still needed,[143] but it can now also be concluded that genetic markers sometimes change because "proteins and enzymes degrade in unpredictable ways."[144]

A closer look at the scientific findings is in order. It will be especially instructive first to examine the evidence about the reliability of the detection by forensic scientists of the ABO groups and other antigens. The antigen procedures have existed longer than the electrophoretic ones. Consequently, the reliability of the antigen procedures have been more fully explored than the reliability of the newer tests. It is especially disturbing, then, to find out how little is actually known about the typing of antigens in dried blood.

B. *Reliability of ABO Typing in Dried Blood*

The reliability of ABO detection in bloodstains seems certain to forensic scientists. One scientist reports that, "The ABO system appears to be more persistent and stable than other antigen, enzyme or protein systems, and the procedures are more simple and straightforward than for the other antigens."[145] The group that has made the most thorough study of how aging conditions affect the persistence of genetic markers concludes that, "The evidence implies no significant substratum, humidity, or temperature effects on the persistence of ABO antigens, at least up to 26 weeks."[146]

---

[141] *See* Zajac, *supra* note 17, at 160.

[142] *Id.* at 167.

[143] *See* Baxter, *supra* note 2, at 292.

[144] Grunbaum, *supra* note 37, at 103.

[145] Zajac, *supra* note 17, at 164. *See also* BLOOD GROUP, *supra* note 11, at 408.

[146] Denault Study, *supra* note 32, at 485. The substratum means the material upon which the blood dried. In the Denault study, those materials were "glass, wool, nylon, and three types of cotton (plain cotton, permanent press, and denim)." *Id.* at 480. Other scientists "have reported A and B antigens to be detectable on bloodstained fabric at ambient conditions after two years. They also successfully typed a 34-year-old Type A stain. Their analyses were also conducted with the absorption-elution method on stains prepared on different fabrics substrata that included cotton, wool, silk, and synthetic materials such as

Thus, one of the pioneers in the study of phenotyping in dried blood has concluded that, "In experienced hands, the grouping of dried bloodstains in this [ABO] system presents relatively few problems."[147]

As we shall see, however, even ABO detection is beset with pitfalls. For one thing, trained personnel working carefully are essential.[148] Careless laboratory work can lead to wrong ABO groupings in bloodstains—that is, to false positives. For instance, "too large a sample or thread for antigen determination increases the chances of nonspecific absorption of antiserum and subsequent insufficient washing of the unabsorbed antisera. The consequence is a false positive reaction."[149] Wrong groupings may also result from the antiserum. "Antiserum that is too concentrated may produce agglutination in negative controls, and antiserum that is too dilute may result in insufficient absorption and subsequent false negatives."[150] The forensic serologist "must run casework stains in duplicate under rigidly controlled conditions with appropriate standards and controls."[151]

False ABO positives, however, are perhaps most often caused by contaminations from other secretions. The ABO antigens appear not only in blood, but also in other body fluids in most people:

---

nylon and rayon." *Id.* at 485 (summarizing Fiori, Marigo, & Benciolini, *Modified Absorption Elution Method of Siracusa for ABO and MN Grouping of Bloodstains*, 8 J. FORENSIC SCI. 419 (1963)).

[147] B. CULLIFORD, *supra* note 3, at 75.

[148] These caveats also apply to the grouping of fresh blood:

> The fact that blood typing has been routinely performed in the armed forces and in blood banks has led to a misconception as to the ease with which such tests may be made. It has been estimated that approximately 10% of the ABO groupings entered on the "dogtags" of World War II servicemen were erroneous. Only specially trained technicians are qualified to carry out blood tests that can be relied on.

2 AM JUR. PROOF OF FACTS, *Blood Types* 608 (1959). It is reported, however, that "errors in blood grouping today are rare, and when they do occur, can invariably be related to a clerical error, an error in identification rather than to an error in grouping at the laboratory bench." Baird, *supra* note 1, at 89.

[149] Zajac, *supra* note 17, at 165.

[150] Denault Study, *supra* note 32, at 481. Zajac similarly reports that "false positive reactions may result with the absorption-elution procedure when an antiserum of too high a titer is used." Zajac, *supra* note 17, at 165.

[151] Zajac, *supra* note 17, at 165.

> Approximately 80-85% of the population, known as secretors, have blood group substances in their saliva, tears, perspiration, semen, vaginal fluids, mucus, gastric contents, etc. The quantity of blood group antigen in semen and saliva, for example, is much greater than that found in red blood cells. . . . The remaining portion of the population do not secrete these blood group substances.[152]

Mistyping of the blood can occur if an apparent bloodstain is actually a mixed stain of dried blood and another fluid from a secretor.[153] In one study, for example, urine and perspiration from a secretor of the B type was put over bloodstains of different groups. The subsequent analyses discovered that

> both urine and perspiration from a known type B secretor caused [the] three type O donors to react as though these bloods were type B. The effect of the urine was more pronounced than the effect of perspiration. With type A donors the effect of B secretor urine was to give an absorption effect as would be obtained with an AB stain. With both known AB stains there was no observable effect of either contamination.[154]

The scientific investigators concluded that, "The effect of contamination from body secretions of a secretor donor was exactly the same as if his blood were added to the stain. . . . [S]uch contamination by secretor fluids can produce serious error in absorption tests."[155] They concluded that special care must be given in report-

---

[152] A. MOENSSENS & F. INBAU, *supra* note 10, at § 6.11. *See also* B. CULLIFORD, *supra* note 3, at 98. Culliford states that: "Classification into secretor and non-secretor is usually a very simple matter, but occasionally it is difficult to assess. There is also evidence that some individuals secrete in one body fluid and not another." *Id.* In addition, he warns, a small percentage of people show a different group in their secretions than in their blood. *Id.*

[153] *See* BLOOD GROUP, *supra* note 11, at 410. "False conclusions may be reached if the possibility of a mixture is not realized." *Id. See also* B. CULLIFORD, *supra* note 3, at 75; Baxter, *supra* note 2, at 295; Zajac, *supra* note 17, at 164-65.

[154] Marsters & Schlein, *Factors Affecting the Deterioration of Dried Bloodstains*, 3 J. FORENSIC SCI. 288, 297-98 (1958). In another study, type A bloodstains were contaminated with type B sweat and type B bloodstains were contaminated with type A sweat. The subsequent grouping found both type A and type B in most of the stains. *See* Aye, *The Reliability of ABO Grouping of Bloodstains Contaminated with Sweat*, 18 J. FORENSIC SCI. SOC'Y 193 (1978).

[155] Marsters & Schlein, *supra* note 154, at 299.

ing a bloodstain as type AB, the rarest of the ABO groups,[156] for it might be the combination of blood and some other fluid.[157]

Because of the possibility of such contamination, "the necessity for adequate controls from an unstained portion of the item cannot be over-emphasized. These should be taken from as near as possible to the bloodstain."[158]

> Usually, a washing stage is incorporated in this method. Because stains are ordinarily encountered on fabric, it is necessary to test unstained material in order to establish the presence or absence of contaminants and to determine whether the washing has been complete.[159]

If the control samples show no antigenic activity or only a weak reaction, the serologist may be able safely to type the stain. If the control samples show a strong reaction, however, the typing of the bloodstain should not be considered reliable.[160] Consequently, ABO groupings of bloodstains not done in conjunction with suitable controls should be considered untrustworthy.

---

[156] *See* A. MOENSSENS & F. INBAU, *supra* note 10, at § 6.11 ("Roughly 40 percent of the population in the United States is type A, 43 percent is type O, 14 percent is type B and 3 percent is type AB."). *Id.*

[157] A report of type AB blood in a stain may also be considered suspect for another reason:

> [T]here is difficulty in the interpretation of negative results. Such negative readings may indeed mean that the particular factor was never present in the blood, but negative results may also be caused by denaturation of the bloodstain, deterioration or masking of the reaction, or insufficient antigen. Consequently, the forensic analyst working with aged or dried blood specimens should regard only positive results as significant.

Zajac, *supra* note 17, at 166. This must be considered if the "Lattes method" is used to type the dried blood:

> The disadvantage of the method is that the agglutinins, ie. [sic], the antibodies, are not stable. They are very quickly altered or denatured in dried stains and no longer react. Also a group AB bloodstain does not contain anti-A or anti-B antibodies and hence when group A or B indicator red cells bearing the antigens are brought into contact, no reaction, ie [sic], agglutination, will take place. It is dangerous to report the group of a bloodstain, however fresh, based on negative findings.

Baird, *supra* note 1, at 101.

[158] B. CULLIFORD, *supra* note 3, at 75.

[159] Baxter, *supra* note 2, at 294.

[160] *See* Aye, *supra* note 154, at 195.

According to at least one forensic serologist, controls are also necessary because of possible contamination from sources other than body secretions:

> False positive ABO antigen reactions can also result from contamination with such substances as molds, bacteria, household dusts, detergents, and the decomposition of the samples, and from such substrates as leather. The effects of such substances on a bloodstain are unpredictable. Loss, addition and substitution of antigenic activity upon decomposition of the sample have been observed and reported. . . . It is essential to run unstained portions of the substrate, taken near a bloodstained area, and to use the Lattes test in conjunction with the absorption-elution tests for confirmation of the ABO phenotypes.[161]

This passage should be considered carefully. If the role of the legal system is to be more than just a passive recipient of the scientist's assertions, questions certainly should be asked at this point. First, conditions that can produce false results are noted. Then it is stated that it is crucial to run control samples and to run various tests. However, how will these steps cure or prevent the problems? Nothing indicates that the ABO antigens are contained in molds, bacteria, household dusts, detergents, or leather. If those substances sometimes do produce false positives as the scientific literature states, it is not because, like the secretions, they are adding their own antigens to the dried blood. Instead, it seems plausible that these contaminants produce erroneous readings because they actually change the antigenic activity of the blood. If that is the case, it is hard to see how running a control on a piece of fabric without the stain will show that the dried blood elsewhere on the fabric has changed. Since the effects of these contaminants are "unpredictable," and since no reason is presented to believe that the wrong results can be detected, how then are we to know if the ABO typing of dried blood is correct?

---

[161] Zajac, *supra* note 17, at 165. Concerning the same area, Culliford merely delivers this cryptic comment without further elucidation: "The possibility of obtaining false positive [ABO] reactions caused by bacterial contamination cannot be ignored but, in practice, this is not a problem." B. CULLIFORD, *supra* note 3, at 75.

Furthermore, why should the control detect the "addition and substitution of antigen activity upon decomposition" of the blood? Although the Zajac passage quoted above indicates that decomposed blood can lead to false ABO typing, it does not suggest any way to detect or prevent that change. Since blood starts decomposing the moment it is shed,[162] does the quoted statement indicate that the typing of all ABO stains is potentially unreliable? Do we know what aging conditions bring on these false positives or how often the wrong results occur? Zajac offers her observations that contamination can produce wrong readings and that decomposition has been observed to produce substitutions without explanation. These are crucial questions, but no answers are given.

One major study indicates that far from being idle concerns, these issues should be considered important by both forensic scientists and lawyers. In *Detectability of Selected Genetic Markers in Dried Blood on Aging* (Denault Study),[163] the authors tested the detectability of several genetic markers, including the ABO system, by depositing blood on six kinds of material, aging some of the resultant samples at room temperature of 23° C. and sixty-six percent humidity, and a few at -20° C. Examinations of the stains were then made at four, thirteen, and twenty-six weeks. The scientists concluded that "the evidence implies no significant substratum, humidity, or temperature effects on the persistence of the ABO antigens, at least up to 26 weeks."[164] A close look at their data, however, once again raises questions.

First, the data is very limited. Can authoritative conclusions be drawn from the few samples tested? Through thirteen weeks, twenty room-temperature-dried specimens were examined for the ABO antigens. Are a score of samples truly sufficient to prove reliability? If each permutation of the tested conditions is considered separately, the numbers are miniscule. For example, only two nylon samples dried at sixty-six percent humidity were tested. Does

---

[162] See *supra* text accompanying notes 135-36 for a discussion of the effect of deterioration on the reliability of dried blood testing.

[163] Denault Study, *supra* note 32. See *supra* note 146 for a discussion of the materials used in the Denault Study.

[164] Denault Study, *supra* note 32, at 485.

the conclusion that no problems were found in those two tests establish that no problems will be found in one hundred or more samples? Even if the conclusions about the tested conditions can be correctly drawn from the limited data, do the same conclusions apply to nontested situations? Are ABO antigens reliably detected in a stain found on a sneaker, for example?[165]

Consider another situation. Suppose a women is stabbed to death and two weeks later a man is arrested. A bloodstain is found on his workshirt, and analysis reveals it is type A blood. The victim had type A, while the defendant has type B. The defendant states, however, that the stain was acquired several months ago when he cut himself on a construction job. Since then he has worn the shirt nearly daily and laundered it periodically. If he is telling the truth, the stain has aged part of almost every day at the temperature of his working body; it has been repeatedly soaked with perspiration and dirt;[166] it has been dried at room temperature at night; and it has been subjected several times to household detergents.[167] If the stain had been subjected to these not far-fetched conditions, does the data of this study truly indicate anything about whether the typing of it would be without error?

A closer look at this data indicates not only its limited nature, but also that the scientists' conclusions are not always supported by their own reported results. At twenty-six weeks, eighteen samples dried at room temperature were examined for ABO antigens.[168] The investigators reported false positives in four of these samples,[169] meaning that they "found" an antigen which was the

---

[165] *See* State v. Fulton, 299 N.C. 491, 263 S.E.2d 608 (1980) (evidence crucial to case was bloodstain found on tennis shoe).

[166] One scientist knowledgeable about dried blood has concluded that heat and humidity, especially in combination, adversely affect ABO detection. *See* Outteridge, *The Biological Individuality of Dried Bloodstains*, 5 J. FORENSIC SCI. SOC'Y 22, 28-29 (1965). Several scientists suggest that "the blood in general should be preserved in a dry environment. This conclusion confirms the experience of criminalists, although no study attempting to verify this has been reported." Denault Study, *supra* note 32, at 487.

[167] See *supra* text accompanying notes 155-57 for a discussion of potential sources of contamination of bloodstains.

[168] *See* Denault Study, *supra* note 32, at 487, Figure 1.

[169] *Id.* at 493, Table 6.

wrong one more than twenty percent of the time.[170] Since this study was done under laboratory conditions, actual forensic samples could be expected to present more problems.[171] This could mean that the twenty-two percent false positive rate is a lower limit, with the actual rate of false blood group identification in casework to be even higher.

And yet, the Denault Study concluded that they found no substratum effect on the ABO system. Of the eighteen samples tested at twenty-six weeks, four were bloodstains on denim. Of those four denim samples, three gave false positive readings.[172] As an explanation for this seventy-five percent error rate, the authors stated that, "Denim regularly unraveled upon being washed, and the attempt to avoid this may have induced insufficient washing."[173] This proposition was not tested, and no solution to prevent these errors was proposed. This data, instead of proving that no problems exist with the persistence of the ABO antigens up to twenty-six weeks, actually indicates considerable unreliability at that time. Clearly, the results show that the typing on denim at twenty-six weeks lacks reliability.

Denim is hardly an unpopular fabric, and actual casework can be expected frequently to encounter stains on it. Because these tests have shown a problem with denim, one should expect to find further research into this area. For example, the false positives were first found at twenty-six weeks, not at four or thirteen. Does that mean that the problem only occurs after three months, or were earlier problems just missed because of the limited number of denim samples (four)? Were the errors really caused by insufficient washing (if so, why weren't false positives found on earlier tests

---

[170] A different study reports proficiency testing that shows a 1.6% error rate for ABO typing. The details of this testing, however, were not given, so these results cannot be compared with the data produced by the authors of the Denault Study. *See* Shaler, Hagins, & Mortimer, *MN Determinations in Bloodstains—Selective Destruction of Cross-Reacting Activity*, 23 J. FORENSIC SCI. 570 (1978) [hereinafter cited as *MN Determinations*].

[171] *See supra* text accompanying note 133 for a discussion of the laboratory conditions which limited the practical applicability of the Denault Study.

[172] The fourth was a bloodstain on nylon aged at 20% humidity. *See* Denault Study, *supra* note 32, at 493.

[173] *Id.* at 491.

when the same washing problem also existed) or from some other source? If from another source, would that cause also affect stains on other materials? Even though this major study identifies what appears to be a considerable problem with ABO testing, the subsequent literature shows no attempt to answer such questions raised by the study.

Besides establishing an area of unreliability, this study should also point out the dangers of drawing conclusions from limited data. If the examinations had stopped at thirteen weeks, no false positives in the ABO system would have been found. Although three out of four denim samples showed false positives at twenty-six weeks, if only one sample had been examined and it was that fourth denim piece, no reason would have been presented to suspect unreliable antigen reports for bloodstains on denim. It must be stressed that this study is a major one which summarizes and tries to build upon previous test results. Unfortunately, the study has not been duplicated. If this is the last word on ABO detection in dried blood, the genetic marker system that has been the most used and studied, the research reveals not the stated conclusions of its authors, but rather the paucity of information about the detection of those markers in bloodstains. Furthermore, the limited data assembled has already uncovered areas in which false positives are produced. Certainly it is fair to ask whether such scientific information really proves that the tests are reliable.

C. *Reliability and the Other Antigens*

In 1965, one expert in the individualization of dried blood stated that since the results were not reliable for other systems, "many writers consider that the extension of bloodstain grouping outside the major divisions of the ABO system would be a dangerous practice."[174] Even today, the forensic scientist still often limits his grouping of antigens to the ABO factors.

> In the crime laboratory, bloodstain evidence is most frequently typed for the ABO group system, and often there is no determination of the other antigens. . . . Analysts are de-

[174] Outteridge, *supra* note 166, at 40.

> terred from making antigen determinations other than ABO for several reasons in addition to the cost or lack of time and training. The determinations call for complicated and lengthy procedures requiring numerous manipulations. A large amount of sample is needed compared with the ABO and enzyme systems. Sensitive, specific, reliable and inexpensive antisera for forensic use are not always available for the other antigens. There is a necessity for rigid testing of all reagents and strict quality control. The variability in antigenic concentrations causes problems with some systems.[175]

The antigen-antibody reaction in some systems is so slight that their typing is inherently suspect. "The antigens of the Kidd, Duffy, and Kell systems are extremely difficult to type because of weak agglutination between the antigens and their antibodies."[176] Thus, these "agglutination tests involve a degree of subjective judgment . . . [and] are prone to laboratory misclassification."[177] Not surprisingly, little scientific literature exists on the detectability of these systems in dried blood.[178]

Several forensic serologists have suggested, however, that antigen systems besides the ABO can be typed reliably under certain circumstances. For instance, although "[t]he complexity of the [Rh] system and its nomenclature was so great, that for many years it was considered an unsuitable system for forensic purposes,"[179] noted British scientist Bryan Culliford has stated that,

---

[175] Zajac, *supra* note 17, at 163-64. For discussions of the problems with antisera, see Denault Study, *supra* note 32, at 483; Grunbaum, Selvin, Myhre, & Pace, *Distribution of Gene Frequencies and Discrimination Probabilities for 22 Human Blood Genetic Systems in Four Racial Groups*, 25 J. FORENSIC SCI. 428, 439 (1980) [hereinafter cited as *Gene Frequencies*]. About the forensic usefulness of these other antigen systems, Zajac concludes that: "More reliable information can be obtained from smaller samples and with less time and effort by utilizing several enzyme/protein systems in addition to the ABO system." Zajac, *supra* note 17, at 167. *See also* Baxter, *supra* note 2, at 294-95.

[176] Denault Study, *supra* note 32, at 487-88.

[177] *Gene Frequencies*, *supra* note 175, at 439.

[178] *See* Denault Study, *supra* note 32, at 490.

[179] Baird, *supra*, note 1, at 108-09. Baird indicates that it was not until the late 1960's that a system for the grouping of the Rh factor in dried blood was first prepared and published. *Id.* The intricacy of the Rh system could make it the most useful genetic marker for the individualization of blood. "Because of the complexity of the Rh system, it is potentially capable of providing results of greater significance than any other single red cell antigen system or indeed any other form of blood typing with the possible exception of Gm." B.

"With adequate recent bloodstains that have dried quickly, Rh genotype is highly reliable."[180] Even Culliford, however, goes on to add caveats about RH typing:

> With old stains or those of inadequate quantity, it is only safe to draw conclusions from positive reactions. In these circumstances, one cannot be certain of the absence of an antigen when a negative reaction is obtained. Stains that have dried slowly present the greatest danger of erroneous typing and it is never safe to rely on results from those smelling of putrified blood.[181]

These limitations will, of course, often make the Rh system useless in practice because the aging conditions will not fall within the limited area in which Culliford finds high reliability, or because the aging conditions will not be known.[182] More important, however, the scientist does not give any reasons why he has concluded that Rh typing in dried blood is reliable even under restricted conditions. No published studies have been found to support the naked assertion. Instead, the literature details many problems with Rh typing in bloodstains.[183] Even the limited research done by the Denault Study found false positives using the Rh system, with the majority of these occurring after only two weeks of aging at room temperature.[184]

A scientist's assertion about the reliability of the detection of a genetic marker in dried blood should not be accepted without experimental testing of such a conclusion. The need for reliability testing is evident in the comments about the MN system, another antigen marker, made by J. Brian Baird, a leading Canadian foren-

---

CULLIFORD, *supra* note 3, at 81.

[180] B. CULLIFORD, *supra* note 3, at 96. Baird likewise stated that: "Rhesus bloodstains can be grouped, if they have been properly handled and air-dried at room temperature, for periods of up to one month." Baird, *supra* note 1, at 109.

[181] B. CULLIFORD, *supra* note 3, at 96.

[182] In ascertaining aging conditions to determine whether the Rh typing is reliable, Culliford states that "[t]he appearance of the bloodstain to an experienced eye is a good guide to the suitability for grouping. The behavior of the stain when tested for other components is also a guide to its quality for Rh typing. For example, those which have lost PGM activity are likely to give negative reactions in the Rh systems." *Id.* at 96-97.

[183] For a summary of this literature, see Zajac, *supra* note 17, at 166.

[184] *See* Denault Study, *supra* note 32, at 493, Table 6.

sic serologist:

> For forensic application, however, when one has to group bloodstains the system poses several problems. The M and N antigens are much less stable than the ABO antigens and if the stains are more than several weeks old, grouping is likely to be unsuccessful. . . . A further problem with the MN system is that, if extreme caution is not exercised, a group M stain can be erroneously grouped as MN. Controls must be run at all times and the quality of the antisera must be continually monitored. In the hands of a *competent and experienced serologist reliable results can be obtained on bloodstains.*[185]

Thus, even though problems exist, the right person can produce valid MN results in dried blood. Baird, however, cites nothing to support this final conclusion. After this statement, however, research was done studying the accuracy of MN testing in forensic laboratories. The data did not support Baird's earlier conclusion:

> The results of the proficiency testing demonstrate that an error rate as high as 40% is obtained with MN as compared to 1.6% for ABO. These results suggest that the fault is not necessarily in the ability of crime laboratory personnel to use these techniques but quite possibly in the system itself or the application of these particular techniques to the detection of M and N antigens in dried stains.[186]

Even more recently, Zajac, after reviewing the existing work, concluded that:

> At this time, the MN system must generally be considered unsuitable to forensic work involving phenotyping of dried bloodstains because of inherent unreliability. . . . The MN system, because of its unreliability, probably should not be determined in dried bloodstains until the problems inherent

---

[185] Baird, *supra* note 1, at 108 (emphasis added).

[186] *MN Determinations, supra* note 170, at 570. Culliford details problems encountered with MN typing in dried blood, concluding that these problems "are inherent in the system and a simple solution is unlikely." B. CULLIFORD, *supra* note 3, at 79. One of the major sources of error is that the N antigen is often detected as the M and thus the M type may be misgrouped MN. *See also MN Determinations, supra* note 170, at 570-71.

to the system are identified and clarified.[187]

As the MN system illustrates, reliability of genetic marker detection in dried blood, even for antigens that were typed long before the enzyme and protein systems, cannot be accepted on the basis of a mere assertion, even that of a noted forensic serologist. Instead, reliability can only be established by reliability testing. Little testing such as that done on ABO groups has occurred for the other antigen systems. What tests have occurred either indicate, as the limited Rh studies do, that reliability has not been established and is suspect or, as with the MN system, that the tests in dried blood are unreliable. The scientific literature produces similar conclusions about the detection of the enzyme and protein systems in dried blood.

D. *Reliability of the Electrophoretic Procedures in Dried Blood —An Overview*

Although the reliability of the tests for each genetic marker system must eventually be considered separately, a review of the scientific literature leads to some general conclusions about the markers detected by electrophoresis. First, almost no research has been done on the reliability of the electrophoretic procedures using dried blood. As with the antigen systems, what little can be gleaned about reliability has come primarily from persistence studies. These studies all have notable limitations and do not prove the reliability of the tests for any system.

For example, conclusions to be drawn from the electrophoretic patterns are subjective rather than objective.[188] Therefore, reliability can only be verified with blind tests, and many persistence studies have not employed even this elementary research tool. At a minimum, blind testing requires that the researchers who do the phenotyping in the dried blood be unaware of the samples' true types. If the serologist knows what phenotypes he should find, his eventual readings might be biased.

Some studies, aware of this potential problem, have withheld the

---

[187] Zajac, *supra* note 17, at 166-67.

[188] *See* Denault Study, *supra* note 32, at 485, 493.

expected results from those doing the tests. If, however, the studies are going to show that the procedures are reliable, the studies must go further than just restricting advance knowledge of the specific phenotype from the one analyzing the electrophoretic patterns. The practicing forensic scientist not only should not know what result he is looking for,[189] but also he will not know what aging conditions the blood sample underwent before it was collected. Even in the studies using blind testing, the serologist often knows the circumstances under which the stain aged although he is unaware of the actual phenotype. Significantly, the little research that has used readers ignorant of both the phenotypes present in the stain and also of the aging conditions has reported many false positives.[190]

The persistence studies likewise are deficient because they have not studied stains under the myriad of conditions that real criminal evidence will have undergone, and thus these studies cannot prove reliability in practice. For example, almost no work has been done in such elementary areas as blood dried at body heat, stains subjected to perspiration, blood dried on dirty clothing, stains aged under temperature or humidity fluctuations, and all the possible permutations of frequently encountered aging conditions.

Finally, the meager relevant research has highlighted some conditions that could produce erroneous readings. Moreover, it has produced reports of a high rate of false positives. Even with this evidence pointing time and again to the potential unreliability of the tests, these various results have almost never spawned the further research that would seem necessary before any of the results could be reasonably relied upon. A more detailed examination of the research concerning specific genetic marker systems illustrates

---

[189] Ideally, the practicing forensic serologist has no preconceived notion of what phenotypes he is looking for. However, it is possible for bias to occur. In most criminal cases, three readings are crucial—the dried blood, the blood obtained from the victim, and the blood obtained from the defendant. If the same person does all three readings, he is not really doing a blind test because he may expect that the stain will match one of the other two samples. The better procedure, therefore, is for different serologists to make the readings on each of the samples.

[190] See *infra* text accompanying notes 239-41 for a discussion of the higher rate of false positives that resulted when phenotyping was done completely blind.

these points.[191]

1. *PGM Testing*

In 1978, Baird stated that, "Of the electrophoretic systems used to characterize human bloodstains, the PGM system is probably the most popular."[192] Phosphoglucomatase is typed more often than any other enzyme or protein system. This wide use might lead to the assumption that the reliability of PGM detection for forensic purposes has been widely studied. The scientific literature, however, indicates otherwise.

Though no reports of false PGM positives in bloodstains have been found, only a handful of published studies have examined how long PGM persists in dried blood, and that research has been limited.[193] For example, one of the first studies used only PGM "2-1" samples.[194] This study noted that the phenotype of one genetic marker may degrade differently from that of another phenotype of the same marker. Since each of the different PGM phenotypes seems to deteriorate differently in stored blood,[195] it is questionable how much extrapolation can safely be done from one such work.

The Denault Study, the most comprehensive study done on PGM, is still quite limited. Only twenty-four samples dried at room temperature were tested for PGM. With the various permutations of drying conditions, this means that only one sample of PGM "2-1" from a stain dried on denim at sixty-six percent humidity was analyzed at thirteen weeks.[196] Because no false positive

---

[191] No attempt has been made to discuss every system. More than twenty genetically controlled polymorphic enzymes and proteins have been identified in bloodstains. *See* Grunbaum, *The Grunbaum System for Electrophoresis—Standardization of Equipment*, in HANDBOOK FOR FORENSIC INDIVIDUALIZATION OF HUMAN BLOOD AND BLOODSTAINS 7, 33 (B. Grunbaum ed. 1981). The systems which are discussed are those mentioned in the reported cases; these systems are the ones most often used in routine case work. *See* Baird, *supra* note 1, at 109.

[192] Baird, *supra* note 1, at 112-13.

[193] *See* Denault Study, *supra* note 32, at 496.

[194] *See* Rothwell, *supra* note 136, at 231.

[195] *Id.*

[196] *See* Denault Study, *supra* note 32, at 491, Figure 7.

was found in this single examination, does it follow that false positives would never exist with such aging conditions?[197]

A different study has presented evidence to show that false readings occur with PGM typing. PGM deteriorates rapidly; the effects are seen within a day after the blood is shed.[198] This degradation most often produces diffuse electrophoretic bands.[199] Furthermore, although no false positives have yet been reported in bloodstains, they have been reported in stored blood samples taken from corpses, and in semen stains.

Thus, Zajac notes that "there has [sic] been reported alterations of [PGM] electrophoretic patterns in post mortem samples . . . . "[200] Of course, it will often be crucial to know whether the PGM typings change in blood taken from a dead person. In *People v. Carlson*,[201] for example, the victim's body was found two days after she was last reported alive. Blood was taken from the corpse later on the day it was discovered. A bloodstain subsequently was found on the defendant's jacket. A laboratory analyst testified at trial "that the blood from the stain on Carlson's jacket possessed ABO, PGM, and EAP characteristics identical to the victim's blood . . . . "[202] Obviously, this evidence was truly damning to Carlson only if the PGM phenotypes were typed accurately.[203] If, however, the PGM in the victim's body altered between the time of her death and the laboratory tests, the stain may have indicated nothing about the defendant's guilt.

If the PGM alters in the blood after death, then PGM analysis may become worthless in homicide cases. Thus, it is crucial to

---

[197] A conflict of opinion seems already to have arisen from the sparse study of persistence concerning the effect of the substratum. The Denault Study indicates that permanent press and denim fabrics hinder the detection of certain PGM phenotypes. *Id.* at 494. Culliford indicates no effect from textiles except from the difficulty of handling thin materials. *See* B. CULLIFORD, *supra* note 3, at 120.

[198] *See* Rothwell, *supra* note 136, at 231, 233.

[199] *See* B. CULLIFORD, *supra* note 3, at 119.

[200] Zajac, *supra* note 17, at 168.

[201] 267 N.W.2d 170 (Minn. 1978).

[202] *Id.* at 172.

[203] The evidence was truly damning only if the stain did not match the defendant's blood. The opinion is silent on this point.

know whether such alteration actually takes place, and if it does, under what conditions. The needed research has not followed, however. In spite of a report that such changes occur, no published study has explored this critical area.[204]

A published paper, however, indicates that storing blood can produce false PGM positives.[205] Ten samples of blood were stored at room temperature and analyzed periodically. On day fifty-one, no readings were obtained from two of those samples, five were accurately typed, and the three remaining gave false results. Two samples of PGM "2-1" were typed as "2" and one sample of PGM "1" became "2-1." The authors concluded by suggesting that in blood stored beyond five weeks, PGM should be interpreted cautiously.[206] This warning is curiously weak. If the data correctly indicated that at day fifty-one an error rate greater than one-third occurs, more than caution is needed. Clearly, at that point, the tests are unreliable.

The way to prevent this unreliability would seem to be simply to minimize storage by having the blood of the victim or defendant typed as quickly as possible. This study, however, presents the need for a more thorough consideration of the problem. The researchers were able to accurately group all ten PGM samples through day thirty-two, and consequently, no warning was issued about blood stored for less than five weeks. However, were ten testings enough to support a conclusion of reliability for up to five weeks? The study only examined blood stored at room temperature. Do other storage conditions bring about quicker alterations? Finally, why are false positives reported in the stored blood, but not in the dried? Is there some reason that the bloodstain should show less deterioration than the whole blood? If more testing of stains were done, would erroneous readings also be found? While

---

[204] In support of the statement that postmortem samples have shown alterations, Zajac cites B. Campbell, L. Luke, & M. Cowan, EAP, EsD, and PGM in Post Mortem Blood (1978) (unpublished paper, Cuyahoga County Coroner's Laboratory, Cleveland, Ohio). *See* Zajac, *supra* note 17, at 168.

[205] Rees, Howard, & Strong, *The Persistence of Blood-Group Factors in Stored Blood Samples*, 15 J. FORENSIC SCI. SOC'Y 43 (1975) [hereinafter cited as *Persistence of Blood-Group Factors*].

[206] *Id.* at 48.

these questions and others like them seem important for accurate phenotyping, attempts to answer them do not appear in the literature.

It is well established that PGM can be detected in body tissues other than blood. The enzyme is found in semen, for example. One study has found, however, that semen contaminated with saliva alters the electrophoretic bands of PGM.[207] Once that finding was reported, it would seem logical to ask whether saliva or other contaminants[208] could alter PGM readings in bloodstains. Nevertheless, even though stains might often contain contaminations from other body fluids, no one has studied this potential area of unreliability.

In summary, then, the PGM studies so far have reported no false positives in bloodstains,[209] but the research has been limited. In addition, other areas of PGM study have found erroneous readings that could affect criminal investigations and that should raise questions about the validity of PGM phenotyping in dried blood. Even so, these avenues of research have not been followed, leaving large gaps in knowledge about the detection of this genetic marker.

2. *AK Reliability*

The adenylate kinase research mirrors that done on PGM. Although the study of AK in stains has not been far ranging, no false positives in bloodstains have been reported. The Denault Study consisted of only twelve samples dried at room temperature and two frozen samples, all of which were AK "1."[210] AK, like PGM,

---

[207] *See* Sensabaugh, Blake, & Northey, *Genetic Markers in Semen III: Alteration of the Phosphoglucomutase Isozyme Patterns in Semen Contaminated with Saliva*, 25 J. FORENSIC SCI. 470 (1980) [hereinafter cited as *Genetic Markers in Semen III*].

[208] The authors stress that they only studied semen samples contaminated with saliva and that alterations in the PGM readings may occur from other bodily fluid contaminants. *Id.* at 477.

[209] Zajac reported that "[t]he cathodal bands of some of the more rare phenotypes (the PGM 8-1, for example) lose intensity and degrade sooner than other isoenzyme bands, which may lead to the possible mistyping of the phenotype as a PGM 1." Zajac, *supra* note 17, at 168. Zajac cites nothing in support of this statement.

[210] Denault Study, *supra* note 32, at 493. Similarly, all the blood examined by Rothwell was AK "1." *See* Rothwell, *supra* note 136, at 231. The AK system is not as forensically valuable a system as some others. It does not have a high discriminating power among a

also has exhibited erroneous readings in stored blood.[211] Finally, like PGM, the AK research shows little scientific inquiry into possible areas that might produce unreliable phenotyping in stains.

3. *ADA Reliability*

A little over a decade ago, a leading British forensic serologist stated that, "The assay of ADA is not easy and at present different methods so far published and tried appear to give markedly different results."[212] In spite of that conclusion, the study of possible false adenosine deaminase positives in dried or preserved blood is even less than for the PGM or AK systems. The electrophoretic procedure used by the Denault Study produced no erroneous ADA readings on the fourteen tested samples. That research revealed, however, that ADA was less stable than other enzymes,[213] and persisted best under conditions that do not favor the preservation of other genetic markers.[214] Meanwhile, Zajac reports, without citation to any source, that, "The ADA system has been shown to exhibit alterations in band patterns in dried, stored and bacteria-infected blood samples."[215] Even though this statement would seem to lead to the conclusion that the forensic ADA typing is unreliable, neither Zajac nor any other published author further explores this contention.

---

population since more than 90% of the people have type AK "1." *Id. See* Baird, *supra* note 1, at 113-14.

[211] *See Persistence of Blood-Group Factors, supra* note 205, at 47-48. The authors report that at days 109 and 122 of the storage, two AK "2-1" samples were read as "1" and they therefore concluded that AK results are not reliable for blood stored over twelve weeks. *Id.* AK is generally regarded as a stable enzyme and more reliable than PGM. *See also* B. CULLIFORD, *supra* note 3, at 138; DISTRIBUTION OF HUMAN BLOOD GROUPS, *supra* note 25, at 38; Rothwell & Sayce, *The Stability of PGM and AK Isoenzymes in Human Tissues*, 19 J. FORENSIC SCI. 590 (1974); Zajac, *supra* note 17, at 168.

[212] B. CULLIFORD, *supra* note 3, at 174.

[213] *See* Denault Study, *supra* note 32, at 493-94.

[214] The researchers found that the electrophoretic bands of ADA were more distinct for the samples aged at the higher humidity, but the authors also concluded that "blood samples in general should be preserved at low temperature, where the humidity is also low." *Id.* at 494.

[215] Zajac, *supra* note 17, at 168.

4. *EsD Reliability*

EsD "is not particularly stable."[216] All esterace D phenotypes do not degrade at the same rate.[217] As a consequence, "the EsD system is subject to alterations in the phenotype patterns which could lead to mistypings if the forensic serologist is not aware of such possibility."[218]

The first report of successful EsD typing in dried blood was published in 1975, although this work also found that as the blood aged it became very hard to distinguish EsD "1" from "2-1."[219] In 1979, a pair of researchers published what they termed "a comprehensive stability evaluation" of EsD.[220] They reported that as the bloodstain ages, band 1 in EsD "1" loses intensity more quickly than bands 2 or 3, and that in EsD "2-1," band 2 loses intensity more rapidly than bands 1 and 3.[221] The researchers summarized their results as follows: "With these criteria, no error was encountered in grouping stains up to four weeks of age. Stains older than four weeks could not be grouped."[222] They concluded that if the serologist is aware of these changes during aging, no mistyping will occur, although they implied that erroneous readings will be given if these alterations are unknown to the person typing the stain.[223]

---

[216] Baird, *supra* note 1, at 116.

[217] *See* Rothwell, *supra* note 136, at 233.

[218] Zajac, *supra* note 17, at 168.

[219] Parkin & Adams, *The Typing of Esterase D in Human Bloodstains*, 15 MED. SCI. & L. 102, 103 (1975).

[220] *See* Jay & Philp, *A Stability Study of the Esterase D Isoenzymes*, 24 J. FORENSIC SCI. 193, 193 (1979). Even though labeled "comprehensive," the only drying condition studied was that of room temperature. Certainly blood can age under other conditions. A stain on clothing, for example, may have dried at body temperature and have been subjected to perspiration or washing. The study does not indicate whether the typing of the aged samples was done "blind," that is, with those doing the typing not knowing the phenotypes in the samples. If not done blind, the study also can be faulted on this ground. A study is not comprehensive just because those who produced it say it is. Once again, this study, the most detailed on EsD, really points out how little research has actually been done.

[221] *Id.* at 195.

[222] *Id.* at 196.

[223] It should follow that if the person doing the bloodstain analysis is not familiar with these characteristic degradation effects, the subsequent phenotyping is not reliable. No reported cases indicate whether the forensic personnel were aware of these conditions. This is true even in People v. Bush, 103 Ill. App. 3d 5, 430 N.E.2d 514 (5th Dist. 1981), in which the blood tests must have been performed only shortly after the article detailing the deteriora-

5. *Hp Reliability*

A decade ago Culliford stated that haptoglobin stains had to be less than three days old to be reliably typed and that Hp "2-1" can appear as "1-1."[224] Apparently shortly thereafter a new procedure and new equipment gave forensic scientists increased confidence in their ability to type Hp in dried blood.[225] In 1979, however, another new procedure was developed. "Bloodstains ranging in age from approximately six weeks to two years . . . were accurately phenotyped in haptoglobin by electrophoresis after the samples were prepared by a newly adapted chloroform procedure."[226] Stolorow and Wraxall, the developers, concluded that without this new technique, haptoglobin was uninterpretable in stains after six weeks.

The developers' conclusion is interesting when compared to the analysis done in *State v. Fulton*.[227] In this case, following a violent robbery on February 7, 1979, the defendant was arrested on that night and a bloodstain was seen on his shoe. The shoe, however, was first sent to a forensic laboratory for analysis of the stain on April 10, 1979, making the minimum age of the dried blood approximately nine weeks at the time of the forensic tests. According to Stolorow and Wraxall, without their newly developed method, Hp analysis should have produced uninterpretable results. Their method, however was not published until after April 1979, making it doubtful that it was applied in *Fulton*. Even so, a forensic serol-

---

tion characteristics was published. The study had a publication date of January, 1979, the bloodstains and other blood samples in *Bush* were collected on May 30, 1979, and presumably analyzed shortly thereafter. One can only speculate whether the absence of any mention of the Jay & Philp article also indicates a lack of knowledge of its contents. *Bush* also is curious because, though other enzymes are more stable than EsD and problems with EsD typing have been reported, the only phenotyping reported was of the ABO and EsD systems. Nothing in the opinion indicates why the analysis was so limited.

[224] B. CULLIFORD, *supra* note 3, at 216. Unlike the other genetic markers discussed here which are typed through electrophoresis, Hp is not an enzyme, but a serum protein. *See* Baird, *supra* note 1, at 118.

[225] *See* Baird, *supra* note 1, at 118. Baxter likewise states that: "It is only with the development of sophisticated equipment that haptoglobins can now be reliably typed in bloodstains." Baxter, *supra* note 2, at 287.

[226] Stolorow & Wraxall, *An Efficient Method to Eliminate Streaking in the Electrophoretic Analysis of Haptoglobin in Bloodstains*, 24 J. FORENSIC SCI. 856, 861 (1979).

[227] 229 N.C. 491, 263 S.E.2d 608 (1980).

ogist stated at trial that the bloodstain contained Hp "2-1," which was the same phenotype as the victim's.

Several things should be noted from this. Although nothing specific is mentioned in the opinion, perhaps there is a way to reconcile this apparent conflict between the researchers and the state's forensic expert. The court's inattention to this conflict may have resulted from the absence of a defense challenge to the admissibility of the tests. The prosecution's witness, no doubt, indicated that her Hp analysis of the bloodstain was reliable, even though not all forensic serologists would have agreed with her. Surely this ought to give pause again to defense counsel who do not challenge such assertions and to courts that accept these assertions without insisting upon thorough reliability studies.

6. *EAP Reliability*

In *State v. Washington*,[228] the court, presented with conflicting evidence about the reliability of the tests for erythrocyte acid phosphatase in bloodstains, upheld the admissibility of the procedures.[229] A closer look at the merits of that dispute is in order.[230] The court was impressed with the state's expert who testified that the EAP "enzyme was less likely to deteriorate when dried than if kept in liquid form. He testified that EAP enzyme analysis on dry blood stains was reliable, and that this reliability was well established."[231] This opinion was supported by the research done by the

---

[228] 229 Kan. 47, 622 P.2d 986 (1981).

[229] See *supra* text accompanying notes 110-22 for an analysis of the contradictory expert testimony presented in *Washington* concerning the reliability of enzyme typing of blood samples taken from the crime scene.

[230] Errors can easily occur in reading EAP electrophoretagrams if done by insufficiently skilled personnel: "The EAP system is sometimes difficult to interpret since the phenotypes depend not only on a pattern of relative mobility of the isoenzyme bands but also on the relative intensities of the bands." Zajac, *supra* note 17, at 167. Therefore, "the results are somewhat more difficult to read than those of some of the other systems and the serologist should have a considerable amount of experience in using this system before applying it to routine casework." Baird, *supra* note 1, at 115.

[231] State v. Washington, 229 Kan. 47, 51, 622 P.2d 986, 990 (1981). For the problems of typing EAP in aged whole blood, see DISTRIBUTION OF HUMAN BLOOD GROUPS, *supra* note 25, at 41. The authors therein state that electrophoresis "in fresh blood samples, produces two main bands and some fainter ones. In older blood specimens additional bands appear which render typing difficult." *Id.*

Denault Study, who, according to the court, "concluded that EAP enzymes did not deteriorate for a period of up to thirteen weeks."[232]

The basis for the "well established" reliability opinion was not given. Other forensic serologists have reached less dogmatic conclusions. Baird, for example, states that, "One disadvantage is that the [EAP] enzyme is not particularly stable in dried bloodstains and hence they have to be not more than 2-3 weeks old for successful typing. Older stains can give spurious results."[233]

Even the study by the Denault Study does not support the blanket conclusion that EAP bloodstains remain without deterioration for thirteen weeks. Instead, those scientists gave a limited conclusion to this phase of their work. They only placed their findings within the framework of the other EAP literature after first noting that "a large discrepancy exists in the detectability of the EAP isoenzymes in stains. Reported time limits for their identification vary from a few days to eight or nine weeks. This study has shown that EAP isoenzymes can be detected after 13 weeks of storage under a variety of conditions."[234] This research does not conclude that no deterioration occurs; instead, it establishes that on twenty-four stains (two of which were frozen), accurate phenotyping was done at thirteen weeks with the specified aging conditions. Indeed, this study actually indicates that deterioration happens by thirteen weeks. The authors reported that "[a]t 13 weeks, greater enzymatic activity was consistently observed for the specimens aged under the higher humidity level (66%) than for those stored at the low level (20%)."[235]

Deterioration, however, does not necessarily mean unreliable tests. The crucial concern is whether deterioration results in false positives. While the Denault Study reported no erroneous readings, Grunbaum testified at the *Washington* trial that EAP analysis often gave erroneous readings. The court rejected Grunbaum's

---

[232] 229 Kan. at 52, 622 P.2d at 990.
[233] Baird, *supra* note 1, at 115.
[234] Denault Study, *supra* note 32, at 494.
[235] *Id.*

work, however, noting that "Dr. Grunbaum used a different method than is used in the method developed by Wraxall."[236] This assertion is particularly interesting, since Grunbaum's published work has shown that false EAP positives are obtained on both mediums—cellulose acetate membranes and starch gel.[237]

In Grunbaum's study, done in conjunction with Zajac, samples of whole fresh blood with known EAP phenotypes, were applied to clean cotton cloth and dried at room temperatures for nineteen to twenty-six days. Portions of the liquid blood were also stored at body temperature for up to two days and then frozen to prevent further deterioration until the phenotyping was done.

> The blood samples were prepared in this manner to simulate the adverse conditions which may take place in actual casework submitted to the crime laboratory. It is not uncommon to have dead bodies and bloodstained clothing or objects subjected to heat and humidity for several days prior to discovery and collection. Air temperatures are frequently near or above body temperature (37°C) and can be considerably higher in desert or tropical climates, in closed rooms, closed automobiles, or in direct sunlight.[238]

In spite of what the *Washington* opinion indicated, the typing was done with both cellulose acetate membranes and starch gel. The starch gel procedure was "essentially that described by Wraxall and Emes with minor modifications and improvements developed at the University of California's White Mountain Research Station."[239] The phenotyping was done in a completely blind manner by competent people:

> The results were read independently by four individuals having extensive experience with both methods of EAP phenotyping, in research and casework situations. The readers

---

[236] State v. Washington, 229 Kan. 47, 55, 622 P.2d 986, 993 (1981). The state's witness indicated that the medium used by Wraxall was a starch gel base while that used by Grunbaum was a cellulose acetate membrane.

[237] *See* Zajac & Grunbaum, *Problems of Reliability in the Phenotyping of Erythrocyte Acid Phosphatase in Bloodstains*, 23 J. FORENSIC SCI. 615, 617 (1978).

[238] *Id.* at 616.

[239] *Id.* at 616 n.3 (referring to Wraxall & Emes, *Erythrocyte Acid Phosphatase in Bloodstains*, 16 J. FORENSIC SCI. SOC'Y 127 (1976)).

> had no knowledge of the phenotypes, the age of the samples, or the way in which they were prepared prior to analysis.[240]

The starch gel readers did a total of twenty-eight readings of stains and heat degraded samples; ten were wrong.[241] Three out of twelve stains were incorrectly phenotyped, while false positives were registered for seven of twelve heat degraded samples.

A total of forty-six readings of stains and heat treated liquid samples were done on cellulose acetate. Seven of those forty-six were labelled inconclusive. Six more were called inconclusive, but with tentative readings given. Three of those six tentative readings were wrong. "Successful" phenotyping was therefore done on twenty-seven samples; three of these were wrong.[242] The total rate of false positives in this study thus exceeded twenty-five per cent. These false positives occurred not because the readings bordered on the inconclusive but because of changes in the blood samples. "The problem with misidentification of the samples in this study was not due to weak or indistinct band patterns. Rather, discrete bands were present and readable, but they had been altered to indicate erroneous phenotypes."[243]

The Denault Study reported no false EAP positives; Zajac and Grunbaum did. These differences cannot be explained merely by stating that each used a different medium, because Zajac and Grunbaum used both mediums.[244] Although neither scientific team commented on the results of the other, Zajac and Grunbaum gave a possible reason for the discrepancy. The two authors first stated that others besides themselves had reported mistypings in aged EAP samples, but also noted that "Wraxall and Emes report no mistyping on 387 bloodstains aged from two days to six weeks."[245] Zajac and Grunbaum pointed to a distinction between that study

---

[240] Zajac & Grunbaum, *supra* note 236, at 616.

[241] This statistical data comes from Zajac & Grunbaum, *supra* note 237, at 616, Table 1.

[242] These three incorrect readings were all in heat degraded samples. *Id.* at 616.

[243] Zajac & Grunbaum, *supra* note 237, at 617.

[244] Zajac & Grunbaum thus concluded that: "The results . . . indicate that there can be a definite problem with the EAP phenotyping no matter which electrophoretic supporting medium is used." *Id.*

[245] *Id.* at 615.

and their own: "Apparently the bloodstains they used were prepared under laboratory conditions and the history of drying and preservation were well known."[246]

Using a study designed to approximate closely actual casework situations (something other researchers failed to do), Grunbaum and Zajac thus reported false EAP positives. At the least, this raises questions about the forensic EAP phenotyping. The Zajac and Grunbaum research, of course, does more than merely raise some questions. Their work indicates that subjecting liquid blood to body temperature heat makes the subsequent EAP phenotyping unreliable since, of the thirty-two readings of the heat treated samples, fourteen were wrong.[247]

This conclusion about the effect of heat is supported by a more recent report from actual casework. "In a recent homicide case, several items of evidence from the victim, three suspects, and the crime scene were submitted for analysis. A 'liquid' blood sample, collected from the ground upon which the victim was found, was submitted as a standard for comparison."[248] This liquid blood was exposed "to temperatures exceeding 32°C (90°F) for several hours."[249] It was then kept at room temperature for several days. The subsequent EAP phenotyping gave a false positive. The author concluded that the erroneous reading was caused by the heat and stated that, "Generally, when a blood sample is exposed to excessive conditions it may undergo an alteration of the EAP isozyme patterns and appear to be some type other than that originally coded."[250]

The scientific evidence, therefore, demonstrates that heat of approximately body temperature, even for a short period, degrades the blood in ways that often produce false positive EAP readings. This conclusion is important not only because liquid blood collected in casework may have been subjected to such conditions,

---

[246] *Id.*

[247] Seven were typed on starch gel and three on cellulose acetate. *Id.* at 616.

[248] Yeshion, *Thermal Degradation of Erythrocyte Acid Phosphatase Isozymes in a Case Sample*, 25 J. FORENSIC SCI. 695, 696 (1980).

[249] *Id.*

[250] *Id.* at 697.

therefore making the subsequent EAP phenotyping unreliable, but also because of the questions that it ought to trigger. For example, if heat produces unreliable EAP readings, how does heat affect the typing of other genetic markers in liquid blood? The scientific literature is silent. How does heat affect EAP phenotyping in bloodstains? The importance of the answer to this cannot be overestimated. Blood, on hot days or in a closed room or automobile or under many other circumstances, will be exposed to warm temperatures as it dries or as it ages. Does that result in false positives, and thus unreliable phenotyping? Doesn't blood on clothing worn by a person dry or age at above room temperature? A high percentage of bloodstains submitted for forensic analysis must fall into this category. Even so, the scientific literature is silent about a stain that dried or aged on clothing worn by a person. Finally, how do warm temperatures during the drying or aging of a bloodstain affect the phenotyping of genetic markers other than EAP? Again, the scientific literature is silent.

E. *The Reliability of Forensic Genetic Marker Phenotyping*

To state that gaps exist in the knowledge about the phenotyping of genetic markers in forensic casework is to understate the situation. Even so, as evidenced by the testimony in the cases and by the widespread use by forensic laboratories, many feel that the tests are reliable. Part of this confidence comes, perhaps, from the fact that few of the people who use the tests have ever listed the possible ways that the typing might produce untrustworthy results and then scientifically examined these possibilities.

Even those scientists who have identified problems with the genetic marker typing indicate that it can be done reliably. Thus, Zajac, at the conclusion of her survey of possible areas of unreliability, states that: "Alterations in degraded blood do not pose severe problems in interpretation if they are recognized for what they are. Unrecognized, they can lead to errors in identification."[251] This apparent minimization of the unreliability problem is incomprehensible when it is realized that it was Zajac who earlier posited

---

[251] Zajac, *supra* note 17, at 168.

that no way exists to recognize the alterations:

> Phenotypes in some contaminated, dried, or putrified blood may be in an altered and degraded state which gives an unambiguous but false reading. Since it is impossible to subject a standard to the unknown conditions of the environment and aging that has altered the sample under examination, there is no true basis of comparison and no empirical way to detect that the phenotype has altered. Consequently, the analyst should be hesitant to accept even "unambiguous" readings for certain comparatively unstable systems if the history of the sample is unknown, or if it has been subjected to aging and adverse conditions.[252]

Thus, even though it is not possible to tell when the phenotype has altered, reliability is somehow aided by knowledge of the history of the sample.[253] Looking at a reported case makes one wonder why this is true.

In *State v. Hampton*,[254] the prosecution's witness testified that on the night of January 21, 1977, he was with the defendant and the victim when the defendant beat the victim to death with a car jack. The witness then stated that he and the defendant placed the body in an automobile trunk, drove to a lake, and dumped the body in the water. The body was recovered on February 27, 1977.

---

[252] *Id.* at 163.

[253] Zajac also states that:

> The history of the sample affects the typing results and samples subjected to extreme and adverse conditions may appear other than what they are. There are many variables which must be accounted for and the forensic serologist should not be reluctant to report an inconclusive result when working with samples which have been subjected to unknown or adverse conditions.

Id. at 168-69. Zajac and Grunbaum similarly note that:

> Erythrocyte acid phosphatase is a useful system for the crime laboratory for both fresh and degraded blood and bloodstains, provided the inherent problems of phenotyping this particular enzyme system are recognized. Because of the great number of variables affecting this enzyme system in vitro, phenotyping should not be attempted unless the complete history of origin and handling of the sample is known.

Zajac & Grunbaum, *supra* note 237, at 617. *See also* Denault Study, *supra* note 32, at 496 ("With some knowledge of the history of a bloodstain, the criminalist can now determine the reasonableness of performing certain analyses.").

[254] 294 N.C. 242, 239 S.E.2d 835 (1978).

> The State also offered expert testimony tending to show that the deceased's blood type was AB, EAP, Group B, a rare blood type which normally occurs in only two or three percent of the population. The same type blood was found on the bumper jack allegedly used in the killing and on a leather jacket found near the scene where the killing supposedly took place.[255]

If histories of the samples are invaluable to accurate phenotyping, then what are the histories of these samples? The serologist was not present when the blood was shed. His knowledge of the samples' aging conditions before they were collected[256] can only be learned if the prosecution's witness has been telling the truth. The very nature of a trial, however, means that the defendant is challenging that witness' truthfulness. Since the serologist has no scientific way of determining veracity, the blood analysis cannot be done solely under the hypothesis that the witness is honest. If he is not truthful, what are the histories of the samples? Clearly, they are not determinable. This same sort of problem will be evident with any blood sample when the defendant does not concede the same history as the prosecution theory proclaims.[257] A conclusion that the blood sample's history must be known before the reliable detection of genetic markers can be done really is a conclusion that the forensic serologist seldom, if ever, can do reliable phenotyping.

Even if it can be assumed that the witness was telling the truth and all the aging conditions can be learned, this does not assure error free analysis. A scientist can only guarantee that the typing is accurate if controlled experiments have been done that match, or at least closely approximate the case's actual conditions. Knowledge of the sample's history is useful only if science knows how those historical conditions affect genetic marker detection. The history of the blood sample in *Hampton* thus only begins the questioning. What is known about phenotyping in a six week old

---

[255] *Id.* at 245, 239 S.E.2d at 838.

[256] "[B]lood found at the scene of a crime degrades before and after collection." Grunbaum, *supra* note 37, at 103.

[257] "Frequently, the analyst is unaware of the age and history of the blood or bloodstain and has no way of obtaining this information." Zajac, *supra* note 17, at 160.

corpse?[258] Or a submerged corpse? Or a corpse submerged at the temperature of that lake? Did any exhaust fumes get into the trunk while the body was transported? If so, how do those gases effect the subsequent phenotyping? Did the trunk become warm? If so, what effect did that temperature have?

Perhaps the corpse seems an extreme example, but similar questions abound about the jacket and the jack. If the jack has oil or dirt on it, how did that affect the phenotyping? What effects were caused by the temperature changes that the jack must have undergone? The jacket was of leather. How does that substrate affect the phenotyping? If the jacket was dirty or contained sweat or bacteria, how was the phenotyping affected? How did the weather conditions affect the jacket as it lay at the crime scene?

Similar questions should be legion in almost any criminal investigation in which blood is typed. The data concerning any of these matters is minimal or more often nonexistent; certainly it is nonexistent on the crucial cumulative effect of such conditions.[259] In fact, almost the only data supporting reliability of genetic marker detection in forensic detection consists of blood deposited on clean cloth and dried at room temperature in a laboratory. Only a very considerate criminal would leave such evidence behind for the police serologist.[260]

No research has been done which comes close to duplicating the conditions that can actually be expected to be encountered in a criminal case. The forensic analysis, therefore, can only be considered reliable if, from the few tested conditions, extrapolations can

---

[258] Although the case does not say so, the phenotypes ascribed to the victim no doubt came from analyses done on the corpse's blood, since there is seldom a medical reason to phenotype the red cell enzymes.

[259] Zajac has noted that:

> Age alone is not critical to the reliability of phenotyping a blood sample. Rather, it is essential to consider the cumulative effects of substrate, biological quality and quantity, environmental effects such as heat, humidity and contamination, and the preservation and transportation of the samples prior to submission to the crime laboratory.

Zajac, *supra* note 17, at 167.

[260] *See* Baird, *supra* note 1, at 115 ("Most perpetrators of crime are not always so considerate as to leave behind adequate quantities of their blood to ensure that it can be reliably grouped in the EAP or any other blood grouping system.").

be made to actual case conditions. Such extrapolation is, of course, not a matter of proven science, but truly a matter of faith—that is, the scientists, even though it has not been proven, believe that because something was true under one circumstance, it will be true under other circumstances. No reason exists for such faith here. Even under the usual laboratory conditions, reports of false positives exist. Nevertheless, no research has answered even the simple question of the rate at which these errors occur.[261] Thus, the reliability of genetic markers detection even under the controlled laboratory conditions is questionable. Certainly, then, no reason exists to believe that under other circumstances the procedures are trustworthy, and many good reasons exist to believe that they are unreliable. The little scientific research that has considered conditions other than the usual laboratory ones repeatedly has reported inaccurate phenotyping. For instance, in research previously discussed which would be relevant to the *Hampton*[262] case, there is evidence to show false positives in postmortem samples, in blood subjected to heat, in stains dried on leather, and in stains contaminated with bacteria. The evidence certainly does not establish that the genetic marker tests are reliable as applied in actual criminal investigations; indeed, many reasons for their unreliability have already been suggested. Knowing the history of the sample does not change these facts.

V. CONCLUSION

Forensic serology has attempted to perfect methods of detecting genetic markers in blood samples discovered during criminal investigations. These tests are new and different from traditional blood grouping tests. In spite of this newness, the techniques to individualize blood have already been widely adopted by forensic laborato-

---

[261] In contrast to the lack of research into error rates in the context of criminal investigations, research into error rates in paternity bloodtesting has shown a 200% decrease in five years. *See* Terasaki, *supra* note 2, at 555. Terasaki argues that HLA bloodtesting is "a very powerful, effective new tool in cases of disputed paternity," *id.* at 548, but also can report that, after extensive testing under conditions which should prevail in disputed paternity blood tests, the error rate had dropped from 1.08% in 1971 to 0.35% in 1976 with even increased accuracy in the future. *Id.*

[262] 294 N.C. 242, 239 S.E.2d 835 (1978). See *supra* text accompanying note 255 for the factual situation in *Hampton*.

ries. These procedures, however, differ in important ways from other scientific tests since they are often ends in themselves and not merely means to other scientific advances. Consequently, widespread acceptance and use by the forensic community does not guarantee that the genetic marker procedures are reliable.

The cases that have considered the admissibility of the genetic tests uniformly have placed the burden to prove unreliability of this new evidence upon the defendants. This is the wrong approach because even if the forensic procedures are unreliable, the defendant will be unable to establish this fact. The burden of proof, rather, should be on the prosecution who should have to show that the procedures are reliable.

The reliability of the tests is not revealed by the uses to which the results are put. Instead, reliability can only be established by careful scientific experimentation. The scientific literature contains few studies that bear on reliability. In fact, little is known about the reliability of these tests as they are used in actual practice. However, the brief existing literature does contain proven and potential areas of unreliability in the tests' uses. At this stage of development, therefore, the forensic tests for the detection of genetic markers in blood have not been proven to be reliable and should not be admitted into criminal trials.