NIST Scientific Foundation Review on DNA Mixture Interpretation

John M. Butler,
Hari Iyer, Rich Press, Melissa K. Taylor, Peter M. Vallone, Sheila M. Willis*

National Institute of Standards and Technology, Gaithersburg, Maryland USA

*International guest researcher; retired director of Forensic Science Ireland
NIST Draft Report Released in June 2021

DNA Mixture Interpretation: A NIST Scientific Foundation Review

John M. Butler
Hari Iyer
Rich Press
Melissa K. Taylor
Peter M. Vallone
Sheila Willis*

*International Associate under contract; retired director of Forensic Science Ireland

This publication is available free of charge from:
https://doi.org/10.6028/NIST.IR.8351-draft

250 pages
Executive Summary (9 pages)
6 chapters and 2 appendices
528 references cited
47 terms and acronyms defined
29 tables
12 figures
5 boxes
16 principles described
25 key takeaways
8 future considerations

Released for a public comment period (June 9 to August 23, 2021)
Presentation Overview

1. Report Contents and Key Takeaways
   • Why NIST has undertaken this effort
   • Brief summary of our findings

2. Outreach and Public Comments Received
   • Public webinar given on July 21, 2021 (1,000 registrants) – 83 questions/comments
   • Presentations given to FBI SWGDAM (July 14) and NIST/NIJ Human Factors Working Group (July 28)

3. Future Plans
   • A final report will be issued after considering comments received
   • FAQs on a NIST website may also be created in addition to final report

These handouts, which were due to Promega by August 23, do not contain the final slides; for a final version of the presentation, see https://strbase.nist.gov/NISTpub.htm after September 16
Disclaimer & Acknowledgments

Certain commercial equipment, instruments and materials are identified in order to specify experimental procedures as completely as possible. In no case does such identification imply a recommendation or endorsement by the National Institute of Standards and Technology nor does it imply that any of the materials, instruments or equipment identified are necessarily the best available for the purpose.

Acknowledgments (page i): Members of the DNA Mixture Resource Group (listed in Table 1.2) contributed helpful feedback and assistance in the early stages of drafting this report. Katherine Gettings, Nikola Osborne, and Sarah Riman provided valuable input on the text, including the data summaries used in Chapter 4. Jason Weixelbaum, Susan Ballou, Christina Reed, and Kathy Sharpless assisted with copy editing. Kathryn Miller from the NIST Library helped finalize the document for public release.

Acknowledgments: NIST team members and Resource Group for their insights; all those who provided public comments
Requests for **Understanding What Data Exists**

**Supporting Forensic Science Methods and Practices**

- **NRC Report (2009)**
- **NCFS Recommendation (2016)**
- **PCAST Report (2016)**
- **NISTIR 8225 (2020)**

“demonstrating the validity of forensic methods”

(Recommendation #3)

“technical merit evaluation”

Congressional funding uses NCFS language

“establishing foundational validity”

NIST: a “Scientific Foundation Review”
1. **DNA Mixture Interpretation** (initial pilot study)
   -Began in September 2017
   -AAFS 2019, ISHI 2019, ISHI 2020, AAFS 2021 workshops conducted
   -250-page report released for 60-day (extended to 76-day) public comment on June 9, 2021, with a 3-hour webinar held on July 21

2. **Bitemark Analysis**
   -Began in October 2018
   -Workshop held in October 2019

3. **Digital Investigation Techniques**
   -Began in February 2019
   -Interlaboratory “black box” study conducted from June to November 2020

4. **Firearm Examination**
   -Began in October 2019
   -Gathering literature and focusing on error rate studies

https://www.nist.gov/topics/forensic-science/interdisciplinary-topics/scientific-foundation-reviews
DNA Mixture Report Content

In six chapters and two appendices:

• Chapter 1 introduces the topic and challenges of DNA mixture interpretation
• Chapter 2 provides background information on DNA, describes principles and practices underlying mixture measurement and interpretation, and introduces the likelihood ratio (LR) framework and probabilistic genotyping software (PGS)
• Chapter 3 lists data sources used in this study and strategies to locate them
• Chapter 4 and Chapter 5 cover reliability and relevance
• Chapter 6 explores the potential of new technologies to assist mixture interpretation and considerations for implementation
• Appendix 1 reviews the history of how the field has progressed
• Appendix 2 discusses strengthening the field with training & continuing education
• Bibliography includes 528 references cited in the report
Our Desire with This Report is to Help Move the Field Forward to Improved Practices in DNA Mixture Interpretation

From the Executive Summary (page 1):
“As with any field, the scientific process (research, results, publication, additional research, etc.) continues to lead to advancements and better understanding. Information contained in this report comes from the authors’ technical and scientific perspectives and review of information available to us during the time of our study. Where our findings identify opportunities for additional research and improvements to practices, we encourage researchers and practitioners to take action toward strengthening methods used to move the field forward. The findings described in this report are meant solely to inform future work in the field.”
Clarification on What NIST Is and Is Not

• NIST is a Federal government science agency and does not comment on legal admissibility

• NIST is not a regulatory agency, which is why key takeaways are provided in our draft report rather than formal recommendations

• NIST focuses on research and assisting with developing standards (e.g., OSAC or SRMs); NIST does not conduct forensic science casework
We Recognize That There Are Many Different Perspectives and Lenses on This Report…

This is Why Public Comment is so Important!

Image source: https://imgur.com/gallery/1zZ6VSe
### Chapter Mapping

**25 Key Takeaways (KT) and 8 Future Considerations (FC)**

<table>
<thead>
<tr>
<th>Chapter 1</th>
<th>Chapter 2</th>
<th>Chapter 3</th>
<th>Chapter 4</th>
<th>Chapter 5</th>
<th>Chapter 6</th>
<th>Appendix 1</th>
<th>Appendix 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>INTRODUCTION</td>
<td>PRINCIPLES</td>
<td>SOURCES</td>
<td>RELIABILITY</td>
<td>RELEVANCE</td>
<td>TECHNOLOGY</td>
<td>HISTORY</td>
<td>TRAINING</td>
</tr>
<tr>
<td>(none)</td>
<td>KT #2.1</td>
<td>KT #4.1</td>
<td>KT #5.1</td>
<td>KT #6.1</td>
<td>KT #A1.1</td>
<td>FC #A2.1</td>
<td></td>
</tr>
<tr>
<td>KT #2.2</td>
<td>(none)</td>
<td>KT #4.2</td>
<td>KT #5.2</td>
<td>KT #6.2</td>
<td>KT #A1.2</td>
<td>FC #A2.2</td>
<td></td>
</tr>
<tr>
<td>KT #2.3</td>
<td>KT #4.3</td>
<td>KT #5.3</td>
<td>KT #6.2</td>
<td>KT #A1.3</td>
<td>FC #A2.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>KT #2.4</td>
<td>KT #4.4</td>
<td>KT #5.4</td>
<td>KT #A1.3</td>
<td>FC #A2.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KT #2.5</td>
<td>KT #4.5</td>
<td>KT #5.5</td>
<td>KT #A1.4</td>
<td>FC #A2.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KT #2.6</td>
<td>KT #4.6</td>
<td>KT #5.6</td>
<td>KT #A1.5</td>
<td>FC #A2.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16 Principles</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**16 Principles**

<table>
<thead>
<tr>
<th>Tables</th>
<th>Figures</th>
<th>Tables</th>
<th>Figures</th>
<th>Tables</th>
<th>Figures</th>
<th>Tables</th>
<th>Figures</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>9</td>
<td>3</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td></td>
<td>5</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>1</td>
<td>5</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Glossary & Acronyms:** 47 terms

**Bibliography:** 528 references
KEY TAKEAWAY #2.1: DNA mixtures, where the DNA of more than one individual is present in a sample, are inherently more difficult to interpret than single-source DNA samples.

Math Analogy to DNA Evidence

\[
2 + 2 = 4 \quad \text{Basic Arithmetic}
\]

\[
2 x^2 + x = 10 \quad \text{Algebra}
\]

\[
\int_{x=0}^{\infty} f(x) \, dx \quad \text{Calculus}
\]

Higher levels of uncertainty in determining contributing genotypes

Single-Source DNA Profile (DNA databasing)

Sexual Assault Evidence (2-person mixture with high-levels of DNA)

Touch Evidence (>2-person, low-level, complex mixtures perhaps involving relatives)

All DNA Samples Are Not Equivalent
Steps in Processing a DNA Evidence Sample (Single-Source or Mixture)

Collection/Storage/Characterization → Extraction/Quantitation → Amplification/Marker Sets → Separation/Detection → Data → Stats → Report

Measurement

Interpretation

DNA profile from person of interest (POI) compared

Gathering and Generating the Data

Understanding the Results

Electropherogram (EPG)

Written Report

Figure 2.1 (p. 24)
### Table 2.1 (pp. 27-28)

<table>
<thead>
<tr>
<th>Measurable Factor (units)</th>
<th>Validation Experiments to Demonstrate Reliability</th>
<th>Purpose in DNA Mixture Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1a) Peak Position (nucleotides)</td>
<td>Accuracy and precision studies to verify consistency in peak sizing and STR allele calls</td>
<td>To determine limits of peak sizing and accurate allele calls compared to an allelic ladder</td>
</tr>
<tr>
<td>1b) Peak Morphology or Resolution</td>
<td>Examination of peak height and width in allelic ladders and inspecting separation of similar length allelic ladder alleles (e.g., TH01 alleles 9 3 and 10) as quality control of kit and instrumentation</td>
<td>To examine CE separation resolution that can influence ability to accurately designate similar length STR alleles (e.g., Butler et al. 2004)</td>
</tr>
</tbody>
</table>

**Figure 2.2 (p. 29)**

**KEY TAKEAWAY #2.2:** Generating a DNA profile involves measuring the inherent physical properties of the sample. Interpreting a DNA profile involves assigning values that are not inherent to the sample. To do this, the DNA analyst uses their judgment, training, tools (including computer software), and experience, and considers factors such as case context.

Table 2.1. Measurable factors and features in a short tandem repeat (STR) DNA profile electropherogram (EPG) that influence DNA mixture interpretation with binary or probabilistic genotyping software (PGS) approaches.

Figure 2.2. Steps in DNA mixture interpretation first outlined by the UK Forensic Science Service (Clayton et al. 1998) and endorsed by the ISFG DNA Commission (Gill et al. 2006b).
2.3.1 Factors that Contribute to Increased Complexity in DNA Mixtures

There are at least three challenges that are fundamental to DNA mixture interpretation: (1) **stochastic variation**, which impacts recovered quantities of alleles from contributors and can lead to uncertainty in assigning alleles to genotypes and uncertainty in assigning genotypes to contributor profiles when examining small amounts of DNA, (2) **stutter products**, which create uncertainty through minor contributor(s) with alleles in the stutter positions of major contributor(s) alleles, and (3) **sharing of common alleles**, which influences the ability to estimate the number of contributors, particularly when combined with stochastic variation and the existence of stutter products that create uncertainty in deconvoluting mixture components.

**KEY TAKEAWAY #2.3:** The process of generating a DNA profile can produce stochastic or random variation and artifacts that contribute to the challenge of DNA mixture interpretation.
The challenge of genotype assignment increases with the number of contributors in a mixture due to the possibility of allele sharing (Alfonse et al. 2017). In addition, estimating the number of contributors in a DNA mixture becomes more uncertain when there are more contributors as noted in several publications (Paoletti et al. 2005, Buckleton et al. 2007, Coble et al. 2015). The frequency of occurrence for an allele from population data correlates to the degree of allele sharing that is expected if that allele is present in the crime scene DNA mixture. If mixture contributors are related, then even more allele sharing between contributors is expected. Thus, with more contributors to a mixture, more allele sharing occurs, which increases the complexity and ambiguity of interpretation (e.g., Dembinski et al. 2018, Lynch & Cotton 2018).

KEY TAKEAWAY #2.4: DNA mixtures vary in complexity, and the more complex the sample, the greater the uncertainty surrounding interpretation. Factors that contribute to complexity include the number of contributors, the quantity of DNA from each contributor, contributor mixture ratios, sample quality, and the degree of allele sharing.
Different Statistical Approaches Answer Different Questions

This point is emphasized in Principle 15

“Different statistical approaches can produce different numerical results as they utilize different information and/or models and answer different questions.”

Table 2.2. Different approaches used in statistical analysis of DNA and the questions addressed. RMP and MP are calculated for single-source DNA profiles (or deduced major profiles). CPI and LR are calculated for mixtures.

<table>
<thead>
<tr>
<th>Question</th>
<th>Approach (Reference)</th>
<th>Specific Requirements</th>
</tr>
</thead>
<tbody>
<tr>
<td>What is the probability of observing this profile in the population?</td>
<td>Profile Probability (or random match probability, RMP) (NRC 1996 for single-source samples; Bille et al. 2013 for mixtures)</td>
<td>For mixtures, an assumption that the major contributor can be distinguished from minor components so that specific genotypes in the major can be inferred</td>
</tr>
<tr>
<td>What is the probability of observing this profile in the population if we have already observed one person with this profile in this population?</td>
<td>Match Probability (MP) (Balding &amp; Nichols 1994, Weir 2001)</td>
<td>Use of conditional probabilities and a subpopulation correction</td>
</tr>
<tr>
<td>What is the probability that a person selected randomly in the population would be included (or not excluded) as a possible donor of the DNA typing result?</td>
<td>Combined Probability of Inclusion (CPI) (Bieber et al. 2016)</td>
<td>All alleles for all contributors are all present at the reported loci (i.e., cannot cope with allele drop-out that is expected with low quantities of DNA)</td>
</tr>
<tr>
<td>By how much do the DNA typing results support the person of interest (POI) being the donor under specific assumptions and propositions?</td>
<td>Likelihood Ratio (LR) (Evett &amp; Weir 1998)</td>
<td>An assumption as to the number of contributors and a specific pair of propositions</td>
</tr>
</tbody>
</table>
This point is emphasized in **Principle 14** “...continuous models use more information than discrete or binary approaches.”

**Table 2.3.** Comparison of approaches used in DNA mixture interpretation. CPI = combined probability of inclusion, mRMP = modified random match probability, LR = likelihood ratio. Adapted from ISFG 2015 workshop by John Butler and Simone Gittelson available at [https://strbase.nist.gov/training/TSFG2015-Basic-STR-Interpretation-Workshop.pdf](https://strbase.nist.gov/training/TSFG2015-Basic-STR-Interpretation-Workshop.pdf).

<table>
<thead>
<tr>
<th>Takes into account</th>
<th>Mathematically models</th>
</tr>
</thead>
<tbody>
<tr>
<td>Presence/absence of alleles</td>
<td>Possible genotypes based on peak heights</td>
</tr>
<tr>
<td><strong>Binary Approaches</strong></td>
<td></td>
</tr>
<tr>
<td>CPI</td>
<td>X</td>
</tr>
<tr>
<td>mRMP</td>
<td>X</td>
</tr>
<tr>
<td>LR (binary)</td>
<td>X</td>
</tr>
<tr>
<td><strong>Probabilistic Genotyping</strong></td>
<td></td>
</tr>
<tr>
<td>LR (discrete)</td>
<td>X</td>
</tr>
<tr>
<td>LR (continuous)</td>
<td>X</td>
</tr>
</tbody>
</table>

**KEY TAKEAWAY #2.5:** Continuous probabilistic genotyping software (PGS) methods utilize more information from a DNA profile than binary approaches.
Principles Described in Chapter 2

**Principle 1 [Biology]:** Our DNA generally remains unchanged across time and cell type.
- This principle enables meaningful comparison of DNA from a reference sample to an evidence sample deposited and/or collected at a different time and to verify identity in a “biometric” sense, where a previously analyzed DNA profile is checked against a new one for “authentication” purposes.

**Principle 2 [Biology]:** DNA transfers and persists and can be collected and analyzed.
- This principle of direct or primary transfer enables results to be generated from evidentiary DNA profiles to assist in crime-to-crime and crime-to-individual associations.

**Principle 3 [Biology]:** Forensic DNA profiles examine a limited number of specific sites in the human genome.
- This principle is a reminder that the entire DNA sequence is not examined with forensic tests. Statistical assessments of profile rarity are used based on inheritance patterns and population genetics.
Principle 7 [Relevance]: Answers from DNA results depend on questions asked and circumstances of the evidence.

The FBI DNA Advisory Board stated: “Proper statistical inference requires careful formulation of the question to be answered. Inference must take into account how and what data were collected, which, in turn, determine how the data are analyzed and interpreted” (DAB 2000). DNA results typically address questions at the sub-source level of the hierarchy of propositions (i.e., who could be the source of the DNA or is the DNA from the person of interest, Taroni et al. 2013). This principle is a reminder to users that DNA information by itself can only answer “who” questions, that is, questions of source not activity.

These 16 Principles Form the Foundation for DNA Mixture Interpretation

1. DNA stability across time and cell type
2. DNA transferability
3. Forensic profiles only examine a portion of the human genome
4. Established genetic inheritance patterns
5. Strength of evidence calculations use pop. gen.
6. Related DNA more similar than unrelated
7. Answers depend on questions asked
8. PCR can introduce artifacts
9. Peak positions and heights
10. Peak height variance
11. Stochastic variation impacts mixture ratios
12. Stutter peaks impact interpretation
13. Impacts on number of contributors estimate
14. Continuous models use more information
15. Results can differ with various approaches
16. Propositions impact strength of evidence
DNA mixture interpretation is performed in the face of uncertainty. As noted by Ian Evett and Bruce Weir in their 1998 book:

“The origins of crime scene stains are not known with certainty, although these stains may match samples from specific people. The language of probability is designed to allow numerical statements about uncertainty, and we need to recognize that probabilities are assigned by people rather than being inherent physical quantities” (Evett & Weir 1998, p. 21, emphasis added).

**KEY TAKEAWAY #2.6:** Likelihood ratios are not measurements. There is no single, correct likelihood ratio (LR). Different individuals and/or PGS systems often assign different LR values when presented with the same evidence because they base their judgment on different kits, protocols, models, assumptions, or computational algorithms. Empirical data for assessing the fitness for purpose of an analyst’s LR are therefore warranted.
Chapter 4: Reliability of DNA Mixture Measurements and Interpretation

(4.1.1) System Reliability vs Component Reliability
(4.1.2) Definitions of Measurement, Uncertainty, Assessment, and Interpretation
(4.1.3) Empirical Assessments of Reliability
(4.1.4) Factor Space and Factor Space Coverage
(4.1.5) Provider-User Responsibilities and Examples

(4.2) Data Sources Used to Examine Reliability
(4.3) Review of Publicly Available Data and Factor Space Coverage

(4.4) Discussion

(4.5) Thoughts on a Path Forward

KEY TAKEAWAY #4.1: The degree of reliability of a component or a system can be assessed using empirical data (when available) obtained through validation studies, interlaboratory studies, and proficiency tests.
Factor Space and Factor Space Coverage

- Is a new term but not a new concept
  - FBI QAS 8.3.2.1 requires laboratories to “include samples with a range of the number of contributors, template amounts, and mixture ratios expected to be interpreted in casework”

- Table 4.1 lists influencing factors with DNA mixture measurements and interpretations using PGS systems

- Factor space coverage is summarized for 3 STR kit developmental validation studies (Table 4.2), 60 published PGS studies (Table 4.3), 11 publicly available internal validation summaries (Table 4.5), 83 proficiency test data sets (Tables 4.6 and 4.7), and 18 interlaboratory studies (Table 4.8)
8 PGS studies were available and cited in the 2016 PCAST report. We examined and summarized 60 published PGS studies.

Table 4.3. Factor space coverage for published PGS validation data from peer-reviewed literature. Studies are grouped by PGS system and publication date. Studies listed on rows #6, #7, #10, #11, #12, #13, and #49 were part of the PCAST 2016 review. Nikola Osborne and Sarah Riman (NIST Associates) assisted with early versions of these summaries. NoC = number of contributors; N.E.S. = not explicitly stated in the referenced publication; N/A = not applicable; *comparison of multiple PGS systems are discussed in Table 4.4. †Inclusion of ranges is not meant to imply that all combinations of DNA quantities and mixture ratios were covered. §a 31-laboratory compilation (Bright et al. 2018) contained data from eight different STR kits: GlobalFiler, Identifiler Plus, NGM SEnet, PowerPlex Fusion 5C, PowerPlex Fusion 6C, PowerPlex ES17 Pro, PowerPlex ES17 Fast, and PowerPlex 16 HS.

<table>
<thead>
<tr>
<th>#</th>
<th>Reference</th>
<th>PGS System STR Kit</th>
<th>NoC Range</th>
<th># samples by NoC</th>
<th>Total DNA Quantity Range (pg)*</th>
<th>Mixture Ratio Range†</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Perlin &amp; Sinelnikov 2009</td>
<td>TrueAllele PowerPlex 16</td>
<td>2</td>
<td>40</td>
<td>125 to 1000</td>
<td>1:1 to 9:1</td>
</tr>
<tr>
<td>2</td>
<td>Perlin et al. 2011</td>
<td>TrueAllele Pro+Cofiler</td>
<td>2</td>
<td>16 adjudicated cases</td>
<td>N.E.S.</td>
<td>N.E.S.</td>
</tr>
<tr>
<td>3</td>
<td>Perlin et al. 2013</td>
<td>TrueAllele Pro+Cofiler</td>
<td>2/3</td>
<td>73/14 adjudicated cases</td>
<td>N.E.S.</td>
<td>N.E.S.</td>
</tr>
<tr>
<td>4</td>
<td>Ballantyne et al. 2013 (proof of concept)</td>
<td>TrueAllele Identifiler</td>
<td>2</td>
<td>2</td>
<td>N.E.S.</td>
<td>1:1</td>
</tr>
<tr>
<td>5</td>
<td>Perlin et al. 2014</td>
<td>TrueAllele PowerPlex 16</td>
<td>2/3/4</td>
<td>40/65/8 adjudicated cases</td>
<td>N.E.S.</td>
<td>N.E.S.</td>
</tr>
<tr>
<td>6</td>
<td>Perlin et al. 2015</td>
<td>TrueAllele Identifiler Plus</td>
<td>2/3/4/5</td>
<td>10/10/10 (5 donors)</td>
<td>200, 1000</td>
<td>1:1 to 32:16:15:2:1</td>
</tr>
<tr>
<td>7</td>
<td>Greenspoon et al. 2015</td>
<td>TrueAllele PowerPlex 16</td>
<td>1/2/3/4</td>
<td>11/18/15/7 (11 donors)</td>
<td>10 to 1000</td>
<td>1:1 to 17:1:1:1</td>
</tr>
<tr>
<td>59</td>
<td>You &amp; Balding 2019 (data from Steele et al. 2016)</td>
<td>*multiple NGM SEnet</td>
<td>1/2/3</td>
<td>36/24/12 (36 donors)</td>
<td>4 to 328</td>
<td>1:1 to 16:1; 1:1:1 to 16:4:1</td>
</tr>
<tr>
<td>60</td>
<td>Riman et al. 2021</td>
<td>*multiple GlobalFiler</td>
<td>2/3/4</td>
<td>154/147/127 (PROVEDI data)</td>
<td>30 to 750</td>
<td>1:1 to 1:9; ... 1:1:1:1 to 1:9:9:1</td>
</tr>
</tbody>
</table>
## Published PGS Comparison Studies

11 + 1 NIST study (conducted during our review)

### Table 4.4

<table>
<thead>
<tr>
<th>PGS Systems Compared</th>
<th>Samples Tested</th>
<th>Observations Made</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Reference</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A pre-print version is available at <a href="https://www.biorxiv.org/content/10.1101/2021.05.26.445891v1">https://www.biorxiv.org/content/10.1101/2021.05.26.445891v1</a></td>
<td><strong>Supplementary Tables 4 and 5 contain all LR values and provide an example</strong></td>
<td>Provided LR values for 1279 Hp-true tests (Supplemental Table 4) and 1279 Hd-true tests (Supplemental Table 5) for each software; explored LR distributions observed and used ROC plots, scatter plots, histograms with distribution of differences; evaluated apparent discrepancies between PGS models, adventitious exclusionary and inclusionary support, and verbal equivalent discordance; the authors reported: “in certain cases differences in numerical LR values from both software resulted in differences in one or more than one verbal categories (Table 8). These differences were substantially more with low template minor contributors and higher [number of contributors]…”</td>
</tr>
<tr>
<td>EuroForMix (v2.1.0)</td>
<td>Examined 154 two-person, 147 three-person, and 127 four-person mixtures from the PROVEDIt dataset; see Supplemental Table 4 in their article</td>
<td></td>
</tr>
<tr>
<td>STRmix (v2.6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Riman et al. 2021</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Chapter 5: Context and Relevance Related to DNA Mixture Interpretation

High-Sensitivity Methods Impact Scientific Relevance

KEY TAKEAWAY #5.3: Highly sensitive methods increase the likelihood of detecting contaminating DNA that might affect an investigation. Contamination avoidance procedures should be robust both at the crime scene and in the laboratory.

Case Context is Important to Scientific Relevance

KEY TAKEAWAY #5.5: The fact that DNA transfers easily between objects does not negate the value of DNA evidence. However, the value of DNA evidence depends on the circumstances of the case.
Public Comments Received

• Perspectives on our report have been offered from practitioners, researchers, prosecutors, defense attorneys, consultants, PGS vendors, and multiple stakeholder organizations

• We are processing these public comments and will post the public comments on our website to coincide with the release of the final report
AAFS 2019 Workshop with NIST DNA Team and Resource Group Members

Thank you to our Resource Group members and their agencies who permitted them to assist us in this study

Email received after our last meeting: John and NIST colleagues: thank you very much for the invitation to participate in this illustrious group. I gained a great deal from our robust discussion and enjoyed it thoroughly. Many viewpoints always makes the product stronger.
Thank you for your attention!

John Butler

john.butler@nist.gov

https://www.nist.gov/topics/forensic-science