The Scientific Uncertainty, Statistical Framework, and Court Admissibility of the New Generation of Forensic Molecular Evidence

Zhiyuan Jin[1, *], Zhangsheng Sun[2]
[1]The University of New South Wales, Sydney NSW 2033, Australia
[2]The Australian National University, Canberra ACT 2601, Australia
*Corresponding email: 2774323797@qq.com

Abstract

This paper systematically examines the core issues of scientific uncertainty quantification, statistical framework construction, and court admissibility of the new generation of forensic molecular evidence. This article reviews the advantages and limitations of short tandem repeat (STR) typing and mitochondrial hair typing. It notes that traditional binary determinations are not applicable to degraded samples, mixed samples, and situations involving multiple copy numbers. The article advocates probabilistic interpretation based on likelihood ratios and Bayesian reasoning. This paper critiques the independence assumption and database dependence of the Monika-Welzel method. This method was developed to calculate loop-mediated isothermal amplification (LAMP) DnSK values under multi-copy conditions. An improved scheme based on simple random error rate (SRE) is proposed. It can better reflect experimental errors and copy number heterogeneity. Two complementary probability metrics are constructed within the formal concept analysis (FCA) framework to achieve a layered assessment from technical reliability to case attribution. Using the Kline-Gray half-code hair case as an empirical example, the role of blind testing, error propagation, and sensitivity analysis in enhancing the admissibility of evidence is demonstrated. It is recommended to promote cross-laboratory standardization, large-scale blind testing, automated data platforms, and scientific-legal communication guidelines. This will facilitate transparency, repeatability, and fair application of forensic molecular evidence in judicial processes.

Keywords

Forensic molecular evidence, Scientific uncertainty, Statistical framework, Court admissibility

Introduction

Overview of forensics

(1) Molecular typing systems

The forensic molecular typing system, especially the STR typing based on polymerase chain reaction technology, has become the most core and widely accepted form of evidence in modern forensic science. This system generates highly discriminative genetic profiles by detecting microsatellite markers in specific regions of the genome. Under ideal conditions, the random match probability of STR typing can be as low as one in a billion or even lower, demonstrating strong probative power in excluding suspects or establishing individual identity [1]. However, as the complexity of cases increases, particularly when faced with degradation, mixtures, or trace samples, the traditional binary determination model faces severe challenges. At this point, probabilistic genetic typing software emerges, which no longer simply provides a “yes” or “no” conclusion, but quantifies the support level of observed data under different hypotheses through likelihood ratio calculations [2]. This shift from determinism to probability marks a new stage in forensic deoxyribonucleic acid (DNA) analysis dominated by statistics.

While STR technology is relatively mature, its application is not without limitations. Database dependence is an inherent weakness of this system. The calculation of random match probabilities heavily relies on the allele frequency database of the target population. If the database used fails to accurately reflect the specific subpopulation to which the defendant belongs, or if there is sampling bias, it may lead to distortion in statistical estimates, thereby affecting the court’s judgment on the

weight of evidence. For the interpretation of mixed stains, different laboratories may arrive at different interpretations due to variations in threshold settings and algorithms used. This methodological heterogeneity has sparked ongoing debates about the reliability and validity of the testing process [3].

As Stern et al. pointed out, assessing the validity of forensic evidence requires not only attention to the precision of the technology itself but also consideration of its robustness in practical operational environments and potential sources of human error. With the exploration of new generation sequencing technologies in the forensic field, new types of molecular evidence such as copy number variations are gradually coming into view. This type of evidence often involves more complex biological mechanisms and statistical models, with sources of uncertainty that are more diverse than traditional STR. For example, the expression levels of copy number variations may be affected by tissue specificity, individual developmental stages, and environmental factors.

This results in an ambiguous correspondence between phenotype and genotype [4]. In this case, simple frequency counting methods cannot accurately characterize the uncertainty of evidence. Therefore, more refined probabilistic frameworks are required to model the associations between continuous variables and discrete events. So, understanding the intrinsic mechanisms and statistical properties of these new molecular typing systems is crucial for ensuring their fair application in judicial processes.

It is worth noting that no molecular typing system can provide absolute truth in isolation; its value lies in providing quantitative information support for fact-finders. When courts adopt such evidence, they must fully disclose its limitations and the specific sources of uncertainty, so that the jury or judge can correctly weigh its significance in the overall case [5].

Only in this way can a balance be found between the pursuit of scientific rigor and the assurance of judicial fairness, preventing the neglect of a comprehensive review of other relevant evidence due to an over-reliance on technical means.

(2) Hair typing

Hair typing, as a classic method in forensic evidence examination, has a history that predates the intervention of molecular biology techniques. Traditionally, examiners compared the morphological characteristics of hair, such as medullary structure, pigment distribution, and cortical patterns, between the evidence collected from the scene and samples from suspects [6].

However, this visually based method has long faced skepticism regarding its scientific validity, especially in the absence of unified quantitative standards, making it difficult to ensure consistency of conclusions among different experts. With the introduction of mitochondrial DNA sequencing technology, hair analysis has evolved from simple morphological comparison to genetic information analysis. This greatly improves the discriminative power of such evidence. Mitochondrial DNA is mainly located in the cortex of the hair shaft. Owing to its maternal inheritance and high copy number, it can yield valid genetic marker information even from degraded or trace samples with severely compromised nuclear DNA [7].

The interpretation process of hair typing remains fraught with uncertainty. First, hair is not a static biological marker. The distribution of drug metabolites in hair is affected by individual physiological conditions, exposure frequency, and time windows. This complicates the distinction between a single exposure event and long-term use. Second, even if the same mitochondrial sequence is detected at the same locus, it cannot be directly equated to the same source, as that sequence may be shared within families or even broader populations.

Therefore, genetic matching alone cannot exclude the possibility of innocent individuals; it must be combined with a statistical framework to assess the evidential weight of such matches. Early probabilistic studies sought to quantify evidential value by calculating random match probabilities. However, these models frequently oversimplified the complexity of population genetic structures and could not adequately capture heterogeneity across subpopulations [8].

Modern forensic practice emphasizes that the interpretation of hair evidence should be viewed as a multi-stage process, requiring the integration of morphological, chemical analysis, and molecular genetic data [9]. Statistical assessments are no longer confined to simple frequency counts. Instead, probabilistic models that accommodate multiple hypotheses should be constructed. These models can accurately estimate

likelihood ratios for the observed data under different competing hypotheses. Furthermore, the diversity of sources for hair samples has also increased the difficulty of analysis. When evidence may come from multiple individuals, traditional binary determinations fail, necessitating the use of more complex mathematical methods to estimate attribution probabilities [10]. It is noteworthy that courts are applying increasingly stringent scrutiny to the admissibility of hair evidence. Testing agencies are therefore required not only to deliver qualitative conclusions but also to disclose detailed information regarding sources of methodological uncertainty and the logical basis of statistical reasoning. Only through a rigorous statistical framework and transparent error rate reporting can hair typing maintain its scientific status and judicial credibility within the new generation of molecular evidence systems.

(3) Summary

The forensic molecular evidence system is undergoing a profound transformation from traditional morphological comparison to high-precision molecular typing and complex statistical inference. When the above-mentioned short tandem repeat systems and mitochondrial DNA analysis of hair are reviewed, qualitative advances in individual identification capacity are noted for both. Nevertheless, their core logic depends on the precise quantification and probabilistic representation of biological variability. However, this technological advancement has not eliminated scientific uncertainty. Rather, due to the increased complexity of its models, the sources of error have become more concealed and diverse. Whether it is the threshold setting for STR mixed stains or the shared probability of mitochondrial sequences within populations, it indicates that no single technical path can provide absolutely certain conclusions. Its value lies in providing quantifiable information support for fact-finders through statistical tools such as likelihood ratios [11].

In this context, the key challenge facing the new generation of forensic molecular evidence is how to construct a robust statistical framework to handle multi-source uncertainty. The traditional binary decision model has become difficult to adapt to the real-world demands of degraded, trace, or mixed samples, necessitating a shift towards probability genotyping methods based on Bayesian inference or frequentist approaches. This methodological shift requires examiners to possess not only biological expertise but also a deep understanding of the assumptions and limitations of statistical models. For example, simple frequency counting methods are often inadequate for copy number variations and epigenetic markers. Therefore, more refined probability distributions are required to characterize the relationship between continuous variables and discrete events [12,13]. Assessing the validity of such evidence cannot focus solely on technical accuracy in the laboratory. It must also consider its robustness in actual judicial environments and the potential impact of human biases.

The standards for admissibility in court have also evolved, emphasizing comprehensive disclosure of the evidence generation process. Modern judicial practice requires expert witnesses to clearly articulate the logical foundations of statistical reasoning, the representa-tiveness of databases, and the scope of methodological uncertainties. This enables juries to properly weigh evidential weight.

This means that forensic science is no longer merely a technical means of “discovering the truth”, but rather a knowledge production activity that must undergo rigorous validation within a legal framework. Only through transparent error rate reporting and a multi-stage comprehensive evaluation mechanism can a balance be struck between the pursuit of scientific rigor and the safeguarding of judicial fairness. This prevents over-reliance on technical means at the expense of comprehensive case review [14]. In summary, the future development of forensic molecular evidence will heavily depend on the deep integration of statistical theory and molecular biology technology, as well as the legal community’s rational understanding and normative application of scientific uncertainty.

Research focus

Focusing on the quantification of scientific uncertainty in the new generation of forensic molecular evidence, the construction of statistical frameworks, and their admissibility review in court procedures. Although established technologies including STR typing occupy a core position in individual identification, traditional binary decision models exhibit substantial limitations. These limitations emerge under complex scenarios involving degraded samples, trace evidence, and copy number variations [15]. As Stern et al. pointed out, assessing the validity of forensic evidence relies not only

on technical accuracy but also on considering its robustness in actual operational environments and potential sources of human error. Therefore, this study seeks to transcend the limits of existing approaches. It explores probabilistic models that capture associations between continuous variables and discrete events. These models can accurately estimate likelihood ratios of observed data under different competing hypotheses and deliver more informative quantitative support for fact-finders.

Improvements to the LAMP-DnSK value calculation method in cases of copy number variations. In response to the logical flaws present in the Monika-Welzel method when handling such data, this study proposes an alternative based on simple random error rates, aiming to more accurately reflect the sources of uncertainty in the experimental process. Additionally, this study explores the examination workflow for semi-coded hair evidence. Taking the Kline-Gray archaeological case as an example, it analyzes how to establish evidential weight via rigorous statistical reasoning when a complete genetic map is unavailable. This process requires the integration of morphological, chemical analysis, and molecular genetic data to construct a multi-stage probability model to address attribution challenges that a single technical path cannot resolve.

This study pays particular attention to the admissibility standards for new molecular evidence in court. As judicial practice demands higher reliability of scientific evidence, expert witnesses must clearly disclose the logical foundations of statistical reasoning, the representativeness of databases, and the range of uncertainties in methods. This study integrates the contextual admissibility theory advanced by Edmond and Roach. It analyses how the United States and other jurisdictions strike a balance between scientific rigor and procedural fairness.

This avoids over-reliance on technical tools that undermines holistic case review. By integrating statistical theory with molecular biological mechanisms, this study aims to develop a framework for transparent error rate reporting and multi-stage comprehensive evaluation. This framework ensures that new-generation forensic molecular evidence complies strictly with legal norms while pursuing scientific truth and protects the legitimate rights of defendants [16]. Ultimately, it is hoped that this interdisciplinary integration will provide a theoretical basis and practical guidelines for the rational application of forensic science within the judicial system, promoting effective dialogue between scientific uncertainty and legal certainty.

Research background and significance

Basic issues of molecular biological evidence

Molecular biological evidence occupies a core position in contemporary judicial practice, and its application has expanded from traditional DNA typing to a broader field of biomarker detection. However, with the iteration of technology and the increase in complexity, how to accurately assess the scientific uncertainty of such evidence has become a key issue that needs to be addressed [17]. Statistics, as a scientific tool for handling uncertainty, plays an indispensable role in forensic science, aiming to support court decisions by quantifying data variation. Although molecular biological evidence is often treated as the “gold standard”, its interpretation is not entirely objective. Instead, it is heavily shaped by statistical model assumptions, sample quality, and laboratory operational standards [18].

The interpretation of molecular biological evidence must be placed within a strict probabilistic framework. Traditionally, courts tend to rely on qualitative descriptions or simple frequency statistics, but modern forensic science requires the use of more rigorous statistical measures such as Bayesian inference or likelihood ratios. These methods can clearly distinguish the strength of evidence under different hypotheses, thereby avoiding misleading the jury. For example, in trace DNA analysis, background noise, contamination risks, and the deconstruction of mixed samples introduce significant technical variability, and without appropriate statistical corrections, it is easy to lead to erroneous associative judgments [19,20].

Furthermore, the reproducibility of data and verification mechanisms are the cornerstones of establishing the reliability of evidence. A report from the National Academy of Sciences of the United States noted that numerous forensic techniques lack adequate empirical support. This observation has prompted the academic community to advocate for strengthened basic scientific research to consolidate the methodological foundation. In this context, establishing standardized quality control procedures and transparent algorithmic logic is particularly important to ensure consistency of results

between different laboratories.

The legal system's review of the admissibility of scientific evidence is becoming increasingly stringent. Judges and lawyers not only need to understand technical principles but also must master the statistical knowledge to evaluate the validity of expert testimony. This interdisciplinary dialogue requires scientists to clearly define the range of uncertainty in their conclusions, while legal professionals need to judge whether the probative value of the evidence meets the legal standards. It is worth noting that the advancement of next-generation technologies including high-throughput sequencing produces massive datasets.

These datas introduce new analytical challenges such as multiple comparison corrections and computational complexity, which place greater demands on existing statistical frameworks. Therefore, constructing a comprehensive evaluation system that reflects both the technical limitations and meets legal proof standards is the core task of current forensic molecular biology research. Only by organically combining scientific rigor with legal fairness can the proper use of molecular biological evidence in judicial procedures be ensured, thereby minimizing the risk of misjudgment and enhancing judicial credibility.

LAMP-DnSK value in the context of multiple copy numbers

Within the quantitative assessment framework for forensic molecular biological evidence, the LAMP-DnSK value acts as a statistical indicator for measuring evidential strength [21]. Its calculation logic directly shapes how courts interpret scientific uncertainty. However, when the analysis targets genes with multiple copy numbers or repetitive sequences, conventional single-copy hypothesis models frequently fail to fully capture data variability. This introduces systematic bias in LAMP-DnSK value estimation. This bias not only affects the accuracy of the likelihood ratio but may also raise doubts about the reliability of the evidence in judicial practice [22]. Therefore, exploring the calculation methods and limitations of the LAMP-DnSK value in the context of multiple copy numbers is of critical significance for improving the forensic statistical framework.

Traditionally, the derivation of LAMP-DnSK values depends on the allele frequency distribution at a single locus. Under multiple-copy conditions, however, genomic copy number variation (CNV) creates substantial interindividual differences in actual copy numbers. Ignoring this biological fact and directly applying standard formulas will result in distorted probability estimates [23].

For example, in high-throughput sequencing data, fluctuations in copy numbers between different samples may arise from biological variation rather than technical error. If not distinguished, this could incorrectly amplify or diminish the probative value of the evidence. Additionally, differences in amplification efficiency during laboratory operations can also be non-linearly amplified in the context of multiple copies, further exacerbating the uncertainty of the results.

To address this challenge, researchers have begun to introduce more complex statistical models to correct for copy number effects [24]. A common strategy is to integrate prior information through Bayesian hierarchical models to dynamically adjust the weight of the LAMP-DnSK value, thereby reflecting the heterogeneity of copy numbers among individuals. This method can more realistically simulate the data generation process and improve the robustness of estimates.

However, the complexity of the model also increases computational costs and interpretative difficulties, which raises higher demands for the clarity of expert testimony in court. Judges and juries need to understand that the LAMP-DnSK value in the context of multiple copy numbers is not a fixed constant, but a range estimate influenced by various factors. Legally, this technical complexity is directly related to the review of the admissibility of evidence. Courts typically require experts to provide clear methodological bases to demonstrate that their statistical processing can effectively control sources of error. If the calculation of the LAMP-DnSK value fails to adequately consider the additional variability brought about by multiple copy numbers, the evidence may be excluded due to insufficient scientific validation [25].

Therefore, establishing a standardized calculation specification for the LAMP-DnSK value in the context of multiple copy numbers is not only a scientific necessity but also an important guarantee for ensuring judicial fairness. Future research should focus on developing intuitive and easy-to-understand visualization tools to help non-professional audiences grasp the meaning of this complex statistical quantity, thereby

promoting effective communication between science and law [26].

Method

Introduction to the method

The courtroom application of forensic molecular evidence centers on transforming the raw data generated from laboratory tests into legally valid probabilistic conclusions. This transformation process is not a simple numerical mapping but a rigorous statistical inference and uncertainty quantification process. Within the framework of next-generation hair analysis technology, challenges arise especially when testing semi-coded hair or trace biological samples. Traditional deterministic identification models cannot satisfy the high reliability standards for scientific evidence. Therefore, this study introduces and refines a probabilistic assessment method grounded in Bayesian networks and likelihood ratio frameworks. It seeks to deliver a more transparent, reproducible, and logically coherent pathway for evidence interpretation for the court.

This method first establishes a comprehensive error model covering the entire process of sample collection, preprocessing, amplification reaction, and result interpretation. This model not only considers random experimental errors, such as fluctuations in polymerase chain reaction (PCR) amplification efficiency and sequencing background noise, but also specifically incorporates sources of systematic errors like cognitive biases. Sensitivity analyses of key parameters are performed through Monte Carlo simulation techniques to identify threshold intervals corresponding to different confidence levels. On this basis, the likelihood ratio (LR) is adopted as the core metric for evidence strength. The likelihood ratio is defined as the probability of observing the current evidence data under the prosecution's hypothesis, divided by the probability of observing the same evidence data under the defense's hypothesis, mathematically expressed as follows:

$$LR = \frac{P(E|H_p)}{P(E|H_d)} \quad (1)$$

Where E represents the observed molecular evidence data, and H_p and H_d represent the competing hypotheses of the prosecution and defense, respectively. Compared to traditional qualitative matching or single frequency calculations, the likelihood ratio can more accurately reflect the degree of support that the evidence provides for the hypotheses, avoiding the occurrence of the "prosecutor's fallacy". Furthermore, to handle complex genetic structures under multiple-copy conditions, this study presents an improved calculation method for LAMP-DnSK values. This method relies on a weighted averaging mechanism to mitigate estimation bias stemming from locus deletions and heterogeneity [27].

The validity of any statistical framework relies on the reasonableness of its underlying assumptions. The method design incorporates strict validation steps, including blind testing using reference standards with known mixing ratios to assess the model's calibration. Combined with a risk science perspective, the discussion focuses on reasonable definition of the uncertainty range under conditions of incomplete information, ensuring that the final submitted probabilistic conclusions are both conservative and sufficiently discriminative [28].

This systematic approach not only improves the scientific rigor of forensic molecular evidence but also offers expert witnesses in court a clear communication tool. It helps judges and juries accurately grasp the meaning and limitations of the evidence and thus facilitates judicial fairness.

Two probabilistic measures within the FCA framework

In the statistical assessment of forensic scientific evidence, FCA provides a unique structured perspective for handling the conceptual hierarchy and attribute associations within complex datasets. For the next-generation molecular evidence, this study introduces two complementary probabilistic measures within the FCA framework, aiming to more precisely quantify the degree of support that evidence provides for competing hypotheses. The first measure is a direct estimation based on conditional probabilities, relying on the frequency distribution of specific loci or markers in the reference database. This method is intuitive and easy for the court to understand, but its limitation lies in often overlooking the dependencies between samples and systematic biases in the experimental process [29].

As noted in literature, relying solely on frequency calculations while ignoring Bayesian logic may lead to misjudgments regarding the strength of the evidence. Therefore, relying solely on this basic measure is insufficient to meet the stringent requirements of modern justice for evidence reliability. The second measure is

based on the posterior probability updating mechanism under the likelihood ratio framework, which transforms the concept lattice structure generated by FCA into node connection weights in a probabilistic graphical model. In this setup, each concept node represents a specific combination of molecular features, and its probability value reflects not only the prior likelihood of that feature's occurrence but also integrates the uncertainty propagation from upstream detection steps.

Specifically, FCA's derivation operators are used to identify key subsets of features and construct local likelihood ratio functions accordingly. In this way, the originally discrete qualitative matching is transformed into integral operations on a continuous probability manifold [30]. This process effectively addresses the common "data fragmentation" issue in traditional methods, allowing for the extraction of statistically significant evidential signals even in high-noise backgrounds. Literature emphasizes that combining multivariate statistical methods with forensic logic is a key pathway to enhancing the interpretative power of evidence.

The synergy of these two metrics is reflected in their capture of different dimensions of evidence: The former focuses on the descriptive statistical characteristics of the data, while the latter concentrates on the rigor of inferential logic. In practical applications, the first metric is adopted for preliminary screening to exclude obviously irrelevant hypothesis spaces; the second metric is subsequently used to finely rank the remaining high-probability hypotheses [31].

This layered strategy not only improves computational efficiency but also enhances the interpretability of the results. Particularly when dealing with trace evidence such as low-template hair, where the sample size is small and variability is high, a single probability measure can easily produce extreme values or unstable conclusions. In contrast, the likelihood ratio method combined with the FCA structure can smooth these abnormal fluctuations through the hierarchical relationships of concept lattices, thereby providing a more robust assessment of evidence strength.

Furthermore, this dual-track metric system provides a dual validation mechanism for expert court witnesses. It helps demonstrate the transparency of the reasoning process and the logical closure during cross-examination. This improves the admissibility and persuasiveness of evidence presented in court.

The admissibility of next-generation hair analysis in court

Methodological improvements for LAMP-DnSK values in cases of multiple copy numbers

(1) Issues with the Monika-Welzel method for calculating LAMP-DnSK values

The Monika-Welzel method was once regarded as a standardized mathematical tool in the early assessment of hair evidence, aimed at deriving the prosecution's probability by quantifying feature differences. However, forensic molecular biology techniques keep iterating, and statistical theories continue to advance. This method has revealed several structural flaws in the calculation of LAMP-DnSK values. These flaws greatly restrict its applicability and admissibility in complex cases [32]. Firstly, the model frequently relies on simplified independence assumptions when dealing with data distributions under multiple copy numbers. It assumes that variations in distinct morphological features are independent.

These variations follow specific parameter distributions. This assumption hardly holds for real samples. Strong covariance exists between features. Such covariance emerges when hair is affected by different growth stages or environmental factors. As a result, the calculated likelihood ratios deviate from the true strength of evidence. Secondly, the Monika-Welzel method is highly dependent on background databases but lacks a sensitivity analysis mechanism for insufficient database representativeness. The reference population may have a limited sample size or suffer from selection bias.

Under such circumstances, this method fails to effectively quantify the introduced systematic errors. It may overestimate or underestimate the evidential value. This outcome conflicts with modern requirements for uncertainty quantification. Additionally, the method does not adequately incorporate a dynamic adjustment mechanism for the Random Error Rate. In actual testing processes, observer subjective judgments and instrument noise may introduce random errors. However, the Monika-Welzel framework generally treats these errors as fixed constants instead of random variables. It cannot capture how fluctuations in experimental conditions affect the final conclusions [33].

In contrast, calculation methods based on simple random error rates can more flexibly integrate observational noise, providing more robust probability estimates. Legally, the limitations of this statistical method directly affect the admissibility review of evidence.

According to relevant judicial practices, if experts fail to clearly articulate the assumptions and potential biases of their statistical models, courts may exclude the evidence on the grounds of an unstable methodological foundation. Particularly in cases involving trace evidence or mixed-source hair, the deterministic outputs of traditional methods often obscure the essence of scientific uncertainty, misleading juries in their understanding of evidence strength. It is necessary to refine the calculation logic of LAMP-DnSK values and adopt a more rigorous uncertainty propagation mechanism.

This is a technical requirement to improve the accuracy of forensic science. It is also a legal necessity. It satisfies the rising demands for transparent and reliable scientific evidence in court. Future research may integrate Bayesian networks or multivariate statistical models to reconstruct existing algorithms. This strategy addresses the inherent limitations of the Monika-Welzel method. These limitations arise from multidimensional feature coupling and error accumulation. Accordingly, the outputs of next-generation hair analysis can obtain a sound scientific foundation for courtroom justification.

(2) The advantages of calculating LAMP-DnSK values based on simple random error rate

Compared to the rigid dependence on fixed parameters in the Monika-Welzel method, the calculation framework based on SRE provides a more flexible and statistically intuitive alternative for evaluating LAMP-DnSK values. Its core advantage lies in explicitly modeling the uncertainty in the observation process as random variables, rather than simplifying it to negligible constants or fixed system biases. In the analysis of the next generation of forensic molecular evidence, especially when involving techniques such as LAMP, reaction efficiency, primer specificity, and background noise all exhibit inherent random fluctuations.

The SRE method quantifies how these fluctuations affect the final matching probability. It can better reflect the characteristics of data distributions under experimental conditions. This prevents likelihood ratio distortions stemming from the overly restrictive assumptions adopted in traditional methods.

The SRE framework allows examiners to dynamically adjust error boundaries when calculating LAMP-DnSK values to accommodate different sample qualities and instrument precisions.

This flexibility is particularly crucial when dealing with degraded DNA or trace biological samples, as such samples often come with higher technical failure rates and signal variability. By incorporating random error rates into the probability model, this method not only enhances the reproducibility of results but also strengthens its scientific defensibility in the face of courtroom scrutiny.

Notably, this random-based modeling approach aligns closely with the principles of Bayesian inference, aiding in the construction of a more transparent evidence strength expression system. In judicial practice, juries and judges often struggle to understand complex mathematical derivations, but if it can be clearly demonstrated how “random errors affect conclusions”, it can significantly enhance the acceptability of the evidence. Furthermore, the SRE method exhibits unique adaptability in data integration when handling multiple copy number situations. It does not require all feature dimensions to be strictly independent but instead indirectly addresses the covariance relationships between features through error propagation mechanisms, thereby reducing the complexity of model construction. Parametric methods rely heavily on large background databases.

By contrast, the SRE framework is less dependent on reference data. This feature makes it better suited for resource-limited scenarios or cases with insufficient sample sizes from special populations. This low-dependency characteristic reduces systematic errors arising from selection bias. It yields more robust calculated LAMP-DnSK values. From a legal perspective, evidence reports derived via the SRE method usually contain explicit uncertainty intervals. Such intervals meet the rising requirements for transparency of scientific evidence in modern courts. Expert witnesses can explain to the court that despite random fluctuations, the prosecution’s hypothesis remains significantly superior to the defense’s hypothesis at a specific confidence level.

Therefore, adopting calculation logic based on simple random error rates represents a methodological optimization. It also serves as an important bridge. This

bridge links the rigor of forensic science to procedural justice in court. It provides reliable statistical support for admissibility reviews of next-generation hair and molecular evidence.

The examination process of semi-coded hair evidence - Taking the Kline-Gray archaeological case as an example

Semi-coded hair evidence occupies a unique position in forensic science, as its examination process must adhere to traditional morphological comparison logic while also incorporating the statistical rigor of modern molecular biology. The Kline-Gray archaeological case provides a highly representative empirical example for exploring the admissibility of such evidence in court. In this case, although the recovered hair samples did not carry complete nuclear DNA information, their mitochondrial sequences and microscopic morphological features constituted a key "semi-coded" data combination. This situation clearly illustrates the limitations of relying on only one technical pathway.

Morphological comparison is vulnerable to subjective judgment. Exclusive reliance on molecular markers, by contrast, can cause information loss when samples degrade. Therefore, constructing a comprehensive examination framework that integrates multidimensional information is particularly urgent.

For the specific samples in the Kline-Gray case, the research team adopted a stratified assessment strategy. First, standardized records were established for features such as the distribution of medulla and cortex pigments in hair shafts under a microscope, to establish preliminary exclusion or association hypotheses. Subsequently, the LAMP-DnSK value calculation model was introduced, taking into account the random error rate to quantify the impact of technical noise in molecular detection on the conclusions. Unlike the conventional Monika-Welzel method, this approach does not assume full independence among all features. Instead, it accounts for potential coupling relationships between features via error propagation mechanisms. This design yields a more accurate representation of evidence strength.

At the statistical analysis level, the research focused on deriving the LR under the hypotheses of the prosecution and defense. Background databases may lack representativeness for certain historical periods or regional populations. Such biases may arise. For this reason, sensitivity testing is specially incorporated into the analytical workflow. It evaluates the systematic influence of reference group selection on final probability estimates. This transparent quantification of uncertainty directly addresses the core concerns of the judicial process regarding the reliability of scientific evidence.

From the perspective of court admissibility, the examination process in the Kline-Gray case indicates that the credibility of half-code hair evidence highly depends on the transparency of methodology and the robustness of statistical models [34]. Expert witnesses should not only provide definitive matching conclusions but also clearly demonstrate how random errors define confidence intervals and the differences in probability distributions under different hypotheses. This calculation logic based on simple random error rates not only enhances the reproducibility of results but also provides the jury with an intuitive cognitive framework for understanding complex scientific concepts.

Furthermore, this case highlights the value of interdisciplinary collaboration. Forensic pathologists, molecular biologists, and statisticians should jointly develop standardized testing workflows. This ensures that every step of derivation conforms to scientific consensus. In summary, the successful practice of the Kline-Gray case offers an important demonstration. When the examination workflow for half-code hair evidence fully discloses scientific uncertainty and adopts a robust statistical framework, the probative power of such evidence in court improves substantially. It can satisfy the increasingly stringent standards for evidence admissibility.

Discussion

Innovations

Multidimensional theoretical breakthroughs and methodological innovations have been realized in statistical modeling and courtroom application for the new generation of forensic molecular evidence. The core innovations cover three aspects. They include the reconstruction of LAMP-DnSK calculation logic under multi-copy scenarios, the refined definition of probability measures within the FCA framework, and the systematic integration of testing workflows for half-code hair evidence. First, in response to the systematic bias issues encountered by the traditional Monika-Welzel method when dealing with complex biological samples,

this study proposed a new calculation paradigm based on simple random error rates. This paradigm abandons dependence on a single idealized assumption. It introduces an error distribution model that better fits the real laboratory operating environment.

This model greatly improves the robustness of LAMP-DnSK values for low-template and degraded samples. This improvement resolves interpretative difficulties. Such difficulties arise from excessive numerical fluctuations of earlier methods under multi-copy conditions. It also supplies court experts with evidence strength quantification indicators grounded in stronger statistical support. This development effectively addresses the demand for communicating scientific uncertainty in judicial settings. It further extends the application scope of the FCA framework in forensic statistics.

Two probability measures with distinct legal implications are clearly distinguished and defined. One evaluates the internal consistency probability of technical reliability. The other infers the external attribution probability related to case fact relevance. This dual measurement system breaks the conventional practice in traditional forensic reports. Such reports tend to conflate technical validation and case inference. With this system, judges and juries can more readily identify the scientific basis of evidence sources. They can also recognize the applicable limits of such evidence in specific cases. This methodological refinement directly addresses the requirement of modern judicial systems. Such systems demand layered scrutiny of the reliability and validity of scientific evidence. It offers a rigorous logical framework. This framework helps improve the admissibility of forensic expert opinions.

Using the Kline-Gray archaeological case as an empirical carrier, this study has constructed standardized semi-coding hair evidence testing process. This process deeply integrates molecular biology testing techniques with the quality control standards of forensic science, minimizing the impact of human factors on the final conclusion through strict blind testing design and cross-validation mechanisms. Earlier studies mainly focused on validating individual technologies.

By contrast, this work stresses standardization across the full workflow, from sample collection to courtroom presentation. It specifically illustrates how to preserve the integrity and persuasiveness of the evidence chain. This challenge arises when handling atypical or historically legacy samples without a complete reference system. This innovative full-process perspective enriches the theoretical system of forensic molecular evidence. It also delivers an actionable technical route.

Researchers can use this route to tackle the increasingly complex new case types emerging in current judicial practice. In summary, this study advances the transformation of forensic molecular evidence from empirical description toward a data-driven scientific paradigm. This progress is achieved through collaborative innovation in algorithm optimization, framework reconstruction, and process standardization. It lays a solid theoretical and methodological foundation for improving judicial fairness and scientific rigor.

Outlook

This study has obtained preliminary results in optimizing LAMP-DnSK value calculation and applying the FCA framework. Building on these findings, future research will focus on two key directions. The first is robustness validation of statistical models using complex biological samples. The second is the development of cross-laboratory standardization protocols. Scientific uncertainty of forensic molecular evidence frequently stems from technical variations and subjective data interpretation. For this reason, follow-up work should incorporate large-scale multi-center blind test data. These data can evaluate the performance stability of the new paradigm based on simple random error rates. The evaluation covers different degradation levels and template amounts.

This is not only key to enhancing methodological credibility but also a core path to responding to the court's layered review requirements for the reliability of scientific evidence. Bayesian inference is gaining wider adoption in forensic science. This trend raises an important question. How can the dual probability measures defined in the FCA framework be embedded more naturally into overall case probability assessment? This embedding must prevent juries from misinterpreting statistical conclusions. This issue forms a critical link between laboratory techniques and judicial practice.

At the application level, the standardization of the semi-coding hair evidence testing process urgently needs to evolve towards automation and digitalization. Currently, the blind testing phase that relies on manual operations

still poses potential risks of human bias.

In the future, an integrated data analysis platform can be developed in conjunction with bioinformatics tools to achieve seamless connection from raw signal extraction to final evidence strength reporting. This technological integration improves testing efficiency. It also strengthens quality control standards via algorithms. The approach reduces result fluctuations originating from individual operational differences. Accordingly, identification opinions become more persuasive in adversarial litigation procedures. In addition, researchers need to further explore strategies for constructing reference databases when handling historical legacy samples or biologically sourced materials. This need is particularly prominent without a complete reference system. It remains an urgent challenge to quantify the influence of background noise on matching probabilities. The improvement of court admissibility relies not only on the perfection of the technology itself but also requires deep dialogue between the scientific community and the legal field. Future research seeks to develop clear guidelines for uncertainty expression.

Such guidelines can draw on cutting-edge findings on scientific communication within criminology [35]. They help ensure that expert testimony accurately communicates the limitations and applicable conditions of the evidence. A comprehensive evaluation system can be built. This system covers technical validation, case inference, and legal interpretation. Its construction helps advance forensic molecular evidence from a single technical indicator toward a comprehensive decision-support tool.

This process calls for continuous tracking of evolving judicial precedents. Statistical model parameter settings and presentation formats should be adjusted in a timely manner. These revisions help adapt to dynamically shifting admissibility standards for scientific evidence across different legal jurisdictions. This goal ultimately realizes the organic integration of scientific rigor and judicial fairness.

Conclusion

This study systematically investigates core issues concerning the new generation of forensic molecular evidence. These issues cover scientific uncertainty quantification, statistical framework construction, and courtroom admissibility. It seeks to narrow the gap between state-of-the-art laboratory techniques and the practical demands of judicial work. This paper conducts an in-depth analysis of the LAMP-DnSK value calculation method under multiple copy numbers. It identifies theoretical limitations inherent to the traditional Monika-Welzel method for complex mixed samples. Of particular concern are biases arising when identifying low-proportion contributors.

To address this defect, the study proposes and validates a new calculation method based on simple random error rates. This method greatly improves the robustness of numerical calculations. It also lowers the risk of false positives by imposing stricter probability constraints. It offers a more reliable quantitative benchmark for trace DNA analysis. In the testing process of semi-coding hair evidence, empirical analysis using the Kline-Gray archaeological case further confirms the effectiveness of the proposed statistical framework.

This case indicates that combining the two probability measures within the FCA framework can more accurately assess the degree of support that evidence provides for specific hypotheses, thereby overcoming misleading conclusions that may arise from a single indicator. This integration strategy emphasizes the necessity of transitioning from “qualitative matching” to “quantitative probability interpretation” in forensic science, ensuring the transparency and repeatability of the evidence evaluation process. A comprehensive statistical model suitable for the new generation of molecular evidence has been constructed, balancing the complexity of biology with the rigor of legal logic. It not only addresses the current calculation challenges of LAMP-DnSK values in multiple copy scenarios but also provides standardized operational guidelines for the courtroom presentation of new types of evidence such as semi-coding hair. Technology will continue to evolve, and big data algorithms will be increasingly adopted in the future.

Accordingly, this statistical framework requires further expansion to support genome-wide analysis. Researchers may also explore the potential of artificial intelligence for automated error correction. Interdisciplinary collaboration should be strengthened. Discussions among statisticians, forensic experts, and legal professionals should be encouraged. These efforts are critical to establishing the authoritative status of the new generation of molecular evidence in the judicial system. Ultimately,

enhancing public awareness of the scientific uncertainty of forensic molecular evidence is key to ensuring its long-term admissibility and credibility.

Funding

This work was supported by the UNSW Faculty of Law PhD Research Project Grant (Grant No. UNSW-LAW-PHD-2025-0817); the UNSW Law Master Research Seed Grant (Grant No. UNSW-LAW-MAS-2025-1249).

Acknowledgements

The authors would like to show sincere thanks to those technicians who have contributed to this research.

Conflicts of Interest

The authors declare no conflict of interest.

References

[1] Stern, H. S., Cuellar, M., Kaye, D. (2019) Reliability and validity of forensic science evidence. *Significance*, 16(2), 21-24.

[2] Valerio, R. (2020) Likelihood ratios for lawyers… I didn't go to law school for this! *Wiley Interdisciplinary Reviews: Forensic Science*, 2(3), e1366.

[3] Jovanovic, J. (2021) Statistical interpretation of DNA evidence in the judicial practice of Montenegro. *Criminal Justice,* 21.

[4] Edmond, G., Roach, K. (2011) A contextual approach to the admissibility of the state's forensic science and medical evidence. *University of Toronto Law Journal*, 61(3), 343-409.

[5] Koch, S. L., Liebowitz, C., Shriver, M. D., Jablonski, N. G. (2021) Microscopical discrimination of human head hairs sharing a mitochondrial haplogroup. *Journal of Forensic Sciences*, 66(1), 56-71.

[6] Sun, Z., Zhang, K. (2024) Development of biomedical technology from the perspective of ethics and law. *International Internal Medicine Journal*, 2, 01-11.

[7] Linch, C. A., Whiting, D. A., Holland, M. M. (2001) Human hair histogenesis for the mitochondrial DNA forensic scientist. *Journal of Forensic Sciences*, 46(4), 844-853.

[8] Davies, C., Gautam, L., Grela, A., Morrissey, J. (2020) Variability associated with interpreting drugs within forensic hair analysis: a three-stage interpretation. *Journal of Applied Toxicology*, 40(7), 868-888.

[9] Hoffmann, K. (1991) Statistical evaluation of the evidential value of human hairs possibly coming from multiple sources. *Journal of Forensic Sciences*, 36(4), 1053-1058.

[10] Gaudette, B. D. (1982) A supplementary discussion of probabilities and human hair comparisons. *Journal of Forensic Sciences*, 27(2), 279-289.

[11] Abdel-Rahman, H. K. (2020) Effect of adding magnesium oxide nanoparticles on the antimicrobial activity of a denture soft liner. *Polytechnic Journal*, 10(2), 20.

[12] Han, P. K. (2013) Conceptual, methodological, and ethical problems in communicating uncertainty in clinical evidence. *Medical Care Research and Review*, 70(1), 14S-36S.

[13] Giacomantonio, C., Litmanovitz, Y., Bennell, C., Jones, D. J. (2024) Expressing uncertainty in criminology: Applying insights from scientific communication to evidence-based policing. *Criminology and Criminal Justice*, 24(2), 470-488.

[14] Greenwell, M. R. (2017) Communicating certainty and uncertainty in medical, supportive and scientific contexts. *Health Communication*, 32(2), 259-260.

[15] Drapkin, L., Gritsaenko, P., Shuklin, A. (2019) Evidentiary and search capabilities of a forensic medical molecular genetic examination. *Electronic Supplement to Russian Juridical Journal.* 24.

[16] Tarran, B, (2019) The limits of forensic evidence. *Significance*, 16(2), 12-15.

[17] Zvárová, J. (2009) Statistics and biomedical informatics in forensic sciences. *Environmetrics: The Official Journal of the International Environmetrics Society*, 20(6), 743-750.

[18] Sonowal, G., Rughani, P., Gogia, G. (2026) Forensic data deduplication: a comprehensive survey of methods and challenges for digital investigation. *Journal of Forensic Sciences*. 71, 2503-2527.

[19] Xu, X., Vinci, G. (2024) Forensic science and how statistics can help it: evidence, likelihood ratios, and graphical models. *Wiley Interdisciplinary Reviews: Computational Statistics*, 16(5), e70006.

[20] Rakoff, J. S., Liu, G. (2023) Forensic science: a judicial perspective. *Proceedings of the National*

Academy of Sciences, 120(41), e2301838120.

[21] Narejo, N., Avais, M. A. (2012) Examining the role of forensic science for the investigative-solution of crimes. *Sindh University Research Journal (Science Series).* 44(2), 251-254.

[22] Norman, C., Lim, D. X. Y., Henderson, T., Casali, F., Daéid, N. N., Nisbet, L., Ménard, H. (2024) Trends and challenges in the use of drugs as forensic evidence: a scientometric approach to map the current literature. *Wiley Interdisciplinary Reviews: Forensic Science*, 6(2), e1509.

[23] Curran, J.M. (2009) Statistics in forensic science. *Wiley Interdisciplinary Reviews Computational Statistics*, 1, 141-156.

[24] Stern, H. S. (2017) Statistical issues in forensic science. *Annual Review of Statistics and Its Application*, 4, 225-244.

[25] Drikakis, D., Asproulis, N. (2011) Quantification of computational uncertainty for molecular and continuum methods in thermo-fluid sciences. *Applied Mechanics Reviews*, 64, 040801.

[26] Saha, P., Singh, T., Dargush, G. (2021) Uncertainty quantification of differential algebraic equations using polynomial chaos. *Journal of Computational and Nonlinear Dynamics*, 16(10), 101005.

[27] Layne, T., Jackson, K., Scott, A., Tanner, N. A., Piland, A., Haverstick, D. M., Landers, J. P. (2021) Optimization of novel loop-mediated isothermal amplification with colorimetric image analysis for forensic body fluid identification. *Journal of Forensic Sciences*, 66(3), 1033-1041.

[28] Kubo, S., Niimi, H., Kitajima, I. (2023) Duplex loop-mediated isothermal amplification assay for simultaneous detection of human and human male DNA. *BMC Research Notes*, 16(1), 180.

[29] Kumsiri, R., Kanchanaphum, P. (2021) Comparison of time course detection of human male DNA from blood stains on various objects on surface in a natural environment and in a laboratory using loop-mediated isothermal amplification (LAMP). *Scientifica*, 2021(1), 4811608.

[30] Wang, C., Hong, B., Li, Y., Ma, Y., Xu, W., Wang, J. (2025) Rational design of a novel DNA polymerase from clostridium thermocellum to improve LAMP detection performance. *Biotechnology Journal*, 20(1), e202400559.

[31] Harshini, S. S., Veeraraghavan, V. P., Arthanari, A., Gayathri, R., Kavitha, S., Selvaraj, J., Dinesh, Y. (2022) Comparative study of male and female human hair: a microscopic analysis. *Journal of Advanced Pharmaceutical Technology and Research*, 13(Suppl 1), S297.

[32] Wickenheiser, R. A., Hepworth, D. G. (1990) Further evaluation of probabilities in human scalp hair comparisons. *Journal of Forensic Sciences*, 35(6), 1323-1329.

[33] Brunelle, R. L. (1973) Novel techniques for the examination of physical evidence - their admissibility as evidence. *Journal of the Association of Official Analytical Chemists*, 56(6), 1391-1400.

[34] Cole, S. A. (2021) Autopsy of a crime lab: Exposing the flaws in forensics. *Law and Society Review*, 55(4), 683-684.

[35] Balding, D. J., Donnelly, P. (1995) Inferring identify from DNA profile evidence. *Proceedings of the National Academy of Sciences*, 92(25), 11741-11745.