Printed in the Republic of Korea
ANALYTICAL SCIENCE & TECHNOLOGY
Vol. 39 No. 3, 176-197, 2026
https://doi.org/10.5806/AST.2026.39.3.176

Review

Trace evidence: A critical review of forensic analysis, statistical evaluation, and emerging technologies

Songhie An★, Kyoung-Min Kim, and Wooyong Park

Forensic Chemistry Division, National Forensic Service, 10 Ipchun-ro, Wonju, Gangwon-do 26460, Republic of Korea

(Received March 30, 2026; Revised May 18, 2026; Accepted May 26, 2026)

Abstract: Trace evidence encompasses the physical remnants that arise from contact between persons, objects, and locations, which may be recovered to support the reconstruction of criminal events. This review aims to provide an integrated and critical overview of six major categories of trace evidence: fiber, hair, paint, soil, glass, and tape. For each category, forensic significance (including transfer and persistence considerations), collection strategies, tiered examination methodologies, and emerging artificial intelligence (AI)-assisted applications are addressed within a unified framework. This review covers the institutional infrastructure that governs examination practices, including the standards and guidelines produced by the Scientific Working Group for Materials Analysis, Organization of Scientific Area Committees, European Network of Forensic Science Institutes, INTERPOL, Asia Forensic Science Network, and American Society for Testing and Materials. We address data interpretation, with particular attention paid to the likelihood ratio (LR) framework and the conditions required for its defensible implementation, alongside emerging approaches, including Bayesian networks, score-based LR, and conformal prediction, and maintain a critical perspective. Advances in analytical capabilities and AI do not automatically translate into stronger forensic evidence, and the adoption of any method in operational casework must be preceded by rigorous validation, interlaboratory studies, and judicial scrutiny. The tiered analytical workflow and comparative method performance assessment presented here provide a structured basis for evidence-based instrument selection and responsible LR implementation in forensic casework.

Keywords: trace evidence, forensic chemistry, likelihood ratio, artificial intelligence

1. Introduction

Trace evidence refers to physical remnants, such as materials, marks, or residues, that arise from contact between persons, objects, and locations. Edmond Locard articulated the foundational principle underlying forensic application, that is, the exchange principle which states that any contact event inevitably produces a bidirectional transfer of material. This review addresses six categories of trace evidence: fibers, hair, paint, soil, glass, and tape. Each trace type is distinguished by characteristic transfer mechanisms,

★ Corresponding author
Phone : +82-(0)33-902-5431 Fax : +82-(0)33-902-5933
E-mail : stellaluna@korea.kr

persistence behavior, and analytical requirements; together, they constitute the core of trace evidence examination practices in forensic laboratories worldwide.[1]

Forensic examinations of trace evidence have undergone substantial methodological developments

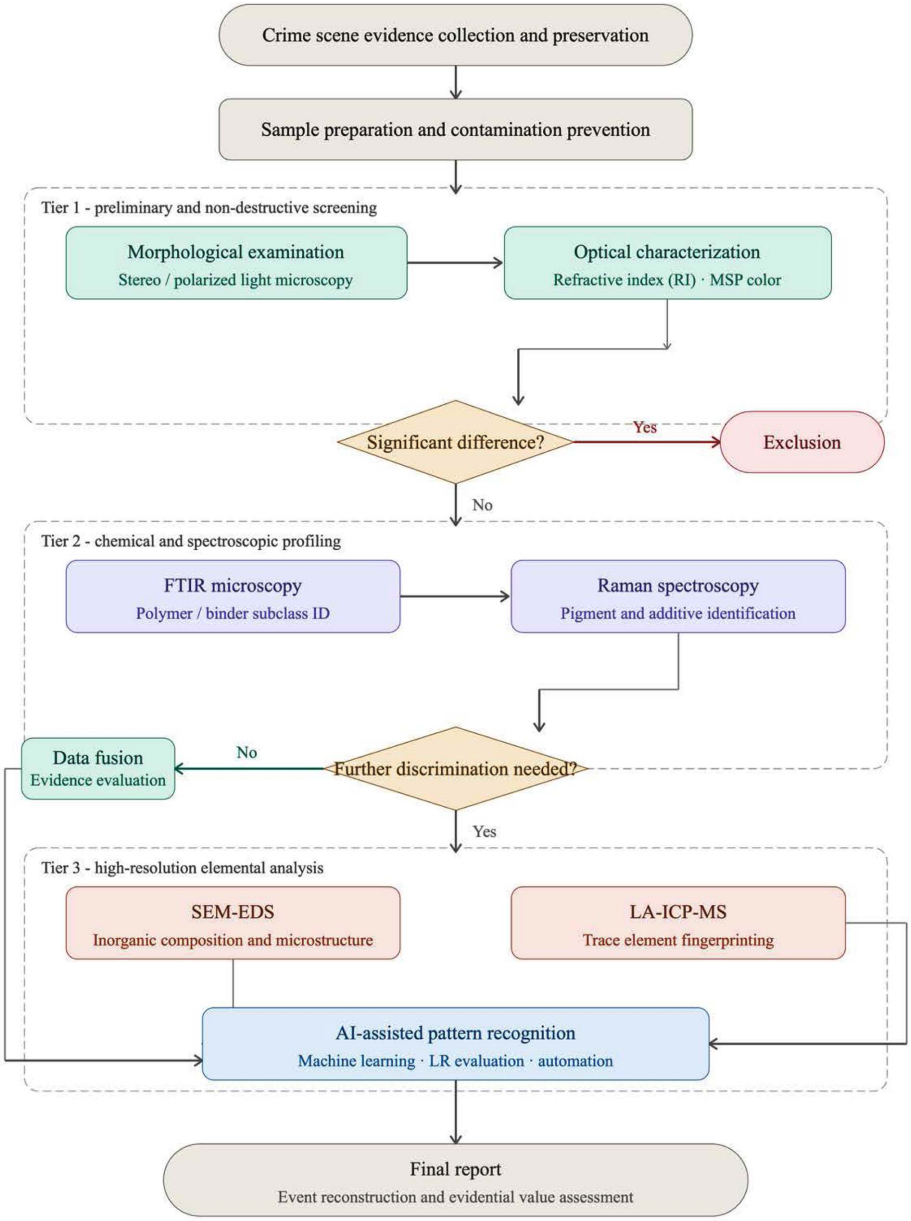

Fig. 1. Tiered analytical workflow for forensic trace evidence examination. The flowchart illustrates the sequential progression from crime scene collection through three analytical tiers: preliminary nondestructive screening (Tier 1), chemical and spectroscopic profiling (Tier 2), and high-resolution elemental analysis (Tier 3), followed by artificial intelligence (AI)-assisted data processing and final reporting. Decision points at each tier reflect the principle of proportionate resource use; escalation proceeds only when the preceding tier does not provide sufficient discrimination. FTIR=Fourier-transform infrared microscopy; SEM-EDS=scanning electron microscopy with energy-dispersive X-ray spectroscopy; LA-ICP-MS=laser ablation inductively coupled plasma mass spectrometry.

in recent decades. Instrumentation has progressed from optical microscopy and microspectrophotometry through Fourier transform infrared (FTIR) and Raman spectroscopy to high-resolution elemental platforms such as laser ablation inductively coupled plasma mass spectrometry (LA-ICP-MS) and scanning electron microscopy–energy dispersive X-ray spectrometry (SEM-EDS), with each advance bringing increased discrimination capability at the cost of greater data complexity. This complexity has driven a parallel development in interpretation; visual comparison has given way to multivariate statistical evaluation, probabilistic frameworks such as the likelihood ratio (LR), and, most recently, artificial intelligence (AI)-assisted pattern recognition. Simultaneously, a robust infrastructure of standard-setting organizations (Scientific Working Group for Materials Analysis (SWGMAT), American Society of Trace Evidence Examiners (ASTEE), Organization of Scientific Area Committees (OSAC), European Network of Forensic Science Institutes (ENFSI), INTERPOL, Asia Forensic Science Network (AFSN), and American Society for Testing and Materials (ASTM)) has been established to govern practice, promote interlaboratory harmonization, and ensure that new methods are validated before operational deployment.[2,3]

This review provides a comprehensive and critical overview of six major trace evidence categories. Section 2 outlines the tiered examination workflow; Section 3 covers each trace type (forensic significance, collection, examination methods, and emerging AI applications); Section 4 addresses data analysis, the LR framework, Bayesian networks, and the Korean implementation path; and Section 5 surveys the governing professional organizations. The tiered workflow, comparative method performance, and LR-readiness criteria are summarized in *Fig.* 1, *Table* 1, and *Fig.* 2.

2. Tiered Analytical Workflow

The examination of trace evidence is most efficiently structured as a tiered sequence in which nondestructive and rapid methods are first applied, escalating to more resource-intensive approaches only when the preceding tier does not provide sufficient discrimination. The principle of proportionate resource use is formalized in the workflow shown in *Fig*. 1.

Tier 1 (preliminary and nondestructive screening) encompasses morphological examination using stereo- and polarized light microscopy, color assessment via UV-Vis microspectrophotometry (MSP), and physical screening, including refractive index measurements of glass. In many cases, these methods are rapid, preserve sample integrity, and provide class-level information that is sufficient to support exclusion. Tier 2 (chemical and spectroscopic profiling) applies FTIR microscopy and Raman spectroscopy to

Table 1. Criteria used for comparative assessment of analytical methods in *Fig.* 2

Criterion	Definition	Score 1	Score 3	Score 5
Discriminative power	Ability to distinguish samples from different sources based on measured features	Class-level only	Subclass level	Individual-source level
Nondestructiveness	Degree to which sample integrity and mass are preserved during analysis	Sample consumed	Partial damage	No damage
Speed	Relative throughput and turnaround time in routine forensic casework	1 d/sample	h	30 min
Accessibility	Availability of instrument and expertise in standard forensic laboratories	Specialist labs only	Regional labs	Widely available
Multimaterial applicability	Range of trace types (fibers, hair, paint, soil, glass, tape) addressable using the method	1–2 types	3–4 types	5–6 types
Likelihood ratio-readiness	Operational maturity of LR-based evaluation, including validated background databases	No framework	Research stage	Validated and casework-ready

Scores (1–5) represent the assessment of the authors based on the reviewed literature and casework experience; 1=lowest, 5=highest

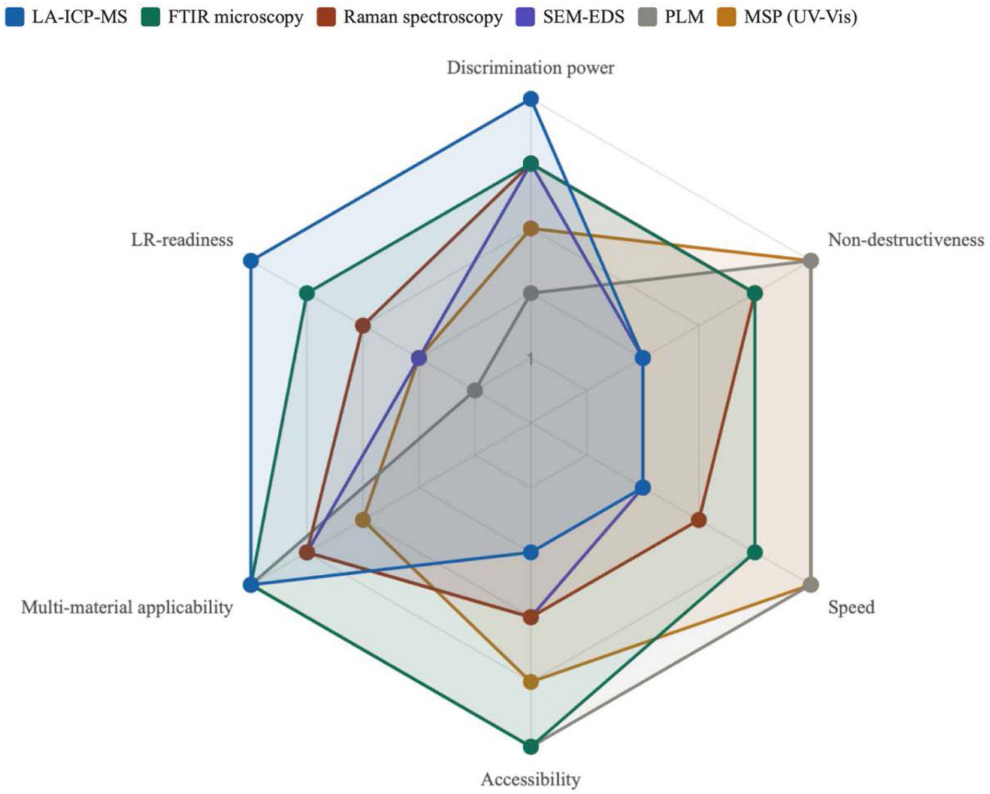

Fig. 2. Comparative radar plot of six analytical methods across six performance criteria. Scores (1–5) represent the authors' assessment based on the reviewed literature and casework experience. Criteria: Discrimination power (ability to distinguish samples from different sources); Non-destructiveness (5 = fully non-destructive); Speed (throughput in routine casework); Accessibility (5 = widely available in routine labs), Multi-material applicability (range of trace types addressable); LR-readiness (5 = validated and casework-ready). LA-ICP-MS=laser ablation inductively coupled plasma mass spectrometry; FTIR=Fourier-transform infrared microscopy; SEM-EDS=scanning electron microscopy with energy-dispersive X-ray spectroscopy; PLM= polarized light microscopy; MSP=UV-Vis microspectrophotometry.

characterize the polymer, binder, and pigment/additive composition. Tier 3 (high-resolution elemental analysis) uses SEM-EDS and LA-ICP-MS when the preceding tiers fail to resolve the comparison or when the case significance warrants a quantitative evaluation. AI-assisted data processing and LR-based evaluation are applied to the outputs of tiers 2 and 3 to produce defensible probabilistic statements of evidential strength.[3-6]

This tiered structure applies to all six trace types reviewed herein, although the specific methods within each tier vary by material class. The comparative performances of the principal analytical methods across six criteria (discrimination power, nondestructiveness, speed, accessibility, multimaterial applicability, and LR-readiness) are summarized in *Table* 1 and *Fig.* 2.

3. Trace Evidence: Significance, Collection, and Examination

Recovery of trace evidence must follow a strict sequence: visual documentation and photography first, then nondestructive collection, then methods that alter the surface or substrate. The spatial distribution of trace material, its orientation, and its relationship to other evidence can be as informative as the material itself, and this information is irreversibly lost once collection begins. Whole-item submission is strongly preferred across all trace types wherever

logistics permit, as it preserves spatial context, prevents cross-transfer, and allows the laboratory examiner to make informed recovery decisions; field collection tools should be used only when whole-item submission is genuinely impractical.

Forceps are used for discrete solid items, including paint chips, tape fragments, large glass shards, hair bundles, and fiber tufts, with clean tips contacting the least analytically significant part of the item. Tape lifting is the primary recovery method for fibrous trace materials, single fibers and hairs on fabric or hard surfaces, and fine glass fragments on clothing too small to collect individually; lifts should be submitted intact on their backing, labeled with the exact recovery location, and submitted in a rigid container. Tape lifting is not suitable for loose particulate materials, including soil, dust, sediment, and fine mineral residues, as recovering such material from an adhesive matrix without mechanical disruption or chemical interference is practically infeasible.

Packaging follows material-specific requirements: paper is preferred over plastic for most trace evidence, since plastic generates static charge that redistributes fibrous material and condensation promotes biological degradation; items are packaged individually to prevent cross-transfer; fragile items such as glass and paint chips are placed in rigid containers; and adhesive surfaces are protected with clean inert interleaves. Chain of custody documentation begins at the point of collection, not at laboratory receipt.

For each trace type, the following subsections address forensic significance, collection strategy, and the tiered examination sequence. Examination proceeds from nondestructive screening (Tier 1) through chemical and spectroscopic profiling (Tier 2) to high-resolution elemental analysis (Tier 3), with each tier applied in sequence and more resource-intensive methods reserved for cases unresolved by the preceding tier. General collection principles apply across all trace types and are not repeated in full for each category. The deployment of AI-assisted examination tools is conditional on the validation requirements described in Section 4.4, and the trace-type-specific state of that infrastructure is assessed at the end of each examination subsection.

3.1. Fibers

3.1.1. Forensic significance

Fibers are among the most frequently encountered trace materials in violent crimes involving physical contact. Natural fibers (e.g., cotton and wool) and synthetic polymers (e.g., polyester, polyamide, and acrylic) are commonly recovered. Their forensic value arises from the potential for transfer during contact between garments or furnishings and the discriminating power of combined morphological, color, and chemical examinations. The evidential significance of a fiber association depends on the rarity of the fiber type within the relevant population, number recovered, and plausibility of alternative transfer pathways. Population studies of man-made fibers have demonstrated that when polymer class, color, and chemical subclass are considered jointly, the probability of coincidental correspondence between unrelated garments is extremely low, providing a scientific basis for evidential significance even where source-level individualization is not achievable.[7,8] Transfer studies have demonstrated that fiber exchange can occur from a single brief contact and that synthetic fibers may persist on recipient garments for hours to days, depending on substrate type and physical activity, with persistence decreasing markedly after laundering.[3,4]

3.1.2. Collection

Fibers are recovered from garments and surfaces primarily, but not exclusively, using tape lifting. Depending on the context of the scene and, positioning of a particular fiber etc, other collection tools such as forceps etc should be considered. Items should be packaged individually on paper to prevent cross-transfer, and the recovery location should be documented in sufficient detail to anchor laboratory results to the recovery context. Wherever feasible, all items should be submitted intact rather than subjected to in situ tape-lifting, to ensure preservation of spatial distribution information.[3,4]

3.1.3. Examination

PLM characterizes fiber morphology, optical properties, and refractive behavior, enabling identification of the nine generic synthetic fiber classes through assessment of birefringence, sign of elongation, surface texture, cross-sectional shape, and refractive index relative to the mounting medium.[8-12] Transverse cross-sections provide additional information on fiber shape (round, trilobal, ribbon) and internal structure, applied when longitudinal PLM is insufficient and case significance warrants the destructive step.[13,14] PLM is nondestructive in its primary configuration; its limitation is class-level rather than source-level discrimination.

UV-Vis microspectrophotometry (MSP) provides objective color measurement by recording transmittance or absorbance spectra, occupying a central position in the examination workflow because color is the most immediately discriminating feature for many fiber types and the only reliable discriminator for some natural fibers.[15-17] Metamerism is a recognized limitation: chemically distinct dyes may produce indistinguishable spectral profiles.[18] First-derivative preprocessing and chemometric classification strengthen discrimination beyond visual assessment.[16,19]

FTIR microscopy and Raman spectroscopy provide complementary chemical characterization: ATR-FTIR identifies polymer class and resolves subclass distinctions from comonomer ratios and additive packages nondestructively, while Raman interrogates organic dyes directly at concentrations below the MSP detection threshold, with confirmed interlaboratory reproducibility.[20-28] Principal limitations are FTIR insensitivity to dye-only differences and Raman fluorescence interference.[23,27,28]

Targeted dye analysis using **LC-MS** provides the highest discrimination tier, resolving individual dye components by molecular mass and differentiating dyes with indistinguishable UV-Vis profiles.[29-32] HPLC-MS/MS extends applicability to limited-quantity samples; the method is destructive and reserved for high-significance cases where nondestructive tiers have established a meaningful preliminary association.[30,31,33,34]

Taken together, each tier provides class-level or subclass-level discrimination that approaches but cannot reach individualization. At each transition between methods, the unresolved similarity becomes the object of probabilistic evaluation rather than further analytical discrimination. For fibers, LR-based evaluation has been demonstrated in research settings using MSP spectral data and FTIR chemometric outputs, and population frequency data for common fiber types published by the ENFSI European Fibres Group provide a partial foundation for background population modeling.[7,8] However, as Vooijs et al. concluded, numerical evidential values for fiber examinations are not yet feasible given the limited population data available for most fiber types,[35] and jurisdiction-specific reference databases reflecting local garment manufacturing are absent for most of Asia and absent entirely for Korea.

3.1.4. Emerging AI applications

Machine and deep learning approaches offer several promising avenues for automating and accelerating fiber examinations. Convolutional neural networks (CNNs) trained on polarized light microscopy images have the potential to classify fiber type and cross-sectional morphology without manual feature extraction, which could reduce subjectivity in the initial screening step. FTIR spectra treated as one-dimensional signals or transformed into two-dimensional image representations are well-suited for CNN-based classification, enabling automated polymer subclass identification from spectral libraries at speeds far exceeding manual comparison. Multimodal fusion architectures that jointly process morphological images and spectral data represent a promising research direction for forensic fibers, where color, morphology, and polymer class must be considered together, but have not yet been validated on forensic-representative fiber datasets.[36]

3.2. Hair

3.2.1. Forensic significance

Hair is a keratin-based biological fiber encountered in many forensic cases. Two complementary analytical approaches are employed in forensic hair examination.

Traditional morphological hair comparison relies on microscopic examination of structural features, such as scale pattern, cortical texture, medullary index, and pigment distribution. This information is used to assess the consistency between a questioned hair and a known source and to determine its somatic origin. However, morphological comparisons have been subjected to sustained scientific scrutiny. A systematic review by the US Federal Bureau of Investigation identified erroneous testimony in a substantial proportion of examined cases, and interlaboratory proficiency data confirmed that the reproducibility of microscopic feature assessments remains limited.[37-39] Therefore, morphological findings should be interpreted conservatively, framed as class-level consistency rather than source attribution, and followed by mitochondrial or nuclear DNA analyses wherever possible. A complementary dimension involves the chemical characterization of exogenous residues on hair shafts (e.g., cosmetic products, environmental contaminants, and occupational deposits), which can support activity-level propositions regarding the exposure history or environmental associations of an individual.[40,41]

3.2.2. Collection

Hairs are collected using clean forceps or tape, packaged individually in paper folds, and the recovery location is documented. Minimizing friction and mechanical contact during collection is important for surface residue analysis because thin exogenous deposits are readily dislodged.[40]

3.2.3. Examination

Macroscopic and comparison microscopy constitute the primary tier of forensic hair examination.[37,39,42] Macroscopic screening characterizes gross features (color, length, diameter, curl, body region) for initial inclusion or exclusion.[37] Comparison microscopy assesses morphological features across the cuticle, cortex, and medulla using a two-headed microscope; a reference sample of 25 or more full-length hairs is required to characterize intra-source variation.[37,41] Microscopic association establishes class-level consistency only; population studies confirm that fewer than one pair in 4,500 brown Caucasian scalp hairs is indistinguishable.[38,39,43,44]

mtDNA sequencing complements microscopic comparison for rootless hairs, resolving associations that microscopy cannot and excluding contributors whose microscopic profile is superficially consistent; a systematic FBI evaluation confirmed the methods are complementary rather than interchangeable.[40,45] mtDNA establishes lineage-level rather than individual-level consistency, as it is maternally inherited and shared among maternal relatives.[40]

Chemical examination addresses exogenous residues (hair dyes, bleaching agents, styling treatments, contaminants) providing treatment and exposure history information that morphological comparison cannot resolve.[37,41] ATR-FTIR spectroscopy nondestructively classifies dyed and treated hairs by surface polymer and dye functional group composition, reliably discriminating between treated and untreated hairs and between dye types at the class level.[46] SERS enables direct in situ detection of artificial colorants without extraction, achieving identification accuracy exceeding 95% across commercial dye types and brands.[47]

LA-ICP-MS provides a segmental elemental profile reflecting endogenous metabolic status and exogenous environmental exposures, with the distribution of trace elements (Zn, Cu, Fe, As, Pb, Se, Hg) along the shaft serving as a temporal exposure record supporting forensic comparison and toxicological interpretation.[48] Substantial within-individual variability limits source-level comparison on elemental grounds alone without a rigorous statistical framework and representative reference data.[41,48]

The specific challenge of hair evidence is that the highest tier of molecular analysis (mtDNA) provides lineage-level but not individual-level discrimination, and the lower tiers (comparison microscopy, chemical characterization) provide class-level consistency assessments whose subjectivity has been empirically documented. No LR framework for hair comparison microscopy has been validated for operational use. For mtDNA, LR-based probabilistic reporting has been described in research contexts using population

frequency databases for mitochondrial haplotypes. A Korean mtDNA control region database has been published (n = 704 individuals),[49] but whole mitochondrial genome (mtGenome) data for Koreans at the depth required for next-generation LR calculations remain limited, and the database does not yet support the population frequency estimates needed for defensible casework LR reporting at the individual haplotype level.

3.2.4. Emerging AI applications

Machine learning approaches offer the potential to reduce interexaminer variability, thereby addressing the recognized limitation of subjectivity in morphological hair comparison. Unsupervised clustering of microscopic hair shaft images using color models (RGB, CIE L*a*b*) and deep feature extraction has demonstrated discriminating power comparable to that of expert visual assessment for color-based comparisons, with the advantage of consistent and documentable results. Such tools are currently in the research stage and require extensive validation on casework-representative samples before operational deployment. However, they signal a plausible path toward a more objective and reproducible morphological assessment.[50,51]

3.3. Paint

3.3.1. Forensic significance

Forensic paint evidence is mostly found in vehicle collision investigations and burglary and assault cases. The discriminating power of paint is derived from its hierarchical structure, which includes layer sequence, binder chemistry, pigment and extender packages, and additives, collectively providing a multidimensional chemical profile. Automotive paints are useful because they consist of multiple functional layers, including electrodeposited primers, surfaces, basecoats, and clearcoats, each with a characteristic composition. Architectural and industrial paints, which are less complex, can provide meaningful associations when examined using orthogonal analysis. A key distinction exists between factory-applied (OEM) coatings and aftermarket repaints, which may differ in their layer architecture and composition. Paint transfer in vehicle collisions can involve primary deposition and secondary redistribution; the directionality and layering of transferred smears can inform the reconstruction of the impact geometry.[52,53]

3.3.2. Collection

Paint chips and smears are collected using a clean scalpel or forceps and packaged in small rigid containers to prevent crushing or loss. For vehicle collision cases, wherever feasible, the entire painted surface, panel or vehicle itself should be submitted. The recovery location and direction of transfer should be documented to support later interpretation of the layer sequence.[52]

3.3.3. Examination

Stereo microscopy and PLM document layer count, sequence, color, and thickness nondestructively, providing the primary basis for initial association or exclusion: samples differing in layer number or color sequence are excluded without further analysis.[42,52] Layer count and the OEM versus aftermarket repaint distinction carry direct evidential significance for vehicle collision reconstruction.[52,53] Layer architecture alone cannot distinguish chemically distinct samples sharing the same physical profile, making chemical analysis essential for any positive association.[54]

FTIR microscopy and Raman spectroscopy are complementary: ATR-FTIR classifies binder and resin systems by layer, resolving major classes (alkyd, polyurethane, acrylic, melamine, polyester) and subclass distinctions from comonomer ratios, with combined FTIR imaging enabling spatially resolved profiling across a full chip cross-section.[57-60] Raman identifies organic and inorganic pigments in situ without extraction, targeting the pigment phase FTIR undercharacterizes; the combination substantially exceeds the discriminating power of either alone.[62-65]

SEM-EDS provides elemental profiles of pigment and extender components by layer, particularly valuable for metallic and inorganic pigments that produce limited infrared signal but strong X-ray emission; its recognized limitation is poor sensitivity to trace and

minor elements, motivating escalation to LA-ICP-MS.[54,58,61]

LA-ICP-MS and XRF provide trace and minor element discrimination beyond SEM-EDS capability, with LA-ICP-MS time-resolved analysis quantifying up to 30 or more elements per layer simultaneously, yielding high-dimensional output suited for multivariate statistical treatment and LR-based probabilistic reporting when used alongside reference databases.[52-54,62,63]

Pyrolysis GC-MS characterizes organic binders when FTIR cannot resolve chemically similar formulations, thermally fragmenting polymer matrices into characteristic pyrograms distinguishing binder classes by minor component differences; the method is destructive and applied selectively at the highest examination tier when spectroscopic methods fail.[64,65]

These methods collectively provide the highest multidimensional discriminating power of any trace type in this review, by virtue of the hierarchical layer structure that multiplies the independent chemical dimensions available for comparison. Despite this discriminating power, LR-based evaluation of paint evidence remains operationally immature. LR approaches applied to Raman and FTIR spectra of automotive paints have been demonstrated at the research level,[66,67] but variability introduced by weathering, aftermarket repair, and manufacturing batch variation is not yet captured in reference databases of sufficient geographic and temporal scope, and no interlaboratory validation of paint LR pipelines comparable to those conducted for glass has been published.

3.3.4. Emerging AI applications

The multilayered and multimodal nature of forensic paint evidence makes it well-suited for AI-assisted fusion analysis. Machine learning classifiers (random forest and support vector machine) applied to combined FTIR and SEM–EDS datasets have demonstrated improved subclass discrimination compared to single-technique approaches. CNNs applied to FTIR spectral images have the potential to automate binder classification across multiple layers, and transfer learning from pretrained spectral models may reduce training data requirements for new paint systems, though neither has been demonstrated specifically for forensic automotive paint. For automotive paint databases, AI-assisted matching against manufacturer-specific layer sequences and elemental profiles can triage potential source vehicles more efficiently than manual database lookup, provided that the underlying database is sufficiently representative and the matching algorithm is validated against known source samples.[68]

3.4. Glass

3.4.1. Forensic significance

Glass fragments are transferred when glazed surfaces break, making glass evidence relevant in forced entry, vehicle collisions, and assault cases. The primary analytical strategy exploits variations in minor and trace element concentrations introduced during raw material selection and manufacturing. The refractive index (RI) serves as a primary screening tool; differences support exclusion, but similarity alone is insufficient for association because RI values can be shared across mass-produced panels. High-resolution elemental profiling, particularly LA–ICP–MS, substantially increases discrimination and supports database-assisted and LR-based evaluations. Fragment persistence on clothing is size dependent and decreases rapidly over time, and controlled transfer and persistence studies have established that the probability of recovering fragments decreases substantially within the first hour after breakage, providing critical reference data for the interpretation of fragment counts in casework.[69,70]

3.4.2. Collection

Fragment recovery documents the location and distribution, as this information informs interpretation of the transfer and breakage dynamics. wherever possible the distribution of fragments across a surface should be documented photographically before collection, as fragment distribution patterns can inform interpretation of the breakage event and direction of force. Large fragments are collected individually with clean forceps into rigid containers. Fine fragments on clothing are collected by tape lifting, with the lift

submitted intact. The fragments are collected in rigid containers to prevent further fragmentation. Wherever a physical fit between the fragments is possible, the fragments should be kept separate to preserve the fracture edges for comparison.[70]

3.4.3. Examination

Laboratory examination proceeds from physical fit through RI screening to elemental profiling, with each tier reserved for cases unresolved by the preceding step, consistent with the proportionate resource use principle of Section 2.

Physical fit, achieved when a fragment interlocks continuously along a fracture edge, provides the strongest possible source-level association in glass evidence; the complex three-dimensional fracture geometry is unique to a specific breakage event and cannot be replicated in another object.[71] It is the only glass examination method capable of individualization.

Refractive index (RI) measurement using the thermal oil-immersion method, operationalized in the GRIM instrument, is the primary nondestructive screening step.[72,73] Similarity in RI alone is insufficient for association, as RI values in mass-produced float glass have narrowed due to manufacturing standardization and may be shared by unrelated sources; standardized match criteria are codified in ASTM E1967 and the ENFSI guideline.[73-75]

SEM-EDS and μ-XRF exploit compositional variations introduced during raw material selection and float glass manufacturing.[76,77] SEM-EDS detects major and minor elements characterizing glass type; μ-XRF extends elemental characterization nondestructively to fragments SEM-EDS cannot resolve further, with spatial mapping capability.[76,77] Both methods are limited by sensitivity: trace elements below approximately 10–50 μg/g are not reliably detected, precisely those carrying the highest discriminating information, motivating escalation to LA-ICP-MS.[78,79]

LA-ICP-MS simultaneously quantifies 20 or more trace and minor elements at sub-μg/g detection limits, providing a high-dimensional compositional fingerprint substantially exceeding the discrimination power of RI and SEM-EDS/μ-XRF.[79,80] The multi-element profile is standardized under ASTM E2927 and has undergone extensive interlaboratory validation: interlaboratory validation studies involving up to 16 laboratories confirmed LA-ICP-MS outperforms μ-XRF and solution ICP-MS for discrimination of glass from the same manufacturing plant at different time points, with association/exclusion rates comparable to single-laboratory settings.[78,81] LA-ICP-MS outputs are compatible with score-based LR and multivariate modeling approaches described in Section 4; its primary limitation is instrument accessibility, supporting the tiered approach.[69,79-83]

Glass evidence has the most mature LR infrastructure of any trace type reviewed here. LA-ICP-MS elemental profiles standardized under ASTM E2927 provide high-dimensional, reproducible compositional fingerprints; the 16-laboratory and 10-laboratory interlaboratory studies described above demonstrate reproducibility at a level suitable for probabilistic reporting; and operational LR-based glass evidence reporting has been described in the Netherlands Forensic Institute, UK forensic science laboratories, and several Australian state forensic laboratories.[78,80,83] Score-based LR pipelines combining classifier outputs with kernel density estimation models applied to elemental profiles are the most operationally mature AI-statistical integration for any trace type. The primary remaining gap for Korean casework is the absence of Korea-specific transfer and persistence probability estimates, which should be combined with the NFS internal glass elemental reference database for defensible LR values.

3.4.4. Emerging AI applications

Glass evidence has the most mature AI integration of any trace type, driven by the availability of large, well-characterized elemental databases and standardized LA-ICP-MS analytical protocols. Technique-agnostic machine learning models trained on combined datasets from PIXE, LA-ICP-MS, and SEM-EDS have classified glass fragments into source categories with high accuracy even when test samples were measured by a different technique from the training

data, addressing a key interlaboratory comparability challenge.[84] Random forest and Bayesian Additive Regression Trees consistently outperformed conventional distance-based methods when ≥18 elemental variables were available.[85] AI-assisted LR calculation pipelines, in which a trained classifier generates a similarity score that is fed into a score-based LR model, have been proposed as a route to more efficient probabilistic reporting in high-volume casework.[5,6,84–86]

3.5. Soil

3.5.1. Forensic significance

Soil is a heterogeneous mixture of mineral particles, organic matter, biological components, and anthropogenic materials, the composition of which reflects the geological and land-use history of the location. This site specificity is the forensic basis for soil comparison. Samples from the footwear, clothing, or vehicle of a suspect may be compared with soil from the location of interest. The forensic value of soil is enhanced by its ubiquity, and its composition can vary markedly over short distances, making a close association potentially informative even in geologically uniform areas. The transfer of soil to footwear occurs readily through ground contact, and persistence studies indicate that the bulk of the transferred material is shed within the first hour of activity, with residual quantities detectable for considerably longer periods.[87,88]

3.5.2. Collection

Soil and loose particulate material from footwear tread recesses is collected by dry brushing with a clean soft-bristled brush over clean paper or directly into a clean rigid container, working tread recess by tread recess. Compacted material in deep tread recesses may require scraping with a clean scalpel or stainless steel spatula. Reference or control samples from the hypothesized contact location are collected with a clean trowel for small volumes or a spade for larger sampling areas, at a recorded depth, since elemental and mineralogical composition varies with soil depth and comparisons between questioned and control samples are only valid when collected at equivalent depths. Reference samples should be replicated across multiple microlocations within the hypothesized contact area rather than taken as a single grab sample, to characterize spatial variability within the site.[87,88]

3.5.3. Examination

Laboratory examination combines physical, mineralogical, organic, and elemental characterizations in a tiered sequence, proceeding from rapid nondestructive screening to resource-intensive analyses when earlier tiers are insufficient; the combination of multiple independent parameters substantially increases discriminating power.[87,88]

Color assessment, particle size distribution, PLM and X-ray diffraction (XRD) constitute the first nondestructive screening tiers.[87-92] Munsell color chart comparison discriminates approximately 70 % of soil pairs from geographically distinct locations; instrumental UV-Vis spectrophotometry provides a more reproducible alternative for small samples.[89] Particle size distribution characterizes sand, silt, and clay fractions reflecting geological origin and pedogenic history.[90,91] PLM enables particle-by-particle mineral identification and quantification, providing a mineralogical fingerprint of the parent formation and local pedogenic alteration.[91,92] XRD detects bulk mineral phase composition across the full size range; quantitative Rietveld refinement converts diffractogram data into mineral weight percentages, resolving comparisons that PLM alone cannot.[92]

Organic matter analysis and diatom examination contribute discriminating dimensions independent of the mineral fraction.[93-95] Pyrolysis GC-MS and HPLC profiling of non-volatile organic compounds characterize site-specific organic signatures as a function of vegetation and land use.[93,94] Diatom frustule assemblages are highly site-specific for soils from aquatic or semi-aquatic environments and require specialist taxonomic expertise; both methods are destructive and applied as higher-tier analyses when mineral characterization is insufficient.[95]

Elemental profiling using **ICP-MS or XRF** quantifies major oxides, minor elements, and trace rare-earth elements capturing the geochemical signature of the

parent rock and effects of pedogenic weathering.[96,97] ICP-MS simultaneously quantifies 30 or more elements at sub-μg/g sensitivity, generating high-dimensional profiles suited for multivariate statistical analysis and LR-based evaluation; geospatially referenced elemental databases enable regional provenance inference analogous to the glass database infrastructure.[87,88,96,98] XRF provides a rapid nondestructive screening alternative; spatial heterogeneity and depth variation in soil elemental composition require consistent sample preparation for valid comparisons.[87,97]

These examination methods are the most heterogeneous of any trace type: each contribute independent discriminating information that no single method provides alone. Despite this multidimensionality, LR-based evaluation of soil evidence is less developed than for glass or fibers. Published LR calculations for soil comparisons have applied multivariate normal and Bayesian interpolation models to elemental profiles, with a systematic study across 685 georeferenced topsoil samples in Australia demonstrating the approach on blind simulated casework samples; performance depends critically on the representativeness of the reference database relative to the geographic context of the case.[99] For Korea, the Korea Institute of Geoscience and Mineral Resources (KIGAM) national geochemical databases provide a regionally specific elemental reference that is largely absent from the international soil forensics literature and represents an underexploited resource for LR-based soil evidence evaluation.[100]

3.5.4. Emerging AI applications

Forensic soil comparison benefits from large national geochemical survey datasets that exist in many jurisdictions and provide geographically referenced elemental and mineralogical reference data. Machine learning models trained on such datasets, including random forest, gradient boosting, and neural network classifiers, have been applied to geographic provenance tasks, assigning questioned soil samples to regions of origin based on their elemental profiles, with a discrimination performance that exceeds that of conventional threshold-based methods. Deep learning applied to X-ray diffraction patterns enables automated mineralogical phase identification, thereby reducing the time required for quantitative Rietveld analysis. These approaches are particularly promising for cases where a large number of reference samples must be compared against the questioned soil, as the computational cost of AI-based comparison scales far more favorably than manual examinations.[99]

3.6. Tapes

3.6.1. Forensic significance

Adhesive tapes are used in restraints, package sealing, improvised devices and crime scene staging. Forensic examinations exploit physical and chemical characteristics. Physical end matching, where torn or cut ends from two tape pieces are compared, can provide highly individualized associations comparable to physical fracture matching in glass when the fracture morphology is preserved. Chemical examination of the backing polymer and pressure-sensitive adhesive (PSA) using FTIR enables classification into meaningful subgroups, whereas elemental profiling of the backing additives using LA-ICP-MS or laser induced breakdown spectroscopy (LIBS) provides additional discrimination, particularly for electrical and packaging tapes. Unlike fibrous trace materials, adhesive tape does not shed passively. Its transfer is intentional, and its persistence is governed by adhesive properties and substrate characteristics, which influence the inferential weight attributable to its recovery.[101,102]

3.6.2. Collection

Tape evidence is collected using clean forceps gripping the non-adhesive edge only, never across the adhesive surface, to preserve any trace material embedded in the adhesive and to prevent contamination. Where tape is already adhered to an object, whole-item submission is strongly preferred over field removal; if removal is unavoidable, the tape should be peeled slowly at a low angle to minimize stretching and distortion of fracture edges. Torn or cut ends, which carry the most individualized physical fit information, are protected with a clean inert interleave

and packaged separately from other tape pieces to prevent inadvertent contact or bonding between fracture surfaces. The adhesive face is similarly protected with a clean polyethylene or silicone-release paper interleave during transport to prevent pickup of extraneous material.[101]

3.6.2. Examination

Physical characteristics and end matching are assessed before chemical analysis.[101-106] Width, thickness, color, surface texture, and scrim thread count serve as class-level screening parameters; samples differing in any of these are excluded without further analysis.[101,102] Physical end matching of torn or cut ends provides an individualized association, with validation studies confirming false-positive rates of 0–3 % for duct and vinyl electrical tapes.[103-105] The edge similarity score (ESS) provides a reproducible quantitative metric supporting score-based LR calculation.[106,107]

FTIR spectroscopy identifies the backing polymer class and PSA type using ATR configurations that allow nondestructive analysis of both surfaces without sample preparation.[109-111] Backing identification distinguishes among PVC, polyethylene, polypropylene, polyester, and cellulosic substrates; PSA identification resolves natural rubber, synthetic rubber, and acrylic formulations; backing polymers and PSAs within the same class may remain spectrally indistinguishable, requiring elemental analysis for further discrimination.[101,110-112]

Elemental profiling of the backing by LA-ICP-MS or LIBS quantifies trace element concentrations of fillers, stabilizers, and inorganic additives.[112,113] LA-ICP-MS achieved 94% correct discrimination and 100 % correct association across 90 electrical tape rolls; a 7-laboratory interlaboratory exercise confirmed that both LIBS and LA-ICP-MS reliably distinguish different rolls while associating fragments from the same roll.[112,113]

Chemometric methods including PCA and LDA are used to structure the multi-dimensional data generated by FTIR and elemental analysis, converting spectral fingerprints into scores compatible with score-based LR frameworks.[101,102,107]

These examination methods present a unique evidential hierarchy: physical end matching is the only method capable of individualization, and chemical comparisons provide probabilistic evidence of common source only when a fit cannot be established. For end matching, the ESS metric provides a quantitative foundation for score-based LR calculation that is the most operationally ready AI-statistical integration in tape examination.[106,107] For chemical comparison by FTIR and elemental profiling, LR-based evaluation has been demonstrated in research settings[101,102] but lacks the reference population infrastructure for operational deployment in most jurisdictions.

3.6.3. Emerging AI applications

AI-assisted tape examinations are at an earlier stage than glass examinations; however, several applications have been explored. Machine learning classifiers applied to FTIR spectral data can automate backing polymer and PSA subtype assignment, thereby reducing the analysis time during routine classification. For duct tapes, CNNs applied to high-resolution scrim images could in principle automate thread count and weave pattern characterization, which are traditionally assessed using microscopy, though this application has not been reported in the published forensic literature. Hyperspectral imaging combined with deep learning offers the potential to simultaneously acquire morphological and chemical information across an entire tape cross-section, enabling the rapid screening of multiple evidential features in a single analytical pass. As with all trace types, training data coverage of the diversity of tape products in the market is a key limiting factor for model generalizability.[114]

4. Data Analysis, Statistical Evaluation, and Artificial Intelligence

4.1. Validation infrastructure as a prerequisite for operational deployment

The central argument of this review, implicit in the tiered analytical workflows described in Sections 2

and 3, and made explicit here, is that the principal challenge facing forensic trace evidence analysis is not a shortage of statistical or computational tools. Likelihood ratio frameworks, Bayesian networks, score-based LR pipelines, machine learning classifiers, and multivariate chemometric methods are all available; their research-stage capabilities are well documented. The critical shortage is of the conditions under which these tools earn their place in casework: representative reference populations of sufficient geographic and temporal scope; interlaboratory validation studies demonstrating reproducibility under realistic casework conditions; and calibrated performance metrics evaluated at the degraded sample quality and limited sample quantities that characterize operational evidence. A method that is highly discriminating on curated research samples but lacks published error rates or interlaboratory validation does not strengthen forensic evidence, and rather introduces an unquantified and potentially misleading risk.[115–117]

4.2. Likelihood ratio-based evaluation

The LR is a logically correct framework for expressing the evidential value of a forensic comparison. It is defined as the ratio of the probability of the observed analytical findings given the prosecution hypothesis (H_p: the questioned and known samples share a common source) to the probability of the same findings given the defense hypothesis (H_d: the samples originate from different sources). An LR greater than one supports H_p, an LR less than one supports H_d, and the magnitude expresses the discriminative power of the evidence. This Bayesian framework correctly separates the evaluative role of the scientist, which quantifies the weight of the analytical evidence, from the role of the court in assessing prior probabilities and reaching a verdict.[115,116]

In practice, the calculation of a defensible LR requires three elements: (i) a representative background or reference population from which the distribution of the analytical features under Hd can be estimated; (ii) reproducible and validated feature measurements; and (iii) a statistical model that appropriately accounts for within-source and between-source variability. For continuous multivariate data, such as elemental profiles, models based on multivariate normal distributions or kernel density estimation are commonly used. The performance of these models must be empirically validated under case-like conditions before operational deployment.[6,115,116]

Internationally, the deployment of LR-based evaluation varies substantially by trace type and jurisdiction. For glass, LR-based reporting is operationally deployed in the Netherlands Forensic Institute, UK forensic science laboratories, and several Australian state forensic laboratories, supported by dedicated reference databases validated under ASTM E2927 protocols.[78,80,83] For fibers, ENFSI population frequency data underpin LR calculations in several European jurisdictions, though operational deployment remains limited even within Europe; for paint, demonstration has been confined to research settings; and Bayesian network models have been applied in Swiss and Dutch fiber casework.[118,119] For soil, hair, and tape, LR frameworks remain at the research stage in all jurisdictions. The consistent pattern is that LR-based reporting is operational only for glass, and only where dedicated reference databases exist; the preconditions are achievable, but must be built before deployment.

The LR outputs depend on the modeling assumptions, representativeness of the background database, and choice of features included in the model. Poorly representative databases or inappropriate models can produce misleading LR values; transparent reporting should therefore include an explicit description of the background population, the statistical model and its validation performance, and the features on which the calculation is based. Verbal equivalents for LR ranges (e.g., “moderate support,” “strong support”) should be aligned with published frameworks to ensure consistent communication to courts and to avoid the well-documented risks of forensic overstatement.[69,115,116]

4.3. Emerging interpretation frameworks

Several complementary frameworks for evidence interpretation have gained traction alongside the

classical LR. Bayesian Networks (BNs) extends the LR approach to complex multivariable inference problems by modeling the conditional dependencies among transfer, persistence, recovery, and background variables as directed acyclic graphs. This is particularly valuable for activity-level propositions, in which the relationship between a trace and an alleged act is contested and multiple probabilistic nodes must be evaluated simultaneously. Recent studies have introduced template BN models for evaluating transfer evidence when the item–activity link is uncertain, and object-oriented BNs have been applied to secondary fiber transfer scenarios in which conventional LR calculations become intractable. These frameworks extend the LR approach in directions that are scientifically promising but share a common requirement: they cannot be responsibly deployed without the validation infrastructure described in Section 4.4.[118,119]

Score-based LR methods offer a practical bridge between high-dimensional analytical data and probabilistic evaluations, reducing data to a single similarity score derived from a distance metric or classifier output and estimating the LR from its distribution across same-source and different-source pairs; this approach is computationally tractable for high-dimensional data and naturally integrated into machine learning pipelines. An emerging complement is conformal prediction, a distribution-free framework that constructs statistically valid prediction sets or uncertainty intervals for any pretrained model without requiring parametric assumptions, addressing the recognized limitation of black-box AI models in forensic contexts: the inability to quantify uncertainty in a principled and model-agnostic manner.

4.4. Validation Requirements and critical perspective

Analytical and statistical advances do not automatically translate into stronger forensic evidence, and a critical perspective on method adoption is warranted. The admissibility of expert evidence in many jurisdictions requires the underlying methodology to be scientifically valid, reliably applied, and capable of withstanding cross-examination. Methods that are highly discriminating in research settings but lack published error rates, interlaboratory validation, or documented performance on casework-quality samples may introduce risk rather than reduce it. A poorly calibrated LR, a BN whose conditional probability tables are populated with unvalidated estimates, or an AI classifier whose training data are unrepresentative of the casework population can produce outputs that overstate the strength of the evidence and mislead the trier of fact. Accordingly, the adoption of any new interpretation method in operational casework should be preceded by rigorous empirical validation under realistic conditions and independent peer review irrespective of its theoretical appeal or reported performance on curated datasets. A method earns its place in evidence-based casework not through computational sophistication but through the quality of its validation.[68,115–117]

4.5. Implementation in the Korean forensic science environment

4.5.1. Institutional and legal context

Korea operates a continental inquisitorial legal system in which judges evaluate evidence and determine facts in criminal proceedings.[120] This structure has an important implication for the adoption of probabilistic evidence reporting: professional judges are in principle better equipped to engage with LR-based conclusions than lay juries, as empirical studies have consistently demonstrated systematic misinterpretation of LR values even among educated lay participants.[121-123] The Korean judicial system's established practice of giving substantial weight to forensic expert testimony, combined with the recognized status of physical evidence as the most reliable evidentiary basis for conviction, creates a legal environment that is in principle receptive to probabilistic reporting, provided that the method is presented with transparent documentation of its validation basis and limitations.

The National Forensic Service (NFS), the primary forensic institution providing casework analysis in Korea under the Ministry of the Interior and Safety, is KOLAS-accredited and already employs high-resolution analytical platforms, including ICP-MS

for elemental trace work and FTIR/Raman systems for trace material characterization, in its Forensic Chemistry Division. This analytical infrastructure is a necessary precondition; what is currently absent is the jurisdiction-specific reference population framework required to calculate defensible LR values.

4.5.2. The reference population gap and priority materials

For Korea specifically, three trace types present the most urgent and tractable priority for reference population development. Glass has the most operationally mature LR methodology internationally (LA-ICP-MS with ASTM E2927, as described in Section 3.4), and the NFS internal glass elemental reference database provides the necessary compositional foundation for LR implementation. The primary remaining gap is Korean-specific transfer and persistence data; combining these with the existing database would provide the foundation for defensible LR reporting in relevant cases. Automotive paint presents a similar opportunity: the Korean vehicle fleet, dominated by Hyundai, Kia, and Genesis vehicles, has a paint layer chemistry partially distinct from European or North American databases, and the NFS already has the analytical capability to characterize paint pigment and extender compositions. Fiber is addressed through the AFSN Trace Evidence Working Group (TEWG), within which Korea is an active member; the TEWG should prioritize initiating collaborative fiber database development among member states including Thailand, China, Singapore, and Korea, as such a regionally specific database is a prerequisite for defensible LR-based fiber evidence reporting in the region.[2,124]

4.5.3. A staged implementation path

The implementation path should proceed in stages, with each conditioned on completion of the preceding one, reflecting the principle that operational deployment without validation introduces errors that are harder to detect than those of the qualitative methods they replace.

In the immediate term, the most productive investment is not tool adoption but reference population construction. Systematic sampling programs for the three priority materials, textiles, automotive paint, and glass, should be initiated with sampling designs that reflect the Korean casework context: vehicle types, glass manufacturers, and garment provenance distributions encountered in actual cases. This data collection phase can proceed in parallel with NFS participation in AFSN-TEWG collaborative exercises, which provide both interlaboratory validation experience and the opportunity to test Korean-specific samples against regional reference materials.[124]

In the medium term, LR implementation can begin with glass evidence, the trace type with the most mature international methodology under ASTM E2927 and the highest Korean casework frequency (traffic cases, forced entry). The initial implementation should use kernel density estimation or multivariate normal models applied to LA-ICP-MS elemental profiles, with the model's calibration and performance validated empirically on a hold-out sample before operational deployment. Bayesian network implemen- tations for activity-level propositions require not only source-level LR values but also transfer and persistence probability estimates, and should therefore be regarded as a longer-term development, conditional on the availability of Korean-specific transfer and persistence data for each trace type.

In the longer term, the development of Korean probabilistic reporting frameworks should be coordinated through AFSN and supported by bilateral technical exchange with ENFSI member laboratories that have operational LR reporting experience. The ENFSI Evaluative Reporting Guidelines provide a practical template that can be adapted to the Korean judicial context, and AFSN's Quality Assurance and Standards Committee provides a regional governance mechanism for harmonizing probabilistic reporting practices across member states.[121,124] The goal is a Korean implementation grounded in Korean reference populations, validated under Korean casework conditions, and presented in a form that Korean courts can evaluate and scrutinize; such an implementation is also a genuine scientific contribution, given that Northeast Asian LR performance data are almost

entirely absent from the published record and NFS is positioned to close that gap.

5. Professional Organizations, Working Groups, and Standardization Bodies

Trace evidence examination practices are governed by a network of professional organizations and standard development bodies. In the United States, the SWGMAT established foundational best-practice guidelines for fibers, hair, paint, glass, and tape before transitioning to the OSAC framework; the OSAC Trace Materials Subcommittee now develops formally peer-reviewed standards under NIST administration, including the Standard Guide for Forensic Trace Evidence Recovery (OSAC 2023-N-0027) and discipline-specific examination and reporting standards.[116,125-128] In Europe, the ENFSI Trace Evidence Working Group produces Best Practice Manuals, organizes interlaboratory proficiency exercises, and has published evaluative reporting guidelines that provide a widely adopted framework for LR-based evidence.[4,88,121] At the international level, INTERPOL's triennial Forensic Science Managers Symposium produces systematic literature reviews across all trace types, and ASTM Committee E30 formalizes voluntary consensus standards, including ASTM E1610 for paint analysis.[33,129,130] In the Asia-Pacific region, the AFSN and its Trace Evidence Working Group support harmonization among member states; however, regionally specific reference databases remain largely absent, and their systematic construction is a prerequisite for defensible LR-based evidence reporting in the region.[2,124]

6. Conclusion

This review provides an integrated survey of six categories of forensic trace evidence (fibers, hair, paint, soil, glass, and tapes) within a tiered analytical framework that treats analytical method, transfer and persistence context, and evidential interpretation as inseparable. That integration is the central argument of the review. A fiber is not simply a polymer to be identified by FTIR; it is a material whose significance depends on how readily it transfers between garments, how long it persists on a given substrate, and how unusual its combination of color, morphology, and chemistry is within the relevant population. The same logic applies to every trace type surveyed: glass fragments whose persistence on clothing decreases sharply within the first hour of breakage; soil whose discriminating power arises from the combined pattern of its mineral and anthropogenic constituents rather than any single component; and tape whose physical end match can individualize an association that chemical profiling alone cannot achieve. Analytical sophistication serves these questions. It does not replace them.

The tiered workflow provides a hierarchical strategy that ensures advanced, resource-intensive methods are reserved for cases where their discriminatory power is essential. *Fig.* 2 reinforces this principle: no single platform is optimal across all six evaluation criteria, and method selection must be driven by the propositions at issue, not by instrument availability alone.

Statistical evaluation frameworks (e.g. LR, BNs, and emerging score-based and conformal prediction methods as discussed in this paper) provide the principled language for expressing how strongly analytical findings support one proposition over another. Their adoption is scientifically correct, but their value is conditional. An LR derived from a nonrepresentative background database, or a model validated only on curated laboratory samples, can misrepresent the evidence it is intended to quantify. The critical perspective maintained throughout this review is therefore not a counsel of caution about new methods but a reminder that the primary challenge for forensic trace evidence science is not tool acquisition but validation infrastructure. Therefore, building the representative reference populations, interlaboratory harmonization frameworks, and performance benchmarks under casework conditions that give any method, analytical or statistical, its evidential legitimacy is essential. Professional organizations such as OSAC, ENFSI, AFSN, and ASTM exist precisely

to operationalize these conditions across the international community, and NFS is positioned to contribute jurisdiction-specific reference data from the Korean forensic science context that is currently absent from the published record.[115,116,124]

Current forensic trace evidence analysis operates under structural constraints that apply across all methods reviewed here: even high-resolution elemental analysis establishes consistency rather than individualization, with shared manufacturing processes and background population overlap setting a fundamental limit that no analytical advance alone can overcome. Forensic reference datasets are also typically several orders of magnitude smaller than those on which deep learning architectures were validated, and casework samples are often degraded, limited in quantity, and unavailable for replicate analysis; reference populations built from one region or time period may not represent compositions encountered elsewhere, as formulations for paints, glass, and adhesive tapes change across product generations and production regions. Addressing these constraints will require longitudinal casework datasets, collaborative development of region- and time-specific reference collections, and validation protocols designed explicitly for degraded and limited-quantity materials.

Across all six trace types, the interpretive value of laboratory findings ultimately depends on what was preserved at the point of recovery. Careful scene documentation, correct packaging, and unbroken chain of custody are not administrative requirements peripheral to the analytical work; they are the foundations on which every finding in this review rests. Locard's principle, that no criminal act can be carried out with sufficient intensity without leaving traces of those involved, is the starting point of trace evidence analysis, but that sign must first be found, preserved, and correctly interpreted before it can speak to what actually happened, and it is at the crime scene as much as in the laboratory that the quality of trace evidence analysis is ultimately determined.

Acknowledgements

This work was supported by a grant from the National Forensic Service (NFS2026CHE-01) of the Ministry of Interior and Safety of the Republic of Korea.

References

1. C. Roux, B. Talbot-Wright, J. Robertson, F. Crispino, and O. Ribaux, *Philos. Trans. R. Soc. B*, **370**, 20140260 (2015). https://doi.org/10.1098/rstb.2014.0260
2. J. R. Almirall and T. Trejos, *Forensic Sci. Int. Synergy*, **6**, 100306 (2023). https://doi.org/10.1016/j.fsisyn.2022.100306
3. T. Trejos, S. Koch, and A. Mehltretter, *Forensic Chem.*, **18**, 100223 (2020). https://doi.org/10.1016/j.forc.2020.100223
4. ENFSI, Best Practice Manual for the Forensic Examination of Fibres. BPM-THG-04, ENFSI, Brussels (2022).
5. G. Zadora and D. Ramos, *Sci. Justice*, **50**, 32 (2010). https://doi.org/10.1016/j.scijus.2009.11.028
6. C. G. G. Aitken and D. Lucy, *Appl. Stat.*, **53**(1), 109-122 (2004). https://doi.org/10.1046/j.0035-9254.2003.05271.x
7. M. C. Grieve, T. W. Biermann and K. Schaub, *Sci. Justice*, **45**(1), 13-28 (2005). https://doi.org/10.1016/S1355-0306(05)71616-6
8. SWGMAT, *Forensic Sci. Commun.*, **2**(1) (2000).
9. M. C. Grieve, *Forensic Sci. Rev.*, **6**(1), 59-80 (1994).
10. S. F. Stoeffler, *J. Forensic Sci.*, **41**(2), 297-299 (1996).
11. J. A. Reffner, B. W. Kammrath, and S. Kaplan, *J. Forensic Sci.*, **65**(3), 744-750 (2020). https://doi.org/10.1111/1556-4029.14252
12. M. C. Grieve and K. G. Wiggins, *J. Forensic Sci.*, **46**(4), 835-843 (2001). https://doi.org/10.1520/JFS15055J
13. N. Petraco, *J. Police Sci. Admin.*, **9**, 448 (1981).
14. M. C. Grieve and T. M. Kotowski, *J. Forensic Sci. Soc.*, **26**(1), 29-34 (1986). https://doi.org/10.1016/S0015-7368(86)72443-2
15. M. C. Grieve, T. W. Biermann, and M. Davignon, *Sci. Justice*, **41**(4), 245-260 (2001). https://doi.org/10.1016/S1355-0306(01)71908-9
16. K. G. Wiggins, R. Palmer, W. Hutchinson, and P. Drummond, *Sci. Justice*, **47**, 9-18 (2007). https://doi.org/

10.1016/j.scijus.2006.11.001
17. M. C. Grieve, T. W. Biermann, and M. Davignon, *Sci. Justice*, **43**(1), 5-22 (2003). https://doi.org/10.1016/S1355-0306(03)71736-5
18. M. Eng, P. Martin, and C. Bhagwandin, *J. Forensic Sci.*, **54**(4), 841-845 (2009). https://doi.org/10.1111/j.1556-4029.2009.01079.x
19. J. Was-Gubala and R. Starczak, *Spectrochim. Acta A*, **142**, 118-125 (2015). https://doi.org/10.1016/j.saa.2015.01.116
20. M. W. Tungol, E. G. Bartick, and A. Montaser, *Appl. Spectrosc.*, **47**(10), 1655-1658 (1993). https://doi.org/10.1366/0003702934065715
21. E. G. Bartick, M. W. Tungol, and J. A. Reffner, *Anal. Chim. Acta*, **288**(1-2), 35-42 (1994). https://doi.org/10.1016/0003-2670(94)85114-X
22. M. C. Grieve, *Sci. Justice*, **35**(3), 179-190 (1995). https://doi.org/10.1016/S1355-0306(95)72659-4
23. D. K. Laing, A. W. Hartshorne, and R. J. Harwood, *Forensic Sci. Int.*, **30**(1), 65-77 (1986). https://doi.org/10.1016/0379-0738(86)90179-9
24. G. Massonnet et al., *J. Forensic Sci.*, **50**(5), 1028-1038 (2005). https://doi.org/10.1520/JFS2004329
25. L. Lepot, K. De Wael, F. Gason, and B. Gilbert, *Sci. Justice*, **48**(3), 109-117 (2008). https://doi.org/10.1016/j.scijus.2007.09.013
26. G. Massonnet et al., *Forensic Sci. Int.*, **222**(1-3), 200-207 (2012). https://doi.org/10.1016/j.forsciint.2012.05.025
27. P. Buzzini and G. Massonnet, *J. Forensic Sci.*, **58**(6), 1593-1600 (2013). https://doi.org/10.1111/1556-4029.12298
28. P. Buzzini and G. Massonnet, *J. Forensic Sci.*, **60**(3), 712-720 (2015). https://doi.org/10.1111/1556-4029.12654
29. M. Huang, J. Yinon, and M. Sigman, *J. Forensic Sci.*, **49**(2), 238-249 (2004). https://doi.org/10.1520/JFS2003295
30. L. M. Petrick, T. A. Wilson, and W. R. Fawcett, *J. Forensic Sci.*, **51**(4), 771-779 (2006). https://doi.org/10.1111/j.1556-4029.2006.00155.x
31. J. V. Goodpaster and E. A. Liszewski, *Anal. Bioanal. Chem.*, **394**, 2095-2103 (2009). https://doi.org/10.1007/s00216-009-2786-8
32. M. Huang, R. Russo, B. G. Fookes, and M. E. Sigman, *J. Forensic Sci.*, **50**(3), 526-534 (2005). https://doi.org/10.1520/JFS2004223
33. F. Bianchi, M. Careri, E. Marengo, and M. Musci, *J. Chromatogr. A*, **975**(1), 113-121 (2002). https://doi.org/10.1016/S0021-9673(02)00881-6
34. C. Hu, J. Zhu, H. Mei, H. Shi, H. Guo, G. Zhang, P. Wang, L. Lu, and X. Zheng, *Forensic Chem.*, **11**, 1-6 (2018). https://doi.org/10.1016/j.forc.2018.08.001
35. C. Vooijs, P. Vergeer, and J. van der Weerd, *Forensic Sci. Int.*, **250**, 57-67 (2015). https://doi.org/10.1016/j.forsciint.2015.02.009
36. A. Aljannahi, R. A. Alblooshi, R. H. Alremeithi, I. Karamitsos, N. A. Ahli, A. M. Askar, I. M. Albastaki, M. M. Ahli, and S. Modak, *Molecules*, **27**(13), 4281 (2022). https://doi.org/10.3390/molecules27134281
37. SWGMAT, *Forensic Sci. Commun.*, **7**(2) (2005).
38. W. F. Rowe, *Sci. World J.*, **1**, 452513 (2001). https://doi.org/10.1100/tsw.2001.358
39. C. T. Oien, *Forensic Sci. Commun.*, **11**(2) (2009).
40. P. Pienpinijtham, C. Thammacharoen, S. Naranitad, and S. Ekgasit, *Spectrochim. Acta A*, **197**, 230-236 (2018). https://doi.org/10.1016/j.saa.2018.01.084
41. D. S. H. Funes, K. Bonilla, M. Baudelet, and C. Bridge, *Forensic Sci. Int.*, **346**, 111622 (2023). https://doi.org/10.1016/j.forsciint.2023.111622
42. ASTM E3316-22, Standard guide for forensic examinations of hair by microscopy, ASTM International (2022).
43. B. D. Gaudette and E. S. Keeping, *J. Forensic Sci.*, **19**(3), 599-606 (1974).
44. R. A. Wickenheiser and D. G. Hepworth, *J. Forensic Sci.*, **35**(6), 1323-1329 (1990).
45. M. M. Houck and B. Budowle, *J. Forensic Sci.*, **47**(5), 964-967 (2002).
46. M. S. Boll, K. C. Doty, R. Wickenheiser, and I. K. Lednev, *Forensic Chem.*, **6**, 1-9 (2017). https://doi.org/10.1016/j.forc.2017.08.001
47. D. Kurouski and R. P. Van Duyne, *Anal. Chem.*, **87**(5), 2901-2906 (2015). https://doi.org/10.1021/ac504405u
48. D. Pozebon, G. L. Scheffler, and V. L. Dressler, *Anal. Chim. Acta*, **992**, 1-23 (2017). https://doi.org/10.1016/j.aca.2017.09.017
49. S. B. Hong, K. C. Kim, and W. Kim, *Forensic Sci. Int. Genet.*, **17**, 99-103 (2015). https://doi.org/10.1016/j.fsigen.2015.03.017
50. M. Airlie, J. Robertson, and E. Brooks, *Aust. J. Forensic Sci.*, **56**(5), 551-565 (2024). https://doi.org/10.1080/00450618.2023.2254337
51. M. Airlie, J. Robertson, W. Ma, and E. Brooks, *Aust. J.*

Forensic Sci., **57**(3), 340-353 (2025). https://doi.org/10.1080/00450618.2024.2343368
52. B. Caddy, *Forensic Examination of Glass and Paint: Analysis and Interpretation*, Taylor & Francis, London (2001).
53. P. Buzzini and J. M. Curran, *Forensic Paint Analysis*, NCJRS Report 311039 (2025).
54. SWGMAT, *Forensic Sci. Commun.*, **1**(2) (1999).
55. ASTM E2809-22, Standard Guide for Using Scanning Electron Microscopy/Energy Dispersive X-Ray Spectroscopy (SEM/EDS) in Forensic Polymer Examinations, ASTM International (2022). https://doi.org/10.1520/E2809-22
56. P. Buzzini and E. Suzuki, *J. Raman Spectrosc.*, **47**(1), 16-27 (2016). https://doi.org/10.1002/jrs.4818
57. J. Zieba-Palus, *J. Mol. Struct.*, **511-512**, 327-335 (1999). https://doi.org/10.1016/S0022-2860(99)00175-1
58. J. Zieba-Palus, in 'Encyclopedia of Analytical Chemistry', R. A. Meyers, and T. A. Brettell (eds.), Wiley (2006). https://doi.org/10.1002/9780470027318.a1117
59. M. A. Malek, T. Nakazawa, H.-W. Kang, K. Tsuji, and C.-U. Ro, *Molecules*, **24**(7), 1381 (2019). https://doi.org/10.3390/molecules24071381
60. P. Buzzini, G. Massonnet, and F. Monard Sermier, *J. Raman Spectrosc.*, **37**(9), 922-931 (2006). https://doi.org/10.1002/jrs.1522
61. J. Almirall, T. Trejos, and K. Lambert, *Forensic Sci. Int. Synergy*, **2**, 404-415 (2020). https://doi.org/10.1016/j.fsisyn.2020.01.010
62. A. L. Hobbs and J. R. Almirall, *Anal. Bioanal. Chem.*, **376**, 1265-1271 (2003). https://doi.org/10.1007/s00216-003-1918-x
63. D. M. Wright, M. J. Bradley, and A. H. Mehltretter, *J. Forensic Sci.*, **58**(2), 358-364 (2013). https://doi.org/10.1111/1556-4029.12074
64. J. Zieba-Palus and B. M. Trzcinska, *J. Forensic Sci.*, **58**(5), 1359-1363 (2013). https://doi.org/10.1111/1556-4029.12183
65. J. Zieba-Palus, G. Zadora, J. M. Milczarek, and P. Koscielnak, *J. Chromatogr. A*, **1179**(1), 41-46 (2008). https://doi.org/10.1016/j.chroma.2007.09.044
66. A. Michalska, A. Martyna, and G. Zadora, *J. Raman Spectrosc.*, **46**, 1041-1048 (2015). https://doi.org/10.1002/jrs.4719
67. A. Martyna, A. Michalska, and G. Zadora, *Anal. Bioanal. Chem.*, **407**, 5931-5942 (2015). https://doi.org/10.1007/s00216-015-8558-9
68. S. Farber, *J. Forensic Sci.*, **70**(1), e70035 (2025). https://doi.org/10.1111/1556-4029.70035
69. J. M. Curran, C. M. Triggs, J. R. Almirall, J. S. Buckleton, and K. A. J. Walsh, *Sci. Justice*, **37**(4), 245-249 (1997). https://doi.org/10.1016/S1355-0306(97)72198-1
70. R. Palmer, K. Sheridan, J. Puckett, N. Richardson, and W. Lo, *Forensic Sci. Int.*, **278**, 334-337 (2017). https://doi.org/10.1016/j.forsciint.2017.07.035
71. E. Brooks, M. Prusinowski, S. Gross, and T. Trejos, *Forensic Sci. Int.*, **313**, 110349 (2020). https://doi.org/10.1016/j.forsciint.2020.110349
72. J. Locke, *Microscope*, **33**(3), 169-178 (1985).
73. ENFSI, Guideline for the Determination of Refractive Index, ENFSI Trace Evidence Working Group (2022).
74. R. D. Koons and J. Buscaglia, *J. Forensic Sci.*, **44**(3), 496-503 (1999). https://doi.org/10.1520/JFS14500J
75. E. J. Garvin and R. D. Koons, *J. Forensic Sci.*, **56**(2), 491-500 (2011). https://doi.org/10.1111/j.1556-4029.2010.01634.x
76. J. A. Buscaglia, *Anal. Chim. Acta*, **288**, 17-24 (1994). https://doi.org/10.1016/0003-2670(94)85112-3
77. S. G. Ryland, *J. Am. Soc. Trace Evid. Examiners*, **2**(1), 2-12 (2011).
78. T. Trejos et al., *Anal. Bioanal. Chem.*, **405**, 5393-5409 (2013). https://doi.org/10.1007/s00216-013-6978-y
79. T. Trejos, *WIREs Forensic Sci.*, **5**(2), e1476 (2023). https://doi.org/10.1002/wfs2.1476
80. T. Trejos, S. Montero, and J. R. Almirall, *Anal. Bioanal. Chem.*, **376**(8), 1255-1264 (2003). https://doi.org/10.1007/s00216-003-1968-0
81. R. Corzo et al., *Forensic Chem.*, **22**, 100307 (2021). https://doi.org/10.1016/j.forc.2021.100307
82. J. M. Curran, *Sci. Justice*, **56**(5), 380-382 (2016). https://doi.org/10.1016/j.scijus.2016.05.005
83. T. Trejos et al., *J. Anal. At. Spectrom.*, **28**, 1270-1282 (2013). https://doi.org/10.1039/c3ja50098a
84. O. Kaspi, O. Israelsohn-Azulay, Z. Yigal, H. Rosengarten, M. Krmpotić, S. Gouasmia, I. Bogdanović Radović, P. Jalkanen, A. Liski, K. Mizohata, J. Räisänen, Z. Kasztovszky, I. Harsnyi, R. Acharya, P. K. Pujari, M. Molnár, M. Braun, N. Shabi, O. Girshevitz, and H. Senderowitz,

J. Chem. Inf. Model., **63**(1), 227-240 (2023). https://doi.org/10.1021/acs.jcim.2c01362
85. S. Park and A. Carriquiry, *Ann. Appl. Stat.*, **13**(2), 1068-1102 (2019). https://doi.org/10.1214/18-AOAS1211
86. T. Trejos and J. R. Almirall, *Anal. Chem.*, **76**(5), 1236-1242 (2004). https://doi.org/10.1021/ac034933o
87. R. M. Morgan and P. A. Bull, *Sci. Justice*, **47**(3), 107 (2007). https://doi.org/10.1016/j.scijus.2007.02.001
88. ENFSI, Best Practice Manual for the Forensic Comparison of Soil, ENFSI-BPM-APS-02, ENFSI, Brussels (2019).
89. R. Sugita and Y. Marumo, *Forensic Sci. Int.*, **83**(3), 201-210 (1996). https://doi.org/10.1016/S0379-0738(96)02038-5
90. R. Sugita and Y. Marumo, *Forensic Sci. Int.*, **122**(2-3), 155-158 (2001). https://doi.org/10.1016/S0379-0738(01)00490-X
91. N. Petraco, T. A. Kubic, and N. D. K. Petraco, *Forensic Sci. Int.*, **178**(2-3), e23-e27 (2008). https://doi.org/10.1016/j.forsciint.2008.03.008
92. A. Ruffell and P. Wiltshire, *Forensic Sci. Int.*, **145**(1), 13-23 (2004). https://doi.org/10.1016/j.forsciint.2004.03.017
93. J. M. L. Mazzetto, V. F. Melo, E. J. Bonfleur, P. Vidal-Torrado, and J. Dieckow, *Sci. Justice*, **59**(6), 635-642 (2019). https://doi.org/10.1016/j.scijus.2019.07.003
94. G. McCulloch, L. A. Dawson, M. J. Brewer, and R. M. Morgan, *Forensic Sci. Int.*, **272**, 127-141 (2017). https://doi.org/10.1016/j.forsciint.2017.01.009
95. N. G. Cameron, *Geol. Soc. Spec. Publ.*, **232**, 277-280 (2004). https://doi.org/10.1144/GSL.SP.2004.232.01.20
96. K. Pye and S. J. Blott, *Sci. Justice*, **49**(3), 170-181 (2009). https://doi.org/10.1016/j.scijus.2009.02.007
97. B. Woods, K. P. Kirkbride, C. Lennard, and J. Robertson, *Forensic Sci. Int.*, **245**, 195-201 (2014). https://doi.org/10.1016/j.forsciint.2014.10.018
98. A. J. Jeffery, D. Pirrie, J. K. Pringle, A. Ruffell, and I. G. Stimpson, *Geol. Today*, **40**, 139-145 (2024). https://doi.org/10.1111/gto.12484
99. M. G. Aberle, P. de Caritat, J. Robertson, and J. A. Hoogewerff, *Forensic Sci. Int.*, **353**, 111883 (2023). https://doi.org/10.1016/j.forsciint.2023.111883
100. Korea Institute of Geoscience and Mineral Resources (KIGAM), National Geochemical Survey of Korea, KIGAM, Daejeon, Republic of Korea (2019).
101. M. Prusinowski, Z. Andrews, C. Neumann, and T. Trejos, *Forensic Sci. Int.*, **343**, 111567 (2023). https://doi.org/10.1016/j.forsciint.2023.111567
102. M. Prusinowski, A. Mehltretter, C. Martinez-Lopez, J. Almirall, and T. Trejos, *Forensic Chem.*, **13**, 100146 (2019). https://doi.org/10.1016/j.forc.2019.100146103. SWGMAT, *J. Am. Soc. Trace Evid. Exam.*, **2**(1), 88-97 (2011).
104. SWGMAT, *J. Am. Soc. Trace Evid. Exam.*, **2**(1), 98-105 (2011).
105. M. J. Bradley, R. L. Keagy, P. C. Lowe, M. P. Rickenbach, D. M. Wright, and M. A. LeBeau, *J. Forensic Sci.*, **51**(3), 504-508 (2006). https://doi.org/10.1111/j.1556-4029.2006.00106.x
106. M. J. Bradley, J. M. Gauntt, A. H. Mehltretter, P. C. Lowe, and D. M. Wright, *J. Forensic Sci.*, **56**(3), 606-611 (2011). https://doi.org/10.1111/j.1556-4029.2011.01736.x
107. M. Prusinowski, E. Brooks, and T. Trejos, *Forensic Sci. Int.*, **307,** 110103 (2020). https://doi.org/10.1016/j.forsciint.2019.110103
108. M. Prusinowski et al., *Forensic Sci. Int.*, **343**, 111567 (2023). https://doi.org/10.1016/j.forsciint.2023.111567
109. R. A. Merrill and E. G. Bartick, *J. Forensic Sci.*, **45**(1), 93-98 (2000). https://doi.org/10.1520/JFS14645J
110. A. H. Mehltretter, M. J. Bradley, and D. M. Wright, *J. Forensic Sci.*, **56**(1), 82-94 (2011). https://doi.org/10.1111/j.1556-4029.2010.01560.x
111. A. H. Mehltretter, M. J. Bradley, and D. M. Wright, *J. Forensic Sci.*, **56**(6), 1493–1504 (2011). https://doi.org/10.1111/j.1556-4029.2011.01873.x
112. C. Martinez-Lopez, T. Trejos, A. H. Mehltretter, and J. R. Almirall, *Forensic Chem.*, **4**, 96-107 (2017). https://doi.org/10.1016/j.forc.2017.03.003
113. C. Martinez-Lopez et al., *Forensic Chem.*, **12**, 66–77 (2019). https://doi.org/10.1016/j.forc.2019.01.001
114. J. Stacey, R. Fleming, D. Sheppard, J. Sheppard, G. Dobbie, and D. Karunakaran, *Forensic Sci. Int.*, **375**, 112548 (2025). https://doi.org/10.1016/j.forsciint.2025.112548
115. I. W. V. Evett, *Forensic Sci.*, **9**(3), 209-217 (1977). https://doi.org/10.1016/0300-9432(77)90093-0
116. OSAC, Standard Guide for Interpretation and Reporting in Forensic Comparisons of Trace Evidence, 2022-S-0029, OSAC (2022).
117. I. Hefetz, *Forensic Sci. Int. Synergy*, **11**, 100645 (2025). https://doi.org/10.1016/j.fsisyn.2025.100645
118. Y. C. Lim-Hitchings, F. Taroni, and G. Massonnet, *Forensic*

Sci. Int., **361**, 112086 (2024). https://doi.org/10.1016/j.forsciint.2024.112086
119. M. Vink and M. J. Sjerps, *Forensic Sci. Int. Synergy*, **9**, 100546 (2024). https://doi.org/10.1016/j.fsisyn.2024.100546
120. M. Jang, *Forensic Sci. Int. Synergy*, **9**, 100544 (2024). https://doi.org/10.1016/j.fsisyn.2024.100544
121. ENFSI, Guideline for Evaluative Reporting in Forensic Science, Version 3.0, ENFSI, Brussels (2015).
122. W. C. Thompson and E. L. Schumann, *Law Hum. Behav.*, **11**(3), 167-187 (1987). https://doi.org/10.1007/BF01044641
123. K. A. Martire, R. I. Kemp, M. Sayle, and B. R. Newell, *Forensic Sci. Int.*, **240**, 61-68 (2014). https://doi.org/10.1016/j.forsciint.2014.04.005
124. A. Yap and N. Cheng, *Forensic Sci. Int.*, **359**, 112042 (2024). https://doi.org/10.1016/j.forsciint.2024.112042
125. SWGMAT, *Forensic Sci. Commun.*, **2**(1) (2000).
126. SWGMAT, *Forensic Sci. Commun.*, **1**(3) (1999).
127. OSAC, Standard Guide for Forensic Trace Evidence Recovery, OSAC 2023-N-0027 Version 2.0, NIST, Gaithersburg, MD (2024).
128. OSAC, Standard Practice for Forensic Trace Evidence Recovery, 2023-N-0027, NIST (2023).
129. INTERPOL, INTERPOL Review of Trace Evidence 2019-2022, INTERPOL, Lyon (2022).
130. ASTM E30 Committee, *Standard Guide for Forensic Paint Analysis and Comparison*, ASTM E1610, ASTM International, West Conshohocken, PA (2018).