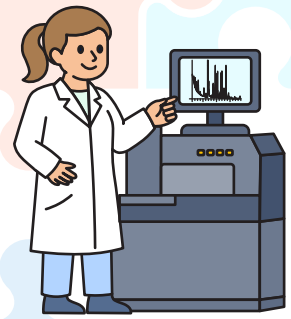
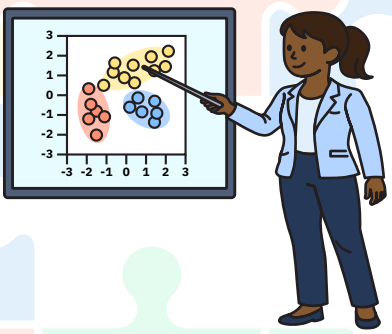
Chemical Forensics

Report of the Scientific Advisory Board's Temporary Working Group





SAB/REP/2/26
April 2026

ACKNOWLEDGEMENTS

The Temporary Working Group (TWG) on Chemical Forensics would like to thank all the people who have contributed during the two-year mandate of the Group. It is most appreciative to all the invited external speakers who gave up their time to come and make presentations to the TWG, and their significant inputs during the meetings. These have been most valuable and informative, have covered various and complementary topics, and have provided considerable insights for the Group. The TWG is grateful to the Director-General of the OPCW for his leadership in establishing this Group on Chemical Forensics in support of the Chemical Weapons Convention. Further to this, the TWG would like to acknowledge the Technical Secretariat for its help and support during the Group's mandate. In particular, the interns for image generation, Peter Hotchkiss for his support of the TWG, input and guidance when requested, and Sarah Clapham for her invaluable help in distilling the technical contributions to the end-of-mandate report. A special mention to Ernesa Ademagić for the support she has provided the TWG over its mandate. No request was too big, and we truly appreciated everything she has done for the Group. Finally, the TWG thanks the European Union and the United States of America for the financial support that enabled the Group's activities, and the Finnish Institute for Verification of the Chemical Weapons Convention (VERIFIN) for hosting the final plenary TWG meeting in Helsinki, Finland.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

Abbreviation	Definition
AI	Artificial intelligence
API	Application programming interface
Blue Book	Recommended Operating Procedures for Analysis in the Verification of Chemical Disarmament
CBE	Confidence-building exercise
CFITWG	Chemical Forensics International Technical Working Group
CFM-ID	Competition Fragmentation Modeling-Identification
CFSDB	Chemical forensics spectral database
ChemTech Centre	Centre for Chemistry and Technology
Convention	Chemical Weapons Convention
CWA	Chemical warfare agent
DAT	Declaration Assessment Team
DC	Methylphosphonyl dichloride
DF	Methylphosphonyl difluoride
DGA	French Defence Agency
DL	Designated Laboratory
DNA	Deoxyribonucleic acid
ENFSI	European Network of Forensic Science Institutes
FFM	Fact-Finding Mission
FOI	Swedish Defence Research Agency
GBM	Gradient boosting methods
GC	Gas chromatography
GC-HRMS	Gas chromatography coupled to high-resolution mass spectrometry
GC-MS	Gas chromatography-mass spectrometry
HRMS	High-resolution mass spectrometry
IC	Ion chromatography

ICP-MS	Inductively coupled plasma mass spectrometry
IEC	International Electrotechnical Commission
IIT	Investigation and Identification Team
InChIKey	International Chemical Identifier Key
INRAE	Institut national de recherche pour l'agriculture, l'alimentation et l'environnement
INTERPOL	International Criminal Police Organization
IRMS	Isotope ratio mass spectrometry
ISO	International Organization for Standardization
JIM	(OPCW-United Nations) Joint Investigative Mechanism
LA-ICP-MS	Laser ablation-inductively coupled plasma mass spectrometry
LC	Liquid chromatography
LC-HRMS	Liquid chromatography coupled to high-resolution mass spectrometry
LC-MS	Liquid chromatography-mass spectrometry
LR	Likelihood ratio
MALDI	Matrix-assisted laser desorption ionisation
ML	Machine learning
MS	Mass spectrometry
NIST	National Institute of Standards and Technology
NMR	Nuclear magnetic resonance (spectroscopy)
OCAD	OPCW Central Analytical Database
OPCW	Organisation for the Prohibition of Chemical Weapons
OPLS-DA	Orthogonal partial least squares discriminant analysis
OSM	Office of Special Missions
PCA	Principal component analysis
PLS-DA	Partial least squares discriminant analysis
PSIA-NMR	Position-specific isotope analysis by nuclear magnetic resonance (spectroscopy)
PT	Proficiency Test
QA	Quality assurance

QC	Quality control
QMS	Quality management system
RMP	Random match probability
R/SOP	Recommended and/or standard operating procedure
RF	Random forest
SAB	Scientific Advisory Board
Secretariat	Technical Secretariat
SMILES	Simplified molecular-input line-entry system
TATP	Triacetone triperoxide
TNO	Netherlands Organisation for Applied Scientific Research
TNT	Trinitrotoluene
TOR	Terms of reference
TWG	Temporary Working Group
TWG on IST	Temporary Working Group on Investigative Science and Technology
UN	United Nations
UNICRI	United Nations Interregional Crime and Justice Research Institute
VERIFIN	Finnish Institute for Verification of the Chemical Weapons Convention
VGWD	Validation Group Working Database
VR	*S*-[2-(Diethylamino)ethyl] *O*-(2-methylpropyl) *P*-methylphosphonothioate

PLAIN-LANGUAGE SUMMARY

1. Chemical forensics combines advanced chemical analysis with contextual and comparative information to extract chemical information from evidence (i.e., samples) collected during investigations. For the OPCW, chemical forensics goes beyond identifying chemicals covered by the Chemical Weapons Convention (the Convention). It can help determine how a substance was produced, its potential precursors, how it was handled or stored, and whether samples from different locations or incidents are related. Within a structured investigative framework, it strengthens the scientific evidence used to understand alleged chemical weapons incidents.

2. Rapid advances in analytical chemistry, data science, and forensic methods are expanding what chemical analysis can reveal. However, chemical warfare agent-related chemical forensics faces challenges: incidents are rare, reference datasets are limited, and the sensitive nature of the information can restrict data sharing, limiting the data available to develop, validate, and augment forensic methods.

3. Recognising the importance of this field, the Director-General established a Temporary Working Group on Chemical Forensics to review current scientific capabilities and advise on how these approaches could best support the Convention.

4. The Group brought together experts from the Scientific Advisory Board, the Designated Laboratories network, the OPCW Technical Secretariat, and external specialists in forensic science, chemistry, and data analysis. It examined analytical techniques, the potential role of artificial intelligence and machine learning, and the secure management and sharing of forensic data.

5. The Group determined that many necessary skills and capabilities already exist within the network of Designated Laboratories and the broader scientific community. Further development will require collaboration with experts in traditional chemical forensics and related disciplines, adoption of new analytical and data-driven methods, and secure sharing and curation of relevant data. Collaboration and confidence-building exercises will be key to strengthening chemical forensics capabilities.

6. Based on its work, the Temporary Working Group developed a set of recommendations (Table 1) to help the OPCW build and sustain the scientific and technical capabilities needed to support investigations and strengthen implementation of the Convention.

EXECUTIVE SUMMARY

7. Chemical forensics is a highly specialised branch of forensic science that refers to the systematic application of chemical analysis, in combination with powerful data analysis methods, to extract chemical information from evidence (i.e., samples) collected during an investigation. It represents just one element within a broader, multi-source investigative framework. From the standpoint of the OPCW, chemical forensics extends beyond the initial identification of chemicals relevant to the Convention, providing insights into sample provenance, potential precursors, handling or storage, dissemination methods, and potential relationships between samples collected across different locations or incidents. When applied within a robust investigative framework, chemical forensics can contribute to a more comprehensive understanding of alleged chemical weapons use and strengthen the scientific basis underpinning investigative findings.

8. The increasing relevance of chemical forensics has become apparent through recent investigations and rapid advances in analytical chemistry, data science, and forensic methodologies. These developments have expanded the range of questions that can potentially be addressed through chemical analysis, while also introducing new challenges related to data interpretation, quality assurance, and organisational preparedness. The previous Temporary Working Group on Investigative Science and Technology (TWG on IST) identified chemical forensics as a knowledge gap for the OPCW. Consequently, Recommendation 18 of that TWG's end-of-mandate report stated that the Director-General should "consider establishing a new TWG on the provenancing of samples of chemicals relevant to the Convention".

9. This recommendation was implemented through the establishment of the TWG on Chemical Forensics. The TWG was tasked with reviewing the state of the art in science and technology relevant to chemicals, identifying gaps and challenges to ensure that the OPCW can make best use of the information chemical forensics can yield, and advising on measures required to ensure that chemical forensics capabilities can be accessed and applied effectively in support of the Convention. In undertaking this work, the TWG remained cognisant of the findings and recommendations of the previous TWG on IST, which helped frame its deliberations.

10. The work of the TWG on Chemical Forensics was informed by a broad range of expertise, including members of the Scientific Advisory Board (SAB), representatives of the network of Designated Laboratories (DLs), staff from the Technical Secretariat (the Secretariat), and invited external experts drawn from forensic science, analytical chemistry, artificial intelligence (AI) and machine learning (ML), and related disciplines. In its deliberations, the Group considered experience gained from past

and ongoing investigations, existing verification practices, and the practical constraints encountered during contingency operations in the Syrian Arab Republic.

11. The TWG on Chemical Forensics comprised four subgroups, each focusing on specific questions posed by the Director-General, as outlined in the terms of reference.

12. Subgroup 1 examined the state of the art with respect to analysis, including analytical techniques and methodologies. To guide these discussions, a set of foundational chemical forensic questions was proposed that would be applicable to investigations of alleged chemical weapons incidents. Consideration was given to methods enabling batch matching, synthetic route determination, precursor-product matching, chemical profiling to identify key chemical impurities, isotope and trace element analysis, and the assessment of environmental and storage impacts on chemical profiles. A further consideration was analytical methods that would allow for an understanding on how an agent was disseminated. Recent advances in chemical warfare agent-based (CWA-based) chemical forensics were also discussed, highlighting the research undertaken by four recent doctoral graduates, and the Icarus exercise coordinated through the Chemical Forensics International Technical Working Group (CFITWG).

13. Subgroup 2 focused on future capabilities, emphasising the application of AI/ML and deep learning to chemical forensics. Discussions addressed the analysis of both large and limited datasets and the implications for model validity, strategies for creating and expanding datasets for model training, and the role of targeted research informed by optimised computational models. The potential to adapt AI/ML techniques and strategies developed in metabolomics—a field characterised by large datasets—for CWA-related chemical forensics, where datasets are comparatively small, was also examined. The subgroup additionally considered anticipated developments over the next five to ten years and how advances in analytical science may influence this.

14. Subgroup 3 addressed methods and procedures, with particular emphasis on in-field sampling and the creation and curation of a chemical forensics database. Discussions highlighted overarching themes that resonated across the TWG. Significant attention was given to database architecture, data types and formats, metadata requirements, data anonymisation, quality oversight, and governance. Sampling protocols and methodologies were examined in the context of balancing verification requirements with chemical forensic objectives. The fusion of data from disparate analytical techniques was explored, as was the need for appropriate quality management systems (QMSs) based on relevant International Organization for Standardization (ISO) standards and/or frameworks.

15. Subgroup 4 examined requirements for augmenting the OPCW's capabilities in chemical forensics and identified key areas for consideration. The subgroup explored

how forensic science can be effectively integrated into investigations of alleged chemical weapons use. A range of realistic scenarios were considered—from small-scale terrorist attacks to mass-casualty events, lone-wolf incidents, and hybrid threats—illustrating the diverse challenges in applying chemical forensic methods. Discussions considered the reporting requirements for analytical results in chemical forensic investigations and leveraging expertise beyond the DL network to augment existing DL capabilities and support OPCW chemical forensic investigations. The importance of training—including for first responders who may be first to encounter a chemical event—was highlighted, and sampling protocols, situational awareness, and communication among all relevant parties, including the OPCW and laboratories undertaking the chemical forensic analysis, were considered in detail. The subgroup also explored approaches to upskilling the DL network in chemical forensic analysis.

16. During discussions, a number of common themes emerged across subgroups. These included the relationship and key differences between verification and chemical forensics, sampling strategies suitable for both purposes, the importance of continuous two-way communication between investigators and DLs, and the growing impact of AI/ML on analytical workflows.

17. The TWG identified that building and sustaining a chemical forensics capability within the OPCW will require systematic engagement with the DL network. Activities promoting capability development—including workshops and both laboratory- and data-focused exercises—should be considered. Opportunities to partner with external organisations and initiatives, such as the CFITWG, were noted. Approaches to sharing potentially sensitive data to support DL knowledge development and database population, should also be considered.

18. As a result of its deliberations, the TWG formulated 19 recommendations. These were derived from both individual subgroup and plenary TWG discussions, and broad consensus was reached across the Group. Consequently, many recommendations are referenced in multiple sections of the report. A summary of the TWG is provided in Figure 1.

OPCW Scientific Advisory Board's Temporary Working Group on Chemical Forensics

1 Establishment of the TWG on Chemical Forensics

- Examined the state of the art in chemical forensics
- Assessed the impact of AI
- Determined how the OPCW can improve its capabilites

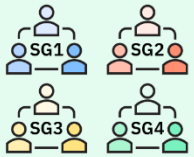
2 Composition of the Group

- 17 members (SAB, Designated Laboratories, government, academia)
- Chair: Dr Anne Bossée
- Vice-Chair: Dr Simon Ovenden

3 Timeline

- 2-year mandate (January 2024 to December 2025)
- 5 plenary meetings
- 1 final report drafting meeting
- Virtual subgroup meetings

4 Subgroups and Approach



SG1: State of the art

SG2: Future capabilities

SG3: Methods and procedures

SG4: Augmenting the OPCW's capabilties

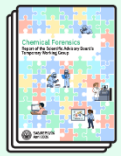
5 External Expertise

- Presentations from 32 invited experts
- Topics included: forensic research, evidence collection, instrumentation, analytical techniques, databases, chemometrics, and AI/ML

6 Drafting and Consensus Building

- Consolidated outputs through iterative drafting and review
- Final text reviewed collectively and agreed by consensus

7 Findings and Recommendations



- Final report (2026)
- 19 recommendations covering: capability and capacity building, procedures and methodology, and chemical forensics data and databases

Key Themes


Data Sharing


Database Building


Sampling Strategies


Reporting Criteria


Quality Management


Collaborative Exercises


Continuous Communication

Figure 1: TWG on Chemical Forensics in summary

RECOMMENDATIONS

19. The full set of recommendations made by the TWG is listed in Table 1, along with summary paragraphs outlining the rationale for their formulation. The TWG has highlighted seven recommendations—shown in green—that it considers of particular importance. The 19 recommendations are also integrated into the findings of the TWG and appear throughout the text. Some recommendations relate directly to a specific subgroup's findings, while others reflect cross-cutting issues arising from multiple subgroups. They are presented in the most appropriate section and may therefore not follow ascending numerical order.

Table 1: Summary of recommendations

Capability and capacity building	
1	**Rationale:** While the DL network possesses baseline skills and knowledge in chemical forensics, the critical role of chemical forensics in OPCW investigations highlights the need to strengthen and broaden these capabilities. In particular, specialised expertise and techniques available in laboratories outside the network, working in fields beyond CWAs, represent a valuable resource. Targeted initiatives—such as topical workshops—can enhance expertise in key areas of chemical forensic analysis, including sample matching and route sourcing, thereby ensuring DLs are fully prepared to support OPCW investigations. **Recommendation: The Secretariat should promote capability building among the network of DLs in performing chemical forensic analyses of Convention-relevant chemicals. This should include engagement with specialised laboratories that are currently outside the network. The organisation of topical workshops in partnership with DLs should also be considered.**

2

Rationale: Confidence-building exercises (CBEs) could play an important role in further developing the technical skills and capabilities of DLs in chemical forensics. Examples include the two CFITWG exercises—Icarus, a laboratory-based ("wet") exercise and the data-focused ("dry") exercise coordinated by the OPCW. Furthermore, the OPCW Proficiency Test (PT) Programme could provide a valuable opportunity to assess and strengthen DL capabilities in this area.

While a previous attempt of including a supplementary trace sample in PTs for optional analysis had limited success due to poor participation, there remains value in exploring alternative approaches to increase participation and maximise effectiveness. Such approaches could include providing a defined and relevant scenario, specific chemical forensic questions, and sufficient time following completion of the PT to support analysis and reporting. In addition, the OPCW could consider how laboratories that successfully analyse chemical forensics samples and demonstrate proficiency could be recognised.

Recommendation: The OPCW should promote, organise, and leverage current and future CBEs to strengthen the skill set of the DL network. As an initial step, this could include the provision of supplementary samples with defined chemical forensic questions into the existing PT Programme.

3

Rationale: In the event of a chemical weapons incident, it is likely that the affected State Party's national response forces will be first at the scene. To preserve the integrity of forensic evidence and avoid disturbance of the scene, a coordinated response and adequate situational awareness are required. It is therefore essential that forensic requirements are integrated into recommended and/or standard operating procedures (R/SOPs), ensuring that sample collection and scene management support robust chemical forensic investigations.

Recommendation: The OPCW should promote, through its capacity-building activities, forensics-based approaches to investigations of alleged use of chemical weapons and the continued coordination of national and regional chemical forensics capabilities. This may include forensic awareness training for first responders, crime scene investigators acting in a potentially contaminated environment, and laboratory personnel. Consideration should also be given to best practices and procedures for on-site sampling.

4

Rationale: To strengthen the OPCW's chemical forensics capability, the expertise and resources of other international organisations engaged in research and capacity building on forensics should be leveraged. In addition, accessing forensic science networks and drawing on knowledge from other fields that use similar techniques is considered essential.

Recommendation: **The OPCW is encouraged to leverage relationships with international organisations, such as INTERPOL and UNICRI, as well as initiatives like the CFITWG, to incorporate forensic-based approaches into CWA investigations. A capabilities gap evaluation of the network of DLs should also be conducted to complement the expertise that already exists in dedicated forensic laboratories.**

Procedures and methodology

5 **Rationale:** In forensic investigations, a chemical sample can contain a broad spectrum of signatures, meaning that a wide range of analytical instrumentation and methodologies—including data science—should be considered. High-resolution mass spectrometry, already widely available in most DLs, is a powerful technique employed for chemical forensics. However, complementary methods can provide additional information on inorganic elements, organic chemicals, and isotope ratios, strengthening the overall evidential picture. The TWG recognises that expertise in these complementary techniques—including isotope ratio mass spectrometry, position-specific isotope analysis, inductively coupled plasma-mass spectrometry, explosive residue analysis, polymer analysis, data science, AI, ML, and chemometrics—may exist in laboratories beyond the current DL network.

Recommendation: **The OPCW should continue to monitor and evaluate emerging analytical methods relevant to chemical forensics and engage with expert laboratories both within and beyond the DL network that possess specialised capabilities. DLs should be encouraged to establish their own collaborations with external expert laboratories to further strengthen their chemical forensics capabilities.**

6 **Rationale:** The recent developments in AI and ML have led to their successful application in various scientific fields, including omics and forensic sciences, to analyse complex chemical datasets, classify samples, and uncover patterns in limited and/or heterogeneous data. Chemical forensics of CWAs faces similar analytical challenges, making these approaches highly relevant for enhancing investigative capabilities and supporting robust conclusions.

Recommendation: **The Secretariat, in conjunction with DLs, should evaluate current and emerging AI/ML tools and software from relevant scientific and forensic fields that could be applied to chemical forensics. Compliance with national and OPCW security policies should be considered as part of this evaluation. The OPCW Laboratory and DLs should also be encouraged to share and evaluate preferred tools and software in CBEs.**

7

Rationale: Chemical forensics is well established and research is widely published in traditional forensic fields—such as fire debris analysis, illicit drug investigations, and verifying the authenticity of artworks or food. By contrast, publications on chemical forensics applied to CWAs are scarce. While some DLs are unable to publish due to security or policy constraints, sharing methods and workflows for CWA analyses—particularly through publication in peer-reviewed journals—is essential to advance the field and strengthen OPCW capabilities. Other forums for sharing methods and workflows, such as topical workshops and conference sessions, should also be leveraged.

Recommendation: The Secretariat should encourage DLs to further develop, publish, and share chemical forensic methods and workflows for Convention-relevant chemicals, establishing a strong scientific foundation for their application. Designated Laboratories are urged to contribute to updates of the recommended operating procedures in the Blue Book and publication of relevant research in peer-reviewed journals is strongly encouraged.

8

Rationale: The nature of chemical forensic investigations is dynamic, with the forensic questions often evolving as more data become available. In addition, all CWA incidents requiring chemical forensics begin with identification and verification activities. To ensure that analyses are focused on relevant forensic questions and minimise redundant efforts—such as re-identifying chemicals previously confirmed during verification analysis—clear and continuous communication between investigators and the DLs conducting the analysis is essential.

Recommendation: When chemical forensic approaches are applied to authentic samples, the Secretariat should maintain an open and continuous line of communication with the DLs involved. This will ensure clarity on the forensic questions and the application of the most appropriate techniques, approaches, and methods to generate the most relevant information for the investigation.

9

Rationale: It is inevitable that access to the site of an alleged chemical weapons incident will be limited. An inspection may require significant preplanning and preparation, and sampling activities will likely be conducted under tight timelines. Furthermore, since a chemical forensic investigation may follow the initial identification and verification activities, it is imperative that any samples collected meet the requirements for all eventual investigative purposes. This highlights the need for R/SOPs and working instructions on sample collection to explicitly account for chemical forensics requirements.

Recommendation: **The Secretariat should ensure that chemical forensic requirements are incorporated into R/SOPs and working instructions for on-site sample collection. In doing so, relevant experience from the contingency operations in the Syrian Arab Republic should be leveraged.**

10

Rationale: Confidence in analytical results from an international investigation depends not only on laboratory capability and validated methods, but also on the robustness and defensibility of the quality management framework under which the work is conducted. While laboratories in the DL network are accredited to ISO/International Electrotechnical Commission (IEC) 17025:2017—which ensures technical competence in calibration and the analysis of suspected CWA samples for verification purposes—chemical forensic investigations may require additional considerations specific to forensic science, such as interpretation and reporting.

The guidance provided in the ISO 21043 series and by other relevant bodies—including the ANSI National Accreditation Board, the European Network of Forensic Science Institutes, the International Laboratory Accreditation Cooperation, and the International Forensic Strategic Alliance—offers a framework for addressing these forensic-specific requirements and strengthening the overall reliability and credibility of investigative findings. Notably, ISO 21043 provides guidance rather than constituting an accreditation standard. As such, national accreditation bodies do not currently assess the competency of laboratories against this standard. When undertaking a CWA-based chemical forensic investigation, the Secretariat—in conjunction with the partnering DLs—should define the specific forensic science requirements.

Recommendation: **Laboratories involved in an international investigation should operate under an overarching QMS that ensures the implementation of regular quality control measures (including personnel competence, equipment calibration, lot (batch) documentation, appropriate validation and documentation of methods, and regular error analysis) such as those outlined in ISO/IEC 17025:2017. This system may be supplemented by forensic-specific requirements, for example through guidance provided in ISO 21043 and other relevant forensic science guidelines. Procedures applied in an international investigation must be technically robust and reproducible, as demonstrated through internal validation, international analytical forensic exercises, and/or peer-reviewed publications. Accreditation of a specific method should not be regarded as an absolute requirement for its use in a forensic investigation.**

11

Rationale: While well-established criteria exist for reporting verification analyses—where results can support unambiguous identification—this is not the case for chemical forensics due to the nature of investigations. Verification typically provides categorical outcomes (for example, the presence or absence of a CWA). In contrast, chemical forensic analyses integrate multiple techniques and statistical evaluation, with results requiring interpretation rather than yielding definitive identification. Forensic investigations therefore emphasise interpretation and the expression of confidence rather than categorical determination, with conclusions reported in terms of uncertainty and probabilistic weight. Clear, proportionate, and harmonised criteria are therefore needed to guide the reporting of forensic findings, including how validity, uncertainty, and confidence are communicated.

Recommendation: The OPCW Laboratory, in partnership with DLs, should define criteria for the reporting of results of chemical forensic analyses. A working group would be one way to support the development of these criteria. Confidence-building exercises could also be used to further develop and refine them. The criteria should include the methods used, key chemical forensic data, and the associated level of confidence.

Chemical forensics data and databases

12

Rationale: While conducting verification work, DLs use the OPCW Central Analytical Database (OCAD) to support chemical identification. Although OCAD is well suited to verification activities, it is unlikely to provide sufficient support for chemical forensic applications, which have distinct requirements. A dedicated chemical forensics database should therefore be developed, incorporating spectral data and additional information on relevant chemicals—including substances not directly related to CWAs—along with chemical profiling data. Each spectral entry should be augmented with additional information, such as relevant literature and any information that links a substance to a CWA or associated processes. This resource would assist in addressing key forensic questions such as route sourcing, sample/batch matching, precursor-product matching, and could extend to identifying storage container types and dissemination methods.

Recommendation: The OPCW should establish an overarching chemical forensics database platform for CWA-related investigations. The platform would comprise two linked modules: a spectral databases module and a knowledge hub module containing synthetic route data, a list of key chemical impurities, and relevant metadata. The spectral databases module would integrate existing spectral resources alongside a newly developed database, either standalone or incorporated into a modified Validation Group Working Database.

13 **Rationale:** The establishment of a robust chemical forensics database for CWAs requires access to high-quality ground-truth data, much of which resides within the DL network. Sharing information on the chemical forensics of scheduled chemicals from States Parties could greatly enhance both the population and authenticity of the database. While sensitivities around data sharing—such as identifying originating laboratories or revealing synthesis routes—must be carefully managed, leveraging approaches such as CBEs can enable effective database development.

Recommendation: It is imperative that the OPCW obtain chemical forensics-related data to populate the proposed chemical forensics database platform and support ongoing and future investigatory efforts. This can be achieved via multiple, non-competing approaches, including:

a. **encouraging States Parties to share chemical forensics data on scheduled chemicals; and**
b. **making use of CBEs focused on chemical forensics**.

14 **Rationale:** During TWG discussions on sharing data from States Parties, it was acknowledged many are hesitant to provide information related to Schedule 1 chemicals. Sharing sensitive reference data, particularly on synthesis routes, could pose a proliferation risk. To address this, a procedure and clear criteria for anonymising and protecting such data are needed to enable its use while mitigating security concerns.

Recommendation: The OPCW, in collaboration with DLs, is encouraged to define a common framework—including criteria and procedures—to anonymise data relevant to chemical forensics, reducing security risks while supporting database development.

15 **Rationale:** The Validation Group plays an essential role in curating spectral data for inclusion in the OCAD, ensuring that only technically validated, high-quality data are incorporated. A similar approach would be beneficial for the proposed overarching chemical forensics database platform. An equivalent group could therefore be established to oversee the inclusion of data in both the spectral databases and knowledge hub modules, defining robust criteria for data submission, validation, and management. Such a group could also contribute to addressing data sensitivity concerns, including the anonymisation of shared data where appropriate.

Recommendation: The OPCW should establish a validation group responsible for defining a framework for data management and setting criteria for the inclusion and curation of data within the proposed chemical forensics spectral database and knowledge hub. This group could also facilitate communication between DLs during the initial construction phase of the database, promoting contributions from these laboratories.

16 **Rationale:** Chemical forensic datasets for CWA-related analyses are inherently limited due to the rarity of incidents and the sensitivity of the associated data. As a result, obtaining sufficient entries for chemical forensics databases is challenging. One potential approach to help address these gaps is the use of *in silico* data generated through AI/ML methods to augment existing datasets. As demonstrated in resources such as the Human Metabolome Database, predicted entries could be incorporated alongside experimental data and clearly identified as such.

Recommendation: Designated Laboratories and the OPCW Laboratory are encouraged to investigate the feasibility, reliability, and relevance of expanding chemical forensic datasets using *in silico* data generated by AI/ML methods, including physicochemical properties, spectral data, impurity datasets, and potential synthesis (by-)products. Where predicted data are incorporated into chemical forensics databases, they should be clearly identified as *in silico* entries and efforts should be taken to validate them.

17 **Rationale:** The increasing role of data science—including AI/ML, chemometrics, and robust data curation and management—in chemical forensics highlights the need for dedicated expertise to support the development and maintenance of resources such as the proposed overarching chemical forensics database platform.

Recommendation: The OPCW is encouraged to recruit a data scientist with expertise in ML, chemometrics, and scientific data management to support the integration of data-driven approaches into the Organisation's activities, including the development, curation, and governance of the proposed overarching chemical forensics database platform and its modules.

18 **Rationale:** There is a knowledge gap within the DL network regarding the application of AI/ML in chemical forensics. Tools and databases developed through initiatives such as the OPCW AI Research Challenge provide a practical foundation to help laboratories apply these methods effectively and address challenges related to limited datasets, data fusion, and predictive analysis.

Recommendation: The Secretariat should share AI/ML tools and reference databases developed through the OPCW AI Research Challenge to support DLs in applying these technologies to chemical forensics.

19 **Rationale:** While the importance of chemical forensics for DLs is clear, these laboratories face competing priorities and the multi-disciplinary nature of the work makes collaboration across multiple laboratories essential. Funding for such collaborative programmes is critical, yet identifying sources is challenging due to the geographical dispersion of DLs and external expertise. The OPCW should take a leading role in identifying funding streams to support these initiatives.

Recommendation: The OPCW should help identify funding streams to support interdisciplinary chemical forensic research and development, enabling DLs to collaborate on common syntheses, analytical work, and data science to drive knowledge acquisition, data generation, and the verification or enhancement of forensic results.

INTRODUCTION

20. This introductory section provides a brief overview of the concepts and processes relevant to chemical forensics. The investigative process provides a structured framework for examining incidents, collecting and analysing evidence, and drawing scientifically supported conclusions. Defining core concepts such as chemical forensics, chemical profiling, and chemometrics, establishes a common understanding and ensures consistent use of terminology throughout the report. Key technical terms used in the report are defined in the Glossary (Annex 1).

21. **Forensic science** is the application of scientific methods and techniques to examine evidence for investigative purposes, typically in legal or compliance contexts. These methods and techniques are drawn from diverse scientific fields, such as biology, physics, and chemistry. Common types of traditional forensic evidence include fingerprints, traces of deoxyribonucleic acid (DNA), and ballistics. Their analysis—which may involve chemical techniques—can reveal distinctive signatures that form a vital part of a forensic investigation. A brief introduction to forensic science is provided in Annex 2.

22. The **investigative process** is a structured and methodical framework used to determine what occurred, how it occurred, and who may be responsible. It begins with the initial police response and scene security, followed by systematic crime scene investigation involving documentation, assessment, and the identification, recovery, and preservation of evidence in accordance with forensic protocols. In parallel, investigators conduct interviews with witnesses and relevant experts to obtain contextual, testimonial, and specialist information that can guide investigative direction and support evidential interpretation. Recovered material is subjected to laboratory analysis, where screening, confirmatory testing, and interpretation are conducted using validated scientific methods. Findings are evaluated within the case context to establish evidential value, limitations, and potential associations. The process is iterative and intelligence led. Laboratory results and interview findings inform ongoing investigations by generating leads, supporting the identification of persons of interest, and enabling the linkage of individuals, objects, and locations. This interaction between forensic scientists and investigators is central to developing a coherent evidential framework. Evidence is then integrated to reconstruct events and support the prosecution case. Forensic findings are formally reported and ultimately presented in court, where expert evidence assists judicial decision-making. This end-to-end forensic investigative process is depicted in Figure 2.

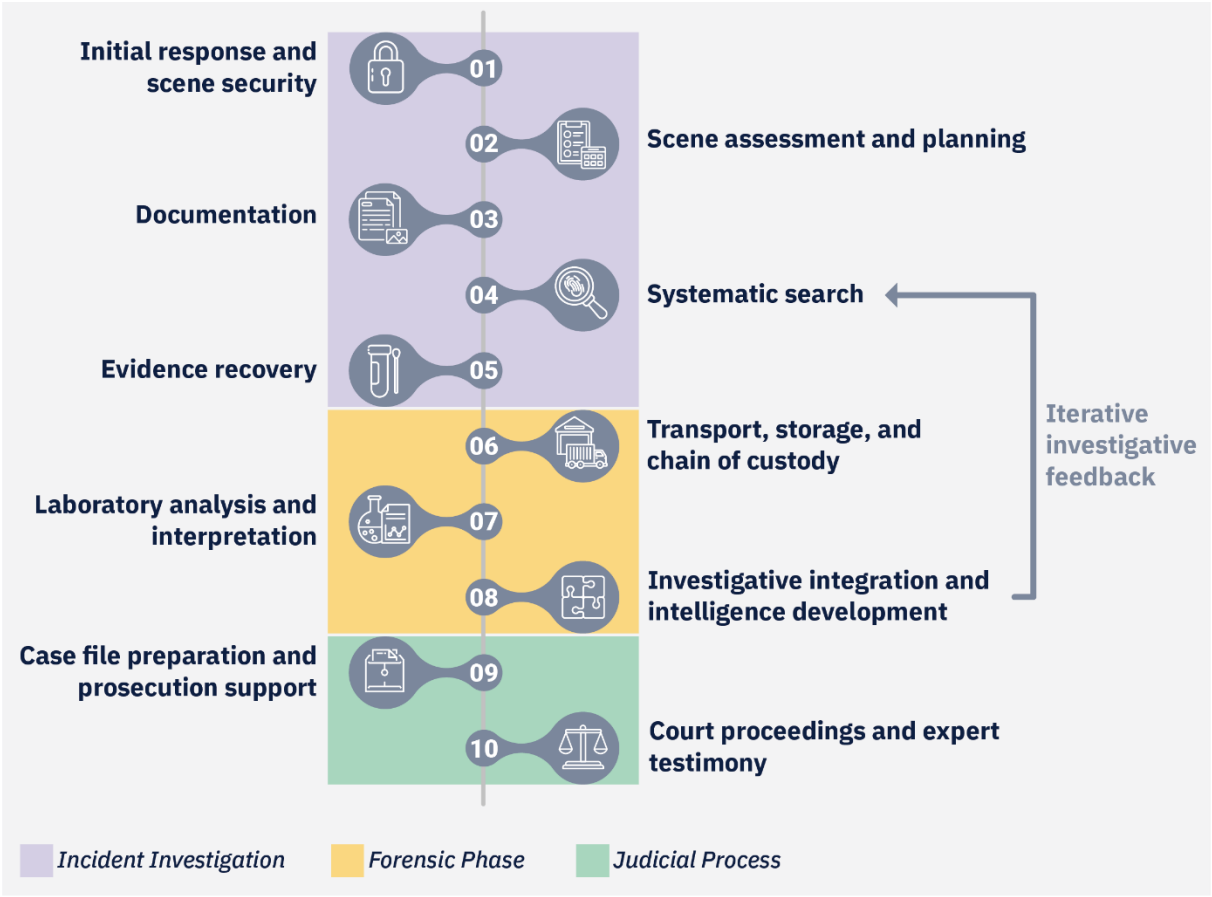


Figure 2: End-to-end forensic investigative process

23. The investigative process is important because it protects the integrity and reliability of both the evidence and the final conclusions. A systematic investigation ensures that evidence is properly handled, preserved, and analysed, which is essential for producing results that are scientifically sound and legally admissible. It also helps investigators avoid bias and maintain transparency in how they reach their conclusions. Ultimately, the investigative process provides a defensible framework that supports an accurate reconstruction of events and helps decision-makers—such as courts—trust the findings.

24. Within this framework, there are overarching **forensic questions** that play a central role by guiding the direction and focus of the investigation. These are the key questions investigators seek to answer, such as *What happened? How did it happen? Where did it occur? When did it take place? Who was involved?* and, in some cases, *Why did it happen?* From an OPCW standpoint, the forensic questions act as the roadmap of the investigation of a chemical event. It determines what evidence needs to be collected, what analytical methods should be used, and how the findings will ultimately be interpreted. By structuring the investigation around these fundamental questions, investigators ensure that their work remains purposeful, thorough, and aligned with the goals of uncovering the truth. Ultimately, aspects of these forensic questions can be broken down into specific **chemical forensic questions**.

25. **Chemical forensics** is a specialised branch of forensic science. It may be defined as the application of analytical chemistry, synthetic chemistry, and chemometric methods and techniques to examine chemical evidence for investigative purposes. Chemical forensics focuses on the chemical aspects of an investigation. It identifies the primary substance (for example, a drug of abuse) in a sample and provides additional chemical data that may be used to draw conclusions regarding the origin or production route, and where possible the identity of additional constituents in the sample, including solvents and plasticisers.

26. In an OPCW context, identification and chemical forensics are two distinct processes. Identification of the primary substance is the process of confirming the identity of a chemical warfare agent (CWA) in a sample, and is classed as a verification activity. This activity is performed in accordance with established OPCW verification processes. To support the OPCW's verification regime and its capacity to investigate allegations of chemical weapons use, two Designated Laboratories (DLs) in the network perform off-site analysis of samples to determine the presence of, and unambiguously identify, any chemicals relevant to the Chemical Weapons Convention (the Convention). This process, known as **chemical verification**, confirms a chemical's identity by either comparing it with an authentic standard or matching it to a chemical in a spectral database such as the OPCW Central Analytical Database (OCAD).

27. Chemical forensics extends this verification process by analysing additional characteristics such as by-products, degradation products, other impurities—including contaminants and stabilisers, some of which may not be directly related to substances covered by the Convention—and unreacted starting materials found in the sample. Critical chemical information can then be extracted from the collected analytical data using chemometrics and other appropriate data analysis methodologies. Chemical forensics focuses exclusively on obtaining additional data on a sample's chemical composition and how this relates to the previously identified CWA. The additional information that chemical forensics of CWAs provides could enable chemical forensic questions—such as *How sophisticated was the synthetic production method? Is the dispersion method compatible with the use of an explosive charge? Was the product stored in a metal container?*—to be answered. This information may provide insight into how the sample was stored or could support sample or batch comparisons. Stages two to eight of the end-to-end forensic investigative process in Figure 2 are relevant to chemical forensics. A more detailed breakdown of the activities involved in chemical forensic investigations of CWAs, together with the associated challenges and considerations, is provided in Figure 3.

28. A key component of chemical forensics is **chemical profiling**. This refers to the analysis and characterisation of chemical samples to detect key chemical impurities and/or isotope ratio information, commonly referred to in the scientific literature as

“chemical attribution signatures”. No chemical is ever 100% pure, and a sample of a particular target chemical will always contain small amounts of other substances, originating from its production or environmental conditions. These **impurities** may include unreacted starting materials, by-products, degradation products, solvent residues, stabilisers, additives, and environmental contaminants. They serve as **extrinsic signatures** because they come from external sources, i.e., substances that are not the target chemical itself. Signatures—such as stable isotope ratios—may also be **intrinsic**. These are features that come from the target chemical itself and are inherent to its molecular structure or composition. In a chemical forensic investigation, these extrinsic and intrinsic signatures may provide insights into how, where, and when a chemical was produced and/or stored and/or dispersed. The interpretation of results from chemical profiling often relies on **chemometrics**, which is the application of multivariate statistical analysis to analytical data to aid in extracting relevant chemical information.

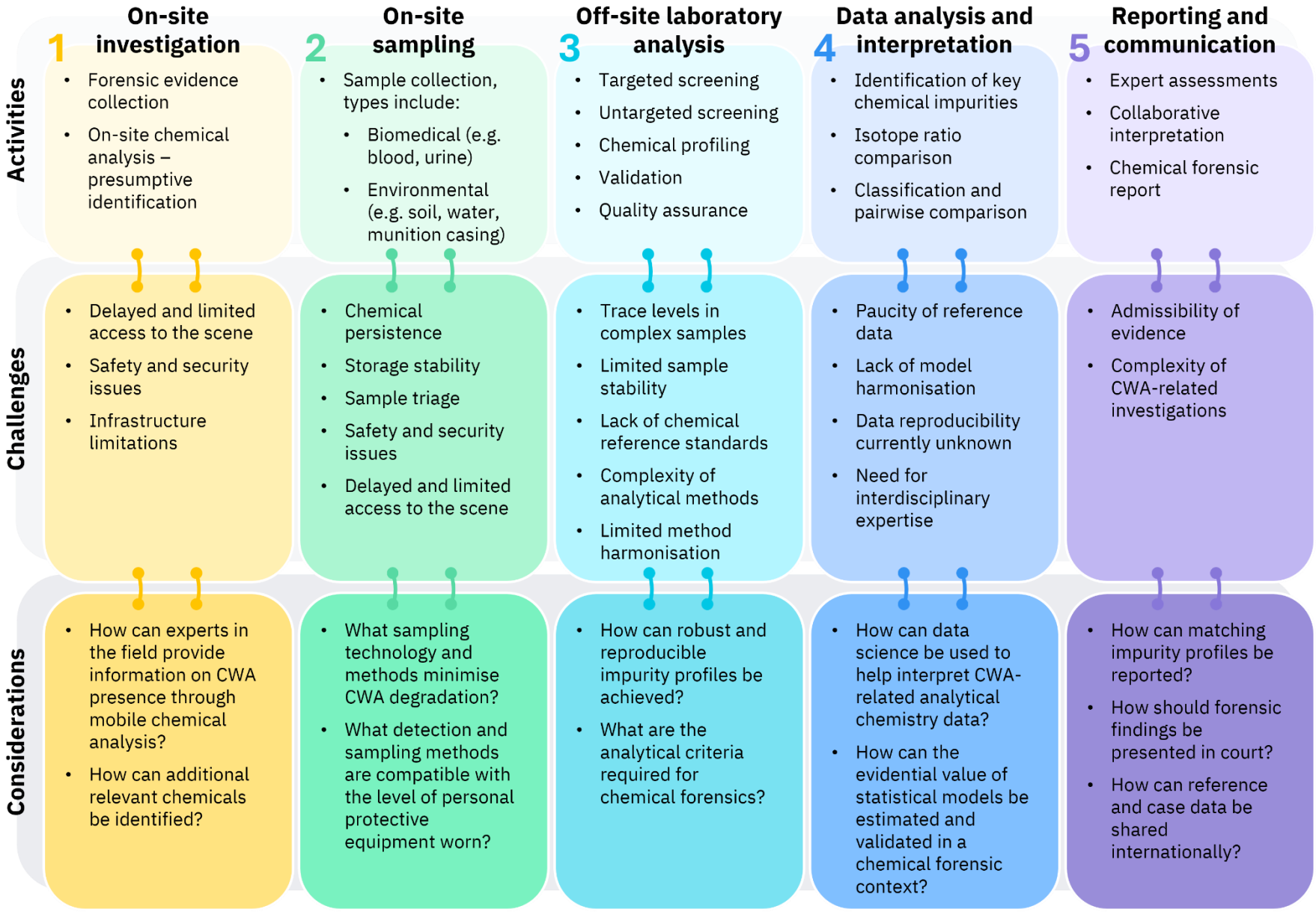


Figure 3: Key activities, challenges, and considerations relating to chemical forensic investigations of CWAs

Background

29. Chemical forensics is an established field applied across many investigative domains, and it plays a particularly critical role in combating illicit drug production and trafficking.[1,2,3,4] In this context, it enables investigators to determine the origin and manufacturing methods of substances such as cocaine or heroin by identifying chemical signatures inherited from natural sources—such as minerals in irrigation water or altitude-related growth factors—and by analysing impurities like solvent residues, intermediates, and by-products.[5] These insights can link seizures to specific production sites, track trafficking routes, and identify supply chains. Chemical forensics also supports the pharmaceutical sector through the identification of counterfeit medicines.[6] In the art world, it can be leveraged to determine the age and geographic origin of materials, supporting the authentication of artworks,[7] while in the food industry, it verifies the authenticity and provenance of products such as olive oil.[8] Similar approaches are also applied in the cosmetics sector to identify

---

1 Lurie, Ira S., Arthur L. Berrier, John F. Casale, Reiko Iio, and Joseph S. Bozenko. "Profiling of Illicit Fentanyl Using UHPLC–MS/MS." *Forensic Science International* 220, no. 1–3 (July 2012): 191–96. https://doi.org/10.1016/j.forsciint.2012.02.024.

2 Dams, Riet, Tom Benijts, Willy Lambert, Désiré Luc Massart, and Andreas De Leenheer. "Heroin Impurity Profiling: Trends throughout a Decade of Experimenting." *Forensic Science International* 123, no. 2-3 (December 2001): 81-88. https://doi.org/10.1016/S0379-0738(01)00541-2.

3 Andersson, Kjell, Eric Lock, Kaisa Jalava, Henk Huizer, Sten Jonson, Elisabet Kaa, Alvaro Lopes, et al. "Development of a Harmonised Method for the Profiling of Amphetamines VI." *Forensic Science International* 169, no. 1 (June 2007): 86–99. https://doi.org/10.1016/j.forsciint.2006.10.020.

4 Nielsen, Louise Stride, Palle Villesen, and Christian Lindholst. "Stability of Cocaine Impurity Profiles during 12 Months of Storage." *Forensic Science International* 264 (July 2016): 56 - 62. https://doi.org/10.1016/j.forsciint.2016.03.012.

5 Collins, Michael, Juuso Huttunen, Ian Evans, and James Robertson. "Illicit Drug Profiling: The Australian Experience." *Australian Journal of Forensic Sciences* 39, no. 1 (June 2007): 25 - 32. https://doi.org/10.1080/00450610701324924.

6 Maira Kerpel dos Santos, Nayara Araujo dos Santos, Joao Francisco Allochio Filho, Layla Paixao Santos, Wanderson Romao, Rafael Scorsatto Ortiz. "Paper Spray Ionization Coupled to Fourier Transform Ion Cyclotron Resonance Mass Spectrometry as a Tool to Fight the Counterfeiting of Medicines." *International Journal of Mass Spectrometry* 468 (June 2021): 116649. https://doi.org/10.1016/j.ijms.2021.116649.

7 Geddes da Filicaia, Eugenia, Richard P. Evershed, and David A. Peggie. "Review of Recent Advances on the Use of Mass Spectrometry Techniques for the Study of Organic Materials in Painted Artworks." *Analytica Chimica Acta* 1246 (March 2023): 340575. https://doi.org/10.1016/j.aca.2022.340575.

8 Saadat, Saeida, Hardi Pandya, Aayush Dey, and Deepak Rawtani. "Food Forensics: Techniques for Authenticity Determination of Food Products." *Forensic Science International* 333 (April 2022): 111243. https://doi.org/10.1016/j.forsciint.2022.111243.

counterfeit products and detect hazardous ingredients.[9] These diverse, well-established applications demonstrate the maturity and reliability of chemical forensic methods.

30. Over the past decade, the OPCW has recognised the importance and utility of chemical forensics approaches with respect to CWAs and has developed its capabilities in this regard. The Scientific Advisory Board's (SAB) Workshop on Chemical Forensics, which took place in 2016, was instrumental in this process.[10] This workshop explored the capabilities of the field and potential applications to support implementation of the Convention. Presentations on OPCW contingency operations in the Syrian Arab Republic noted that chemical forensic techniques were already being employed in these activities—albeit in an elementary capacity—and would continue to be required to meet their objectives, as well as those of similar operations in the future.

31. The Syrian Arab Republic acceded the Convention in September 2013, became a State Party to the treaty on 14 October 2013, and submitted its initial declaration on 23 October 2013. Over the months that followed, Syria submitted a series of amendments to its initial declarations. Despite these amendments, questions remained about the accuracy and completeness of its declaration, and this led to the formation of additional contingency capabilities to ensure Syria's compliance with the Convention. In April 2014, the Declaration Assessment Team (DAT)[11] was established to verify the accuracy and completeness of the Syrian declarations. The same month, the Fact-Finding Mission (FFM)[12] was established to determine whether toxic chemicals had been used as weapons in Syria. In 2018, the Investigation and Identification Team (IIT)[13] was established to identify the perpetrators of specific instances of chemical weapons use in the Syrian Arab Republic.[14]

32. To fulfil its mandate, the IIT is required to investigate with the aim of establishing the facts regarding the perpetrators of the use of chemical weapons and, when feasible,

---

9 Teodoro, Janaina A., Hugo V. Pereira, D. N. Correia, Marcelo M. Sena, Evandro Piccin, and Rodinei Augusti. "Forensic Discrimination Between Authentic and Counterfeit Perfumes Using Paper Spray Mass Spectrometry and Multivariate Supervised Classification." *Analytical Methods* 9, no. 34 (May 29, 2017): 4979–87. https://doi.org/10.1039/c7ay01295k.

10 Report of the Scientific Advisory Board's Workshop in Chemical Forensics (SAB-24/WP.1, dated 14 July 2016).

11 "Declaration Assessment Team." OPCW, accessed December 15, 2025.

12 "Fact-Finding Mission." OPCW, accessed December 15, 2025.

13 "Investigation and Identification Team (IIT)." OPCW, accessed December 15, 2025.

14 (a) The Office of Special Missions (OSM) was established on 1 June 2025 and is responsible for coordinating all contingency operations. It integrates into a single unit the functioning of the DAT, FFM, and IIT in the Syrian Arab Republic. (b) Opening Statement By The Director-General to the 109th Session of the Executive Council (Paragraph 29 of EC-109/DG.25, dated 8 July 2025).

their identity, in the incidents within its purview. The Team therefore comprises investigators and analysts with relevant qualifications and experience in complex investigations, analysis, and forensics. The Note by the Secretariat EC-91/S/3 (dated 28 June 2019) stressed that the IIT would "ensure the security, integrity, preservation, and chain of custody of the information and material in its possession from the moment of collection or receipt, and analyse and store technical and scientific information meeting the highest technical standards, as well as through the meticulous employment of forensic processes".[15]

33. The reports issued by the IIT clearly demonstrate that it has employed a range of both chemical and traditional forensic techniques to reach its conclusions.[16] In terms of chemical forensics, this has included comparing chemical profiles to determine the production route, comparing chemical signatures of samples to indicate the same source, synthesising CWAs to verify the production route, and various chemical analyses to identify and/or quantify inorganic species and contaminants such as explosives. For a detailed breakdown of the chemical forensics and traditional forensic techniques used by the IIT, see Annex 3 to this report.

34. In light of the mandates of the contingency operations in Syria, the SAB recommended at its Twenty-Fourth Session that a temporary working group (TWG) should be established to conduct an in-depth review of methods and technologies that could be used by the OPCW in investigative work.[17] It noted that capabilities enabled through advances in investigative science and technology would benefit the robustness of information and analysis associated with contingency operations.

---

15 Work of the Investigation and Identification Team Established by Decision C-SS-4/DEC.3 (EC-91/S/3, dated 28 June 2019).

16 (a) First Report by the OPCW Investigation and Identification Team (IIT) Pursuant to Paragraph 10 of Decision C-SS-4/Dec.3 "Addressing the Threat From Chemical Weapons Use" Ltamenah (Syrian Arab Republic) 24, 25, and 30 March 2017 (S/1867/2020, dated 8 April 2020). (b) Second Report by the OPCW Investigation and Identification Team Pursuant to Paragraph 10 of Decision C-SS-4/DEC.3 "Addressing the Threat from Chemical Weapons Use" Saraqib (Syrian Arab Republic) – 4 February 2018 (S/1943/2021, dated 12 April 2021). (c) Third Report by the OPCW Investigation and Identification Team Pursuant to Paragraph 10 of Decision C-SS-4/DEC.3 "Addressing the Threat from Chemical Weapons Use", Douma (Syrian Arab Republic) - 7 April 2018 (S/2125/2023, dated 27 January 2023). (d) Fourth Report by the OPCW Investigation and Identification Team Pursuant to Paragraph 10 of Decision C-SS-4/DEC.3 "Addressing the Threat from Chemical Weapons Use", Marea (Syrian Arab Republic) – 1 September 2015 (S/2255/2024, dated 22 February 2024). The Fifth Report by the Investigation and Identification Team was issued after the end of the mandate of the TWG.

17 See paragraphs 1.2 and 8.12 to 8.17 of the Report of the Scientific Advisory Board at its Twenty Fourth Session (SAB-24/1, dated 28 October 2016).

35. As a result of this recommendation, the Director-General requested the establishment of the TWG on Investigative Science and Technology (IST) in 2017.[18] This TWG dedicated a considerable amount of time to exploring how the provenance of a sample could be determined.[19] It found that determining the provenance of a CWA could be achievable through the analysis of both extrinsic signatures (such as impurities and additives) and intrinsic signatures (such as stable isotope ratios). Furthermore, the TWG on IST identified that having the resources and methodologies to perform sample matching from seemingly disparate incidents of CWA use, and then link the agent to a production facility, would make significant contributions to the forensic investigation of an alleged incident.

36. The TWG on IST also recognised the need for the Secretariat to ensure that sampling at sites of alleged chemical weapons use meets current best forensic practice. This would include clearly articulating forensic issues in recommended and/or standard operating procedures (R/SOPs) and work instructions, ensuring that information collected for investigative purposes is managed separately from that for routine verification activities. The TWG also noted that not all required expertise for forensic investigations resides within the DL network. As a result of these findings, the TWG on IST recommended that the Director-General "Consider establishing a new TWG on the provenancing of samples of chemicals relevant to the Convention" (Recommendation 18).[19] At its Thirty-Seventh Session, the SAB also recommended that the Director-General establish a new TWG that would focus on determining the provenance of chemical samples relevant to the Convention, including methods of chemical forensic profiling.[20] The timeline of key events and outputs leading up to the establishment of the TWG on Chemical Forensics is shown in Figure 4.

---

18 See paragraphs 8 to 9 of Response to the Report of the Twenty-Fourth Session of the Scientific Advisory Board (EC-84/DG.9, dated 18 January 2017).

19 Final Report of the Scientific Advisory Board's Temporary Working Group on Investigative Science and Technology (SAB/REP/1/19, dated December 2019).

20 Report of the Scientific Advisory Board at its Thirty-Seventh Session (SAB-37/1, dated 1 September 2023).

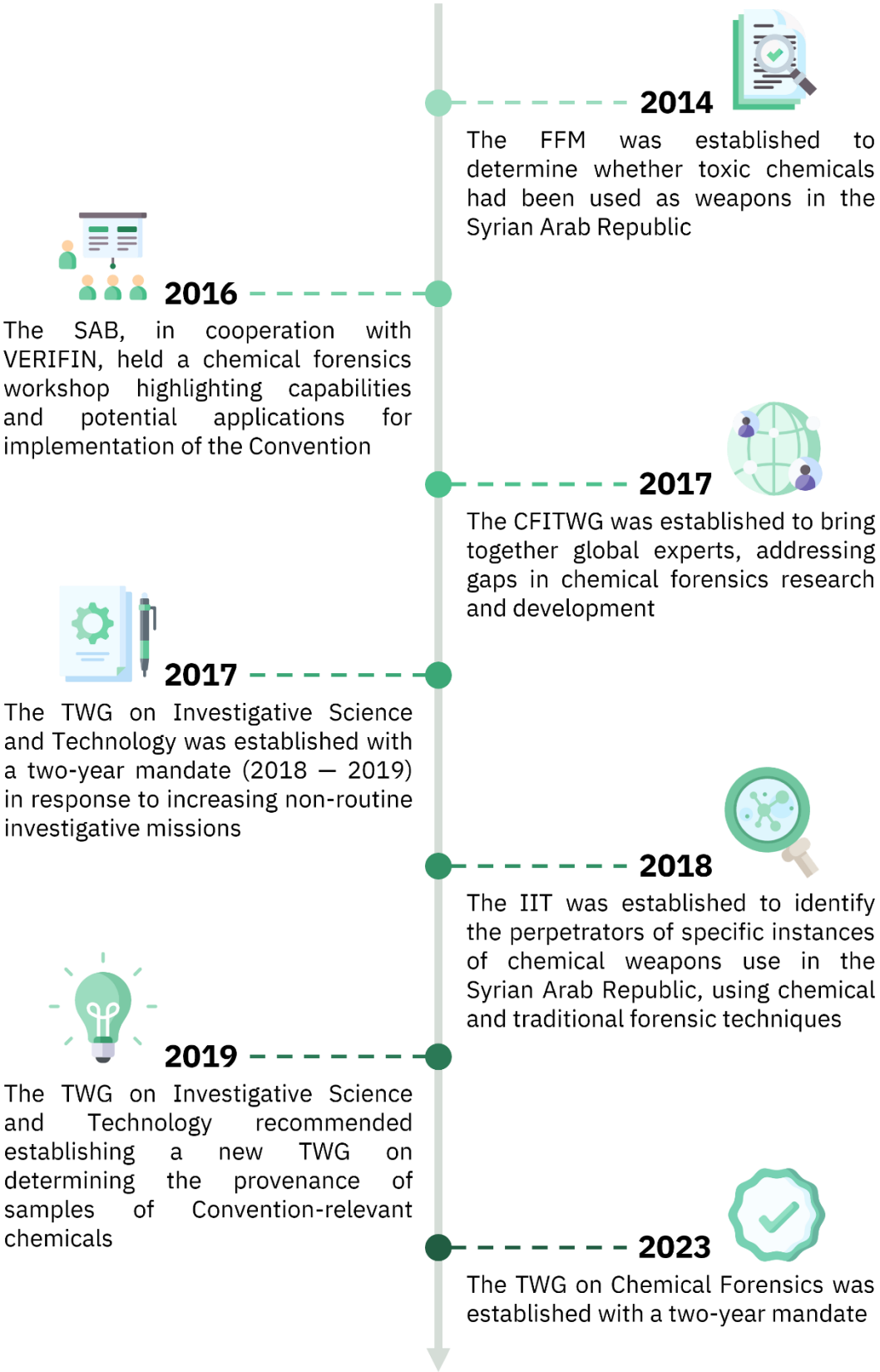


Figure 4: Timeline of key events and outputs leading up to the establishment of the TWG on Chemical Forensics

Establishment of the TWG on Chemical Forensics

37. In his response to the SAB's recommendation, the Director-General decided to establish the TWG on Chemical Forensics.[21] He recognised that the ability to ascertain additional information about a given sample related to alleged use of a chemical weapon has become increasingly relevant. This information can be derived from a sample's unique chemical signatures, obtained through distinct but complementary chemical analyses. The Director-General noted that the types of methods and analyses used to understand these unique chemical signatures are captured in the term "chemical forensics".

38. In determining the terms of reference (TOR) of the TWG on Chemical Forensics, it was noted that the threat of chemical weapons use remains, and the ways in which chemicals may be used as weapons is evolving. While the threat of large-scale attacks by State actors persists, there is growing concern about other scenarios of misuse, including by non-State actors and through more targeted uses by States. The potential for the use of non-traditional agents—such as biotoxins[22,23] and central nervous system-acting chemicals[24]—is also increasing. Given this expanding threat spectrum and the range of possible attack scenarios (including small-scale use—by a non-State actor, a "lone wolf", or a cult— targeted assassinations, and mass casualty events), it is essential to obtain as much information as possible from any samples collected after an alleged attack. This requires the application of a range of chemical forensic methods and techniques. It is therefore critical for the OPCW to fully understand, develop, and apply these approaches, and ensure that validated results

---

21 Response to the Report of the Thirty-Seventh Session of the Scientific Advisory Board (EC-104/DG.22, dated 27 September 2023).

22 Ricin (a) In 1978, Bulgarian dissident Georgi Markov was assassinated using a pneumatic umbrella that injected a metal pellet containing ricin. Crompton, Rufus, and David Gall. "Georgi Markov — Death in a Pellet." *Medico-Legal Journal* 48, no. 2 (June 1980): 51–62. https://doi.org/10.1177/002581728004800203. (b) There have been a number of cases involving the use or attempted use of ricin by non-State actors, notably in the UK (2002 and 2003), USA (1991, 1995, 2003, 2004, 2011, and 2013) and Germany (2018). Moshiri, Mohammad, Fatemeh Hamid, and Leila Etemad. "Ricin Toxicity: Clinical and Molecular Aspects." *Reports of Biochemistry & Molecular Biology* 4, no. 2 (April 2016): 60-65.

23 Trichothecenes: "Yellow rain" incidents in Laos, Cambodia, and Afghanistan (1974 – 1981). During the Vietnam War (1975 – 1981), the Soviet Union was alleged to have provided mycotoxins to the armies of Vietnam and Laos for use against resistance forces in Laos and Cambodia. During the Iran-Iraq War (1985 – 1989), there were reports of mycotoxin shipments (in powder and smoke form) to Iraq. Wannemacher, Robert W., and Stanley L. Wiener. "Trichothecene Mycotoxins." *Medical Aspects of Chemical and Biological Warfare*, 6, (1997): 655–76.

24 Riches, James R., Robert W. Read, Robin M. Black, Nicholas J. Cooper, and Christopher M. Timperley. "Analysis of Clothing and Urine from Moscow Theatre Siege Casualties Reveals Carfentanil and Remifentanil Use." *Journal of Analytical Toxicology* 36, no. 9 (September 20, 2012): 647–56. https://doi.org/10.1093/jat/bks078.

can be used effectively in any investigation into the misuse of chemicals. The development of the TOR also took into account recommendations made by the TWG on IST regarding this new TWG on Chemical Forensics.

39. The objective of the TWG on Chemical Forensics was to review the science and technology relevant to chemical forensics and identify remaining gaps and challenges such that the OPCW can make best use of the information chemical forensics can yield moving forward. The TOR required that the work and recommendations from the previous TWG on IST be considered, as well as ongoing work in this area within the Secretariat. Annex 4 contains the complete TOR of the TWG.

40. The TWG on Chemical Forensics comprised 17 members, bringing together SAB members, representatives of DLs, and other experts in chemical forensics and profiling—as recommended by the TWG on IST. The TWG members were from 15 States Parties and one international organisation. Dr Anne Bossée of the SAB was appointed as the Chairperson of the TWG on Chemical Forensics and was supported by Dr Simon Ovenden as the Vice-Chairperson. A list of the TWG members is provided in Annex 5.

41. The full TWG convened five times during its two-year mandate (January 2024 to December 2025), with one virtual and four in-person meetings.[25] Of these four in-person meetings, the first three were held at the OPCW Headquarters in The Hague, the Netherlands, and the final meeting was held at the Finnish Institute for Verification of the Chemical Weapons Convention (VERIFIN) in Helsinki, Finland. Summaries of the first four meetings were reported to the SAB at its Thirty-Eighth and Thirty-Ninth Sessions.[26] The summary of the fifth meeting[27] and this end-of-mandate report will be presented at the Fortieth Session of the SAB in May 2026.

---

25 A sixth meeting took place at the OPCW Main Building in November 2025. This meeting was only attended by the TWG Chairperson, Vice-Chairperson, and subgroup leads, and was dedicated to drafting this end-of-mandate report.

26 (a) Summary of the First Meeting of the Scientific Advisory Board's Temporary Working Group on Chemical Forensics (SAB-38/WP.1, dated 15 May 2024). (b) Summary of the Second Meeting of the Scientific Advisory Board's Temporary Working Group on Chemical Forensics (SAB-39/WP.1, dated 21 August 2024). (c) Summary of the Third Meeting of the Scientific Advisory Board's Temporary Working Group on Chemical Forensics (SAB-39/WP.2, dated 16 January 2025). (d) Summary of the Fourth Meeting of the Scientific Advisory Board's Temporary Working Group on Chemical Forensics (SAB-39/WP.3, dated 25 March 2025).

27 Summary of the Fifth Meeting of the Scientific Advisory Board's Temporary Working Group on Chemical Forensics (SAB-40/WP.2, dated 27 August 2025).

42. In addition to these plenary meetings, a series of biweekly virtual meetings took place from September 2025 to February 2026 to support preparation of this end-of-mandate report. The timeline of the TWG's mandate is shown in Figure 5.

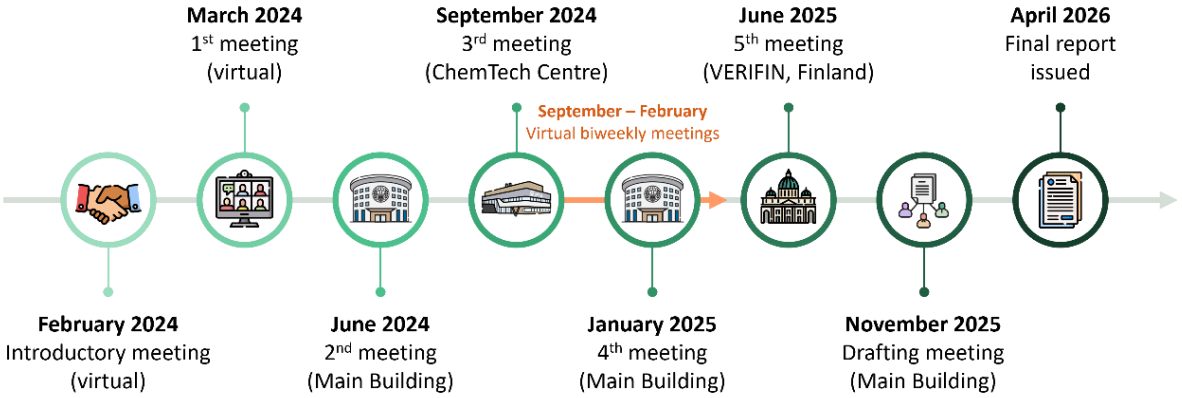


Figure 5: Timeline of the TWG's mandate

43. The work of the TWG was structured to address the TOR. First, to inform its discussions, the TWG received a total of 34 presentations: four from Secretariat staff, six from TWG members, and 24 from invited experts. The list of invited speakers and their areas of expertise is available in Annex 6. The presentations from Secretariat staff (OPCW Laboratory and OSM) provided the TWG with additional context and understanding relating to its current practices, the needs of the OPCW, and its operational constraints. The technical presentations by TWG members and invited experts covered a broad range of topics across all aspects of forensic research, including chemical, environmental, fire, and nuclear forensics, evidence collection and documentation, chain of custody, advances in instrumentation, novel analytical techniques for sample matching and chemical profiling, chemometrics (including concepts such as machine learning (ML) and multivariate statistical analysis), artificial intelligence (AI), block-chain, and forensics databases.

44. Second, the TWG formed four subgroups—aligning with the expertise of its members—to address the TOR in the most efficient and effective manner. The overarching topics and specific questions assigned to the subgroups are outlined in Table 2. While each subgroup focused on its assigned topics, these were also discussed collectively in plenary sessions. This approach ensured that the findings and recommendations, although originating from a particular subgroup, were agreed by consensus across the entire TWG. The findings and recommendations of the TWG on Chemical Forensics will be instrumental in augmenting the Secretariat's capability in this field, particularly in the current threat landscape. The TWG's recommendations will also play an important role in facilitating the work of current and future OPCW contingency operations.

Table 2: Subgroups of the TWG and their assigned questions from the TOR

Subgroup	Assigned Questions from the TOR	
1 – State of the Art	6(a)	What is the current state of the art related to determining the life cycle of a given chemical sample?
	6(b)	Can analysis of other materials that may be found, such as a storage container, or parts of a dissemination device, yield relevant information compared to liquid and soil samples?
	6(c)	What information is available related to the ability to conduct chemical forensics analysis on non-traditional agents which could be applied to CWA? Are there any restrictions or limitations?
2 – Future Capabilities	7(a)	What impact will the increased power and integration of ML and deep learning have on the field of chemical forensics?
	7(b)	What impact will the limited size of datasets available have on CWA forensic analysis?
	7(c)	What will the field look like in five to ten years, particularly in regard to the capabilities around the specific areas mentioned in Subgroup 1?
3 – Methods and Procedures	8(a)	How can applied analytical methods have an impact on the results related to trace analysis and the chemical footprint?
	8(b)	How can data, methods, and procedures to conduct the in-depth CWA analyses expected in chemical forensics be standardised and shared?
	8(c)	What information is needed to ensure trust and reproducibility in the analysis and the results?
	8(d)	How can analytical data from different techniques be combined in forensic analysis?
	8(e)	Would curated/shared database(s) of relevant reference data be useful? Are there any recommendations/restrictions to secure these types of database(s)?
	8(f)	What best practices exist for on-site sampling to ensure the validity of subsequent forensic laboratory analysis and what challenges/gaps remain?

Subgroup	Assigned Questions from the TOR	
4 – Augmenting the OPCW's capabilities	9(a)	How can the OPCW enhance its capability to capture and utilise chemical forensics-related information in the context of alleged use?
	9(b)	How will the OPCW ensure the validity of the forensic results?
	9(c)	How will the forensic results be reported? (criteria, level of confidence)
	9(d)	Are there other organisations or information that could be leveraged to augment the capability of the OPCW?
	9(e)	How can forensic analysis be promoted and enhanced at DLs? How might other organisations or laboratories contribute?

FINDINGS OF THE TWG ON CHEMICAL FORENSICS

Chemical forensic questions

45. At the outset of its work and to frame its discussions, the TWG defined three key chemical forensic questions that may be posed during an investigation relating to the alleged use of a CWA and identified the processes required to answer them.

    i. **Is there a link between samples?** To answer this question, the chemical profiles of two or more samples suspected to share a common origin need to be compared and their similarity assessed. This can be achieved through **sample matching**, which compares individual samples to see if they could have the same source or production process, or **batch matching**, which links samples to a particular production batch.[28] The presence of impurities such as unreacted starting materials and any associated contaminants, by-products from synthesis, stabilisers, and other components in a CWA sample enable chemical profiles to be compared. Such comparisons are a common application of chemical forensics, helping investigators connect events or seized materials to each other.

    ii. **Was a particular chemical produced using a specific starting material?** This may be addressed by **precursor-product matching**, in which the chemical profile of a starting material is compared with that of a product (for example, a CWA). Published examples demonstrate that chemical impurities present in starting materials may be carried through the entire synthesis pathway and remain detectable in the chemical impurity profile of the product, which may be a CWA precursor or final product.[29,30] Chemical profiling experiments using isotope ratios have also been shown to be a powerful tool for precursor-product

---

28 For consistency and brevity, references to both sample matching and batch matching are combined as "sample/batch matching" throughout this report.

29 Webster, Renée L., Simon P. Ovenden, Lyndal J. McDowall, Genevieve H. Dennison, Melissa J. Laws, Nathan W. McGill, Jilliarne Williams, and Shannon D. Zanatta. "Chemical Forensic Profiling and Attribution Signature Determination of Sarin Nerve Agent Using GC–MS, LC–MS and NMR." *Analytical and Bioanalytical Chemistry* 414, no. 13 (April 8, 2022): 3863–73. https://doi.org/10.1007/s00216-022-04027-1.

30 Säde, Solja, Lina Mörén, Karin Höjer Holmgren, Hanna Hakulinen, Andreas Larsson, Magnus Engqvist, Linnea Ahlinder, Rikard Norlin, Harri Kiljunen, Crister Åstot and Paula Vanninen. "Chemical Impurity Profiling: Linkage of Starting Materials and an Intermediate Synthesis Product of a Carbamate Chemical Warfare Agent." *Forensic Chemistry* 39 (July 2024): 100581. https://doi.org/10.1016/j.forc.2024.100581.

matching.[31] The combination of impurity profiles and isotope ratios may further increase the confidence of a match.

iii. **How was this CWA produced?** This forensic question may be answered through **route sourcing**, which uses chemical profiling techniques to identify the synthetic pathway or method employed to produce a particular chemical. Route sourcing depends on impurities such as route-specific by-products and remnants of the starting materials (or their associated impurities), as well as on reliable databases for by-product identification, and the expertise of analytical and synthetic chemists to interpret these data. The combination of these impurities gives rise to a unique chemical profile that is specific to the production route used.

46. Chemical impurity profiles are critical for addressing the three chemical forensic questions described above. Figure 6 shows a gas chromatography-mass spectrometry (GC-MS) chromatogram of a crude sulfur mustard sample. The main chromatogram is dominated by the peak corresponding to sulfur mustard—highlighted in orange for clarity—which represents the primary component in the sample. The sample also contains multiple impurities at levels below 1% of the sulfur mustard peak intensity. These impurity peaks are barely perceptible in the light blue band of the main chromatogram but are more clearly visible in the magnified inset, which expands the vertical axis to show the true complexity of the sample. Only some of these impurities will be relevant for chemical forensic analysis (shown in blue), while the remainder (shown in black) are not considered informative.

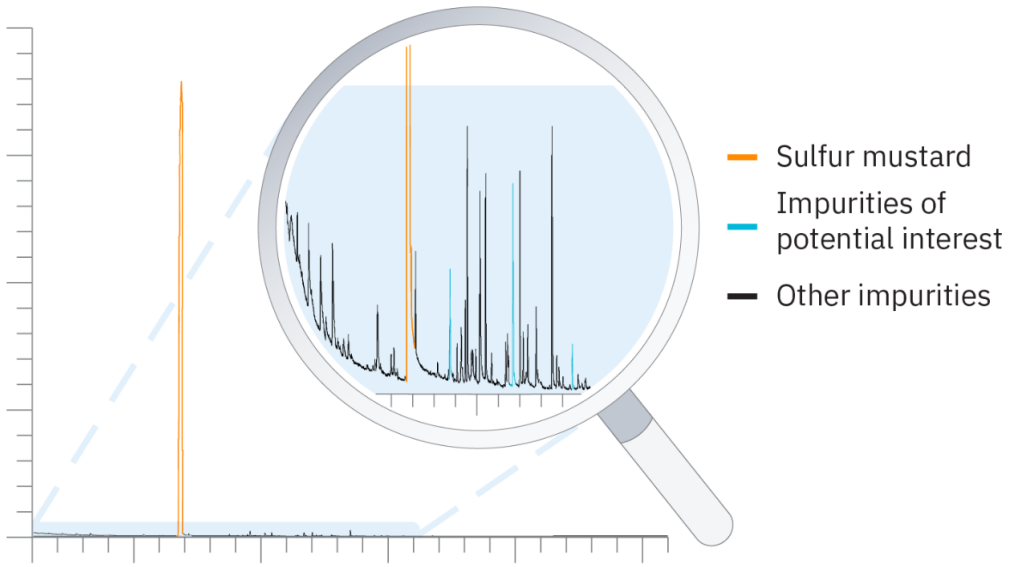


Figure 6: Simplified GC-MS chromatogram of a sulfur mustard sample

[31] Lindberg, Sandra, Magnus Engqvist, Lina Mörén, Crister Åstot, and Rikard Norlin. "Source Attribution of the Chemical Warfare Agent Soman Using Position-Specific Isotope Analysis by ^{2}H NMR Spectroscopy: From Precursor to Degradation Product." *Analytical Chemistry* 93, no. 36 (September 1, 2021): 12230–36. https://doi.org/10.1021/acs.analchem.1c01271.

47. A key component of chemical forensic analysis is the use of **ground-truth data** and classification models to support the interpretation of chemical profiles. Ground-truth data are well-characterised reference samples or datasets with known composition, origin, or production route, providing a reliable benchmark for comparison. **Classification models** are systematic frameworks—often statistical or computational—trained on ground-truth data that use patterns in chemical profiles, including impurities and isotope ratios, to assign unknown samples to a category such as a specific production route, batch, or precursor source. These tools enable the interpretation of complex chemical information in a scientifically robust and reproducible manner. However, the availability of ground-truth data relating to CWAs is limited, which poses a significant challenge in chemical forensic investigations.

48. When ground-truth data are not available, **retrosynthesis** may be used to support route sourcing and help address the forensic question regarding how a particular CWA was produced. This approach, commonly employed by synthetic chemists, involves working backwards from the target product to identify potential precursors and plausible reaction pathways. Through its application, reaction by-products identified in a CWA sample could then be explained from a synthesis perspective, enabling the identification of a plausible production route and, consequently, the starting materials used. Where feasible, the proposed synthetic route is tested in the laboratory to verify the hypothesis. Effective application of retrosynthesis requires expertise in organic synthesis, as interpretation depends on understanding chemical reaction mechanisms and pathways.

Current research and capability development in CWA chemical forensics

49. While chemical forensics is well-established in a number of fields and supported by extensive literature, there remains a comparatively small body of published research on CWA-based chemical forensics. The TWG notes that this is gradually changing: four recent doctoral studies, together with two international collaborative exercises, are strengthening international efforts, accelerating research, and expanding the knowledge base in this important area. These endeavours were presented to the TWG, which found them particularly valuable and insightful for informing its recommendations, and are briefly described below, focusing exclusively on aspects relevant to chemical forensics, particularly their work on methods for chemical profiling.

50. In her doctoral research programme, Dr Karin Höjer Holmgren (Linköping University and the Swedish Defence Research Agency (FOI), Sweden, in collaboration with Lawrence Livermore National Laboratory, United States of America) undertook studies on the CWAs *S*-[2-(diethylamino)ethyl] *O*-(2-methylpropyl) *P*-methylphosphonothioate (VR) and sulfur mustard, as well as on the nerve agent

precursor methylphosphonyl dichloride (DC).[32] Across three studies, classification methods for route sourcing were developed based on impurity profiles derived from ground-truth CWA samples.[33,34,35] The studies demonstrated that the classification models were capable of distinguishing between six and 11 production routes for VR and sulfur mustard, respectively. The work further highlighted the power of combining non-targeted analytical approaches with high-resolution mass spectrometry (HRMS) to enable detection of low-concentration impurities in sulfur mustard samples. The performance of two ML classification methods was successfully validated with sample test sets, and key chemical markers relevant to route differentiation were identified. A fourth study reported the results of an interlaboratory comparison of DC samples jointly conducted by eight defence research laboratories based in Europe, North America, Asia, and Australia.[36] In this study, impurity profiles in DC samples were matched between laboratories, demonstrating that harmonised laboratory GC-MS methods for CWA profiling can produce highly comparable analytical data across different facilities.

51. Dr Mirjam de Bruin-Hoegée (University of Amsterdam and the Netherlands Organisation for Applied Scientific Research (TNO), the Netherlands), during her doctoral research, investigated chemical impurity profiling and route sourcing of fentanyl and its analogues in samples that had been incubated with liver microsomes

---

32 Höjer Holmgren, Karin. "Route Attribution of Chemical Warfare Agents: Retrospective Classification of Unknown Threat Samples." *Linköping Studies in Science and Technology. Dissertations*, (December 8, 2022). https://doi.org/10.3384/9789179295844.

33 Höjer Holmgren, Karin, Carlos A. Valdez, Roger Magnusson, Alexander K. Vu, Sandra Lindberg, Audrey M. Williams, Armando Alcaraz, Crister Åstot, Saphon Hok, and Rikard Norlin. "Part 1: Tracing Russian VX to its Synthetic Routes by Multivariate Statistics of Chemical Attribution Signatures." *Talanta* 186 (August 15, 2018): 586–96. https://doi.org/10.1016/j.talanta.2018.02.104.

34 Höjer Holmgren, Karin, Saphon Hok, Roger Magnusson, Andreas Larsson, Crister Åstot, Carolyn Koester, Daniel Mew, et al. "Synthesis Route Attribution of Sulfur Mustard by Multivariate Data Analysis of Chemical Signatures." *Talanta* 186 (August 15, 2018): 615–21. https://doi.org/10.1016/j.talanta.2018.02.100.

35 Höjer Holmgren, Karin, Lina Mörén, Linnea Ahlinder, Andreas Larsson, Daniel Wiktelius, Rikard Norlin, and Crister Åstot. "Route Determination of Sulfur Mustard Using Nontargeted Chemical Attribution Signature Screening." *Analytical Chemistry* 93, no. 11 (March 12, 2021): 4850–58. https://doi.org/10.1021/acs.analchem.0c04555.

36 Höjer Holmgren, Karin, Hanna Hakulinen, Rikard Norlin, Mirjam de Bruin-Hoegée, Marie Spiandore, Samantha Qi Shu See, Renee Webster, et al. "Interlaboratory Comparison Study of a Chemical Profiling Method for Methylphosphonic Dichloride, a Nerve Agent Precursor." *Forensic Chemistry* 33 (May 2023): 100473. https://doi.org/10.1016/j.forc.2023.100473.

in vitro to mimic human metabolism.[37,38] The study demonstrated that, even after metabolic transformation, it was possible to correctly classify samples according to their production route. Two further studies focused on the elemental profiling of polymers.[39,40] This technique has potential value for analysing plastic items used to store or disseminate CWAs and for linking samples originating from the same source. During this work, novel standards were developed for the quantitative analysis of trace element profiles in polymers using laser ablation-inductively coupled plasma mass spectrometry (LA-ICP-MS). Using these standards, LA-ICP-MS was demonstrated to be a promising approach for comparing and classifying polymeric materials, providing an orthogonal analytical method that complements existing chemical forensic techniques.

52. During her doctoral studies, Dr Carla Orlandi (Institut national de recherche pour l'agriculture, l'alimentation et l'environnement (INRAE), in collaboration with the French Defence Agency (DGA), France) applied metabolomics-based methodologies to the chemical forensic analysis of CWAs. Using the pesticide chlorpyrifos as a model CWA, non-targeted metabolomics-based approaches were successfully used to classify chlorpyrifos samples based on a wide variety of key chemical impurities (such as unreacted precursors, synthesis by-products, or degradation products). Analyses using multiple techniques—including trapped ion mobility spectrometry coupled to liquid chromatography coupled to high-resolution mass spectrometry (LC-HRMS)[41] and a chromatography-free approach based on matrix-assisted laser desorption ionisation (MALDI) coupled to HRMS with trapped ion mobility spectrometry[42]—

---

37 de Bruin-Hoegée, Mirjam. "Revealing the Origin of Chemical Weapons." PhD diss., University of Amsterdam, 2024.

38 Vangerven, Daan, Mirjam de Bruin-Hoegée, Fleur Kerstens, Meike Kerklaan, Rowdy P.T. Bross, Alex Fidder, Marcel J. van der Schans, Daan Noort, and Arian C. van Asten. "Post-Metabolism Impurity Profiling of Carfentanil, Remifentanil, Sufentanil, and Benzylfentanyl." *Forensic Chemistry* 40 (September 2024): 100587. https://doi.org/10.1016/j.forc.2024.100587.

39 de Bruin-Hoegée, Mirjam, Jorien Schoorl, Peter Zoon, Marcel J. van der Schans, Daan Noort, and Arian C. van Asten. "A Novel Standard for Forensic Elemental Profiling of Polymers by LA-ICP-TOF-MS." *Forensic Chemistry* 35 (September 2023): 100515. https://doi.org/10.1016/j.forc.2023.100515.

40 de Bruin-Hoegée, Mirjam, Ruthmara Corzo, Peter D. Zoon, Peter Vergeer, Jorien Schoorl, Marcel J. van der Schans, Daan Noort, and Arian C. van Asten. "Evaluating the Strength of Evidence of Elemental Profiling of Polymers with LA-ICP-MS." *Forensic Chemistry* 38 (May 2024): 100570. https://doi.org/10.1016/j.forc.2024.100570.

41 Orlandi, Carla, Grégoire Delaporte, Christine Albaret, Emmanuel Joubert, Anne Bossée, Laurent Debrauwer, and Emilien Jamin. "Unveiling Impurity Profiling of Synthetic Pathways of Organophosphorus Chlorpyrifos through LC–HRMS Metabolomics-Based Approaches" *Rapid Communications in Mass Spectrometry* (January 31, 2025): e9996. https://doi.org/10.1002/rcm.9996.

42 Orlandi, Carla, Justine Ferey, Grégoire Delaporte, Christine Albaret, Emmanuel Joubert, Anne Bossée, Laurent Debrauwer, and Emilien Jamin. "Chromatography-Free High Throughput Analysis by Matrix Assisted Laser Desorption Ionization Mass Spectrometry, Applied to Chlorpyrifos Synthetic Sources Evidence." *Talanta* 301 (May 1, 2026): 129291. https://doi.org/10.1016/j.talanta.2025.129291.

combined with chemometric analysis, proved effective in addressing key analytical challenges associated with chemical forensics. The robustness of the approach was then assessed using crude samples of chlorpyrifos synthesised independently after one year, as well as environmental samples (water and soil) spiked at varying concentrations without aging; all were successfully classified. Data were introduced into the pre-existing multivariate statistical methods, and the synthetic route used to produce the chlorpyrifos was predicted. Importantly, it was determined that working with a larger number of samples covering a wider variability would improve the robustness of statistical models. Additionally, approaches such as integrating data from several analytical sources, exploitation of metadata, or the use of adapted validation methods could be explored to increase fidelity of results.

53. In her doctoral work, Dr Solja Säde (VERIFIN, University of Helsinki, Finland) researched concepts and methods that aimed to advance the field of chemical forensics of CWAs. As part of initial standardisation efforts, a proposed quality control (QC) test sample—consisting of 27 compounds—was analysed by 11 participating laboratories to assess its applicability and robustness in evaluating the performance of GC-MS instruments.[43] Data from the participating laboratories were assessed for their suitability for interlaboratory comparison. As a result, the QC test sample has been included in the Recommended Operating Procedures for Analysis in the Verification of Chemical Disarmament, 2023 Edition (the "Blue Book") [44] and employed in the Icarus exercise (see paragraph 54). In a second study, a new experimental methodology was developed for the impurity profiling of carbamate CWA precursors and for linking them to their starting material suppliers.[30] This study was a collaboration between FOI and VERIFIN and investigated the differences in data resulting from conducting the syntheses in two separate laboratories. The data generated in the carbamate research were subsequently comprehensively evaluated. Six relevant multivariate classification methods were applied, alongside three variable selection methods commonly used in CWA chemical forensic studies. This approach demonstrated that the supervised classification methods produced similar results, whereas the variable selection methods showed significant variability.

54. Finally, Icarus is a collaborative international exercise coordinated through the Chemical Forensics International Technical Working Group (CFITWG),[45] which was

---

43 Säde, Solja, Grégoire Delaporte, Carlos G. Fraga, Hanna Hakulinen, Karin Höjer Holmgren, Marie Spiandore, Crister Åstot, et al. "Interlaboratory Development and Proposition for a New Quality Control Sample for Chemical Forensics Analysis of Chemical Warfare Agents." *Talanta Open* 8 (December 2023): 100249. https://doi.org/10.1016/j.talo.2023.100249.

44 *Recommended Operating Procedures for Analysis in the Verification of Chemical Disarmament.* University of Helsinki, 2023.

45 "Chemical Forensics International Technical Working Group." U.S. Department of State, accessed March 17, 2026.

established to address gaps in chemical forensics science and capabilities through international partnerships of experts from scientific, policy, academic, law enforcement, and export-control organisations. Icarus is a multi-laboratory chemical forensic exercise using wet (laboratory-based) sample matching, which began in 2024 and will conclude in 2026.

55. The organising team, led by Dr Carlos Fraga (Air Force Research Laboratory, United States of America) and including members from four DLs (DGA, FOI, Spiez Laboratory, and VERIFIN), developed a coherent scenario for a chemical forensic investigation over the course of a year. Samples for the exercise were produced in three DLs (FOI, Spiez, and VERIFIN) using both a common synthesis pathway and alternative synthesis routes. To eliminate bias, the organising team was separated from the teams performing blind analysis of the samples.

56. In total, 14 laboratories are involved in the analysis of 19 unknown samples. The participating laboratories include DLs and academic partners, as well as the OPCW Laboratory. The analyses require adapted sample treatment procedures, and numerous analytical methods are being tested, including GC-MS, gas chromatography coupled to high-resolution mass spectrometry (GC-HRMS), inductively coupled plasma mass spectrometry (ICP-MS), position-specific isotope analysis by nuclear magnetic resonance spectroscopy (PSIA-NMR), isotope ratio mass spectrometry (IRMS), and LC-HRMS. Chemometric data analysis is also being applied. Participants are leveraging this exercise to further test the recently developed QC test sample (see paragraph 53).[43]

57. To assist participating laboratories and enable comparison of results obtained from GC-MS analysis, data pre-treatment was performed centrally by a single laboratory (DGA). For comparison, laboratories may also carry out pre-processing individually. Regular technical team meetings are being organised to share preliminary results, discuss challenges, and review the methods used.

58. This exercise aims to determine whether samples of a CWA simulant—diisopropyl methylphosphonate—can be matched to other samples from the same source (a "Tier 1" forensic question involving nine samples) and to its precursors, DC (five samples) and isopropyl alcohol (five samples), from specific sources (a "Tier 2" forensic question). The exercise is designed to simulate real-world conditions and address realistic chemical forensic questions.

59. The results generated in this exercise will be shared among the participating laboratories, and the chemical forensic questions posed will be addressed collectively. This approach will enable the outcomes of individual laboratories to be compared with those of an international team comprising members from all participating laboratories. Moreover, the exercise will explore the benefits of

combining data from different techniques and compare approaches for data integration. Specifically, chemical forensics results derived from merging raw data from two analytical techniques (using a multi-block approach) will be compared with results obtained by pooling chemometric outputs from two datasets generated by different techniques. This comparison will help assess the value of these strategies in chemical forensic investigations.

60. The outcomes of Icarus are expected to provide a clear assessment of the current state of the art in chemical forensic analysis related to CWAs and identify areas where further research and improvement of methodologies are needed. These insights will help guide future development, collaboration, and standardisation efforts. In addition, the publication of validated methods and approaches resulting from the exercise in peer-reviewed journals will expand the relatively limited body of published research in this field. These validated methods and approaches will also be incorporated into the next edition of the Blue Book, [46] ensuring that new methodologies and best practices are being shared across the DLs and other relevant communities.

[46] Fraga, Carlos G., Crister Åstot, Karin Höjer Holmgren, Rikard Norlin, Solja Säde, Marie Spiandore, and Grégoire Delaporte. "Section 5. Attribution Analyses." In *Recommended Operating Procedures for Analysis in the Verification of Chemical Disarmament*, 1109-51, University of Helsinki, 2023.

Subgroup 1: State of the art

61. Subgroup 1 addressed three questions from the TOR, provided below. In addressing these questions, it specifically considered the areas of sample/batch matching; impurity and trace element profiling; route sourcing; geographic context; and environmental factors and storage conditions.

    i. **Question 6(a)** What is the current state of the art related to determining the life cycle of a given chemical sample?

    ii. **Question 6(b)** Can analysis of other materials that may be found, such as a storage container, or parts of a dissemination device, yield relevant information compared to liquid and soil samples?

    iii. **Question 6(c)** What information is available related to the ability to conduct chemical forensics analysis on non-traditional agents which could be applied to CWAs? Are there any restrictions or limitations?

62. In order to respond to these questions, Subgroup 1 reviewed the current state of the art across several fields of forensics (including environmental, fire debris, and illicit drugs) where chemical forensic techniques have been applied. TWG members and invited expert speakers alike shared their knowledge and experience, providing insights into how chemical forensic information may yield leads and evidence in support of forensic investigations. In the area of CWAs, a number of relevant publications were considered.[29,30,34,47] The application of chemical forensics to investigations of chemical weapons use in the Syrian Arab Republic and for assessments of its declared chemical weapons programme—specifically by the FFM, IIT, and DAT—was of particular interest to this subgroup.

---

47 Examples include: (a) Fraga, Carlos G., Gabriel A. Pérez Acosta, Michael D. Crenshaw, Krys Wallace, Gary M. Mong, and Heather A. Colburn. "Impurity Profiling to Match a Nerve Agent to Its Precursor Source for Chemical Forensics Applications." *Analytical Chemistry* 83, no. 24 (November 22, 2011): 9564–72. https://doi.org/10.1021/ac202340u. (b) Lu, Xiaogang, Zixuan Zhang, Runli Gao, Hongmei Wang, and Junhua Xiao. "Recent Progress in the Chemical Attribution of Chemical Warfare Agents and Highly Toxic Organophosphorus Pesticides." *Forensic Toxicology* 39, no. 2 (May 13, 2021): 334–49. https://doi.org/10.1007/s11419-021-00578-7. (c) Strozier, Erich D., Douglas D. Mooney, David A. Friedenberg, Theodore P. Klupinski, and Cheryl A. Triplett. "Use of Comprehensive Two-Dimensional Gas Chromatography with Time-of-Flight Mass Spectrometric Detection and Random Forest Pattern Recognition Techniques for Classifying Chemical Threat Agents and Detecting Chemical Attribution Signatures." *Analytical Chemistry* 88, no. 14 (June 29, 2016): 7068–75. https://doi.org/10.1021/acs.analchem.6b00725.

Question 6(a): What is the current state of the art related to determining the life cycle of a given chemical sample?

63. As a powerful chemical forensics tool, chemical profiling is widely applied in the field to identify unique features within unknown samples. Given that sample/batch matching, precursor-product matching, and route sourcing are the three chemical profiling applications required to answer the most likely forensic questions posed in investigations of alleged use of a chemical weapon, Subgroup 1 focused on the state of the art in these three areas.

64. **Sample matching** is a core chemical forensic method used in criminal investigations, including forensic areas such as illicit drug analysis.[48] Chemical impurity profiling of a large number of seized narcotic drug samples is performed at national forensic institutes worldwide in order to find linkages between seizures, thereby deducing common origins of drug samples to map trading networks.[49] Illicit drug production using similar or identical production routes and conditions can make the chemical impurity profiles very similar. Therefore, the application of harmonised chemical analytical methods is essential to produce chemical impurity profiling data with sufficient quality to resolve small differences in the impurity profiles. In addition, samples are often run simultaneously to minimise analytical variation, which is particularly important given the trace-level nature of many impurities.[50]

65. Chemical impurity profile data from CWA samples are typically obtained using GC-MS, liquid chromatography-mass spectrometry (LC-MS), and nuclear magnetic resonance (NMR) spectroscopy. The use of HRMS generally offers greater sensitivity and allows the detection of impurities at trace levels.[35] The analytical techniques used for impurity profiling form the backbone of techniques used in chemical identification analysis and, as such, the DL network is well equipped to undertake chemical impurity

---

48 Ahmed, Reem, Mohamad J. Altamimi, and Mayssa Hachem. "State-of-the-Art Analytical Approaches for Illicit Drug Profiling in Forensic Investigations." *Molecules* 27, no. 19 (October 5, 2022): 6602. https://doi.org/10.3390/molecules27196602.

49 (a) Broséus, Julian, Sami Huhtala, and Pierre Esseiva. "First Systematic Chemical Profiling of Cocaine Police Seizures in Finland in the Framework of an Intelligence-Led Approach." *Forensic Science International* 251 (June 2015): 87–94. https://doi.org/10.1016/j.forsciint.2015.03.026. (b) Czyż, Anna, Katarzyna Pawlak, Emilia Waraksa, and Tomasz Bieńkowski. "Comprehensive Profiling of Illicit Amphetamines Seized in Poland: Insights from Gas Chromatography–Mass Spectrometry and Chemometric Analysis." *Molecules* 30, no. 3 (January 2025): 579. https://doi.org/10.3390/molecules30030579. (c) Salouros, Helen. "Illicit Drug Chemical Profiling: Current and Future State." *Australian Journal of Forensic Sciences* 50, no. 6 (January 2018): 689–96. https://doi.org/10.1080/00450618.2018.1424244.

50 Lock, Eric, Laura Aalberg, Kjell Andersson, Johan Dahlén, Michael D. Cole, Yvonne Finnon, Henk Huizer, et al. "Development of a Harmonised Method for the Profiling of Amphetamines V." *Forensic Science International* 169, no. 1 (June 2007): 77–85. https://doi.org/10.1016/j.forsciint.2006.10.019.

profiling analysis. The TWG on IST underscored the importance of chemical profiling reference data for determining the provenance of a sample.[19]

66. In the area of CWA forensics, an impurity profiling method for DC was developed[47a,51] and subsequently assessed by seven different DLs in an interlaboratory comparison study.[36] The method was shown to be robust, and comparison of the GC-MS impurity profiles generated by all participating laboratories produced consistent results. The ongoing collaborative exercise Icarus (described in paragraphs 54 to 60) will also assess the robustness of laboratories to undertake chemical impurity profiling and generate similar results, this time using a more complex forensic scenario. The results from this exercise will provide additional information on how sample matching based on impurity profiles can be applied to the field of CWA forensics.

67. The OPCW IIT has applied chemical profiling in its investigations to identify those involved in the use of chemical weapons in the Syrian Arab Republic. As a result, the IIT was able to match the nerve agent sarin used in the chemical attacks in Ltamenah on 24 and 30 March 2017, to the sarin previously developed by the Syrian Arab Republic as part of its chemical weapons programme.[16a] The sample-matching approach identified specific chemical impurities and unreacted remnants of the binary sarin agent, and the impurity profile was similar to the samples collected at the site of the Khan Shaykhun sarin attack on 4 April 2017, as reported by the OPCW-United Nations (UN) Joint Investigative Mechanism (JIM).[52] This approach provided evidence that the perpetrators of the two chemical attacks investigated had used sarin with a common origin.

68. A negative sample-matching result can be just as informative as a positive one, as demonstrated in the IIT's investigation of the sulfur mustard attack in Marea, where it was alleged that toxic chemicals originating from the stockpile of the Syrian Arab Republic had been seized by a non-State actor operating in the area where the incident occurred.[16d] Based on its chemical impurity profile, the sulfur mustard used in Marea was shown to have been produced using an improvised Levinstein method. In contrast, the sulfur mustard samples from the Syrian government stockpile lacked the Levinstein-specific chemical impurities and instead exhibited a chemical impurity profile characteristic of the Meyer production route. The available evidence therefore allowed the IIT to discount the hypothesis that the chemical agent used in the attack may have originated from a State stockpile as extremely unlikely.

---

51 Hoggard, Jamin C., Jon H. Wahl, Robert E. Synovec, Gary M. Mong, and Carlos G. Fraga. "Impurity Profiling of a Chemical Weapon Precursor for Possible Forensic Signatures by Comprehensive Two-Dimensional Gas Chromatography/Mass Spectrometry and Chemometrics." *Analytical Chemistry* 82, no. 2 (December 16, 2009): 689–98. https://doi.org/10.1021/ac902247x.

52 Seventh report of the Organisation for the Prohibition of Chemical Weapons-United Nations Joint Investigative Mechanism (S/2017/904, dated 26 October 2017).

69. **Precursor-product matching** may use both impurity profiling and stable isotope analysis to link a seized sample of a product (for example, a nerve agent) to a suspected precursor (such as DC) used in its production.[53]

70. Carbon, hydrogen, nitrogen, oxygen, sulfur, phosphorus, and fluorine are the main elements in most CWAs. While all are widespread in nature, phosphorus and fluorine are monoisotopic, meaning they only have one stable isotope. The other elements have multiple stable isotopes, whose relative abundances vary geographically and environmentally. The integrity of this intrinsic isotopic information, as reflected in natural-abundance isotope profiles within a chemical of interest, is typically conserved from precursors to products during a chemical reaction, although minor changes may occur due to isotopic fractionation during the synthetic process. Consequently, isotope analysis of CWAs can provide forensic insights—such as precursor-product matching—only for elements with multiple stable isotopes.

71. The determination of stable isotope ratios in CWAs is performed using techniques such as IRMS[54,55,56] and PSIA-NMR,[31,57] which are compared in Figure 7. Isotope ratio mass spectrometry provides the overall (bulk) isotope composition of an element within a molecule, representing an average across all atoms of that element. In contrast, PSIA-NMR provides isotopic information for individual atomic positions within a molecule.

72. Isotope ratio mass spectrometry is often coupled with a separation technique such as gas chromatography (GC), enabling the measurement of stable isotope ratios for specific CWAs within complex samples.[56] The enhanced sensitivity and selectivity

---

53 Lippmann, Martin, Christoph Schaefer, Clara Schindler, Michiel Beukers, Niels Beijer, Moritz Hitzemann, Ben van de Kamp, Ruud Peters, Jaap Knotter, and Stefan Zimmermann. "Revealing Illicit Drug Laboratories by Gas Chromatography-Ion Mobility Spectrometry." *Forensic Science International* 377 (December 2025): 112661. https://doi.org/10.1016/j.forsciint.2025.112661.

54 Mirjankar, Nikhil S., Carlos G. Fraga, April J. Carman, and James J. Moran. "Source Attribution of Cyanides Using Anionic Impurity Profiling, Stable Isotope Ratios, Trace Elemental Analysis and Chemometrics." *Analytical Chemistry* 88, no. 3 (January 8, 2016): 1827–34. https://doi.org/10.1021/acs.analchem.5b04126.

55 Csernica, Timothy, James J. Moran, Carlos G. Fraga, and John M. Eiler. "Simultaneous Observation of ^{2}H and ^{13}C Enrichment of Methyl Phosphonic Acid via Orbitrap-IRMS with Applications to Nerve Agent Forensics" *Talanta* 281 (January 2025): 126802. https://doi.org/10.1016/j.talanta.2024.126802.

56 Moran, James J., Carlos G. Fraga, and Megan K. Nims. "Stable-Carbon Isotope Ratios for Sourcing the Nerve-Agent Precursor Methylphosphonic Dichloride and Its Products." *Talanta* 186 (August 2018): 678–83. https://doi.org/10.1016/j.talanta.2018.04.021.

57 Rasmussen, Cornelia, and David W. Hoffman. "Novel Nuclear Magnetic Resonance Method for Position-Specific Carbon Isotope Analysis of Organic Molecules with Significant Impurities." *Analytical Chemistry* 94, no. 43 (October 2022): 15124–31. https://doi.org/10.1021/acs.analchem.2c03356.

provided by GC-IRMS make it particularly well-suited for analysing real-world samples.

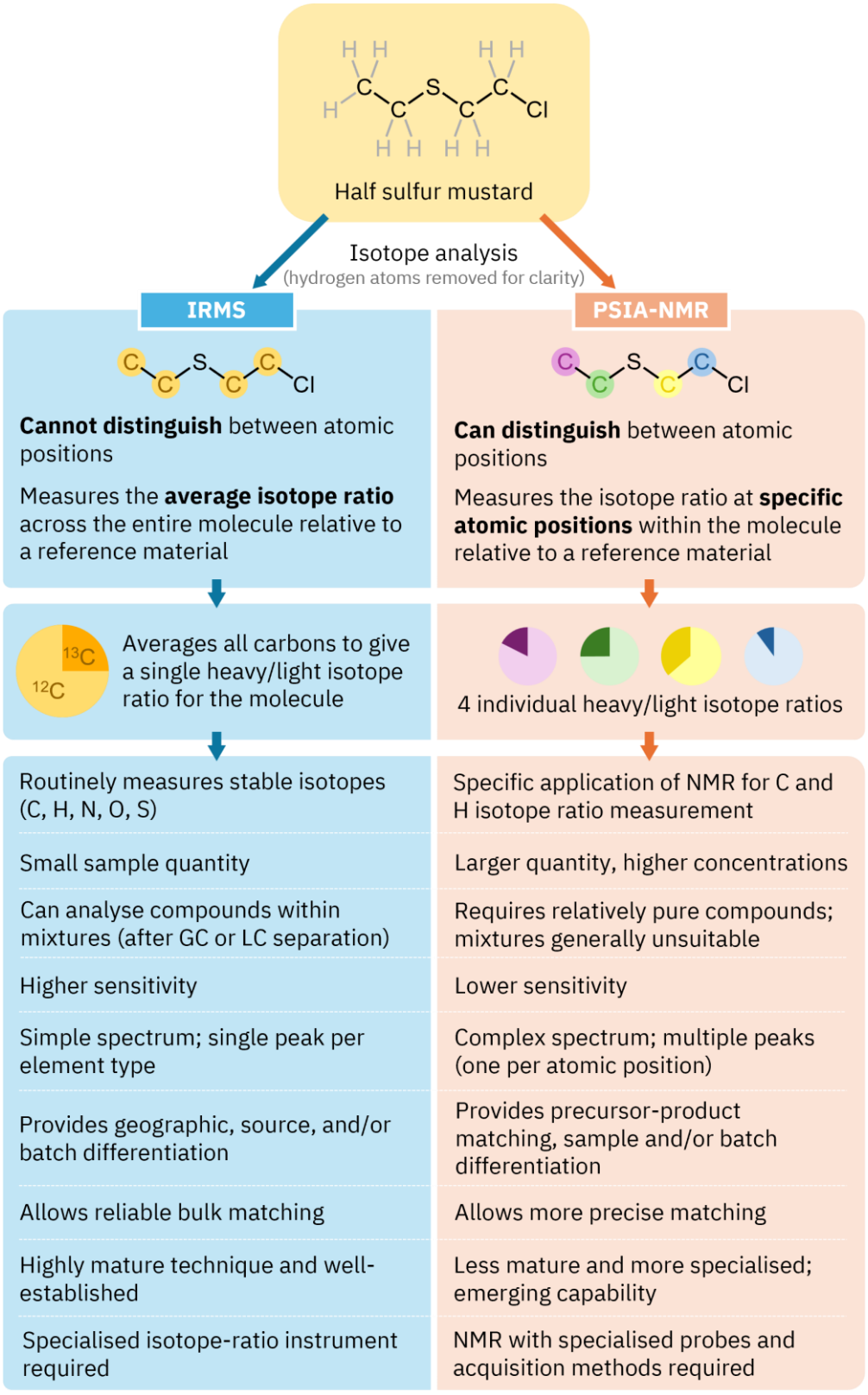


Figure 7: Comparison of IRMS and PSIA-NMR

73. Position-specific isotope analysis by NMR can determine $^{2}H/^{1}H$ and $^{13}C/^{12}C$ ratios. The isotopic distribution of deuterium (^{2}H) in the nerve agent precursors pinacolyl[31] and isopropyl alcohols,[58] and in their corresponding nerve agents soman and sarin, has been measured using PSIA-NMR. The results indicate that this technique can be used to match CWA products and precursors successfully.

74. For carbon isotopes, PSIA-NMR can determine $^{13}C/^{12}C$ isotope ratios either indirectly, using satellite peaks in ^{1}H and/or ^{31}P one-dimensional spectra, or directly combining ^{13}C NMR spectra with IRMS measurements.[59] Compared to deuterium-based PSIA-NMR, $^{13}C/^{12}C$ analysis carries a lower risk of peak overlap due to the wider range of chemical shifts and sharper peak resolution. However, the natural variation in $^{13}C/^{12}C$ ratios is about one order of magnitude smaller than that of hydrogen isotopes, making measurements more challenging. Consequently, specific acquisition methods and processing are required to achieve the accuracy needed for reliable forensic interpretation.

75. Compared with IRMS, NMR analysis requires a higher CWA concentration to ensure high-quality data acquisition. In particular, PSIA-NMR requires that the CWA be the primary component of the sample. To meet these requirements, if neat chemical is not available, samples that contain CWA precursors and/or degradation products undergo preparation to extract the analytes of interest and increase their concentration and/or purity.

76. It should be noted that while developments are being made in the application of HRMS for PSIA measurements, the technique has not yet been demonstrated in an operational application in chemical forensics.[60,61]

77. Impurity profiling has also been reported to be suitable for CWA precursor-product matching. Provided they are chemically stable, impurities present in a given precursor batch may be transferred into the corresponding CWA, and their detection can

---

58 Meier, Urs C. "Forensic Analysis of the Deuterium/Hydrogen Isotopic Ratios of the Nerve Agent Sarin, Its Reaction By-Product Diisopropyl Methylphosphonate and Their Precursors by ^{2}H SNIF-NMR." *Talanta* 253 (February 2023): 123890. https://doi.org/10.1016/j.talanta.2022.123890.

59 Hoffman, David W. and Cornelia Rasmussen. "Position-specific Carbon Stable Isotope Analysis of Glyphosate: Isotope Fingerprinting of Molecules within a Mixture." *Analytical and Bioanalytical Chemistry* 416, no. 16 (May 2024): 3847-56. https://doi.org/10.1007/s00216-024-05326-5.

60 Weiss, Gabriella M., Alex L. Sessions, Maxime Julien, Timothy Csernica, Keita Yamada, Alexis Gilbert, Katherine H. Freeman, and John M. Eiler. "Analysis of Intramolecular Carbon Isotope Distributions in Alanine by Electrospray Ionization Orbitrap Mass Spectrometry." *International Journal of Mass Spectrometry* 493 (November 2023): 117128. https://doi.org/10.1016/j.ijms.2023.117128.

61 Wilkes, Elise B., Alex L. Sessions, Sarah S. Zeichner, Brooke Dallas, Brian Schubert, A. Hope Jahren, and John M. Eiler. "Position-Specific Carbon Isotope Analysis of Serine by Gas Chromatography/Orbitrap Mass Spectrometry, and an Application to Plant Metabolism." *Rapid Communications in Mass Spectrometry* 36, no. 18 (September 2022): e9347. https://doi.org/10.1002/rcm.9347.

support a precursor-product link.[47a] In the investigation of the 2017 sarin attack in Khan Shaykhun by the OPCW-UN JIM,[52] the sarin samples were reported to contain three specific impurities: hexafluorophosphate, isopropyl phosphates, and isopropyl phosphorofluoridates. These three impurities were also found in samples of the nerve agent precursor methylphosphonyl difluoride (DF), collected during the 2014 OPCW-UN Joint Mission to remove and destroy Syria's chemical weapons stockpiles. The presence of these specific chemical impurities provided a strong indication that the sarin used in Khan Shaykhun was produced using DF from the Syrian stockpile.

78. **Route sourcing**—used to classify the production method of a chemical of interest—was first pioneered in illicit drugs forensics. The production route for a given chemical consists of the precursors used and the reaction conditions applied; in many cases, multiple routes are possible for the same chemical. Methods for route sourcing of synthetic drugs, including amphetamines and analogues,[62] and fentanyl,[63,64,65,66] have been published. Illicit drug manufacturing laboratories are increasingly innovative, developing alternative production routes to circumvent the regulatory restrictions on critical precursors. Consequently, monitoring the chemical profiles of seized illicit drugs is important for providing intelligence to law enforcement organisations.

---

62 Stojanovska, Natasha, Shanlin Fu, Mark Tahtouh, Tamsin Kelly, Alison Beavis and K. Paul Kirkbride. "A Review of Impurity Profiling and Synthetic Route of Manufacture of Methylamphetamine, 3,4-Methylenedioxymethylamphetamine, Amphetamine, Dimethylamphetamine and P-Methoxyamphetamine." *Forensic Science International* 224 no. 1-3 (January 2013): 8-26. https://doi.org/10.1016/j.forsciint.2012.10.040.

63 Casale, John F., Patrick A. Hays, Steven G. Toske and Jennifer R. Mallette. "Unique Bipiperidinyl Impurities Produced from the "One-Pot" Synthesis of Fentanyl." *Forensic Chemistry* 17 (March 2020): 100203. https://doi.org/10.1016/j.forc.2019.100203.

64 Mayer, Brian P., Alan J. DeHope, Daniel A. Mew, Paul E. Spackman and Audrey M. Williams "Chemical Attribution of Fentanyl Using Multivariate Statistical Analysis of Orthogonal Mass Spectral Data." *Analytical Chemistry* 88, no 8 (March 2016): 4303-10. https://doi.org/10.1021/acs.analchem.5b04434.

65 Toske, Steven G., Jennifer R. Mitchell, James M. Myslinski, Andrew J. Walz, David B. Guthrie, Elizabeth M. Guest, Charlotte A. Corbett and Emily D. Lockhart "Organic Impurity Profiling of Fentanyl Samples Associated with Recent Clandestine Laboratory Methods." *Journal of Forensic Science* 68 no. 5 (September 2023): 1470-83. https://doi.org/10.1111/1556-4029.15281.

66 Ovenden, Simon P.B., Lyndal J. McDowall, Hugh E. McKeown, Nathan W. McGill, Oliver A.H. Jones, James R. Pearson, Marija Petricevic, et al. "Investigating the Chemical Impurity Profiles of Fentanyl Preparations and Precursors to Identify Chemical Attribution Signatures for Synthetic Method Attribution." *Forensic Science International* 321 (April 2021): 110742. https://doi.org/10.1016/j.forsciint.2021.110742.

79. Studies of route sourcing of CWAs produced by different routes have been reported for VR,[33] VX,[67] sulfur mustard,[34,35] and tabun.[68] The publications show how specific impurity profiles from samples produced by different routes can be used to classify unknown samples according to their production route (method). Route sourcing has provided essential information about the chemical weapons used in the Syrian Arab Republic, as reported by the OPCW-UN JIM and the OPCW IIT. Based on analysis performed by DLs, the OPCW-UN JIM concluded in its seventh report that the sarin disseminated in Khan Shaykhun on 4 April 2017 was produced by the binary route, reacting DF with isopropyl alcohol in the presence of hexamine.[52] A second route sourcing example may be found in the fourth report of the IIT, which focused on the incident in Marea on 1 September 2015.[16d] The IIT determined that the sulfur mustard—disseminated as a black liquid—had been produced by an improvised Levinstein method. The sulfur mustard was shown to contain Levinstein route-specific chemical impurities. Furthermore, its black colour and the presence of polychlorinated mustard species indicated that the sulfur mustard had been produced under improvised conditions, with a shortage of ethylene, which led to the formation of black tar.

80. A number of **considerations and limitations** with respect to CWA chemical profiling were considered. Chemical warfare agents (for example, chemicals listed in Schedule 1) and many of the by-products contained within a production batch sample are highly reactive chemicals. This is an important consideration when their chemical impurity profiles are analysed in sample matching and route sourcing processes. When applicable, a set of chemically stable impurities are preferred as they offer reproducible profiles over extended time periods.[69] However, it is important to note that some highly reactive by-products may also contain valuable information for route sourcing. The successful application of chemical forensics to an incident involving chemical weapons may therefore require that samples have not been unduly exposed to extreme environmental conditions (such as high temperatures or humidity).

---

67 Ovenden, Simon P.B., Renée L. Webster, Eva Micich, Lyndal J. McDowall, Nathan W. McGill, Jilliarne Williams, and Shannon D. Zanatta. "The Identification of Chemical Attribution Signatures of Stored VX Nerve Agents Using NMR, GC-MS, and LC-HRMS." *Talanta* 211 (May 2020): 120753. https://doi.org/10.1016/j.talanta.2020.120753.

68 Lu, Xiaogang, Xiaxia Zhu, Runli Gao, Hui Tang, Chengxin Pei, Hongmei Wang, and Junhua Xiao. "Chemometrics-Assisted Analysis of Chemical Impurity Profiles of Tabun Nerve Agent Using Comprehensive Two-Dimensional Gas Chromatography-Time-of-Flight Mass Spectrometry." *Journal of Chromatography A* 1685 (December 6, 2022): 463643. https://doi.org/10.1016/j.chroma.2022.463643.

69 Aalberg, Laura, Kjell Andersson, Christina Bertler, Michael D. Cole, Yvonne Finnon, Henk Huizer, Kaisa Jalava, et al. "Development of a Harmonised Method for the Profiling of Amphetamines: II. Stability of Impurities in Organic Solvents." *Forensic Science International* 149, no. 2–3 (May 10, 2005): 231–41. https://doi.org/10.1016/j.forsciint.2004.06.019.

81. Consequently, it is important to gain access to the incident site to collect environmental and/or material samples as soon as practicable after an alleged CWA incident. The longer the time between the incident and sample collection, the greater the risk that the CWAs and many of their related impurities will degrade due to environmental conditions. This makes sample matching to a common origin more difficult. While degradation products (diagnostic markers) may still exist in these affected samples, the resolving power of chemical forensics may be reduced. In these circumstances, the ability to acquire analytical data on isotope ratios (for example, using PSIA-NMR) of key degradation products could help match relevant isotope profiles between samples.

82. Ground-truth data obtained from the chemical analysis of reference samples with a known synthetic route is important. Such reference data is a prerequisite for route sourcing methods as it helps define the chemical impurity profile of CWA samples produced by a specific synthesis route. The reference data also enables classification models that would help elucidate a possible production route for unknown samples to be built. In sample matching, a set of independent reference samples is needed to assess the variation in the chemical profiles of CWA samples produced using the same route. The absence of sufficient reference data limits the effectiveness of chemical forensics as a tool in investigations of chemical weapons use.

83. Geographical sourcing is a method used to link plant-derived materials to their geographical origin based on variations in their metabolic profiles.[70] It has been used to determine the origin of illicit plant-based drugs and biotoxins.[71] To date, the application of this technique to CWAs is limited, although approaches using profiles of trace elements, such as strontium isotope ratios ($^{87}Sr/^{86}Sr$), [72] have been implemented. It is important to emphasise, however, that although determining the geographical origin of a CWA is theoretically possible, no known methods can link a sample to a particular production facility without access to specific reference samples for sample matching.

---

[70] Abdelrahman, Mostafa, Sho Hirata, Takuya Mukae, Tomohiro Yamada, Yuji Sawada, Magdi El-Syaed, Yutaka Yamada, Muneo Sato, Masami Y. Hirai, and Masayoshi Shigyo. "Comprehensive Metabolite Profiling in Genetic Resources of Garlic (*Allium Sativum L.*) Collected from Different Geographical Regions." *Molecules* 26, no. 5 (March 2021): 1415. https://doi.org/10.3390/molecules26051415.

[71] Pigott, Eloise J., Warren Roberts, Simon P. Ovenden, Simone Rochfort, and David J. Bourne. "Metabolomic Investigations of *Ricinus Communis* for Cultivar and Provenance Determination." *Metabolomics* 8, no. 4 (August 26, 2011): 634–42. https://doi.org/10.1007/s11306-011-0355-7.

[72] Kant, Lisa B., Megan K. Nims, Abdullah D. Shouaib, Kelly C. McHugh, Tyler D. Schlieder, Eirik J. Krogstad, and Elizabeth H. Denis. "Geographic Source Attribution of Honey by Strontium Isotope Analyses: Latvia and India Measurements Compared to Model Predictions in a Feasibility Study." *Journal of Food Composition and Analysis* 145 (September 2025): 107822. https://doi.org/10.1016/j.jfca.2025.107822.

84. The successful application of chemical forensics to the analysis of real-world CWA samples from the Syrian Arab Republic highlights the valuable contribution these forensic tools may bring to investigations of chemical weapons use. Currently, the complexity of such investigations calls for experts with extensive forensic analysis experience relating to a particular CWA and it is challenging for the OPCW and its Member States to maintain preparedness over time. The actions proposed by the TWG in this report address this challenge and will help ensure that this capacity is maintained and further developed in the future. (**Recommendation 1**)

Question 6(b): Can analysis of other materials that may be found, such as a storage container or parts of a dissemination device, yield relevant information compared to liquid and soil samples?

85. Chemical forensics can provide information to indicate how a CWA was stored or dispersed. For example, chemical profiling methods may be employed to identify a specific plasticiser profile from a polymeric material[39] or a trace metal profile,[73] linking a CWA sample to a particular storage container. This information may be used to strengthen a sample match based on chemical profiling analysis of organic chemicals. (**Recommendation 5**)

The OPCW should continue to monitor and evaluate emerging analytical methods relevant to chemical forensics and engage with expert laboratories both within and beyond the DL network that possess specialised capabilities. DLs should be encouraged to establish their own collaborations with external expert laboratories to further strengthen their chemical forensics capabilities.

[73] Ulrich, Andrea, Christoph Moor, Heinz Vonmont, Hans-Rudolf Jordi and Martin Lory. "ICP-MS Trace-Element Analysis as a Forensic Tool." *Analytical and Bioanalytical Chemistry* 378 no. 4 (February 2004): 1059-68. https://doi.org/10.1007/s00216-003-2434-8.

Question 6(c): What information is available related to the ability to conduct chemical forensics analysis on non-traditional agents which could be applied to CWAs? Are there any restrictions or limitations?

86. State-of-the-art chemical forensics capabilities using profiling techniques are found in forensic laboratories that perform sample matching and production route sourcing in areas such as illicit drugs, fire debris analysis, and environmental forensics, as well as in research laboratories advancing this field. However, applying these capabilities to CWA-related chemical forensics is not straightforward. Unlike most materials analysed in these forensic fields, CWAs and many of their associated impurities are highly chemically reactive. As a result, a narrower set of chemically stable impurities is suitable for profiling, which limits the resolution and robustness of sample matching analyses compared with those achievable for less reactive substances.

87. An OPCW investigation of chemical weapons use begins with the verification of the presence or absence of a Convention-relevant chemical(s). This verification provides an unambiguous identification of a CWA, forming the basis for subsequent investigations. Chemical forensic analysis, while interpretive, is grounded in objective chemical data and represents one component of a comprehensive forensic investigation, contributing evidence alongside other investigative approaches. To ensure robust forensic outcomes, the investigation team should pose clear chemical forensic questions to the scientists at the outset, revising them iteratively as evidence is collected. (**Recommendation 8**)

88. Furthermore, the multivariate nature of chemical impurity profiles collected as part of the chemical forensic investigation calls for the use of advanced data evaluation tools based on AI and ML—such as those that have found application in omics research—for both sample matching and the construction of production route classification models. Additionally, probabilistic evaluation, which applies statistical methods to quantify the likelihood that samples are related, taking into account variability and uncertainty in the data, is now being applied in forensic science (see Annex 2). Probabilistic evaluation can help strengthen conclusions in chemical forensics—for example, by assessing the similarity between chemical impurity profiles between samples—and represents an emerging state-of-the-art approach in the field. Expert assessments of chemical forensic data acquired by scientifically accepted methods constitute the accepted procedure for presenting chemical evidence in a court of law. The further development of tools for classification of unknown samples and for probabilistic evaluation of forensic findings will improve the scientific basis for these expert assessments made, enabling more robust and defensible conclusions in cases of chemical weapons use. (**Recommendation 6**)

The Secretariat, in conjunction with DLs, should evaluate current and emerging AI/ML tools and software from relevant scientific and forensic fields that could be applied to chemical forensics. Compliance with national and OPCW security policies should be considered as part of this evaluation. The OPCW Laboratory and DLs should also be encouraged to share and evaluate preferred tools and software in CBEs.

Subgroup 2: Future capabilities

89. Although substantial progress has been made in chemical forensics since its inception, the state of the art in data analysis within this field lags considerably behind that of the omics sciences, a group of biological research fields that study the complete set of molecules of a particular type within a cell, tissue, organism, or environment. This disparity can be attributed to a combination of scientific and organisational factors that complicate the implementation of modern data analysis tools and workflows. Scientifically, sample numbers in CWA-based chemical forensic studies are very limited, and samples are often difficult to obtain. Organisationally, laboratories undertaking CWA-based chemical forensics have additional security and confidentiality constraints that can restrict large-scale data sharing and collaboration. In this context, it is necessary to explore the opportunities offered by advanced AI/ML methods and workflows—particularly those already proven in other relevant scientific fields such as the omics sciences. Drawing on insights from invited expert speakers, Subgroup 2 therefore considered the applicability of these approaches to chemical forensics and how the current scientific and organisational constraints might be overcome.

90. In order to explore the opportunities offered by advanced AI/ML tools, methods, and workflows, Subgroup 2 considered the following three questions from the TOR:

    i. **Question 7(a)** What impact will the increased power and integration of ML and deep learning have on the field of chemical forensics?

    ii. **Question 7(b)** What impact will the limited size of datasets available have on CWA forensic analysis?

    iii. **Question 7(c)** What will the field look like in five to ten years, particularly in regard to the capabilities around the specific areas[74] considered in Subgroup 1?

91. Machine learning is a subset of AI, which refers more broadly to the goal of achieving some form of intelligence in machines. Machine learning focuses specifically on algorithms that learn from data to make predictions or classifications. This topic has been considered in detail by the recent TWG on AI. In chemical forensics, ML methods are increasingly relevant, as they can identify patterns in complex datasets, enabling samples to be linked to specific precursors, additional samples, or potential production routes.

92. Machine learning methods can be applied using different strategies. Supervised ML requires labelled training data—for example, identifying the synthetic pathway

[74] Sample/batch matching; impurity and trace element profiling; route sourcing; geographic context; environmental factors and storage conditions; and isotope ratio techniques.

utilised to generate each sample—which are used to develop a model (i.e., a set of rules) that can then be applied to unlabelled test data, such as classifying the synthetic route of an unknown sample. Unsupervised ML does not require labels and identifies patterns without explicit instruction, while self-supervised learning is a strategy where models generate their own labels from the input data's inherent structure, enabling training without manual annotation.[75] Of these approaches, supervised ML is the most common. Its effectiveness depends on the availability of sufficiently large and representative datasets, as models trained on limited or non-representative data can overfit: performing well on data similar to the training set but generalising poorly to other data.[76] Consequently, when training sets are sparse, smaller and simpler models are often preferred.

93. Classification workflows are particularly important, as they are widely used in chemical forensics. These workflows are structured, iterative, and systematic, and are used to build models that predict discrete classes or categories. Most workflows follow a common structured process designed not only to achieve optimal model performance but also to identify and mitigate potential statistical bias. In this approach, training and validation data are derived from a ground-truth dataset, while model performance is evaluated using test data from an independent ground-truth dataset (see Figure 8). This strict independence is particularly critical in chemical forensics, where high variable-to-sample ratios increase the risk of overfitting during data analysis.

94. Model performance is evaluated using standard metrics to quantify accuracy and highlight potential misclassifications. A commonly used tool is the confusion matrix, which summarises the number of correct and incorrect predictions across classes. Further technical details on ML strategies and model validation are provided in Annex 7.

Question 7(a): What impact will the increased power and integration of ML and deep learning have on the field of chemical forensics?

95. Over the past 15 years, the implementation of advanced data analysis workflows, together with developments in instrumentation, has enabled significant progress in scientific fields. This progress has been particularly notable in the omics sciences. Some omics sciences—principally proteomics, metabolomics, and their derivatives—

---

75 Carpenter, Gail A., and Stephen Grossberg. "Adaptive Resonance Theory." In *Encyclopedia of Machine Learning and Data Mining*, 24–40, Springer US, 2017. https://doi.org/10.1007/978-1-4899-7687-1_6.

76 Webb, Geoffrey I. "Overfitting." In *Encyclopedia of Machine Learning and Data Mining*, 947–948, Springer US, 2017. https://doi.org/10.1007/978-1-4899-7687-1_960.

utilise the same analytical techniques as chemical forensics. Advances in these areas can be attributed, in part, to the incorporation of AI/ML.

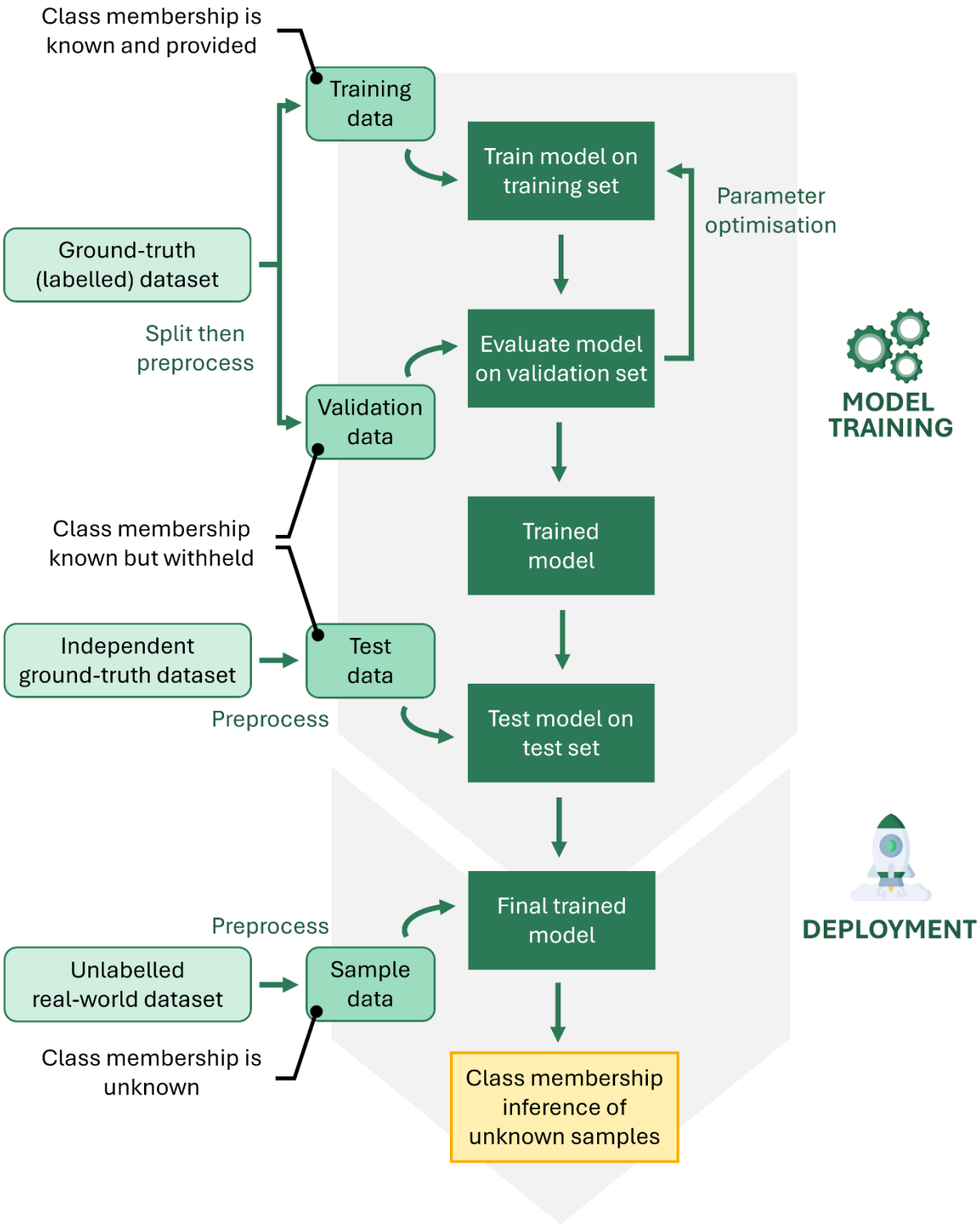


Figure 8: ML classification workflow

96. Methods using AI/ML are now employed extensively across the omics sciences, supporting data analysis tasks such as classification, sample differentiation, and

sample discrimination. They also help analysts handle the large volumes of data generated in these fields, enabling the discovery of significant patterns in complex datasets through techniques such as data mining. In addition, AI/ML methods support the creation of large-scale databases (for example, MetaboLights[77] and the Human Metabolome Database[78]) and are used to assist in populating these resources by augmenting datasets where experimental data is scarce. Examples include generating mass spectra from chemical structures[79] and predicting collision cross-section values in ion mobility mass spectrometry.[80] These developments clearly highlight the potential of AI/ML approaches to help address complex scientific questions.

97. Consideration was given to how AI capabilities could be leveraged to help address the three key chemical forensic questions defined by the TWG at the start of its mandate (see paragraph 45). It was noted that the scientific questions addressed in the omics field are often similar in type to those in chemical forensics. For instance, pairwise comparisons of biological samples correspond to sample and precursor-product matching, while classification of exposed/control samples in an epidemiological study is analogous to route sourcing. Selecting the most appropriate data analysis method is important, as well as adapting the workflow to the forensic question(s) being addressed. Providing relevant data science training to chemists has become essential in this increasingly interdisciplinary field.

98. In addition to similarities in question types, the analytical chemistry challenges in both the omics sciences and chemical forensics are often comparable: there are large numbers of unknown compounds, a wide dynamic range of analyte concentrations, limited availability of experimental reference data, and datasets that are both high-dimensional and complex. These challenges highlight the need for effective data analysis strategies to support analytical chemists.

---

77 (a) Yurekten, Ozgur, Thomas Payne, Noemi Tejera, Felix Xavier Amaladoss, Callum Martin, Mark Williams, and Claire O'Donovan. "MetaboLights: Open Data Repository for Metabolomics." *Nucleic Acids Research* 52, no. D1 (November 16, 2023). https://doi.org/10.1093/nar/gkad1045. (b) "MetaboLights." European Molecular Biology Laboratory's European Bioinformatics Institute, accessed December 15, 2025.

78 (a) Wishart, David S., Dan Tzur, Craig Knox, Roman Eisner, An C. Guo, Nelson Young, Dean Cheng, et al. "HMDB: The Human Metabolome Database." *Nucleic Acids Research* 35, Database issue (January 2007): D521–26. https://doi.org/10.1093/nar/gkl923. (b) "HMDB: The Human Metabolome Database." The Human Metabolome Database, accessed December 15, 2025.

79 Wang, Fei, Jaanus Liigand, Siyang Tian, David Arndt, Russell Greiner, and David S. Wishart. "CFM-ID 4.0: More Accurate ESI-MS/MS Spectral Prediction and Compound Identification." *Analytical Chemistry* 93, no. 34 (August 17, 2021): 11692–700. https://doi.org/10.1021/acs.analchem.1c01465.

80 Bouwmeester, Robbin, Keith Richardson, Richard Denny, Ian D. Wilson, Sven Degroeve, Lennart Martens, and Johannes P.C. Vissers. "Predicting Ion Mobility Collision Cross Sections and Assessing Prediction Variation by Combining Conventional and Data Driven Modeling." *Talanta* 274 (July 2024): 125970. https://doi.org/10.1016/j.talanta.2024.125970.

99. To date, the AI/ML methods used to address chemical forensics-related questions have mostly relied on linear algebra, with the most commonly applied being unsupervised principal component analysis (PCA) and supervised partial least squares discriminant analysis (PLS-DA). These methods are considered the workhorses of chemical forensics, especially for route sourcing determinations, because they are highly interpretable, very flexible (adaptable to a variety of situations, including multi-block analysis from several analytical datasets), computationally efficient, and easy to implement. However, PCA and PLS-DA are not always capable of handling the most complex situations and chemical forensics could therefore benefit from access to a larger selection of validated AI/ML methods. These new methods could be variations of linear algebra methods[81] or use ML methods from other families such as those described in Table 3.[82] The importance of ensuring that these new methods are able to analyse data from several analytical sources (so-called multi-block data) simultaneously in a way that is relevant for chemical forensics was underscored (see also paragraph 135).

100. Guided by areas of significant development reported in the scientific literature, database searching and data mining, analytical and chemical property prediction, and data augmentation were considered in detail. As a result of the development of large-scale analytical databases—such as the Wiley Registry/NIST[83] Mass Spectral Library 2023[84]—combined with the increasing volume and complexity of data generated by modern instrumentation, the use of advanced data analysis methods and strategies has become essential. For example, advanced data analysis methods have been successfully applied in forensic analysis of illicit drugs. AI-assisted sample matching has been used to detect fentanyl analogues by training an ML model on the NIST mass spectral database to recognise fentanyl-like spectra, even when the specific chemical is absent from the reference dataset.[85]

---

81 Forsgren, Edvin, Benny Björkblom, Johan Trygg, and Pär Jonsson. "OPLS-Based Multiclass Classification and Data-Driven Interclass Relationship Discovery." *Journal of Chemical Information and Modeling* 65, no. 4 (February 3, 2025): 1762–70. https://doi.org/10.1021/acs.jcim.4c01799.

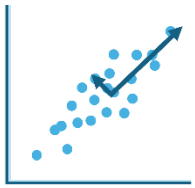
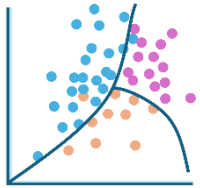
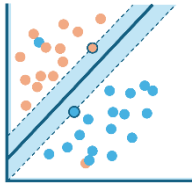
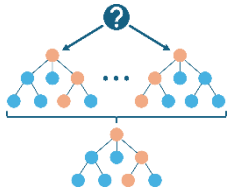
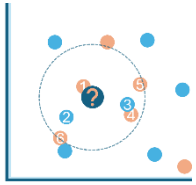
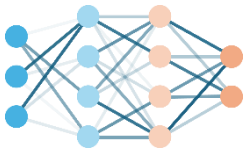
82 Reproduced with permission from Tiggelaar, Emma F. D. "From Molecules to Models: The Implementation of Machine Learning in the Field of Chemical Warfare Agents." Master's thesis, University of Amsterdam, 2024.

83 National Institute of Standards and Technology.

84 "Wiley Registry/NIST Mass Spectral Library 2023." Wiley Science Solutions, accessed December 15, 2025.

85 Koshute, Phillip, Nathan Hagan, and N. Jordan Jameson. "Machine Learning Model for Detecting Fentanyl Analogs from Mass Spectra." *Forensic Chemistry* 27 (March 2022): 100379. https://doi.org/10.1016/j.forc.2021.100379.

Table 3: Summary of ML methods

Method family	Linear methods	Probabilistic methods	Hyperplane methods	Decision tree methods	Cluster methods	Neural networks
Illustrative representations						
Examples	• PCA • PLS-DA • Linear discriminant analysis • Logistic regression	• Naive Bayes classifier • Kernel density estimation	• Support vector machines	• Random forest (RF) • Gradient boosting methods (GBM)	• k-nearest neighbours	• Artificial neural network • Convolutional neural network
Applications	• Classification • Prediction • Data visualisation	• Classification	• Classification • Prediction	• Classification • Prediction	• Classification • Prediction	• Image processing • Large language models
Chemical forensics application example	Synthetic route classification	Threat detection, outlier detection (unpublished)	Not applied directly to chemical forensics to date Used to help database matching of unknowns in CFM-ID[79]	Synthetic route classification, mass spectra class prediction (RF) for drug analysis GBM not applied to chemical forensics to date but is known to often outperform RF	Sample matching	Not applied to chemical forensics to date Deep neural networks used in mass spectral prediction tools such as CFM-ID[79]

Method family	Linear methods	Probabilistic methods	Hyperplane methods	Decision tree methods	Cluster methods	Neural networks
Theoretical basis	Linear algebra, matrix decomposition	Probability theory	Maximising margin with support vectors	Combines multiple weak classifiers (decision trees) to obtain reliable output	Distance-based similarity	Mimics neurone/ synapse network connection, can be built with several neurone layers
Advantages	Simple, interpretable	Handles non-linear relationships, parametric and non-parametric methods	Effective on high-dimensional data	Little overfitting, handles missing values, small decision trees are interpretable Some GBM offer improved interpretability as well (XGBoost)	Simple, non-parametric	Some models can accept natural language inputs, can handle a large diversity of requests
Disadvantages	Relies on strong statistical assumptions, limited to linear relationships	Can be computationally intensive, requires large datasets, can be difficult to tune	Interpretability is often poor, designed for two-class problems	Requires large datasets, small decision trees perform poorly and are unstable, interpretability is sacrificed for performance and stability in bigger models (RF or GBM)	Computationally expensive on large data, need to recompute entire model at each query, natively not interpretable	Require large datasets, computationally heavy, interpret-ability often very poor, consumer market oriented and behave poorly without proper training on a specific scientific subject

101. Chemical and analytical property prediction is now an established area of analytical chemistry. Earlier prediction tools developed in the 1990s and 2000s relied largely on cheminformatics approaches,[86] whereas more recent approaches increasingly leverage AI/ML methods.[87] Over the past few years, numerous studies have demonstrated the use of AI/ML for predicting retention times[88] and indices,[89] mass spectra,[79] and collision cross-section values.[80] This has been facilitated by the availability of large training datasets—particularly for mass spectral prediction—since the NIST/EPA/NIH Mass Spectral Library 2023[90] and Wiley Registry of Mass Spectral Data 2023 GC-MS (EI)[91] databases contain approximately 400,000 and 900,000 spectra, respectively.

102. However, it is important to consider the representativeness of these databases for CWA-related analyses, as CWA-relevant molecules likely constitute only a very small fraction of their total entries, even though some chemicals of interest—such as plasticisers and explosives—are included. The only curated database of significant size for CWAs is the OCAD, which contains more than 14,700 data entries (including spectra and retention indices) covering 10,720 unique chemicals.[92] Despite its relatively large size and coverage of six analytical techniques, the OCAD contains very few entries for tandem mass spectrometry (MS) (315 entries) and GC-HRMS (35 entries), in particular. The prediction of mass spectra using *in silico* methods has been successfully demonstrated in large-scale metabolomics repositories, such as the Human Metabolome Database, and could therefore be applied to existing CWA databases to help address current data gaps. (**Recommendation 16**)

103. Through its discussions, and drawing on insights from experts, the TWG concludes that the areas of chemical forensics likely to see the greatest impact from AI/ML integration are the automated annotation of unknown compounds (including the *in*

---

86 Gasteiger, Johann, Wolfgang Hanebeck, and Klaus P. Schulz. "Prediction of Mass Spectra from Structural Information." *Journal of Chemical Information and Computer Sciences* 32, no. 4 (July 1992): 264–71. https://doi.org/10.1021/ci00008a001.

87 Russo, Francesco F., Yannek Nowatzky, Carsten Jaeger, Maria K. Parr, Phillipp Benner, Thilo Muth, and Jan Lisec. "Machine Learning Methods for Compound Annotation in Non-targeted Mass Spectrometry—a Brief Overview of Fingerprinting, in Silico Fragmentation and de Novo Methods." *Rapid Communications in Mass Spectrometry* 38, no. 20 (August 2024). https://doi.org/10.1002/rcm.9876.

88 Haddad, Paul R., Maryam Taraji, and Roman Szücs. "Prediction of Analyte Retention Time in Liquid Chromatography." *Analytical Chemistry* 93, no. 1 (October 21, 2020): 228–56. https://doi.org/10.1021/acs.analchem.0c04190.

89 Vrzal, Tomáš, Michaela Malečková, and Jana Olšovská. "DeepReI: Deep Learning-Based Gas Chromatographic Retention Index Predictor." *Analytica Chimica Acta* 1147 (February 22, 2021): 64–71. https://doi.org/10.1016/j.aca.2020.12.043.

90 "NIST/EPA/NIH Mass Spectral Library 2023." Wiley Science Solutions, accessed December 15, 2025.

91 "Wiley Registry of Mass Spectral Data 2023." Wiley Science Solutions, accessed December 15, 2025.

92 Release of the OPCW Central Analytical Database (S/2477/2026, dated 19 January 2026).

silico generation of analytical data) and the increasing capability to combine analytical data from multiple sources. The Group notes that the current state of the art of AI/ML tools in other scientific fields demonstrates a high potential to help chemists extract the maximum information from chemical data during investigations. (**Recommendation 16**)

104. To effectively harness developments in AI/ML, advances related to analytical chemistry and their applications in areas of chemical forensic analysis beyond CWAs need to be continuously monitored. Rapid progress in related fields, such as the omics sciences, demonstrates how AI/ML can transform the analysis of complex, high-dimensional datasets. Chemical forensics could benefit from similar approaches, but only if developments are actively tracked and adapted to the field's specific challenges. Without such monitoring, chemical forensics, particularly in the context of CWAs, risks falling behind the state of the art.

Question 7(b): What impact will the limited size of datasets available have on CWA forensic analysis?

105. In traditional chemical forensics fields, such as illicit drugs, abundant authentic sample data is available due to the high number of seizures, enabling extensive sample analysis and dataset development. In contrast, CWA-related chemical forensics faces a very different situation. The relative rarity of CWA incidents, combined with the sensitive nature of the data, means far fewer authentic samples are available. For example, CWA-related datasets may only contain tens of samples, whereas datasets in traditional chemical forensics fields or large-scale metabolomics studies can contain hundreds or even thousands. The sparsity of ground-truth reference samples introduces the additional challenge of ensuring they are representative of real-world variability, which may arise from factors such as synthesis scale, batch-to-batch variation, and environmental weathering. This key difference means that investigations involving CWAs require different approaches and methodologies to those employed in traditional chemical forensics fields.

106. To address these challenges, two avenues may be pursued. First, AI/ML methods that are well suited to small-scale datasets should be explored. Second, the feasibility and relevance of expanding datasets using modern data analysis approaches—both AI-based and otherwise—should be examined. (**Recommendation 16**)

107. Invited data science experts expressed confidence in the ability of modern ML strategies to address the challenges associated with the limited size of CWA-related datasets. This confidence is based on the successful application of such strategies in the omics sciences, including the fine-tuning of models to predict metabolic pathways or analytical properties, as well as the use of carefully designed and

controlled data augmentation techniques to address experimental limitations. One expert did not consider the limited size of CWA-related databases to be a particular challenge, noting that some ML methods—especially the simpler ones—can perform well even on small datasets. For successful application, it is paramount that the analyst has a thorough understanding of the dataset and AI/ML methodology in order to select the most suitable method(s) to meet the objectives of the analysis. The experts emphasised, however, that care should be taken when applying data augmentation methods, as any pre-existing biases in the datasets will be amplified.

108. After extensive discussions with experts, it was decided to expand question 7(b) to include the potential challenges and constraints—both technical and organisational—associated with the implementation of AI/ML methods for chemical forensics applications, as described below:

a. The implementation of some ML methods requires a specific level of competency in IT areas, such as high-performance computing and large data repository management. These skills are often not available within DLs, as analytical and synthetic chemists typically do not possess them. Conversely, AI/ML experts are not often familiar with chemistry. The invited experts agreed unanimously that training chemists in data science, together with close collaboration between chemists and data scientists, is critical for effective implementation of AI/ML methods in chemical forensics.

b. Due to the sparsity of data in CWA-related chemical forensic studies—and the potentially high impact of the results—method interpretability is essential. As AI/ML methods are not always easily interpretable, interpretability should be a key criterion when selecting methods for application in this context. This issue is becoming increasingly important in modern AI/ML applications [93] and close monitoring of the scientific literature can help inform selection of the most appropriate methods.

c. Tools originating from fields such as the omics sciences are often accessed as web-based services or via application programming interfaces (APIs), which may not be compatible with laboratories handling CWA-related data. In addition, any data leakage into the public domain could pose a proliferation risk. These tools may also lack the flexibility required for chemical forensic data analysis workflows. According to experts, these practical limitations could be addressed by training chemists in programming and data science, enabling the development of standalone data analysis workflows within forensic institutions. From a technical perspective, local implementation of data science tools should be

[93] Cardoso Rial, Rafael. "AI in Analytical Chemistry: Advancements, Challenges, and Future Directions." *Talanta* 274 (July 2024): 125949. https://doi.org/10.1016/j.talanta.2024.125949.

feasible, although some additional funding may be required to upgrade computing infrastructure within DLs.

d. As data relevant to chemical forensics of CWAs is limited, data sharing is critical. Effective sharing and subsequent analysis using ML methods rely on the standardisation of dataset management systems and their adoption across all DLs—an approach commonly used in large-scale metabolomics data repositories.[94] The issue of data sharing was also discussed at length by Subgroup 3 (see paragraphs 143 and 153 to 157).

e. Experts from established forensic fields, including drug sourcing and fire debris analysis, noted that the way questions or hypotheses are formulated for analysts can greatly influence the assessment of the reliability of the final results—as the question posed influences both the statistical hypothesis and the likelihood ratio (LR)—and that a balance is needed between the specificity of the question and its investigative or evidentiary value.

109. The recent OPCW AI Research Challenge[95] provides an excellent starting point toward addressing the challenges and constraints discussed above, and the four projects that were funded in 2025 are summarised in Table 4. The TWG recognises that the outcomes could deliver some valuable tools for CWA-related chemical forensic analysis[96] and encourages the OPCW to continue its endeavours in this area and share any resultant tools with the DLs. (**Recommendation 18**)

The Secretariat should share AI/ML tools and reference databases developed through the OPCW AI Research Challenge to support DLs in applying these technologies to chemical forensics.

94 Hajjar, Ghina, Millena C. Barros Santos, Justine Bertrand-Michel, Cécile Canlet, Florence Castelli, Nicolas Creusot, Sylvain Dechaumet, et al. "Scaling-up Metabolomics: Current State and Perspectives." *TrAC Trends in Analytical Chemistry* 167 (October 2023): 117225. https://doi.org/10.1016/j.trac.2023.117225.

95 "OPCW AI Research Challenge: Harnessing AI Tools to Enhance Global Chemical Security." OPCW, April 29, 2025.

96 Wishart, David. "Using Machine Learning to Facilitate the Identification of Harmful Chemicals." Presentation, 15th CBRNe Protection Symposium, Malmö, September 30 – October 2, 2025.

Table 4: Summary of AI Research Challenge projects

Machine Learning for Novel Chemical Weapons Identification	**AI-Driven Prediction and Identification of Toxic Nerve Agents**
University of Alberta, Canada	*Korea Military Academy, Republic of Korea*
Using advanced ML technologies based on chemical language models, this project predicted the structures of novel, chemically feasible toxic compounds. Modified graph neural networks were then used to predict the mass spectra and retention indices of these compounds, generating a comprehensive reference library. This library could enable the OPCW to rapidly identify both known and unknown toxic chemicals, significantly enhancing its ability to detect and monitor potential CWAs.	This project leveraged AI to predict the toxicity and vapour pressure of existing and new organophosphate compounds, including recently scheduled nerve agents. Quantum chemistry techniques were applied to predict spectroscopic information (such as NMR spectra), enhancing chemical identification. This research aimed to support the OPCW's mission by strengthening its capabilities for detecting and verifying chemical threats.
ARTISTiC: Application of Artificial Intelligence to Support the Chemical Weapons Convention	**Adapting, Developing, and Evaluating AI Tools for Characterising and Comparing Chemical Profiles**
TNO, The Netherlands	*Defence Science and Technology Laboratory (Dstl), United Kingdom of Great Britain and Northern Ireland*
After the suspected release of a CWA, OPCW DLs aim to determine the composition of unknown samples. Currently, data interpretation and combining information from various techniques are mostly manual. This project aimed to develop AI models for automatic identification of scheduled chemicals, as well as combine different instrumentation types with data fusion to provide a more comprehensive picture. AI tools to extract characteristic chemical forensic information were explored.	This project aimed to identify the most suitable AI tools for characterising and comparing chemical profiles using MS data. A range of open-source AI models were tested and evaluated. The results were compared against existing non-AI methods, such as PCA. This research aimed to enhance the OPCW's ability to identify emerging threats and strengthen its verification capabilities.

Question 7(c): What will the field look like in five to ten years, particularly in regard to the capabilities around the specific areas considered by Subgroup 1?

110. To answer this question, selected experts from the field of chemical forensics, as well as experts working at the intersection of analytical chemistry and data science were consulted. The experts were asked for their views on which technological innovations or novel applications could be game changers for chemical forensics over the next five to ten years. They were also invited to share any additional perspectives on the future of the field.

111. According to the experts, potential game changers for chemical forensics may emerge from several areas:

    a. Instrumentation: Increases in the sensitivity of analytical systems could benefit chemical forensics. However, there is currently no consensus on which analytical techniques would benefit most from this increased sensitivity. In particular, the potential for further gains in tandem MS remains uncertain, whereas analytical techniques orthogonal to MS (such as NMR, ultra-violet, and infrared spectroscopies) may offer greater opportunities, especially when combined with developments in AI/ML-based signal processing. Future advances in analytical instrumentation should therefore be monitored closely, and their relevance to chemical forensics should be systematically assessed.

    b. Computational power: Improvements in computational power, particularly in low-cost systems that enable intensive computation on more compact and easy-to-use platforms, may also benefit chemical forensics, especially for applications such as analytical data prediction. However, because the AI/ML methods currently used to address chemical forensic questions are often relatively low in computational demand, such advances may be less transformative for chemical forensics than for other scientific fields.

    c. Omics sciences: Advances in the omics sciences and their application to chemical forensics may represent significant game changers, particularly through the development of automated yet flexible data analysis workflows integrated with high-quality databases.

112. In addition to technological improvements, experts noted that some of the most significant developments in chemical forensics may be non-technical in nature. In particular, the emergence of a new generation of forensic chemists who are more comfortable with data analysis tools and who have a stronger understanding of method selection and limitations, was highlighted as a key factor. Moreover, emerging AI-assisted tools, such as large language models, may support chemical forensics by helping analysts become more proficient in data analysis tasks, for example by

writing scripts or accelerating data analysis method implementation, provided that such tools are used wisely.

113. Experts highlighted the value of having access to techniques orthogonal to NMR and MS in order to cover as much of the chemical space of samples as possible. To fully leverage the complex datasets generated, AI/ML methods are expected to be highly valuable. The need for close collaboration between forensic chemists and data scientists was again strongly emphasised by the experts, making such collaboration essential for future developments. As these methods are implemented in more operational contexts, the reliability of chemical forensic workflows will depend on validation through peer-reviewed publications, as well as by the establishment of reporting guidelines by the scientific community. (**Recommendation 7**)

The Secretariat should encourage DLs to further develop, publish, and share chemical forensic methods and workflows for Convention-relevant chemicals, establishing a strong scientific foundation for their application. Designated Laboratories are urged to contribute to updates of the recommended operating procedures in the Blue Book and publication of relevant research in peer-reviewed journals is strongly encouraged.

Subgroup 3: Methods and procedures

114. Methods and procedures form a critical part of a chemical forensics capability. It is important that they are well validated and are performed by laboratories with quality assurance (QA) and QC measures in place to ensure the quality and validity of the results for chemical forensic investigations. In addition, databases of relevant reference data are essential components of these methods and procedures and must be carefully curated.

115. Subgroup 3 examined these considerations through the following six questions from the TOR, and these are discussed separately below:

    i. **Question 8(a)** How can applied analytical methods have an impact on the results related to trace analysis and the chemical footprint?

    ii. **Question 8(b)** How can data, methods, and procedures to conduct the in-depth CWA analyses expected in chemical forensics be standardised and shared?

    iii. **Question 8(c)** What information is needed to ensure trust and reproducibility in the analysis and the results?

    iv. **Question 8(d)** How can analytical data from different techniques (e.g., GC-MS, LC-HRMS, and NMR) be combined in forensic analysis?

    v. **Question 8(e)** Would curated/shared database(s) of relevant reference data be useful? Are there any recommendations/restrictions to secure these types of database(s)?

    vi. **Question 8(f)** What best practices exist for on-site sampling to ensure the validity of subsequent forensic laboratory analysis and what challenges/gaps remain?

Question 8(a): How can applied analytical methods have an impact on the results related to trace analysis and the chemical footprint?

116. There is a multitude of analytical instrumentation, techniques, and methods that can be employed to uncover the total population of intrinsic and extrinsic chemical signatures present in a sample. However, it is important to recognise different analytical instrumentation is, by design, more selective for a particular aspect of a sample's chemical signature, including organic and inorganic chemicals, polar and non-polar compounds, and elemental composition. Instrument sensitivity—dependent on the scan types, instrument parameters, and performance checks—as well as the nature of the matrix and the sample preparation methods, affect the

number and quantities of trace components that can be detected. These factors influence the number and abundance of chemicals identified in a sample which can be harnessed for chemical forensic purposes. It should be noted that as instrument technologies advance, the inherent sensitivity of analytical techniques will increase, leading to improved resolving power in chemical profiling.

117. It may not always be necessary to identify a comprehensive range of chemicals in a sample to obtain the required chemical forensic information, such as the synthetic route or the batch of precursor used. However, analysing a broader range of chemicals can provide greater discriminatory power, as certain chemicals may be essential for distinguishing the samples under investigation from other seized samples.

118. Given the diversity of the individual intrinsic and extrinsic profiles that make up a sample's unique chemical signature for chemical forensic investigations, a variety of analytical instrumentation and techniques should be considered. Laboratories—DLs and beyond—could contribute data from orthogonal techniques to build a comprehensive profile of the contributions of both extrinsic (such as chemical impurities and side reaction products) and intrinsic (such as isotope ratios) signatures.

119. Quality control measures, including criteria related to chemical impurity profiling and QC samples, should be established to ensure the reproducibility of chemical impurity profiles. Quality control criteria are measures used to evaluate instrument performance such as signal-to-noise ratio, peak height, peak tailing, isotope ratios, and mass accuracy. Compliance with these criteria should be verified before analysis commences. (**Recommendation 10**)

120. Quality control samples should contain a range of representative chemicals spanning different concentrations and polarities because samples relevant to chemical profiling contain chemicals across a wide range of concentrations (some of which may be very low) and polarities. For example, the QC sample recently developed for chemical forensic purposes (see paragraph 53) contains a total of 27 chemicals, which were specifically selected based on their properties.[43] The OPCW, in collaboration with DLs, should evaluate the suitability of this proposed QC sample and acceptance criteria through chemical forensics confidence-building exercises (CBEs). Additionally, QC samples and acceptance criteria for other instruments such as GC-HRMS, LC-MS, and LC-HRMS should be established.

Question 8(b): How can data, methods, and procedures to conduct the in-depth CWA analyses expected in chemical forensics be standardised and shared?

121. To ensure quality, robustness, and repeatability, harmonised methods and procedures for sample preparation, instrumental analysis, and data analysis are critical prior to collating any data to be used in chemical forensics. These methods and procedures could be documented in a recommended operating procedure (ROP) and shared with DLs to develop capability across the network. A section devoted to chemical forensics in the Blue Book[44] already provides a number of ROPs for procedures such as sample matching, route sourcing, position-specific isotope analysis, QC for GC-MS analysis, and chemometrics. (**Recommendation 5**)

122. With ROPs established, interlaboratory CBEs—such as Icarus (see paragraphs 54 to 60)—would provide an excellent avenue for further validation and enhancement of the methods. These exercises could also provide opportunities to strengthen chemical forensics capabilities and capacity in participating laboratories. (**Recommendation 2**)

123. To maximise the value of CBEs, dedicated opportunities—such as workshops held in conjunction with the exercise evaluation meeting—should be created to enable participating laboratories to share their experiences in the exercises and to ensure these insights are used to refine methodologies and strengthen laboratory capabilities. The refined methods and procedures should then be shared with participating laboratories and incorporated into new editions of the ROP in the Blue Book for dissemination across the chemical forensics community. (**Recommendations 1 & 7**)

The Secretariat should promote capability building among the network of DLs in performing chemical forensic analyses of Convention-relevant chemicals. This should include engagement with specialised laboratories that are currently outside the network. The organisation of topical workshops in partnership with DLs should also be considered.

124. While the chapter on chemometrics in the Blue Book provides a general overview of chemometric tools, it was noted that it does not address their application in chemical forensic investigations. Therefore, DLs should develop a chapter on the use of classification tools specifically for chemical forensic investigations, drawing on experience from the evaluation and comparison of different classification tools in CBEs.

Question 8(c): What information is needed to ensure trust and reproducibility in the analysis and the results?

125. The level of confidence in the results depends on the proven capability of the laboratory, the validity of the adopted methods, as well as the QA and QC measures in place. While accreditation of comprehensively validated methods according to national and international standards provides the highest credibility for the results, overly restrictive guidelines may be counterproductive due the rapid pace with which techniques and methods evolve. The ability to assess DLs based on their performance in the forensic analysis of CBE and/or Proficiency Test (PT) samples is also valuable for determining the level of confidence.

126. Standards relevant to forensic science practice—including International Organization for Standardization/International Electrotechnical Commission (ISO/IEC) 17025:2017,[97] ISO/IEC 17020:2012,[98] and ISO 21043[99]—were reviewed, as well as forensic science guidelines published by the ANSI National Accreditation Board,[100] the European Network of Forensic Science Institutes (ENFSI),[101,102] the International Laboratory Accreditation Cooperation,[103] and the International Forensic Strategic Alliance.[104]

---

97 *General Requirements for the Competence of Testing and Calibrations Laboratories* (ISO/IEC 17025:2017). International Organization for Standardization and International Electrotechnical Commission, November 2017.

98 *Conformity Assessment — Requirements for the Operation of Various Types of Bodies Performing Inspection* (ISO/IEC 17020:2012). International Organization for Standardization and International Electrotechnical Commission, March 2012.

99 (a) *Forensic Sciences — Part 1: Vocabulary* (ISO 21043-1:2025). International Organization for Standardization, June 2025. (b) *Forensic Sciences — Part 2: Recognition, Recording, Collecting, Transport and Storage of Items* (ISO 21043-2:2018). International Organization for Standardization, August 2018. (c) *Forensic Sciences — Part 3: Analysis* (ISO 21043-3:2025). International Organization for Standardization, June 2025. (d) *Forensic Sciences — Part 4: Interpretation* (ISO 21043-4:2025). International Organization for Standardization, June 2025. (e) *Forensic Sciences — Part 5: Reporting* (ISO 21043-5:2025). International Organization for Standardization, June 2025.

100 *Accreditation Requirements for Forensic Testing and Calibration* (AR 3125). ANSI National Accreditation Board, January 2023.

101 *ENFSI Guideline for Evaluative Reporting in Forensic Science.* European Network of Forensic Science Institutes, 2016.

102 *ENFSI Guideline for the Use of Chemometrics in Forensic Chemistry.* European Network of Forensic Science Institutes, April 2020.

103 *Modules in a Forensic Science Process* (ILAC-G19:06/2022). International Laboratory Accreditation Cooperation, June 2022.

104 *Minimum Requirements for Crime Scene Investigation.* International Forensic Strategic Alliance, January 2021.

127. Currently, many forensic service providers seek accreditation to ISO/IEC 17020:2012 (Conformity assessment — Requirements for the operation of various types of bodies performing inspection) for crime scene investigations, or to ISO/IEC 17025:2017 (General requirements for the competence of testing and calibration laboratories) for laboratory-based activities. However, neither standard is specific to forensic science. The guidelines (see paragraph 126) provide guidance to forensic service providers implementing ISO/IEC 17020:2012 or ISO/IEC 17025:2017, supplementing these particular standards to strengthen forensic-specific practices.

128. In contrast, the ISO 21043 series published in 2025 is specific to forensic science and can be considered to be better for providing requirements and recommendations directly relevant to forensic activities, covering both investigative and evidential applications.[105] An overview of ISO 21043 standards and the key differences from ISO/IEC 17020:2012 and ISO/IEC 17025:2017 are depicted in Figure 9. Notably, ISO 21043 provides guidance on forensic requirements but is not an accreditation standard. As such, national accreditation bodies do not currently assess the competency of laboratories against this standard. However, it can be incorporated into a quality management system (QMS) to strengthen it, particularly with regard to scene procedures, evidence handling, and the interpretation and reporting of results.

129. The need for laboratories involved in international forensic investigations to operate under an overarching QMS, such as ISO/IEC 17020:2012 and/or ISO/IEC 17025:2017, is underscored. The QMS should incorporate routine QC measures—such as personnel competence, equipment calibration, lot (batch) documentation, appropriate validation and documentation of methods, and regular error analysis—and be supplemented with requirements specific to forensic science. When selecting laboratories for investigations, the Secretariat—possibly with support from the OPCW Forensic Adviser in particular—could play an important role in assessing their suitability based on forensic science-specific requirements relevant to the investigation. To facilitate this role, the Forensic Adviser should be familiar with ISO 21043. (**Recommendation 10**)

130. The procedures used in an international investigation should be technically robust and demonstrated through internal validation, international analytical or forensic exercises, and/or peer-reviewed publications. Accreditation of the specific method to be applied in an investigation is not strictly necessary. (**Recommendation 10**)

---

105 Morrison, Geoffrey Stewart, Simon Elliott, June Guiness, Lisa Sonden, and Denise Syndercombe Court. "A Guide to ISO 21043 Forensic Sciences from the Perspective of the Forensic-Data-Science Paradigm." *Science & Justice* 65, no. 5 (September 2025): 101304. https://doi.org/10.1016/j.scijus.2025.101304.

10

Laboratories involved in an international investigation should operate under an overarching QMS that ensures the implementation of regular quality control measures (including personnel competence, equipment calibration, lot (batch) documentation, appropriate validation and documentation of methods, and regular error analysis) such as those outlined in ISO/IEC 17025:2017. This system may be supplemented by forensic-specific requirements, for example through guidance provided in ISO 21043 and other relevant forensic science guidelines. Procedures applied in an international investigation must be technically robust and reproducible, as demonstrated through internal validation, international analytical forensic exercises, and/or peer-reviewed publications. Accreditation of a specific method should not be regarded as an absolute requirement for its use in a forensic investigation.

131. The Secretariat should undertake interpretation and reporting in consultation with the laboratories conducting forensic analyses. Both the Secretariat and the laboratories should be familiar with the requirement for interpretation and reporting. (**Recommendation 11**)

Figure 9: Key overlap and differences between ISO/IEC 17020, ISO/IEC 17025, and ISO 21043

Question 8(d): How can analytical data from different techniques (e.g., GC-MS, LC-HRMS, and NMR) be combined in forensic analysis?

132. As previously discussed, analytical data from multiple techniques may be essential to provide the comprehensive chemical information required to address chemical

forensic questions. The simultaneous analysis of data from several analytical techniques could help identify subtle variations and enhance understanding of how different analytical techniques contribute to the overall chemical forensic analysis.

133. Datasets from different analytical techniques are multimodal (multiple measurement types) and multivariate (many variables per technique). They are challenging to analyse simultaneously due to differences in structure and statistical metrics (for example, differing numbers of variables, means, and standard deviations). One approach is to analyse the datasets separately, then cross-validate the interpretations and conclusions. This approach has the advantages of being easy to implement and closely aligned with the current practices of DLs. However, this can be tedious as model output for each dataset has to be interpreted separately.

134. An alternative approach is to perform multi-block data analysis, in which data from different techniques are integrated and analysed together within a single multivariate model.[106] The standard single-block techniques commonly employed in chemical forensics—such as PCA or orthogonal partial least squares discriminant analysis (OPLS-DA)—cannot readily be applied to multi-block data. Specific multi-block techniques are therefore required to extract the complementary information from data generated in multiple modes. This field has been particularly active in industrial chemometrics, where large volumes of data from different sensors are analysed, as well as in the omics sciences, particularly metabolomics studies.[106] The development of interactive data visualisation tools dedicated to multi-block data analysis could benefit the chemical forensics field, and may even enable non-experts to better understand their data.[107] Examples of multi-block techniques in chemical analysis have been published, including the matching of paint samples using complementary data from two analytical techniques.[108] However, to date there has been no reported use of multi-block methods for chemical analyses related to CWAs, even though

---

106 Mishra, Puneet, Jean-Michel Roger, Delphine Jouan-Rimbaud-Bouveresse, Alessandra Biancolillo, Federico Marini, Alison Nordon, and Douglas N. Rutledge. "Recent Trends in Multi-Block Data Analysis in Chemometrics for Multi-Source Data Integration." *TrAC Trends in Analytical Chemistry* 137 (April 2021): 116206. https://doi.org/10.1016/j.trac.2021.116206.

107 Skotare, Tomas, Rickard Sjögren, Izabella Surowiec, David Nilsson, and Johan Trygg. "Visualization of Descriptive Multiblock Analysis." *Journal of Chemometrics* 34, no. 1 (July 31, 2018). https://doi.org/10.1002/cem.3071.

108 Lambert, Danny, Cyril Muehlethaler, Pierre Esseiva, and Geneviève Massonnet. "Combining Spectroscopic Data in the Forensic Analysis of Paint: Application of a Multiblock Technique as Chemometric Tool." *Forensic Science International* 263 (June 2016): 39–47. https://doi.org/10.1016/j.forsciint.2016.03.049.

laboratories have published impurity profiling studies that combine several analytical techniques.[29,109]

135. Given their benefits and the availability of software tools, multi-block approaches should be explored to capitalise on the information offered by different analytical techniques in addressing chemical forensic questions. Proven methods for orthogonal analytical data integration rely mainly on traditional multivariate statistics, as they are generally extensions of single-block methods (for example, there is a multi-block variant of PCA called MB-PCA). Recent advances in ML have led to the development of multi-block, or multi-view, analytical approaches that can integrate diverse datasets within a unified framework. Although these approaches are increasingly used in multi-omics research[110,111] and offer the potential to enhance chemical forensics capabilities, they have not yet been applied in CWA-related chemical forensics. (**Recommendation 6**)

Question 8(e): Would curated/shared database(s) of relevant reference data be useful? Are there any recommendations/restrictions to secure these types of database(s)?

Database utility and approaches for development

136. The utility of a CWA-based chemical forensics database was discussed extensively by both the individual subgroups and the TWG as a whole. The Group recognises that such a database would play a pivotal role in developing and enhancing the chemical forensics capability within DLs and the OPCW Laboratory. Approaches to developing a suitable CWA-based chemical forensics database, as well as the actions required to achieve this, were explored. (**Recommendation 12**)

137. It was considered how the two current OPCW analytical databases—the OCAD and the Validation Group Working Database (VGWD)—may be leveraged for chemical impurity identification when undertaking chemical profiling. The primary use of OCAD is to support OPCW verification activities, including on-site inspections and

---

109 Fredriksson, Sten-Åke, David S. Wunschel, Susanne Wiklund Lindström, Calle Nilsson, Karen Wahl, and Crister Åstot. "A Ricin Forensic Profiling Approach Based on a Complex Set of Biomarkers." *Talanta* 186 (August 2018): 628–35. https://doi.org/10.1016/j.talanta.2018.03.070.

110 Acharya, Debabrata, and Anirban Mukhopadhyay. "A Comprehensive Review of Machine Learning Techniques for Multi-Omics Data Integration: Challenges and Applications in Precision Oncology." *Briefings in Functional Genomics* 23, no. 5 (April 10, 2024): 549–60. https://doi.org/10.1093/bfgp/elae013.

111 Wekesa, Jael Sanyanda, and Michael Kimwele. "A Review of Multi-Omics Data Integration through Deep Learning Approaches for Disease Diagnosis, Prognosis, and Treatment." *Frontiers in Genetics* 14 (July 20, 2023). https://doi.org/10.3389/fgene.2023.1199087.

equipment testing and calibration, as well as for off-site verification analysis at DLs. It may also be used by any laboratory in a State Party, principally for the analysis of unknown samples and in OPCW PTs. The OCAD contains spectral data of scheduled chemicals, their derivatives, and non-scheduled chemicals relevant to the Convention, such as starting materials and degradation products. Prior to inclusion in the OCAD, in accordance with the mechanism proposed in EC-IV/DEC.2 (dated 5 September 1997),[112] new spectra are technically validated by the Validation Group, which comprises diverse experts from the field of analytical chemistry. Following this technical validation process, spectral data meeting the stringent criteria of the Validation Group are subsequently presented to the OPCW Executive Council for approval.[113] Only data evaluated and accepted by the Validation Group and approved by the Executive Council may be added to the OCAD.

138. The VGWD is a more comprehensive database, consisting of all spectral data found in the OCAD in addition to data relating to other chemicals that are relevant to the Convention but not directly related to OPCW verification activities (i.e., relevant chemicals that are not scheduled). While these additional data have also been evaluated by the Validation Group and meet its stringent validation criteria, they have not been approved by the Executive Council for addition to the OCAD.

139. While the OCAD contains CWA-related data, it was not created for chemical forensic purposes. Considering the content of this database and its intended end use, it was recognised that any database used for forensic purposes would need to perform a distinctly different function, helping to deduce information that can aid in sample matching, synthetic route sourcing, and precursor-product matching. Recalling Recommendation 19 from the TWG on IST—that the OPCW "should consider developing a chemical profiling database"[19]—the TWG considers that this recommendation remains highly relevant and concludes that a dedicated database platform for forensic purposes is required. (**Recommendation 12**)

140. To support the development of such a platform, the TWG proposes that, first, a module comprising a set of linked spectral databases could be established. This module would integrate existing spectral databases, such as the VGWD and NIST mass spectral libraries. Given that no existing database has been designed for, or meets the requirements of, CWA-related chemical forensics, the module would also include a new data source: the chemical forensics spectral database (CFSDB). The CFSDB should adopt a common, fixed ontology inspired by large-scale omics repositories, improving the interpretability and applicability of analytical data,

---

112 This mechanism was approved by the Conference of the States Parties at its Second Session (subparagraph 11.2(c) of C-II/8, dated 5 December 1997).

113 Note by the Technical Secretariat. Procedures for the Evaluation of Data to be Included in the OPCW Central Analytical Database (S2296/2024, dated 17 June 2024).

including raw data and data analysis workflows, for chemical forensic investigations. (**Recommendation 12**)

141. The CFSDB would complement existing spectral databases by containing spectral data for key chemical impurities not currently available elsewhere. Operationally, the CFSDB could be implemented as a standalone system or incorporated into the VGWD. Doing so would require assessing the VGWD's suitability for supporting new fields and enhancing the current search tool to accommodate specific forensic queries. (**Recommendation 12**)

142. The TWG proposes that the next step in developing an overarching chemical forensics database platform would be to create a second module: a newly curated and populated "knowledge hub". This hub would contain chemical impurity profiles, synthetic routes, a list of key chemical impurities, relevant literature references, and other metadata for each individual Schedule 1 chemical synthesis pathway, and would exist alongside the spectral databases module described above within a single overarching platform (see Figure 10). (**Recommendation 12**)

143. Acknowledging that sharing detailed synthesis routes for Schedule 1 chemicals could pose a significant proliferation risk, the Group proposes that this could be mitigated by only sharing the minimum essential data to support chemical forensics work. Additional risks and mitigations are discussed in paragraphs 153 to 156.

The OPCW should establish an overarching chemical forensics database platform for CWA-related investigations. The platform would comprise two linked modules: a spectral databases module and a knowledge hub module containing synthetic route data, a list of key chemical impurities, and relevant metadata. The spectral databases module would integrate existing spectral resources alongside a newly developed database, either standalone or incorporated into a modified Validation Group Working Database.

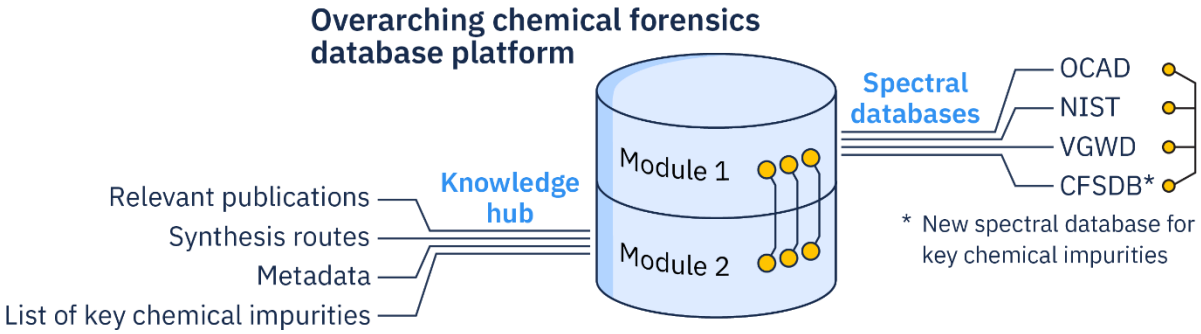


Figure 10: Overarching chemical forensics database platform and its modules

144. The interoperability of the two modules would ensure that entries are fully cross-referenced, facilitating access to spectral data for impurity identification and enabling an impurity or degradation product to be linked to a certain chemical and its production route. For example, this could include fields that note links between particular chemical impurities or precursors and the production of a Schedule 1 chemical, as well as fields listing key chemical impurities associated with specific synthesis routes. Together, the modules would function as a unified infrastructure supporting forensic queries and comparative analyses. Figure 11 shows how the database platform would fit in the overall chemical forensics workflow.

145. Within the knowledge hub, the OPCW should consider compiling and maintaining anonymised chemical impurity profiles from different synthesis routes of Schedule 1 chemicals, enabling comprehensive searches for route matching. This suggestion comes with several challenges:

    a. whether DLs can be requested to produce or provide such data, potentially in connection with OCAD spectral data production;

    b. how laboratories can provide information on the reproducibility—both intra-laboratory and interlaboratory—of analytical data;

    c. how laboratories can evaluate the influence of synthesis scale (such as microsynthesis and large-scale production) on impurity profiles;

    d. how metadata should be incorporated into the VGWD—for example, whether to use simplified molecular-input line-entry system (SMILES) notation or International Chemical Identifier Key (InChIKey), noting that InChIKey offers advantages for chemical coding; and

    e. how secure mechanisms can be established for sharing and storing sensitive data among DLs.

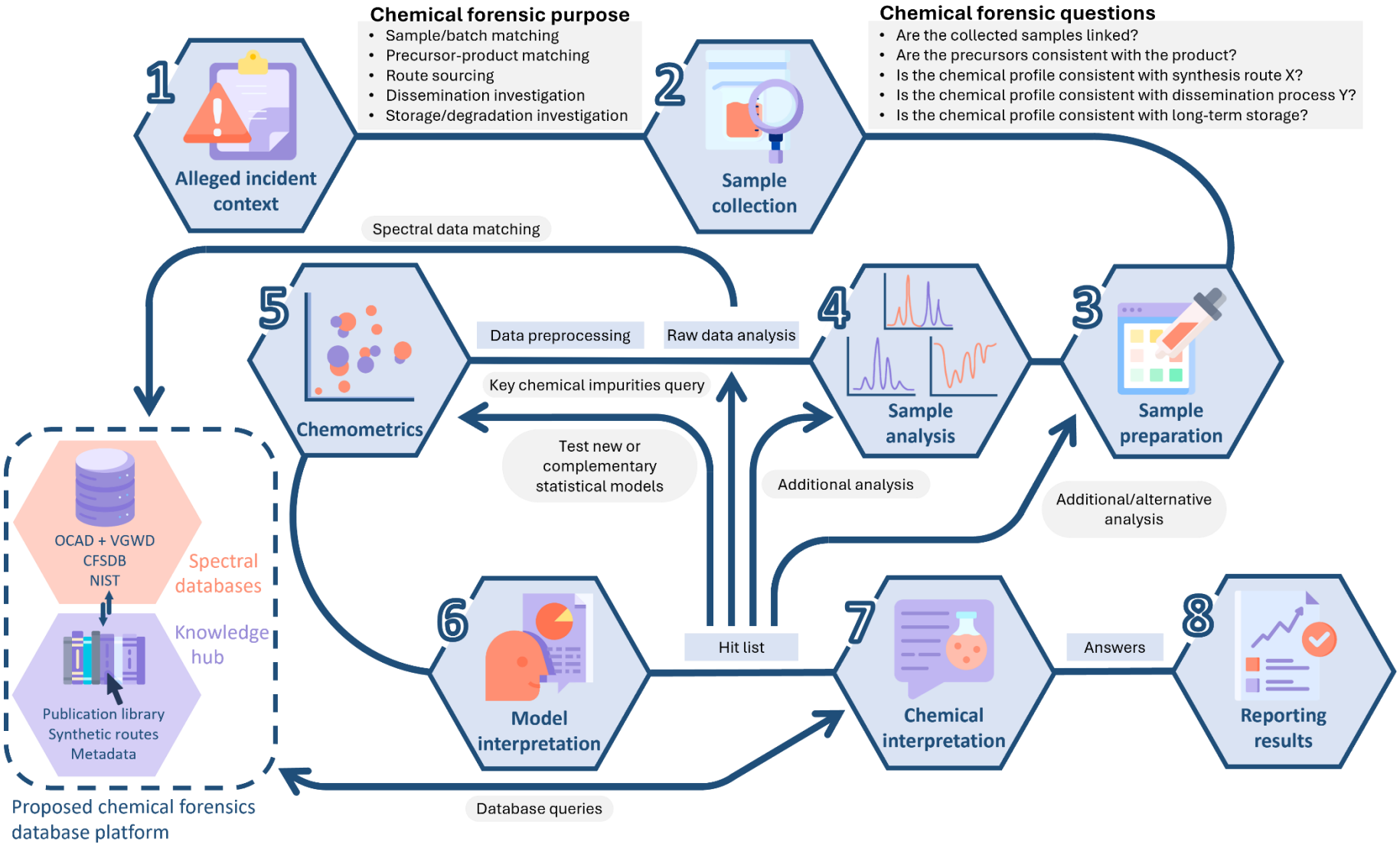


Figure 11: The integration of the database platform in the forensic workflow

Data curation and management

146. Various aspects of data curation and data management relating to a chemical forensics database were discussed. It was agreed that for the spectral data in the CFSDB and information in the knowledge hub to be useful, the curation process is critical. This is a focused, content-specific process that ensures data are accurate, relevant, organised, and usable. The first stage in this process involves selecting and validating data to ensure its quality and relevance.

147. It was agreed that data from traditional methods of chemical analysis (such as hyphenated MS) are important and would form the foundation of any forensics database. However, orthogonal data such as Raman and Fourier-transform infrared spectra, isotope ratio analysis, and elemental analysis conducted via ICP-MS or optical detection could also be valuable in creating a comprehensive database. While data would principally be acquired through off-site laboratory analysis, the potential value of data collected on-site at the scene of an incident should also be highlighted.

148. The essential role the Validation Group plays in terms of curating spectral data for the OCAD was noted. The TWG advocates the establishment of an analogous validation group responsible for defining robust criteria for data inclusion in both spectral databases and knowledge hub modules of the proposed chemical forensics database platform. These criteria are vital for the inclusion of relevant, high-quality data. (**Recommendation 15**)

149. Pertinent scientific publications were discussed,[33,34,35,36,43,46,66,114,115,116,117] and various criteria were proposed for data inclusion relating to instrumental analysis, the identification of key chemical impurities, and the selection of chemical profiles (ratio between impurities). Examples of criteria for data inclusion are found in Table 5. It

---

114 Jansson, Daniel, Susanne Wiklund Lindström, Rikard Norlin, Saphon Hok, Carlos A. Valdez, Audrey M. Williams, Armando Alcaraz, Calle Nilsson, and Crister Åstot. "Part 2: Forensic Attribution Profiling of Russian VX in Food Using Liquid Chromatography-Mass Spectrometry." *Talanta* 186 (August 15, 2018): 597–606. https://doi.org/10.1016/j.talanta.2018.02.103.

115 Williams, Audrey M., Alexander K. Vu, Brian P. Mayer, Saphon Hok, Carlos A. Valdez, and Armando Alcaraz. "Part 3: Solid Phase Extraction of Russian VX and Its Chemical Attribution Signatures in Food Matrices and Their Detection by GC-MS and LC-MS." *Talanta* 186 (August 15, 2018): 607–14. https://doi.org/10.1016/j.talanta.2018.03.044.

116 Mörén, Lina, Johanna Qvarnström, Magnus Engqvist, Robin Afshin-Sander, Xiongyu Wu, Johan Dahlén, Christian Löfberg, Andreas Larsson, and Anders Östin. "Attribution of Fentanyl Analogue Synthesis Routes by Multivariate Data Analysis of Orthogonal Mass Spectral Data." *Talanta* 203 (October 1, 2019): 122–30. https://doi.org/10.1016/j.talanta.2019.05.025.

117 Bruin-Hoegée, Mirjam de, Djarah Kleiweg, Daan Noort, and Arian C. van Asten. "Chemical Attribution of Fentanyl: The Effect of Human Metabolism." *Forensic Chemistry* 24 (June 2021): 100330. https://doi.org/10.1016/j.forc.2021.100330.

was recognised that any criteria proposed at the outset should be routinely reviewed for relevance and refined as additional data become available.

150. The TWG envisages that this new validation group would also curate published scientific data on key chemical impurities for inclusion in the knowledge hub and CFSDB, if applicable. (**Recommendation 15**)

151. The key chemical impurities associated with different synthesis pathways for a Schedule 1 chemical would be based on published studies with adequate replication of both the synthesis and the analyses. The introduction of two levels of confidence for impurities is considered to be a valuable enhancement. The first level could be designated "presumptive", based on data from a single laboratory, and the second "confirmed", based on data from two or more laboratories. Examples within the scientific literature that could be incorporated into such a database include reports on VR[33,114] and sulfur mustard.[34] Relevant metadata information for constructing a list of key chemical impurities, such as data acquisition parameters, data analysis, raw data preprocessing workflows, and chemometrics, should also be included. (**Recommendation 13**)

152. Population of the knowledge hub was considered, and relevant methods and procedures for chemical forensics were discussed. The knowledge hub would establish a critical shared resource of reference materials, curated from open-source profiling data as well as outputs from collaborative projects specifically designed to generate data for inclusion.

Table 5: Examples of criteria for data inclusion in the proposed chemical forensics database

Phase	Category	Criteria
Chemical Analysis – Data Collection	Instrument performance to ensure data reliability	• Analysis of blanks (process blanks, if applicable) and QC samples before, during, and after analysis of samples to ensure analytical performance and no carry-over • Analysis of QC mix after every 10th to 15th sample analysed and fulfilment of QC criteria
	Reproducibility & repeatability	• At least three reference samples per source / route (different days and chemists), duplicate per synthesis • At least three replicate injections per sample
Peak Table Construction – Data Processing and Tabulation	Peak detection	• Based on signal-to-noise ratio (range from 5 to 20) • Based on peak area of extracted ion chromatogram of base peak (e.g., 30,000) • Combination of both peak area and signal-to-noise ratio
	Unique chemical constituents	• Compared with blanks => peaks not in blanks • Peak matching through retention time alignment or use of retention index
	No. of replicates constituents detected in	• Depends on the signal-to-noise ratio of the peak, for example, less stringent criteria for trace components
	Identification of chemical constituents	• Known: based on minimum match factor (≥ 80) and retention index (± 20), mass accuracy (5 ppm) for HRMS, if possible • Unknown: Full identity is not required but will strengthen the evidence if backed by science
	Table output	See Figure 12
	Selection of chemical constituents through chemometric feature selection	• Constituents that differentiate reference samples from each other • Can also be based on relative analytical signals or amounts of a constituent in different reference samples • For example, Fisher-ratio or degree-of-class separation

Peak ID	For identified constituents				For non-identified constituents	RT/RI	(Relative) intensity
	Name	CAS Registry Number	Match factor	Relevance (if any)	mass/charge ions (1 quantitative & 1 confirmation)		

Figure 12: Proposed table output (CAS= Chemical Abstracts Service; RT=retention time; RI=retention index)

153. Regarding the synthesis of scheduled chemicals, the TWG notes that a significant pool of ground-truth data resides within the DL network, which is vital for the creation of a comprehensive chemical forensics database. During discussions on the chemical forensics database platform, two principal data sensitivity concerns were considered at length, and potential solutions were proposed. First, there is a recognised reluctance among States Parties to share raw data relating to Schedule 1 chemicals. This stems from concerns that raw data may reveal unique, laboratory-specific signatures that could lead to the identification of the originating laboratory and associated State Party. These risks however can be mitigated if multiple laboratories across different States Parties synthesise the same Schedule 1 chemical using a particular route.

154. The TWG concurs that implementing a forensics database based on raw data would be challenging until these concerns are addressed. Consequently, the Group proposes that the initial database platform should exclude raw data, with the possibility of including it at a later stage once robust data protection measures are in place.

155. As noted in paragraph 143, sharing sensitive data could also pose a proliferation risk by revealing the synthesis routes of Schedule 1 chemicals. It is therefore necessary to strike a considered balance between making valuable information available and preventing proliferation. Other possible concerns of the contributing laboratories, including accessing and managing the modules of the overarching chemical forensics database platform, were also considered.

156. To address some of these concerns, the TWG underscores the importance of appointing the OPCW as the trusted forensic data custodian to encourage contributions from laboratories. To mitigate proliferation risks, mechanisms to facilitate sharing these ground-truth data with the OPCW in a secure manner should

be explored. One option could be to use blockchain-based technologies to anonymise these entries for inclusion into the knowledge hub. A framework for sharing sensitive data between laboratories within the DL network that will satisfy States Parties needs to be established. (**Recommendation 14**)

The OPCW, in collaboration with DLs, is encouraged to define a common framework—including criteria and procedures—to anonymise data relevant to chemical forensics, reducing security risks while supporting database development.

157. The TWG also proposes that the new validation group should establish a framework for data management and security (such as data anonymisation), in addition to facilitating communication between DLs in the initial chemical forensics database construction phase, promoting contributions from these laboratories. (**Recommendation 15**)

The OPCW should establish a validation group responsible for defining a framework for data management and setting criteria for the inclusion and curation of data within the proposed chemical forensics spectral database and knowledge hub. This group could also facilitate communication between DLs during the initial construction phase of the database, promoting contributions from these laboratories.

158. To assist in the design of the chemical forensics database platform and in all aspects of the data curation requirements—especially for the knowledge hub—the OPCW should recruit a data scientist. The TWG considers that knowledge and expertise in AI/ML would be particularly beneficial to this role. (**Recommendation 17**)

The OPCW is encouraged to recruit a data scientist with expertise in ML, chemometrics, and scientific data management to support the integration of data-driven approaches into the Organisation's activities, including the development, curation, and governance of the proposed overarching chemical forensics database platform and its modules.

Database content and data generation

159. Despite the concerns around sharing raw data, it was proposed that a list of possible key chemical impurities that meet an agreed set of criteria—along with the analytical methods and conditions under which the data were produced—should be included in the knowledge hub. To start populating the knowledge hub while the data management framework is being set up, the TWG proposes that data could be mined from the existing scientific literature, in addition to any data that DLs are willing to share with the OPCW. The origin of the data in the knowledge hub would also be included—specifically whether the data are experimental or generated using AI methods. (**Recommendation 14**)

160. Designated laboratories and the OPCW Laboratory should investigate whether *in silico* dataset expansion by AI methods is feasible, reliable, and relevant to chemical forensic analyses (**Recommendation 16**). This includes:

a. the generation of physiochemical and spectral data using AI methods to compensate for the lack of experimental data and support the population of large-scale databases;

b. the expansion of experimental impurity datasets using known, real samples to increase statistical power during data analysis; and

c. the generation and prediction of synthesis pathways, including an evaluation of the potential threats posed by AI methods from a chemical forensics perspective, whether in terms of proliferation risks or regulatory circumvention.

Designated Laboratories and the OPCW Laboratory are encouraged to investigate the feasibility, reliability, and relevance of expanding chemical forensic datasets using in silico data generated by AI/ML methods, including physicochemical properties, spectral data, impurity datasets, and potential synthesis (by-)products. Where predicted data are incorporated into chemical forensics databases, they should be clearly identified as in silico entries and efforts should be taken to validate them.

161. To address the limited information available on CWA-based chemical forensics in the current scientific literature, proposed solutions include exploring open-source or AI-predicted synthesis routes, and conducting syntheses at the OPCW Centre for Chemistry and Technology (the ChemTech Centre) to compare AI-generated and experimental data for chemical profiling.

162. To further augment the generation of a chemical forensics database, it is strongly recommended that the OPCW instigate CBEs as a useful mechanism to both generate data for inclusion into the knowledge hub and CFSDB, and broaden the skill set within the DL network. (**Recommendations 2 & 14**)

163. In support of this recommendation, the preparation of samples for CBEs, including the synthesis of Schedule 1 chemicals, could be undertaken either at the ChemTech Centre, or on a voluntary basis by States Parties that have a single small-scale facility. If coordinated by the ChemTech Centre, any potential concerns by a State Party regarding revealing the synthesis routes of Schedule 1 chemicals would be reduced. The syntheses of Schedule 1 chemicals may not only include different routes but also elicit variations within a particular route, including reaction conditions, matrices, and production scale, for incorporation into chemometric models and the knowledge hub. It is essential that information regarding these synthesis routes is handled appropriately to mitigate any proliferation concerns.

164. The TWG proposes that laboratories participating in CBEs could receive samples of Schedule 1 chemicals and guidelines on the sample preparation and chemometric methods of analysis, as well as the criteria for data inclusion. These laboratories would then be required to provide a list of the key chemical impurities in line with predetermined criteria for reporting. It was noted that two different types of CBE could be explored: dry exercises, which are data-driven and rely on simulated or pre-existing datasets, and wet exercises, which involve laboratory-based activities using actual chemical samples. Each approach has distinct advantages and limitations that should be considered when determining the appropriate type of CBE. (**Recommendation 2**)

165. To validate the chemical forensic data obtained from any such CBE and confirm the key chemical impurities identified, the ChemTech Centre and/or single small-scale facilities could produce the same scheduled chemical using harmonised synthetic routes.

Question 8(f): What best practices exist for on-site sampling to ensure the validity of subsequent forensic laboratory analysis and what challenges/gaps remain?

166. An on-site sampling protocol is required to ensure the validity of forensic analyses. Important considerations include determining who is authorised to collect samples and ensuring that background samples are collected to account for matrix effects, which can influence impurity levels. The requirements for sample collection are set out in ISO/IEC 17020:2012 and, as noted earlier, could be supplemented with more forensic science-specific requirements from ISO 21043-2:2018.

167. The Secretariat should determine the requirements needed to ensure that chemical forensic sample methodologies are included in the R/SOPs and working instructions, including those related to on-site sample collection, handling, curation and storage, and annotation, in accordance with forensic best practices. The valuable experience gained through contingency operations in the Syrian Arab Republic should be leveraged. Additionally, the R/SOPs should regularly be reviewed and updated. (**Recommendation 9**)

The Secretariat should ensure that chemical forensic requirements are incorporated into R/SOPs and working instructions for on-site sample collection. In doing so, relevant experience from the contingency operations in the Syrian Arab Republic should be leveraged.

168. In the event of a chemical weapons incident, the national response forces in the affected State Party will respond first. It is critical that these agencies incorporate forensic requirements into their R/SOPs to ensure that the samples collected are sufficient and valid for forensic investigations. This was also discussed by Subgroup 4 (see paragraphs 176 to 178).

169. OPCW sampling R/SOPs and best practices should be promulgated to national response forces through capacity-building activities to ensure that collected samples fulfil forensic requirements. Beyond initial training, it is also important that the competency is continually maintained through regular national and/or international exercises. Such exercises also provide opportunities to share best practices, identify areas for improvement, and harmonise sampling protocols. (**Recommendation 3**)

Subgroup 4: Augmenting the OPCW's capabilities

170. Subgroup 4 considered how the OPCW's capabilities in chemical forensics could be augmented and identified steps that could be taken. Discussions, supported by the invited expert speakers, explored how forensic science can be effectively integrated into investigations of alleged chemical weapons use. These discussions reflected a deep engagement with the scientific, operational, and strategic dimensions of chemical forensics. The subgroup approached the topic through scenario-based analysis, identifying a range of realistic situations—from small-scale terrorist attacks to mass-casualty events, lone-wolf incidents, and hybrid threats—that illustrate the diverse challenges and opportunities in applying chemical forensic methods.

171. With these scenarios in mind, the following five questions from the TOR were considered:

i. **Question 9(a)** How can the OPCW enhance its capability to capture and utilise chemical forensics-related information in the context of alleged use?

ii. **Question 9(b)** How will the OPCW ensure the validity of the forensic results?

iii. **Question 9(c)** How will the forensic results be reported? (criteria, level of confidence)

iv. **Question 9(d)** Are there other organisations or information that could be leveraged to augment the capability of the OPCW?

v. **Question 9(e)** How can forensic analysis be promoted and enhanced at DLs? How might other organisations or laboratories contribute?

172. Across its discussions, the subgroup identified several cross-cutting themes relevant to augmenting the OPCW's chemical forensics capabilities. These included the need to enhance collaboration within and beyond the DL network, to develop tailored forensic reporting criteria, and to strengthen training and exercise design. These themes are addressed in detail under the relevant TOR questions below.

173. A central point underpinning many of the discussions was the need to distinguish chemical forensics from verification analysis (Figure 13). While verification seeks to confirm the presence of toxic chemicals, chemical forensics aims to address the important forensic questions related to the chemical agent by leveraging processes such as sample/batch matching, precursor-product matching, and route sourcing. This requires a broader analytical toolkit, including chemometrics expertise and a forensic interpretive framework.

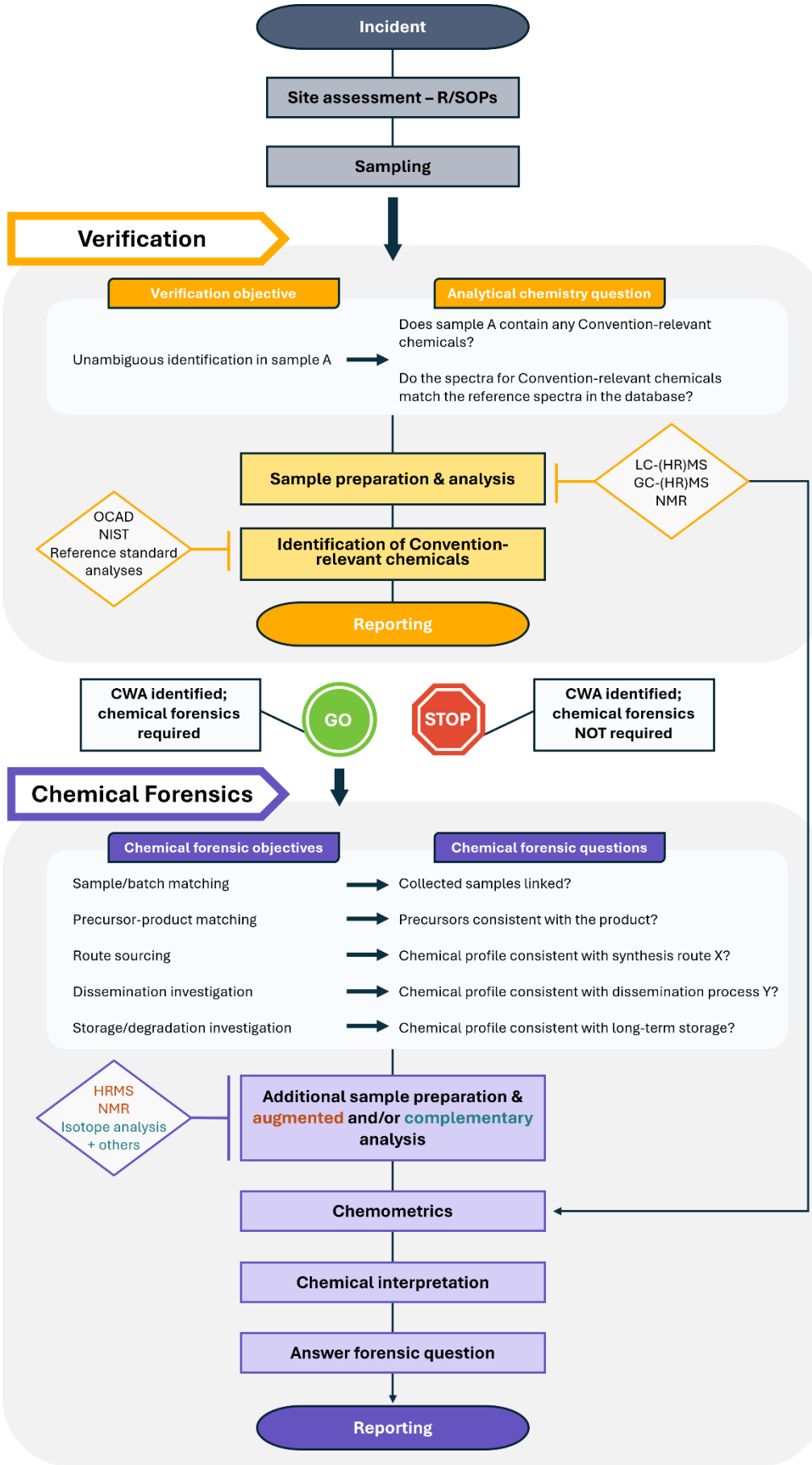


Figure 13: Overview of CWA-related verification and chemical forensics

Question 9(a): How can the OPCW enhance its capability to capture and utilise chemical forensics-related information in the context of alleged use?

174. During their work in the Syrian Arab Republic, the FFM and IIT developed substantial knowledge, experience, and expertise relevant to chemical forensics. With the formation of the OSM, it is imperative that the Secretariat ensure these capabilities are transferred in full. Furthermore, the OPCW should implement measures to maintain these capabilities and should seek to apply them more broadly to ensure that all relevant information may be leveraged to assist with an investigation. As chemical forensics approaches may increasingly support OPCW assistance to States Parties, they should be considered as an option from the outset, as applicable.

175. The importance of a collaborative mindset among DLs was emphasised, as well as the need for increased communication between DLs and the Secretariat during all investigations involving chemical forensics. Having access to additional context on the purpose of the analysis—including the suspected CWA and any specific chemical forensic questions—would enable DLs to carry out their work more efficiently and effectively, providing targeted support. Facilitating open communication ensures that forensic questions are continually reviewed and refined, thereby minimising redundant efforts such as re-identifying chemicals previously confirmed during verification analysis. In this context, the option of extending the secrecy agreement between DLs and the Secretariat to include this possibility was considered. Open lines of communication should also exist between those DLs and non-DLs undertaking the chemical forensic analysis for the same reasons. (**Recommendation 8**)

When chemical forensic approaches are applied to authentic samples, the Secretariat should maintain an open and continuous line of communication with the DLs involved. This will ensure clarity on the forensic questions and the application of the most appropriate techniques, approaches, and methods to generate the most relevant information for the investigation.

176. There are often features in a forensic case that limit what can be analysed, concluded, and reported. This may be compounded by the actions of first responders—who may be first at an incident scene—and the sampling strategy. It was highlighted that situational awareness in an alleged use case is of utmost importance and is challenging in multiple ways. Operational challenges arise from variability in incident-site response depending on the circumstances. Coordinating the response at the scene requires harmonised protocols among first responders, law enforcement, investigators, and other authorities, as well as ensuring that medical personnel

attending to casualties maintain situational awareness and understand how their actions may impact the integrity of forensic evidence.

177. In this context, the scenarios considered by the subgroup (see paragraph 170) are of particular importance. Although the relevant chemical forensic methodologies were determined to vary little according to the scenario, the incident response can vary significantly, causing a range of challenges. Deviations from sampling SOPs, sampling bias, or the absence of a clear sampling plan could compromise evidential value. Additionally, the response may lead to disturbance or irreversible alteration of the scene, leading to potential loss of probative material. Other risks include chemical exposure of first responders, inappropriate decontamination practices, and medical interventions that are not informed by forensic considerations.

178. It is important to note that every CWA-related incident begins with identification and verification activities, which may require different sampling procedures. Therefore, sampling procedures should explicitly incorporate requirements for chemical forensic purposes to ensure that appropriate sampling strategies are embedded within R/SOPs. (**Recommendation 3**)

The OPCW should promote, through its capacity-building activities, forensics-based approaches to investigations of alleged use of chemical weapons and the continued coordination of national and regional chemical forensics capabilities. This may include forensic awareness training for first responders, crime scene investigators acting in a potentially contaminated environment, and laboratory personnel. Consideration should also be given to best practices and procedures for on-site sampling.

Question 9(b): How will the OPCW ensure the validity of the forensic results?

179. Traditionally, as part of the existing OPCW identification and verification workflow, a sample preparation and analysis ROP is considered validated following the successful completion of the following steps:

a. a known chemical is spiked into a particular sample matrix (solid, liquid, or swab);

b. the sample is subjected to the corresponding sample preparation ROP;

c. chemical analysis is conducted, and the spiked chemical is identified using the OCAD; and

d. the identity of the spiked chemical is chemically verified through comparison with an authentic reference standard.

180. In chemical forensics, the process for confirming the presence of a key chemical impurity is more complex. While it is not possible to have authentic reference standards for all components of a chemical profile, the TWG underscores that the availability of a chemical forensics database would greatly assist in the confirmation of forensic results. The need for a chemical forensics database was discussed in depth, and it is the considered opinion of the TWG that all necessary actions should be taken to achieve this, while balancing non-proliferation concerns. There was also extensive discussion on the appropriate form of the proposed database and the types of information it should include. These discussions are reported under Subgroup 3, see paragraphs 136 to 165.

181. To ensure the validity of forensic results, the importance of DLs and the OPCW Laboratory operating under an overarching QMS is reiterated. It was also noted that their performance can be assessed through the forensic analysis of CBE and/or PT samples, and that procedures can be validated through peer-reviewed publications (see paragraphs 125 to 130). (**Recommendation 10**)

182. Furthermore, harmonised criteria to guide the reporting of forensic findings will also be crucial for helping ensure the validity of forensic results.

Question 9(c): How will the forensic results be reported? (criteria, level of confidence)

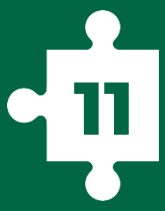
183. The OPCW Laboratory has well-established criteria for reporting analytical results in verification activities, namely the use of two independent, information-rich methods, typically supported by comparison with an authentic reference standard. However, this approach is less straightforward in chemical forensic analysis. Compounds relevant to such analysis are often present in trace quantities, and the amount of sample available for preparation may be limited. Consequently, it may only be feasible to employ a single analytical technique. Nevertheless, a single technique can still provide valuable support for the forensic investigation.

184. An additional challenge concerns the availability of authentic reference standards for comparative purposes, given the nature of many relevant compounds (for example,

plasticisers and stabilisers). Reporting criteria for chemical forensics should therefore be developed with these factors in mind.

185. Designated laboratories must therefore be made aware that the requirements for chemical forensic analysis are different from those for verification activities. For forensics, DLs are expected to provide all chemical information they consider relevant, clearly specifying the associated degree of certainty. Unlike during PTs, the performance score of laboratories would not be adversely affected during evaluation if they were unable to confirm the identity of a compound using two independent identification techniques. However, when analytically feasible, the use of two techniques remains desirable, as it increases the confidence level regarding the presence of an impurity.

186. The TWG encourages the OPCW to harness the expertise available within the DL network to help frame and establish reporting criteria for chemical forensic results. These criteria should be practical, proportionate, and readily achievable by the DLs. Elements to be considered include validity, uncertainty, and measures of confidence (for example, tentative or confirmed, LRs, and expert assessments). Confidence-building exercises could provide a valuable opportunity to test, further develop, and refine these reporting criteria. (**Recommendation 11**)



The OPCW Laboratory, in partnership with DLs, should define criteria for the reporting of results of chemical forensic analyses. A working group would be one way to support the development of these criteria. Confidence-building exercises could also be used to further develop and refine them. The criteria should include the methods used, key chemical forensic data, and the associated level of confidence.

187. An assessment of the evidential value and significance of findings derived through the application of ML methods is critical when such results are intended for use in a court of law. While LRs are commonly used in other forensic disciplines to evaluate evidential value through statistical testing (see Annex 2), no directly comparable method is currently available for CWA chemical forensics. The ENFSI Guideline for Evaluative Reporting in Forensic Science notes that such metrics should only be used to "assist the court in relation to the strength of the findings" during discussion with experts.[101] The ENFSI Guideline also addresses the presentation of results in court, including both statistical evaluation and expert chemical interpretation, as well as appropriate wording associated with different levels of confidence. Although this approach cannot yet be readily applied to chemical forensics involving CWAs, developments in this area should continue to be closely monitored.

188. Clear and specific reporting criteria that address the proposed chemical forensic questions is important, as well as the need to balance scientific rigour with legal and operational clarity.

Question 9(d): Are there other organisations or information that could be leveraged to augment the capability of the OPCW?

189. As already discussed, the TWG considers non-DLs to represent an additional source of relevant expertise that should be leveraged. Drawing on the expertise of non-DLs is viewed as an asset.

190. The current capabilities available across the DL network were assessed through a survey that was sent out by the OPCW Laboratory. Unsurprisingly, the DL network demonstrates strong proficiency in MS-based techniques and basic NMR spectroscopy, however other capabilities (such as PSIA-NMR, ICP-MS, IRMS, metals analysis, explosives analysis, and advanced chemometrics and AI-assisted analysis) are limited and lack sufficient coverage. It was recognised that while many requisite skills already exist within the DL network—though they may not always be fully utilised—certain specialised techniques from non-DLs will also be needed to bolster capabilities. Furthermore, non-DLs may have access to the larger sample quantities required for developing statistical analysis methods, which could be valuable in the context of CWA forensics.

191. The possibility of developing a roster of non-DLs with specialised analysis capabilities that address analytical gaps in the DL network was explored. This roster would be supported by clear, standardised criteria to guide both the selection process and the framework for engagement, ensuring transparency and consistency while fostering collaboration across diverse expertise. These non-DLs should also be aware of the specific reporting criteria for forensic analyses that is required by the OPCW.

192. The ongoing Icarus exercise already involves non-DLs providing analytical expertise typically beyond DLs' existing capabilities. This approach should be considered for future CBEs and could inform the development of the roster. Regional collaboration and outreach were emphasised as critical for building a resilient and responsive chemical forensics capability.

193. The OPCW's capabilities could also be augmented by leveraging the expertise and resources of other international organisations with a mandate to conduct research and build capacity in CWA forensics, such as the International Criminal Police Organization (INTERPOL)[118] and the United Nations Interregional Crime and Justice

[118] "Chemical and Explosives Terrorism." INTERPOL, accessed December 15, 2025.

Research Institute (UNICRI).[119] Collaboration among such global law enforcement agencies, and the sharing of forensic intelligence, could provide valuable insights and enhance the OPCW's operational framework. It is also important to engage with existing initiatives and groups working in the field of chemical forensics, such as the CFITWG,[19] and to identify common ground with other forensic science networks, including fields that utilise similar analytical techniques. Figure 14 shows key stakeholders that could help augment the OPCW's chemical forensics capabilities. (**Recommendation 4**)

The OPCW is encouraged to leverage relationships with international organisations, such as INTERPOL and UNICRI, as well as initiatives like the CFITWG, to incorporate forensic-based approaches into CWA investigations. A capabilities gap evaluation of the network of DLs should also be conducted to complement the expertise that already exists in dedicated forensic laboratories.

119 "Prosecution and Adjudication of CBRN Crimes." UNICRI, accessed December 15, 2025.

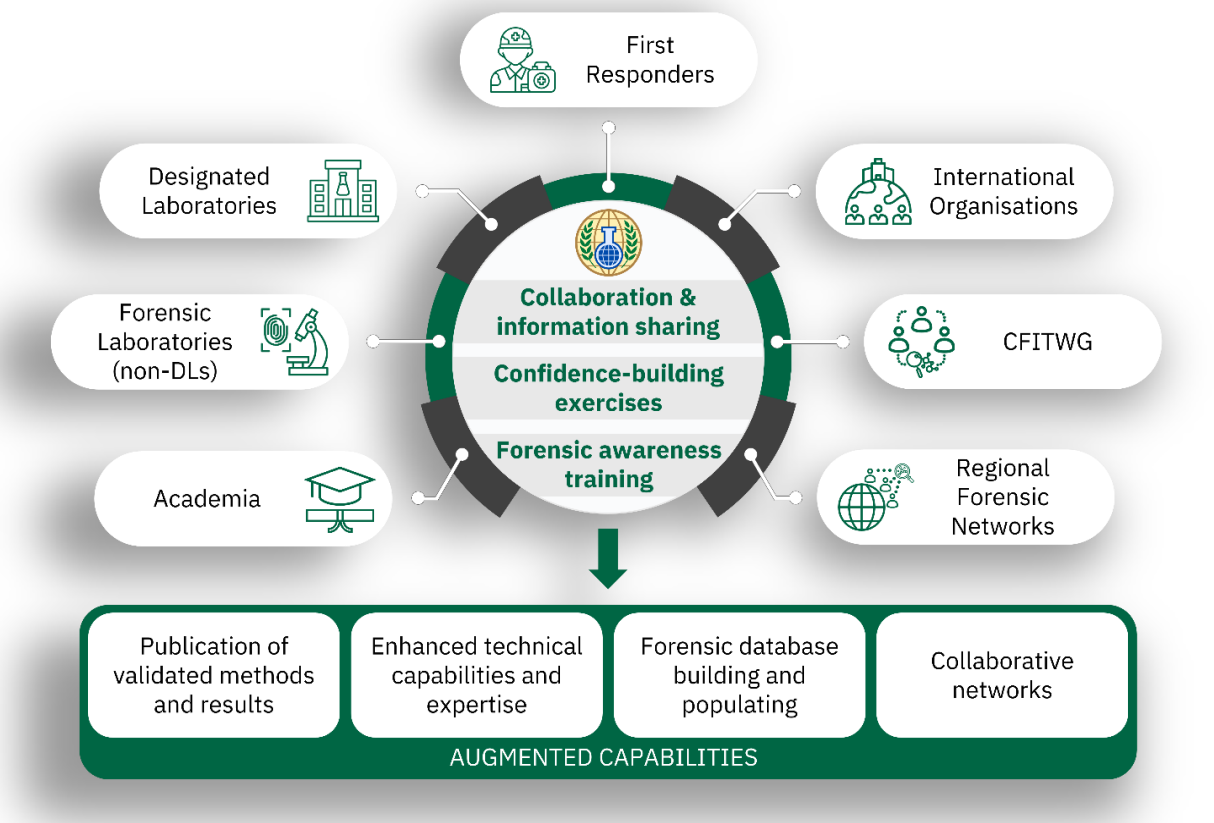


Figure 14: Stakeholders that could help augment chemical forensics capabilities

Question 9(e): How can forensic analysis be promoted and enhanced at DLs? How might other organisations or laboratories contribute?

194. Confidence-building exercises and other chemical forensics tests are essential for identifying gaps in the capabilities of DLs and for building capacity in chemical forensic techniques. The current OPCW PT Programme, which DLs undertake to obtain and maintain designation, is the cornerstone of the verification regime and must be preserved, as anonymous identification in independent laboratories is essential. However, a team-focused approach, in which laboratories combine expertise and complementary techniques, should be considered to enhance chemical forensic outcomes. For example, following a verification activity, several laboratories could collaborate on chemical forensic analyses, with additional laboratories involved as needed to provide specialised expertise. A series of forensic questions could be posed, with the laboratory teams reporting back their analyses and findings.

195. Current exercises, such as the Icarus wet exercise and the OPCW-coordinated dry exercise, are crucial for developing the requisite skills in chemical forensics.

Confidence-building exercises can also serve as valuable training opportunities, allowing laboratories with established expertise to support those seeking to expand their capabilities. While this approach is currently being tested in the Icarus exercise, results from any CBE would also contribute to the development and testing of the proposed chemical forensics database platform, including the associated knowledge hub. It is therefore important that future CBEs encompass the entire chemical forensic workflow in their design. (**Recommendations 2, 13, & 14**)

It is imperative that the OPCW obtain chemical forensics-related data to populate the proposed chemical forensics database platform and support ongoing and future investigatory efforts. This can be achieved via multiple, non-competing approaches, including:

a. *encouraging States Parties to share chemical forensics data on scheduled chemicals; and*
b. *making use of CBEs focused on chemical forensics.*

196. During CBEs, the Blue Book methods for chemical forensic analyses and chemometrics should be widely tested and improved or augmented as required.[44] It is crucial that these ROPs remain up to date to support the harmonisation of analytical methods across laboratories, as recommended for chemical forensics. In the Icarus exercise, participating laboratories should apply, where applicable, Blue Book chemical forensic ROPs, although these currently cover only GC-MS and NMR analyses. Whenever possible, DLs should seek to publish their analytical and chemometrics methodologies, as well as harmonisation efforts, thereby increasing the information available in the scientific literature.[36,43,120] (**Recommendation 7**)

197. Future CBEs should assess and report on the usability of the proposed chemical forensics database. They should be designed to be as realistic as possible to maximise their effectiveness and to identify capability gaps that need to be addressed, ensuring that DLs are ready to respond when chemical forensics is required.

198. The additional capabilities required by existing DLs—beyond those needed for verification—should be explored in a structured manner. As noted previously, many

120 Säde, Solja, Martin Härkönen, Leena Kalliovirta and Hanna Hakulinen. "Comparison of Statistical Multivariate Analysis Methods for Chemical Forensics Profiling of a Carbamate Chemical Warfare Agent Precursor." *Forensic Chemistry* 46 (December 2025): 100700. https://doi.org/10.1016/j.forc.2025.100700.

of these analytical capabilities—such as PSIA-NMR, ICP-MS, and IRMS—already exist to a certain extent within the network but are utilised in different contexts.

199. Practical approaches may be leveraged for promoting chemical forensics within the DL network. For example, a roadmap for the technologies required to advance chemical forensics—including analytical techniques, data analysis tools, statistical methods, and reporting guidelines—could be developed. First, the key steps should be defined, such as identifying required actions and corresponding achievable milestones. Next, required techniques and methods must be explored and developed, ensuring they are tested and refined. Finally, methods should be harmonised and shared to enable future implementation and practical use across the chemical forensics framework.

200. A well-defined roadmap, developed and promulgated by the OPCW, will play a critical role in supporting method development and capacity building across the chemical forensics network. It will also facilitate access to funding opportunities, which are essential for advancing technology development, implementing new methods, and ensuring sustainable growth in analytical capabilities. Furthermore, such a roadmap could serve as a strong motivator for States Parties to allocate resources for building capabilities in chemical forensics. (**Recommendation 19**)

The OPCW should help identify funding streams to support interdisciplinary chemical forensic research and development, enabling DLs to collaborate on common syntheses, analytical work, and data science to drive knowledge acquisition, data generation, and the verification or enhancement of forensic results.

201. OPCW investigations of alleged use of chemical weapons often need to go beyond simply verifying the presence of scheduled chemicals. In this context, the existing PT regime—including scenarios and scope of analyses (such as sensitivity and range of chemicals)—could be reviewed and adapted to ensure it remains fit for purpose and responsive to the evolving needs of the OPCW. This expansion beyond basic verification increasingly relates to the analysis of authentic samples, including chemical forensic investigations. Incorporating an optional chemical forensics component into the PT construct would allow DLs that choose to participate to develop broader capabilities, enhancing their support for OPCW investigations. The OPCW could consider how laboratories that successfully analyse chemical forensics samples and demonstrate proficiency could be recognised. (**Recommendation 2**)

The OPCW should promote, organise, and leverage current and future CBEs to strengthen the skill set of the DL network. As an initial step, this could include the provision of supplementary samples with defined chemical forensic questions into the existing PT Programme.

202. The TWG proposes that the central verification element of PTs remains, but modifications should be made. There are several options that should be considered when undertaking this review. For example, the feasibility of incorporating a scenario—aligned with the current global context—that has fewer chemicals to verify and includes chemical forensic questions (such as sample/batch matching, precursor-product matching, or route sourcing) should be assessed. Dedicated chemical forensic samples—independent of verification samples—will be required.

203. Alternatively, the existing environmental PTs could be adapted to incorporate chemical forensic components. Possible approaches include:

a. providing additional chemical forensic samples that DLs may choose to analyse, with subsequent recognition for successful performance;

b. shipping separate samples to participating laboratories as part of a dedicated CBE exercise; or

c. offering alternating chemical forensic PTs to DLs that have a long-standing record of PT success.

---o---

ANNEX 1: Glossary of Terms

This glossary provides definitions of key technical terms that are used in this report, with illustrative examples of potential relevance to the activities of the OPCW and the Convention. These examples are provided for context and understanding, and do not constitute recommendations for implementation.

Disclosure: This glossary was developed with the assistance of AI technology to help ensure accuracy, clarity, and comprehensiveness of the technical definitions provided.

Term	Definition
Accuracy (metrics)	The number of correct classification predictions divided by the total number of predictions.
Algorithm	A step-by-step set of instructions that tells a computer how to solve a problem or complete a task. AI algorithms are special because they can adapt and improve their performance by learning from data, rather than following the same fixed steps every time.
Application programming interface (API)	A set of rules and protocols that allows different software systems to communicate and exchange information with each other. APIs enable applications to request services, data, or functionality from other systems without needing to understand their internal workings, supporting interoperability and automation.
Artificial intelligence (AI)	Technology that uses mathematical and statistical methods to enable computers to perform tasks that normally require human thinking, such as recognising patterns, understanding language, and making decisions. Unlike traditional computer programs that follow pre-written rules, AI systems analyse data to discover relationships and make predictions based on what they learn.
Bias (algorithmic)	Systematic errors or unfair outcomes in AI systems that may produce skewed results. Bias can emerge from training data, algorithm design, or deployment practices, potentially affecting the reliability and fairness of AI applications in verification and monitoring activities.

Binary/multiclass classification	Binary classification is a type of classification task in which a model predicts one of two mutually exclusive classes. Multiclass classification extends this classification task to three or more distinct categories, assigning a sample to a single class from a wider set of possibilities.
Chain of custody	The chronological documentation of the acquisition, transfer, handling, and disposition of physical or electronic evidence.
Chemical attribution signatures	Residual intrinsic and/or extrinsic information that can provide chemical information to aid in attributing the use of a chemical of interest.
Chemical forensics	A specialised branch of forensic science that applies analytical chemistry, synthetic chemistry, and chemometric methods and techniques to examine chemical evidence for investigative purposes.
Chemical profile	Chemical and/or elemental signatures, which can be used to obtain information about the potential source of a chemical sample and/or its method of synthesis. The profile may include by-products, impurities, and unreacted starting materials found in the sample.
Chemometrics	The application of statistical methodologies to analytical chemical datasets for the purpose of establishing and visualising trends, links, and differences between samples.
Classification model	A machine learning model that predicts the category or class of an input by learning patterns from labelled training data.
Confusion matrix	A table used to evaluate the performance of a classification model by comparing actual classes against predicted classes. In binary tasks, it tracks true/false positives and negatives. In multiclass tasks, it forms a grid where the diagonal represents correct identifications, and the off-diagonal cells highlight specific misclassifications between different substances.

Contingency operations	Specialised missions conducted by the OPCW, such as the Fact-Finding Mission and the Investigation and Identification Team, deployed to respond to allegations of chemical weapons use or threat of use. These operations include investigations of alleged use, challenge inspections, technical assistance visits, activities related to the accession of new possessor States, and requests by States Parties for assistance, protection, and any other type of investigation pursuant to the Convention.
Convolutional neural network	A type of deep learning model particularly suited for analysing spatial or grid-like data, such as images or spectroscopic maps. Convolutional neural networks automatically detect patterns, features, and correlations in complex datasets, making them useful for tasks like spectral classification and chemical imaging analysis.
Cross-validation	A mechanism for estimating how well a model would generalise to new data by testing it against one or more non-overlapping data subsets withheld from the training set.
Data augmentation	Artificially boosting the range and number of training examples by transforming existing examples to create additional examples.
Data curation	The process of actively selecting, evaluating, and transforming raw information from various sources into reliable, well-documented datasets that can be discovered and used for future applications. Data curation involves making decisions about what information is valuable, verifying its accuracy, preserving the context of how it was originally collected or presented, and organising it at the appropriate level of detail to support different analytical needs and use cases.
Data management	The operational implementation of data governance policies through the practical activities of collecting, storing, organising, maintaining, and utilising data throughout their lifecycle. Data management involves the day-to-day processes and technologies that could ensure OPCW information assets remain accurate, accessible, secure, and useful for organisational purposes.

Data model	A blueprint or map that defines how information is organised, what types of data exist, and how different pieces of information connect to each other within a system. Data models serve as a guide for understanding how to navigate from one piece of information to another and establish the rules for how data relate across an organisation. Multiple data models can be linked together to connect different datasets or systems, enabling comprehensive analysis across various domains of information.
Data scientist	A professional who combines statistical analysis, machine learning techniques, and domain expertise to extract insights from complex datasets. In OPCW applications, data scientists may develop predictive models for verification activities or apply AI techniques to support detection and analysis of prohibited activities.
Data security	The protection of data and information systems from unauthorised access, corruption, theft, or loss through technical, administrative, and physical safeguards. Data security focuses on maintaining the integrity, availability, and confidentiality of organisational information assets, ensuring they remain accurate, accessible to authorised users, and protected from cyber threats and other risks.
Data structuring	The technical process of organising curated data into standardised, machine-readable formats that enable efficient storage, retrieval, and automated processing. Data structuring involves defining data types, establishing field formats, creating consistent naming conventions, and ensuring that information is arranged in ways that computer systems can easily access, search, and analyse.
Decision tree	A predictive model that represents decisions and their possible outcomes in a tree-like structure. It classifies or predicts data by following a sequence of simple decision rules based on the input variables.
Designated Laboratory	Laboratories designated by the OPCW for the analysis of authentic samples. Designated Laboratories must be able to perform off-site analysis of chemical samples acquired by OPCW personnel that are collected from storage depots, and other installations, or from the site of an alleged use of chemical weapons.

Detection	The process of confirming the presence of a chemical in a sample using analytical or sensing techniques.
Distance-based similarity	A way of measuring how similar two data points are by calculating the distance between them in a multi–dimensional space. Smaller distances indicate greater similarity.
Dry exercise	A training or evaluation exercise focused on data interpretation and reporting. Participants are provided with raw analytical data (such as spectra) and must identify substances or trends without performing physical laboratory work or sample preparation.
Extrinsic signature	A chemical or physical characteristic introduced from external sources during the production, processing, storage, or handling of a substance, which can provide information about its history or origin.
False negative	A test result which wrongly indicates that a particular condition or attribute is absent.
False positive	A test result which wrongly indicates that a particular condition or attribute is present.
Forensic science	The application of scientific methods and techniques to examine evidence for investigative purposes, typically in legal or compliance contexts.
Fourier-transform infrared spectroscopy	An analytical technique that measures the absorption of infrared radiation by a sample to identify molecular vibrations. Fourier-transform infrared spectroscopy produces a spectrum representing chemical bonds and functional groups, allowing for rapid identification and characterisation of substances.
Gradient boosting methods	A family of machine learning techniques that builds predictive models step by step by combining many simple models, usually decision trees. Each new model focuses on correcting the errors made by the previous ones.
Ground-truth data	Ground-truth data are well-characterised reference samples or datasets with known composition, origin, or production route, providing a reliable benchmark for comparison.

High-dimensional data	Data that contain a very large number of variables or features relative to the number of observations. High-dimensional data arise in areas such as spectroscopy, genomics, and chemical sensing, where each sample can be represented by thousands of measurements, requiring specialised statistical or AI methods to extract meaningful patterns and avoid overfitting.
Hit list	A predefined set of target compounds, chemical impurities, or intrinsic parameters used to screen analytical data. A hit list allows researchers to rapidly identify specific markers that link a sample to a known source, production method, or precursor.
Identification	The process of unambiguously confirming the identity of a specific chemical substance. This involves comparing the analytical data from a sample against known reference standards or library data to distinguish it from all other possible compounds.
***In silico* data**	Data generated through computational modelling or simulation.
Inference	In machine learning, the process of making predictions by applying a trained model to unlabelled examples.
Interpretability	The ability to explain or to present a machine learning model's reasoning in understandable terms to a human.
Intrinsic signature	The unique ratio of stable isotopes and trace elements naturally present within a chemical substance. These proportions are determined by the specific raw materials and environmental conditions at the point of origin.
Kernel density estimation	A statistical method used to estimate the probability distribution of a dataset. It creates a smooth curve representing how data points are distributed, helping to identify patterns such as peaks or clusters.
k-nearest neighbours	A machine learning method that classifies a sample based on the categories of the most similar data points in the dataset. It classifies data based on majority voting or predicts values by averaging neighbour values.

Large language model	An AI program that mimics the decision-making abilities of human experts in specific domains. These systems use knowledge bases and inference rules to provide recommendations or analysis. They could be used by the OPCW to support chemical weapons identification or verification procedures.
Linear algebra	A branch of mathematics that studies vectors, matrices, and linear equations.
Linear discriminant analysis	A statistical method used to classify samples into predefined groups. It finds linear combinations of features that best separate different classes in the data.
Linear model	A model that assumes a straight-line relationship between the input features and the predicted outcome. It assigns a specific weight to each feature, making the model easy to interpret but limited in its ability to capture complex, curved patterns in the data.
Logistic regression	A statistical modelling method used to estimate the probability that a sample belongs to a particular category. It is commonly used for classification tasks where the outcome is binary or involves a small number of classes.
Machine learning	A type of AI where computers learn to perform tasks by finding patterns in examples and data, rather than being programmed with specific rules. The system improves its performance as it processes more examples, enabling it to make predictions or decisions about new situations it has not seen before, potentially supporting various OPCW analytical tasks.
Mass spectrometry	An analytical technique used to identify chemical substances by measuring the mass-to-charge ratio of their ions. By fragmenting molecules into a unique pattern of ions, it provides characteristic molecular features that enable the definitive identification and quantification of chemical warfare agents, precursors, and degradation products, even at trace levels.
Matrix decomposition	A mathematical technique that breaks down a complex matrix into a set of simpler matrices to facilitate faster, more stable, and more efficient computations.

Metadata	Data that describe other data, providing contextual information such as origin, format, author, date of creation, classification level, or processing history. Metadata help organise, manage, and retrieve information efficiently and supports traceability, governance, and quality control within data systems.
Naive Bayes classifier	A probabilistic classification method which assumes that the variables used for classification are independent from each other, which simplifies calculations while often still providing effective predictions.
Nerve agents	Toxic chemicals that disrupt the mechanisms by which nerves transfer signals throughout the central nervous system. By interfering with the breakdown of chemical messengers, they cause continuous overstimulation of the body's muscles and vital organs.
Neural network	A type of AI system designed to process information in a way loosely inspired by how the human brain works. These systems can recognise complex patterns and relationships in data, making them useful for tasks like image recognition or data analysis.
Non-linear model	A model that can capture complex, curved, or multi-dimensional relationships between variables. Unlike linear models, these can account for interactions where the effect of one feature depends on another, allowing for higher accuracy in sophisticated forensic datasets.
Non-parametric methods	Statistical or machine learning approaches that make no strict assumptions about the specific form or distribution of the data being analysed. These methods are more flexible and can adapt to complex, irregular patterns within a dataset, making them less susceptible to errors caused by incorrect preliminary hypotheses.
Off-site analysis	Chemical analysis that takes place in a laboratory away from the site at which the sample was collected.
Omics sciences	Collective of biological disciplines that study complete molecular sets within a biological system.
On-site analysis	Chemical analysis that takes place at the site at which the sample was collected.

Ontology	A structured framework that defines the concepts, categories, and relationships used to organise information within a domain. In databases, a fixed ontology provides a shared and consistent vocabulary that ensures data are described, stored, and interpreted in the same way across systems and users.
Organic samples	Samples containing organic compounds, meaning carbon-based molecular substances typically composed of carbon bonded to hydrogen and other elements such as oxygen, nitrogen, or halogens.
Outliers	Data points that differ significantly from the majority of observations in a dataset. They may result from measurement errors, unusual events, or natural variability and can influence statistical analyses or model performance.
Overfitting	A modelling problem that occurs when a model learns the details and noise of the training data too closely. As a result, it performs well on the training data but poorly on new, unseen data.
Parametric methods	Statistical or machine learning approaches that assume the underlying data follows a specific, predefined distribution or functional form. These are often computationally efficient and easier to interpret, provided the initial assumptions about the data are correct.
Partial least squares discriminant analysis	A statistical classification method that identifies patterns in complex data and separates predefined groups. It combines dimensionality reduction with discrimination between classes.
Precision (metrics)	A performance metric that measures the accuracy of the positive predictions made by a model. It represents the proportion of instances identified as positive that actually belong to the positive class. It is defined by this formula: $Precision = \frac{True\ Positives}{True\ Positives + False\ Positives}$.
Precursor	A chemical compound used as a starting material in the production of another substance. Precursors undergo chemical reactions or transformations to form the desired target compound, serving as the initial building blocks in synthetic pathways.

Precursor-product matching	The process of comparing the chemical profile of a starting material with that of an end product. This analysis aims to establish a direct link between a specific precursor and the final product, helping to identify the synthetic origin and the materials used in the production process.
Pre-trained model	A machine learning model that has already been trained on a substantial dataset to establish foundational patterns and features. These models or their internal components serve as a functional base that can be further adapted to specific tasks, significantly reducing the amount of data and computational time required for new applications.
Principal component analysis	A statistical method used to reduce the complexity of large datasets while preserving the most important information. Principal component analysis identifies patterns by transforming many correlated variables into a smaller number of new variables, called principal components.
Probability theory	The mathematical framework used to quantify the likelihood of random events and the associated uncertainties.
Provenance	The verifiable history and origin of a chemical sample, including its association with a specific source, event, or production pathway. It may be established or supported through analytical signatures, such as impurity profiles and isotope ratios, establishing an evidentiary link between a chemical profile and its specific source, production batch, or environmental background.
Raman spectroscopy	An analytical technique that measures the inelastic scattering of light by molecules to provide information on vibrational, rotational, and other low-frequency modes. Raman spectra complement Fourier-transform infrared spectra, and are useful for identifying chemical composition and structural features.
Random forest	A machine learning method that builds many decision trees and combines their results to make predictions. By aggregating multiple models, it improves accuracy and reduces the risk of overfitting.

Retrosynthesis	The process of determining possible synthetic pathways to produce a target chemical compound by working backward from the final product to identify starting materials and reaction sequences. In a forensic context, it is used to evaluate the feasibility of different production routes, predict necessary reagents, and identify the chemical blueprints used to manufacture a specific agent.
Route sourcing	The process of determining the synthetic pathway used to produce a chemical substance by analysing its chemical profile, including impurities, by-products, and remnants of the starting materials. Route sourcing aims to infer the methods, precursors, and conditions involved in production.
Sample/batch matching	The process of comparing the chemical profiles of two or more samples to determine if they share a common origin, such as the same production batch, storage container, or geographical source.
Selectivity	The ability of an analytical method to identify and quantify a target chemical in the presence of other substances. High selectivity ensures that the analysis is not compromised by other chemicals in the sample that might produce a similar signal.
Sensitivity	The degree to which an analytical method can detect and distinguish small differences in the concentration of a chemical. A sensitive method produces a clear, measurable change in output even for a small change in the amount of the substance present, allowing for high-resolution measurement of trace levels.
Spectral data	Information obtained from analytical techniques that measure the interaction of matter with electromagnetic radiation, such as infrared, nuclear magnetic resonance, or Raman spectroscopy. Spectral data contain molecular features that reveal chemical composition, structure, and properties of substances, enabling identification, verification, and quality control.
Supervised machine learning	A machine learning approach where the model is trained using a dataset that contains both input features and their corresponding, ground truth. The model can then predict the correct category or value for new, unseen examples.

Support vector machine	A machine learning method used for classification and regression tasks. It separates different groups of data by finding the boundary that best distinguishes between them in the dataset.
Synthesis	The deliberate chemical process of producing a target compound from simpler starting materials or precursors. Synthesis involves planning reaction pathways, selecting reagents and conditions, and executing laboratory procedures to generate a desired chemical product, whether for research, pharmaceutical, or industrial applications.
Test dataset	A final, independent dataset that the model has never seen during training or validation. It is used only once—after the model is finished—to provide an unbiased evaluation of how the system will perform on real-world, unseen samples.
Training dataset	The primary set of examples used to teach the AI model. The algorithm analyses these data to learn patterns and adjust its internal parameters.
Transformer model	A class of deep learning architecture designed to process sequential data using attention mechanisms. Transformers can capture long-range dependencies and contextual relationships in data, making them well suited for tasks such as natural language processing, spectral prediction, and multimodal chemical data integration.
Unsupervised machine learning	A machine learning approach where the model analyses data that have no pre-existing labels. Instead of predicting a specific target, the algorithm identifies inherent patterns, structures, or clusters within the dataset.
Validation	The process of testing whether an AI system meets its requirements and performs as expected for its intended use. Validation involves evaluating the system's accuracy, reliability, and performance using independent data to confirm it works properly in real-world situations before deployment.
Validation dataset	A set of unseen data used during development to fine-tune the model's settings. It helps the developer see how well the model generalises to new information so they can optimise the model.

Weak classifier	A simple machine learning model that performs only slightly better than random guessing. While limited individually, many weak classifiers can be mathematically combined through ensemble methods to create an accurate classifier.
Wet exercise	A comprehensive training or evaluation exercise involving the physical handling of chemical samples. This includes the analytical process, such as sample preparation, extraction, and derivation, followed by instrumental analysis and data interpretation.

ANNEX 2: A short introduction to forensic chemistry and the statistical evaluation of forensic evidence

This annex is based on a background paper provided by Dr Anders Nordgaard (National Forensic Centre, Sweden) and Prof. Arian van Asten (University of Amsterdam, the Netherlands).

Criminalistics is the Science of Individualisation — Paul L. Kirk

1. Introduction to forensic chemistry

First introduced in the twentieth century, forensic science is the application of science in matters of law, where scientific knowledge, insights, and methods are applied to criminal investigations and interpreted in courts of law to assist judges and magistrates in their decisions. Areas of expertise in forensic science, widely known today across the world, range from (but not limited to) DNA analysis, fingerprint recovery, toxicology, and drugs analysis to firearms and ballistics. Physical evidence, known as the 'silent witness', falling under these expert areas and more, is recovered from crime scenes, suspects, and victims by personnel trained in identification, documentation, collection, preservation, packaging, and transportation—while keeping an auditable trail of their movement through the chain-of-custody documentation. The main aim of the crime scene investigation is to reconstruct the event(s) and link suspects to victims and crime scenes. In this multidisciplinary field, forensic experts are trained to operate professionally within the criminal justice system. They use the latest scientific developments in their area of expertise and develop and validate their own methods under strict QA conditions, applying them to most criminal investigations as part of their case work in specialised laboratories. After examination and analysis of the physical evidence brought from the crime scene, these experts report their findings to their customer (mainly the police and the public prosecution office). When summoned to testify in court, forensic scientists have a duty to maintain impartiality in presenting and interpreting their findings.

Forensic chemistry covers a wide spectrum of specialised expertise grounded in chemical science and analytical methodology. [121] Its primary function is the identification, characterisation, and interpretation of chemical substances that are relevant to criminal investigations. Many substances encountered in this domain are intrinsically linked to criminal conduct. These include explosives (energetic materials) and drugs of abuse (psychoactive substances), the possession of which is frequently criminalised under

[121] van Asten, Arian. *Chemical Analysis for Forensic Evidence.* Amsterdam, Netherlands: Elsevier, 2023.

statutory control regimes. Such compounds are typically listed explicitly in legislation or regulated schedules. In addition, less common or highly specialised toxic agents may be used to perpetrate a crime, including substances capable of poisoning or incapacitating individuals or larger populations, such as CWAs. However, forensic chemistry is not confined to inherently illicit materials. Numerous chemicals with legitimate industrial, domestic, or commercial applications may be misused for criminal purposes. Their possession remains lawful until criminal intent or use is established. Ignitable liquids such as gasoline or lamp oil exemplify this dual-use context: while lawfully available, their detection in fire debris may indicate deliberate rather than accidental ignition.

The discipline also addresses materials that are not directly instrumental in the offence but are produced, transferred, or deposited as a consequence of it. This is particularly relevant in the examination of forensic micro traces—minute material fragments that can assist in event reconstruction and the evaluation of competing accounts. Examples include gunshot residue particles, automotive paint transfers in hit-and-run incidents, and textile fibres exchanged during violent or sexual offences. The systematic detection, collection, and chemical characterisation of such traces are often evidentially significant.

Chemistry further underpins a range of forensic enhancement techniques. Chemical reactions are routinely employed to visualise latent evidence, such as the use of ninhydrin to develop amino acid residues in latent finger marks on porous substrates. In these instances, chemical principles form the foundation of the forensic methodology applied to physical evidence. Central to the discipline is analytical chemistry, through which forensic experts determine the composition, structure, and distinguishing characteristics of materials. Advanced instrumental techniques, including MS, nuclear magnetic resonance spectroscopy, gas and liquid chromatography, infrared and Raman spectroscopy, and X-ray diffraction and fluorescence, are employed to identify unknown substances, resolve complex mixtures, detect trace impurities, and establish features such as isotopic signatures. Through the rigorous application of these methods, forensic chemistry provides scientifically robust data to support investigative, intelligence, and judicial decision-making.

2. Criminalistic principles

The question of whether forensic science constitutes a true scientific discipline has been debated since science-based methods were first introduced into criminal investigations. These early discussions led to foundational principles that continue to underpin forensic practice. One of the most widely cited is the concept attributed to French forensic pioneer Edmond Locard, who established the world's first dedicated forensic laboratory in 1910. Although he did not formulate it verbatim as "every contact leaves a trace", his work articulated the principle that criminal acts, particularly violent encounters, inevitably generate material traces reflecting interaction between individuals, objects, and environments.

This principle remains robust across societal and technological evolution. It implies that traces are generated through contact and transfer, and that it is the responsibility of the forensic expert to apply appropriate scientific methods to detect, recover, and interpret them—provided they have not been deliberately removed or degraded. The logic extends beyond physical evidence; digital interactions similarly produce residual traces, although digital data introduce complexities such as infinite replicability that differ fundamentally from physical trace behaviour.

Modern forensic interpretation, however, recognises that the mere presence of trace material does not automatically indicate the actions that led to its presence. Courts frequently consider multiple scenarios, including secondary or tertiary transfer mechanisms in which individuals may present incriminating traces without direct involvement in the offence. Addressing such challenges requires methodological refinement: quantitative assessments (i.e., how much material is present), spatial distribution analysis, and advanced imaging of transfer patterns. These developments reflect the scientific imperative to move beyond binary presence–absence assessments toward probabilistic evaluations grounded in empirical data.

The transfer principle also explains procedural safeguards in crime scene work. Investigators wear protective clothing not primarily for self-protection, but to prevent contamination of the scene. Every contact made during examination risks altering evidentiary patterns irreversibly. In hazardous contexts—such as investigations involving CWAs—personal protective measures serve the dual function of safeguarding health and preserving evidential integrity.

The scientific identity of forensic science was advocated by Paul L. Kirk in his 1962 essay, *The Ontogeny of Criminalistics*.[122] Kirk argued that criminalistics merits recognition as an independent science due to its distinct specific objective: not merely identifying or describing materials but establishing their origin. He summarised this with the statement, “Criminalistics is the science of individualisation”. In forensic reasoning, the aim is to associate a trace with its specific source. For example, an unidentified bloodstain at a crime scene must be linked to a particular individual. Short tandem repeat DNA profiling exemplifies this principle by generating highly discriminating genetic profiles that function as individualising markers.

In forensic chemistry, the principle of individualisation translates into the generation of distinctive chemical features. Once the nature of a material is established, further analysis targets trace impurities and other specific chemical characteristics that enable comparison with known reference samples. Such comparative analysis can associate seized materials such as drugs, explosives, or toxic agents with specific production batches or sources. This

---

122 Kirk, Paul L. “The Ontogeny of Criminalistics.” *The Journal of Criminal Law, Criminology, and Police Science* 54, no. 2 (June 1963): 235. https://doi.org/10.2307/1141173.

can provide critical associative value, both independently and alongside biological or fingerprint evidence.

Within the international chemical forensic community, particularly in the context of CWA investigations, this process is commonly referred to as chemical attribution. The resulting chemical attribution signature may encompass impurity profiles, by-products, degradation products, and isotopic characteristics. Advanced techniques, such as IRMS, enable measurement of light element isotope ratios, while specialised NMR methods can determine position-specific isotopic distributions. These approaches are especially valuable when analysing highly purified substances with minimal detectable impurities.

3. Criminalistics is the science of individualisation

Building on Kirk's principle that criminalistics is the science of individualisation, forensic chemists must follow a structured sequence of analytical steps before attributing a trace to a specific source. The first step is fundamental: the trace material must be accurately identified and chemically characterised. Without establishing what the material is, no meaningful comparison can be undertaken. This initial phase answers the essential question of "what". It also prevents unnecessary or scientifically unsound comparisons. For example, if an improvised explosive device at a crime scene contains triacetone triperoxide (TATP), while trinitrotoluene (TNT) is recovered from a suspect's residence, the materials are chemically distinct and cannot share a common origin. Although the suspect may still be implicated through other evidence, no chemical link can be established. Proper characterisation therefore serves both evidential clarity and analytical efficiency. A second consideration is the reliability of contextual information. Crime scenes are uncontrolled environments; labels, packaging, or apparent descriptions cannot be assumed accurate. Mislabelled containers or substituted contents are not uncommon. Therefore, independent chemical verification is essential prior to further interpretation. Once the principal compound or material type has been established, the forensic chemist can proceed to further material characterisation. At this stage, classification becomes relevant. Classification seeks to determine the category to which the material belongs by identifying class characteristics—features shared by all members of a given group. In manufactured materials, such characteristics are often deliberately created to achieve specific product performance and are maintained through rigorous QC.

In (semi)biological materials, such as cocaine derived from coca leaves or heroin derived from poppy plants, class characteristics typically reflect inherent biological or chemical properties of the source species. Classification enhances investigative intelligence but does not individualise. It cannot distinguish between different batches of the same product type. Nonetheless, it can generate valuable leads, particularly when no suspects have been identified through tactical investigation and case context. For instance, identifying a specific precursor chemical used in a large synthetic drug production may direct investigators toward supply chains, purchase records, or import data.

Once a suspect batch is located, the analytical focus shifts from classification to individualisation: do two samples share a common origin? Individualisation relies on profiling methods that target individual characteristics—features arising from uncontrolled or stochastic processes that do not recur. These may include trace impurities, synthesis by-products, degradation compounds influenced by ageing conditions, or temperature-sensitive reaction artefacts. Unlike class characteristics, these features are not intentionally created and therefore may vary between production batches. If two samples display the same rare chemical attribution signature, the evidential weight of association increases. Yet, the identification of significant differences may provide exculpatory value, demonstrating that two materials do not share a common source. However, even a clear analytical match does not automatically establish common origin. A match merely indicates the possibility of shared origin; its forensic value depends on contextual evaluation and population-based understanding of how frequently such a profile occurs within the relevant class. Therefore, the final and most critical step is evaluative reporting.

Only when scientific findings are contextualised and their weight conveyed can triers of fact determine whether the evidence contributes meaningfully to establishing involvement beyond reasonable doubt. In this structured progression, from identification to classification, profiling, and evaluative interpretation, lies the scientific integrity of forensic chemistry.

4. The individualisation fallacy

Kirk's assertion that "criminalistics is the science of individualisation" does not imply that true individualisation is routinely achievable. The term itself is conceptually problematic. Strictly interpreted, successful individualisation would mean linking an evidentiary material to its one and only possible source. Such a conclusion would require the exclusion of all alternative sources, an assumption that, in practice, is rarely demonstrable. For this reason, the concept has been subject to sustained criticism within both forensic science and legal communities.[123] In a closed-set comparison, individualisation is theoretically feasible. The potential sources are limited and defined in advance. A hypothetical example is the classic "murder on the Orient Express" scenario: assuming no one entered or exited the train, the perpetrator must be among the passengers. If a DNA profile from a bloodstain matches only one passenger's reference profile, the evidential conclusion appears definitive within that restricted universe of candidates. Similar controlled scenarios occasionally arise in forensic chemistry, but they are rare.

Most real-world investigations involve open-set comparisons, in which the true source may lie outside the known reference samples. In such contexts, it is challenging to examine every conceivable alternative. Instead, forensic scientists estimate the broader context by

[123] Saks, Michael J., and Jonathan J. Koehler. "The Individualization Fallacy in Forensic Science Evidence." *Vanderbilt Law Review* 61 (2008): 199-219. https://scholarship.law.vanderbilt.edu/vlr/vol61/iss1/4.

analysing representative samples from a relevant population. This allows estimation of the random match probability (RMP)—the probability that two materials with matching profiles originate from unrelated sources. By studying ground-truth comparisons between unrelated materials and measuring the frequency of coincidental matches, scientists can estimate the discriminating power of a profiling method. For highly selective methods targeting independent and characteristic features, the RMP may be extremely low—approaching zero—but it can logically never reach zero in an open-set framework. A match therefore provides probabilistic support for a same-source hypothesis; it does not establish uniqueness. Despite this, some forensic domains have historically expressed conclusions in categorical terms. For example, in jurisdictions where fingerprint identification relies on a threshold number of matching minutiae (e.g., 10 or 12 points), examiners may declare exclusion of all other sources. From a scientific standpoint, however, such claims must be approached cautiously. Unless the fingerprints of every individual in the relevant population have been examined, the RMP cannot be asserted as zero. When uniqueness is claimed in an open-set context without empirical support, it constitutes the individualisation fallacy—which is the erroneous inference that high discriminative power equates to absolute uniqueness.

Chemical impurity profiling introduces more complexity. Unlike DNA databases or fingerprint repositories, relevant reference populations for chemical materials are often difficult to define, obtain, or characterise. The available material universe may be constrained by production processes, supply chains, or market distribution, yet still remains significantly larger than the examined reference set.

An example arises in arson investigations. Suppose gasoline residue recovered from a fire scene is chemically profiled using GC-MS and matches a sample obtained from a specific fuel station. Even if security footage confirms that the suspect visited that station, the evidential inference remains limited. Fuel dispensed from a single storage tank will be chemically homogeneous across potentially thousands of litres, supplied to numerous customers. Some may have filled vehicles; others may have used containers such as jerrycans. Consequently, a compositional match does not imply exclusive origin from a single purchaser.

This highlights two additional concerns: batch size and duplicate sub-samples. Materials originating from a common production batch may be chemically indistinguishable yet widely distributed. To mitigate these issues, forensic chemists often study manufacturing processes, industrial supply chains, and trade patterns. Understanding how materials are produced, blended, transported, and distributed provides essential context for interpreting profiling results. Only by integrating analytical findings with knowledge of batch magnitude and market flow can experts provide scientifically defensible evaluations.

5. Evidence evaluation in the Bayesian framework

The above reasoning can be formalised using probability calculus. Chemical warfare agents are used here as an illustrative example, while a more comprehensive treatment of forensic

CWA analysis is provided in Section 7. Suppose there are reasons to believe that the CWA used in an attack at site A has the same origin as the CWA used in an attack at site B. This belief can be expressed as a hypothesis: *"The CWA used in an attack at site A has the same origin as the CWA used in an attack at site B"*. This hypothesis is associated with a prior probability, reflecting the initial degree of belief that it is true. That prior probability is established based on intelligence and investigative information related to the two attacks and, at this stage, is independent of the chemical characteristics of the CWAs themselves.

To evaluate how chemical findings inform this hypothesis, it is useful to introduce symbolic notation. Let the hypothesis be denoted as H, and its prior probability as $P(H)$. Both arise from the preliminary assessment of whether the two CWAs may share a common origin. The analytical findings can be denoted as E, representing the chemical evidence. The overall probability of observing the evidence in isolation is not central, apart from considerations of measurement uncertainty—which are typically negligible in validated analytical systems. The critical quantities are the conditional probabilities: the probability of observing the evidence if the two CWAs have a common origin and the probability of observing the same evidence if they originate from different sources. If these probabilities differ, the evidence discriminates between the competing propositions. In other words, the extent to which the evidence is more probable under one hypothesis than under the alternative constitutes its probative value.

The probability of observing the evidence if the two CWAs have a common origin is denoted $P(E|H)$ and is the conditional probability of the evidence E, given the hypothesis H, ('*The probability of obtaining the given findings if the considered hypothesis is true*'). The probability of observing the evidence if the two CWAs have different origins is denoted $P(E|\neg H)$, where "$\neg H$" stands for "not-H". Generally, one would introduce an alternative hypothesis, A, that expresses the relevant alternative to H. For instance, this would be the case when there are two distinct statements about the CWAs. One could be that the two CWAs were produced at laboratory L_1, and the other could be that the two CWAs were produced at laboratory L_2. Note that in this case, the second statement is not the opposite of the first, as other laboratories could be involved. However, for the present case, it may be argued that laboratories other than L_1 and L_2 are not relevant. For the alternative "not-H", the hypothesis pair becomes exhaustive, meaning that the sum of the probabilities equals 1; the samples are either related or they are not.

At this stage, there are three probabilities, $P(H)$, $P(E|H)$ and $P(E|\neg H)$. The fourth probability is as follows: $P(\neg H)$, since this must be $1 - P(H)$. If the CWAs do not have a common origin, they must have different origins. The probability sought is therefore the probability that the two CWAs have a common origin *given* the chemical findings (i.e. given the evidence). This probability is denoted $P(H|E)$ and it is also possible to consider the probability $P(\neg H|E)$, i.e., the probability that the two CWAs have different origins *given* the evidence (which is $1 - P(H|E)$). These probabilities are related through the so-called *Bayes' theorem:*

$$\frac{P(H|E)}{P(\neg H|E)} = \frac{P(E|H)}{P(E|\neg H)} \times \frac{P(H)}{P(\neg H)}$$

The right-most term in this expression, i.e., $\frac{P(H)}{P(\neg H)}$ is referred to as the *prior odds* of the hypothesis H. Odds are equivalent to probability but measured on another scale (from 0 to infinity, instead of from 0 to 1). It may be more practical in the sense that one could assess how much is speaking of H and how much is speaking against H; but it is always possible to deduce the probability from the odds and vice versa.

The left-most term of Bayes' theorem is referred to as the *posterior odds* of H, hence the odds of H given the evidence obtained. From these odds, the desired probability $P(H|E)$ can be deduced. The middle term, i.e., $\frac{P(E|H)}{P(E|\neg H)}$ is usually referred to as the LR for H in the light of the evidence. The term comes from the fact that both the numerator and denominator are seen as likelihoods for the respective hypotheses. The probability $P(E|H)$ is here interpreted as a measure of H rather than a measure of E (since E is known), and analogously for $P(E|\neg H)$.

Likelihood ratios can take values from 0 to infinity, where the value 0 occurs when $P(E|H) = 0$, i.e., when the evidence is not possible to obtain if H is true. This would be the case if the CWA used in the attack at site A showed to be mustard gas and the CWA used in the attack at site B showed to be sarin. An infinite value occurs when $P(E|\neg H) = 0$, i.e., when the evidence is not possible to obtain if $\neg H$ is true. However, to be able to assign a value of zero to this probability requires knowledge of all potential sources of the two CWAs, which is practically impossible. A value of 1 means that the evidence is neutral with respect to the two hypotheses, i.e., $P(E|H)$ is equal to $P(E|\neg H)$. The chemical findings are equally probable under both hypotheses and therefore do not provide any additional information with respect to which hypothesis is true. A value greater than 1 indicates that the evidence is supporting the hypothesis that the two CWAs have a common origin, and the higher the value, the greater the support. A value lower than 1 means that the evidence is supporting the hypothesis that the two CWAs have different origins, and the lower the value, the greater that support is. To move from the LR to a probability statement about a common origin, the prior probabilities of the competing hypothesis must also be taken into account through Bayes' theorem. The LR alone does not provide a posterior probability. Instead, it updates prior odds to posterior odds, which are equal to the LR only when the prior odds are equal to one (i.e., when both hypotheses are considered equally probable *a priori*).

As a numerical example, assume for simplicity that we have strong reasons to believe that the two CWAs have a common origin. The prior probability for this hypothesis is assigned to be higher than 0.75. The chemical findings are deemed to be at least 99% probable if the CWAs have a common origin, but only 0.2% probable if they have different origins. The LR then becomes $0.99/0.002 = 495$, hence the findings are 495 times more probable if the two

CWAs have a common origin than if they have different origins. Using Bayes' theorem, the prior odds are $0.75/0.25 = 3$, and the posterior odds are then $495 \times 3 = 1485$. From these, the posterior probability of $1485/(1485 + 1) \approx 0.9993$ is obtained. Hence, the chemical findings have increased the probability that the two CWAs have a common origin from 75% to 99.93%. The stronger the chemical evidence, the higher the LR value and, consequently, the closer the posterior probability will approach 100%. However, it will never reach 100%, as the final determination regarding whether the two common CWAs share a common origin ultimately rests with the triers of fact. This fundamental aspect of forensic science can be challenging for non-experts, especially when a designated expert is unable to provide a definitive and clear answer to what appears to be a straightforward question.

6. Feature-based vs score-based methods for evaluation

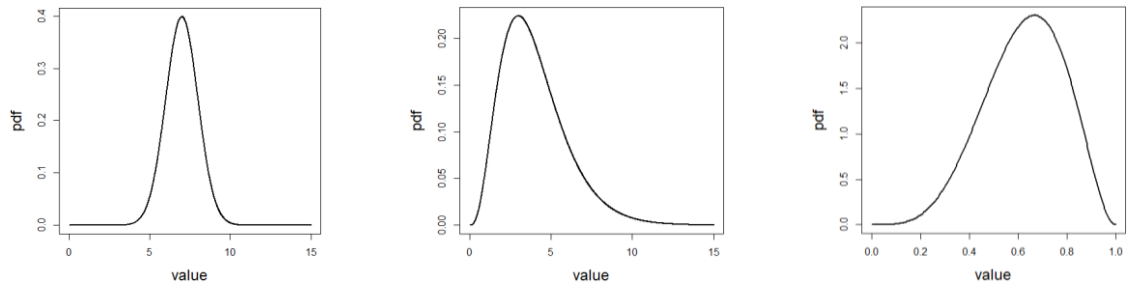
Continuous measurements

The foundational form of Bayes' theorem is based on calculation with distinct (conditional) probabilities for the evidence obtained. In forensic chemistry, the evidence comes from continuous-valued measurements, making it impractical to assign probabilities due to the extensive range of potential values—which are so numerous that they cannot be counted. In such cases, probability density functions are used rather than probabilities. A probability density function is a continuous function that is positive for every potential value that can be the outcome of a chemical analysis, for all other values it equals zero. Integrating the probability density function over the range of possible values leads to a value of 1. There are many families of continuous probability distributions, each characterised by the shape of its probability density function. Most chemists are familiar with the 'normal' (or Gaussian) distribution, with its bell-shaped probability density function. Its name refers to normal or natural variation, which is largely correct. Nevertheless, there is natural variation that cannot be described with a normal distribution (e.g., age at death). Many analytical chemistry results also deviate from the normal distribution and are therefore subject to different transformations to ensure an accurate representation. Figure A2.1 illustrates probability density functions for three common distributions.

Bayes' theorem can be written using probability density functions instead of probabilities in the LR. Instead of using the general symbol E for the evidence, the implicitly involved measurement makes it natural to use a typical symbol for a variable x:

$$\frac{P(H|E)}{P(\neg H|E)} = \frac{f(x|H)}{f(x|\neg H)} \times \frac{P(H)}{P(\neg H)}$$

where $f(x|H)$ denotes the probability density function valid under the assumption that H is true, and $f(x|\neg H)$ denotes the probability density function under the assumption that H is not true.



(a) (b) (c)

Figure A2.1: Examples of probability density functions: (a) a normal distribution with mean of 7 and standard deviation of 1; (b) a gamma distribution with shape parameter 4 and rate parameter 1; (c) a beta distribution with shape one parameter 5 and shape two parameter 3.

When a specific alternative hypothesis (A) is formulated, $\neg H$ is replaced by A and the corresponding probability density function can be denoted $f(x|A)$.

While distinct probabilities may be directly assigned based on subject knowledge and experience from both personal and professional contexts, it is not realistic to make an analogous assignment of a probability density function. If a particular distribution family—such as normal distribution, gamma distribution, or beta distribution—is assumed, then the probability density function can be assigned using sufficient reference data to estimate its parameters. When that is not the case, the *kernel density estimation* is typically used. However, this requires a large amount of available reference data.

Different objectives

Forensic investigations have multiple objectives, and the objective of a particular investigation will stem from the specific question posed by the commissioner. In forensic chemistry, the most prevalent questions concern *identification*, *classification*, *comparison*, and *quantitation*. These are described in the ENFSI Guideline for the use of Chemometrics in Forensic Chemistry.[102] Identification and classification do not typically involve significant uncertainty, and results from quantitation are reported with uncertainty margins to communicate the inevitable uncertainty originating from measures such as repeatability and reproducibility.

However, when classes overlap due to their chemical characteristics, classification cannot be reported with a categorical answer. Instead, the chemical findings must be probabilistically analysed. Legal rules for evidence evaluation set aside, probabilistic classification can be done using classical or more recent statistical or chemometric methods, such as discriminant analysis or support vector machines. These methods, however, do not consider prior probabilities assigned by commissioners or factfinders, but are assigned from reference data, thereby reflecting the general distribution of classes.

There is a difference between the question *To which class does this material belong?* and *Does this material belong to class C*? In the former, traditional chemometrics can be used, while in the latter LRs can be used. The hypotheses then are H: *The material belongs to class C* and $\neg H$: *The material belongs to another class than C*.

The comparison of two materials (typically seizures of drugs, oils, explosive agents, and CWAs) can be conducted through various methods; however, for forensic findings to be applicable in a legal context, the use of LRs is recommended. The previous section outlined the general procedure when two materials have been recovered, focusing on the question of whether they originate from the same source or different ones. It is also possible that there is one recovered material and one suspected origin, raising the question of whether it comes from this origin or another one.

Multivariate data

In forensic chemistry, chromatographic analysis typically produces measurements across multiple dimensions, which are statistically referred to as multivariate observations. These dimensions are often not independent, and substantial correlations may exist between them. Examples include when an impurity in a sample is the result of the chemical reaction of another impurity; when impurities are both created in the same degradation reaction; or when both impurities originate from the same additive. This complicates the assignment of LRs. However, in the case of negligible correlations, the combined LRs can be obtained simply by multiplying the LRs from each measured dimension.

In *feature-based* evaluation, the multivariate observations should be modelled under each of the two competing hypotheses, which involves estimating the probability density functions of multiple variables. A measurement is denoted by $\boldsymbol{x}$, which is vector-valued so that $\boldsymbol{x} = (x_1, x_2, \dots, x_m)$, where m is the number of dimensions measured (e.g., number of monitored peaks in a chromatogram), referred to as the *features*.

For classification with overlapping classes, the LR for the hypothesis H: The material belongs to class *C* is then:

$$LR = \frac{f(\boldsymbol{x}|H)}{f(\boldsymbol{x}|\neg H)}$$

Assigning such an LR is challenging as the hypothesis $\neg H$ (i.e., "not-H") comprises several classes (all but class *C*). The probability density $f(\boldsymbol{x}|\neg H)$ must then be a mixture of the probability densities from each of these classes. Assuming there are n_c possible classes and they are numbered from 1 to n_c, with $C = C_1$, the numerator should be written:

$$f(\boldsymbol{x}|\neg H) = w_2 \cdot f(\boldsymbol{x}|C_2) + w_3 \cdot f(\boldsymbol{x}|C_3) + \cdots + w_c \cdot f(\boldsymbol{x}|C_c)$$

where $w_2 + w_3 + \cdots + w_c = 1$. However, the weights $w_2, \dots, w_c$ may be difficult to assign from a forensic perspective, since they refer to how common each class would be in a case like

the one under investigation. Consequently, a conservative assignment may be made in which the highest of $f(\boldsymbol{x}|C_2), \ldots, f(\boldsymbol{x}|C_c)$ is substituted for each $f(\boldsymbol{x}|C_j), j = 2, \ldots, c$, and the LR satisfies the inequality:

$$LR \geq \frac{f(\boldsymbol{x}|H)}{\max_{C' \neq C}\{f(\boldsymbol{x}|C')\}}$$

When comparing two recovered materials (common source question) or one recovered material against a known source (specific source question), it is possible to denote the measurements from one material by $\boldsymbol{x}_a$ and the measurements from the other material by $\boldsymbol{x}_b$. The hypothesis is either H: "*The two materials have a common origin*" or H: "*The recovered material originates from the known source S*". The LR is then:

$$LR = \frac{f(\boldsymbol{x}_a, \boldsymbol{x}_b \mid H)}{f(\boldsymbol{x}_a, \boldsymbol{x}_b \mid \neg H)}$$

which can be shown to be the same as:

$$LR = \frac{f(\boldsymbol{x}_a \mid \boldsymbol{x}_b, H)}{f(\boldsymbol{x}_a \mid \neg H)}$$

To estimate the probability density functions $f(\boldsymbol{x}_a \mid \boldsymbol{x}_b, H)$ and $f(\boldsymbol{x}_a \mid \neg H)$, a substantial amount of reference data must be used, and an increase in the number of dimensions (features) considered necessitates a corresponding increase in the required data. This is a significant challenge in many disciplines, and particularly applicable to the comparison of CWAs. Moreover, even when an adequate number of reference samples is available, determining which probability distribution is valid remains challenging. The probability distribution associated with the probability density function $f(\boldsymbol{x}_a \mid \boldsymbol{x}_b, H)$ (i.e., the numerator function of the LR) describes the variation between materials with a common origin (or the specific source, *S*). This distribution is expected to be close to a normal distribution, but the probability distribution having $f(\boldsymbol{x}_a \mid \neg H)$ as probability density function (i.e. the denominator function) is typically non-normal.

Instead of modelling the features, the two materials can be compared by measuring the distance (or similarity) between the two feature vectors. Common distance measures are:

- Euclidean distance $\sqrt{\sum_{i=1}^{m}(\boldsymbol{x}_{1,i} - \boldsymbol{x}_{2.i})^2}$;
- Canberra distance $\sum_{i=1}^{m} |\boldsymbol{x}_{1,i} - \boldsymbol{x}_{2.i}| / (|\boldsymbol{x}_{1,i}| + |\boldsymbol{x}_{2.i}|)$; and
- cosine similarity $\sum_{i=1}^{m} \boldsymbol{x}_{1,i} \cdot \boldsymbol{x}_{2.i} / \left(\sqrt{\sum_{i=1}^{m} \boldsymbol{x}_{1,i}^2}\right) \cdot \left(\sqrt{\sum_{i=1}^{m} \boldsymbol{x}_{2,i}^2}\right)$.

The measured distance is referred to as the *score* of the comparison. This score should then be transformed into an LR. For the common origin question, a set of pairs of materials known to have a common origin and a set of pairs of materials known to have different origins are

needed. These data are referred to as *ground-truth data*. In each set, the scores from all pairwise comparisons are calculated. The resulting sets of scores are used to model the probability distribution of a score from the comparison of two materials with a common origin and the probability distribution of a score from the comparison of two materials with different origins respectively. Since the common distance measures produce values on a continuous scale, it is necessary to estimate the probability density functions of these two distributions. Identifying distribution families for this estimation is often challenging and rarely successful, therefore kernel density estimation can be used instead. This method involves smoothing a histogram into a continuous curve, such that the area under the curve (its integral) equals 1. Figure A2.2 highlights an example of this process for the two sets of scores (ground truth).

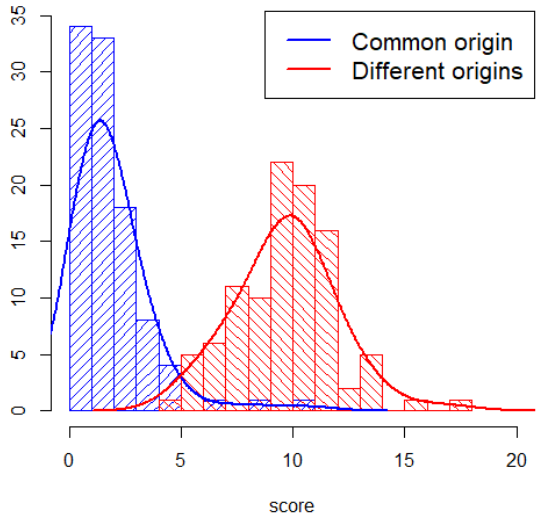


Figure A2.2: Histograms of scores from common origin pairs (blue) and from different origin pairs (red), together with kernel density estimates (solid curves).

The resulting curves depend on the bandwidth selected for smoothing. In the absence of evidence suggesting the existence of subgroups within the ground-truth data—which may lead to multimodal distributions—any additional peaks caused by overfitting the data should be smoothed out by increasing the bandwidth.

The kernel densities are now used as probability density functions for the probability distribution of a score given the materials compared have a common origin, and for the probability distribution of a score given the materials compared have different origins. A score can be denoted $s(\boldsymbol{x}_1, \boldsymbol{x}_2)$—with $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ being the feature vectors as described above—and the two probability density functions can be denoted as $f(s(\boldsymbol{x}_1, \boldsymbol{x}_2)|H)$ and $f(s(\boldsymbol{x}_1, \boldsymbol{x}_2)|\neg H)$ respectively. The LR for H given an obtained score is then:

$$LR = \frac{f(s(\boldsymbol{a}, \boldsymbol{b})|\, H)}{f(s(\boldsymbol{a}, \boldsymbol{b})\,|\,\neg H)}$$

This is illustrated in Figure A2.3.

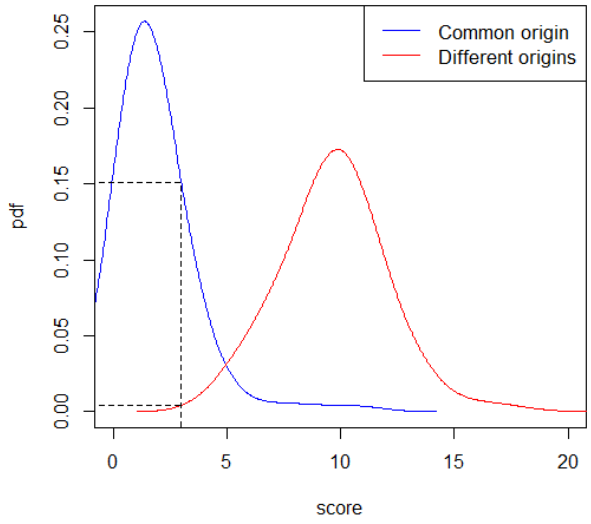


Figure A2.3: Illustration of how to obtain the LR from a score of 3 (horizontal axis). On the blue curve (probability density function of a score from common source comparisons), this score corresponds to 0.151, and on the red curve (probability density function of a score from different origins comparison) it corresponds to 0.004. For a common origin, the LR corresponds to $0.151/0.004 \approx 38$.

For the specific origin question, ground truth is required in terms of a set of materials known to originate from the specific source and a set of materials known to have originated from elsewhere. To determine the distributions, scores are then calculated from the feature vectors of pairs formed by combining a material known to originate from the specific source and the specific source itself (specific source scores to assess the within-source distribution), and pairs formed by combining a material known to have another origin than the specific source and the specific source itself (another source scores to assess the between-source distribution).

Validation

Likelihood ratio calculation methods that are based on the modelling of reference data should systematically be validated. This means that when the ground truth is known, the LR should be consistent with this truth. Therefore, if two materials are known to have a common origin, or if a material is from a specific source, the LR should be greater than 1. Similarly, if two materials are known to have different sources, or if a material is known to have a source other than the specific one, the LR should be less than 1. Considering this, one QA step involves analysing the rates of 'false positives' and 'false negatives', with the former being crucial in legal applications. In the Bayesian framework, false positive and false negative refer to a rate of misleading evidence—in other words, an LR value supporting the wrong hypothesis. However, even if these rates are low, this does not mean that the method for calculating LRs is entirely valid. In addition to low error rates, LRs should vary appropriately

with the strength of the evidence. A key distinction arises between feature-based and score-based evaluation. Feature-based approaches account for both similarity and typicality: two materials that are both similar and rare provide stronger support than two materials that are similar but common. In contrast, score-based approaches, at least in the form considered here, do not explicitly account for typicality. As a result, two pairs of materials may receive similar scores despite differing substantially in their rarity.

The aim is not to delve into validation methods, nonetheless the interested reader could be referred to the paper by Leegwater et al.[124]

7. CWA detection, characterisation, and attribution

Reaching this stage, a central question arises: how can criminalistic reasoning and the Bayesian framework for evaluating evidential strength be applied in the domain of CWA analysis and attribution? Equally important are the specific operational constraints of this expertise area and the related consequences when investigating a suspected CWA incident.

In the context of an OPCW investigation into alleged CWA use or suspicious activities, such as illicit production or storage, the primary task is the detection and identification of scheduled chemicals. Within Kirk's framework, this corresponds to answering the fundamental "what" question: which chemicals are present, and do they relate to CWA production or deployment? Due to the extreme toxicity, volatility, and reactivity of many CWAs, this step is technically challenging. The intact agent may no longer be present, particularly when investigations occur long after the reported event. Consequently, analysts frequently focus on associated substances, including precursor chemicals, hydrolysis products, and other degradation markers. The detection of such compounds may indirectly indicate CWA use and may assist in narrowing down the specific agent involved. Demonstrating the presence of the intact CWA itself is often considerably more difficult. For this reason, investigators may rely on matrices capable of retaining trace-level residues, such as biomedical samples from exposed individuals or environmental samples, including vegetation collected at the scene.

When analytical results indicate the presence of a CWA or related compounds, the next phase involves further characterisation. This includes assessing potential synthesis routes, precursor materials, stabilisers, and degradation pathways, as well as estimating production timeframes. Under favourable conditions, such signatures may be associated with specific production methodologies linked to particular states or time periods, given that historical production processes are relatively well documented. In other forensic domains, production-

[124] Leegwater, Anna Jeannette, Peter Vergeer, Ivo Alberink, Leen V. van der Ham, Judith van de Wetering, Rachid El Harchaoui, Wauter Bosma, Rolf J.F. Ypma, and Marjan J. Sjerps. "From Data to a Validated Score-Based LR System: A Practitioner's Guide." *Forensic Science International* 357 (April 2024): 111994. https://doi.org/10.1016/j.forsciint.2024.111994.

route information may have limited discriminative value; however, in CWA forensics it can be highly significant. The feasibility of such profiling is strongly dependent on sample quality. In complex biological matrices, such as blood, detailed impurity or process markers may be heavily obscured or degraded, although emerging studies indicate that some degree of profiling may still be achievable. In severely degraded environmental samples, meaningful chemical characterisation may be nearly impossible. By contrast, high-quality samples may enable the generation of chemical features capable of supporting batch differentiation and comparative analysis against reference or seized materials.

When profiles differ, the interpretation must proceed with caution. Observable discrepancies do not automatically exclude a common origin, particularly where degradation, environmental exposure, or matrix effects may account for compositional changes. Any inference of relatedness despite differences must be supported by chemically plausible and internally consistent explanations. At the same time, the simplest explanation—namely that the materials are unrelated—must always remain under consideration. On the other hand, even a close or apparently perfect profile match does not, in itself, establish evidential weight. Under a Bayesian framework, the probative value depends on the LR, which requires estimation of how frequently such a profile would occur among unrelated sources. For both score-based and feature-based models, this necessitates a sufficiently large and representative reference dataset.

In the CWA domain, this requirement constitutes a major structural limitation. Incidents are rare, sample availability is constrained by toxicity, security restrictions, and diplomatic sensitivity, and systematic reference collections are difficult and resource intensive. Score-based approaches may be comparatively efficient because they require less extensive population data; nevertheless, even these models may be limited due to insufficient reference materials. In such circumstances, the expert may be able to state that profiles are consistent with a possible common origin but cannot reliably quantify the evidential strength of that conclusion.

Recognising these constraints, the TWG on Chemical Forensics has explored options to facilitate broader compilation and exchange of reference data. One potential solution is the creation of shared, anonymised profile databases among accredited laboratories, containing essential analytical metadata without disclosing sensitive contextual information. Alternatively, another promising approach is the development of realistic *in silico* profile datasets generated through advanced data science techniques. Both approaches aim to strengthen the empirical foundation necessary for probabilistic evaluation in CWA attribution, while respecting security and diplomatic considerations.

ANNEX 3: Summary table of IIT reports

During its mandate, the TWG reviewed the first four reports issued by the IIT and summarised the areas where forensics approaches were applied in Table A3.1. In each investigation, the IIT first outlined working hypotheses as to how the incidents may have occurred and subsequently developed concrete scenarios based on all available information to test them. The approaches applied to support this process involved the analysis and integration of diverse sources of evidence—including chemical analyses, engineering assessments, witness testimony, and operational data—to assess whether proposed scenarios were plausible and consistent with the available information. These approaches may be chemical or drawn from other forensic fields, highlighting that chemical forensics is just one element within a broader multi-source investigative framework. Key examples are presented in the table and should be considered illustrative rather than exhaustive.

For context, it should be noted that while Table A3.1 focuses on investigative tasks, sample analysis had already been conducted beforehand, unambiguously confirming the identification of CWAs, including sarin-related compounds and sulfur mustard degradation products. In the context of the Convention, this type of identification is considered to be a verification task rather than a forensic task. Verification tasks—using GC-MS, LC-MS, and NMR—are typically carried out prior to forensic tasks, though they may also be conducted in parallel.

Lastly, the analysis of biomedical samples for forensic purposes was not leveraged in the investigations covered by these four reports. This is because it is significantly more challenging to identify key impurities of CWAs, as well as their associated profiles, in biomarkers of exposure in biomedical samples than in environmental samples.

Table A3.1: Summary of forensic approaches used by the IIT in its first four reports

Forensic or investigative tasks	Chemical forensics approaches	Approaches from other forensic fields	Examples
Chemical analysis and profiling			
Comparison of chemical profiles across collected samples	Analysis of environmental and material samples (including soil, clothing, and metal fragments) to determine key impurity profiles		1st IIT report: Sarin used in the chemical attacks in Ltamenah on 24 and 30 March 2017 was matched to the sarin previously developed by the Syrian Arab Republic as part of its chemical weapons programme. The sample-matching approach identified specific chemical impurities and unreacted remnants of the binary sarin agent. 1st IIT report: Samples from the 25 March 2017 incident were screened using ICP-MS and IC (both anion and cation) to detect and quantify inorganic species. The concentrations of the inorganic species were determined using certified reference standards. The samples collected closer to the chlorine cylinders showed consistently higher chloride concentrations than the ones collected further away, consistent with the release of gas from those cylinders.

Forensic or investigative tasks	Chemical forensics approaches	Approaches from other forensic fields	Examples
Provenance determination of the CWA	Analysis of chemical profiles and laboratory synthesis experiments enabled production routes to be determined	Analysis of intelligence, witness testimonies, and satellite imagery of storage facilities	1st IIT report: One DL conducted sarin synthesis experiments with different samples of DF (a sarin precursor and used in binary systems), including two subsamples from the Syrian stockpile. The chemicals observed in these experiments strongly correlated with those identified in samples from Ltamenah on 24 and 30 March 2017, as well as those from the Khan Shaykhun incident on 4 April 2017.The analytical results from these incidents were consistent with sarin resulting from a binary process using the DF from the Syrian Arab Republic stockpile. 4th IIT report: The IIT determined that the sulfur mustard—disseminated as a black liquid—had been produced by an improvised Levinstein method. The sulfur mustard was shown to contain Levinstein route-specific chemical impurities. Furthermore, its black colour and the presence of polychlorinated mustard species indicated that the sulfur mustard had been produced under

Forensic or investigative tasks	Chemical forensics approaches	Approaches from other forensic fields	Examples
			improvised conditions, with a shortage of ethylene, which led to the formation of black tar. Based on the production route, the IIT therefore assessed the possibility that the chemical incident in Marea may have been carried out using sulfur mustard originating from a State stockpile, rather than indigenously produced by a non-State actor, as extremely unlikely.
Analysis of chemical residues on munition fragments	Chemical analysis		1st IIT report: The interiors of three metal fragments (consistent with a part of the mixing system of a M4000 chemical aerial bomb, in addition to fragments consistent with the front and rear filling plug and housing) were checked for the presence of chemicals, enabling the IIT to conclude that neither sarin nor its precursors had been added to remnants of conventional aerial bombs.

Forensic or investigative tasks	Chemical forensics approaches	Approaches from other forensic fields	Examples
Delivery method and munitions reconstruction			
Identification of delivery method and impact analysis		Trajectory simulations, terminal ballistic analysis, experimental drop trials, impact phenomena assessments, and crater interpretation to evaluate possible aerial delivery methods.	3rd IIT report: The simulations (almost 80,000) performed by a missile trajectory expert, supported by the analysis of a terminal ballistics expert and a munitions expert, provided the IIT with reasonable grounds to believe that the cylinders found at Location 2 and Location 4 in Douma were dropped by a helicopter on 7 April 2018.
Munition identification			
Analysis of munition remnants		Morphological and metallurgical analyses, and photogrammetric studies of munition fragments	Physical examination of fragments, features, and deformation patterns helped determine the type and configuration of the munitions used. 1st IIT report: Specialists concluded that some munition remnants from the incident in Ltamenah on 30 March 2017 had more than likely originated from a M4000 chemical bomb. Photogrammetric studies of

Forensic or investigative tasks	Chemical forensics approaches	Approaches from other forensic fields	Examples
			two munition fragments from the 24 March 2017 incident in Ltamenah established that they were likely part of M4000 air-delivered munitions.
Dispersion and environmental analysis			
Gas dispersion modelling		Atmospheric modelling using environmental data and chemical properties	Atmospheric modelling was used to estimate gas dispersion patterns and potential casualty zones based on environmental conditions. 3rd IIT report: The two independent gas dispersion models considered by the IIT for Location 2 indicated that the accounts of the witnesses and the rapidity at which the symptoms began were reliable, and that those persons were affected by chlorine gas used as a weapon.

Forensic or investigative tasks	Chemical forensics approaches	Approaches from other forensic fields	Examples
Meteorological analysis		Exploitation of data from the World Meteorological Organization	Meteorological data were analysed to determine whether environmental conditions were favourable for chemical agent use. 1st IIT report: It was considered that the weather conditions on 24 March 2017 in the area of Ltamenah were favourable to the use of a non-persistent agent such as sarin.
Analysis of documentary, observational, and digital evidence			
Analysis of documents and material		Analysis of information obtained from States Parties, official documents, open-source information, flight data, intercepted communications, and operational documentation	1st IIT report: Flight data, communications, and documents were analysed to reconstruct operational activity around the attacks. 4th IIT report: Information on the dark web was analysed. This information included statements and discussions relating to the manufacturing and development of chemical weapons, as well as records, videos, and photographs documenting relevant military activities by the Islamic State in Iraq and the Levant (ISIL).

Forensic or investigative tasks	Chemical forensics approaches	Approaches from other forensic fields	Examples
Analysis of videos and photographs		Authentication through geolocation, metadata assessment, and other techniques, and subsequent exploitation of the information Reconstruction of chains of custody	Photographs and videos were leveraged extensively across all investigations and functioned as an important contextual and corroborative evidence stream. They enabled investigators to document scenes, verify locations, reconstruct timelines, test physical hypotheses, and authenticate digital materials. 3rd IIT report: Authenticated videos and photographs of the immediate aftermath of the incident in Douma on 7 April 2018 showed a white discolouration visible on one of the cylinders. This was assessed to be frost arising from the auto-refrigeration phenomenon that occurs when a liquefied gas is rapidly released from its cylinder. The white discolouration was not present the following day. 4th IIT report: Based on authenticated videos and images and relevant testimonies, the IIT was able to reconstruct the chain of custody of previously collected samples

Forensic or investigative tasks	Chemical forensics approaches	Approaches from other forensic fields	Examples
			from the date of collection to the date on which they were analysed by the OPCW DL.
Satellite imagery analysis		Geospatial analysis correlating satellite imagery with interviews, video evidence, and intelligence to identify storage sites and incident locations 3D building reconstruction	1st IIT report: Satellite imagery confirmed the presence of Su-22 aircraft at Shayrat airbase in late February 2017. 3rd IIT report: In its consideration of the hypothesis that the cylinders at Locations 2 and 4 had been dropped from adjacent buildings, the heights of these buildings were determined independently via 3D reconstruction from satellite images.
Witness and victim interviews		Exploitation of information from interviews	Testimonies from witnesses and victims provided descriptions of aircraft activity, gas release, and exposure events. 1st IIT report: Witnesses described a helicopter approaching from the west and circling the Ltamenah hospital before dropping a barrel. This testimony, which included a hand-drawn flight path, was used by experts to calculate that the "circling trajectory" and 24-kilometre flight time from

Forensic or investigative tasks	Chemical forensics approaches	Approaches from other forensic fields	Examples
			Hama airbase perfectly matched the reported timing of the attack.
Medical evidence assessment			
Assessment of clinical symptoms and medical reports		Clinical and toxicological analysis of medical records, reported symptoms, and treatment data	1st IIT report: A toxicologist found it plausible that the patients allegedly affected by the sarin attack on 24 March 2017 in Ltamenah displayed signs and symptoms consistent with sarin poisoning. 2nd IIT report and 3rd IIT report: Toxicologists and/or specialists concluded that the accounts of the victims and documentation were consistent with exposure to a toxic gas like chlorine. 4th IIT report: A toxicologist assessed that the clinical symptoms reported by the victims of the attack were characteristic of sulfur mustard exposure.

Forensic or investigative tasks	Chemical forensics approaches	Approaches from other forensic fields	Examples
Military and operational intelligence			
Assessment of the military and tactical situation on the ground		Reconstruction of the tactical situation and positions of actors using intelligence analysis, satellite imagery, and witness testimony	The general situation in the area was assessed for each incident investigated, including the broader context of military activities. 3rd IIT report: Military specialists assessed that Douma was a strategic hub at the M5 highway junction, where Syrian government forces and special forces units held surrounding air bases and infrastructure to conduct a high-intensity offensive against the armed opposition groups controlling the enclave.
Identification of military authority and chain of command		Reconstruction of the decision-making and command hierarchy by analysing communications, military orders, and operational records	Intelligence and communications analysis were used to examine possible command responsibility for the attacks. 2nd IIT report: Analysts identified a signed military order from the Minister of Defence and intercepted radio communications where a brigadier-general and a colonel gave direct orders to pilots from the Tiger

Forensic or investigative tasks	Chemical forensics approaches	Approaches from other forensic fields	Examples
			Forces' operations room at al Mujanzarat during the offensive to retake Saraqib.
Biological forensics			
Fingerprints and DNA analysis		Biological forensic analysis including fingerprint examination and DNA swab analysis	1st IIT report: Inspection of filling plugs recovered did not reveal any fingerprints; a swab was preserved for possible DNA testing and reference.

ANNEX 4: TWG on Chemical Forensics – Terms of Reference

1. The threat of use of chemical weapons remains. However, the ways in which chemicals may be used as weapons is evolving. While the threat that chemical weapons may be misused by a State actor in a large-scale attack still exists, there is growing concern around other scenarios of misuse of chemicals, such as by non-State actors as well as more targeted uses by States. In addition, there is an increasing potential for non-traditional agents to be used. These include biotoxins and central nervous system-acting chemicals, among others. Given the expanding threat spectrum, and the myriad attack uses and scenarios, it is more important than ever that as much information as possible be derived from any available samples from an alleged attack. It is imperative that the OPCW be able to fully understand and harness the utility of chemical forensics approaches and be able to use validated results in any investigation of misuse of a chemical.

2. In order to strengthen and augment the Technical Secretariat's capability in this regard, an in-depth review of the methods and technologies used in the field of chemical forensics would be a particularly valuable exercise and highly relevant in the current threat landscape. Further to his response to the report of the Thirty-Seventh Session of the Scientific Advisory Board (SAB), (EC-104/DG.22, dated 27 September 2023) and in accordance with paragraph 9 of the terms of reference of the SAB (Annex to C-II/DEC.10/Rev.1, dated 2 December 2004), the Director-General has decided to establish a Temporary Working Group (TWG) on Chemical Forensics and has appointed Dr Anne Bossée as the Chairperson of the Group.

3. The objective of the TWG is to review the science and technology relevant to chemical forensics and identify remaining gaps and challenges such that the OPCW can make best use of the information chemical forensics can yield moving forward. Considerations should be given to the work and recommendations from the SAB's previous TWG on Investigative Science and Technology (SAB/REP/1/19, dated 1 December 2019) as well as ongoing work in this area within the Secretariat. The findings will be considered by the SAB and recommendations will be provided to the Director-General.

4. The TWG will consist of individuals who have expertise in: chemical forensics, analysis, and/or synthesis; OPCW proficiency tests; in-field sampling; machine learning; or experience of implementation of the Chemical Weapons Convention. The TWG will comprise qualified members of the SAB as well as representatives from relevant academic, scientific and international organisations. Guest speakers will be invited regularly to assist the TWG in its collection of data and information, and formulation of advice. The TWG may also, when necessary, draw upon the expertise

of the Technical Secretariat, in particular the OPCW Laboratory, Inspectorate, and non-routine missions.

5. The TWG will report to the SAB, and will address specific questions relating to four subtopics, detailed below.

6. **State of the art.** Considering these specific areas:

- Batch-matching (correlating different samples to the same production batch and the influence of matrices)
- Impurity and trace element profiling (determining the trace impurities in a sample)
- Synthesis route (deriving the synthesis route used to make a specific chemical in a sample)
- Geographic context (pinpointing a geographic area where a sample was created or stored)
- Environmental factors and storage conditions (determining whether a sample was exposed to a certain environment, such as extreme heat or cold, dust or humidity)
- Isotope ratio techniques (specifically related to phosphorus, carbon, and deuterium)

a. What is the current state of the art related to determining the life cycle of a given chemical sample?
b. Can analysis of other materials that may be found, such as a storage container, or parts of a dissemination device, yield relevant information compared to liquid and soil samples?
c. What information is available related to the ability to conduct chemical forensic analysis on non-traditional agents which could be applied to CWAs? Are there any restrictions or limitations?

7. **Future capabilities**. Considering the following:

- The ability to analyse large datasets
- The ability to analyse limited datasets and the impact on model validity
- The ability to create or expand datasets to train models
- Targeted research based off data and results that have been optimised with computer models

a. What impact will the increased power and integration of machine learning and deep learning have on the field of chemical forensics?
b. What impact will the limited size of datasets available have on CWA forensic analysis?

c. What will the field look like in five to ten years, particularly in regard to the capabilities around the specific areas mentioned in 6?

8. **Methods and procedures**.

a. How can applied analytical methods have an impact on the results related to trace analysis and the chemical footprint?
b. How can data, methods, and procedures to conduct the in-depth CWA analyses expected in chemical forensics be standardised and shared?
c. What information is needed to ensure trust and reproducibility in the analysis and the results?
d. How can analytical data from different techniques (e.g., GC-MS, LC-HRMS, and NMR) be combined in forensic analysis?
e. Would curated/shared database(s) of relevant reference data be useful? Are there any recommendations/restrictions to secure these types of database(s)?
f. What best practices exist for on-site sampling to ensure the validity of subsequent forensic laboratory analysis and what challenges/gaps remain?

9. **Augmenting the OPCW's capabilities**.

a. How can the OPCW enhance its capability to capture and utilise chemical forensics-related information in the context of alleged use?
b. How will the OPCW ensure the validity of the forensic results?
c. How will the forensic results be reported? (criteria, level of confidence)
d. Are there other organisations or information that could be leveraged to augment the capability of the OPCW?
e. How can forensic analysis be promoted and enhanced at Designated Laboratories? How might other organisations or laboratories contribute?

10. In addition, the TWG will provide advice, as requested, on Technical Secretariat proposals for methodologies, procedures, technologies, and equipment for chemical forensics.

11. The Director-General might pose other relevant questions to the TWG, through the SAB.

12. The TWG will exist for a period of two years starting 1 January 2024. Thereafter, its work will be reviewed by the SAB and the Director-General, and a decision will be made as to whether it should continue its work and, if so, whether these terms of reference should be revised.

ANNEX 5: Members of the TWG on Chemical Forensics

	Participants	Affiliation
1	Prof. Arian van Asten	University of Amsterdam, Netherlands
2	Dr Crister Åstot*	Swedish Defence Research Agency (FOI), Sweden
3	Dr Khaldoun Bachari	Centre de Recherche Scientifique et Technique en Analyses Physico – Chimique (CRAPC), Algeria
4	Capt. Elma Biscotti*	Scientific and Technical Research Institute for Defence (retired), Argentina
5	Dr Anne Bossée* (Chairperson of the TWG)	DGA CBRN Defence, France
6	Dr Grégoire Delaporte	DGA CBRN Defence, France
7	Ms Anne-Marie Fortin	United Nations Office on Drugs and Crime
8	Dr Hanna Hakulinen	Finnish Institute for Verification of the Chemical Weapons Convention (VERIFIN), Finland
9	Ms Ang Lee Hwi	DSO National Laboratories, Singapore
10	Prof. Imee Su Martinez*	University of the Philippines-Diliman, Philippines
11	Dr Simon Ovenden (Vice-Chairperson of the TWG)	Defence Science and Technology Group, Australia
12	Dr Meehir Palit*	Defence Research and Development Organisation (DRDO), India
13	Col. Günter Povoden*	CBRN Defence Centre, Austrian Armed Forces, Austria
14	Prof. Ines Primožič*	University of Zagreb, Croatia
15	Dr Sarah Stubbs	Defence Science and Technology Laboratory (Dstl), United Kingdom of Great Britain and Northern Ireland
16	Dr Hongmei Wang	State Key Laboratory of NBC Protection for Civilian, China
17	Dr Audrey Williams	Lawrence Livermore National Laboratory, United States of America

* Member of the SAB during the mandate of the TWG

ANNEX 6: Invited Speakers at Meetings of the TWG on Chemical Forensics

Speaker	Affiliation	Area(s) of expertise
First meeting		
Dr Karen Howard	Government Accountability Office, United States of America	Analytical chemistry, technology assessments, and investigations on behalf of the US Government
Dr Katie McCormac-Miller	Bureau of Arms Control, Deterrence, and Stability, United States of America	Arms control and policy, CFITWG
Dr Daan Noort	OPCW Laboratory	CWA verification analysis, biomedical analysis for CWA exposure
Second meeting		
Prof. Niamh Nic Daéid	University of Dundee, United Kingdom of Great Britain and Northern Ireland	Forensic scientist focused on fire investigations, terrorist events, and drug manufacture
Prof. Roland Goertz	University of Wuppertal, Germany	Chemical safety and first responder perspectives, fire investigations
Prof. Gerrad Jones	Oregon State University, United States of America	Environmental chemistry and water quality, statistical analysis and ML modelling
Dr Tim Wilson	Verdel Instruments Limited, United Kingdom of Great Britain and Northern Ireland	Total correlation MS, representing emerging analytical techniques
Third meeting		
Dr Lina Chen	State Key Laboratory of NBC Protection for Civilian, China	Chemical forensics of CWAs, analysis of biomedical samples
Dr Laurence Dujourdy	L'Institut Agro Dijon, France	Analytical chemistry, chemical profiling of illicit drugs, statistical analysis

Dr Karin Höjer Holmgren	Swedish Defence Research Agency (FOI), Sweden	Analytical chemistry, impurity profiling of CWAs, statistical analysis
Dr Saer Samanipour	University of Amsterdam, Netherlands	Analytical chemistry, prediction of chemical signatures, toxicity and chemical space, statistical analysis
Ms Lisa Scharrenbroch	Federal Criminal Police Office, Germany	Forensic science, analytical chemistry, identification and molecular forensics of ricin
Mr Nihad Alihodžić	Declaration Assessment Team, OPCW	Syria-related investigations
Dr Sami Barrek	Fact-Finding Mission, OPCW	
Ms Valentina Falco	Investigation and Identification Team, OPCW	
Mr Moez Hani	Declaration Assessment Team, OPCW	
Ambassador Hernán Salinas	Investigation and Identification Team, OPCW	
Fourth meeting		
Dr Ivo Alberink	Netherlands Forensic Institute, Netherlands	Statistics
Dr Mirjam de Bruin-Hoegée	TNO, Netherlands	Profiling of CWAs, chemical markers
Dr Georgios Gkatzelis	Forschungszentrum Jülich GmbH, Germany	Source attribution of air pollutants
Dr Michiel Grutters	Netherlands Forensic Institute, Netherlands	Chemical forensics of drugs and fire debris, toxicology, and environmental investigations
Mr Sami Huhtala	National Bureau of Investigation, Finland	Chemical forensics of drugs and oil spills
Dr Helen Salouros	National Measurement Institute, Australia	Impurity profiling of drugs
Prof. Mingxun Wang	University of California, Riverside, United States of America	Metabolomics, identification of molecules

Dr Nina Welti	Commonwealth Scientific and Industrial Research Organisation (CSIRO), Australia	Forensic applications in agriculture
Prof. David Wishart	University of Alberta, Canada	Prediction of spectral and chemical properties and supporting identification using AI and ML methods
Fifth meeting		
Dr Cornelia Rasmussen	University of Texas at Austin, United States of America	Analytical geochemistry, PSIA-NMR
Prof. Sermet Sezigen	University of Health Sciences, Türkiye	Tracking chemical casualties, biomedical samples in supporting investigations
Dr Anders Nordgaard	Swedish National Forensic Centre, Sweden	Probabilistic assessment of forensic information and statistical methods relevant to forensic questions
Dr Maria Wallenius	European Commission – Joint Research Centre	Nuclear forensics
Dr Carlos Fraga	Air Force Research Laboratory, United States of America	CFITWG, chemical impurity profiling of CWAs
Dr Thierry Rakotoarivelo	Commonwealth Scientific and Industrial Research Organisation (CSIRO), Australia	Privacy and sensitivity issues in AI systems

ANNEX 7: Classification Methods

All classification workflows involve several datasets for model optimisation, training, and validation. Training datasets are used to adjust model parameters, and a subset is often reserved for internal validation. A common internal validation strategy is cross-validation, which involves splitting the training dataset into two subsets. Splitting is usually carried out with strategies and algorithms to ensure that as much variability as possible is captured. The model is trained on one subset and tested on the other, with the process repeated several times. Once the model parameters have been optimised, an external validation dataset—ideally completely independent of the training dataset—is used to assess the overall model performance.

Model evaluation relies on metrics that can vary according to the mathematical principles underlying each method. A common, model-agnostic tool is the confusion matrix,[125] which summarises the number of correct and incorrect predictions made by a classification model. From this matrix, metrics such as true positive, false positive, true negative, and false negative rates can be computed. An example of a confusion matrix for a three-class model is provided in Table A7.1.

Table A7.1: Example of a confusion matrix for a three-class model[126]

	Setosa (**predicted**)	Versicolor (**predicted**)	Virginica (**predicted**)
Setosa (**ground truth**)	88	12	0
Versicolor (**ground truth**)	6	141	7
Virginica (**ground truth**)	2	27	109

In this example, the model is pre-trained, meaning that predictions on unknown data are made without referring directly to the reference dataset. This approach is more computationally efficient than training the model from scratch for each query, though it must be retrained whenever the reference dataset is updated.

Before implementing ML methods, datasets generally require preprocessing to address statistical biases. Techniques include data normalisation, scaling, transformation, and missing value imputation. The choice of preprocessing method depends on the data

125 Brereton, Richard G. "Contingency Tables, Confusion Matrices, Classifiers and Quality of Prediction." *Journal of Chemometrics* 35, no. 11 (April 14, 2021). https://doi.org/10.1002/cem.3331.

126 "Machine Learning Glossary Google for Developers." Google, accessed December 15, 2025. https://developers.google.com/machine-learning/glossary?hl=en#confusion-matrix.

characteristics (for example, size, distribution, descriptive statistics, and sparsity) and the forensic question being addressed.[127] Analytical (MS and NMR) data preparation has been extensively described in the scientific literature of the field of metabolomics,[128] as well as in Blue Book ROP for CWA analysis.[46]

There are many methods in ML and those of potential relevance to chemical forensics are provided in Table 3 under Subgroup 2's findings. These methods can be divided into six families based on their underlying mathematical and algorithmic principles, with some methods capable of hybridisation to combine strengths. The choice of method also depends on the data characteristics and the forensic question being addressed.[127] Linear methods such as PLS-DA or PCA are widely used due to their interpretability and flexibility, though they are theoretically limited to linear phenomenon and rely on strict statistical assumptions, which can be mitigated through preprocessing.

Some classification methods were originally designed for two-class problems. Extending these to multi-class problems requires careful consideration, though suitable algorithms and strategies exist to address these challenges.[81]

For a general audience, "ML" often implies artificial neural networks or similar methods such as deep learning. These methods are extremely powerful for certain applications, including self-driving cars or image and voice recognition. However, their underlying structure—multiple interconnected artificial neurones arranged in layers—results in a high number of parameters to tune. As noted previously, models with many parameters require large training datasets to avoid overfitting, a challenge that is particularly pronounced with deep learning due to its multiple layers. Despite these demands, some applications indirectly relevant to chemical forensics, such as predicting physico-chemical or analytical properties, [129] successfully employ these algorithms.

---

127 Mörén, Lina. "Different Strategies and Methods Used for Matching of Chemical Profiles/Impurity Profile Comparison." Presentation, Informal Workshop on Chemical Forensics Capabilities of the OPCW Designated Laboratories, Helsinki, June 10, 2025.

128 Berg, Robert A. van den, Huub C.J. Hoefsloot, Johan A. Westerhuis, Age K. Smilde, and Mariët J. van der Werf. "Centering, Scaling, and Transformations: Improving the Biological Information Content of Metabolomics Data." *BMC Genomics* 7, no. 1 (June 8, 2006). https://doi.org/10.1186/1471-2164-7-142.

129 Wang, Fei, Dana Allen, Siyang Tian, Eponine Oler, Vasuk Gautam, Russell Greiner, Thomas O Metz, and David S Wishart. "CFM-ID 4.0 – a Web Server for Accurate MS-Based Metabolite Identification." *Nucleic Acids Research* 50, no. W1 (May 24, 2022). https://doi.org/10.1093/nar/gkac383.