



# THE ROLE OF STATISTICS IN FORENSIC SCIENCE CASEWORK AND RESEARCH

Marjan SJERPS

*Netherlands Forensic Institute, the Hague, the Netherlands*

## Abstract

Statistics plays a crucial role in obtaining, interpreting and reporting results. Sampling, statistical quality control, chemometrics, biometrics, design of experiments, and the Bayesian framework for interpreting evidence are areas of statistics that are of key importance to many forensic scientists. However, most forensic scientists need additional training and support to make full use of the wealth of statistical results that is available. The role of statistics is becoming increasingly important as forensic reports are currently scrutinised by courts and accreditation bodies. Statistics and forensic science have a lot to offer to each other, and more interaction is desirable.

## Key words

Evidence evaluation; Forensic statistics, Bayes theory.

*Received 16 April 2006; accepted 15 May 2006*

## 1. Introduction

Forensic science laboratories usually have an extremely wide range of areas of expertise. Statistics and probability theory is usually not one of them. Nevertheless, statistics plays a crucial role in obtaining, interpreting and reporting the results in many of the areas. The rapid development of DNA analysis and the assessment of probabilities associated with this type of evidence have boosted the use of probabilistic reasoning, in particular the use of the so-called Bayesian framework for interpreting evidence. Many publications are available on the application of this framework to all sorts of forensic evidence (see [3] for a recent overview). The statistical literature also contains a wealth of other methods that are useful for forensic scientists. Unfortunately, many of these results are not used in practical forensic casework. For example, a special book has been published [6] on the inter-

pretation of glass evidence, but many experts still rely on rules of thumb and “eyeball” their data. I have the impression that this is because most forensic scientists lack sufficient statistical training or support, and that there is a strong need for easily accessible software.

This paper explores the role of statistics in forensic science in obtaining, interpreting and reporting results, and concludes that more interaction between statisticians and forensic scientist is desirable.

## 2. The role of statistics in obtaining results

A forensic investigation often starts by taking samples. For example, some items are selected from a pile of tablets or packets containing illicit drugs and sent to the laboratory for analysis, or some paint chips or glass fragments are taken from the car that was supposedly involved in a hit and run case in order to compare them

with fragments found on the victim, or samples are taken from huge amounts of solid wastes in forensic environmental casework. If it is known that the source is completely homogeneous, one sample is sufficient to draw conclusions about the whole source. Unfortunately, the source is generally not homogeneous, and the forensic scientist has to decide on the number, size, and location of the samples. These decisions should be based on statistical considerations, and this is considered a difficult task by many forensic scientists. For sampling discrete units from a large bulk, like the tablets mentioned above, a booklet has been written on the topic [7]. Lucy ([12], chapter 16) and Aitken and Taroni ([3], chapter 6) provide additional information and references for sources consisting of discrete units. For environmental casework, there is an extensive literature on sampling (see, e.g. [11] for environmental sampling in general, and Stelling and Sjerps [20, 21], for forensic casework). The literature on sampling also is of key importance when forensic databases are constructed.

Another area in which statistics plays an important role when obtaining results is quality control. Many forensic departments regularly use tools like control charts or calibration methods. Furthermore, introducing a new analytical method requires validation, say, by comparing the results to the results of another method. Quite complex situations can arise here, for example when there are unknown matrix effects. Hence, the relevant literature for the forensic scientist again includes a lot of statistics, e.g. Miller and Miller [13], Lucy [12] or Ryan [19].

A forensic research project, or a special experiment performed to solve a case require yet another specialised area of statistics called “design of experiments” (e.g. Montgomery [14], Robinson [18]). For example, a DNA scientist interested in the effect of the chemical composition of luminol (a substance that is used to find blood on crime scenes) on the performance of this fluid, should make an experimental design where the different compositions are varied and combined in a systematic and pre-defined way. The firearms expert may want to design an experiment to investigate the statement of the suspect that the gun fired even though the safety was on.

Hence, the statistical methods developed in the specialised areas of statistics like sampling, statistical quality control, chemometrics, and design of experiments are needed by forensic scientists in order to have their research meet current scientific standards.

### 3. The role of statistics in interpreting results

Once the forensic scientist has obtained the data, he has to interpret their evidential value for the case at hand. The rapid developments in forensic DNA analysis in the last twenty years have led to many questions concerning the interpretation of the data. For example, how should one calculate the match probability? What is the appropriate reference population to use in the calculation? What is the evidential value of a mixture of cell material from two or more persons, resulting in a so-called DNA mixture profile? How can we establish relatedness between persons based on their DNA profiles? Can we predict the ethnic group of the unknown donor of a semen stain? How does the number of comparisons affect the evidential value of a DNA match? These questions have led to heated discussions and numerous publications (see e.g. [4, 5, 9, 16]).

The questions raised in DNA evidence interpretation boosted the development of the “Bayesian approach to evidence interpretation” (some proponents of this approach prefer the name “logical approach”, or “likelihood ratio approach”), a general framework to assess the evidential value of all kinds of forensic evidence. This approach is accepted as the method of choice by most statisticians working in the forensic field, but is not generally used in casework by forensic scientists, except in the field of forensic DNA analysis. The main textbooks on this approach targeted at forensic scientists in general are Aitken and Stoney [2], Robertson and Vignaux [17], Aitken and Taroni [3], and Lucy [12].

These texts point out several errors of thinking that are common in forensic reports, e.g. the prosecutor’s fallacy, defender’s fallacy, base rate fallacy, double coincidence fallacy etc. Furthermore, they explain how the evidential value of a piece of forensic evidence can in principle be expressed as a number, called the likelihood ratio ( $LR$ ). This way of thinking has great merits even if this number cannot be produced due to lack of adequate data, as is often the case. The framework is useful as a guide to the expert to address the proper questions, thereby forcing the expert to consider not only the hypothesis of the prosecution, but also alternative hypotheses. Moreover, it can facilitate the harmonisation of expert conclusions, because the numerical scale provides a common language. Finally, it can also be used to draw conclusions about the evidential value of the combination of different kinds of forensic evidence.

Consider as an example of the latter the following imaginary scenario, based on a real case from the Netherlands Forensic Institute. A boy has been run

over by a jeep. The suspect's jeep is investigated, and red material is found on the exhaust. This material is compared to the red trousers of the boy, which are damaged in several places. First, an expert in textiles looks at the trousers and states that the damaged spots are caused by heat. A second expert compares the shapes of the material on the exhaust and the damaged spots on the trousers, and concludes that the shapes are similar. A third expert compares the chemical composition of the material and the trousers and concludes that they are made of the same kind of PET material. None of these statements is conclusive evidence that the material originated from the boy's trousers, but their combination obviously is very convincing. The experts therefore included a joint statement in their report, which addressed the following hypotheses: (a) the material on the exhaust originated from the trousers, versus (b) the material originated from some other source. They concluded that their combined findings were much more likely if the first hypothesis is true than if the second hypothesis is true<sup>1</sup>. The court made it clear that they appreciated this kind of combination by the experts.

Combining evidence and linking different sets of hypotheses can be facilitated by the use of Bayesian Belief Networks (also called Probabilistic Expert Systems). These kinds of models are implemented in probabilistic software which allows the expert to explore, for instance, the sensitivity of the LR to certain parameter estimates. Several forensic applications have been published for this relatively new technique (see e.g. [15, 22, 23], and much current research concentrates on this area.

Another area of applied statistics which has very useful results for forensic science, but which is not widely used is chemometrics, or the application of (multivariate) statistics to the results of chemical analyses (see e.g. [24] for a recent example). Over the years, the chemical or technical analyses have become very sophisticated, with techniques like ICP-MS. The statistical techniques to analyse the large amounts of data produced have also become sophisticated, and require statistical software. Techniques like principal components or graphics like score and loading plots, or their combination, the biplot, are powerful tools for the forensic scientist to wade through the data and grasp the main structure in a minute: for example, which samples are similar and which elements are most useful to discriminate? Subsequently, the scientist may use multivariate hypotheses tests to decide whether there is a significant difference between the

sample groups (for the statistical analysis of forensic glass data, see e.g. [6]). Even better, he may calculate a multivariate LR to express the evidential strength that the samples came from two different sources [1].

Unfortunately, there are currently very few forensic scientists who use the wealth of statistical techniques available to analyse their data and draw conclusions from them. It seems that the current procedure for most forensic practitioners is to glance over the large tables of data, investigate closely only a few of the variables produced and draw their conclusion using rules of thumb or an educated guess. It can only be a matter of time before this is seriously contested in court as examples of poor statistics.

#### 4. The role of statistics in reporting results

The Bayesian framework for interpreting evidence also has implications for the way forensic evidence is reported. First of all, information about the circumstances of the case which are relevant for the interpretation of the evidence should be mentioned. For example, in a DNA case the circumstances affect the population database that is used to calculate the match probability, or the interpretation of a sample containing a mixture of DNA. In environmental casework, for example in a case where the question is whether or not an industrial chimney produced a deposit on objects in the neighbourhood, circumstances like the wind direction, height of the chimney etc. may be of the utmost importance.

Some experts are reluctant to make statements about the circumstances of the case, arguing that if for some reason the expert was misinformed and the circumstances were in fact different then the report contains erroneous statements. However, if this is the case and the conclusion of the expert would be different had he known the real circumstances of the case, then it is only an advantage that the report makes it clear that the conclusion of the expert is based on the wrong information and should be reconsidered.

An infamous example of this is a Dutch murder case in a small town called Putten, in which a Dutch professor of gynaecology initially stated that a semen stain found on the thigh of the suspect could originate from (consensual) intercourse at an earlier time and could have been dragged out of the vagina when she was raped by the offender. This statement was crucial because it meant that the sperm could be from an unknown boyfriend and was not necessarily left by the offender. The two suspects, neither of whom matched the DNA profile of the sperm stain, were convicted, no

<sup>1</sup> In the real case, the formulation differed somewhat.

doubt partly because they both confessed to having committed the crime. However, after many years the gynaecologist stated for the High Court that he had been misinformed by the police about the circumstances of the case, and now came to a completely different conclusion. This new statement was considered a *novum*, a new fact, which by Dutch law is essential in order to for the Supreme Court to order a revision of the case.

The second step in reporting in a “Bayesian” style is to clearly state which hypotheses (more than one) are considered by the expert. In current practice, most reports will mention only the question from the “customer” (usually the police, the prosecution or an examining judge; the question may be as vague as “perform relevant investigations”), and draw a conclusion on the probability of a single hypothesis, e.g., “the print was probably made by this shoe”. However, this issue deserves much more attention. As Inman and Rudin [10] put it “probably the biggest favour most analysts could do for themselves and our profession would be to start formulating and explicating the scientific hypotheses that they will be testing with their analyses”. For example, if a crime is committed on a military base and a shoe print is compared to the shoe of a suspect soldier, the alternative hypothesis could be “the print was made by some unknown shoe” but it may be more appropriate to consider “the print was made by the shoe of some other soldier”.

In the third step, the evidence is described, that is, the results of the investigations. Statistics presented in tables or plots may be very useful in this step. Statistics also play an important role when numerical measurement results have to be reported. For instance, the level of alcohol found in the blood of the deceased, or the concentration of some contaminant in a waste lot. The uncertainties in the outcomes are not always reported, because they are not known exactly for the specific case, or because the expert finds it hard to convey concepts like a standard deviation. However, at the instigation of the Daubert case in the United States, a case in which the US Supreme Court considered how a judge can determine whether scientific evidence is admissible, accrediting bodies and courts will no longer accept the omission of uncertainty in forensic reports.

The fourth step, the interpretation of the evidence in light of the hypotheses considered and the circumstances of the case, is sometimes completely missing in forensic reports. In this step the expert explains how the results lead to the conclusion. This is therefore the essential part of the report and should not be left out, unless the results trivially lead to the conclusion. For

example, an expert may report as results that the elemental compositions of two samples taken at a fire scene are similar, but from this it does not trivially follow that therefore they probably came from the same source (e.g. the same jerry can with gasoline). The power of the analytical method to distinguish between different sources is essential information to interpret the results and simply should not be missing in a scientific report.

In the final step, the expert draws a conclusion. This is often a statement about the probability of a hypothesis, given the evidence. However, as numerous publications have argued (see e.g. [8]), this should be vice versa, the probability of the evidence given the hypotheses, or the LR expressing the strength of the evidence. For example, the expert should not report “this letter was probably written by the suspect”, but should report e.g. “there is strong evidence to support that the suspect wrote this letter”. Most laboratories use a standard set of terms to express their opinion. Like these “classical” conclusion scales, the Bayesian framework also has a wide variety of scales (see [3], chapter 3.5.3. for a recent overview). For example, the above conclusion could also be stated as “the observations are much more likely if the suspect wrote this letter than if somebody else wrote it”.

## 5. Conclusion

Statistics and forensic science have a lot to offer to each other. More interaction is desirable for both.

## References

1. Aitken C. G. G., Lucy D., Evaluation of trace evidence in the form of multivariate data, *Applied Statistics* 2004, 53, 109–122.
2. Aitken C. G. G., Stoney D. A., The use of statistics in forensic science, Ellis Horwood Limited, Chichester 1991.
3. Aitken C. G. G., Taroni F., Statistics and the evaluation of evidence for forensic scientists, Wiley & Sons, Chichester 2004.
4. Balding D. J., Weight-of-evidence for forensic DNA profiles, Wiley & Sons, Chichester 2005.
5. Buckleton J., Triggs C. M., Walsh S. J., Forensic DNA evidence interpretation, CRC Press, Boca Raton 2005.
6. Curran J. M., Hicks T. N., Buckleton J. S., Forensic Interpretation of Glass Evidence, CRC Press, Boca Raton 2000.
7. ENFSI Drugs Working Group, guidelines on representative drug sampling, Opmeer Drukkerij, Den Haag, the Netherlands, [www.enfsi.org](http://www.enfsi.org).

8. Evett I. W., Avoiding the prosecutor's fallacy, *Science & Justice* 1995, 35, 127–131.
9. Evett I. W., Weir B. S., *Interpreting DNA evidence – statistical genetics for forensic scientists*, Sinauer Associates, Sunderland 1998.
10. Inman K., Rudin N., *Principles and practice of criminalistics – the profession of forensic science*, CRC Press, Boca Raton 2001.
11. Keith L. H., *Principles of environmental sampling*, American Chemical Society 1996.
12. Lucy D., *Introduction to statistics for forensic scientists*, John Wiley & Sons, Chichester 2005.
13. Miller J. N., Miller J. C., *Statistics and chemometrics for analytical chemistry*, Pearson Education Limited, Harlow 2000.
14. Montgomery D. C., *Design and analysis of experiments*, John Wiley & Sons Inc., New York 1984.
15. Mortera J., Dawid A. P., Lauritzen S. L., Probabilistic expert systems for DNA mixture profiling, *Theoretical Population Biology* 2003, 63, 191–205.
16. National Research Council Committee on DNA forensic science: an update. The evaluation of forensic DNA evidence, National Academy Press, Washington 1996.
17. Robertson B., Vignaux G. A., *Interpreting evidence – evaluating forensic science in the courtroom*, John Wiley & Sons, Chichester 1995.
18. Robinson G. K., *Practical strategies for experimenting*, John Wiley & Sons, Chichester 2000.
19. Ryan T. P., *Statistical methods for quality improvement*, John Wiley & Sons, New York 2000.
20. Stelling M., Sjerps M., A sampling and hypothesis testing method for the legal control of solid hazardous waste, *Environmetrics* 1999, 10, 247–259.
21. Stelling M. A., Sjerps M., Heterogeneity of contaminants in solid waste materials sampled for compliance testing, *Environmental Forensics* 2005, 6, 231–239.
22. Taroni F., Aitken C. G. G., Garbolino P., Biedermann A., *Bayesian networks and probabilistic inference in forensic science*, John Wiley & Sons, New York 2006.
23. Taroni F., Biedermann A., Garbolino P. [et al.], A general approach to Bayesian networks for the interpretation of evidence, *Forensic Science International* 2004, 139, 5–16.
24. Waddell R. J. H., NicDaéid N., Littlejohn D., Classification of ecstasy tablets using trace metal analysis with the application of chemometric procedures and artificial neural network analysis, *The Analyst*, 2004, 129, 235–240.

---

**Corresponding author**

Marjan Sjerps  
Netherlands Forensic Institute  
P.O. Box 24044, 2490 AA, the Hague,  
the Netherlands  
e-mail: m.sjerps@nfi.minjus.nl

---

## ROLA STATYSTYKI W PRAKTYCE BIEGŁEGO SĄDOWEGO

### 1. Wstęp

Laboratoria pracujące dla potrzeb wymiaru sprawiedliwości wykonują badania i wydają opinie, wykorzystując wiele różnych dziedzin wiedzy. Statystyka jest jedną z nich. Niemniej jednak podstawową rolą tej nauki jest umożliwienie wyciągania wniosków na podstawie rezultatów przeprowadzonych badań. Szybki rozwój analizy DNA i konieczność oceny wartości dowodowej uzyskiwanych wyników zwiększyły zapotrzebowanie na zastosowanie metod statystycznych, szczególnie związanych z tzw. podejściem bayesowskim. Liczne publikacje opisują zastosowanie tego podejścia do analizy danych uzyskanych w trakcie przeprowadzonych badań różnego rodzaju materiału dowodowego (zob. np. [3]). Literatura dotycząca statystyki zawiera liczne informacje o innych metodach, które również są użyteczne w praktyce opiniodawczej. Niestety wiele z nich nie jest stosowanych. Na przykład w książce Currana i in. [6] opublikowano opis metod, które mogą być zastosowane w przypadku badań okruchów szkła, ale wielu biegłych ciągle interpretuje tego typu dowody na podstawie danych wizualnych. Taki stan rzeczy spowodowany jest w dużej mierze brakiem wiedzy statystycznej u biegłych sądowych oraz brakiem wsparcia ze strony statystyków. Ponadto wciąż brakuje łatwo dostępnych programów komputerowych.

W niniejszym artykule przedstawiono rolę statystyki w naukach sądowych, tzn. jej przydatność w interpretacji danych uzyskanych w trakcie badań materiału dowodowego oraz formy przedstawiania tych danych w opiniach. Wniosek końcowy brzmi, że konieczna jest intensyfikacja współpracy pomiędzy statystykami i biegłymi sądowymi.

### 2. Rola statystyki w pozyskiwaniu danych

Analiza materiału dowodowego często zaczyna się od pobrania reprezentatywnej próbki. Należy na przykład wybrać kilka sztuk tabletek lub opakowań zawierających nielegalne substancje z dużej partii zakwestionowanego materiału, które trafiają do laboratorium w celu ich przebadania. Innym przykładem jest sytuacja, gdy z samochodu, którym podejrzany prawdopodobnie spowodował wypadek, oddalając się z miejsca zdarzenia, zostaje zabezpieczonych kilka fragmentów szkła w celu porównania ich z fragmentami szkła znalezionymi na odzieży ofiary. W sytuacji, gdy próbki badane są np. pod kątem skażenia środowiska, pobiera się je z dużej partii odpadów toksycznych. Jeżeli można założyć, że cała partia materiału jest jednorodna, to w celu wydania pełnej opi-

nii wystarczy przekazać do analizy jedną próbkę. Niestety materiał, z którego pozyskiwana jest próbka, często nie jest jednorodny i biegły musi podjąć decyzję o liczbie, wielkości i miejscu pobrania odpowiednich próbek. Takie decyzje powinny być zatem oparte na analizie statystycznej. W przypadku pobierania próbek obiektów policzalnych z dużej partii materiału, takich jak wspomniane wcześniej tabletki, informacje o sposobie postępowania można znaleźć w poradniku Grupy Roboczej ds. Badania Narkotyków działającej w ramach Europejskiej Sieci Instytutów Nauk Sądowych [7]. Dodatkowe informacje o tym, jak rozwiązywać podobny problem, można odszukać w książce Lucy'ego [12, rozdział 16] oraz Aitkena i in. [3, rozdział 6]. Natomiast problem pobierania reprezentatywnych próbek środowiskowych jest omówiony w wielu innych publikacjach (zob. np. [11] – ogólnie o teorii pobierania próbek, [20] – pobieranie próbek w przypadku nauk sądowych). Należy dodać, że problem pobierania próbek do analizy jest również bardzo ważny w przypadku tworzenia baz danych.

Inny formą działań biegłych sądowych, w których statystyka odgrywa ważną rolę, jest kontrola otrzymanych danych analitycznych. W wielu pracowniach instytutów nauk sądowych regularnie używa się takich narzędzi, jak karty kontrolne lub krzywe kalibracyjne. Proces kalibracji jest czasem skomplikowany, szczególnie wtedy, gdy nie jest znany tzw. efekt matrycy. Ponadto przy wprowadzaniu nowych technik analitycznych wymagane jest, aby była ona poddana procedurze walidacyjnej poprzez porównanie uzyskanych za jej pomocą wyników z wynikami otrzymanymi inną metodą analityczną. W związku z tym odpowiednie metody, które mogą być pomocne biegłym sądowym, również oparte są na metodach statystycznych [12, 13, 19].

Projekty badawcze prowadzone w celu rozwiązywania problemów istotnych dla nauk sądowych, jak też eksperymenty wykonywane w celu wydawania opinii w konkretnej sprawie, wymagają zastosowania metod z innej dziedziny statystyki, tj. planowania eksperymentów [14, 18]. Na przykład biegli sądowi analizujący DNA, którzy zainteresowani są wpływem składu chemicznego luminolu na ujawnienie śladów DNA w świetle widzialnym, powinni zaplanować eksperyment w taki sposób, aby różne składniki tego związku występowały w różnych próbkach w różnym składzie w sposób systematyczny. Planowanie eksperymentu pozwala bowiem ograniczyć ilość przeprowadzanych badań, czyli zredukować koszty, a jednocześnie umożliwia wyciągnięcie pełnych wniosków. Ekspert z zakresu balistyki może zaprojektować eksperyment w celu zbadania, czy wyjaśnienia podejrzanego, który twierdzi, że broń wystrzeliła pomimo jej

zabezpieczenia, są prawdziwe. Również w tej sytuacji zastosowanie metod planowania eksperymentów powoli na minimalizację liczby koniecznych do wykonania doświadczeń przy jednoczesnym uzyskaniu pełnych wniosków z tak przeprowadzonych badań.

Podsumowując, należy stwierdzić, że metody statystyczne rozwinięte w ramach specjalistycznych dziedzin statystyki, takich jak pobieranie próbek, statystyczna kontrola jakości, chemometria czy planowanie eksperymentów, muszą być stosowane przez biegłych sądowych po to, by ich badania były wykonane zgodnie z obecnie przyjętymi standardami naukowymi.

### 3. Rola statystyki w interpretacji danych eksperymentalnych

Gdy biegły sądowy uzyska dane w trakcie analizy materiału dowodowego, to wówczas musi je zinterpretować tak, aby na ich podstawie oszacować wartość poszczególnych dowodów. Szybki rozwój metod analizy DNA, jaki dokonał się podczas ostatnich dwudziestu lat, stał się źródłem licznych problemów i pytań związanych z interpretacją tego typu danych. Na przykład, jak powinno być liczone prawdopodobieństwo zgodności profili DNA? Jaka jest odpowiednia populacja odniesienia, którą uwzględnić się w obliczeniach? Jaka jest wartość dowodowa mieszaniny komórek pochodzących od dwóch lub więcej osób, tzn. jak interpretować rezultaty analizy mieszaniny DNA? Jak określać związek pomiędzy osobami na podstawie ich profili DNA? Czy możemy określić grupę etniczną, do której należy nieznanemu dawca spermy? Jak wpływa liczba wykonanych porównań na wartość dowodową zgodnego DNA? Te i inne pytania są przedmiotem dyskusji w wielu publikacjach [4, 5, 9, 16].

Problemy związane z interpretacją profili DNA spowodowały również rozwój tzw. „oszacowania wartości dowodowej DNA na podstawie teorematu Bayesa” (określanego często jako „podejście logiczne” lub „iloraz wiarygodności”), które jest również metodą służącą do oceny wartości różnych materiałów dowodowych. Podejście to akceptuje większość statystyków zajmujących się naukami sądowymi, ale (z wyjątkiem opinii genetycznych [2, 3, 12, 17]) nie jest ono powszechnie stosowane w opiniowaniu.

W przytoczonych publikacjach wymienionych jest kilka błędów popełnianych przy interpretacji danych, powszechnie spotykanych w opiniach biegłych, np. błąd prokuratora czy błąd obrony. Ponadto we wspomnianych opracowaniach ich autorzy opisują, jak wartość dowodowa analizowanego materiału może być w praktyce wyrażona w postaci liczbowej, czyli w postaci ilorazu wiarygodności (*LR*). Wnioskowanie oparte tym ilorazem posiada wiele zalet, choć nierzadko zdarza się, że iloraz wiarygodności nie może być obliczony z powodu braku

danych. Użyteczność metody *LR* w naukach sądowych związana jest również z tym, że jest ona swego rodzaju przewodnikiem dla biegłych sądowych, mówiącym, jak poprawnie formułować hipotezy odnoszące się do problemu oszacowywania wartości dowodowej analizowanej próbki. Wymusza ona bowiem na biegłym sądowym analizowanie rezultatów wykonanych badań nie tylko z punktu widzenia hipotezy prokuratora, ale również z punktu widzenia obrony (tzn. aby biegły rozważył hipotezę alternatywną). Co więcej, podejście oparte na ilorazie *LR* pozwala na ujednoczenie wydawanych opinii przez biegłych sądowych, co w dużym stopniu przyczynia się do standaryzacji opinii wydawanych przez różnych biegłych. Numeryczna skala *LR* wprowadza bowiem punkt odniesienia, który posiada uniwersalne zastosowanie. Ponadto podejście to może zostać użyte przy wnioskowaniu o wartości dowodowej w przypadku, gdy rozważanych jest jednocześnie kilka dowodów.

Rozważmy jako przykład następującą autentyczną historię, która została oparta na przypadku analizowanym przez Holenderski Instytut Nauk Sądowych. Samochód terenowy potrafił chłopca. Auto, które mogło uczestniczyć w wypadku, poddano oględzinom. Na rurze wydechowej ujawniono fragment czerwonego materiału, który porównano z materiałem czerwonych spodni chłopca uszkodzonych w kilku miejscach. W pierwszej kolejności biegły z zakresu analizy włókien dokonał oględzin spodni i stwierdził, że uszkodzenia te zostały spowodowane przez działanie wysokiej temperatury. Drugi ekspert porównał kształty materiału zabezpieczonego z rury wydechowej z uszkodzeniami na spodniach i stwierdził, że są one podobne. Trzeci biegły porównał skład chemiczny materiału znalezionego na rurze wydechowej z materiałem, z którego wykonano spodnie i stwierdził, że są one wykonane z tego samego rodzaju tkaniny. Żadna z powyższych opinii nie jest opinią stwierdzającą, że materiał znaleziony na rurze wydechowej pochodził ze spodni chłopca, ale ich połączenie jest o wiele bardziej wymowne. Dlatego też biegli opracowujący tę opinię umieścili w niej wnioski, które związane były z następującymi hipotezami: a materiał ujawniony na rurze wydechowej pochodzi ze spodni chłopca, b materiał ujawniony na rurze wydechowej pochodzi z innego źródła. Biegli sądowi stwierdzili, że dane uzyskane w trakcie wykonanych badań silnie wspierają pierwszą hipotezę niż drugą<sup>1</sup>. Sąd uznał jednoznacznie, że docenia ten sposób opiniowania, tj. łączenie informacji pochodzących z badania różnego rodzaju dowodów.

Łączenie informacji o różnego rodzaju dowodach z różnego rodzaju hipotezami może być wykonane za pomocą sieci bayesowskich (zwanymi również jako probabilistyczne systemy eksperckie). Tego rodzaju modele istnieją w postaci programów komputerowych pozwalających na analizowanie np. czułości ilorazu wiarygod-

<sup>1</sup> W rzeczywistości twierdzenie to brzmiało nieco inaczej.

ności na wartość konkretnego parametru. Kilka przykładów zastosowania tego rodzaju metod w naukach sądowych można znaleźć w literaturze [15, 22, 23], a obecnie prowadzonych jest wiele badań, na których są oparte.

Inna dziedzina statystyki, która dostarcza cennych rezultatów, choć nie jest powszechnie stosowana, to chemometria lub zastosowanie statystycznej analizy danych wielowymiarowych [24]. W ostatnich latach analiza chemiczna znacznie się udoskonalila (np. technika ICP-MS). Statystyczne procedury analizy dużej liczby danych również stały się skomplikowane i wymagają zastosowania programów statystycznych. Takie techniki, jak analiza głównych składowych lub techniki graficzne, np. różnego rodzaju wykresy rozrzutu, są bardzo użytecznymi narzędziami dla osób zajmujących się naukami sądowymi. Umożliwiają one analizę struktury danych wielowymiarowych w przeciągu bardzo krótkiego czasu i np. pozwalają udzielić odpowiedzi na pytania, które próbki są podobne, a które pierwiastki najbardziej użyteczne w celach dyskryminacyjnych, tj. pozwalających na rozróżnianie porównywanych obiektów. Jednocześnie biegły sądowy może stosować wielowymiarowe testy istotności w celu podjęcia decyzji, czy istnieją istotne statystycznie różnice pomiędzy analizowanymi próbkami (np. [6] w przypadku statystycznej analizy szkła). Ponadto biegły sądowy może obliczyć iloraz wiarygodności dla danych wielowymiarowych w celu ustalenia, czy próbka pochodzi z tego samego, czy z dwóch różnych źródeł [2].

Niestety obecnie niewielu biegłych sądowych używa którejś ze wspomnianych metod statystycznych w celu analizy danych i wyciągnięcia na tej podstawie wniosków dotyczących wartości dowodowej analizowanego materiału. Procedura ta kończy się z reguły na przeglądnięciu (czasem obszernej) tabeli z danymi, a następnie na dokładniejszej analizie tylko kilku zmiennych oraz wyciągnięciu wniosków na podstawie subiektywnych odczuć i przekonań. Niemniej jednak w niedługim czasie takie podejście będzie kwestionowane w sądzie, stając się przykładem złej praktyki badań przydatnych dla potrzeb wymiaru sprawiedliwości.

#### 4. Rola statystyki w opracowaniu opinii

Zastosowanie podejścia bayesowskiego do interpretacji materiału dowodowego wpływa również na to, w jaki sposób dane analityczne prezentowane są w opiniach biegłych. Po pierwsze, w opinii powinna zostać uwzględniona informacja o okolicznościach zdarzenia, ponieważ posiada ona istotne znaczenie w ocenie materiału dowodowego. Na przykład w przypadku analizy DNA lub jej mieszaniny okoliczności zdarzenia mają wpływ na dobór bazy danych, która zostanie zastosowana w celu obliczenia prawdopodobieństwa zgodności profili. Podczas ana-

lizy próbek środowiskowych, gdy pytanie dotyczy np. problemu, czy dymy wydobywające się z komina zakładu przemysłowego mogły spowodować powstanie osadu na obiektach w jego sąsiedztwie, udzielenie odpowiedzi wymaga danych o kierunku wiatru lub wysokości komina, bowiem wspomniane informacje mogą mieć istotne znaczenie przy udzielaniu poprawnej odpowiedzi.

Część biegłych niechętnie wypowiada się o okolicznościach sprawy, argumentując, że jeżeli z jakiś powodów biegły został źle poinformowany, a tym samym okoliczności sprawy były inne, to jego opinia zawierać będzie błędne wnioski. Oczywiście w takim przypadku pierwotne wnioski biegłego będą różne od tych, które sformułowałby, posiadając wiedzę o prawdziwych okolicznościach zdarzenia, ale zmienione okoliczności nie umniejszą jego wiarygodności, ponieważ pierwotne wnioski były poprawne w świetle wcześniejszych danych o zdarzeniu. Jedyne, co powinien wykonać, to zinterpretować dane uzyskane w trakcie analizy materiału dowodowego w świetle nowych okoliczności zdarzenia.

Jako niechlubny przykład można przytoczyć przypadek morderstwa w Putten, małym miasteczku w Holandii. Powołany jako biegły profesor ginekologii wnioskował początkowo, że sperma znaleziona na udzie kobiety mogła być wynikiem wcześniejszego stosunku, tj. mogła wypłynąć z pochwy w momencie, kiedy ofiara była gwałcona przez napastnika. Ta opinia okazała się decydująca, ponieważ oznaczała, że sperma może pochodzić od innego mężczyzny, niekoniecznie od gwałciiciela. Skazano dwóch podejrzanych, których profile DNA nie pasowały wprawdzie do profilu DNA zabezpieczonej plamy nasienia, lecz przyznali się oni do popełnienia tego przestępstwa. Po kilku latach ten sam ginekolog stwierdził jednak przed sądem apelacyjnym, że został źle poinformowany przez policję o okolicznościach zdarzenia i w chwili obecnej jego wnioski są całkowicie odmienne. Nowe wnioski uznano za nowe fakty, co w prawie holenderskim uznaje się za podstawę dla sądu najwyższego do ponownego rozpatrzenia sprawy.

Drugim etapem wydawania opinii w oparciu o podejście bayesowskie jest dokładne określenie, jakie hipotezy są rozważane przez biegłego. Obecnie w praktyce w większości opinii znajdują się tylko pytania i polecenia wymienione w postanowieniu wydanym przez organ zlecający jej wydanie (zazwyczaj jest to policja, prokuratura lub sąd). Mogą być one czasem niejasne, np. „należy wykonać istotne badania”. Wnioski opinii formułowane są natomiast w formie wyrażenia prawdopodobieństwa zajścia opisanego zdarzenia z punktu widzenia jednej tylko hipotezy, np. „odcisk spodu podeszwy buta jest prawdopodobnie śladem buta dowodowego”. Niemniej jednak takie podejście metodologiczne zasługuje na szczególną uwagę. Jak piszą Innam i Rudin [10], „prawdopodobnie największą przysługą, jaką analitycy mogą oddać sobie i naszej profesji, może okazać się początek



formułowania i tłumaczenia hipotez naukowych, które będą testowane w danym przypadku". Na przykład, jeżeli przestępstwo zostało popełnione w jednostce wojskowej i odcisk podeszwy buta był porównywany do odcisku podeszwy buta podejrzanego żołnierza, to wówczas alternatywna hipoteza może brzmieć: „odcisk podeszwy buta pozostawiła nieznaną osobą”, choć bardziej celowe byłoby rozważenie następującej hipotezy: „odcisk spodu buta pozostawił inny żołnierz”.

W trzecim etapie powstawania opinii opisywane są rezultaty analizy, zaś dane statystyczne prezentuje się w postaci tabel lub wykresów. Statystyka odgrywa tu istotną rolę, szczególnie gdy w opinii muszą zostać podane wartości uzyskanych pomiarów (na przykład poziom alkoholu we krwi zmarłego bądź zawartość zanieczyszczeń w uszkodzonej partii produktów). Nie zawsze jest określana niepewność uzyskanych pomiarów, np. wtedy, gdy nie jest ona dokładnie znana lub wówczas, gdy biegły sądowy uważa za trudne czy też niewłaściwe wyrażanie niepewności w formie odchylenia standardowego. Niemniej jednak po sprawie Dauberta w Stanach Zjednoczonych (Sąd Najwyższy Stanów Zjednoczonych rozważał, czy i jak sędzia może stwierdzić poprawność dowodu naukowego), sądy nie mogą już akceptować pomijania informacji o niepewności wyników w opiniach biegłych sądowych.

Czwarty etap powstawania opinii to interpretacja dowodu w świetle rozważanych hipotez i okoliczności zdarzenia, co jest często całkowicie pomijane. Biegły powinien wyjaśnić, w jaki sposób uzyskane rezultaty doprowadziły go do konkretnych wniosków. Ta część opinii jest niezwykle istotna i można z niej zrezygnować tylko wówczas, gdy wyniki badań w oczywisty sposób prowadzą do wniosków. Na przykład biegły może napisać w opinii, że skład chemiczny dwóch próbek pobranych z miejsca pożaru jest podobny, ale nie wynika z tego bezpośrednio (i jednoznacznie), że pochodzą one z tego samego źródła, np. z benzyny wylanej z konkretnego kanistra, ponieważ metody analityczne służące do analizy tego typu próbek nie dostarczają informacji, które byłyby wystarczające do sformułowania w opinii tak kategorycznego wniosku.

W ostatnim etapie pisanie opinii biegły sądowy opisuje wnioski końcowe. Najczęściej posiadają one formę wyrażającą stopień prawdopodobieństwa, że dana hipoteza jest poprawna w świetle danych, które uzyskano na temat analizowanego dowodu. Dzieje się tak pomimo licznych publikacji [8], które dowodzą, iż powinno być odwrotnie, tzn. powinno się oszacowywać prawdopodobieństwo, że dowód wspiera (bądź nie) daną hipotezę i (lub) przedstawia je w postaci ilorazu wiarygodności, który, jak wspomniano wcześniej, wyraża siłę danego materiału dowodowego. Na przykład biegły nie powinien formułować wniosku w następujący sposób: „ten list został najprawdopodobniej napisany przez podejrz-

nego”, ale powinien raczej napisać, że „jest to mocny dowód wspierający hipotezę, iż podejrzany napisał ten list”. W celu sformułowania wniosków używa się zestawu klasycznych zwrotów, które w tzw. podejściu bayesowskim występują w różnorodnych zestawieniach [1]. Na przykład wnioski w powyższym przykładzie można wyrazić następująco: „dostępne dane o dowodzie są znacznie bardziej wiarygodne w przypadku, gdy podejrzany napisał ten list, niż gdy go nie napisał”.

## 5. Wniosek

Statystyka i nauki sądowe mają sobie wzajemnie dużo do zaoferowania, ale też konieczne jest, aby lepiej przebiegała współpraca pomiędzy biegłymi sądowymi a statystykami.