

Forensic metrology: its importance and evolution

Alessandro Ferrero^{1,*} and Veronica Scotti^{1,**}

¹ Politecnico di Milano - DEIB - Piazza Leonardo da Vinci 32 - 20133 Milano, Italy

Abstract. Measurements play a fundamental role in nowadays forensic activities, especially in criminal justice. Guilt or innocence and the severity of a sentence do often depend upon the results of measurements. On the other hand, it is well-known that measurement results are always affected by uncertainty, so that any decision based on measurement results carries an implicit risk of being wrong: uncertainty helps in estimating this risk. If this concept is translated into the juridical world, it is quite immediate to recognize that uncertainty may quantify how reasonable is the doubt with which a verdict is rendered by the trier of facts. Unfortunately, forensic disciplines are still largely unaware of the basic, fundamental concepts of metrology. This paper is aimed at showing how forensic metrology is developing in different countries and reporting a few cases where metrology played a key role.

1 Introduction

The law reports of the last few decades have consistently shown that justice, and in particular criminal justice, is more and more often referring to scientific evidence to solve the cases. Scientific evidence is the result of experimental activity, so that guilt or innocence and the severity of a sentence do often depend upon the results of measurements. It is common knowledge that sanctions against speed limit violations are issued on the basis of speed measurements performed by radars or other instruments, that sanctions (and sentences to imprisonment) for driving under the influence (DUI) of alcohol or drugs are issued on the basis of breath or blood tests, and that the perpetrators of crimes are very often identified through DNA analysis, bitemark analysis or hair analysis.

The main reason for this is that science is generally considered by non scientists, such as the triers of facts are, as capable of providing a more certain reconstruction of the facts than the one that could be obtained with more traditional methods, such as testimonies. On the other hand, scientists are well aware that science has intrinsic limitations, and cannot represent reality with full certainty. In particular, when scientific evidence is obtained through measurement results, it is well-known that these results can never provide the true value of the measurand, but only incomplete information about the measurand itself. Metrology quantifies this lack of complete information with measurement uncertainty [1, 2].

It can be therefore assumed, as the Guide to the Expression of Uncertainty in Measurement (GUM) [3] clearly states, that uncertainty in measurement is a synonym for *doubt*, that is it means doubt about the validity

of the result of a measurement. On the other hand, one of the principles on which all law systems are based is that any decision rendered by the trier of facts must be rendered *beyond any reasonable doubt*. Since one of the pillars on which metrology is based is the capability of evaluating and expressing uncertainty, a measurement result, if correctly expressed together with its associated uncertainty, provides also an evaluation of the inherent doubt about its validity. Consequently, when the trier of facts bases his or her decision on a measurement result, if uncertainty is provided, he or she is given an important piece of information to consider whether the remaining doubt about the correctness of the decision is reasonable or not.

Scientific evidence has gained such an importance in the courts of justice, that the validity of the forensic science methods, as applied in tribunals, has been widely discussed in the USA, where a recent document has been published by the President's Council of Advisors on Science and Technology (PCAST) [4]. Yet, this important document neglects the importance of metrology in assessing the validity of these methods, and neglects the recent application field of metrology, *forensic metrology* [5–8], that is more and more referred to by all figures working in the forensic field.

This paper is aimed at proposing a short survey on how forensic metrology developed and how it can provide good answers to the questions raised by the PCAST report. At last, a few examples of cases where scientific evidence played a key role and have been solved thanks to metrology will be reported.

2 The evolution of forensic metrology

Forensic disciplines have been, and still are, largely unaware of the basic, fundamental concepts of metrology.

*e-mail: alessandro.ferrero@polimi.it

**e-mail: veronica.scotti@gmail.com

Law people (judges, prosecutors and lawyers) tend to trust science and the results of scientific tests (such as DNA analysis) as ultimate and conclusive evidence and do believe that verdicts based on such results are always rendered “beyond any reasonable doubt”. This was the firm common belief in the late 70s of the 20th century, when the International Association for Identification (I.A.I.) adopted a resolution, as reported in [9], where providing courtroom testimony describing an identification through latent prints as “possible, probable or likely”, rather than “certain”, was considered a professional misconduct for its members.

This conviction that experts called to testify in a trial must provide certainty to the trier of facts remained deep-rooted for a long period of time, at least until judge Blackmun of the US Supreme Court, in the famous Daubert case, stated, in 1993, that “[I]t would be unreasonable to conclude that the subject of scientific testimony must be ‘known’ to a certainty; arguably there are no certainties in science” [9].

As a consequence of this sentence, also the US lower courts began to consider the experts who testified taking into account uncertainty, or a safety margin, more reliable than those who did not. According to this new trend, some US lawyers started to openly refer to measurement uncertainty whenever measurement results were considered as pieces of evidence [10]. It is not surprising that the cases where uncertainty has first become an important additional piece of evidence were those related to driving under the influence (DUI) of alcohol or drugs, since the amount of alcohol or drug in the driver’s blood is measured by dedicated instruments, whose uncertainty should be preliminarily evaluated and considered in assessing whether the alcohol or drug content has exceeded the tolerated limit [11].

Such cases represent a typical example of how, absent an uncertainty value, the measured values may lead to an incorrect decision. Let us suppose, for instance, that a value x_m is measured for a quantity x on a given scale, as shown in Fig. 1, where uncertainty is not given.

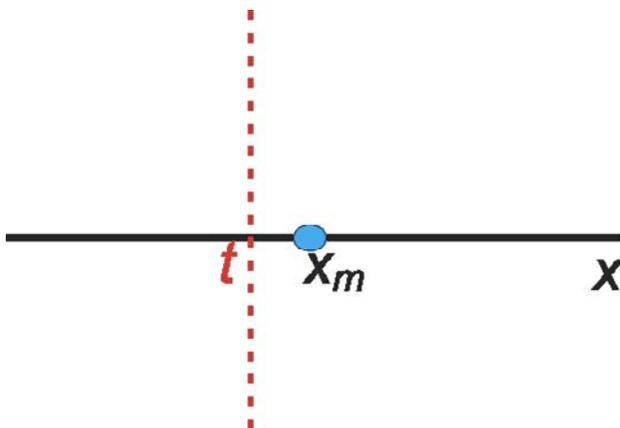


Figure 1. Example of measured value with no uncertainty declared. t : threshold value

If quantity x is prescribed not to exceed a threshold value t , such situation is apparently showing that the threshold value has been exceeded with full certainty, so that the trier of fact will quite likely render a verdict of guilt beyond any reasonable doubt.

However, once the standard uncertainty is evaluated and expressed, it is possible, according to the GUM [3] recommendation, to build an interval, about the measured value, into which the measurand value is supposed to lie with a given coverage probability, that depends on the assumed coverage factor and probability distribution. Fig. 2 shows the same situation as that of Fig. 1, where uncertainty is declared, and a normal probability distribution is assumed for the distribution of values that could be reasonably attributed to the measurand x .

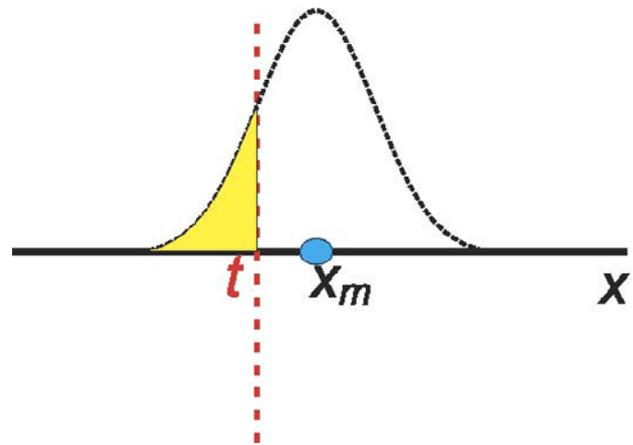


Figure 2. Example of measured value where uncertainty is declared. t : threshold value. Yellow area: probability that the value of the measurand falls below the threshold

It can be immediately recognized that a probability exists that the measurand value is lower than the given threshold, and this probability is given by:

$$P\{x \leq t\} = \int_{-\infty}^t p(x)dx \quad (1)$$

where $p(x)$ is the considered probability density function. This probability is graphically represented, in Fig. 2, by the yellow-shaded area under $p(x)$.

It can be readily perceived that, the closer is the measured value to the threshold and the higher is uncertainty, the higher is the probability that the measurand is below the threshold even if the measured value is above the threshold. In other words, the trier of facts might not be able to render a decision beyond any reasonable doubt, and uncertainty, if properly evaluated and used, can help him or her in quantifying this doubt. It is evident that not providing uncertainty is the same as hiding an important piece of evidence to the trier of facts.

3 Forensic metrology and the PCAST report

As stated in the last section, the methods of the forensic metrology have been initially adopted by lawyers, par-

ticularly in fields where the triers of facts base their decisions directly on measurement results (such as in DUI cases), and have been positively considered mainly by the courts of common law countries, where the adopted law system empowers judges with more discretionary power than other law systems.

More recently, a debate has started in the US, about the validity of the forensic science methods adopted in the courts of justice. The already mentioned PCAST report [4] is the most recent and pondered answer to the questions raised. Nevertheless, this report does not yet consider metrology in an explicit way, and this is, probably, its major weak point.

Indeed, this report was originated by a 2009 report on forensic science, issued by the National Research Council [12], which was quite critical on the real validity of the employed methods and caused a sensation. In the attempt to answer to the critical comments moved by this report, the PCAST report considers the most used feature-comparison methods and considers general rules to establish the scientific validity of DNA analysis, latent fingerprint analysis, bitemark analysis, firearm analysis, footwear analysis and hair analysis.

In particular, scientific validity is defined considering two different aspects: *foundational validity* and *validity as applied*. The following definitions are given.

- *Foundational validity* for a forensic-science method requires that it be shown, based on empirical studies, to be *repeatable*, *reproducible*, and *accurate*, at levels that have been measured and are appropriate to the intended application. Foundational validity, then, means that a method can, *in principle*, be *reliable*.
- *Validity as applied* means that the method has been reliably applied *in practice*.

Despite the extensive discussion on how the validity of the considered feature-comparison methods can be assessed in terms of the two above definitions, the report fails to provide a quantitative evaluation on *how valid* these methods are. However, this would have been relatively easy if also *uncertainty* were considered, together with the already mentioned concepts of *repeatability*, *reproducibility* and *accuracy*.

Indeed, both given definitions of foundational validity and validity as applied can be interpreted in terms of uncertainty contributions. Let us consider, first, the definition of *definitional uncertainty* given by the International Vocabulary of Metrology (VIM) [13]: *component of measurement uncertainty resulting from the finite amount of detail in the definition of a measurand*. Considering that this component of measurement uncertainty represents also the practical minimum measurement uncertainty achievable in any measurement of a given measurand, since it depends on the limited knowledge about the measurand, it does quantify also how a *method can, in principle, be reliable*. Therefore, the definitional uncertainty can be seen as the metrological counterpart of the definitional validity defined by the PCAST report, with the additional advantage that methods are available to quantify it.

Let us now consider the definition of *instrumental uncertainty* given by the VIM [13]: *component of measurement uncertainty arising from a measuring instrument or measuring system in use*. It is evident that this uncertainty contribution is the most relevant one in assessing whether a method can be reliably applied in practice and is, therefore, the metrological counterpart of the validity as applied. Strictly speaking, the validity as applied can be quantified by the combination of the definitional and instrumental uncertainty contributions.

To better qualify these concepts, let us briefly consider the validity of the DNA analysis used in DNA profiling, that represents, nowadays, one of the most relevant pieces of evidence in crime solution. Validity of DNA profiling is generally discussed only in terms of wrong match probability, that is the probability that two individuals share the same DNA components used to profile a single individual. It has been proven [14] that the wrong match probability represents only the definitional uncertainty, and that this contribution might be order of magnitude lower than all other uncertainty contributions that might be originated in the references used in the analysis, as well as in the possible sample contamination, amplified by the DNA amplification methods employed when the available material is not enough to run the test [14, 15].

It is then possible to conclude that metrology can provide a solid contribution toward a quantitative evaluation of the validity of forensic science methods and the provided results. Consequently, it helps the trier of facts to estimate the remaining doubt on the validity of a decision rendered on the basis of these results, as the cases reported in the following section show.

4 Case studies

The following three cases, two based on DNA profiling and one on breath alcohol tests, show, in a quite significant way, the problems that neglecting uncertainty may cause.

4.1 The case of the bartender in Liverpool

In Litherland, a small village north of Liverpool, UK, a bartender was quietly attending at his duties when, on a day of February 2003, the police collared him under indictment for murdering a 24 year-old Italian girl during a robbery attempt in Castiglioncello, a small village on the Tuscany coast, not far from Florence, Italy, on the previous August. The DNA found in the abundant trail of blood left on the crime scene by the mugger, after the victim's boyfriend hit his face with a stone, in the attempt to protect the victim, was found compatible with that of the bartender with full certainty, and the poor bartender was considered guilty beyond any reasonable doubt [16].

The bartender protested his innocence, stating that he never went to Italy, but nobody believed him, since the DNA was considered conclusive and undeniable evidence. Luckily for the bartender, the regular customers of the pub where he worked gave evidence that they saw him at work on the day the murder was committed. Even more luckily,

on that same day, the pub's owner was at the funeral ceremony of a relative. The pub was open, the bartender was the only employee, so he was the person behind the bar, on that day, without any possible doubt. In the end, he was cleared from the charge, despite the DNA analysis.

This is one of the best-known examples that scientific tests may provide incorrect results. In this case, measurement uncertainty was not considered, and nobody considered whether the performed test had scientific validity or not. However, after this case, the foundational validity of the DNA profiling tests was reconsidered, and the number of DNA alleles considered by UK labs was increased from 9 to 13, and uniformed to the number used in the USA to profile individuals.

A more careful consideration of both the definitional and instrumental uncertainty contributions would have saved the bartender from a big fear and, most importantly, could have avoided a big miscarriage of justice that only the presence of so many undeniable witnesses avoided.

4.2 The Perugia case

This case had a large resonance on the crime news in Italy, UK and US, due to the nationality of the involved people [7]. A young British student, Meredith, attending the University in Perugia, Italy, was killed, in 2007, with several stab wounds inflicted by a kitchen knife. The initial investigation led to the discovery of such a knife in the house of two school mates of Meredith, Raffaele, an Italian student, and Amanda, his American girl-friend. The coroners who examined the victim assessed that this knife was compatible with her wounds, so it could have been the murder weapon.

A DNA analysis was ordered and the report filed by the expert who was asked to perform the analysis stated that Raffaele's and Amanda's DNA were present on the knife handle (which was not surprising at all, since this was their kitchen knife), and that Meredith's DNA was present on the blade. The report stated also that the quantity of biological material present on the knife blade was quite small and insufficient to repeat the test. The important point, here, is that uncertainty was estimated, but was reported without the necessary emphasis, so that it was not considered during the trial.

As a consequence of this test and the way it was reported to the trier of facts, the presence of Meredith's DNA on the blade of Amanda's and Raffaele's knife was considered conclusive evidence that Amanda and Raffaele killed Meredith, and they were convicted of murder and sentenced to twenty-six years imprisonment.

The defense appealed against this verdict, and a second trial started in front of a different judge. According to the Italian law system, the judge of the second trial can order the repetition of the experimental tests considered during the first trial or, if this is not possible, a new analysis of the results can be requested.

Since the repetition of the DNA analysis was not possible, the judge asked a team of experts to re-analyze the report filed by the expert who performed the DNA analysis

and provide their opinion on the results. This team considered the estimation of measurement uncertainty done by the expert who performed the DNA analysis and discussed it and the different contributions to uncertainty in deep details. The conclusion was that the values measured during the test were below the uncertainty value, and therefore, it was not possible to state whether the DNA on the blade was that of Meredith beyond the reasonable doubt quantified by measurement uncertainty. Surprisingly enough, this conclusion was already present, though not clearly exposed, in the original report, but was totally ignored by the trier of facts of the first trial.

The judge accepted the fact that the DNA test could not be considered conclusive evidence and returned a verdict of not guilty for Amanda and Raffaele, who were released after four years of imprisonment. Unfortunately, this was not the end of the story, because three more proceedings were celebrated before a final verdict of not guilty was returned.

For the sake of brevity, we do not consider, here, the judicial details of these proceedings, since, in the end, they confirmed the verdict of the first appeal trial. On the other hand, it is worth noting that the absence of a clear indication about measurement uncertainty led the trier of facts to assign the DNA test the importance of fully certain conclusive evidence, thus rendering a verdict beyond any reasonable doubt, whilst the doubt existed and was way more than reasonable, as recognized by the judge of the appeal trial and those of the Italian Supreme Court (Corte di Cassazione) who, finally, accepted those conclusions and rendered a verdict of not guilty.

In other words, measurement uncertainty, in this case, represented a major and important piece of evidence, and having neglected it led to an unfair verdict toward the two defendants.

4.3 The case of the alcohol test

As aforementioned, the DUI cases represented the first cases of successful application of forensic metrology in the USA [10, 11, 17]. This is no exception also in Europe. In the past, the punishment was reduced, or not at all inflicted, only when tests could prove that the instrument used to perform a breath alcohol test was not working properly. A recent Italian verdict, however, reduced a DUI charge on the basis of mere metrological considerations about uncertainty associated with the measurement result [18].

A driver was stopped and a breath alcohol concentration (BAC) test was run. The resulting BAC was 1.51 g/l of equivalent blood alcohol concentration. The Italian law considers imprisonment up to 12 months, confiscation of the driving license up to 2 years, confiscation of the vehicle and a fine up to 6,000 € if the alcohol concentration in blood exceeds 1.5 g/l.

The defense proved that, according to the laws in force and absent any valid calibration of the instrument used for the BAC test, the manufacturer specification could be used to evaluate how good the measured alcohol concentration was.

The expert witness called by the defense proved that, according to the manufacturer specifications, the instrument might drift for up to ± 0.02 g/l per month after calibration. Moreover, the manufacturer declared a standard uncertainty of 0.023 g/l for the considered range. It was easy to prove that one month after calibration (whose date was unknown), the instrument might have measured an alcohol concentration value above the 1.5 g/l threshold even with a concentration value of 1.49 g/l, below the 1.5 g/l threshold.

Moreover, assuming a normal probability distribution for the values that can be reasonably attributed to the measurand, the uncertainty value provided by the manufacturer returns a probability of 32.5% that the value of the measurand is below the 1.5 g/l threshold for a measured value of 1.51 g/l.

The trier of facts accepted the defense reasoning, supported by the related uncertainty evaluation, and reduced the charge as if the measured blood alcohol concentration was in the range immediately below 1.50 g/l, thus proving that, when uncertainty is correctly evaluated and explained, its importance is correctly understood by the triers of facts.

5 Conclusions

Forensic metrology is a new field of application of metrology and is gaining importance as scientific evidence is becoming more and more relevant in establishing facts to identify the perpetrators of a crime or misconduct.

People working in the administration of justice are slowly becoming aware that science cannot always provide a fully certain interpretation of the facts, especially when this interpretation is based on measurement results. On the other hand, as proved by the PCAST report, evaluating how “reliable” a forensic science method is, in order to quantify the risk of unfair decision, is not an easy task.

In this respect, metrology is a very helpful tool, since, through the evaluation of the definitional and instrumental uncertainty contributions, it is capable of quantifying the *foundational validity* and the *validity as applied* of forensic science methods.

Therefore, it can be concluded that measurement uncertainty is an important piece of evidence and does always play in favor of good and fair decisions and, therefore, it favors the certainty of law. On the other hand, disregarding uncertainty can be seen as if an important piece of evidence is kept hidden to, or not at all considered by the trier of facts. This opens important ethical, and not only technical or juridical issues.

References

- [1] A. Ferrero, Instrumentation Measurement Magazine, IEEE **18**, 7 (2015)
- [2] A. Ferrero, S. Salicone, Instrumentation Measurement Magazine, IEEE **9**, 44 (2006)
- [3] JCGM 100:2008, *Evaluation of Measurement Data – Guide to the Expression of Uncertainty in Measurement, (GUM 1995 with minor corrections)*, Joint Committee for Guides in Metrology (2008), <http://www.bipm.org/en/publications/guides/gum.html>
- [4] President’s Council of Advisors on Science and Technology (PCAST), *Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature–Comparison Methods* (2016), https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/PCAST/pcast_forensic_science_report_final.pdf
- [5] T. Vosk, in *Encyclopedia of forensic sciences*, edited by J.A. Siegel, P.J. Saukko (Academic Press, Waltham, MA, USA, 2013), Vol. 3, pp. 322–331, ISBN 9780123821652
- [6] T. Vosk, A.F. Emery, E. Fitzgerald, *Forensic Metrology: A Primer on Scientific Measurement for Lawyers, Judges, and Forensic Scientists* (Taylor & Francis, New York, NY, USA, 2014), ISBN 9781439826195
- [7] A. Ferrero, V. Scotti, Instrumentation Measurement Magazine, IEEE **16**, 14 (2013)
- [8] V. Scotti, *Forensic Metrology: Where Law Meets Measurements*, in *20th IMEKO TC4 Symposium Measurements of Electrical Quantities* (2014), pp. 385–389, ISBN 9781634394253
- [9] E.J. Imwinkelried, UC Davis Legal Studies Research Paper No. 317 pp. 1–33 (2012)
- [10] T. Vosk, Journal of Physics: Conference Series **772**, 1 (2016)
- [11] T. Vosk, The champion pp. 48–56 (2010)
- [12] N.R.C. Committee on Identifying the Needs of the Forensic Sciences Community, *Strengthening Forensic Science in the United States: A Path Forward* (2009), <http://www.nap.edu/catalog/12589.html>
- [13] JCGM 200:2012, *International Vocabulary of Metrology – Basic and General Concepts and Associated Terms (VIM 2008 with minor corrections)*, Joint Committee for Guides in Metrology (2012), <http://www.bipm.org/en/publications/guides/vim.html>
- [14] A. Ferrero, Instrumentation Measurement Magazine, IEEE **20**, 4 (2017)
- [15] C.M. Cale, M.E. Earll, K.E. Latham, G.L. Bush, Journal of Forensic Sciences **61**, 196 (2016)
- [16] A. Ferrero, V. Scotti, Instrumentation Measurement Magazine, IEEE **18**, 18 (2015)
- [17] T. Vosk, E.J. Imwinkelried, Criminal Law Bulletin **53**, 532 (2017)
- [18] A. Ferrero, V. Scotti, Tutto_Misure (in Italian) **18**, 259 (2016)