AUTHORS' POSTPRINT

Sironi, E., Bozza, S. and Taroni, F., Chapter 4.3: Age estimation of living persons: a coherent approach to inference and decision. In: Z. Obertova, A. Stewart, C. Cattaneo (Eds), Statistics and Probability in Forensic Anthropology. Academic Press, London (2020) 183-208.

5.1 Age estimation of living persons: a coherent approach to inference and decision

Emanuele Sironi[1], Silvia Bozza[2,1], Franco Taroni[1]

[1] School of Criminal Justice, University of Lausanne, Switzerland

[2] Department of Economics, University Ca' Foscari of Venice, Italy

Abstract

Evidence interpretation is a fundamental aspect of forensic science: it is basically a problem of inference and decision. Forensic age estimation is no exception to this reality. Evidence related to the biological development of an individual is often relevant from a legal perspective, such as when examining the probability that a person is younger or older than a given age threshold, for instance the age of majority. Provided that uncertainty in forensic evidence should be measured by means of probability, the Bayesian approach represents the ideal solution for both inference and decision. This contribution aims to illustrate how this perspective operates in age estimation from both theoretical and operational points of view.

Keywords

Bayesian approach
Forensic Age estimation of living persons
Inductive inference
Normative decision theory
Forensic evidence evaluation and interpretation
Subjective probability
Forensic inference and decision

Glossary

Subjective probability: the probability of an event is interpreted as the expression of a degree of belief in that particular event. Subjective probabilities are personal and conditional on the individual's experience and knowledge. As all probabilities, subjective probabilities take values in the range between 0 and 1.

Frequentist probability: The frequentist probability is defined as the limit of the relative frequency of a target event that has occurred in a large number of trials if it is conceivable that the same experiment may be repeated under identical conditions a very large number of times.

Utility function: the utility measures on some numerical scale the desirability of decision consequences $C_{ij} = c(d_i, \theta_j)$ that take place when a decision d_i is taken and θ_j turns out to be the true state of nature. Whenever a '0-1' scale is used, a value equal to 0 and 1 is assigned to the least and to the most desirable consequences, respectively. All the remaining consequences are assigned a value within this interval with the sole constraint of coherence: if a consequence is more desirable than another, it must have a greater utility, and vice versa.

Loss function: the loss measures on some numerical scale the undesirability of decision consequences $C_{ij} = c(d_i, \theta_j)$ that take place when a decision d_i is taken and θ_j turns out to be the true state of nature. Whenever a '0-1' scale is used, a value equal to 1 and 0 is assigned to the least and to the most desirable consequences, respectively. All the remaining consequences are assigned a value within this interval with the sole constraint of coherence: if a consequence is less desirable than another, it must have a greater loss, and vice versa. Note that that the loss function is obtained as the difference between the utility of the best consequence under the state of nature at hand and the utility for the consequence of interest. Stated otherwise, the loss measures the penalty for choosing a non-optimal decision,

Maximizing expected utility: the desirability of alternative decisions is measured by their corresponding expected utility which is obtained by combining utilities $u(C_{ij})$ associated with the consequences of decisions C_{ij} and probabilities for states of nature $Pr(\theta_j)$ as $\bar{u}(d_i) = \sum_{j=1}^{n} u\left(C_{ij}\right) \times Pr(\theta_j)$, $i = 1, \ldots, m$. A standard decision rule instructs one to select the action which maximizes the expected utility.

Minimizing expected loss: the undesirability of alternative decisions is measured by their corresponding expected loss which is obtained by combining losses $l(C_{ij})$ associated with the consequences of decisions C_{ij} and probabilities for states of nature $Pr(\theta_j)$ as $\bar{l}(d_i) = \sum_{j=1}^{n} l\left(C_{ij}\right) \times Pr\left(\theta_j\right), i = 1, \ldots, m$. A standard decision rule instructs one to select the action which minimizes the expected loss.

1. Introduction

Nowadays age estimation of living persons is a recognized discipline in the forensic panorama. Age is a fundamental piece of information in our society for the exercise of personal rights and duties. Thus, faced with persons unable or unwilling to declare their age, judicial or administrative authorities often request an expert opinion. Such request have been on the rise recently, since the number of individuals of questioned age has increased, due to the tremendous tide of migration movements, ease of world travel, but also due to the professionalization of criminal organizations involved in human smuggling or trafficking (Law et al., 2010). As highlighted by Schmeling et al. (2007), a question to be answered in case of age diagnostics for living persons mostly concerns the probability of a person being younger or older than a legally relevant age threshold, such as the age of majority. In other cases, a point estimate of the real age can be of interest. In any case, the two pieces of information are strictly related, since the former logically depends on the latter.

Although European Asylum Support Office (EASO, 2018) recommends a holistic perspective for age estimation in living individuals (including psychological assessments), we believe that scientific evidence-based anthropological/medico-legal methods should provide the basis of an age estimate. The Study Group on Forensic Age Diagnostic (AGFAD) has provided, particularly in Europe, a strong contribution to this field, by publishing a set of recommendations mainly focused on operational perspectives (Schmeling et al., 2008). These recommendations include the choice of reference studies, the examination steps that ideally ought to be performed during a medico-legal age estimation appraisal, and the structure of expert reports. Similar recommendations were published by the Forensic Anthropology Society of Europe (Cunha et al., 2009).

An evaluation of multiple items of evidence is highly recommended in order to increase the accuracy of the age estimate (Schmeling et al., 2003, Schmeling et al., 2008, Bassed, 2012). However, in the early stages of applied forensic age estimation , the domain suffered from a lack of adequate (statistical) methods that would allow to comprehensively evaluate the age-related evidence (Ritz-Timme et al., 2000). The aim of this contribution is to illustrate that forensic age estimation is a problem of inference and decision, and should not be considered only from a statistical perspective. Provided that uncertainty is unavoidable and should be measured by probability, the Bayesian paradigm represents a formidable tool to combine different sources of information that are at the disposal to the different actors involved in the legal disputes regarding a person's age.

2. Uncertainty and inference in forensic age estimation

In forensic age estimation the forensic experts needs to translate their scientific findings into legal information by evaluating and interpreting the evidence (Sironi et al., 2017). Typically, a judicial authority needs to decide whether a person is an adult (e.g. 18 years old or more) or a minor, within the meaning of the law. In order to make a decision, the authority must obtain information about the chronological age of the individual, i.e., the quantity that measures – in years, months and days – the time since the person was born (EASO, 2018). However, age-related and developmental evidence collected by the forensic expert consists of data on the individual's biological age, i.e., the developmental step reached by the individual reflected by achieving a given specific physical attribute at the time of the examination (Hackman et al., 2010). This biological (scientific) information has to be correctly used in order (1) to infer the chronological age, i.e., the information needed by the mandating authority, and (2) for the authority to make a decision about the matter (i.e., in most cases whether the person in question shall be declared a minor or an adult).

Although physical development is a continuous biological process, it is generally described in categorical steps due to the extreme difficulty in appreciating globally all changes which occur during aging in a continuous scale (Lucy et al., 2002). A body of age-related evidence consists therefore in the detection (by an expert) of the developmental steps reached by specific physical attributes (or age indicators) in an examined individual constituting the biological age. It is recognized that the developmental steps are universally reached in their totality and in the same sequence, but the chronology of the developmental phases varies considerably between individuals (Boldsen et al., 2002, Cameron and Jones, 2010). This is mainly due to the influence of a large panoply of individual and environmental factors (i.e., social context or ethnic origin), which may affect the developmental process (Kemkes-Grottenthaler, 2002, Cameron and Jones, 2010) with varying extent for each biological system, such as the skeletal or dental systems (Schmeling et al., 2005). Interpretation of age-related evidence must be provided by assessing the uncertainty regarding the relationship between biological age (i.e. the scientific information) and chronological age (i.e. the legal information). In the presence of available data collected from subjects whose age is known, the uncertainty about such relationships could be modeled by means of appropriate statistical models. However, it must be added that an *ad-hoc* statistical model is not the end of the matter, as highlighted by Taroni and Biedermann (2014, p. 3948) recalling a statement of I. W. Evett:

> "[…] Statistics concentrates primarily on data, whereas the retrospective meaning of an observation relies on the more general concept of inference which focuses on the notion of uncertainty. […]"

In forensic science, the decision-maker seeks to evaluate a hypothesis or a feature of interest in the light of scientific findings (Taroni and Biedermann, 2015). This is typically an inductive line of reasoning; it is said to be ampliative, since the conveyed conclusions contain elements that are not present in the premises. One's knowledge is extended by inference throughout data (findings, observations) acquisition. Such amplification naturally implies uncertainty: the key elements that allows one to move from an initial belief about the feature of interest (i.e. the age of a given individual) to an updated belief is data acquisition (so often called 'evidence'), which is generally incomplete, imprecise and rarely conclusive (Schum, 1994). This is also the case in forensic age estimation: the age-related evidence concerning the biological age is incomplete by nature, since it informs us about a given moment in time of a continuous evolutionary process. Moreover, the conclusions in the legal process refer to the chronological age. However, the relationship between biological and chronological age is affected by many sources of uncertainty. It is therefore fundamental to handle such uncertainty in a logical way, in order to provide meaningful findings to the mandating authority.

Uncertainty can be qualified and quantified by means of probability. In fact, probability is the standard measure for uncertainty (Lindley, 1991). Probability is defined as the measure of one's degree of belief on a given event or statement (Taroni et al., 2001). Such an interpretation is generally referred to as subjective (in the meaning of personal, and not arbitrary) and plays an important role in the forensic context (Berger et al., 2011, Biedermann, 2015, Taroni et al., 2015, Biedermann et al., 2017b). An ideal solution to handle the uncertainty in inductive reasoning is provided by the Bayesian paradigm (Taroni and Biedermann, 2014), that formalizes the general approach for thinking about evidence (Robertson and Vignaux, 1998). The usefulness of Bayesian perspective for forensic evidence interpretation has been recognized (Aitken and Taroni, 2004, Comitee of forensic experts, 2011, ENFSI, 2015, Robertson et al., 2016) to the point that Evett (2015, p. 10) states that "The nature of forensic science is now firmly founded in the Bayesian paradigm […]".

3. Bayesian perspective in forensic age estimation

The Bayesian paradigm states that all uncertainties characterizing an issue of interest must be described by means of probabilities or probability distributions. Probabilities are interpreted as a conditional measure of uncertainty associated with the unknown feature of interest (e.g., the chronological age) given the available information. The learning process about the feature of interest is described as the modification of the uncertainty in the light of new information, the scientific findings; the Bayes theorem explains how this should be done, formalizing the common notion 'learning from experience' (Jeffreys, 1961). The algebraic expression of the Bayes theorem depends on the nature of the involved variables. Suppose we have a set $\boldsymbol{E}$ of categorical age-related evidence and that we consider the chronological age to be the realization of a continuous variable A. A standard application of the Bayes' theorem allows one to obtain the posterior distribution of the chronological age as

$$f_A(a|\boldsymbol{E},I) = \frac{Pr(\boldsymbol{E}|a,I) \times f_A(a|I)}{\int Pr(\boldsymbol{E}|a,I) \times f_A(a|I)da}, \quad (1)$$

where I indicates the background information related to the examined person, such as the sex and the ethnic origin and any other relevant pieces of information. A discussion on the meaning and the role of background information can be found in Aitken and Nordgaard (2018). The formulae state that the updated belief, i.e., the posterior probability distribution on the age $f_A(a|\boldsymbol{E},I)$ results from a normalized combination of the initial beliefs about the age, represented by the prior probability distribution $f_A(a|I)$ and the information originating from the available evidence, quantified as $Pr(\boldsymbol{E}|a,I)$.

Note that analytically solving Bayesian models can be a tedious and time-consuming procedure, which is unsuitable for daily practice. Nonetheless, specific statistical tools exist in order to simplify this aspect, especially so-called Bayesian

Networks (BNs). BNs are probabilistic graphical models that present the dual advantage of graphically describing the relationship between variables describing the inferential model, as well as of directly providing automatic computations following the rule of the probability theory (Taroni and Biedermann, 2013). BNs are widely used in forensic science (Taroni et al., 2014), including age estimation (Sironi et al., 2016). They can be easily extended to take into account the decisional aspect (Taroni et al., 2014).

4.1 Posterior probability distribution on the chronological age

Posterior density is the target outcome of the inductive process: it encapsulates all available information related to the specific case, ranging from the collected evidence, the available background information up to the specific expert knowledge at a given time. It is therefore possible to quantify and combine all sources of uncertainty about the target quantity (i.e., the chronological age) in a rational way, as clearly requested in the age estimation domain (EASO, 2018, Malmqvist et al., 2018).

The posterior distribution can provide point and interval estimates of the chronological age. Furthermore, it can be used to inform us about the probability of competing propositions such as whether the examined individual is older or younger than a specific threshold. For instance, the probability that an individual is 18 years or older can be obtained by the integration of the *posterior* density function $f_A(a|\boldsymbol{E}, I)$ over the age-space of interest (Thevissen et al., 2010, Sironi et al., 2016):

$$Pr(\theta_1|\boldsymbol{E}, I) = \int_{\theta_1} f_A(a|\boldsymbol{E}, I)\, da = \int_{\theta_1} \frac{Pr(\boldsymbol{E}|a, I) \times f_A(a|I)}{\int Pr(\boldsymbol{E}|a, I) \times f_A(a|I) da}\, da \qquad (2)$$

where θ_1 is the interval that covers the age space equal to and greater than 18 years of age. The (posterior) probability that the individual is younger than 18 years, θ_2, can be computed analogously.

4.2. Prior probability distribution on the chronological age

The prior probability distribution should reflect prior beliefs on the chronological age of the examined individual before the evaluation of the collected age-related evidence. Though strategies suitable for formalizing prior beliefs in terms of a probability distribution have been proposed both in statistical and forensic literature (O'Hagan et al., 2006, Taroni et al., 2010, Sironi et al., 2017, Bolstad and Curran, 2017), the elicitation of a prior in age estimation can be a challenging task, since there is generally little initial information at the expert's disposal (Schmeling et al., 2003). From a Bayesian perspective this must not be felt to be an insurmountable drawback, as the prior probability distribution "[…] reflects [one's] belief about the subject matter, conditioned as these will presumably be by [one's] available background evidence […]" (Howson, 2002, p. 53). The expert should be able to qualify or quantify their personal belief on the age of the examined person based on preliminary information on the case at hand.

In the framework suggested by the AGFAD, the prior probability distribution could be assigned by the expert based on information collected during an interview or the physical examination of the person under investigation, since it is generally acknowledged that data gathered in this initial step of the examination sequence should not be used for the effective age estimation (Schmeling et al., 2006a, Schmeling et al., 2011). Due to the lack of available information, many authors have proposed adopting so-called non-informative or vague prior probability distributions, such as a uniform distribution over a given age range, according to which all possible age values are considered, *a priori*, equally likely (Braga et al., 2005, Thevissen et al., 2010, Cameriere et al., 2016, Bleka et al., 2018). This way of thinking should result in posterior conclusions that will be minimally dependent on prior distribution. This is not intrinsically wrong, but information-less priors may be misleading as such priors actually do not exist (Howson, 2002). The introduction of a uniform distribution over a given age range is far from being without information, as it conveys the belief that all ages in the chosen interval are considered equally likely, whilst those outside the range are considered as not possible (Sironi et al., 2017). The elicitation of such a uniform distribution may be reasonable in some scenarios (Sironi et al., 2018b), when the age interval is chosen based on the available knowledge (Sironi et al., 2018b, Konigsberg et al., 2019). Other choices of probability distributions over the chronological age are clearly possible. For instance, in case of forensic age estimation of a young adult, a probability distribution centered around the legal threshold may be preferred. Asymmetrical distribution has been proposed in the scientific literature (Sironi et al., 2016), nonetheless Konigsberg et al. (2019) argued that it would be more beneficial to rather assign a symmetrical one, as the normal or the Laplace distribution, located around the mentioned legal threshold, which leads to an equal support of the competing intervals under consideration (here named θ_1 – a given person is aged 18 years or older; and θ_2 – a given person is younger than 18 years) (Konigsberg et al., 2019).

Examiner may be concerned about the sensitivity of the posterior distribution to alternative prior distributions that fit just as well the prior beliefs. The sensitivity analysis is a powerful tool for investigating the robustness of the posterior inference on prior assignments (Sironi and Taroni, 2015, Sironi et al., 2015, Sironi et al., 2018a, Sironi et al., 2018b, Konigsberg et al., 2019).

4.3. Likelihood function

The likelihood function models the relationship between the age-related evidence (i.e., the biological age) and the chronological age. The choice of standard statistical models (such as classical regression models) is generally unfeasible. It must be acknowledged that in the age estimation scenario, there are generally multiple items of evidence available to authorities, and that the quantified variables do not necessarily have identical scales of measurement. A statistical model must clearly provide a coherent assessment of the uncertainty about the relationship between the biological and the chronological age, but must also be capable of dealing with multiple items of evidence.

Several statistical methods have been proposed in the literature for age estimation purposes, although, they generally present some drawbacks, such as the unsuitability of being employed for the evaluation of multiple items of evidence having different scales of measurement (Thevissen et al., 2010, Hillewig et al., 2013, Sironi et al., 2018a) or operational limitations (Braga et al., 2005).

The assumption of conditional independence between pieces of evidence given the age has often been retained (Corradi et al., 2013a, Corradi et al., 2013b, Gelbrich et al., 2015, Fieuws et al., 2015, Tangmose et al., 2015, Bleka et al., 2018). If this assumption were reliable, the likelihood function could be obtained as the product of individual likelihoods for each considered item of evidence, when n items of evidence are evaluated:

$$Pr(\boldsymbol{E}|a,I) = \prod_{k=1}^{n} Pr(E_k|a,I). \tag{3}$$

However, this assumption does not seem to meet the biological reality, since the development of a single part of the body is rarely independent from the others (Boldsen et al., 2002). The consequence may be that the estimated posterior density is too narrow compared to what it should be (Fieuws et al., 2015). In this perspective, Boldsen et al. (2002) have suggested a statistical procedure for correcting the posterior interval estimates, and Fieuws et al. (2015) have extended the procedure in order to allow one to correct directly the posterior density. The assumption of conditional independence may be strengthened by considering other pieces of information provided by the examined person. Examples of such information may be the knowledge of diseases or lower socio-economic status of the examined person during development, which may affect the skeletal and dental development to a different degree (Schmeling et al., 2005, Schmeling et al., 2006b).

When considering a single item of evidence, the use of regression models specifically developed for the treatment of categorical dependent variables is beneficial. Notably, models from both *probit* and *logit* families have been used (Konigsberg (2015) and references therein). Specifically, *unrestricted cumulative* models (Sironi and Taroni, 2015, Sironi et al., 2015, Sironi et al., 2018a, Konigsberg et al., 2019) and *continuation ratio* models (Fieuws et al., 2015, Tangmose et al., 2015, Sironi and Taroni, 2015, Sironi et al., 2016, Bleka et al., 2018) have been proposed. Note that the models employed by Sironi and Taroni (2015), Sironi et al. (2016) and Bleka et al. (2018) are sometimes also referred to as *stopping ratio models* (Konigsberg et al., 2019). *Proportional-odds* (or *restricted cumulative*) models have also been employed (Bleka et al., 2018). These models are based on assumptions that do not meet the biological reality of the developmental process, thus they are not appropriate in this field (Boldsen et al., 2002). As pointed out by Konigsberg et al. (2019), there is little practical difference between the *logit* or *probit* models: traditionally *logit* models were preferred because of computational ease compared to the *probit*, but nowadays this is no longer a rational argument, considering the availability of statistical software (Konigsberg, 2015). Moreover, it may be felt that models from the *probit* family may be easier to implement for practitioners, since they refer to normal distribution rather than to logistic distribution, which may be less intuitive (Myers et al., 2002). According to Konigsberg et al. (2019), this aspect is particularly relevant in the forensic field, since scientific results originating from the selected models should be presented to an audience with a limited scientific background. Furthermore, Konigsberg (2015) highlighted that univariate *probit* models can be extended to consider multiple variables by using a multivariate normal distribution, whilst this task would be more complex with logistic distribution.

An attractive feature of the cumulative models is that some consecutive stages can be collapsed into a single one without affecting the estimation of the parameters of the curves of the other stages (Konigsberg et al., 2008). This is particularly interesting in case of age estimation from physical attributes for which development is described by means of several categorical steps. However, these models may generate overlapping regression curves that may lead to inconsistencies associated with a generic developmental stage (Konigsberg and Herrmann, 2002). Sironi and Taroni (2015) showed that different regression models may lead to different quantifications of the posterior probabilities on a given age cohort (such as θ_1 or θ_2). Bleka et al. (2018) adopted the regression model that provided the best fit with the available data. Note that in order to avoid inconsistencies on the value of age, the variable 'age' can be transformed into a logarithmic (Konigsberg et al., 2008) or exponential scale (Bleka et al., 2018). Non-parametric models (such as those based on the Kernel distribution) may also be employed (Lucy et al., 2002).

The main drawback affecting the above models, is that for their operational implementation, there is an urgent need for structured data samples, that unfortunately are unlikely to be available (Konigsberg, 2015). Nonetheless, the lack of adequate data should not be considered as an insurmountable impasse. For instance, the guidelines for evaluative reporting in forensic science published by the European Network of Forensic Science Institutes (ENFSI, 2015) states that:

> "[…] Relevant and appropriate published data will be used wherever possible. If appropriate published data are not available then data from unpublished sources may be used. Regardless of the existence of sources (published or not) of numerical data, personal data such as experience in similar cases and peer consultations may be used, provided that the forensic practitioner can justify the use of such data. […]" (ENFSI, 2015, p. 15).

Furthermore,

> "[…] likelihood […] can be informed by subjective probabilities using expert knowledge. […] Such personal probability assignment is not arbitrary or speculative, but is based on a body of knowledge that should be available for auditing and disclosure." (ENFSI, 2015, p. 16)

Note that the guidelines focus on the assignment of the probabilities that represent the ingredients of the likelihood ratio (i.e., the ratio between the probability of the evidence given two competing propositions), but these statements can be extended in order to choose an appropriate statistical model that is capable of taking into account in a coherent way the uncertain relationship between biological and chronological age. An extensive discussion about the role of the subjectivist approach to the elicitation of probabilities in forensic science can be found in Taroni et al. (2001), Berger et al. (2011), Biedermann et al. (2017b), Taroni et al. (2018).

5. Bayesian inference from an operational perspective

The regression models discussed above represent the ideal approach to forensic age estimation. However, because of the lack of reference data, their application in practice may be problematic. In case of lacking data (see above), forensic practitioners may feel more comfortable in assigning probabilities in the form of relative frequencies rather than probability distributions. The use of relative frequencies to inform subjective probabilities has recently been discussed by Taroni et al. (2018). However, such frequencies cannot logically be assigned for each age of a continuous scale. Since the mandating authority usually asks whether or not an age threshold was exceeded, it may be sensible to consider two alternative propositions in the form:

- θ_1: the examined person is aged 18 years or older;
- θ_2: the examined person is younger than 18 years of age;

Note that in this paper the notation θ_j is used in two different meanings. In Section 4, θ_1 and θ_2 represent a subset of the support of the continuous variable "age", whilst in the current one it represents a discrete event. Note also that the age limits considered as reasonable in the case at hand should be expressed directly in the propositions, whilst in the framework discussed in section 4, these limits are implicitly defined by the choice of the probability distributions of interest. The Bayes' theorem can therefore be formulated as follows:

$$Pr(\theta_j|\boldsymbol{E},I) = \frac{Pr(\boldsymbol{E}|\theta_j,I) \times Pr(\theta_j|I)}{\sum_{j=1}^{2} Pr(\boldsymbol{E}|\theta_j,I) \times Pr(\theta_j|I)} \tag{4}$$

It is worth emphasizing that not only the evidence, but also the chronological age is categorized as a discrete variable. This formulation is rarely addressed in the literature (Lucy, 2010). However, from an operational perspective, the task of the forensic expert is potentially simplified, since it will be limited to the assessment of a probability of discrete events (the exceeding or vice versa of an age threshold). These probabilities can be assigned relying on available relative frequencies of the developmental evidence in the given cohort. Such relative frequencies can sometimes be extrapolated from reference studies, or from unpublished data (ENFSI, 2015).

It must be clarified that this does not amount to equating a conditional probability that according to our view represents a degree of belief, with relative frequency that represents a normalized count of a given quantity. From a frequentist point of view, the probability can indeed be defined as a limiting value of relative frequency, assuming a large repetition of the event under identical conditions is feasible. This is generally not the case for forensic evidence (Lucy, 2010, Curran, 2013), and a subjectivist approach is strongly encouraged (Lindley, 1991). Nevertheless, a personal degree of belief can be informed by relative frequencies, which is even recommended, when data are available (Taroni et al., 2018).

Notably, while the expert needs to assign the probability of the evidence $\boldsymbol{E}$ based on the observation of multiple age indicators, data collected simultaneously from multiple age indicators are infrequently available in forensic literature (Schmeling et al., 2016). However, if the assumption of conditional independence between different pieces of evidence is feasible, the conditional probability $Pr\ (\boldsymbol{E}|\theta_j, I)$ can be simplified as follows:

$$Pr(\boldsymbol{E}|\theta_j, I) = \prod_{k=1}^{n} Pr(E_k|\theta_j, I) \tag{5}$$

Then, conditional probabilities for each type of evidence E_k, $Pr(E_k|\theta_j, I)$, can be derived from relative frequencies accessed from available databases or published reference studies.

Two aspects of the use of relative frequencies to elicit conditional probabilities in Eq. (5) need to be considered. Firstly, the frequency of a specific complex developmental pattern is very low, since the number of observable patterns can be very large. Thus, the posterior inference may be highly sensitive to such assignment. Secondly, the relative frequencies are logically influenced by the structure of the reference sample.

Several studies report the probability of being at least 18 years old given the observed developmental stage (Liversidge and Marsden, 2010, Mincer et al., 1993). However, such probabilities refer to the proposition given the evidence (i.e., $Pr(\theta_1|\boldsymbol{E}, I)$), and not to the evidence given the proposition (i.e., $Pr(\boldsymbol{E}|\theta_1, I)$). Equating these two probabilities would amount to a transposition of conditionals (Evett, 1995). Note that the probability $Pr(\theta_1|\boldsymbol{E}, I)$, whenever available, does not incorporate the prior knowledge about age.

6. Normative approach to decision in age estimation

The problem of evidence interpretation concerns both inference and decision. The importance of a rational approach to decision-making for questions by the different actors of the legal process has gained an increasing attention in forensic literature (Taroni et al., 2005, Taroni et al., 2010, Gittelson, 2013, Biedermann et al., 2016, Biedermann et al., 2018). In age estimation, the expert is asked to make several decisions, including the choice of an appropriate method for assessing the developmental process, or the identification of the developmental status reached by a physical attribute. Nonetheless, the more relevant decision in age estimation concerns indubitably the fact that a person is adult or minor.

Note that a clear distinction is to be made between deciding that an individual being evaluated is younger or older than 18 years of age, and the decision to declare him or her minor or adult within the meaning of the law. Such decision in age estimation will be taken, at a given moment, by one of the participants of the legal or administrative procedure (Biedermann et al., 2017a). The decision maker may be either the scientist or the mandating authority, depending on the framework of the circumstances of a specific case. However, the main focus of this paper is not a discussion about *who* is entitled to make this kind of decision in forensic age estimation framework, but rather *how* such a decision problem should be tackled (see Biedermann et al. (2008) for a wide discussion on this aspect in forensic science). Therefore, the current discussion focuses on the general aspects of the normative approach to making decision in forensic age estimation. From a general point of view, this decision problem can be resumed on evaluating "what is the minimum degree of probability $Pr(\theta_1|\boldsymbol{E},I)$ to be required for accepting that the adulthood of the examined person is established?" (Biedermann et al., 2017a). A study conducted by Polo Grillo et al. (2002) on 47 judicial cases involving immigrants lacking valid ID documents illustrated how decision-makers expressed appreciation for the quantification of uncertainty in terms of probability in age estimation cases. The study pointed out that the judges felt "confident" in declaring an individual to be an adult if the reported probability that he or she is 18 years or older was greater than 0.70. However, the empirical nature of this study must be emphasized, because it is generated by the observation of the behavior of the decision-maker not using a specific decisional criterion. For this reason, it cannot serve as a framework to support a rational decision in specific casework. For this reason, we endorse a normative approach to decision-making, which can be integrated into the Bayesian framework. The normative approach to decision-making has only recently been explored in forensic sciences (Taroni et al., 2005, Biedermann et al., 2008, Gittelson et al., 2013, Gittelson et al., 2014, Biedermann et al., 2016, Gittelson et al., 2016, Biedermann et al., 2017a). More details about the decision theory and its application can be found, among others, in DeGroot (1970), Lindley (2006) and Berger (2010).

A problem of decision can be described in terms of three principal components (Table 1):

- A collection of states of nature, denoted $\theta_1, \theta_2,\ldots, \theta_n$, that represent the events of interest in the decision-making process and about which the decision-maker is uncertain. Assuming that probability is the standard measure of uncertainty, the uncertainty about these events can be quantified in a collection of probabilities $Pr(\theta_1|\cdot)$, $Pr(\theta_2|\cdot)$, …, $Pr(\theta_n|\cdot)$ which are conditioned by all available knowledge in the given as $\sum_{j=1}^{n} Pr(\theta_j|\cdot) = 1$. In age estimation, the events of interest can be stated as the two competing propositions θ_1 and θ_2. To each state of nature a probability will be assigned, namely $Pr(\theta_1|\boldsymbol{E},I)$ and $Pr(\theta_2|\boldsymbol{E},I)$ as in Eqs 2 and 4 (see section 4 and 5).
- A collection of decisions (or actions), denoted d_1, d_2,…, d_m. Decisions must be exhaustive and mutually exclusive, that is, the collection must cover all possible decisions and the decision-maker can choose one, and only one, decision among all the available ones. In the scenario discussed in this Chapter, the available decisions are d_1 : to declare the examined person as adult, and d_2 : to declare the examined person as minor.
- A collection of consequences that result from the combination of the states of nature and the available decisions: the choice of a decision d_i when θ_j is the true state of nature leads to a consequence $C(d_i;\theta_j)$, denoted as C_{ij}.

Table 1
Decision matrix for the given age estimation scenario. d_i with $i = 1{,}2$ denotes the available decisions, θ_j with $j = 1{,}2$ denotes the events of interests, and C_{ij} denotes the possible decision consequences.

Decision	States of nature	
	θ_1 : aged 18 years or older	θ_2 : younger than 18 years of age

d_1 : adult	C_{11} : correct declaration as adult	C_{12} : false declaration as adult
d_2 : minor	C_{21} : false declaration as minor	C_{22} : correct declaration as minor

Decision consequences are characterized by a different level of desirability or undesirability. In the age estimation scenario, the more favorable consequence is when the examined person is correctly declared as being adult or minor (C_{11} and C_{22}), whilst the worst consequence is when a person who is actually younger than 18 years of age, is declared an adult (C_{12}). This type of error is generally referred to as ethically unacceptable (Garamendi et al., 2005). The consequence of declaring a person who actually is older than 18 years of age (C_{21}) to be a minor is generally favorable to the individual, but adverse to the society, because of the additional social expenses that could have been avoided. For this reason, the undesirability of consequence C_{21} can be considered to be intermediary between those referred to the best and the worst consequence. This amounts to the following preference ordering:

$$C_{11} \sim C_{22} \succ C_{21} \succ C_{12}$$

where $\succ$ denotes the preference of the decision-maker for one consequence over another, whilst $\sim$ denotes the indifference on the desirability of two or more consequences.

The desirability of a given consequence C_{ij} can be measured by a numerical value, by means of a function called *utility function*, denoted as $u(d_i;\theta_j) = u(C_{ij})$. Analogously, it is possible to express preferences among decision consequences by means of a *loss function*, denoted as $l(d_i;\theta_j) = l(C_{ij})$, that quantifies on a numerical scale the undesirability of decision consequences. Note that utilities and losses are conceptually and mathematically connected (Berger, 2010), though in the current scenario the more intuitive reasoning would be in terms of losses.

Various strategies can be implemented to build an appropriate loss function (Lindley, 1985, Koller and Friedman, 2009, Berger, 2010), with the sole constraint being that the loss function must correctly reflect the preference ordering. This implies that if $l(\cdot)$ is a loss function and one consequence is felt as less desirable than another one, i.e., $C_{12} \prec C_{22}$, then $l(C_{21}) < l(C_{12})$. The highest loss value, according to the unit scale, will be assigned to the more adverse outcome (i.e., C_{12}), while the smallest will be assigned to the less adverse or more favorable outcome (i.e., C_{11} and C_{22}).

As far as the practical and controversial issue of building the loss function is concerned, both the utility and the loss functions are invariant to linear transformations, so that any particular choice of the unit scale is theoretically allowed (Berger, 2010). A convenient choice is the '0-1' scale, where the minimum and the maximum of the loss scale are fixed at 0 and 1, respectively (Table 2). This implies that a loss equal to 0 is assigned to C_{11} and C_{22} (the most favored outcomes), that is $l(C_{21}) = l(C_{12}) = 0$, and a loss equal to 1 is assigned to C_{12} (the worst outcome), that is $l(C_{12}) = 1$. In this way the construction of a loss function for the age estimation scenario is greatly simplified as the only remaining value that needs to be assigned is the loss associated with the intermediate consequence C_{21}.

A feasible strategy to assign the loss associated with C_{21} refers to a context where one player (i.e., the decision-maker) is asked to choose between a sure event (in this case, the outcome corresponding to the intermediate consequence whose loss needs to be assigned), and a gamble where the worst outcome takes place with probability p and the best outcome takes place with probability $1-p$ (see Lindley (1985) or (Berger, 2010) for a formal approach). The loss of the intermediate consequence C_{21} can be set to be equal to the probability p that makes one indifferent between the sure event and the gamble. Note that this assignment has to be made under the constraint imposed by the personal scale of desirability of the possible consequence: if $C_{11} \sim C_{22} \succ C_{21} \succ C_{12}$, then $l(C_{11}) = l(C_{22}) > l(C_{21}) > l(C_{12})$, thus $l(C_{21})$ has to take a value between 0 and 1 (excluded).

Table 2
The '0-1' loss function for the age estimation scenario, where d_1 and d_2 denote the available decisions, θ_1 and θ_2 denote the events of interest (as illustrated in Table 1), and $l(C_{21})$ denotes the loss associated with the intermediate consequence C_{21}.

	θ_1 : aged 18 years or older	θ_2 : younger than 18 years of age
d_1 : adult	0	1
d_2 : minor	$l(C_{21})$	0

The undesirability of available decisions can be measured by their corresponding (posterior) expected losses:

$$\bar{l}(d_i|\cdot) = \sum_{j=1}^{n} l\left(C_{ij}\right) \times Pr\left(\theta_j|\boldsymbol{E}, I\right) \tag{6}$$

where $Pr\left(\theta_j|\boldsymbol{E}, I\right)$ are the posterior probabilities of the states of nature and $l(C_{ij})$ the losses for the consequences of interest. The optimal decision, also called Bayes decision, is the decision that minimizes the Bayesian expected losses, formally:

$$\underset{i}{arg\,min}\ \bar{l}(d_i|\cdot) = \underset{i}{arg\,min} \sum_{j=1}^{n} l\left(C_{ij}\right) \times Pr\left(\theta_j|\cdot\right) \tag{7}$$

In this way both the undesirability of the consequences (in terms of losses) and the uncertainty of the state of nature (in terms of probabilities) are considered related to the decision.

It is thus possible to quantify the expected loss for the two available decisions in the age estimation scenario. The (posterior) expected loss for the decision d_1 (i.e., to formally declare the person an adult) can be quantified as

$$\bar{l}(d_1|\boldsymbol{E}, I) = \sum_{j=1}^{2} l\left(C_{1j}\right) \times Pr\left(\theta_j|\boldsymbol{E}, I\right) = \underbrace{l(C_{11})}_{0} \times Pr(\theta_1|\boldsymbol{E}, I) + \underbrace{l(C_{12})}_{1} \times Pr(\theta_2|\boldsymbol{E}, I), \tag{8}$$

while the (posterior) expected loss of decision d_2 (i.e., to formally declare the person a minor) can be quantified analogously as

$$\bar{l}(d_2|\boldsymbol{E},I) = \sum_{j=1}^{2} l(C_{2j}) \times Pr(\theta_j|\boldsymbol{E},I) = l(C_{21}) \times Pr(\theta_1|\boldsymbol{E},I) + \underbrace{l(C_{22})}_{0} \times Pr(\theta_2|\boldsymbol{E},I), \tag{9}$$

According to the Bayesian decision criterion in Eq. (7), that prescribes making the decision having the lower expected loss, the decision d_1 is therefore preferable to the decision d_2 when $\bar{l}(d_1|\boldsymbol{E},I) < \bar{l}(d_2|\boldsymbol{E},I)$. Writing expected losses $\bar{l}(d_i|\boldsymbol{E},I)$ in full length, and eliminating the terms in Eqs. (8) and (9) involving zero losses, leads to the following expression

$$l(C_{12}) \times Pr(\theta_2|\boldsymbol{E},I) < l(C_{21}) \times Pr(\theta_1|\boldsymbol{E},I), \tag{10}$$

or equivalently to

$$\frac{Pr(\theta_2|\boldsymbol{E},I)}{Pr(\theta_1|\boldsymbol{E},I)} < \frac{l(C_{21})}{l(C_{12})}. \tag{11}$$

Eq. (11) states that the decision d_1 to declare the examined individual an adult is preferable if, and only if, the posterior odds in favor of θ_2 are smaller than the loss associated to C_{21}, which corresponds to wrongly declaring an individual who has effectively exceeded the age of 18 years as a minor. Note that $l(C_{12}) = 1$ and $Pr(\theta_2|\boldsymbol{E},I) = 1 - Pr(\theta_1|\boldsymbol{E},I)$, thus, the Eq. (11) can be rearranged as follow:

$$Pr(\theta_1|\boldsymbol{E},I) > \frac{1}{1 + l(C_{21})}. \tag{12}$$

That is, the minimum degree of probability $Pr(\theta_1|\boldsymbol{E},I)$ that is required to decide about the adulthood of the examined person is given by the ratio $1/[1 + l(C_{21})]$. A more intuitive way to interpret the principle of minimizing the expected loss in the current scenario is provided in Figure 1, that shows the (posterior) expected losses in Eqs. (8) and (9) as a function of the posterior probability of θ_1, $Pr(\theta_1|\boldsymbol{E},I)$, with $l(C_{21}) = 0.50$ (Figure 1a) and $l(C_{21}) = 0.10$ (Figure 1b).

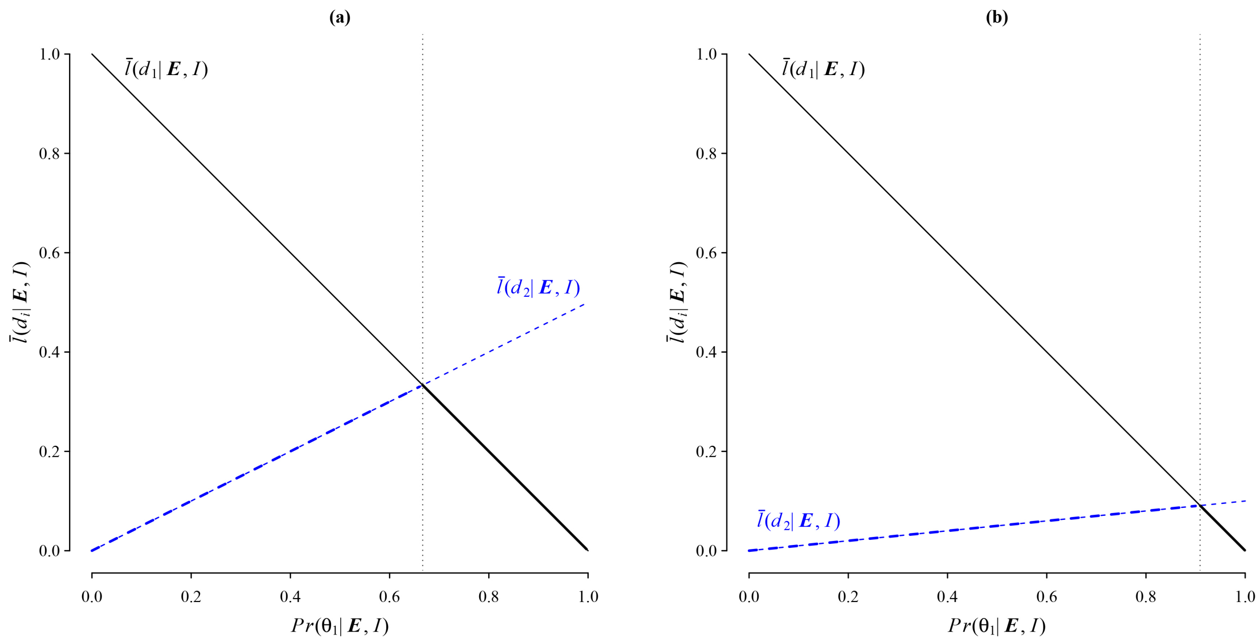

Fig. 1: Expected losses $\boldsymbol{\bar{l}(d_i|E,I)}$, with $\boldsymbol{i = 1, 2}$, for different values of $\boldsymbol{Pr(\theta_1|E,I)}$ with $\boldsymbol{l(C_{21}) = 0.50}$ (a) and $\boldsymbol{l(C_{21}) = 0.10}$ (b). The dotted vertical line indicates the threshold value of $\boldsymbol{Pr(\theta_1|E,I)}$ that inverses the preferability of the decision, in this case $\boldsymbol{Pr(\theta_1|E,I) = 0.\overline{66}}$ (a) and $\boldsymbol{Pr(\theta_1|E,I) = 0.\overline{90}}$

In Figure 1, the optimal decision is the one for which the values of the corresponding expected losses attain the minimum: these values are highlighted in bold. Given $l(C_{21}) = 0.50$, Figure 1a shows that the optimal choice is d_1 if, and only if, $Pr(\theta_1|\boldsymbol{E}, I) > 1/[1 + 0.50] = 0.\overline{66}$, otherwise it is d_2. In the second case (Figure 1b), given $l(C_{21}) = 0.01$, the optimal choice is is d_1 if, and only if, $Pr(\theta_1|\boldsymbol{E}, I) > 1/[1 + 0.01] = 0.\overline{90}$

Let us now consider the problem of choosing a loss function. In the current scenario, the problem is confined to the choice of a meaningful value for $l(C_{21})$. Note that, as pointed out by Biedermann et al. (2016, p. 34), the necessary comparison implied by Eq. (11) " […] is essentially qualitative and reduces to a single factor, call it x for simplicity, that states how much greater one loss value is compared to the other." Given the actual preference ordering, one can define

$$l(C_{12}) = x l(C_{21})\,, \qquad \text{for } x > 1 \tag{13}$$

The practitioner needs to specify how much worse he considers it to wrongly declare an individual who has not effectively exceeded the age of 18 years an adult, with respect to the opposite, that is, to wrongly declare an individual who is effectively older than 18 years a minor. Then, being $l(C_{12})$ set equal to 1 because of the choice of a '0-1' unit scale, the loss associated with C_{21} can be immediately obtained as $l(C_{21}) = {}^{1}\!/_{x}$.

For instance if the decision-maker feels that an erroneous declaration about a person being an adult is two times worse than an erroneous declaration of the person being a minor, then, $x = 2$ and l_{21}=0.50; in case the former would be considered 10 times worse than the latter, then $x = 10$ and l_{21}=0.10. Recalls the example provided by the study of Polo Grillo et al. (2002) and suppose that the evidence evaluation lead to a value of the posterior probability on θ_1 of 0.70, i.e.,

$Pr(\theta_1|\boldsymbol{E}, I) = 0.70$. In this case, from Eq. (12), the rational decision is d_1 (declare the examined person as adult) if and only if $l(C_{21}) > 0.42$ (approximately). Thus, if the decision-maker beliefs that a “false adult” is only two time worse than a “false minor”, then the rational decision is to declare the examined person as adult (d_1). If instead, he or she feels that it is 10 times worse, then the rational decision is to declare the examined person as minor (d_2). Note that the choice of the value of x (as well as any quantifications of the loss value) is subjective and it is based on the personal belief and the personal knowledge of the decision-maker. Elements that can support such choice are the framework of the case (criminal versus asylum cases) or the law system in force in a given country. Analogously to the assignment of subjective probability, these subjective losses are a formalization of the personal belief of the decision-maker, which is informed by all the pieces of information available by the decision-maker. Subjective losses are thus perfectly compatible with the forensic and legal framework, provided that the choice made is coherent and can be justified (Taroni et al., 2010, Biedermann et al., 2016).

A sensitivity analysis is strongly suggested also in this case. Different loss assignments, as well as different prior assignments, will give rise to different expected losses. This must not be considered to be a weakness of the endorsed Bayesian criterion for making decisions. Different prior probabilities, as well as different losses, might fit as well a given degree of belief or a given preference structure, and therefore different expected losses might be entirely justifiable.

7. Discussion and conclusion

The interpretation of scientific evidence is an inferential (inductive) task and thus naturally involves uncertainty. The specific case of age-related evidence in the forensic age estimation framework does not represent an exception to this statement. Assuming that probability is the measure of quantifying uncertainty and that the Bayesian approach provides a logical framework to the problem of induction, the expert plays a central role in the choice of (i) the prior probability distribution on the chronological age, (ii) the statistical model to handle uncertainty about the evidence, and (iii) the relevant database to inform the likelihood function.

It has already been pointed out that the expert must often deal with the absence of adequate databases. This is not surprising, as the collection of the biological age from multiple age indicators for the same individual is extremely difficult, especially in populations generally involved in age estimation procedures. However, the lack of information must not be equated with the impossibility of implementing a probabilistic model. The lack of background data is a common problem shared in various forensic disciplines, at the point that the ENFSI (2015) considers this possibility in its guidelines. In such a context, the role of the expert becomes even more relevant, since the probabilities of the events of interest (i.e., the probability of dental or skeletal evidence given the chronological age) can reasonably be assigned based on a personal (and justified) body of knowledge, by using unpublished data or data published in reference studies (ENFSI, 2015). In this perspective, reference studies including the parameters estimated through regression models applied to unpublished datasets can be extremely useful: such data can be implemented by the expert to quantify likelihood function (Konigsberg, 2015, Konigsberg et al., 2019).

As far as the uncertainty about the chronological age is concerned, an operationally valid alternative to the introduction of a probability density function over a continuous random variable (i.e., age) would be to consider specific age cohorts in the form of competing propositions. This allows us to facilitate the quantification of uncertainty regarding the age in the likely situation of scarce information, as probabilities must be assigned to two discrete events rather than over the entire span of the continuous age range. In this particular case, Taroni et al. (2018) recommend to gather available information (e.g., from the literature) in the form of relative frequencies to inform one's personal beliefs.

In recent years, the normative decisional approach has been promoted in forensic sciences (Taroni et al., 2005, Taroni et al., 2010, Gittelson, 2013, Biedermann et al., 2016, Biedermann et al., 2018), since it offers a structured way of thinking taking advantage of quantitative data, knowledge and experience. This paper illustrates how to use this approach for forensic age estimation in living persons.

Acknowledgment

In memory of David Lucy to whom we owe a debt of gratitude for all that he done for the subject of age estimation and probabilistic reasoning.

9. References

Aitken, C. & Nordgaard, A. 2018. The roles of participants' differing background information in the evaluation of evidence. J Forensic Sci 63, 648-649.

Aitken, C. G. G. & Taroni, F. 2004. Statistics and the evaluation of evidence for forensic scientists. 2nd ed, Chichester: John Wiley & Sons.

Bassed, R. B. 2012. Advances in forensic age estimation. Forensic Sci Med Pathol 8, 194-196.

Berger, C. E. H., Buckleton, J. S., Champod, C., Evett, I. W. & Jackson, G. 2011. Evidence evaluation: A response to the court of appeal judgment in RvT. Sci Justice 51, 43-49.

Berger, J. O. 2010. Statistical decision theory and Bayesian analysis. 2nd edition ed, New York: Springer-Verlag.

Biedermann, A. 2015. The role of the subjectivist position in the probabilization of forensic science. J Forensic Sci Med 1, 140-148.

Biedermann, A., Bozza, S. & Taroni, F. 2008. Decision theoretic properties of forensic identification: Underlying logic and argumentative implications. Forensic Sci Int 177, 120-132.

Biedermann, A., Bozza, S. & Taroni, F. 2016. The decisionalization of individualization. Forensic Sci Int 266, 29-38.

Biedermann, A., Bozza, S. & Taroni, F. 2017a. Analysing and exemplifying forensic conclusion criteria in terms of Bayesian decision theory. Sci Justice 58, 159-165.

Biedermann, A., Bozza, S. & Taroni, F. 2018. Normative decision analysis in forensic science. Artif Intell & Law. Doi: https://doi.org/10.1007/s10506-018-9232-2.

Biedermann, A., Bozza, S., Taroni, F. & Aitken, C. 2017b. The meaning of justified subjectivism and its role in the reconciliation of recent disagreements over forensic probabilism. Sci Justice 57, 477-483.

Bleka, Ø., Wisløff, T., Dahlberg, P. S., Rolseth, V. & Egeland, T. 2018. Advancing estimation of chronological age by utilizing available evidence based on two radiographical methods. Int J Legal Med 133, 217-229.

Boldsen, J. L., Milner, G. R., Konigsberg, L. W. & Wood, J. W. 2002. Transition analysis: A new method for estimating age from skeletons. In: Hoppa, R. D. & Vaupel, J. V. (eds.), Paleodemography. Age distributions from skeletal samples. Cambridge: Cambridge University Press, pp. 73-106.

Bolstad, W. M. & Curran, J. M. (eds.) 2017. *Introduction to Bayesian statistics,* Hoboken: John Wiley & Sons.

Braga, J., Heuze, Y., Chabadel, O., Sonan, N. K. & Gueramy, A. 2005. Non-adult dental age assessment: Correspondence analysis and linear regression versus Bayesian predictions. Int J Legal Med 119, 260-274.

Cameriere, R., Pacifici, A., Pacifici, L., Polimeni, A., Federici, F., Cingolani, M. & Ferrante, L. 2016. Age estimation in children by measurement of open apices in teeth with Bayesian calibration approach. Forensic Sci Int 258, 50-54.

Cameron, N. & Jones, L. L. 2010. Growth, maturation and age. In: Black, S., Aggrawal, A. & Payne-James, J. (eds.), Age estimation in the living: The practioner's guide. Hoboken: John Wiley & Sons, pp. 95-129.

Comitee of Forensic Experts 2011. Expressing evaluative opinions: A position statement. Sci Justice 51, 1-2.

Corradi, F., Pinchi, V., Barsanti, I. & Garatti, S. 2013a. Probabilistic classification of age by third molar development: The use of soft evidence. J Forensic Sci 58, 51-59.

Corradi, F., Pinchi, V., Barsanti, I., Manca, R. & Garatti, S. 2013b. Optimal age classification of young individuals based on dental evidence in civil and criminal proceedings. Int J Legal Med 127, 1157-1164.

Cunha, E., Baccino, E., Martrille, L., Ramsthaler, F., Prieto, J., Schuliar, Y., Lynnerup, N. & Cattaneo, C. 2009. The problem of aging human remains and living individuals: A review. Forensic Sci Int 193, 1-13.

Curran, J. M. 2013. The frequentist approach to forensic evidence interpretation. In: Siegel, J. A. & Saukko, P. J. (eds.), Ecyclopedia of forensic sciences. Waltham: Academic Press, pp. 286-291.

Degroot, M. H. 1970. Optimal statistical decisions. New York: McGraw-Hill.

EASO 2018. Practical guide on age assessment. Available from: https://www.easo.europa.eu/sites/default/files/easo-practical-guide-on-age-assesment-v3-2018.pdf [Accessed 26.11.2019].

ENFSI 2015. ENFSI guideline for evaluative reporting in forensic science: Strengthening the evaluation of forensic results across Europe. Available from: http://enfsi.eu/wp-content/uploads/2016/09/m1_guideline.pdf [Accessed 26.11.2019].

Evett, I. W. 1995. Avoiding the transposed conditional. Sci Justice 35, 127-131.

Evett, I. W. 2015. The logical foundations of forensic science: Towards reliable knowledge. Phil. Trans. R. Soc. B 370, 1-10.

Fieuws, S., Willems, G., Larsen-Tangmose, S., Lynnerup, N., Boldsen, J. & Thevissen, P. 2015. Obtaining appropriate interval estimates for age when multiple indicators are used: Evaluation of an ad-hoc procedure. Int J Legal Med 130, 489-499.

Garamendi, P. M., Landa, M. I., Ballesteros, J. & Solanoa, M. A. 2005. Reliability of the methods applied to assess age minority in living subjects around 18 years old. A survey on a moroccan origin population. Forensic Sci Int 154, 3-12.

Gelbrich, B., Frerking, C., Weiß, S., Schwerdt, S., Stellzig-Eisenhauer, A., Tausche, E. & Gelbrich, G. 2015. Combining wrist age and third molars in forensic age estimation: How to calculate the joint age estimate and its error rate in age diagnostics. Ann Hum Biol 42, 389-396.

Gittelson, S. 2013. Evolving from inferences to decisions in the interpretation of scientific evidence. PhD Thesis, University of Lausanne.

Gittelson, S., Biedermann, A., Bozza, S. & Taroni, F. 2014. Decision analysis for the genotype designation in low-template-DNA profiles. Forensic Sci Int-Gen 9, 118-133.

Gittelson, S., Bozza, S., Biedermann, A. & Taroni, F. 2013. Decision-theoretic reflections on processing a fingermark. Forensic Sci Int 226, e42-e47.

Gittelson, S., Steffen, C. R. & Coble, M. D. 2016. Low-template DNA: A single DNA analysis or two replicates? Forensic Sci Int 264, 139-145.

Hackman, S. L., Buck, A. & Black, S. 2010. Age evaluation from the skeleton. In: Black, S., Aggrawal, A. & Payne-James, J. (eds.), Age estimation in the living: The practioner's guide. Hoboken: John Wiley & Sons, pp. 202-235.

Hillewig, E., Degroote, J., Van Der Paelt, T., Visscher, A., Vandemaele, P., Lutin, B., D'hooghe, L., Vandriessche, V., Piette, M. & Verstraete, K. 2013. Magnetic resonance imaging of the sternal extremity of the clavicle in forensic age estimation: Towards more sound age estimates. Int J Legal Med 127, 677-689.

Howson, C. 2002. Bayesianism in statistics. In: Swinburne, R. (ed.) Bayes's theorem. Oxford: Oxford University Press, pp. 39-69.

Jeffreys, H. 1961. Theory of probability. 3rd ed, Oxford: Clarendon Press.

Kemkes-Grottenthaler, A. 2002. Aging through the ages: Historical perspectives on age indicator methods. In: Hoppa, R. D. & Vaupel, J. V. (eds.), Paleodemography. Age distributions from skeletal samples. Cambridge: Cambridge University Press, pp. 48-72.

Koller, D. & Friedman, N. 2009. Probabilistic graphical models. Principles and techniques. Londre: The MIT Press.

Konigsberg, L. W. 2015. Multivariate cumulative probit for age estimation using ordinal categorical data. Ann Hum Biol 42, 368-378.

Konigsberg, L. W., Frankenberg, S. R. & Liversidge, H. M. 2019. Status of mandibular third molar development as evidence in legal age threshold cases. J Forensic Sci 64, 680-697.

Konigsberg, L. W. & Herrmann, N. P. 2002. Markov chain monte carlo estimation of hazard model parameters in paleodemography. In: Hoppa, R. D. & Vaupel, J. V. (eds.), Paleodemography. Age distributions from skeletal samples. Cambridge: Cambridge University Press, pp. 222-242.

Konigsberg, L. W., Herrmann, N. P., Wescott, D. J. & Kimmerle, E. H. 2008. Estimation and evidence in forensic anthropology: Age-at-death. J Forensic Sci 53, 541-557.

Law, H., Mensah, L., Bailey, S. & Nelki, J. 2010. Immigration, asylum seekers and undocumented identity. In: Black, S., Aggrawal, A. & Payne-James, J. (eds.), Age estimation in the living: The practioner's guide. Hoboken: John Wiley & Sons, pp. 19-29.

Lindley, D. V. 1985. Making decision. Chichester: John Wiley & Son.

Lindley, D. V. 1991. Probability. In: Aitken, C. G. G. & Stoney, D. A. (eds.), The use of statistics in forensic science. Chichester: Ellis Horwood, pp. 27-50.

Lindley, D. V. 2006. Understanding uncertainty. Hoboken: John Wiley & Sons.

Liversidge, H. M. & Marsden, P. H. 2010. Estimating age and the likelihood of having attained 18 years of age using mandibular third molars. Br Dent J 209, E13.

Lucy, D. 2010. The presentation of results and statistics for legal purposes. In: Black, S., Aggrawal, A. & Payne-James, J. (eds.), Age estimation in the living: The practioner's guide. Hoboken: John Wiley & Sons, pp. 267-283.

Lucy, D., Aykroyd, R. G. & Pollard, A. M. 2002. Nonparametric calibration for age estimation. Appl Stat 51**,** 183-196.

Malmqvist, E., Furberg, E. & Sandman, L. 2018. Ethical aspects of medical age assessment in the asylum process: A swedish perspective. Int J Legal Med 132**,** 815-823.

Mincer, H. H., Harris, E. F. & Berryman, H. E. 1993. The A.B.F.O. Study of third molar development and its use as an estimator of chronological age. J Forensic Sci 38**,** 379-390.

Myers, R. H., Montgomery, D. C. & Vining, G. G. 2002. Generalized linear models with applications in engineering and the sciences. New York: John Wiley & Sons.

O'Hagan, A., Buck, C. E., Daneshkhah, A., Eiser, J. R., Garthwaite, Paul H, Jenkinson, D. J., Oakley, J. E. & Rakow, T. 2006. Uncertain judgements: Eliciting experts' probabilities. Chichester: John Wiley & Sons.

Polo Grillo, B., Cattaneo, C., Panigalli, C. & Grandi, M. Can a combined hand–wrist method increase accuracy in age determnation of offenders who claim to be under 18 years of age? 16th Meeting of the International Association of Forensic Sciences, 2–7 September 2002 Montpellier, France.

Ritz-Timme, S., Cattaneo, C., Collins, M. J., Waite, E. R., Schütz, H. W., Kaatsch, H.-J. & Borrman, H. I. M. 2000. Age estimation: The state of the art in relation to the specific demands of forensic practise. Int J Legal Med 113**,** 129-136.

Robertson, B. & Vignaux, G. A. 1998. Explaining evidence logically. New Law Journal 148**,** 159-162.

Robertson, B., Vignaux, G. A. & Berger, C. E. H. 2016. Interpreting evidence. Evaluating forensic science in the courtroom. 2nd ed, Chichester: John Wiley & Sons.

Schmeling, A., Dettmeyer, R., Rudolf, E., Vieth, V. & Geserick, G. 2016. Forensic age estimation - methods, certainty, and the law. Dtsch Arztebl Int 113**,** 44-50.

Schmeling, A., Garamendi, P. M., Prieto, J. L. & Landa, M. I. 2011. Forensic age estimation in unaccompanied minors and young living adults. In: Vieira, D. N. (ed.) Forensic medicine - from old problems to new challenges. Rijeka: InTech, pp. 77-120.

Schmeling, A., Geserick, G., Reisinger, W. & Olze, A. 2007. Age estimation. Forensic Sci Int 165**,** 178-181.

Schmeling, A., Grundman, C., Fuhrmann, A., Kaatsch, H.-J., Knell, B., Ramsthaler, F., Reisinger, W., Riepert, T., Ritz-Timme, S., Rösing, F. W., Rötzscher, K. & Geserick, G. 2008. Criteria for age estimation in living individuals. Int J Legal Med 122**,** 457-460.

Schmeling, A., Olze, A., Reisinger, W. & Geserick, G. 2005. Forensic age estimation and ethnicity. Leg Med 7**,** 134-137.

Schmeling, A., Olze, A., Reisinger, W., König, M. & Geserick, G. 2003. Statistical analysis and verification of forensic age estimation of living persons in the institute of legal medicine of the berlin university hospital charite. Leg Med 5**,** S367-S371.

Schmeling, A., Reisinger, W., Geserick, G. & Olze, A. 2006a. Age estimation of unaccompanied minors. Part i. General considerations. Forensic Sci Int 159 Supl. 1, S61-S64.

Schmeling, A., Schulz, R., Danner, B. & Rösing, F. 2006b. The impact of economic progress and modernization in medicine on the ossification of hand and wrist. Int J Legal Med 120, 121-126.

Schum, D. A. 1994. The evidential foundations of probabilistic reasoning. New York: John Wiley & Sons.

Sironi, E., Gallidabino, M., Weyermann, C. & Taroni, F. 2016. Probabilistic graphical models to deal with age estimation of living persons. Int J Legal Med 130, 475-488.

Sironi, E., Pinchi, V., Pradella, F., Focardi, M., Bozza, S. & Taroni, F. 2018a. Bayesian networks of age estimation and classification based on dental evidence: A study on the third molar mineralization. J Forensic Leg Med 55, 23-32.

Sironi, E., Pinchi, V. & Taroni, F. 2015. Probabilistic age classification with Bayesian networks: A study on the ossification status of the medial clavicular epiphysis. Forensic Sci Int 258, 81-87.

Sironi, E. & Taroni, F. 2015. Bayesian networks for the age classification of living individuals. A study on transition analysis. J Forensic Leg Med 1, 124-132.

Sironi, E., Taroni, F., Baldinotti, C., Nardi, C., Norelli, G.-A., Gallidabino, M. & Pinchi, V. 2018b. Age estimation by assessment of pulp chamber volume: A Bayesian network for the evaluation of dental evidence. Int J Legal Med 132, 1125-1138.

Sironi, E., Vuille, J., Morling, N. & Taroni, F. 2017. On the Bayesian approach to forensic age estimation of living individuals. Forensic Sci Int 281, e24-e29.

Tangmose, S., Thevissen, P., Lynnerup, N., Willems, G. & Boldsen, J. 2015. Age estimation in the living: Transition analysis on developing third molars. Forensic Sci Int 257, 512e1-e7.

Taroni, F., Aitken, C. G. G. & Garbolino, P. 2001. De Finetti's subjectivism, the assessment of probabilities and the evaluation of evidence: A commentary for forensic scientists. Sci Justice 41, 145-150.

Taroni, F. & Biedermann, A. 2013. Bayesian networks. In: Siegel, J. A. & Saukko, P. J. (eds.), Ecyclopedia of Forensic Sciences. Waltham: Academic Press, pp. 351-356.

Taroni, F. & Biedermann, A. 2014. Probability and inference in forensic science. In: Bruinsma, G. & Weisburd, D. (eds.), Encyclopedia of criminology and criminal justice. New York: Springer Science & Business Media, pp. 3947-3957.

Taroni, F. & Biedermann, A. 2015. Uncertainty in forensic science: Experts, probabilities and Bayes' theorem. Statistica Applicata (Ital J Appl Stat) 27, 129-144.

Taroni, F., Biedermann, A., Bozza, S., Garbolino, P. & Aitken, C. 2014. Bayesian networks for probabilistic inference and decision analysis in forensic science. 2ème ed, Chichester: John Wiley & Sons.

Taroni, F., Bozza, S. & Aitken, C. 2005. Decision analysis in forensic science. J Forensic Sci 50, 1-12.

Taroni, F., Bozza, S., Biedermann, A. & Aitken, C. 2015. Dismissal of the illusion of uncertainty in the assessment of a likelihood ratio. Law, Prob & Risk 15, 1-16.

Taroni, F., Bozza, S., Biedermann, A., Garbolino, P. & Aitken, C. 2010. Data analysis in forensic science: A Bayesian decision perspective. Chichester: John Wiley & Sons.

Taroni, F., Garbolino, P., Biedermann, A., Aitken, C. & Bozza, S. 2018. Reconciliation of subjective probabilities and frequencies in forensic science. Law, Prob & Risk 17, 243-262.

Thevissen, P., Fieuws, S. & Willems, G. 2010. Human dental age estimation using third molar developmental stages: Does a Bayesian approach outperform regression models to discriminate between juveniles and adults? Int J Legal Med 124, 35-42.