Age estimation based on different molecular clocks in several tissues and a multivariate approach: An explorative study

J. Becker [1], N. S. Mahlke [1]*, A. Reckert [1], S. B. Eickhoff [2,3], S. Ritz-Timme [1]*

[1] Institute of Legal Medicine, University Hospital Düsseldorf, 40225 Düsseldorf, Germany

[2] Institute for Systems Neuroscience, University Hospital Düsseldorf, 40225 Düsseldorf, Germany

[3] Institute of Neuroscience and Medicine, Brain & Behaviour (INM-7), Research Centre Jülich, 52428 Jülich, Germany

Corresponding author

Nina Sophia Mahlke

nina.mahlke@hhu.de

Tel: 0049-211-8104907

Fax: 0049-211-8119366

*Julia Becker and Nina Mahlke contributed equally to this work and share first authorship.

Abstract

Several molecular modifications accumulate in the human organism with increasing age. Some of these „molecular clocks“ in DNA and in proteins open up promising approaches for the development of methods for forensic age estimation. A natural limitation of these methods arises from the fact that the chronological age is determined only indirectly by analyzing defined molecular changes that occur during aging. These changes are not linked exclusively to the expired life span but may be influenced significantly by intrinsic and extrinsic factors in the complex process of individual aging.

We tested the hypothesis that a combined use of different “molecular clocks” in different tissues results in more precise age estimates because this approach addresses the complex ageing processes in a more comprehensive way. Two molecular clocks (accumulation of D-aspartic acid (D-Asp), accumulation of pentosidine (PEN)) in two different tissues (annulus fibrosus of intervertebral discs and elastic cartilage of the epiglottis) were analyzed in 95 cases, and uni- and multivariate models for age estimation were generated.

The more parameters were included in the models for age estimation, the smaller the mean absolute errors (MAE) became. While the MAEs were 7.5 – 11.0 years in univariate models, a multivariate model based on the two protein clocks in the two tissues resulted in a MAE of 4.0 years.

These results support our hypothesis. The tested approach of a combined analysis of different molecular clocks analyzed in different tissues opens up new possibilities in postmortem age estimation. In a next step, we will add the epigenetic clock (DNA methylation) to our protein clocks (PEN, D-Asp) and expand our set of tissues.

Keywords: Age estimation, pentosidine, D-aspartic acid, machine learning, age prediction model, molecular clocks

Introduction

In times of migration and flight age estimation is becoming an increasingly important issue in Forensic Medicine. This applies not only for age estimation in living young migrants without valid documents, but also for age estimation in the context of the identification of unknown deceased. Postmortem age estimation may be a methodological challenge due to putrefaction and decomposition of a body; in such cases, the applicability of a method may depend on the extent of postmortem changes and the availability of tissues.

Generally, methods of age estimation need to be accurate enough to fulfill the demands in forensic practice [1]. Therefore, the core objective in the development of new methods or the optimization of already established methods for age estimation must be to achieve the highest possible accuracy.

In the last decade, molecular methods of age estimation have attracted much attention. So called “molecular clocks” like the racemization of aspartic acid or the accumulation of pentosidine in proteins (see [2–5]) and, most recently, the methylation of DNA (see [2–4, 6–12]) have been identified and their usability for age estimation has been explored.

DNA methylation (mDNA) is used as basis for epigenetic age estimation. Age-dependent methylation markers have been identified in samples from different tissues and body fluids, and various models for mDNA based age estimation have already been proposed ([13], for review see [6, 8, 10–12, 14]). However, it is well known that epigenetic changes in DNA are influenced by intrinsic factors (e.g. genetic factors, ethnicity, diseases) and extrinsic factors (e.g. environmental conditions, lifestyle) [15, 16]. The implications of such influences on the ultimately achievable accuracy of mDNA based approaches for age estimation are not finally clear and are subject of current research [10, 11, 16–18].

The accumulation of D-aspartic acid (D-Asp) is the result of a non-enzymatic conversion of L-asparagine residues and L-aspartic acid residues into their D-forms (for details see [3, 19, 20]). It has been described in multiple proteins and tissues (for review see [2], data regarding further proteins/tissues in [21–23]). Age estimation based on the D-Asp content in dentinal protein is one of the most accurate methods for age estimation in adults. The application of this approach in forensic scenarios has been proposed decades ago (e.g. [24–26]). In the meantime it has been increasingly established in forensic practice (e.g. [27–30]). This protein clock is very robust as long as the integrity of the corresponding tissue is guaranteed [25]. So far there is no indication of genetic influences on this clock, and ethnicity seems to have no impact [31]. Open research questions concern in particular the applicability of the method to more complex tissues than dentine.

Pentosidine (PEN) is an advanced glycation end product. This group of posttranslational protein modifications is the result of glycation of proteins which is a non-enzymatic reaction of free amino groups (mainly of arginine and lysine) of proteins with glucose or with other reducing carbohydrates [32–35]. It has already been shown that PEN accumulates in permanent and long-living proteins during lifetime in several tissues [36–43]. Theoretically, elevated blood sugar levels in patients with diabetes mellitus may result in an increased formation of PEN [32, 35, 44, 45], which may limit the applicability of this approach to forensic age estimation. However, up to now there are only a few systematic studies regarding this protein clock.

Even promising molecular methods such as methods based on mDNA still have a considerable error range, which currently limits their practical use [7]. In light of the complexity of the biological aging process this it is not surprising.

The best models for age prediction based on mDNA described so far use combinations of up to hundreds of markers and are associated with mean absolute errors (MAE) of 3 - 5 years at best (for review see [6–8, 10, 11]). It has to be taken into account, however, that the deviation between estimated and real age may be considerably higher (theoretically a multiple of the MAE) in single cases [7]. Methods based on the accumulation of PEN have an even larger error range [36, 41]. Age estimation based on the accumulation of D-Asp shows quite accurate results when applied to dentine [5, 24–27, 29]. In contrast, the application of this approach to other, more complex tissues reveal considerably less accurate results; this is especially evident, if tissue samples are analyzed without a preceding protein purification [21–23, 46–52].

If age estimation has to be performed postmortem, it is - dependent on the state of decomposition of a body - possible to analyze several molecular clocks in several tissues. Different molecular clocks in different tissues underlie different influences since they are part of different biological subsystems. A combination of different and at best independent molecular clocks for age estimation may better record complex ageing processes and balance the effects of different influence factors on the accuracy of age estimation.

We tested this hypothesis for the indication “postmortem age estimation”

- by analyzing two molecular clocks in proteins (accumulation of D-Asp, accumulation of PEN) in two different tissues (annulus fibrosus of intervertebral discs (IVD) and elastic cartilage of epiglottis (EPI)) from 95 individuals and
- by generating uni- and multivariate models for age estimation based on the collected data.

The two tissues (IVD and EPI) were chosen since they are relatively robust against putrefaction and can be prepared easily.

Material and Methods

Samples of intervertebral discs (IVD) and epiglottis (EPI) were drawn from 95 individuals (47 females, 48 males, age range 0.06 - 98 years) during autopsy. All available documents (especially autopsy protocols and circumstances of each case) were checked for relevant information regarding the case history; particular attention was paid to evidence of diabetes mellitus. The postmortem intervals of these 95 cases were between 1 and 12 days.

In addition, some cases with special postmortem conditions were examined: 9 cases with advanced putrefaction (postmortem intervals a few days to about three weeks; in one of these cases EPI was not available anymore), 3 burned bodies (one incompletely, and two totally charred) and 3 exhumed bodies (postmortem intervals 54, 59, and 68 days, respectively, with advanced putrefaction).

The following parameters were determined in the IVD and EPI samples:

- IVD: D-Asp in an enzymatically purified collagen fraction,
- IVD: PEN in an enzymatically purified collagen fraction,
- EPI: D-Asp in total tissue,
- EPI: PEN in an enzymatically purified collagen/elastin fraction

The relationship between each of the 4 parameters and age was tested, and uni- and multivariate models for age estimation based on these parameters were developed.

Preparation of the IVD samples

The intervertebral disc L2/3 was removed as intact as possible, washed with water and stored at - 80 °C until further preparation.

Degenerative changes of the discs were documented and classified into the two groups ”no or slightly” and “highly” according to the morphology of the anterior annulus fibrosus. “Highly” degenerative discs exhibited findings that indicate severe alterations (e.g. radial fissures or scar tissue) with a destruction of the typical morphological structure. 41 samples (from individuals with ages between 46 and 98 years) exhibited such “highly” degenerative changes.

For further preparation the IVD samples were thawed. Approximately 0.5 – 0.8 cm of the outer layers of the anterior annulus fibrosus were cut off to remove adjacent tissue and small

samples (1 x 1 cm) were prepared. These samples were coarsely crushed, dipped in liquid nitrogen, crushed again and pulverized. The pulverized material was stored at -20 °C until further processing (purification of collagen).

Preparation of the EPI samples

The upper two thirds of the epiglottis were taken during autopsy and stored at -80 °C until further preparation.

For further preparation the samples were thawed. After removal of the mucosa and surrounding tissue, the cartilage was cut into small pieces. A small piece (about 2 mm in diameter) was put into a sample tube for analysis of D-Asp and stored at -20 °C until analysis. Another piece of tissue (about 8 mm in diameter) was wrapped in aluminum foil, frozen in liquid nitrogen and crushed. The pulverized material was stored at -20 °C until further processing (purification of collagen/elastin).

Enzymatic purification of collagen (IVD) and collagen/elastin (EPI)

IVD and EPI samples were enzymatically treated according to the protocol of Sivan et al. [53], based on the method of Schmidt et al. [54]. In short, a proteolytic enzymatic treatment is applied to remove all non-collagen proteins such as proteoglycans with Chondroitinase ABC (0.125 unit/ml in 0.05 M Tris-Base/0.06 M sodium acetate buffer (pH 8); for 24 h at 37 °C), Streptomyces hyaluronidase (1 unit/ml in 0.05 M Tris-Base/0.15 M NaCl buffer (pH 6); for 24 h at 37 °C) and trypsin (1 mg/ml in 0.05 M Na2HPO4/0.15 M NaCl buffer (pH 7.2); for 16 h at 37 °C). Purified samples were dried by freeze drying and kept at - 80 °C until further use.

The quality of the purification was evaluated by amino acid analysis performed by high performance liquid chromatography (HPLC) according to the method of Dobberstein et al. [21]. An aliquot of each purified sample was transferred in a Pyrex tube and hydrolyzed in 1 ml of a 6 N HCl for 24 h at 110 °C. Liquid in excess was removed overnight under vacuum. The dried hydrolysates were then dissolved in 400 µl 0.01 N HCl. Human collagen type I (Sigma-Aldrich/Merck KGaA, Darmstadt, Germany) and bovine elastin (Sigma-Aldrich/Merck KGaA, Darmstadt, Germany) served as external standards for the adjustment of the amino acid distribution.
Analyzes were performed by HPLC (HPLC 1100 Series, Agilent, CA). OPA reagent (o-phthaldialdehyde) was used for precolumn derivatization of the primary amino acids, secondary amino acids were derivatized using FMOC (9-flourenylmethylchloroformate). Separation was accomplished with a C18 column (Hypersil BDS, C18 250 × 3 mm, particle

size 5 µm; Thermo Electron GmbH, Dreieich, Germany) and a mobile phase consisted of eluents A (40 mM NaH_2PO_4, 1.5 mM sodium azide) and B (45 % methanol, 45 % acetonitrile, 10 % H_2O) according to Heems et al. [55]. After an initial equilibration step of about 5 min, the amino acid derivatives were detected over a period of 50 minutes using a binary gradient. The flow rate was constant at 1.2 ml/min and the column temperature at 40 °C. The OPA derivatives were measured at an excitation wavelength of λ = 335 nm and a detection wavelength of λ = 440 nm. Analysis of the FMOC derivatives was performed after a switching point at 30 min, an excitation wavelength of λ = 260 nm and a detection wavelength of λ = 305 nm were used. Signal identification and quantification were carried out by a calibrated external standard.

AAR analysis

The extent of aspartic acid racemization was determined by HPLC according to the method of Kaufman and Manley [56], modified by Dobberstein et al. [21]. An aliquot of each sample (IVD or EPI) was hydrolyzed in 1 ml 6 N hydrochloric acid for 6 h at 100 °C in Pyrex tubes. The excess liquid was removed overnight under vacuum. Samples were dissolved in 1 ml sample buffer (0.01 M HCL with 1.5 mM sodium azide and 0.03 mM L-homo-arginine). For HPLC analysis, a C18 column from Thermo Scientific (Hypersil BDS C18, 250 x 3 mm, particle size 5 µm) served as the stationary phase. The mobile phase consisted of the eluents A (23 mM sodium acetate, 1.5 mM sodium azide, 1 mM EDTA) and B (92.3 % methanol, 7.7 % acetonitrile). The amino acid enantiomers were detected by a binary gradient over a period of 115 minutes at a constant flow rate of 0.56 ml/min. Amino acids were detected at an excitation wavelength of λ = 230 nm and a detection wavelength of λ = 445 nm. L- and D- aspartic acid residues were identified using the retention times of the amino acids in a standard solution. The accumulation of D-Asp was described by the term ln ((1+D/L)/(1-D/L)) (see [20]).

PEN analysis

PEN concentrations were determined by HPLC as described by Greis et al. [36]. Aliquots of IVD or EPI samples (3.8 - 10.7 mg and 0.7 - 10.8 mg, respectively) were hydrolyzed in 1 mL 6 N HCl for 18 h at 110 °C. The samples were dried overnight in a desiccator, then dissolved in 1 ml 0.01 M heptafluorobutyric acid (HFBA, Thermo Scientific, Rockford, IL, in HPLC-water, HiPerSolv Chromanorm, VWR International), filtrated (syringe filters, pore diameter: 0.45 µm, diameter: 25 mm, VWR International) and dried again overnight in a desiccator. For HPLC analysis, the dried residues were dissolved in 200 µL pyridoxine-HFBA (pyridoxine 2.068815 µmol/mL in 0.01 M HFBA) for IVD and 250 µL for EPI, respectively. Calibration

curve was established with standard PEN samples (pentosidine 0.03303 nmol/mL in 0.01 M HFBA, Cayman Chemical). 50 µL of each sample were injected into the HPLC system (HPLC 1100 Series, Agilent, CA). A semipreparative column (Onyx™ Monolithic Semi-PREP C18, LC Column 100 x 10 mm, Phenomenex, CA) was used as stationary phase. Mobile Phase consists of 0.1 % HFBA (eluent A) and acetonitrile (eluent B, LiChrosolv, Merck KGaA, Darmstadt); a linear gradient program of 10 - 85 % acetonitrile from 0 to 32 min with 0.1 % HFBA was used. The flow rate was 1 mL/min and the column temperature was set to 25 °C. Detection was done at wavelength 335/385 nm. Data were analyzed using „HPChemStation“. PEN was identified by its retention time; its quantification was carried out by determination of the peak area and use of a calibration curve.

Statistics

Classic descriptive statistics

The relationship between chronological age and the accumulation of D-Asp and PEN was tested by rank correlation, and the corresponding correlation coefficients (Spearman) were determined.

Development of multivariate age predicting models by a machine learning approach

Multivariate prediction of individual age was performed using an ensemble of Gaussian Process Regression (GPR) predictors. The key idea behind an ensemble predictor is to repeatedly fit a model for mapping between the input data (features) and the target variable (chronological age) based on only subsets of the training observations, i.e., subjects [57]. That is, we repeatedly draw subsamples of subjects from the training set and then fit the GPR model on this lower number of subjects. This procedure yields a “weak learner” (as it is not based on the entire training sample) that is applied the test-data in order to yield one age prediction. These are then combined into a final prediction allowing to obtain better performance than obtainable from any of the constituent predictors by itself. In practice, we repeatedly sampled (with replacement) 95% of the training cases but retained only one instance of each individual case, which brings the additional advantage of variance in the size of the training set (cf. the number of unique observations contributing to a bootstrap).

For each run, we trained a GPR model [58] to identify the best transformation (for this particular subsample of the subjects in the training set) between the input features and the target, i.e., to predict age based on the respective molecular markers as well as information on the available clinical data. As noted above, training was repeatedly performed on a subset of the training cases, yielding “weak learners” that were later combined to achieve the final prediction. For training each of the weak learners, we used the GPR algorithm implemented

in the MATLAB “statistics and machine learning toolbox” (*fitrgp* function) using a squared exponential kernel with a separate length scale per predictor. Gender and degeneration entered the model as categorical predictors, the remaining continuous features were standardized, i.e., centered and scaled by column mean and standard deviation, respectively. For both fit and prediction, the “exact” Gaussian process regression method was used. Each of these models based on a subset of the training subjects was then applied to predict the age of a held-out, "new" subject, i.e., a case that was not part of the training set. This prediction is recorded, and the procedure repeated 5,000 times with new, independent sampling from the training data, each yielding a new prediction of the test case. These individual predictions are then averaged, a process termed "bagging", to yield the final prediction for the test case. Cycling over all possible test cases then yields an out-of-sample prediction of the age for each subject in the current sample. That is, when predicting the age of an individual subject, we do not use any information about that particular person when training the model.

As a measure of prediction accuracy, we computed the mean absolute deviation (MAE). Importantly, MAE was calculated for “out-of-bag” predictions, i.e., assessing the age predicted for each individual subject when it was not part of the training set. More precisely, as noted above, we averaged the predictions of individual age based on the different weak learners fitted on subsamples of the training data (bagging). This yielded an age prediction for each case based on an ensemble that had no information on this particular subject, as it has not been part on the training set. We then computed the mean (averaged across subjects) absolute differences between the predicted (based on models trained on different cases) and the true chronological age for each subject.

Here, our primary interest related to the prediction of individual age based on all data available for this particular subject, i.e., using both molecular clocks from both tissues as well as the information about sex and IVD degeneration. However, given the scenario of only having either epiglottis or IVD tissue available, we also fitted reduced models based on only one tissue and the information about sex and IVD degeneration.

We developed multivariate models for the following scenarios:

<u>Model1 (only IVD available):</u>

PEN/IVD & D-Asp/IVD & sex & disc degeneration

<u>Model 2 (only EPI available):</u>

PEN/EPI & D-Asp/EPI & sex & disc degeneration

Model 3 (IVD and EPI available):

PEN/IVD & D-Asp/IVD & PEN/EPI & D-Asp/EPI & sex & disc degeneration

Results

Sample Characteristics

The enzymatic purification of the samples according to Sivan et al. [53] destroyed most proteins but collagen and elastin. As a consequence, after the purification step the IVD samples consisted mainly of collagen, the EPI samples mainly of elastin (and collagen), as demonstrated by the results of amino acid analysis (Figures 1a and 1b).

PEN and D-Asp in IVD:

Both PEN and D-Asp accumulate with increasing age in the IVD samples ($r = 0.84$ and $r = 0.73$, respectively; Figures 2a and 2b). However, scattering of the values for both parameters increases substantially with age.

By marking of the samples with highly degenerative changes (n = 41) in Figures 2a and 2b it becomes obvious that these samples contribute substantially to the increased scattering of data in higher ages. On the other hand, the relationship between PEN and D-Asp and age seems to remain close even in old age as long the integrity of the tissue is preserved (no degeneration); due to the low number of these cases (no degeneration in old ages) a further statistical analysis was not performed.

As far as assessable, a diabetic metabolic disorder had been diagnosed in 9 cases (there was no information regarding the quality of therapy), these cases are indicated in Figure 3. Only 1 of these cases exhibited strikingly high PEN values. At the same time, however, there were highly degenerative changes in the tissue of this case.

PEN and D-Asp seem to be robust regarding postmortem influencing factors. In Figure 4, the results for cases of advanced postmortem putrefaction (n = 9), burned bodies (n = 3) and exhumation (n = 3) are presented; the corresponding values do not deviate noticeably from the other values.

PEN and D-Asp in EPI:

The results for the EPI samples were very similar to those for the IVD samples: PEN and D-Asp do also accumulate in EPI with increasing age ($r = 0.86$ and $r = 0.81$, respectively; Figures 5a and 5b), and scattering of PEN and D-Asp-values increased with age.

Again, there was no clear indication for higher PEN values in diabetic cases (n = 9, Figure 6), and PEN and D-Asp seem to be robust regarding postmortem influencing factors (Figures 7a and 7b)

Multivariate models enable more precise age estimates

Table 1 presents the MAEs of age estimation based on univariate predictions (based on either IVD/PEN, either IVD/D-Asp, either EPI/PEN or EPI/D-Asp) and on multivariate predictions, including either 2 parameters of one tissue (model 1 (IVD) and model 2 (EPI)) or 4 parameters of both tissues (model 3). Age estimation based on the univariate models resulted in MAEs between 7.5 and 11.0 years, age estimation using the multivariate models resulted in MAEs of 6.3 years (model 1, IVD), 5.5 years (model 2, EPI) and 4.0 years (model 3, IVD & EPI). Model 3 showed the strongest correlation between chronological age and estimated age ($r=0.95$; Figure 8). All models achieved considerably better results in cases with an age younger than 62 years than in cases older than 62 (age under 62 years: MAEs of 4.7, 5.0 and 3.1 years in models 1, 2 and 3, respectively; age above 62 years: MAEs of 8.0, 6.0 and 4.9 years, respectively).

Discussion

During the last decades it became evident, that several molecular modifications accumulate in the human organism with increasing age. Some of these „molecular clocks“ in DNA and in proteins open up promising approaches for the development of methods for forensic age estimation ([2, 4, 10–12]). Above all, the published data for approaches based on methylation of DNA (mDNA) and accumulation of D-aspartic acid (D-Asp) indicate a high potential of these methods (for review see [2, 6, 8, 12, 14]). However, one should not overlook the fact that even these approaches determine the chronological age only indirectly, namely by an epigenetic clock or a protein clock. The influence of diverse extrinsic and intrinsic factors on the individual aging process and thus also on these clocks should not be underestimated. Accordingly, one of the main problems of all molecular methods of age estimation is an increasing scattering of data with increasing age. It has already been proposed that a

combined use of different biomarkers of aging may address this problem and result in a higher accuracy of age estimation especially in older ages [59] via a better addressing of the complex processes of aging.

We tested this hypothesis by analysis of two molecular clocks (accumulation of D-Asp, accumulation of pentosidine (PEN)) in two different tissues (annulus fibrosus of intervertebral discs (IVD) and elastic cartilage of epiglottis (EPI)) and the development of multivariate models for age estimation based on the collected data.

PEN and D-Asp in collagen from the annulus fibrosus of intervertebral discs (IVD)

An age dependent accumulation of D-Asp and PEN in IVD has already been described [41, 53, 60]. For the first time, we analyzed the two parameters in combination. Furthermore, we analyzed relatively well defined protein fractions after purification of collagen (Figure 1a) instead of total protein [41, 60].

Basically, the PEN and D-Asp data confirm that collagen in the annulus fibrosus of IVD is a permanent protein that accumulates modifications during aging. The increasing scattering of data with age mirrors the individuality of the aging process in this tissue, just as the different extents of morphological degenerative changes do. Degeneration can result in tissue remodeling with breakdown of molecules as well as with a synthesis of new collagen which means an introduction of “young” protein into “old” tissue. Therefore, it is not surprising that even after purification of collagen, a relevant increase in the scattering of data in older ages and in cases with degenerative changes was observed (Figures 2a and 2b).

Theoretically, the accumulation of PEN may be influenced by a diabetic metabolic disorder due the exposition of proteins to higher concentrations of glucose [32, 33, 45, 61]. There was no clear indication for a relevant deviation of the PEN values in diabetic individuals in the collective examined (Figure 3). However, we analyzed only a small number of samples from individuals with diabetes mellitus. Since the formation of PEN is a very slow process [33, 44, 62, 63], it can be assumed that a relevant influence of diabetes mellitus on the overall amount of PEN will only be observed, if metabolic conditions with high levels of glucose persist for a long time. This assumption has to be tested by analysis of a sufficiently large set of cases with known medical history.

Fortunately, PEN and D-Asp seem to be relatively robust against special postmortem conditions (Figures 4a, 4b). However, also this preliminary finding has to be checked by analysis of a large set of samples comprising cases with different postmortem conditions.

The MAEs of 9.5 years (IVD/PEN) and 11.0 years (IVD/D-Asp) for age estimation via univariate predictions (Table 1) are quite high. Theoretically, much better results could be achieved, if samples with degenerative changes are excluded, as proposed by Ritz and Schütz [60] and Pilin et al. [41]. However, this would limit the applicability of the method considerably, since advanced degeneration is a very common finding in older ages.

PEN and D-Asp in elastin/collagen from the elastic cartilage from the epiglottis (EPI)

Since an age-dependent accumulation of D-Asp in the elastic cartilage from the epiglottis has already been described [23], it could be anticipated that this tissue contains permanent proteins that might also accumulate PEN. Indeed, the collagen/elastin protein fraction (Figure 1b) of the tissue exhibited an age-dependent accumulation of PEN (Figure 5a), and the accumulation of D-Asp in EPI was confirmed (Figure 5b). Overall, the results for the EPI samples were very similar to those for the IVD samples and can be interpreted in the same way.

Age estimation based on only one EPI parameter was quite inaccurate as indicated by MAEs of 7.5 years (EPI/PEN) and of 8.6 years (EPI/D-Asp) (Table 1).

Age estimation based on a multivariate approach using AAR and PEN in IVD and EPI

The relatively high MAEs for age estimates based on single parameters reflect the complexity of the aging process in the different biological subsystems.

The inclusion of more than one parameter in the models for age estimation led to considerably lower MAEs (Table 1). While MAEs were 7.5 – 11.0 years in the univariate models, the multivariate model that included all parameters resulted in the lowest MAE of 4.0 years. This is a very promising result – especially with regard to the underlying single data sets for PEN and D-Asp in IVD and EPI with a quite huge scattering of data at least in higher ages (Figures 2 and 5).

The question arises, if this “pure protein approach” promises better results in postmortem age estimation than methods based on DNA methylation. This question is difficult to answer for several reasons. Many authors used MAEs and correlation coefficients as measures for the accuracy of age estimation. MAE and correlation coefficient for our multivariate protein model (4 years and 0.95, respectively) are very close to those of the best models for age estimation based on DNA methylation (ca. 3-5 years, [14, 64–66], for review see [7, 10, 11]). However, neither MAEs nor correlation coefficients do permit any exact statements regarding

the accuracy in individual cases [7]. Apart from that, situations in the context of the identification of an unknown deceased are highly variable (e.g. highly decomposed bodies, skeletons, body parts). The answer to the question, which approach is the best for a single case depends (1.) on the tissues available for the analysis, and (2.) on the applicability of a method under the conditions of advanced putrefaction. The data presented here prove the applicability of the multivariate protein approach to two tissues that are usually well preserved for a relatively long time during postmortem decomposition; moreover, PEN and D-Asp seem to be robust under diverse postmortem conditions. Up to now, there are only a few studies on the applicability of the DNA methylation approach to postmortem degraded tissues [67, 68]. It is the task of future research to further examine the repertoire of methods in terms of their applicability and accuracy under the diverse case scenarios of forensic practice.

Conclusion and vision

The presented results confirm the conclusion that multivariate models, using appropriate parameters, can be used to develop new age estimation methods that are more accurate and broader in scope than univariate methods. Some groups have already introduced multivariate models into the development of methods for mDNA based age estimation [59, 64, 66, 69]. These models include the data for different mDNA markers but remain in one biological level. Moreover, there are already a few promising multivariate approaches that combine parameters concerning different biological levels, e.g. mDNA and morphology [70], or mDNA and signal-joint T-cell receptor excision circles (sjTRECs) [59]. Though the approach of Cho [59] is not applicable to highly decomposed bodies (since it is based on the analysis of blood), it is very interesting, as it shows that the combination of an mDNA model with another biomarker of aging (sjTRECs) may address a major problem of all methods of age estimation, namely the higher scattering of data with increasing age.

In conclusion, the combined analysis of different molecular clocks in different tissues is a very useful approach for several reasons. It may improve predication accuracy especially in older ages [59], it offers new possibilities for age estimation in cases with advanced postmortem putrefaction and decomposition, and - from the perspective of basic sciences - it may contribute to a better understanding of complex aging processes.

Our vision is the development of a system of multivariate models for age estimation that can be used in the multiple scenarios of forensic practice (e.g. advanced putrefaction, skeletons, body parts). In a next step, we will add the epigenetic clock (mDNA) to our protein clocks (PEN, D-Asp) and expand our set of tissues.

Ethical approval

All procedures performed in studies involving human tissue were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards (approved by Ethics Committee at the Medical Faculty of Heinrich-Heine University: 6191R, 3667). This article does not contain any studies with animals performed by any of the authors.

Conflict of interest

The authors declare that they have no conflict of interest.

Fig. 1 Amino acid composition of the purified samples of intervertebral disc (IVD) (**1a**) and the epiglottis (EPI) samples (**1b**), as compared to the amino acid composition of standard samples (human collagen type 1, bovine elastin)

ASX = asparagine and aspartic acid, THR = threonine, SER = serine, GLX = glutamine and glutamic acid, Pro = proline, HYP = hydroxyproline, GLY = glycine, ALA = alanine, VAL = valine, MET = methionine, ILE = isoleucine, LEU = leucine, TYR = tyrosine, PHE = phenylalanine, TRP = tryptophan, ARG = arginine

Fig. 2 Intervertebral disc (IVD) samples (n=95): Accumulation of pentosidine (PEN) (**2a**) and D-aspartic acid (D-Asp, as ln((1+D/L)/(1-D/L))) (**2b**) in relation to age (PEN: r=0.84; D-Asp: r=0.73), samples with highly degenerative changes are marked (n = 41)

D = D-aspartic acid, L = L-aspartic acid

Fig. 3 Intervertebral disc (IVD) samples: Same pentosidine (PEN) data as in Figure **2a** (n = 95, r = 0.84), here with indication of samples from individuals with diabetic metabolic disorder (n = 9)

Fig. 4 Intervertebral disc (IVD) samples: Same data as in Figure 2a and 2b (**4a:** data for PEN, **4b:** data for D-Asp (as ln((1+D/L)/(1-D/L)))**,** data for samples from cases with advanced postmortem putrefaction (n = 9), of burned bodies (n = 3) and of cases with exhumation (n = 3) are added and indicated

D = D-aspartic acid, L = L-aspartic acid

Fig. 5 Epiglottis (EPI) samples (n = 95): Accumulation of pentosidine (PEN; r = 0.86) (**5a**) and D-aspartic acid (D-Asp, as ln((1+D/L)/(1-D/L)); r = 0.81) (**5b**) in relation to age

D = D-aspartic acid, L = L-aspartic acid

Fig. 6 Epiglottis (EPI) samples: Same pentosidine (PEN) data as in Figure **5a** (n = 95, r = 0.86), here with marking of samples from individuals with diabetic metabolic disorder (n = 9)

Fig. 7 Epiglottis (EPI) samples: Same data as in Figure 5a and 5b (**7a:** data for PEN, **7b:** data for D-Asp (as ln((1+D/L)/(1-D/L)))**,** data for samples from cases with advanced postmortem putrefaction (n = 8), of burned bodies (n = 3) and of cases with exhumation (n = 3) are added and indicated

D = D-aspartic acid, L = L-aspartic acid

Fig. 8 Multivariate model 3: Age estimates in relationship to the corresponding chronological ages (n = 95; mean absolute error = 4.0 (3.1 years for ages under 62 years, 4.9 years for ages above 62 years); r = 0.95)

GPR = gaussian progress regression

Table 1

Quality of age estimation by univariate and multivariate predictions

IVD= samples from intervertebral discs, EPI = samples from epiglottis, PEN = accumulation of pentosidine, D-Asp = accumulation of D-aspartic acid, r = coefficient of correlation (Spearman), MAE = Mean absolute error)

Literatur

1. Ritz-Timme S, Cattaneo C, Collins MJ, Waite ER, Schütz HW, Kaatsch HJ, Borrman HI (2000) Age estimation: The state of the art in relation to the specific demands of forensic practise. International journal of legal medicine 113(3):129–136
2. Meissner C, Ritz-Timme S (2010) Molecular pathology and age estimation. Forensic science international 203(1-3):34–43. doi:10.1016/j.forsciint.2010.07.010
3. Ritz-Timme S, Collins MJ (2002) Racemization of aspartic acid in human proteins. Ageing research reviews 1(1):43–59
4. Zapico SC, Ubelaker DH (2013) Applications of physiological bases of ageing to forensic sciences. Estimation of age-at-death. Ageing research reviews 12(2):605–617. doi:10.1016/j.arr.2013.02.002
5. Ritz-Timme S (1999) Lebensaltersbestimmung aufgrund des Razemisierungsgrades von Asparaginsäure: Grundlagen, Methodik, Möglichkeiten, Grenzen, Anwendungsbereiche ; mit 6 Tabellen. Arbeitsmethoden der medizinischen und naturwissenschaftlichen Kriminalistik, Bd 23. Schmidt-Römhild, Lübeck
6. Freire-Aradas A, Phillips C, Lareu MV (2017) Forensic individual age estimation with DNA: From initial approaches to methylation tests. Forensic science review 29(2):121–144
7. Ritz-Timme S, Schneider PM, Mahlke NS, Koop BE, Eickhoff SB (2018) Altersschätzung auf Basis der DNA-Methylierung. Rechtsmedizin 28(3):202–207. doi:10.1007/s00194-018-0249-3
8. Lee HY, Lee SD, Shin K-J (2016) Forensic DNA methylation profiling from evidence material for investigative leads. BMB reports 49(7):359–369
9. Goel N, Karir P, Garg VK (2017) Role of DNA methylation in human age prediction. Mechanisms of ageing and development 166:33–41. doi:10.1016/j.mad.2017.08.012
10. Vidaki A, Kayser M (2018) Recent progress, methods and perspectives in forensic epigenetics. Forensic science international. Genetics 37:180–195. doi:10.1016/j.fsigen.2018.08.008
11. Parson W (2018) Age Estimation with DNA: From Forensic DNA Fingerprinting to Forensic (Epi)Genomics: A Mini-Review. Gerontology 64(4):326–332. doi:10.1159/000486239
12. Zapico SC (2017) Mechanisms Linking Aging, Diseases and Biological Age Estimation. 23-Epigenetics. CRC Press, Portland
13. Horvath S (2013) DNA methylation age of human tissues and cell types. Genome Biol 14(10):R115. doi:10.1186/gb-2013-14-10-r115
14. Jung S-E, Shin K-J, Lee HY (2017) DNA methylation-based age prediction from various tissues and body fluids. BMB reports 50(11):546–553
15. Jones MJ, Goodman SJ, Kobor MS (2015) DNA methylation and healthy human aging. Aging Cell 14(6):924–932. doi:10.1111/acel.12349
16. Spólnicka M, Pośpiech E, Adamczyk JG, Freire-Aradas A, Pepłońska B, Zbieć-Piekarska R, Makowska Ż, Pięta A, Lareu MV, Phillips C, Płoski R, Żekanowski C, Branicki W (2018) Modified aging of elite athletes revealed by analysis of epigenetic age markers. Aging 10(2):241–252. doi:10.18632/aging.101385

17. Gao X, Zhang Y, Breitling LP, Brenner H (2016) Relationship of tobacco smoking and smoking-related DNA methylation with epigenetic age acceleration. Oncotarget 7(30):46878–46889. doi:10.18632/oncotarget.9795

18. Kandi V, Vadakedath S (2015) Effect of DNA Methylation in Various Diseases and the Probable Protective Role of Nutrition: A Mini-Review. Cureus 7(8):e309. doi:10.7759/cureus.309

19. Stephenson RC, Clarke S (1989) Succinimide formation from aspartyl and asparaginyl peptides as a model for the spontaneous degradation of proteins. The Journal of biological chemistry 264(11):6164–6170

20. Geiger T, Clarke S (1987) Deamidation, isomerization, and racemization at asparaginyl and aspartyl residues in peptides. Succinimide-linked reactions that contribute to protein degradation. The Journal of biological chemistry 262(2):785–794

21. Dobberstein RC, Tung S-M, Ritz-Timme S (2010) Aspartic acid racemisation in purified elastin from arteries as basis for age estimation. International journal of legal medicine 124(4):269–275. doi:10.1007/s00414-009-0392-1

22. Klumb K, Matzenauer C, Reckert A, Lehmann K, Ritz-Timme S (2016) Age estimation based on aspartic acid racemization in human sclera. International journal of legal medicine 130(1):207–211. doi:10.1007/s00414-015-1255-6

23. Matzenauer C, Reckert A, Ritz-Timme S (2014) Estimation of age at death based on aspartic acid racemization in elastic cartilage of the epiglottis. International journal of legal medicine 128(6):995–1000. doi:10.1007/s00414-013-0940-6

24. Ohtani S, Yamamoto K (1991) Age estimation using the racemization of amino acid in human dentin. Journal of forensic sciences 36(3):792–800

25. Ritz S, Schütz HW, Schwarzer B (1990) The extent of aspartic acid racemization in dentin: A possible method for a more accurate determination of age at death? Zeitschrift fur Rechtsmedizin. Journal of legal medicine 103(6):457–462

26. Ritz S, Schütz HW, Peper C (1993) Postmortem estimation of age at death based on aspartic acid racemization in dentin: Its applicability for root dentin. International journal of legal medicine 105(5):289–293

27. Ohtani S, Yamamoto T (2010) Age estimation by amino acid racemization in human teeth. Journal of forensic sciences 55(6):1630–1633. doi:10.1111/j.1556-4029.2010.01472.x

28. Chen S, Lv Y, Wang D, Yu X (2016) Aspartic acid racemization in dentin of the third molar for age estimation of the Chaoshan population in South China. Forensic science international 266:234–238. doi:10.1016/j.forsciint.2016.06.010

29. Elfawal MA, Alqattan SI, Ghallab NA (2015) Racemization of aspartic acid in root dentin as a tool for age estimation in a Kuwaiti population. Medicine, science, and the law 55(1):22–29. doi:10.1177/0025802414524383

30. Wochna K, Bonikowski R, Śmigielski J, Berent J (2018) Aspartic acid racemization of root dentin used for dental age estimation in a Polish population sample. Forensic science, medicine, and pathology 14(3):285–294. doi:10.1007/s12024-018-9984-8

31. Ritz-Timme S, Rochholz G, Stammert R, Ritz H-J (2002) Biochemische Altersschätzung Zur Frage genetischer und soziokultureller (ethnischer) Einflüsse auf die Razemisierung von Asparaginsäure in Dentin. Rechtsmedizin 12(4):203–206. doi:10.1007/s00194-002-0152-8

32. Ulrich P, Cerami A (2001) Protein glycation, diabetes, and aging. Recent progress in hormone research 56:1–21

33. Singh R, Barden A, Mori T, Beilin L (2001) Advanced glycation end-products: A review. Diabetologia 44(2):129–146. doi:10.1007/s001250051591

34. Goldberg T, Cai W, Peppa M, Dardaine V, Baliga BS, Uribarri J, Vlassara H (2004) Advanced glycoxidation end products in commonly consumed foods. Journal of the American Dietetic Association 104(8):1287–1291. doi:10.1016/j.jada.2004.05.214

35. Nass N, Bartling B, Navarrete Santos A, Scheubel RJ, Börgermann J, Silber RE, Simm A (2007) Advanced glycation end products, diabetes and ageing. Zeitschrift fur Gerontologie und Geriatrie 40(5):349–356. doi:10.1007/s00391-007-0484-9

36. Greis F, Reckert A, Fischer K, Ritz-Timme S (2018) Analysis of advanced glycation end products (AGEs) in dentine: Useful for age estimation? International journal of legal medicine 132(3):799–805. doi:10.1007/s00414-017-1671-x

37. Odetti P, Rossi S, Monacelli F, Poggi A, Cirnigliaro M, Federici M, Federici A (2005) Advanced glycation end products and bone loss during aging. Annals of the New York Academy of Sciences 1043:710–717. doi:10.1196/annals.1333.082

38. Pokharna HK, Phillips FM (1998) Collagen crosslinks in human lumbar intervertebral disc aging. Spine 23(15):1645–1648

39. Verzijl N, DeGroot J, Oldehinkel E, Bank RA, Thorpe SR, Baynes JW, Bayliss MT, Bijlsma JW, Lafeber FP, Tekoppele JM (2000) Age-related accumulation of Maillard reaction products in human articular cartilage collagen. The Biochemical journal 350 Pt 2:381–387

40. Ramalho JS, Marques C, Pereira PC, Mota MC (1996) Role of glycation in human lens protein structure change. European journal of ophthalmology 6(2):155–161

41. Pillin A, Pudil F, Bencko V, Bezdícková D (2007) Contents of pentosidine in the tissue of the intervertebral disc as an indicator of the human age. Soudni lekarstvi 52(4):60–64

42. Dyer DG, Dunn JA, Thorpe SR, Bailie KE, Lyons TJ, McCance DR, Baynes JW (1993) Accumulation of Maillard reaction products in skin collagen in diabetes and aging. The Journal of clinical investigation 91(6):2463–2469. doi:10.1172/JCI116481

43. Valenzuela A, Guerra-Hernández E, Rufián-Henares JÁ, Márquez-Ruiz AB, Hougen HP, García-Villanova B (2018) Differences in non-enzymatic glycation products in human dentine and clavicle: changes with aging. International journal of legal medicine 132(6):1749–1758. doi:10.1007/s00414-018-1908-3

44. Li H, Yu S-J (2018) Review of pentosidine and pyrraline in food and chemical models: formation, potential risks and determination. J Sci Food Agric 98(9):3225–3233. doi:10.1002/jsfa.8853

45. Stitt AW, Jenkins AJ, Cooper ME (2002) Advanced glycation end products and diabetic complications. Expert Opin Investig Drugs 11(9):1205–1223. doi:10.1517/13543784.11.9.1205

46. Ritz S, Turzynski A, Schütz HW (1994) Estimation of age at death based on aspartic acid racemization in noncollagenous bone proteins. Forensic science international 69(2):149–159

47. Ritz S, Turzynski A, Schütz HW, Hollmann A, Rochholz G (1996) Identification of osteocalcin as a permanent aging constituent of the bone matrix: Basis for an accurate age at death determination. Forensic science international 77(1-2):13–26

48. Ritz-Timme S, Laumeier I, Collins M (2003) Age estimation based on aspartic acid racemization in elastin from the yellow ligaments. International journal of legal medicine 117(2):96–101. doi:10.1007/s00414-002-0355-2

49. Monum T, Jaikang C, Sinthubua A, Prasitwattanaseree S, Mahakkanukrauh P (2017) Age estimation using aspartic amino acid racemization from a femur. Australian Journal of Forensic Sciences 116:1–9. doi:10.1080/00450618.2017.1391330

50. Ohtani S, Matsushima Y, Kobayashi Y, Kishi K (1998) Evaluation of aspartic acid racemization ratios in the human femur for age estimation. Journal of forensic sciences 43(5):949–953

51. Ohtani S, Yamamoto T, Abe I, Kinoshita Y (2007) Age-dependent changes in the racemisation ratio of aspartic acid in human alveolar bone. Archives of oral biology 52(3):233–236. doi:10.1016/j.archoralbio.2006.08.011

52. Tiplamaz S, Gören MZ, Yurtsever NT (2018) Estimation of Chronological Age from Postmortem Tissues Based on Amino Acid Racemization. Journal of forensic sciences 63(5):1533–1538. doi:10.1111/1556-4029.13737

53. Sivan SS, Tsitron E, Wachtel E, Roughley P, Sakkee N, van der Ham F, Degroot J, Maroudas A (2006) Age-related accumulation of pentosidine in aggrecan and collagen from normal and degenerate human intervertebral discs. The Biochemical journal 399(1):29–35. doi:10.1042/BJ20060579.

54. Schmidt MB, Mow VC, Chun LE, Eyre DR (1990) Effects of proteoglycan extraction on the tensile behavior of articular cartilage. J Orthop Res 8(3):353–363. doi:10.1002/jor.1100080307

55. Heems D, Luck G, Fraudeau C, Vérette E (1998) Fully automated precolumn derivatization, on-line dialysis and high-performance liquid chromatographic analysis of amino acids in food, beverages and feedstuff. Journal of Chromatography A 798(1-2):9–17. doi:10.1016/S0021-9673(97)01007-8

56. Kaufman DS, Manley WF (1998) A new procedure for determining dl amino acid ratios in fossils using reverse phase liquid chromatography. Quaternary Science Reviews 17(11):987–1000. doi:10.1016/S0277-3791(97)00086-3

57. Rokach L (2010) Ensemble-based classifiers. Artif Intell Rev 33(1-2):1–39. doi:10.1007/s10462-009-9124-7

58. Rasmussen CE, Williams CKI (2006) Gaussian processes for machine learning. Adaptive computation and machine learning. MIT Press, Cambridge, Mass

59. Cho S, Jung S-E, Hong SR, Lee EH, Lee JH, Lee SD, Lee HY (2017) Independent validation of DNA-based approaches for age prediction in blood. Forensic science international. Genetics 29:250–256. doi:10.1016/j.fsigen.2017.04.020

60. Ritz S, Schütz HW (1993) Aspartic acid racemization in intervertebral discs as an aid to postmortem estimation of age at death. Journal of forensic sciences 38(3):633–640

61. Sell DR, Nagaraj RH, Grandhee SK, Odetti P, Lapolla A, Fogarty J, Monnier VM (1991) Pentosidine: a molecular marker for the cumulative damage to proteins in diabetes, aging, and uremia. Diabetes Metab Rev 7(4):239–251

62. Brownlee M (1995) Advanced protein glycosylation in diabetes and aging. Annu Rev Med 46:223–234. doi:10.1146/annurev.med.46.1.223

63. Semba RD, Nicklett EJ, Ferrucci L (2010) Does Accumulation of Advanced Glycation End Products Contribute to the Aging Phenotype? J Gerontol A Biol Sci Med Sci 65A(9):963–975. doi:10.1093/gerona/glq074

64. Aliferi A, Ballard D, Gallidabino MD, Thurtle H, Barron L, Syndercombe Court D (2018) DNA methylation-based age prediction using massively parallel sequencing data and multiple machine learning models. Forensic science international. Genetics 37:215–226. doi:10.1016/j.fsigen.2018.09.003

65. Jung S-E, Lim SM, Hong SR, Lee EH, Shin K-J, Lee HY (2019) DNA methylation of the ELOVL2, FHL2, KLF14, C1orf132/MIR29B2C, and TRIM59 genes for age prediction from blood, saliva, and buccal swab samples. Forensic science international. Genetics 38:1–8. doi:10.1016/j.fsigen.2018.09.010

66. Naue J, Hoefsloot HCJ, Mook ORF, Rijlaarsdam-Hoekstra L, van der Zwalm MCH, Henneman P, Kloosterman AD, Verschure PJ (2017) Chronological age prediction based on DNA methylation: Massive parallel sequencing and random forest regression. Forensic science international. Genetics 31:19–28. doi:10.1016/j.fsigen.2017.07.015

67. Rhein M, Hagemeier L, Klintschar M, Muschler M, Bleich S, Frieling H (2015) DNA methylation results depend on DNA integrity—role of post mortem interval. Front Genet 6. doi:10.3389/fgene.2015.00182

68. Jarmasz JS, Stirton H, Davie JR, Del Bigio MR (2019) DNA methylation and histone post-translational modification stability in post-mortem brain tissue. Clin Epigenetics 11(1):5. doi:10.1186/s13148-018-0596-7

69. Vidaki A, Ballard D, Aliferi A, Miller TH, Barron LP, Syndercombe Court D (2017) DNA methylation-based forensic age prediction using artificial neural networks and next generation sequencing. Forensic science international. Genetics 28:225–236. doi:10.1016/j.fsigen.2017.02.009

70. Shi L, Jiang F, Ouyang F, Zhang J, Wang Z, Shen X (2018) DNA methylation markers in combination with skeletal and dental ages to improve age estimation in children. Forensic science international. Genetics 33:1–9. doi:10.1016/j.fsigen.2017.11.005