Journal of Forensic and Legal Medicine 36 (2015) 177–184

Contents lists available at ScienceDirect

Journal of Forensic and Legal Medicine

journal homepage: www.elsevier.com/locate/jflm

Original communication

Dental Age Estimation (DAE): Data management for tooth development stages including the third molar. Appropriate censoring of Stage H, the final stage of tooth development

Graham J. Roberts [a, *], Fraser McDonald [a], Manoharan Andiappan [b], Victoria S. Lucas [a]

[a] Department of Orthodontics, King's College London Dental Institute, Floor 25, Tower Wing, Guy's Campus, London SE1 9RT, UK
[b] Biostatistics and Research Methods Centre, Division of Patient and Population Health, Denmark Hill Campus, King's College London Dental Institute, Denmark Hill, London SE5 9RW, UK

ARTICLE INFO

Article history:
Received 30 April 2015
Received in revised form
1 July 2015
Accepted 27 August 2015
Available online 8 September 2015

Keywords:
Third molars
Dental age estimation
Censored data
Three standard deviations

ABSTRACT

Introduction: The final stage of dental development of third molars is usually helpful to indicate whether or not a subject is aged over 18 years. A complexity is that the final stage of development is unlimited in its upper border. Investigators usually select an inappropriate upper age limit or censor point for this tooth development stage.
Materials and methods: The literature was searched for appropriate data sets for dental age estimation and those that provided the count (n), the mean ($\bar{x}$), and the standard deviation (sd) for each of the tooth development stages. The Demirjian G and Demirjian H were used for this study. Upper and lower limits of the Stage G and Stage H data were calculated limiting the data to plus or minus three standard deviations from the mean. The upper border of Stage H was limited by appropriate censoring at the maximum value for Stage G.
Results: The maximum age at attainment from published data, for Stage H, ranged from 22.60 years to 34.50 years. These data were explored to demonstrate how censoring provides an estimate for the correct maximum age for the final stage of Stage H as 21.64 years for UK Caucasians.
Conclusion: This study shows that confining the data array of individual tooth developments stages to ± 3sd provides a reliable and logical way of censoring the data for tooth development stages with a Normal distribution of data. For Stage H this is inappropriate as it is unbounded in its upper limit. The use of a censored data array for Stage H using Percentile values is appropriate. This increases the reliability of using third molar Stage H *alone* to determine whether or not an individual is over 18 years old. For Stage H, individual ancestral groups should be censored using the same technique.

1. Introduction

The concept of dental age estimation (DAE) is not new. The assessment of tooth eruption as a method of estimating age was introduced in the 19th Century[1] and is still occasionally used as a guide to a child's age. The availability of radiographs has enabled visualization of the mineralization of developing teeth which are then classified into identifiable stages from which dental maturity can be determined.[2] The close link between dental maturity and chronological age (CA) justifies the use of dental age (DA) as a surrogate for chronological age.[3] At the 10 year threshold, DA is accurate to within 3 weeks of CA.[4] Despite several publications relating to the 18 year threshold, there is little known about the reliability of estimates made at or near this age.[5–7] A systematic and highly informative attempt has been made to identify and explore these difficulties in the series of papers from the University of Leuven.[8]

The most common way to determine whether or not a subject is over 18 years is to calculate the probability => 18.00 using the mean ($\bar{x}$) and the standard deviation (sd) of the data array for the final stage of tooth development. A typical outcome using this approach is that a subject exhibiting Stage H is 0.92 (92%) probability to be over 18 years.[9] The underlying assumption is that the data for Stage H is Normally distributed. It has been suggested that an approach using percentile data would, on theoretical grounds,

* Corresponding author. Tel.: +44 01732 463636.
E-mail addresses: graham.j.roberts@kcl.ac.uk (G.J. Roberts), fraser.mcdonald@kcl.ac.uk (F. McDonald), manoharan.1.andiappan@kcl.ac.uk (M. Andiappan), victoria.s.lucas@kcl.ac.uk (V.S. Lucas).

http://dx.doi.org/10.1016/j.jflm.2015.08.013
1752-928X/

lead to a conclusion more in line with the distribution of the data.[10] The research team in Germany use both Normal distribution data and some percentile data to display the results for stages of third molar development.[11] The Swiss-German research team proposed that because subjects exhibiting Stage H *may* be below 18 years of age, it is inappropriate to use the presence of Stage H *alone* as an indication that the subject is over 18 years old.[12]

An important requirement of statistical analysis is that the data used are effectively managed.[13] This is to ensure that inappropriate values are not included in the data array. This might be because of a grossly incorrect result such as an age value when the data of birth and data of assessment are reversed, and incorrect assessments that give unjustifiable outliers. These problems are managed by so-called 'cleaning' of the data by visual inspection when checking the data.[14]

Common practice in statistics is to use only the central 95% of the data array, the so called 95% Confidence Interval (CI).[15] It is common to use this 95% confidence interval, sometimes called the reference range, as the data array for a given variable to enable clinicians to determine what is Normal and non-Normal. With regards to DAE, lawyers question the appropriateness of this as it excludes 5% of the data or the equivalent of 1 of 20 subjects.

Given this concern the research presented here is aimed at encompassing as much of the data as possible by using plus or minus three standard deviations (±3sd) as this incorporates 99.9974% of the sample data, and by inference a similar percentage of the population. This is regarded as an age range for subjects of unknown age whose future will depend upon a fair and robust assessment of the likelihood of being within or without this stated age range.

A further problem is that of Stage H, the final stage of development which is unbounded in its upper limit. This raises the issue as to how different investigators have set the upper age limit for the samples investigated. Publications range from 21 years,[16] 22 years,[17] 23 years,[18] 24 years,[19] 26 years,[20] 26.9 years,[21] and 33.9 years.[22]

Clearly this wide range of age limits for the upper boundary of subjects in studies on DAE, spanning almost 13 years, raises the question as to what should be the upper boundary for Stage H of third molars and how this boundary is identified.

Perusal of the literature reveals that investigators do not indicate how the data management has been carried out to ensure that there are robust data sets using the 8 stage system[2] for each of the 250 plus data arrays for each Tooth Development Stage (TDS).

A further problem is that it is difficult to compare published data sets as the number of Tooth Development Stages (TDS) used for assessment of dental development varies from 4 stages,[23] to 21 stages.[24] (Mahtab M. personal communication). The most frequently used number of Tooth Development Stages used to estimate Age at Attainment (AaA) of each stage is the 8 stages described by the Anglo-Canadian research team.[2] Because of its widespread use, the 8 stage system of TDS has been selected for this review. It is important to be aware that the present paper utilizes only the anatomical descriptions of the 8 tooth development stages but does not include the system of mathematical integration to estimate dental maturity and from that dental age.[2] In addition, the 1973 system of dental age assessment was limited to teeth in the mandible and excluded third molars. This limitation was superseded when an American research team applied the 8 stage system to a study of the development of all four third molars.[5] This forms the basis for the system of age estimation at the 18 year threshold currently in use in the United States of America,[25] and also by many investigators in Europe including the DARLInG team at King's College London. The Tooth Development Stages defined in 1973 are referred to by the letters A, B, C, D, E, F, G, and H. Detailed descriptions of each stage are available in the original publication[2] and more recently in a paper at the 16 year threshold.[26]

The presentation of data follows two formats:

1. Normal distribution summary data
2. Percentile data.

The final stage of tooth development is unbounded in its upper age limit. With regard to third molars this means that an enquiry about estimating age using a radiograph showing fully developed roots of a lower left third molar from a male Caucasian who is aged 25 years would return the same age and probability estimate as an enquiry from someone aged 15 years.

The difficulty with the lack of a natural upper limit to the data array for the Age at Attainment (AaA) of Stage H is usually overcome by not using Stage H for age estimation in children who are still developing.[27] This exclusion of Stage H is not possible when estimating age at the 18 year threshold as the third molar Stage H is the principle dental biological marker that can be used for age prediction for late adolescence and emerging adults. This was the genesis of the seminal paper on third molars published by the research team in the USA.[5] In reviewing the articles on this problem, it is clear that almost all the papers published do not indicate in sufficient detail the approach used for the curtailment of the upper limit of the AaA for Stage H. The problem has been touched upon in a paper from South Korea where the authors used different age ranges to estimate the mean value of Stage H.[28] For example, focusing on the Lower Left Third Molar Stage H for males, the AaA for the LL8Hm is 21.6 years for the 14–24 year band, 22.1 years for the 14–25 year age band, and 22.5 years for the 12–26 year year age band.[28] This raises the issue of the amount by which the AaA for Stage H is inflated by using an upper age cut-off or censor point of 26 years when compared to 21 years.

This problem has been explored and led to the concept of 'appropriate censoring'.[29] It is clear from published papers that the issue of censoring the data for the final stage of tooth development has been ignored or overlooked by the majority of investigators in the field of DAE.

A further problem is the way in which the uncertainty of the point estimate is provided. It is conventional to use the 95% confidence interval. The problem with this is that it excludes 1 in 20 subjects, a figure that is not acceptable to lawyers. The use of a 100% confidence interval expressed as a reference range has much to commend it as it places any estimates of age beyond any reasonable doubt. This is important for criminal procedures but for civil procedures it is necessary only to provide the central 50% of the data. This is usually presented as the Interquartile Width (IQW) (Altman 1991).[30]

The purpose of this paper is threefold:

1. To review data management procedures to determine the suitability of a ± 3sd constraint to the lower and upper limits of the data array for each AaA for all the tooth development stages.
2. To describe a simple system of censoring to provide a realistic age to the upper boundary for Stage H.
3. To review extant data sets where the 8 stage system of Demirjian[2] has been used to assess the impact of inappropriate censoring of Stage H on the estimation of the lower and upper limits of AaA for the Lower Left Third Molar (encompassing 100% of the data array), the estimation of the middle 50% i.e. the Inter Quartile Range, and the implications of this for DAE at the 18 year threshold.