People smugglers, statistics and bone age

Desperate voyages, exploitation, naval interceptions and jailings… Who would guess that an obscure 50-year-old medical textbook about bone age in children would become the subject of statistical controversy in court cases about people smuggling across Australia? **Tim Cole** tells the story of how it happened.

Indonesian fishermen are in the news for ferrying illegal immigrants to Australia. Many are arrested by the Australian authorities and held in immigration detention, often for a year or more, charged with people smuggling. Their subsequent treatment depends critically on how old they are. If they are under 18 they are repatriated, but if older they are tried and, if found guilty, sentenced to 5 or more years in jail. As many of those who claim to be minors lack the documentation to prove it, the first stage in these cases involves an age assessment hearing, to decide how old they are.

A ramshackle fishing boat, crewed by four Indonesians and carrying eighty-odd Afghan migrants, was intercepted at sea. How old were the crew?

One typical case involved a ramshackle fishing boat, crewed by four Indonesians and carrying eighty-odd Afghan migrants, which was intercepted by the Royal Australian Navy. The boat was in a sorry state, with the engine room awash and the engine out of action, and it was sailing in circles as the rudder was broken. On being arrested, one of the crew gave his date of birth to indicate he was 17 years old; this led, 15 months later, to his appearance in court for an age assessment hearing.

The issue of age assessment in economic migrants is a global problem, and many countries have developed protocols to address it. The treatment of child migrants raises particular medical, statistical, ethical and human rights questions, and there has been a corresponding drive internationally to protect the interests of minors. In the UK, Professor Sir Al Aynsley-Green, the founding Children's Commissioner, has been tireless in his advocacy for migrant children subject to age assessment procedures, and has identified many ways in which the protocols need to be improved. As part of this he invited me (as a former colleague) to comment on the statistical arguments for age assessment being used in Australia, and this is how I came to be involved.

Deciding someone's age can be hard however old they are, but when they are between their teenage years and young adulthood it is particularly difficult. There are various signs one can use to assess developmental age – physical size (height and weight growth), sexual development (e.g. pubic hair or breast development), bone age or dental age – but they are all imprecise

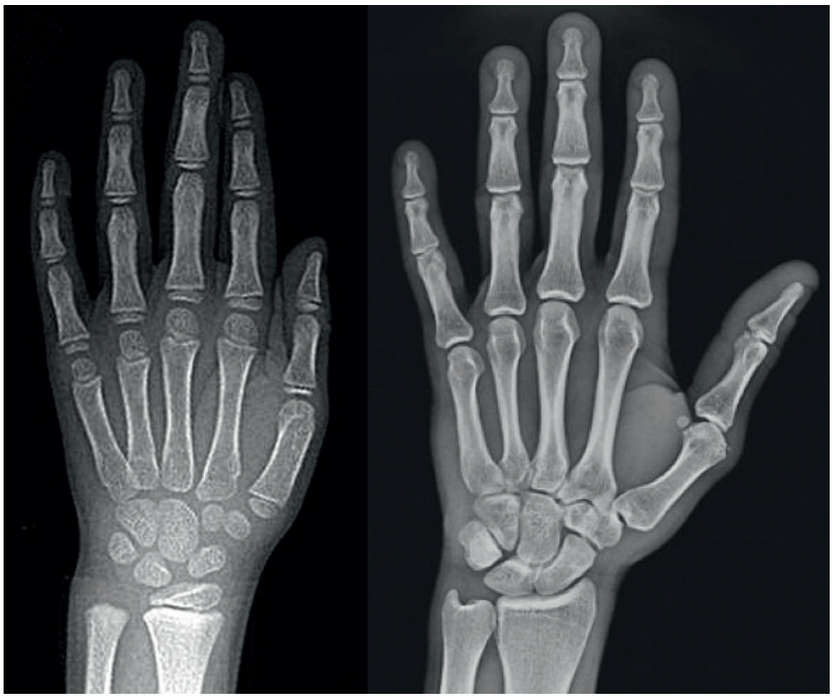

Figure 1. Composite X-rays of, left, an immature hand, and right, an adult hand. The growth plates at the end of the long bones can be clearly seen in the left-hand image. In the right-hand image the growth plates have fused with the long bones. It is impossible to tell the owner's age from this mature X-ray

because developmental age and chronological age are not that closely correlated. In addition, the signs all cease to be useful once the child is physically mature, as they are then indistinguishable from an adult.

Bone age, based on the radiographic appearance of the hand and wrist, is a popular assessment method, as it is easy for a radiologist to take and read an X-ray. Bone growth takes place at the growth plates at the ends of the long bones, and the radiographic appearance of the growth plates changes over time in a well-defined way. So by reading the X-ray one can judge, within a range of uncertainty, how far the child has travelled on their biological journey from birth to adult. The journey ends when growth stops, at which point the child is adult. This is when all the growth plates have fused and no further growth is possible. The appearance of the X-ray is then adult, and remains so throughout life (see Figure 1, right). By calibrating against chronological age the distance travelled from birth to adult in a group of reference children, one can express skeletal maturity as an "age" in units of years, and on average a child's bone "age" should match their chronological age.

The obscure textbook mentioned earlier is the *Radiographic Atlas of Skeletal Development of the Hand and Wrist* (2nd edition) by William W. Greulich and S. Idell Pyle[1], which was published in 1959 to help assess bone age. Greulich and Pyle were US paediatricians with an interest in developmental age in health and disease. Their *Atlas* consists of a series of standard hand–wrist X-rays, each corresponding to a year of age from birth to age 19 in boys and age 18 in girls. Each standard was chosen from a group of 100 middle-class US children of that age to represent the typical level of bone maturity in the group.

To apply the *Atlas* to a particular child, the paediatrician compares their X-ray with the *Atlas* standards and identifies the one most closely matching the child's own. The nominal age for the chosen standard, or an intermediate age if part-way between two standards, is then the child's bone age.

The resulting bone age correlates reasonably well with the child's chronological age, the standard deviation (SD) of the difference between the two ages being around 15 months at age 17. However, there is one key requirement – the X-ray has to be immature. This is because once the X-ray attains mature adult appearance, it is impossible to ascribe an age to the child and the whole concept of developmental age breaks down. For this reason Greulich and Pyle were interested only in skeletally immature children, whose bone ages they could calculate. They found that at age 19 most boys were skeletally mature, and for this reason their standard X-ray for 19 years shows a mature adult X-ray. This is made clear by their comment on the age 19 standard: "The fusion of the radial epiphysis [i.e. the growth plate on the radius bone] with its shaft completes the skeletal maturation of the hand and wrist".

The current furore over bone age arises because Australian immigration authorities are now using the Greulich–Pyle *Atlas* "in reverse". They assess the child's bone age from their X-ray in the usual way, but then treat it as an estimate of their unknown chronological age. Yet with an SD of 15 months the width of the 95% confidence interval around the estimate is ±30 months or ±2.5 years, indicating considerable uncertainty in their likely age.

The process becomes even murkier when the child's X-ray is mature. It shows that their age could be anywhere between teenage and 100+ years, clearly uninformative for administrative purposes. Yet one particular radiologist in Australia – let us call him Dr High – has developed a statistical argument to say that only a minority of young men with a mature X-ray are under age 18. His expert witness evidence has been used in a series of court cases over the past two years to convict young men of people smuggling, on the grounds that they crewed the boat and were probably over 18.

A person whose X-ray shows a mature wrist could be anywhere between teenage and 100 years old – so the X-ray is clearly uninformative to a court

The basis of Dr High's argument is as follows: the Greulich–Pyle *Atlas* shows an adult X-ray at age 19, so age 19 must be the mean chronological age for an adult X-ray. (This sentence probably deserves an

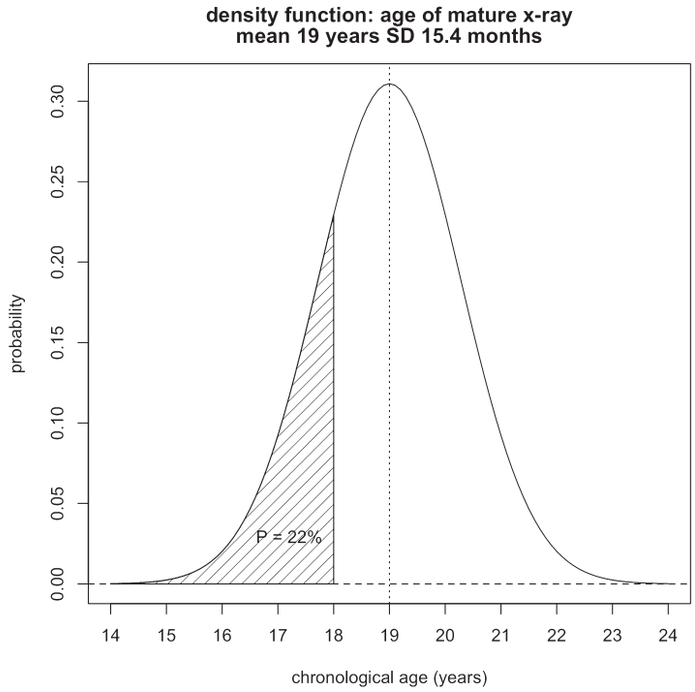

Figure 2. Normal distribution of chronological age, given a mature X-ray

exclamation mark at the end.) In addition, the standard deviation of the difference between bone age and chronological age is 15.4 months for boys aged 17, a statistic also given in the *Atlas*. So assuming that age is normally distributed, it is a simple matter to calculate the probability of the bone age being under 18 from the corresponding normal equivalent deviate of $(18 - 19)/(15.4/12) = -0.78$, that is, a lower tail probability of 22%. The probability of being adult is correspondingly 78%.

Dr High's standard report (it is essentially the same for all his cases) includes a figure of the normal distribution with tail area shaded (like Figure 2), and tables of probabilities corresponding to different age cut-offs, all of which give the report a professional look.

Age assessment hearings are civil cases where the verdict is based "on the balance of probabilities", which can be viewed informally as a probability exceeding one half. (For comparison, criminal cases require evidence "beyond reasonable doubt" corresponding to a much more extreme probability.) Thus evidence of the accused being adult, presented in the form of a probability comfortably exceeding 50%, is attractive to prosecutors and they have won several cases on the strength of it.

However Dr High's argument is specious, as is simply seen. The fact that 19-year-old boys tend to have a mature adult X-ray does not make age 19 the mean age for adult X-rays. Quite the reverse – there is no simple expression for the mean age of an adult X-ray, as the distribution is effectively unbounded at the upper end. Equally the standard deviation of 15.4 months applies to 17-year-olds (not 19-year-olds) with immature (not mature) X-rays, so it is doubly inappropriate. Thus the probability of 22% is based on an age distribution where neither the mean nor the standard deviation is valid.

That 19-year-old boys tend to have a mature adult X-ray does not make age 19 the mean age for adult X-rays

Clearly Dr High's probability is wrong, but can one come up with anything better? Also, and more importantly, is a probability what the court needs to know?

To extract useful information from the mature X-ray we should focus on the timing of an associated event, the age of attainment of skeletal maturity. This is the unobservable age when the X-ray first reaches adult appearance, and it is the lower bound of the child's possible age given that they have a mature X-ray. Knowledge of the distribution of the age of attainment would clarify just how young skeletally mature boys can be. The obvious place to look for this information is the Greulich–Pyle *Atlas*, yet it is silent on the subject. The reason why is not hard to see – mature X-rays were of no interest to Greulich and Pyle, so the age of attainment of mature X-rays was equally irrelevant to them.

Table 1. Centiles for age of attainment of skeletal maturity in boys[2]

Centile	97th	90th	75th
Age (years)	15.1	15.8	16.7

That said, their *Atlas* does contain some relevant information. As already mentioned, there are tables giving the SD of the difference between bone age and chronological age, by year of age up to 17 years. Greulich and Pyle only included children with immature X-rays in these tables, so a reasonable inference is that the tables stopped at 17 because there were too few boys aged 18 or 19 to include. In addition, the sample sizes by year in the tables decreased after 14 years, again suggesting that increasing numbers were being omitted as mature. On this evidence boys as young as 15 may be skeletally mature.

The ages of the standard X-rays provide another perspective. The age 18 standard is immature while the age 19 standard is mature – this suggests that the mean age of attainment is somewhere between 18 and 19 years.

However, for hard evidence on the age of attainment we have to refer to a more recent source. It is the so-called TW3 method of Tanner *et al.*[2] Tanner and Whitehouse (TW) developed a method of assessing bone age by scoring individual bones in the hand and wrist and adding the scores together, where a score of 1000 indicates skeletal maturity; TW3 is the third edition of the TW manual. This is in principle more accurate than the Greulich–Pyle whole-hand-matching procedure. The TW3 manual includes a table with selected centiles for age of attainment of skeletal maturity in boys – just what we want. The centiles are in reverse, reflecting immaturity rather than maturity, so the 97th centile at 15.1 years means that 97% of boys are immature at this age and 3% are mature. By

15.8 years a tenth of boys and by 16.7 years a quarter of boys are skeletally mature (Table 1). What, then, are the centiles of the distribution for later ages?

If we assume a normal distribution, the mean and standard deviation can be estimated from these centiles by regressing the normal equivalent deviates for the centiles against the corresponding ages, using weighted regression. The mean age is derived as minus the intercept divided by the regression coefficient, or 17.6 years, and the standard deviation is the reciprocal of the regression coefficient or 16.5 months. Figure 3 shows the three centiles and the fitted distribution function. The mean age of attainment is 17.6 years, and the probability of being mature before 18 is 61%, with bootstrap 95% confidence interval from 55% to 67%.

Note that if the individual data were available a similar distribution function could be estimated using logistic regression, with the outcome maturity status (i.e. mature versus immature X-ray) and the independent variable chronological age. The graph in Figure 3 would then be a logistic curve. However, in the absence of such data the TW3 table of centiles has to suffice.

The focus of the legal argument can now switch from the individual's current bone age to their (earlier) age of attainment. This increases the probability of their being under 18 from 22% to 61%. Put in words, it states that the majority of boys with a mature X-ray were under 18 when they became mature. Of course there is likely to be a lag between attaining maturity and being X-rayed, and if this lag were to average more than 0.4 years then the mean age when the X-ray was taken would still exceed 18 years. So one can argue about the usefulness of this probability as evidence to the court, but unlike Dr High's probability it is at least validly derived.

There is further uncertainty about the 61% probability, as the Greulich–Pyle *Atlas* on which it relies is based on US children seen in the 1930s. For several reasons this is unlikely to apply to Indonesian fishermen in 2010/11, but as there is no atlas for Indonesian fishermen the Greulich–Pyle *Atlas* remains the standard.

As I mentioned earlier, I was originally drawn into the area of forensic age assessment by Sir Al Aynsley-Green. He had been asked to write an expert witness report on the case of an Indonesian charged with people smuggling, whose bone age had been assessed by Dr High. Sir Al invited me to comment on Dr

A Royal Australian Navy officer stands on the coach house of a boat carrying suspected illegal immigrants near Ashmore Reef, about 850 km west of Darwin. Defence department photograph, April 16th, 2009. REUTERS/Australian Department

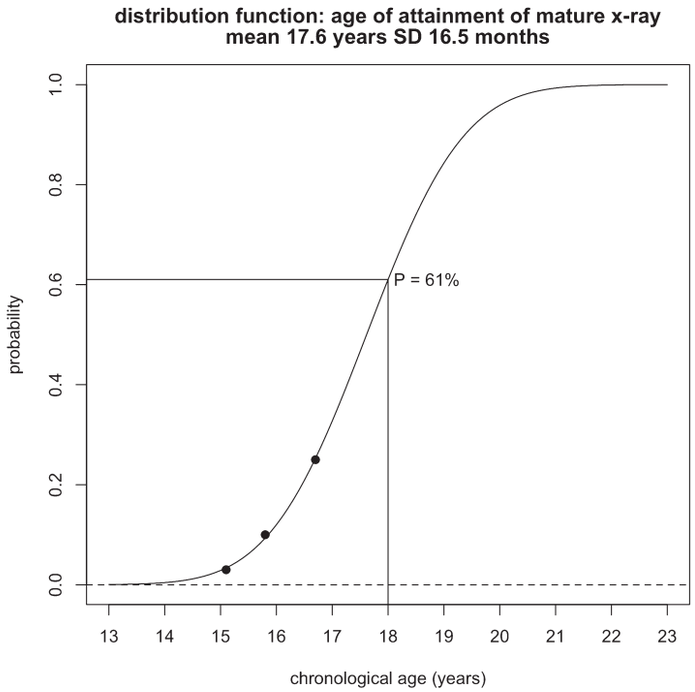

Figure 3. Distribution function for age of attainment of skeletal maturity, based on the three centiles shown in Table 1

High's probability calculation, and I put my conclusions in an appendix to the report. The case was subsequently dismissed. I was then approached by no fewer than 11 different lawyers from all over Australia, representing Indonesians on people trafficking charges with bone ages assessed by Dr High, asking me to act as expert witness to critique Dr High's statistics. Each case involved my writing a report, broadly similar to the story here, and sometimes giving evidence by videolink. Videolink evidence is a strange experience, getting up early or staying up late due to the time difference, and being cross-examined via television. Of the 11 cases for which I provided a report, nine were subsequently dismissed by the prosecution and the young men were repatriated to Indonesia.

The subtext of everything said so far is that a mature X-ray is pretty uninformative. It is compatible with being aged 15 or 50 or 85, and does not discriminate between young men aged between 17 and 19. Yet the probability under discussion, of the accused having a mature X-ray conditional on their being under 18, fails to highlight this lack of information. This is because it is not directly relevant to the case. What is wanted is the "reverse" probability, that of their being under 18 conditional on their X-ray being mature. This confusion of conditioning has a name: it is called the "prosecutor's fallacy".

To expand on this, the court knows that the accused has a mature X-ray, and it must decide whether he is over 18 or under 18. So what it needs to know is *two* probabilities, not one: the probability that he is over 18 versus the probability that he is under 18. Ideally the one probability should be near 100% and the other near 0%. The evidence of a mature X-ray would then clearly separate between adults and minors.

But the problem is that it does not. Figure 3 shows the probability of having a mature X-ray at different ages, rising from near zero at 14 years to near one at 22 years. If we choose two ages, say 17 and 19 years, the corresponding probabilities are 33% and 84%. The probability increases with age, but not that steeply. The ratio of the two probabilities, 2.6, is called the likelihood ratio. A male with a mature X-ray is 2.6 times more likely to be aged 19 than 17. The likelihood ratio is a compact summary of the evidential value of the X-ray for deciding on the individual's age, and the larger it is the better.

With DNA evidence, for example, the likelihood ratio can exceed 1 million. But here the emphasis is different – how *small* can the likelihood ratio be yet still be informative? Taking the analogy of diagnostic tests in medical practice, likelihood ratios less than 5–10 are generally viewed as uninformative. So the value here of just 2.6 is pretty useless – the misclassification rate is just too high.

In practice the age claimed by the accused may vary from under 15 years up to 18, usually based on a rounded date of birth such as January 1st. Clearly the likelihood ratio is higher the younger the claimed age – for 15 years say the likelihood ratio is 30, a clear indication that someone with a mature X-ray is unlikely to be that young. But for ages over 16 the likelihood ratio is less than 7, and hence essentially uninformative.

So the conclusion is this: for young men with a mature hand–wrist X-ray who are accused of people smuggling, and who say they are between 16 and 18 years old, the X-ray provides insufficient evidence for the age hearing to conclude that they are over 18. In practice the X-ray should not be used as evidence, since it raises other ethical issues relating to radiation dose, the nature of consent, and the fact that it penalises skeletally mature individuals.

In November 2011 the Australian Federal Parliament introduced its Crimes Amendment (Fairness for Minors) Bill. Around the same time the Australian Human Rights Commission set up an inquiry into the treatment of individuals suspected of people smuggling offences who say that they are children. Both are concerned about the processes of age assessment, and they have invited submissions from interested parties, to which I have responded. The hope is that when they report, they will recognise the evidential weakness of the hand–wrist X-ray, and recommend it not be used in the future.

References

1. Greulich, W. W. and Pyle, S. I. (1959) *Radiographic Atlas of Skeletal Development of the Hand and Wrist*, 2nd edn. Stanford, CA: Stanford University Press.

2. Tanner, J. M., Healy, M. J. R., Goldstein, H. and Cameron, N. (2001) *Assessment of Skeletal Maturity and Prediction of Adult Height (TW3 Method)*, 3rd edn. London: W.B. Saunders.

Tim Cole is Professor of Medical Statistics at the MRC Centre of Epidemiology for Child Health, UCL Institute of Child Health. His research interests include anything to do with the assessment and interpretation of body size and shape, including child growth and obesity.