People smugglers and statistics

Professor Patty Solomon

25 May 2105

Statistics
School of Mathematical Sciences
University of Adelaide

An ‘adult’ wrist X-ray

From the Greulich-Pyle Radiographic Atlas, 1959

Why wrist X-rays?

- Until recently, Indonesian fisherman were frequently in the news for bringing asylum seekers to Australia by boat.
- Many were arrested and held in detention charged with *people smuggling*.
- Between 2008 and 2012, **1115** crew were held.
- Their subsequent treatment depends on *how old they are.*
- *If they are under 18 they are sent home.*

- *If over 18 and found guilty, they are sentenced to 5 or more years in jail.*
- *Many who claim to be children have no birth certificate.*
- *Between 2000 and 2008 only 55% of Indonesian births were recorded and there are at least 3 different calendars.*
- *The first stage of a defendant's case is an **age assessment hearing** to determine how old they are.*

Where this started …

- In 2001, the Federal Government amended the *1914 Crimes Act* to enable ***age determination*** to be prescribed by regulation.
- At the same time, the *Crimes Regulation 1990* was amended to take **wrist X-rays of defendants** to assist in determining age.
- **No reference tool or methodology** was specified.
- The Government planned to take advice from **Dr Vincent Low** a radiologist.

Dr Low for the prosecution

MALE STANDARD 29

SKELETAL AGE: 17 YEARS

MALE STANDARD 31

SKELETAL AGE: 19 YEARS

GP Atlas

What Dr Low did …

- The *GP Atlas* shows an adult (mature) X-ray at 19, so he states *"19 must be the mean chronological age for an adult X-ray"*.

- **What is wrong with this statement?**

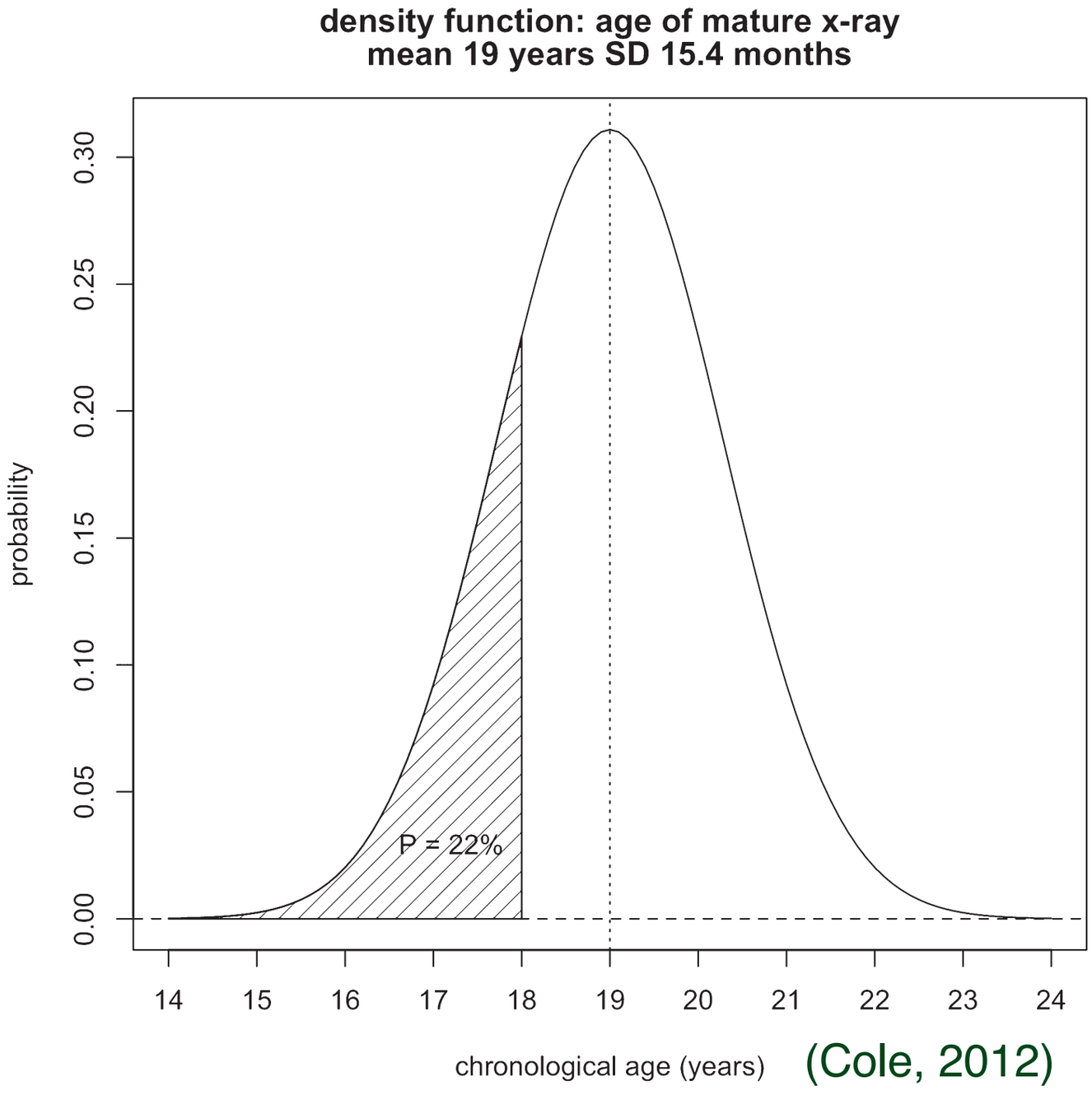
- ***19 years is not the mean age for mature X-rays!*** *Nor does it imply all males aged 19 are skeletally mature.*

What did Dr Low do next?

- ***Having chosen a mean, he then wanted a standard deviation.***
- From the *GP Atlas*, for boys aged **17**, he assumed the *sd of the difference* between bone age and chronological age is **15.4 months**.
- *Ummm,* boys aged **17** have ***immature*** *X-rays*.
- He also assumed age to be ***normally distributed.***

It's not!

He cobbled it all together …



(Cole, 2012)

Immigration authorities treat this as ***the distribution of chronological age given a mature X-ray***.

Dr Low's trouble with numbers

20. During the Jasmin age determination hearing, Dr Low also states that there is a one in a thousand chance of a person with a mature wrist X-ray being under 16, and that there is a probability of zero that a person under the age of 14 could have a mature wrist X-ray (p.47 Jasmin transcript). This point is directly disputed during testimony by Dr Paul Hofman at the AHRC's medical roundtable (p.16, roundtable transcript).

Dr Low's table:

Years of Age	% Probability ≤ this age
14	0
14.5	0.02
15	0.09
15.5	0.32
16	0.97
16.5	2.57
17	5.96
17.5	12.12
18	21.79
18.5	34.84
19	50

0.00005

"1 in 100"

In Civil Law cases …

- The verdict is based on ***the balance of probabilities,*** *i.e. the Judge thinks it is more likely than not.*
- So a ***probability > 0.5 of being an adult*** is attractive to prosecutors.
- *The Commonwealth won many cases on the strength of Dr Low's reports.*
- *With the GP Atlas 18-year standard* ***P=0.5*** *…*

Prof Cole for the defence

- He is a paediatrician with some qualifications in statistics.
- He was engaged to provide expert reports ***disputing*** Dr Low's conclusions.
- Prof Cole rejects the use of the *GP Atlas* and argues it the ***age of attainment of skeletal maturity*** which is important.
- *This has an **unobserved survival distribution.***

Survival distribution

- Defined by an event of interest **attainment of skeletal maturity** and time to the event, called the survival time, **age of attainment.**
- Observations are subject to **censoring** - the event is known to occur only **before or after** a certain time **or within an interval**.
- **Censoring** complicates the usual analysis.
- Survival distributions are typically **right skewed**.

What Prof Cole did

- He used the more recent manual by *Tanner et al* 2001, 3rd Edition *(TW3 Method).*
- *TW3* uses **skeletal maturity scores (SMS)**, which scores individual bones in hand and wrist and adds scores together.
- **SMS = 1000 indicates skeletal maturity.**
- ***TW3 uses children European-American from the 1960s -*** *a very different ethnic and socio-economic group to young Indonesian fishermen!*

TW3 contains the following table:

Table 8 Ages at which very early-maturing (97th centile), early maturing (90th centile) and fairly early-maturing (75th entile) boys and girls ach the RUS score of 1000 (from Tanner *et al.*, 1994)

Centile	Age (year)	
	Boys	Girls
97th	15.1	*13.3*
90th	15.8	*13.9*
75th	16.7	*14.6*

Prof Cole was pretty excited about these …

Where do these centiles come from?

Tanner-Whitehouse bone age reference values for North American children

James Tanner, MD, Dan Oshman, MD, Faranghise Bahhage, MD, and Michael Healy, MA

THE JOURNAL OF PEDIATRICS
JULY 1997 *Note year!!!*

Table I. Maturity scores (SMS) and UK60 bone ages, European-American males

CA group	Mean CA (yr)	Subjects (No.)	Subjects scoring 1000 (No.)	log_{10}SMS		log_{10}SMS, adjusted for censoring		Median SMS	Median BA	BA – CA
				Mean	SD	Mean	SD			
8.0-	8.4	75	0	2.407	0.0401			265	8.4	0.0
9.0-	9.4	100	0	2.453	0.0554			284	9.4	0.0
10.0-	10.4	82	0	2.520	0.0637			331	11.0	0.6
11.0-	11.4	57	0	2.594	0.1034			393	12.8	1.4
12.0-	12.3	68	0	2.666	0.0888			463	13.9	1.6
13.0-	13.3	52	0	2.735	0.0760			543	14.7	1.4
14.0-	14.4	41	0	2.801	0.0611			632	15.3	0.9
15.0-	15.4	57	5	2.886	0.0828	2.889	0.0887	774	16.2	0.8
16.0-	16.3	39	21	2.961	0.0570	2.998	0.0862	995	17.9	1.6

CA, Chronologic age; *BA*, bone age.

Tanner-Whitehouse bone age reference values for North American children

James Tanner, MD, Dan Oshman, MD, Faranghise Babbage, MD, and Michael Healy, MA

Fitted Means and Standard Deviations

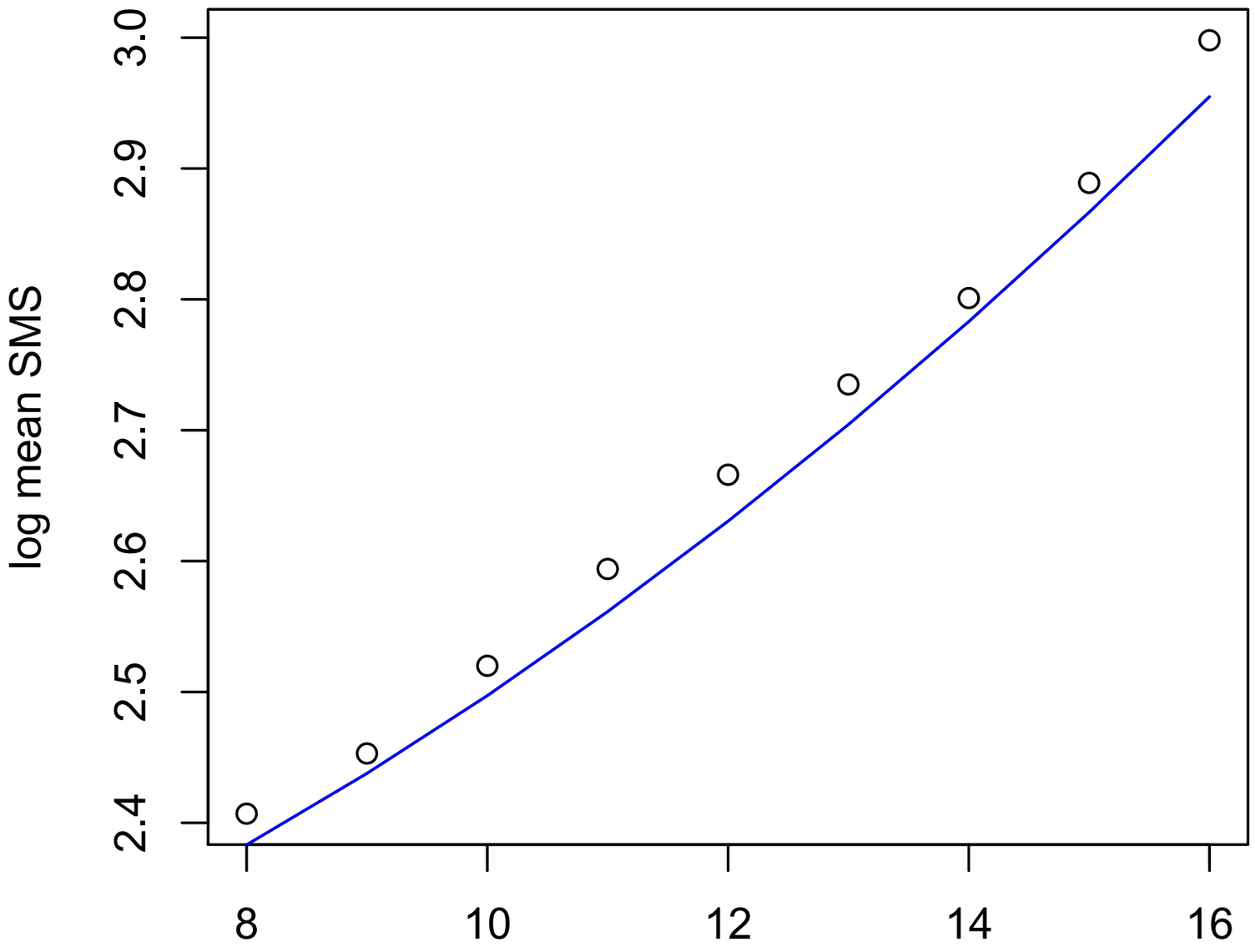
The equations relating means and standard deviations of the logged data to age (in years) were as follows (logarithms to base 10):

1. Boys
   a. $\text{Mean} = 2.120 + 0.1362 \times \text{Age} + 0.002410 \times \text{Age}^2$
   b. $\text{SD} = 0.02154 + 0.004216 \times \text{Age}$.

The coefficient 0.1362 is incorrect. Should it be 0.01362 ?

This is what you get if you fit

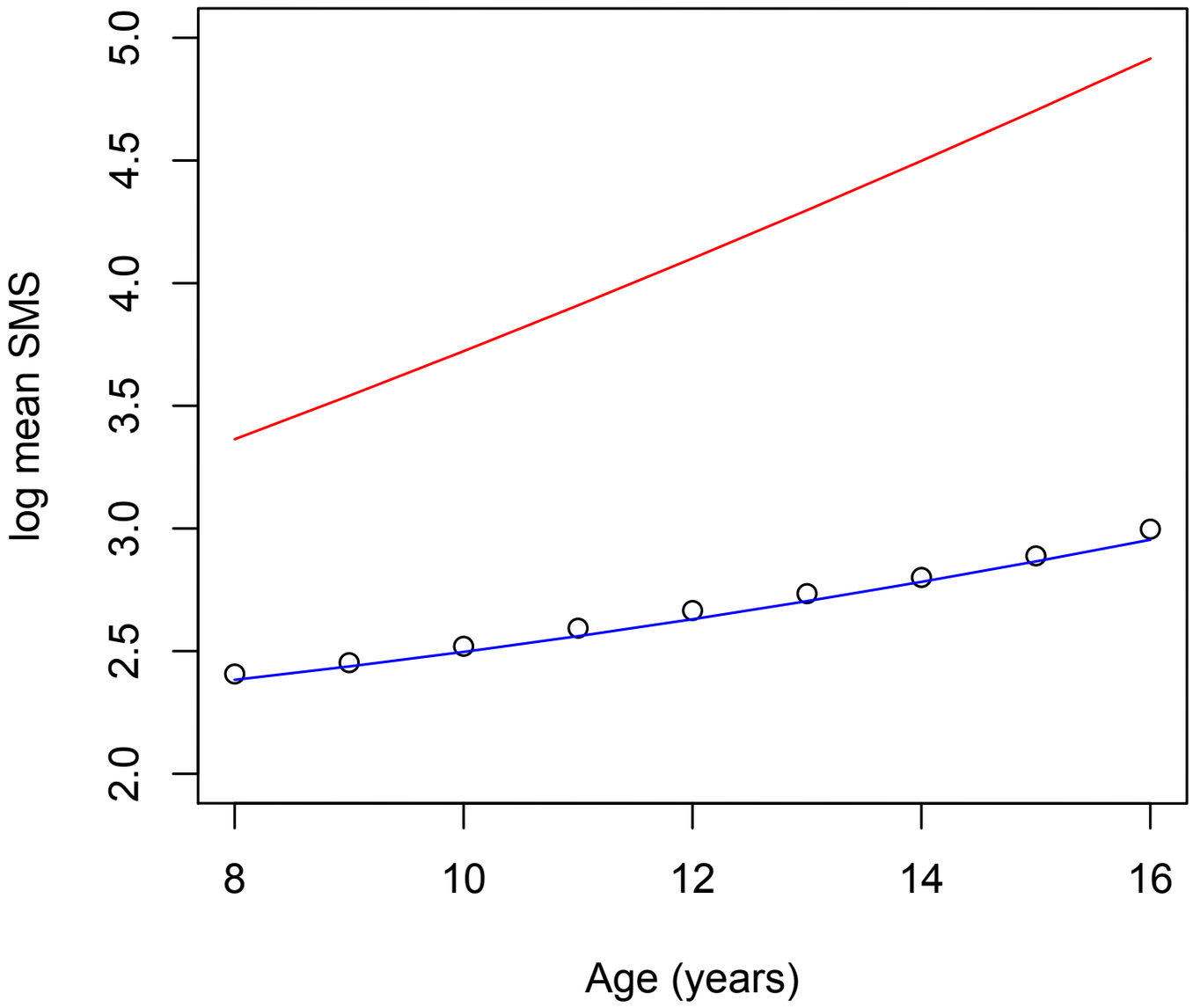
Mean = 2.120 + 0.01362 Age + 0.002410 Age^2



This curve misses all the points!

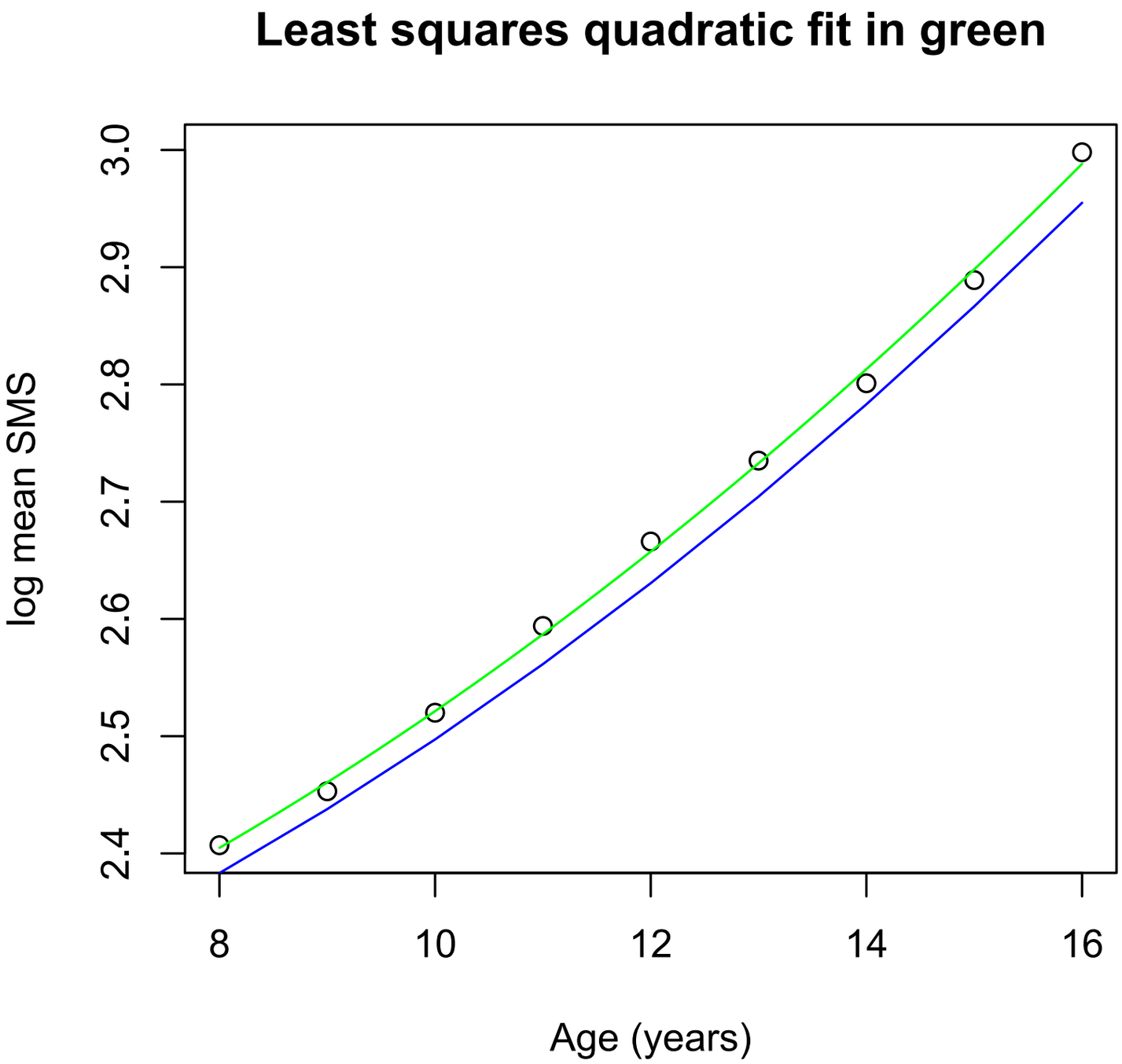
The given equation (in red) definitely misses ...

Mean = 2.12 + **0.1362** Age + 0.002410 Age^2



So let's use *least squares regression* to get the line of best fit.

$$\text{Mean} = 2.1335 + 0.014397 \times \text{Age} + 0.0024383 \times \text{Age}^2$$



If the data in Table 1 are correct, the **blue line** isn't the best fit.

Maybe they used slightly different data??

From the model fits, percentiles were derived using normal deviates (i.e. z)

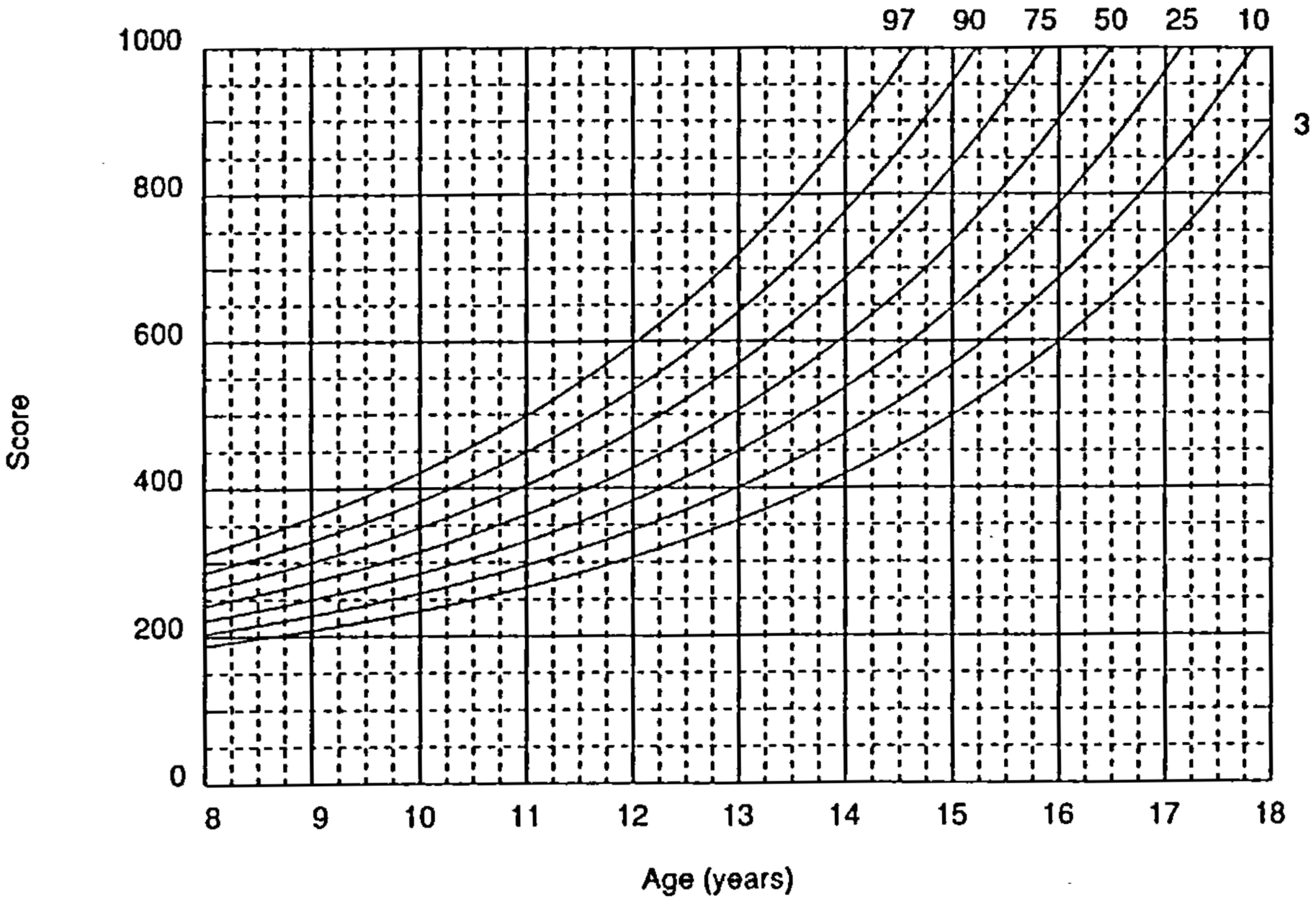


Fig. 3. US90 percentile reference charts, boys.

These charts are used in practice (shudder).

Scores of SMS 1000

- Table 8 of centiles (from *TW3*): *can't read these off the top of the chart because of* ***censoring***.
- *Tanner et al (1997)* say they used the *"assumed model"* to estimate age at which maturity is attained for 15 and 16 year old boys.
- ***We know the statistical models presented are not correct. It gets worse …***

The JOURNAL of PEDIATRICS

Correction

So, they got the Correction wrong for both boys and girls!!!

DOI: http://dx.doi.org/10.1016/j.jpeds.2012.09.021

Article Info

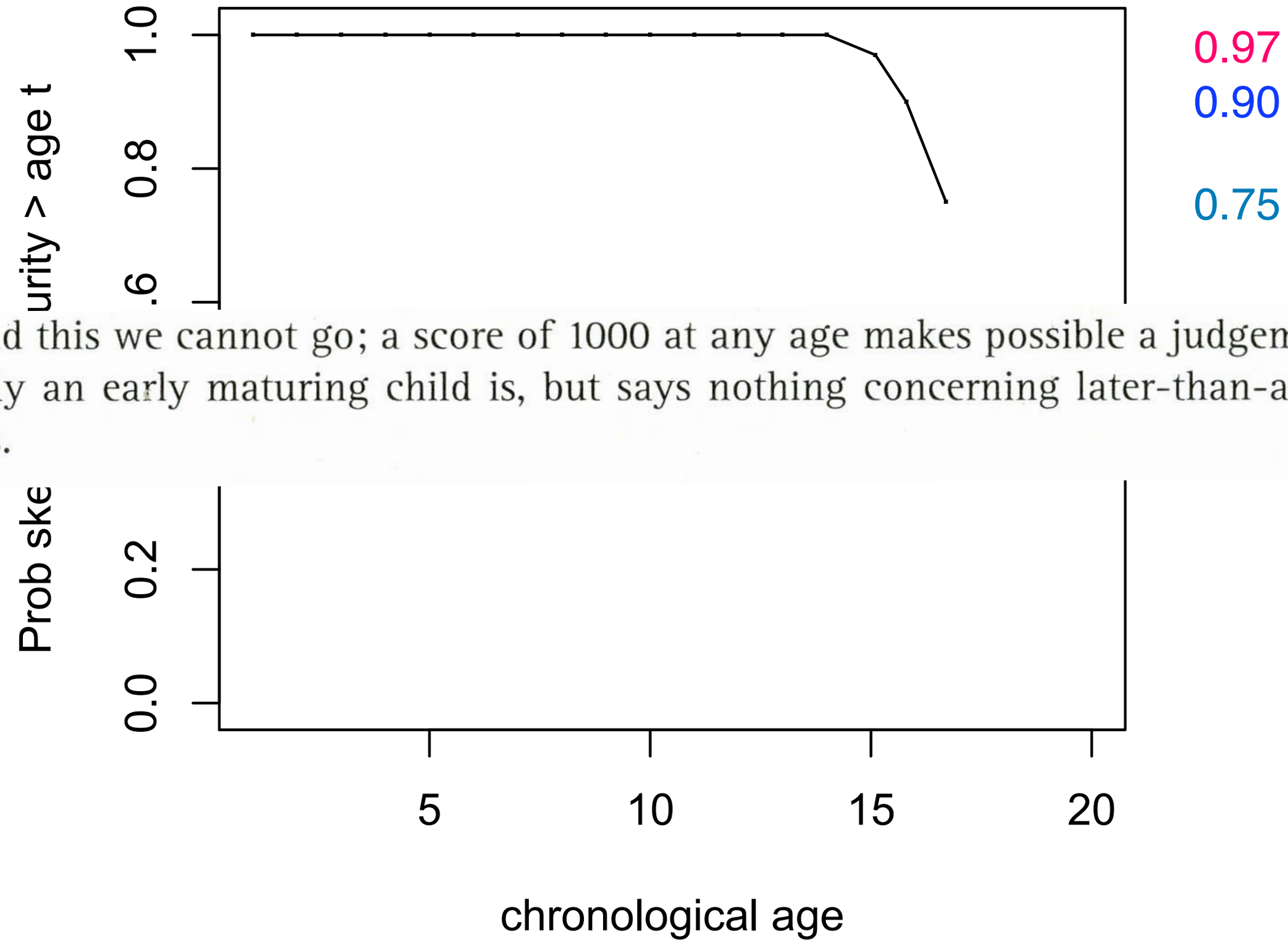
Abstract | Full Text

me

It was recently brought to our attention that corrections may need to be made in the article "Tanner-Whitehouse bone age reference values for North American children," by Tanner et al, J Pediatr 1997;131:8 ... rnal was unable to contact the authors of the article, the article was sent to an independent revie ... d not have access to the raw data, he was able to assess the mean and SD data in Tables ... ssion analysis should give results that are close to those in the published article. The coefficients ... SD equations on page 37 seem correct. However, the coefficients for mean age should be 0.01362 ... ales) and 0.01258 (females) for the equations for mean (log SMS), equations 1(a) and 2(a) (not 0.1362 and 0.1258, respectively, as stated on page 37).

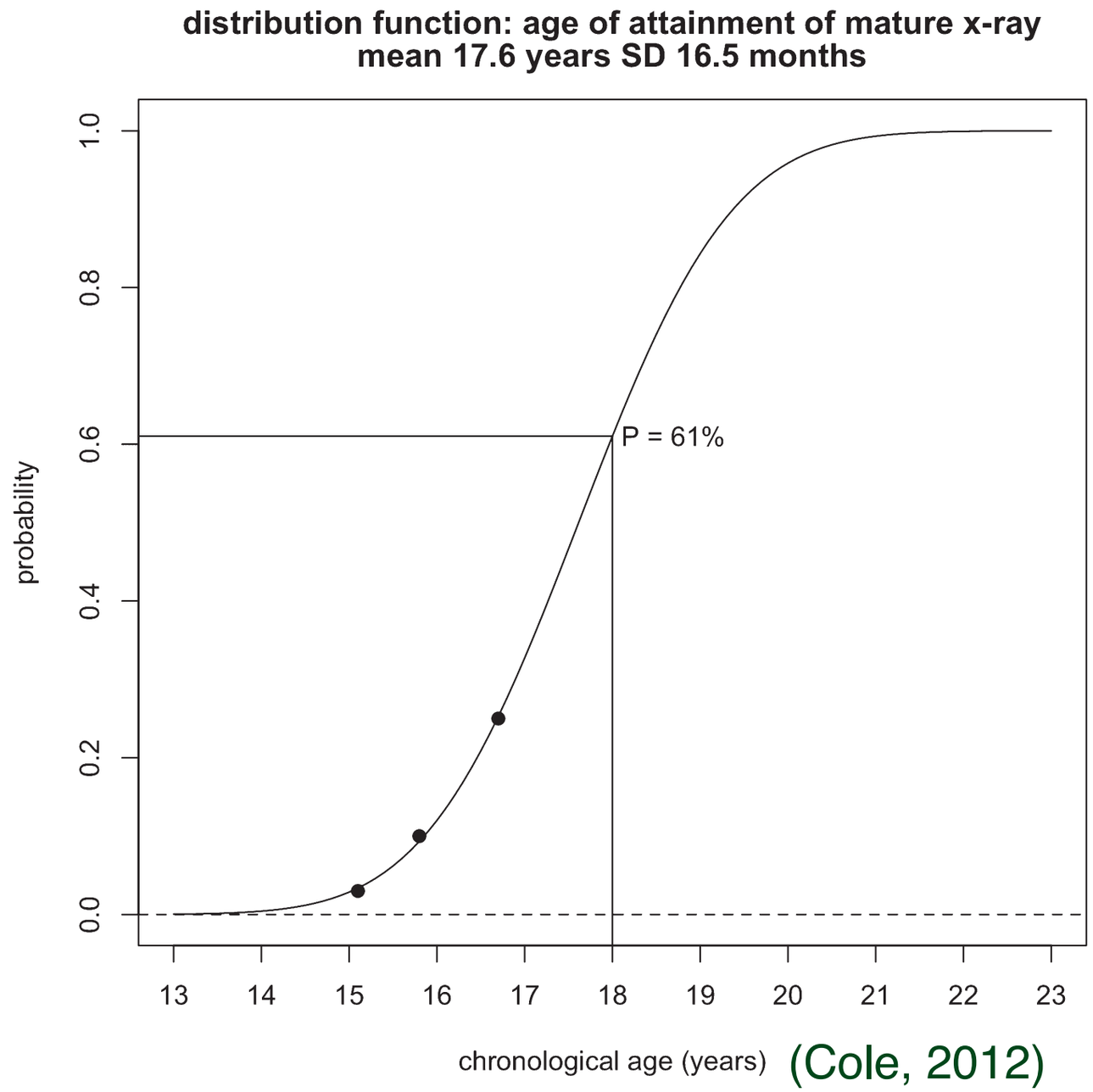
female age coefficient is correct

Probs of early skeletal maturity: Table 8



Beyond this we cannot go; a score of 1000 at any age makes possible a judgement of how early an early maturing child is, but says nothing concerning later-than-average maturers.

But Prof Cole does "go beyond"!



(Cole, 2012)

He "fits" a normal distribution to the three centiles.

Er … some measures of precision?

- Prof Cole didn't include any standard errors in his expert witness reports.
- But can compute ***bootstrap confidence intervals*** for each estimated probability.
- Later, Prof Cole obtained a 95% bootstrap CI of **(0.55, 0.67)**.
- So, Dr Cole claimed the majority of boys with a mature X-ray were under 18 when they became skeletally mature ***and 15 boys were immediately repatriated to Indonesia***.

At the 2012 AHRC hearing there was much argy-bargy over whether the two distributions were different.

Dr Low argues that the discrepancy between his and Prof Cole's statistics are because the last stage in the TW3 manual describes maturity as the *commencement* of fusion, while GP Standard 31 measures when fusion is *complete* (see Dr Low's supplementary submission to the AHRC). Dr Low concludes that in fact his figures are consistent with Prof Cole's, as the process of fusion described in Stage I of the TW3 Manual takes between one and two years, which explains why Prof Cole's statistics indicate that 'maturity' is reached at 17.6 years, and not 19 years. Prof Cole also addresses this issue

How could you compare the two distributions?

Use a two-independent samples z-test or t-test to compare the population means.

Two-sample *t*-test

$H_0 : \mu_{Low} = \mu_{Cole}$ versus $H_A : \mu_{Low} \neq \mu_{Cole}$

We observe $\bar{X}_{Low} = 19$ and $\bar{X}_{Cole} = 17.6$ with "estimated" standard deviations 1.283 and 1.375 years respectively.

The assumption of equal variance looks reasonable. Why?

There is a snag: we have *n_L = 60* but we have to approximate *n_C* = 48.

Now,

$$s_p^2 = \frac{(59)1.283^2 + (47)1.375^2}{60 + 48 - 2} = 1.7545$$

Two-sample *t*-test

Then,

$$t = \frac{17.6 - 19}{\sqrt{(\frac{1}{60} + \frac{1}{48})1.7545}}$$

$$= \quad -5.4580 \text{ on } 106 \text{ df.}$$

Recall the (two-sided) P-value is

$$P(|T| \geq |t|) = P(T \leq -t) + P(T \geq t)$$

$$= \quad 3.1187 \times 10^{-7}.$$

Two-sample *t*-test

Clearly we **reject** the null hypothesis that the two means are the same, and conclude that ***the mean age of skeletal maturity assumed by Prof Cole*** **is significantly lower** ***than the mean age assumed by Dr Low.***

This was broadly interpreted to imply an **"observed variation of two years"** using wrist X-rays to determine chronological age.

A mature wrist X-ray is not very informative ...

- It doesn't discriminate between boys aged 17 and 19.
- So far, we have considered ***the conditional probability of a mature X-ray given the defendant is under 18.***
- What the court wants to know is the ***reverse probability:*** ***of being under 18 given a mature X-ray.***
- This confusion is known as the ***prosecutor's fallacy.***

Bayes' Rule

- Gives a general formula for **updating probabilities** in the light of new information.

Let $P(\leq 18)$ be the probability that the defendant is under 18. (B_1)

Let $P(> 18)$ be the probability that the defendant is over 18. (B_2)

Now suppose the defendant has a mature wrist x-ray. (*A*)

Which probability is better supported by the evidence?

We know

$$P(A \cap B_1) = P(B_1|A)P(A) = P(A|B_1)P(B_1)$$

which implies

$$P(B_1|A) = \frac{P(A|B_1)P(B_1)}{P(A)}$$

We also know that if B_1, B_2 partition A,

$$P(A) = P(B_1)P(A|B_1) + P(B_2)P(A|B_2)$$

by the *Law of Total Probability.*

Then

$$P(B_1|A) = \frac{P(A|B_1)P(B_1)}{P(A|B_1)P(B_1) + P(A|B_2)P(B_2)} \quad .$$

For this we need $P(\text{mature X-ray}| \leq 18)$ and

$P(\text{mature X-ray}| > 18)$.

Let's take the average of 0% and 61% and the average of 61% and 100%.

Then

$$P(\leq 18|\text{mature X-ray})$$

$$= \frac{0.305 \times P(\leq 18)}{0.305 \times P(\leq 18) + 0.805 \times P(> 18)}$$

This is the ***posterior probability*** of being under 18 given the defendant has a mature X-ray.

Suppose there is ***strong corroborating evidence*** that the defendant is less than 18, which we represent by **0.99.**

Then

$$P(\leq 18|\text{mature X-ray}) = \frac{0.305 \times 0.99}{0.305 \times 0.99 + 0.805 \times 0.01}$$
$$= 0.974$$

which is **>> 0.5**

What if the ***prior belief that the defendant is a minor*** is not so strong, say **0.8**?

$$P(\leq 18|\text{mature X-ray}) = 0.602.$$

This is still > 0.5!

Sensitivity analysis demonstrates that the ***prior belief*** about ***chronological age*** is not substantially altered by **a mature X-ray.**

In other words ***there is little useful information in the wrist X-ray outcome*** about whether an individual is under or over 18.

canberratimes.com.au
The Canberra Times

National Times

Govt considers dumping wrist X-rays

May 17, 2013 - 7:33AM

The federal government is considering a halt to using wrist X-rays for determining age - a controversial method which has landed Indonesian boys accused of being people smugglers in adult jails.

The government on Thursday responded to a parliamentary committee into the detention of Indonesian minors in Australia.

It agreed in principle with a recommendation that it consider removing wrist X-rays as a prescribed method for age determination.

The controversial method has been used by federal police in determining the age of Indonesians detained on suspicion of being people smugglers.

The government's response noted advice from Australia's chief scientist Ian Chubb that wrist X-rays did not allow for precise estimation of age and that results varied with ethnic and socio-economic conditions.

Professor Chubb also pointed out that there was an "observed variation" of two years using the procedure.

Soon after, all remaining boys were released and sent home ...