Assessing the optimal time interval between growth measurements using a combined data set of weights and heights from 5948 infants

Charlotte Margaret Wright [1], Caroline Haig,[2] Ulla Harjunmaa,[3] Harshine Sivakanthan,[4] Tim J Cole [5]

► Additional supplemental material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/archdischild-2021-322479).

[1]Department of Child Health, School of Medicine, Nursing and Dentistry, University of Glasgow, Glasgow, UK
[2]Robertson Centre for Biostatistics, Institute of Health and Wellbeing, University of Glasgow, Glasgow, UK
[3]Center for Child, Adolescent and Maternal Health Research, Faculty of Medicine, Tampere University, Tampere, Finland
[4]Department of Human Nutrition, School of Medicine, Nursing and Dentistry, University of Glasgow, Glasgow, UK
[5]Population Policy and Practice department, UCL GOS Institute of Child Health, London, UK

Correspondence to
Professor Charlotte Margaret Wright, Royal Hospital for Children, Office Block CO/2, QE Hospital Campus, Govan, Glasgow G51 4TF, UK; charlotte.wright@glasgow.ac.uk

Received 18 May 2021
Accepted 26 August 2021
Published Online First 14 September 2021

Check for updates

To cite: Wright CM, Haig C, Harjunmaa U, *et al*. *Arch Dis Child* 2022;**107**:341–345.

ABSTRACT

Background Current guidance on the optimum interval between measurements in infancy is not evidence based. We used routine data to explore how measurement error and short-term variation ('noise') might affect interpretation of infant weight and length gain ('signal') over different time intervals.

Method Using a database of weights and lengths from 5948 infants aged 0–12 months, all pairs of measurements per child 2, 4 and 8 weeks apart were extracted. Separately, 20 babies aged 2–10 months were weighed on six occasions over 3 days to estimate the SD of the weight difference between adjacent measurements (=116 g). Values of 116 g and 0.5 cm for 'noise' were then used to model its impact on (a) the estimated velocity centile and (b) the chance of seeing no growth during the interval, in individuals.

Results The average gain in weight and length was much larger than the corresponding SD over 8-week and 4-week time intervals, but not over 2 weeks. Noise tended to make apparent velocity less extreme; after age 6 months, a 2-week velocity that appeared to be on to the ninth centile, would truly be on the second–third centile if measured with no noise. For 2-week intervals, there was a 16% risk of no apparent growth by age 10 months.

Conclusions Growth in infancy is so rapid that the change in measurements 4–8 weeks apart is unlikely ever to be obscured by noise, but after age 6 months, measurements 2 weeks or less apart should be treated with caution when assessing growth faltering.

What is already known on this topic?

- When measurements are collected too close together, measurement error or natural variation (*noise)* may mask the true underlying growth increment (*signal)*.
- Current recommendations on minimum measurement intervals are not evidence based.

What this study adds?

- Weight and length measurements collected 4 or more weeks apart in the first year are unlikely to be obscured by short-term 'noise'.
- Measurements two-week apart collected ater age 6 months are more likely to be obscured by 'noise' and are thus of limited value.

BACKGROUND

Successive weights and lengths measured in childhood are important for monitoring growth in individual children. Assessing growth over short intervals may allow earlier identification of growth faltering, but if the interval is too short, uncertainty in the measurement, or *noise*, may obscure the *signal* which is the true underlying growth increment.[1] In older children, a substantial time interval is recommended between measurements[2] and it has been suggested that over-frequent weighing in infancy may mislead or cause unnecessary anxiety.[3]

The guidance published with the UK-WHO chart stated that babies should be weighed no more than monthly before 6 months and two monthly aged 6–12 months.[4] However, these recommendations, and more recent less restrictive guidance,[3] were based only on expert opinion. There is thus a need for formal evidence.

The *noise* associated with measurement consists of both error and short-term variation. The WHO growth chart project team[5] estimated a technical error of measurement (TEM) for length of around 0.33 cm, using proper equipment, trained staff and regular quality control; in less well-regulated settings, the error will clearly be larger. The WHO growth chart project team did not assess measurement error for weight, presumably assuming it to be minimal with electronic scales. However, weight may vary in the short term, reflecting feeding and voiding patterns, which can be regarded as *noise*. Apart from one small study,[6] there are no published data on its magnitude.

Growth charts describe the average expected growth increment, which in infancy decreases with age and with shorter time intervals, but less is known about how much this increment varies with age and interval duration. As age increases and the interval decreases, there is an increasing chance that the increment will be zero or even negative.

To establish the impact of *noise* on the *signal*, we used two separate data sets to estimate:

1. Short-term variation in weight (*noise*) using survey data collected for the purpose.
2. The distribution of increments in weight and length (*signal*) at different ages and time intervals, using a database of routine growth data.

3. The impact of *noise*, over different ages and time intervals, on the *signal,* measured as the velocity centile or the chance of observing no growth during the interval.

METHODS

Weighing study: to measure *noise*

Mothers of Glasgow babies aged 1–12 months were recruited as part of a student project via social media and word of mouth. At baseline, the student researcher (HS) obtained consent and taught the parents how to weigh using Seca electronic scales, to the nearest 10 g. Both parents and researcher then separately weighed the baby, with both masked to the actual weight by adding to the scale numbered bags of unknown weight.

The families then collected weights at home twice daily over 2 days. The masking bags were not used at home, but to avoid weights being compared, each weight was recorded on a paper slip and posted into a collecting box. On the third day, the family returned and both parent and researcher again weighed the baby, masked as before.

All weights were entered into Microsoft Excel, and at the end of data collection, the numbered bags were weighed, and their weights subtracted from the gross weights. The parent and researcher weights were then compared with assess repeatability. Then, just the parent weights were used to assess variation over time.

Database study: to measure *signal*

Three existing longitudinal growth studies provided data, retrieved mainly from routine records. They had already been cleaned, checked and analysed for other publications.

Newcastle Growth and Development Study: a data set of routine weights of a birth cohort of 3418 children born at term in Newcastle upon Tyne between June 1987 and May 1988. Up to 11 weights, measured with mechanical or electronic scales, were retrieved from clinic records, and 3060 babies (90%) had at least two weights.[7]

Gateshead Millennium Study: a birth cohort of 1029 babies (923 term) born in Gateshead in 1999–2000, representing 81% of eligible births during the recruitment period. Routine weights collected using electronic scales were retrieved from baby clinic records, with a mean of 13 weights per child in the first year.[8 9]

Tampere Study: a data set of routine heights and weights of 2809 children aged 0–4 years born between October 2003 and September 2004 attending child health clinics in Tampere, Finland. Children were weighed by clinical staff on electronic scales. Up to 16 scheduled events were recorded per child, with a mean of 12 per child.[10]

DATA HANDLING

All database weights, plus the lengths in the Tampere Study, collected before age 12 months were combined in a single file. All pairs of measurements per child that were 2, 4 or 8 weeks apart were identified using the following definitions, chosen to maximise the number of intervals while minimising the relative variability:

- 2 weeks=14–15 days apart.
- 4 weeks=26–31 days apart.
- 8 weeks=50–63 days apart.

Intervening measures per child were skipped over to identify more widely spaced pairs. The measurement pairs were exported to per-interval data files along with the origin data set, the child's ID and gender, the two ages of measurement and the two measurements. Each pair was allocated to 3-month age groups in the first year based on the child's average age between the two measurements.

Statistical analysis

Noise was summarised as an SD called SD_{noise}. For weight, SD_{noise} was obtained from the weighing study, where the six parental weights per child were analysed by analysis of variance (ANOVA) to obtain the within-child residual SD, which was multiplied by √2 to give the SD of the weight difference, giving 116 g (see the Results section). In addition, the analysis compared mean weight as measured in the morning, daytime and evening (two each per child).

For length, noise comprised measurement error, based on the WHO TEM of 0.33 cm.[5] For the difference between two length measurements, $SD_{noise}=0.33\times\sqrt{2}=0.5$ cm.

For the database study, for both weight and length, the observed mean ($Mean_{obs}$) and SD (SD_{obs}) of the increment were calculated for each interval and age group. In addition, $Mean_{obs}$ and SD_{obs} were summarised as smooth cubic spline curves plotted against age (see online supplemental appendix).

The analysis then compared three versions of the SD: (1) SD_{obs} as observed (which included SD_{noise}); (2) SD_{obs} with the noise removed=SD_{signal}; (3) SD_{obs} with extra noise added=$SD_{obs+noise}$. Here, SD_{noise} was doubled to 1.0 cm for length and 232 g for weight, to model a context of greater random variation (see online supplemental appendix).

A child's growth increment is expressed as a velocity z-score: z=(increment–$Mean_{obs}$)/SD_{obs} and z is affected by noise via both $Mean_{obs}$ and SD_{obs}. If SD_{noise} rises, then SD_{obs} rises, and this shrinks z towards zero and the velocity centile moves closer to the average. We modelled the impact of adding and subtracting noise on the observed ninth velocity centile ($z=-1.33$) chosen to represent a child with slow, but normal weight gain at different ages.

As the growth rate slows with age, $Mean_{obs}$ decreases and the likelihood of there being no observed growth increases. A convenient milestone in this process is the age when $Mean_{obs}=SD_{obs}$, when a zero increment is 1 SD below the mean. By definition, this corresponds to the 16th centile, so at this age the chance of a zero or negative observed increment is 16%.

RESULTS

Weighing study - to measure noise

Twenty babies (12 female) aged 1.8–9.8 months were recruited in May–June 2018; 12 were exclusively and 4 partially breast fed, and all completed the protocol. Of the 40 immediately repeated measures, 33 (83%) were within 10 g, but 7 differed by up to 40 g. In contrast, only 12 of 100 successive weight pairs (excluding the researcher measurements) differed by less than 10 g. Using ANOVA, the residual within-child weight SD was 82 g, corresponding to an increment SD of 116 g for SD_{noise} (95% CI 102 g to 135 g). There was also a highly significant diurnal trend, with mean weight 42 g higher in the morning and 49 g lower in the daytime compared with the evening ($p<0.001$) (figure 1).

Database study: to measure *signal*

Of 5948 children with measurements in the first 12 months, 2624 had at least one pair of weights 2 weeks apart, 5081 4 weeks apart and 5663 8 weeks apart. The corresponding numbers for length, all from the Tampere Study (N=2809), were 1123, 2323 and 2426. The numbers of pairs available in different age groups for the different time intervals are shown in table 1. The mean increment ($Mean_{obs}$) and SD of the increment (SD_{obs}) were both

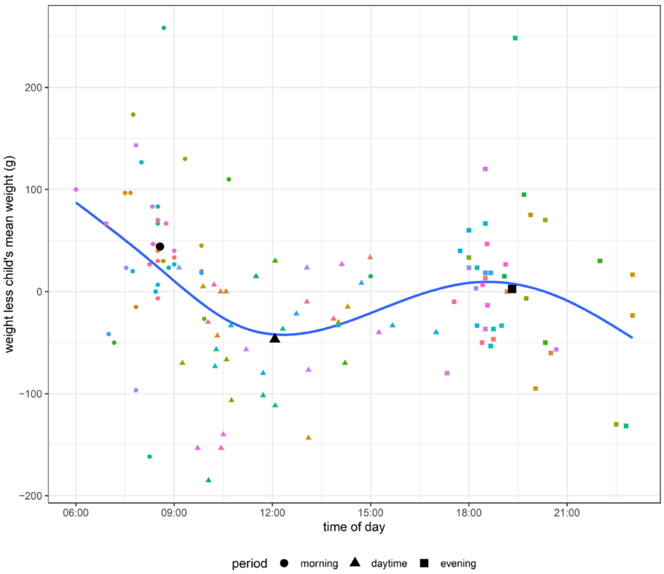

Figure 1 Modelled effect on the 9th velocity centile for weight and length by age, over two-, four- and eight-week intervals, with varying amounts of noise in the measurements. The authors can confirm that we have permission to reuse the image which was created by Professor Tim Cole.

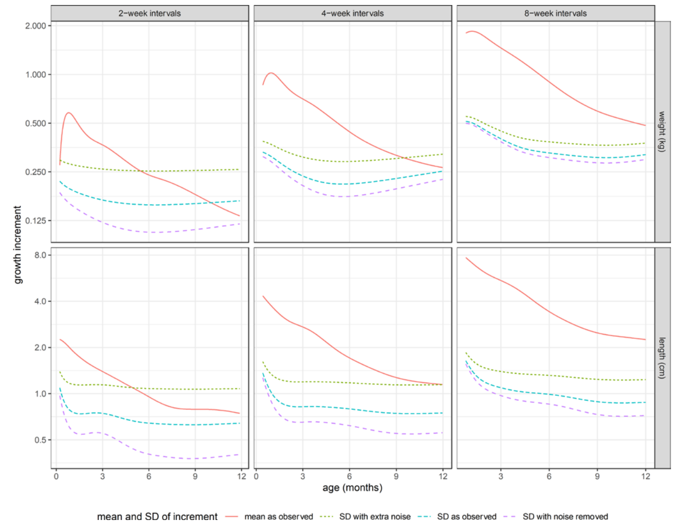

Figure 2 Spline smoothed curves plotted on a log2 scale of the mean growth increments in weight and length by age over 2-week, 4-week and 8-week intervals, along with the observed SD (SD_{obs}), the modelled true SD with noise removed (SD_{signal}) and the modelled SD with extra noise added ($SD_{obs+noise}$). The authors can confirm that we have permission to reuse the image which was created by Professor Tim Cole.

smaller for shorter intervals and fell with increasing age (table 1, figure 2). The effect of noise on the SD_{obs} can be seen in figure 2 as the separation at each age between it and SD_{signal} (ie, SD_{obs} with the noise removed) and $SD_{obs+noise}$ (SD_{obs} with extra noise added); the separation was greatest for the shorter intervals.

Table 2 summarises the risk of an infant failing to gain weight or length for different ages, intervals and amount of noise. For all intervals, the risk of seeing no gain with the SD_{obs} was low throughout the first year and remained so for intervals 4–8 weeks, even with extra noise. For extra-noisy 2-week measurements, the risk of seeing no gain reached 16% as early as 5–6 months, and there was a one in four risk of seeing no gain at 12 months.

Figure 3 uses the curves in figure 2 to model the impact of adding and subtracting noise at different ages on a child with slow/normal growth, corresponding to an observed velocity on the ninth centile (z=−1.33). The effect was greater the shorter the interval and increased in the early weeks, peaking at around 6–9 months. By 6 months, the modelled true velocity for weight and length over 2-week intervals was one centile space lower than the observed velocity; with extra noise this observed centile could have been almost one centile space higher.

DISCUSSION

This study used routinely collected data to assess the age when short measurement intervals become less useful, and how this

Table 1 Summary statistics on the mean and SD of weight and length increment over different time intervals, grouped by mean age

	Weight increment (kg)			Length increment (cm)		
Mean age	**n**	**$Mean_{obs}$**	**SD_{obs}**	**n**	**$Mean_{obs}$**	**SD_{obs}**
2-week intervals						
0–3 months	3857	0.46	0.22	1442	1.93	0.88
4–6 months	1060	0.31	0.17	105	1.22	0.66
7–9 months	290	0.22	0.16	33	0.80	0.73
10–12 months	95	0.16	0.17	17	0.89	0.64
4-week intervals						
0–3 months	8172	0.88	0.32	2984	3.44	1.14
4–6 months	5843	0.58	0.23	3183	2.16	0.89
7–9 months	1074	0.38	0.22	557	1.53	0.80
10–12 months	497	0.29	0.24	198	1.22	0.71
8-week intervals						
0–3 months	11 735	1.73	0.49	4012	6.52	1.5
4–6 months	7766	1.22	0.39	3631	4.62	1.22
7–9 months	3249	0.73	0.33	2067	2.88	0.98
10–12 months	2768	0.54	0.31	1854	2.38	0.87
Total	46 406			20 083		

Data from the Newcastle Growth and Development Study,[7] Gateshead Millennium Study[8 9] and Tampere Study.[10]
$Mean_{obs}$, observed mean; SD_{obs}, observed SD.

Table 2 The age when $Mean_{obs} = SD_{obs}$* for weight or length increment and the probability (%) of seeing no growth at age 12 months, depending on the amount of noise in the measurement. and the time interval.

Interval size		Age (months) when $Mean_{obs}=SD_{obs}$*			Chance of no growth at 12 months		
		2 weeks	4 weeks	8 weeks	2 weeks	4 weeks	8 weeks
Weight	SD as observed	10	12	>12	21%	15%	6%
	SD with extra noise	6	10	>12	30%	20%	10%
Length	SD as observed	>12	>12	>12	12%	6%	1%
	SD with extra noise	5	12	>12	24%	16%	3%

*When chance of no growth reaches 16%.
$Mean_{obs}$, observed mean; SD_{obs}, observed SD.

depends on the quality of the measurement. The first outcome, the risk of seeing no apparent growth, is obviously relevant, as a child failing to grow is concerning. The second outcome, the effect of varying the amount of noise on the growth velocity centile, is more technical, but also has important implications for assessing faltering growth.

To estimate noise, the study drew on published length data and newly collected weight data. Our weighing study was modest in scale, but still represents the largest formal study to date of short-term variation in weight during infancy. One previous study weighed seven children over 2 days with similar results.[6] Our 20 participants supplied 120 successive weights and the protocol used minimised digit preference bias. It was thought unethical to study infants younger than 1 month, but given the rapid rate of early growth, including these, would be unlikely to change our conclusions. Further, larger studies are needed at later ages to examine other factors affecting weight variation.

The combined data set also had some limitations. It was collated from cohorts studied in different eras, and there was some heterogeneity between them. The SD for 2-week weight increments was slightly, but significantly, higher in the Tampere sample (240 g) than the earlier UK samples (210 g), but this difference is unlikely to be important. All the length measurements used were collected in Finland, where there is a culture

Figure 3

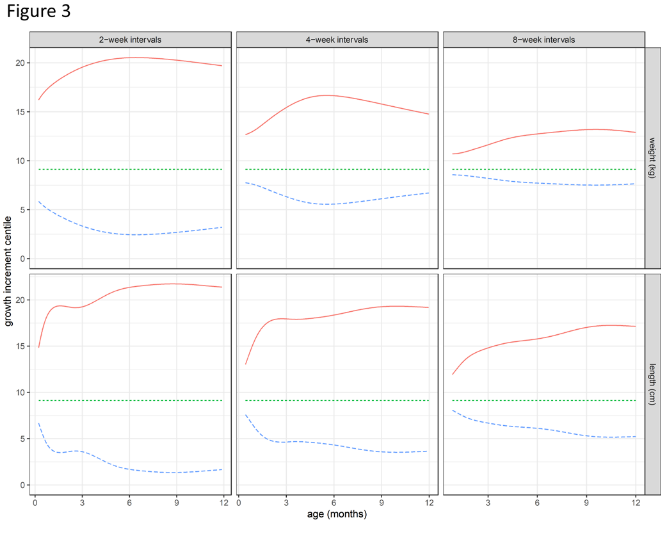

Figure 3 Modelled effect on the ninth velocity centile for weight and length by age, over 2-week, 4-week and 8-week intervals, with varying amounts of noise in the measurements. The authors can confirm that we have permission to reuse the image which was created by Professor Tim Cole.

of routine and widespread length measurement, producing a likely TEM close to the 0.33 cm used here. Where length is not measured frequently, or equipment is inappropriate, the TEM is likely to be larger, and this materially reduces the sensitivity of the assessment.

The database analysis revealed that infant growth is so rapid that even two weekly measurements are unlikely to be materially affected by noise until after the age of 6 months. After that the rate of growth slows, so the mean increment is smaller, while the SD changes little, so that by around 10 months, for 2-week intervals, the amount of variation in weight, the SD, is greater than the average increment. In these circumstances, there is a risk that an apparent small gain may simply reflect short-term weight increase, such as a large feed, or conversely that there might apparently be no gain simply because the child has just emptied their bladder and not yet fed. However, this is still not a high risk; at the point where the mean increment equals the SD, there is, by definition, a 16% chance of no observed gain.

The diurnal trend in weight was striking, being higher in the morning and lower in the day compared with the evening. This suggests that to minimise noise, weight should be measured at the same time each day.

The potential impact on growth assessment is important. Figure 3 shows that at 6 months, the observed ninth velocity centile selected because it represents low but usually acceptable weight gain, corresponds to a true velocity on the second–third centile for 2-week intervals. Thus, what appears to be a low normal level of gain is in fact at the very bottom of the normal range. More troublingly, if the accuracy of measurement were lower, the observed weight gain would be closer to the 25th centile. Thus, both imprecise and over-frequent measurements may obscure detection of growth faltering and falsely reassure. For longer intervals, the effect of noise is much smaller and less likely to be clinically significant.

Measuring length in infancy can be challenging, and assumptions about its likely inaccuracy led the UK-WHO growth charts team to recommend that length should only be measured when there was clinical concern.[4] Our data suggest that successive length measurements are nearly as robust as weight measures, so long as the TEM is as low as 0.33 cm. Even with increased imprecision, lengths collected 4–8 weeks apart are unlikely to be masked by measurement error. Thus, our results for both weight and length suggest that the guidance on current charts[4] may be too conservative.

CONCLUSIONS

For infants growing steadily, measurement intervals of 2 weeks or more are unlikely to result in true growth ('signal') being obscured by measurement error and/or short-term variation ('noise'), where this is of the order of 116 g or 0.5 cm. However, for detecting slow growth, and particularly when length is measured imprecisely, measurements collected only 2 weeks apart should be treated with caution and repeated before being used for any important clinical decision.

Acknowledgements The authors express their gratitude to all study members and their families involved in the three cohorts. They express thanks to all the study teams involved, both past and present, and for the support of all funders. In particular, for GMS they acknowledge the support in conducting the study from the Gateshead Health NHS Foundation Trust, an external Reference Group, the Gateshead Education Authority and local schools. They are also grateful to Professor John McColl for his help in planning this analysis and to Fred Ho for his helpful comments on the manuscript.

Contributors CMW conceived the study, led the analysis and drafted the paper. CH extracted the data for the main analysis and undertook the basic analysis. UH collected the Finnish data, advised on its use and commented on the draft. HS ran the weighing study and undertook the literature review. TJC advised on the design, undertook the main analysis and contributed to the paper drafting.

Funding This work was supported by Chief Scientist Office, Scotland (90549) and Glasgow Children's Hospital Charity (168979-01).

Competing interests None declared.

Patient consent for publication Not required.

Ethics approval Ethical approval was obtained from the University of Glasgow MVLS Ethics Committee (application number 200170123).

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request.

ORCID iDs

Charlotte Margaret Wright http://orcid.org/0000-0001-6256-6315
Tim J Cole http://orcid.org/0000-0001-5711-8200

REFERENCES

1 Cole TJ. Growth monitoring with the British 1990 growth reference. *Arch Dis Child* 1997;76:47–9.
2 Voss LD, Wilkin TJ, Bailey BJ, *et al*. The reliability of height and height velocity in the assessment of growth (the Wessex growth study). *Arch Dis Child* 1991;66:833–7.
3 NICE. Faltering growth: recognition and management of faltering growth in children. London National Institute for health and care excellence; 2017. https://www.nice.org.uk/guidance/ng75
4 Wright CM, Williams AF, Elliman D, *et al*. Using the new UK-WHO growth charts. *BMJ* 2010;340:c1140.
5 WHO Multicentre Growth Reference Study Group. Reliability of anthropometric measurements in the who multicentre growth reference study. *Acta Paediatr Suppl* 2006;450:38–46.
6 Alsop-Shields L, Alexander H. A study of errors that can occur when weighing infants. *J Adv Nurs* 1997;25:587–94.
7 Wright CM, Matthews JN, Waterston A, *et al*. What is a normal rate of weight gain in infancy? *Acta Paediatr* 1994;83:351–6.
8 Parkinson KN, Pearce MS, Dale A, *et al*. Cohort profile: the Gateshead millennium study. *Int J Epidemiol* 2011;40:308–17.
9 Wright CM, Parkinson KN, Drewett RF. How does maternal and child feeding behavior relate to weight gain and failure to thrive? Data from a prospective birth cohort. *Pediatrics* 2006;117:1262–9.
10 Harjunmaa U. Current growth patterns of Finnish children aged 0 to 4 years 2009 https://trepo.tuni.fi/handle/10024/81098

TECHNICAL APPENDIX

SD_{noise} affects SD_{obs} in the following way:

$$SD_{obs}^2 = SD_{signal}^2 + SD_{noise}^2$$

where SD_{signal} is the underlying true but unmeasurable SD, as if based on measurements free of noise. Because the terms are squared, SD_{noise} needs to approach SD_{signal} in size before it has any important effect on SD_{obs}. Rearranging the formula as:

$$SD_{signal}^2 = SD_{obs}^2 - SD_{noise}^2$$

provides an estimate of SD_{signal}. An extra-noisy version of SD_{obs} is obtained by doubling SD_{noise} to give:

$$SD_{obs+noise}^2 = SD_{signal}^2 + (2 \times SD_{noise})^2.$$

For Figure 2 the increment data were grouped by measure and time interval, and their mean and SD were modelled as P-spline curves in √age with 6 degrees of freedom using the NO family in GAMLSS (1). The choices of degrees of freedom and age transformation were guided by the BIC. Table 2 is based on Figure 2, with columns 2-4 corresponding to the ages where the Mean curve crosses each of the SD curves, while columns 5-7 are the % velocity centile corresponding to the z-score z = -Mean/SD at 12 months.

For Figure 3 the increment data were again grouped by measure and time interval, and expected age-specific increments corresponding to the 9th velocity centile were calculated as $Mean_{obs}$ - 4/3 SD_{obs} using the smoothed values in Figure 2. Then each increment was converted to a velocity z-score z = (increment – $Mean_{obs}$) / SD_x where SD_x was respectively SD_{signal} and $SD_{obs+noise}$, and the corresponding velocity centile curves were plotted against age.

1. Rigby, R. A. and D. M. Stasinopoulos (2005). "Generalized additive models for location, scale and shape (with discussion)." Applied Statistics 54(3): 507-544.