Medical Hypotheses 129 (2019) 109234

Contents lists available at ScienceDirect

Medical Hypotheses

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

Weight velocity equations with 14–448 days time separated weights should not be used for infants under 3 years of age

Martin J.C. van Gemert[a,*], Cornelis M.A. Bruijninckx[b], H.A. Martino Neumann[c], Pieter J.J. Sauer[d], D. Martijn de Bruin[a], Ton G. van Leeuwen[a]

[a] Department of Biomedical Engineering & Physics, Amsterdam UMC, Location AMC, University of Amsterdam, Amsterdam, The Netherlands
[b] Surgery Expert Witness, Rotterdam, The Netherlands
[c] Department of Dermatology, Erasmus University Medical Center, Rotterdam, The Netherlands
[d] Department of Pediatrics, University Medical Center, Groningen, The Netherlands

A B S T R A C T

Abnormal growth of infants may indicate disease of the children, thus methods to identify growth disorders are wanted in medicine. We previously showed that two-time-points weight growth velocities at age t, calculated by a commercial software product as [Weight(t) − Weight(t − X)]/X, with X = 448 days, were erroneous due to the long separation of 448 days. We were convinced that shorter X-values would solve this accuracy problem. However, our hypothesis is that: "shorter time separations than 448 days cause a *decreased* accuracy of numerical weight velocity equations in realistic infant weights until an age of about three years". Supporting evidence comes from analyzing how shorter X-values will affect the accuracy of two-time-points weight velocity calculations. We systematically varied X between 1 and 448 days of various P50/0SD-related standard weight curves: (a) P50/0SD with the weights separated by 1 day and X = 1,28,224,448 days; (b) P50/0SD with the weights at variable ages and X = 14–448 days; and (c) case (b) and incorporating weight fluctuations typically occurring in infants. Cases (b) and (c) include details observed in a clinical case. Our results show that the combination of weight fluctuations and varying time intervals between consecutive weights make weight velocity predictions worse for shorter X values in children younger than three years. Because these two causes of failure occur naturally in infants whose weight is regularly measured, other weight velocity equations face the same causes for inaccuracy. In conclusion, our hypothesis suggests that any software that predicts weight velocities should be abandoned in infants < 3 years. Practically, it should require that when (commercial) software weight velocity prediction suggests a medical problem, careful clinical checking should be mandatory, e.g. by linking predicted and exact weight velocities at age t (the latter from the mathematical first derivative at age t of standard weight curves).

Introduction

Abnormal growth of infants and toddlers may be an indication of disease of the children implying that methods to distinguish abnormal from normal growth are needed in neonatal and pediatric medicine. Commercial software exists for this purpose, for example *Growth Analyser Viewer Edition* (*GAVE*) from the *Dutch Growth Research Foundation* [1,2]. To our best knowledge, public documentation and validation of GAVE is lacking. Personal communication with one of the developers informed us that GAVE calculates weight gain velocities (called weight velocities) from the weight difference of two weight points, separated by 448 days. When the body weight is measured prior to 448 days, the weight gain since birth is used.

In previous work [2], we showed that the very long time separation of 448 days between the weight points caused erroneous outcomes of weight velocity calculations. Our analysis was based on the measured weights of a boy, born at term as the 6th child of healthy parents, who failed to gain weight between 2 and 4 months of age due to a cow milk allergy and where weight growth regained strongly when his food was kept free of cow milk [3]. The calculation of weight velocity by GAVE was erroneous [2] and contributed later, around the age of 1 year and 8 months, to the improper diagnosis of the boy's pediatrician (3rd pediatrician mentioned in [3]) that the mother purposely malnourished her son, called Pediatric Condition Falsification or Munchausen Syndrome by Proxy [3], leading to his eight months legal placement in foster care [2].

The rationale for choosing 448 days as the time separation between the two weight points by GAVE, unfortunately, is lacking. Interestingly, already in 2014, Ghaemmaghami et al. [4] concluded that "*It is worth mentioning that – one year is too long for length and weight measurements during infancy,* –". The question then is whether shorter time separations (X, in days) than 448 days between the two weight points would increase the accuracy of weight velocity predictions. If true, this could help preventing devastating events as improper foster care placement from happening again. Because we expected that shorter time separations would increase the accuracy of two-time-points weight velocities,

* Corresponding author.
E-mail address: m.j.vangemert@amc.uva.nl (M.J.C. van Gemert).

https://doi.org/10.1016/j.mehy.2019.109234
Received 16 March 2019; Accepted 18 May 2019

we did not evaluate it in our previous paper [2]. However, the hypothesis to be discussed in this paper claims the opposite of what we expected.

Shorter time separations than X = 448 days between two weight points cause a *decreased* accuracy of equations used to calculate weight velocities in realistic weights of infants until about three years of age.

Methods

We will evaluate the hypothesis by showing its validity in the two-weight-points weight velocity Eqs. (1) below, by systematically reducing time separation X from 448 to 14 days. We have chosen to use the P50(t) (or 0SD) standard weight curve, t is age in days, until three years of age (P stands for Percentile, P50 means that 50% of the children are above- and 50% are below the weights of that curve, in Standard Deviation notation also named 0SD). We included two realistic variations in the P50 curve. First, P50(t_n), t_n is the age of the nth weight measurement, where we incorporated varying time intervals between consecutive weights, representing realistic weight measurements in toddlers and, second, P50fluc(t_n), where we added the relatively large fluctuations ("fluc") that can typically occur in the weights of these young children, due to physiological causes, e.g. an empty or full colon and/or bladder, and pathophysiological causes, e.g. feeding problems, sickness, and increasing weights due to catch-up growth after sickness [5].

Weight curves: P50(t), P50(t_n) and P50fluc(t_n)

The P50(t) standard weight curve we used is for Dutch boys [6], fitted by a 5th degree power series of age [2]. In clinical practice, measured weights are typically available at variable ages t = t_n. We therefore also included the P50(t_n) weight curve, which is the P50(t) curve but sampled at 180 ages t_n measured from 0 to 1131 days as in the clinical case [2] (Fig. 1). Finally, the P50fluc(t_n) weight curve is obtained from the P50(t_n) curve with the typical weight fluctuations for infants included as derived from the measured weights of the clinical case (Fig. 1, see Appendix A for derivation).

Calculation of two-time-points weight velocity

Two-time-points weight velocity at age t (days), expressed in kg/year, is calculated in a similar way as by GAVE, as the difference between the weight at age t and the weight at a fixed time separation of X days earlier (called the "*earlier weight*"), at age (t-X) days, divided by X/365.25 years (assuming one year is 365.25 days). When a weight measurement is not available on day (t-X), the next available weight at day (t-X+ε) is utilized. When age t is less than X days, we used the weight gained since birth divided by t/365.25 years. Thus, similar to Eqs. (1) in previous work [2], we define weight velocity as

$$WeightVelocity(X,\,t) = \frac{Weight(t) - BirthWeight}{t/365.25} \quad t < X \tag{1a}$$

$$WeightVelocity(X,\,t) = \frac{Weight(t) - Weight(t - X + \epsilon)}{\frac{X}{365.25}} \quad t \geq X \tag{1b}$$

We varied time separation X as

X= 1, 14, 28, 56, 112, 224 and 448 days (2)

where X = 1 day is used only for the P50(t) weights.

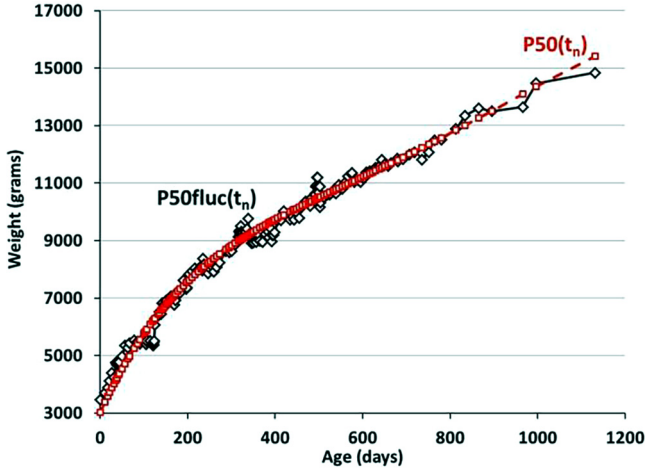


Fig. 1. Weight curves P50(t_n) (red line, open red squares) and P50fluc(t_n) (open lozenges), the latter explained in the Appendix.

Consecutive weight velocities of the P50(t_n) curve: cwvP50(t_n)

From the P50(t_n) weight curve we define the consecutive weight velocities for all 180 weight points as cwvP50(t_n)

$$cwvP50(t_n) = \frac{Weight(t_n) - Weight(t_{n-1})}{(t_n - t_{n-1})/365.25} \tag{3}$$

This weight velocity curve is the gold standard of weight velocities for the P50(t_n) weights.

Average Relative Weight Velocity Difference: ARWVD(X)

Below, we will abbreviate *WeightVelocity* as *"wv"* and we will indicate from which weight curve and, if X-values are included, with what X-value, it is calculated, e.g. for the P50(t_n) weight curve as wvP50(X,t_n).

The contribution to the inaccuracy of weight velocity assessment of: (1) the variable ages that separate consecutive weights, particularly through parameter ϵ in Eq. (1b), and (2) the weight fluctuations of toddlers, was evaluated by comparing weight velocities wvP50fluc (X,t_n), with both elements included, and the gold standard cwvP50(t_n), with element (1) included. To that, we calculated the *Average Relative Weight Velocity Difference* (ARWVD) at all age points t = t_n, using the modulus of [*wvP50fluc(X,*t_n)-cwv*P50(*t_n)]/cwv*P50(*t_n), averaged over all N = 180 weight velocity data points, thus

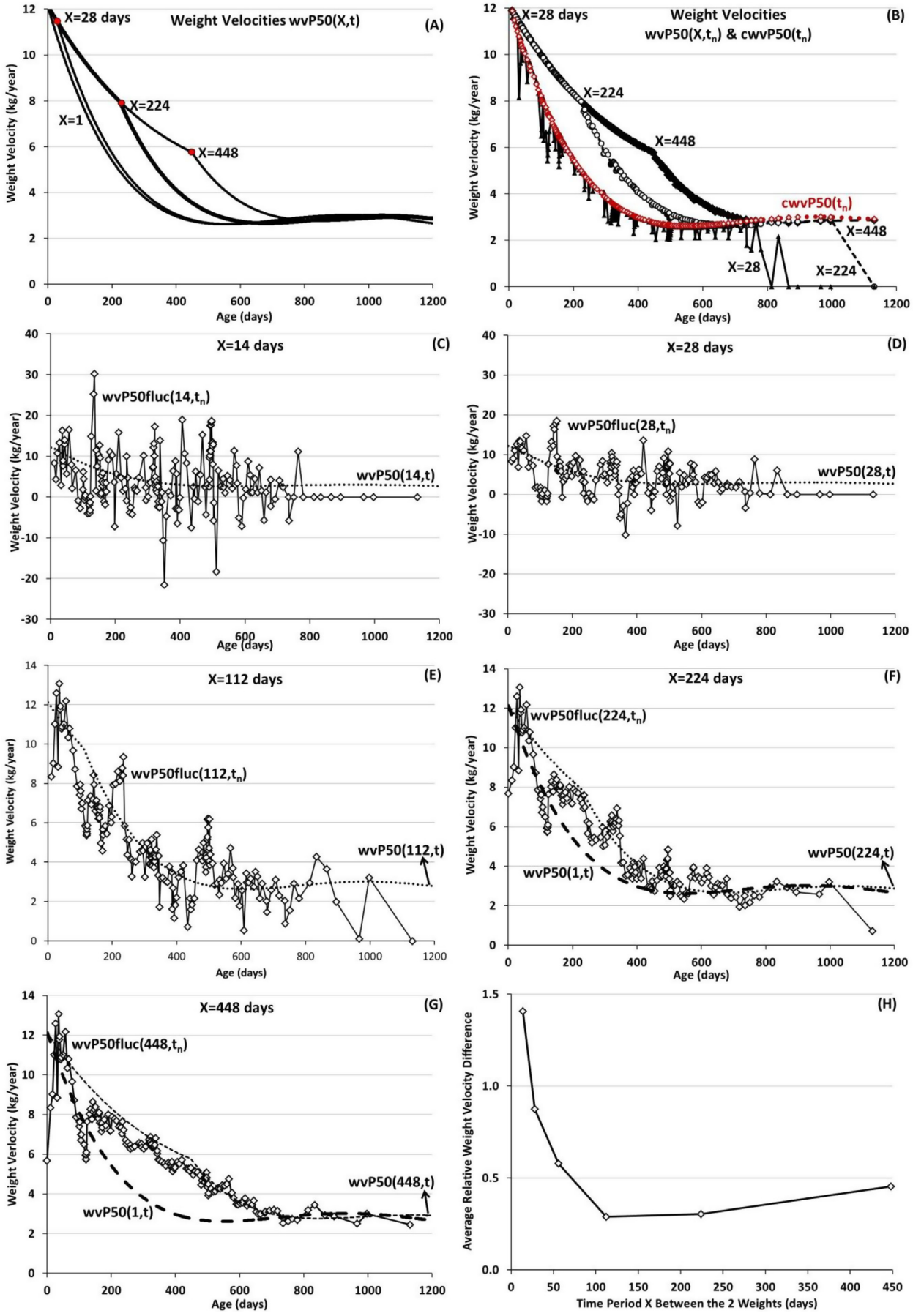
$$ARWVD(X) = \frac{1}{N}\sum_{n=0}^{N}\left|\frac{wvP50fluc(X,\,t_n) - cwvP50(t_n)}{cwvP50(t_n)}\right| \tag{4}$$

We discarded age points t_n where $cwvP50(t_n)$ was zero. Parameter $ARWVD(X)$ is 0 when *wvP50fluc(X,*t_n) = cwv*P50(*t_n) for all ages t_n. So, in reality, it is always positive and measures the combined influences of weight fluctuations in the *wvP50fluc(X,*t_n) curve and the effects of ϵ in both curves on weight velocity inaccuracy.

Evaluation of the hypothesis

Two-time-points weight velocity calculations

Using a selection of the calculations we have made, Fig. 2A shows 4 examples of Eqs. (1) applied to the P50(t) weight curve, using X = 1, 28, 224 and 448 days. The results for X = 1 day represent the exact P50(t) weight velocity curve here, the results for 448 days were shown previously (Fig. 2 of [2]). As expected, compared to the exact weight velocities (X = 1 day), the degree of error increases with increasing values of X, with virtually excellent agreement for X ≤ 28 days towards virtually no agreement for X = 448 days until about two years of age. Fig. 2B shows Eqs. (1) applied to the P50(t_n) curve for X = 28, 224, and 448 days, as well as the cwvP50(t_n) curve (open lozenges), which basically coincides at all ages where t = t_n with the wvP50(X = 1,t) curve. The weight velocities of Fig. 2B basically follow the related wvP50(X,t)



(caption on next page)

Fig. 2. (A) Weight velocities of the P50(t) curve, for X = 1, 28, 224 and 448 days (black lines). The results for X = 1 day are the exact weight velocities here. (B) Weight velocities of the P50(t_n) curve for the same X-values and including the cwvP50(t_n) curve (red dashed line with open red lozenges). (C–G) Weight velocities of the wvP50(X,t) and wvP50fluc(X,t_n) curves for X = 14, 28, 112, 224 and 448 days, for X = 224 and 448 days also including the wvP50(1,t) curve (thick dashed line). (H) The Average Relative Weight Velocity Difference between wvP50fluc(t_n) and cwvP50(t_n), Eq. (4).

curves except when consecutive weight measurements are separated by age intervals comparable to X. This finding implies that the influence of parameter ∈ in Eq. (1b) on weight velocities can become large, for X = 28 days visible at all ages but particularly beyond 750 days when the weights become sparsely available. For X = 224 days, the influence of ∈ is visible beyond 1000 days, while it is not observed for X = 448 days. Figs. 2C–G show the predictions of Eqs. (1) for weight curves P50(t) and P50fluc(t_n) for X = 14, 28, 112, 224 and 448 days. The introduction of weight fluctuations in the P50fluc(t_n) curve results in largely unpredictable weight velocities calculated for shorter values of X, with over- as well as underestimation of the P50(t) weight velocity curve by factors of up to 4 for X = 14 days. For larger X values, the level of over- as well as underestimation of the P50fluc(t_n) weight velocities diminishes. In Figs. 2F,G we also included the wvP50(X = 1,t) curve, the exact P50(t) weight velocities; in the Figs. 2C–E this exact curve is not included because it is virtually undistinguishable from the wvP50(X,t) curves. Fig. 2H shows that the Average Relative Weight Velocity Difference (ARWVD) between wvP50fluc(t_n) and cwvP50(t_n), Eq. (4), decreases from 1.41 to 0.29 for X values from 14 up to 112 days, indicative of a reducing influence of weight fluctuations, and then increases again to 0.45 for X = 448 days, indicative of an increasing influence of ∈ in Eq. (1b).

Our results of Fig. 2 show that none of the X-values used gives accurate estimated weight velocities when Eqs. (1) are applied to the P50(t), P50(t_n) and P50fluc(t_n) weight curves. For X < 112 days, the fluctuating weight variations create unacceptable erroneous weight velocity amplitudes, positive as well as negative, to our best knowledge for the first time identified in Fig. 2H. For X > 112 days, weight velocities become clinically unacceptable erroneous due to increasing deviation from the P50(t) exact weight velocity curve, Fig. 2A and F, G, already explained previously [2]. Nevertheless, Fig. 2H suggests that for this set of weights at ages t = t_n, an optimal X-value of between 100 and 200 days exists for weight velocity calculations. This implies that the combined influences of fluctuating weights and ∈ are at a minimum for producing erroneous weight velocities.

Other weight velocity equations

Our hypothesis basically states that the reasons for failure of the two-time-points weight velocity Eqs. (1) hold for all other weight velocity equations. The basis for this statement is the fact that the two causes for failure are just natural occurring events, namely typical physiological and pathophysiological weight fluctuations of toddlers, and variable time periods between consecutive weight measurements. These same mechanisms will cause failure in any two-time-points method of calculating weight velocity.

Discussion

The fact that Eqs. (1) calculate unreliable weight velocities for infants up till three years of age implies that two-time-points weight velocities, and any other equation used for weight velocity assessment, would need serious and currently unidentified adjustments if intended for clinical use. Interestingly, if also intended for commercial use, there are consequences in view of the new European Union regulations for in vitro diagnostic medical devices, which will be enacted in 2020 [7]. Software that is intended for diagnosis of a medical disease is considered a medical device [7] and thus must meet the strict requirements of medical devices. These requirements include the following. First, the manufacturer identifies the risk class of the software. If the software operates under a higher risk classification, a notified body needs to be involved and clinical evidence must be submitted to acquire a conformity rating. Second, the manufacturer needs to show that the software meets the standards for safety and performance, including technical documentation and the quality management system used. Third, clinical evidence needs to be acquired under strict rules following annex 61 of the regulations, for instance using the IDEAL recommendations for study design [8] and STARD guidelines for reporting diagnostic criteria [9]. Finally, the risk identified for the users carry over to the manufacturer: users need to be informed about the limitations of the software. Obviously, these regulations make it questionable whether any weight velocity software can fulfil these requirements of the European Union.

We emphasize that software-mediated analysis of clinical diagnostic procedures with restricted accuracy in outcome is not limited to the study of infant growth. Semiautomatic assessment of carotid artery stenoses on CT angiography is another example where the software output tends to overestimate the minimal stenosis diameter in case of calcified plaques and the opposite when no calcifications are present [10]. Consequently, a suboptimal treatment method could have been chosen in about 15% of cases. Importantly, the authors also concluded that "*Apparently, the semiautomated method tempts an observer to accept the proposed measurements as true and makes the radiologist less "aware".*" which they confirmed in a separate study (personal communication). Their conclusions support our previous observation that GAVE, applied by a pediatrician to the weights of an infant boy, gave erroneous outcomes with serious consequences for the boy and his family [2].

Finally, extensive literature on weight velocity measurements in children exists, e.g. [11–20], with various formulas for weight velocity related parameters. Three of these equations, but not Eqs. (1), were tested for very preterm and/or very low birthweight babies [10,16], in part also for commercial use (an exponential weight versus age curve was filed for obtaining a patent [13]). However, and despite abundant literature, only the recent paper by Fenton et al. [20] shows similarities with our results. These authors tested the precision of weight velocities in 103 very preterm infants, birthweight range 507–1440 g (for comparison, the P8/-2SD birthweight at 40 weeks gestation is 2400 gr), for X = 1 and 7 days using the average two-point method defined as [(W_2-W_1)/W_{av}/X] (gr/kg/day), where W_1 is the earlier weight, W_2 the final weight separated by time period X, and $W_{av} = (W_2 + W_1)/2$ is their average weight in kg. They found a huge range of weight velocities, comparable to our findings in Figs. 2C,D with strong positive and negative outliers, and concluded that "*Weight gain velocity calculated over 5 to 7 + days have lower variability and less noise than shorter periods*." We hypothesize that lack of access to the weight curve without natural weight variations, as we had, precluded them from explaining their findings, as we did in Fig. 2H.

Conclusions

The two-weight-points method of Eqs. (1), and hence also any other weight velocity equation, are unable to calculate weight velocities accurately from exact as well as realistically adapted P50 weight curves for toddlers, at least until three years of age, and irrespective of time separation X. First, shorter X-values than 100 days produce increasingly larger positive and negative variations in weight velocity amplitudes, and thus make software calculated weight velocity analysis increasingly unusable for all ages considered here (up till 3 years). These outliers are due to a combination of events that are typical for the weights of infants, i.e. natural (patho)physiological weight fluctuations and sparsely

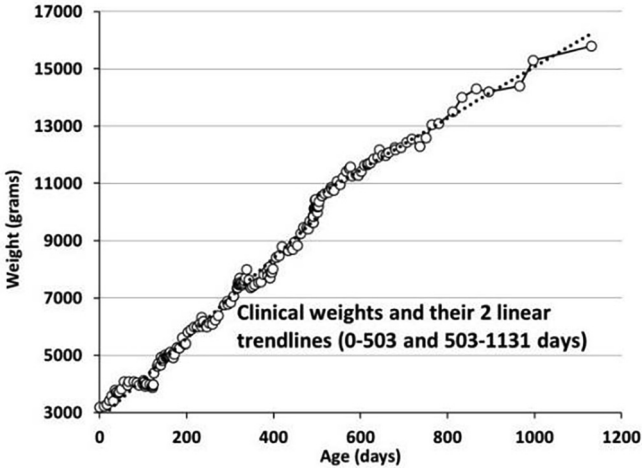


Fig. 3. The clinical weight curve [2] with 2 trendlines (dashed lines): $W1 = 14.081 \cdot t + 2778.6$ g (t = 0–503 days) and $W2 = 8.9961 \cdot t + 6074.2$ g (503–1131 days), see Appendix.

measured weights. Second, X-values > 100 days produce increasingly larger deviations from the exact P50(t) weight velocities as shown previously [2]. In addition, our results, those of Fenton et al. [20] and Marquering's findings that use of diagnostic software make clinicians less critical in identifying clinical facts [10], suggest that computer analysis of infant weight growth should be abandoned in medicine. Practically, if still remaining in use, it unquestionably suggests that when (commercial) software predicts a serious problem, careful clinical checking should be mandatory, e.g. by comparing predicted and exact weight velocity at a certain age. The latter can be determined by the mathematical first derivative of standard weight curves at that age [2].

Acknowledgements

We thank dr. Nick van der Beek, MultiCare, Hilversum, the Netherlands, for his comments on the manuscript and the European Union regulations, and Dr. Henk A. Marquering, Departments of Biomedical Engineering & Physics and Radiology, Amsterdam UMC, for discussions on software analysis for medical diagnostic procedures.

Conflict of interest statement

The authors declare that they have no conflict of interest.

Appendix A

Fluctuating weight variations included in the weight curve P50fluc(t_n) (Fig. 1) was constructed from the clinical weights (Fig. 3) by incorporated 2 linear trendlines in Excel, from ages 0–503 and 503–1131 days. The ratio between clinical and trendline weight points at $t = t_n$ was used as a multiplication factor of the P50(t_n) weights to obtain the P50fluc(t_n) weight curve.

References

[1] https://www.growthanalyser.org.
[2] van Gemert MJC, Bruijninckx CMA, van Leeuwen TG, Neumann HAM, Sauer PJJ. Limitations of weight velocity analysis by commercial computer program Growth Analysis Viewer Edition. Ann Biomed Eng 2019;47(1):297–305.
[3] van Gemert MJC, Vlaming M, Osinga E, Bruijninckx CMA, Neumann HAM, Sauer PJJ. Pediatric condition falsification by misjudged weight growth from the curve of measured weights. Am J Case Rep 2018;19:752–6.
[4] Ghaemmaghami P, Ayatollahi SM, Alinejad V, Haem E. Longitudinal standards for growth velocity of infants from birth to 4 years born in West Azerbaijan Province of northwest Iran. Epidem Health 2015;37:e2015029.
[5] Robertson C. Catch-up growth among very-low-birth-weight preterm infants: a historical perspective (Editorials). J Pediatr 2003;143:145–6.
[6] Gerver WJM, de Bruin R. Paediatric Morphometrics. 2nd ed. Maastricht: University Press Maastricht; 2001.
[7] https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32017R0746&from=EN.
[8] Pennell CP, Hirst AD, Campbell WB, Sood A, Agha RA, Barkun JS, et al. Practical guide to the Idea, Development and Exploration stages of the IDEAL Framework and Recommendations. Br J Surg 2016;103(5):607–15.
[9] Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L, Lijmer JG, Moher D, Rennie D, de Vet HCW, Kressel HY, Rifai N, Golub RM, Altman DG, Hooft L, Korevaar DA, Cohen JF, For the STARD Group. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. BMJ 2015;2015(351):h5527.
[10] Marquering HA, Nederkoorn PJ, Smagge L, Gratama van Andel HA, van den Berg R, Majoie CB. Performance of semiautomatic assessment of carotid artery stenosis on CT angiography: clarification of differences with manual assessment. Am J Neurorad 2012;33:747–54.
[11] Brandt I. Growth dynamics of low-birth-weight infants with emphasis on the perinatal period. In: Falkner F & Tanner JM (eds). Human growth. Springer, Boston. Chapt. 19, pp 557-617, 1978.
[12] Ehrenkranz RA, Younes N, Lemons JA, Fanaroff AA, Donovan EF, Wright LL, et al. Longitudinal growth of hospitalized very low birth weight infants. Pediatr 1999;104:280–9.
[13] Patel AL, Engstrom JL, Meier PP, Kimura RE. Accuracy of methods for calculating postnatal growth velocity for extremely low birth weights infants. Pediatr 2005;116:1466–73.
[14] Danner E, Joeckel S, Michalak S, Phillips S, Goday PS. Weight velocity in infants and children. Nutr Clin Pract 2009;24(1):76–9.
[15] Wu C, Löfqvist C, Smith LEH, VanderVeen DK, Hellström A. Importance of early postnatal weight gain for normal retinal angiogenesis in very preterm infants. A multicenter study analyzing weight velocity deviations for the prediction of retinopathy of prematurity. Arch Ophthalmol 2012;130(8):992–9.
[16] Binenbaum G. Algorithms for the prediction of Retinopathy of Prematurity based upon postnatal weight gain. Clin Perinatol 2013;40(2):261–2.
[17] Kruithof CJ, Gishti O, Hofman A, Gaillard R. Jaddoe VWV. Infant weight growth velocity patterns and general and abdominal adiposity in school-age children. The Generation R Study. Eur J Clin Nutr 2016;70(10):1144–50.
[18] Marinkovic T, Toemen L, Kruithof CJ, Reiss I, van Osch-Gevers L, Hofman A, et al. Early infant growth velocity patterns and cardiovascular outcomes in childhood. J Pediatr 2017;186:57–63.
[19] Fenton TR, Anderson D, Groh-Wargo S, Hoyos A, Ehrenkranz RA, Senterre T. An attempt to standardize the calculation of growth velocity of preterm infants – evaluation of practical bedside methods. J Pediatr 2018;196:77–83.
[20] Fenton TR, Senterre T, Griffin IJ. Time interval for preterm infant weight gain velocity calculation precision. Arch Dis Child Fetal Neonatal Ed 2019;104(2):F218–9.