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

van Gemert, Martin J. C.; Bruijinckx, Cornelis M. A.; Neumann, H. A. Martino; Sauer, Pieter J. J.; de Bruin, D. Martijn; van Leeuwen, Ton G.

Published in:
Medical Hypotheses

DOI:
10.1016/j.mehy.2019.109234

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2019

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):
https://doi.org/10.1016/j.mehy.2019.109234

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.
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. Brujininkx b, H.A. Martino Neumann c, Pieter J.J. Sauer d, D. Martijn de Bruin e, Ton G. van Leeuwen a

Abstract
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 \[ \frac{\text{Weight}(t) - \text{Weight}(t - X)}{X} \], 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 \( X = 1,2,8,224,448 \) 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 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 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,
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 O5D) 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 O5D). We included two realistic variations in the P50 curve. First, P50(\( t_n \), \( t_n \) is the age of the \( n \)th 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(tn) and P50fluc(tn)**

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(\( t_n \), \( t_n \) 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(\( t_n \) 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 (\( t+X+\epsilon \)), the next available weight at day (\( t+X+\epsilon \)) 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

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

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

We varied time separation \( X \) as

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

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

**Consecutive weight velocities of the P50(tn) curve: cwvP50(tn)**

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

\[
\text{cwvP50}(t_n) = \frac{\text{Weight}(t_n) - \text{Weight}(t_n-1)}{(t_n - t_n-1)/365.25} \tag{3}
\]

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 \( \text{wvP50}(X,t_n) \).

The contribution to the inaccuracy of weight velocity assessment of:

1. the variable ages that separate consecutive weights, particularly through parameter \( \epsilon \) in Eq. (1b), and
2. the weight fluctuations of toddlers, was evaluated by comparing weight velocities \( \text{wvP50fluc}(X,t_n) \), with both elements included, and the gold standard \( \text{cwvP50}(t_n) \), 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 \( \text{wvP50fluc}(X,t_n) - \text{cwvP50}(t_n) / \text{cwvP50}(t_n) \), averaged over all \( N = 180 \) weight velocity data points, thus

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

We discarded age points \( t_n \) where \( \text{cwvP50}(t_n) \) was zero. Parameter ARWVD(X) is 0 when \( \text{wvP50fluc}(X, t_n) = \text{cwvP50}(t_n) \) for all ages \( t_n \). So, in reality, it is always positive and measures the combined influences of weight fluctuations in the \( \text{wvP50fluc}(X, t_n) \) curve and the effects of \( \epsilon \) 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 \( X \leq 28 \) 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 \( X = 28, \ 224, \ 448 \) days, as well as the \( \text{cwvP50}(t_n) \) curve (open lozenges), which basically coincides at all ages where \( t = t_n \) with the \( \text{cwvP50}(X = 1,t) \) curve. The weight velocities of Fig. 2B basically follow the related \( \text{wvP50}(X,t) \)
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(tn) curve for the same X-values and including the cwvP50(tn) curve (red dashed line with open red lozenges). (C–G) Weight velocities of the wvP50(X,t) and wvP50flu(X,tn) 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 wvP50flu(X,tn) and cwvP50(tn), Eq. (4).

Discussion

The effect that Eqs. (1) calculate unreliable weight velocities for infants up till three years of age implies that two-point weights 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 fulfill 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 pediatrican 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 g), for X = 1 and 7 days using the average two-point method defined as |(W2 - W1)/Wav/X| (gr/kg/day), where W1 is the earlier weight, W2 the final weight separated by time period X, and Wav = (W2 + W1)/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 fluctuations and sparsely

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 pathophysiologic 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.

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
Appendix A

Fluctuating weight variations included in the weight curve P50flu(t\textsubscript{0}) (Fig. 1) was constructed from the clinical weights (Fig. 3) by incorporated 2 linear trends in Excel, from ages 0–503 and 503–1131 days. The ratio between clinical and trendline weight points at t = t\textsubscript{n} was used as a multiplication factor of the P50(t\textsubscript{0}) weights to obtain the P50flu(t\textsubscript{0}) weight curve.

References


