Fit to WHO weight standard of European infants over time

Citation for published version:
Levin, D, Marryat, L, Cole, TJ, McColl, J, Harjunmaa, U, Ashorn, P & Wright, C 2016, 'Fit to WHO weight standard of European infants over time' Archives of Disease in Childhood, vol. 101, no. 5, pp. 455-60. DOI: 10.1136/archdischild-2015-309594

Digital Object Identifier (DOI):
10.1136/archdischild-2015-309594

Link:
Link to publication record in Edinburgh Research Explorer

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

Published In:
Archives of Disease in Childhood

Publisher Rights Statement:
This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
Fit to WHO weight standard of European infants over time

Daniel Levin,1 Louise Marryat,1 Tim J Cole,2 John McColl,1 Ulla Harjunmaa,3,4 Per Ashorn,3,4 Charlotte Wright5

ABSTRACT

Objectives The 2006 WHO growth charts were created to provide an international standard for optimal growth, based on healthy, breastfed populations, but it has been suggested that Northern European children fit them poorly. This study uses infant weight data spanning 50 years to determine how well-nourished preschool children from different eras fit the WHO standard, and discuss the implications of deviations.

Design Four longitudinal datasets from the UK and one from Finland were used comprising over 8000 children born between 1959 and 2003. Weights from birth to 2 years were converted to age-sex-adjusted Z scores using the WHO standard and summarised using Generalized Additive Models for Location, Scale and Shape.

Results Weights showed a variable fit to the WHO standard. Mean weights for all cohorts were above the WHO median at birth, but dipped by up to 0.5 SD to a nadir at 8 weeks before rising again. Birth weights increased in successive cohorts and the initial dip became slightly shallower. By age 1 year, cohorts were up to 0.75 SD above the WHO median, but there was no consistent pattern by era.

Conclusions The WHO standard shows an acceptable, but variable fit for Northern European infants. While birth weights increased over time, there was, unexpectedly, no consistent variation by cohort beyond this initial period. Discrepancies in weight from the standard may reflect differences in measurement protocol and trends in infant feeding practice.

INTRODUCTION

Growth charts are widely used in child health to identify undernutrition and overnutrition.1 Many countries produce their own charts describing how local children grow. However, these do not necessarily characterise healthy growth; in some countries average weight will be low due to undernutrition, while in high-income countries higher average weights reflect obesity. In the past, charts were based predominantly on bottle-fed infants and did not adequately reflect the growth of breastfed infants who should be the physiological norm.2 Thus, the WHO developed new growth charts based on healthy breastfed children living in optimal circumstances in six world regions.3 The WHO showed that linear growth differed little between the six datasets4 and thus argued that they could be used to define how all children aged 0–5 years should grow, whatever their ethnic origin. However, weights by country for the WHO dataset have never been published. Since the publication of the standard, studies in unselected, healthy, non-deprived populations such as the UK,3 USA,6 Canada,7 Norway, Belgium,8 Italy, Argentina9 and Denmark have generally found a close fit for length, but a tendency for children to become heavier than the WHO standard after the first 6 months. Some authors have argued that these discrepancies represent a fundamental difference that renders the standard unsuitable for high-income countries.8 Others have suggested that variations in fit are to be expected if the WHO charts represent optimal rather than average growth, as few children will be breastfed to age 1 year10 and that the higher weights reflect rising rates of obesity at all ages.3

If the mismatch is due to obesity, it ought to be less evident in historic datasets from eras when obesity was less common. Conversely, if it reflects the nature, prevalence and duration of formula feeding, the fit should be better in more recent cohorts as breastfeeding has increased.

Aims

Our aims are to explore a portfolio of weight datasets from the past 50 years (i) to determine how well real populations of well-nourished preschool children from different eras fit the 2006 WHO growth standard at different ages during infancy;
and (ii) to explore how trends differ by era of the cohort and associated infant feeding patterns.

**METHODS**

**Datasets**

Data came from existing longitudinal growth studies, retrieved mainly from routine records. They had already been cleaned, checked and analysed for other purposes, with four studies already published. Details of the five studies are as follows:

**Widdowson study (1959)**

In a study set up by Dr Elsie Widdowson, routine weights of 1094 babies born in 1959–1965 were obtained contemporaneously from the records of 10 Cambridge Child Welfare Clinics. Weights were measured by clinical staff approximately monthly in the first year of life, with a maximum of 13 weights per child. Of the possible 14,222 measurements, all but 864 (6%) were collected. Although the data were cleaned and analysed at the time of collection, the results were never published.

**Cambridge Infant Growth study (1984)**

Cambridge Infant Growth Study (CIGS) was a research cohort of 255 babies recruited in 1984–1987. Infants were sampled in four cohorts from lists of Cambridge city mothers booked to deliver in particular months, with some filtering by midwives. Measurements were taken mainly by one highly trained auxologist every 4 weeks from 4 to 52 weeks, then at 15, 18, 24, 30, 36 and 48 months. Weight, length, head and arm circumferences, triceps and subscapular skinfolds were measured at each visit. Of the possible 14,222 measurements all but 864 (6%) were collected. Although the data were cleaned and analysed at the time of collection, the results were never published.

**Newcastle Growth and Development study (1987)**

This dataset comprises the 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 by clinical staff in infancy were retrieved from baby clinic records, and 3060 of the babies had at least two weights.

**Gateshead Millennium study (1999)**

Gateshead Millennium Study (GMS) is 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 were retrieved from baby clinic records. There was a mean of 13 weights per child in the first year. Research nurses measured 830 infants at 13 months.

**Tampere study, Finland (2003)**

This dataset comprises the routine heights and weights of 2809 children aged 0–4 years born between October 2003 and September 2004 who attended child health clinics in Tampere. Children were weighed by clinical staff on electronic scales. Up to 16 scheduled events were recorded per child at birth, 1–2 weeks, 6–8 weeks, 2, 3, 4, 5, 6, 8, 10, 12, 18, 24, 36 and 48 months. There was a mean of 12 measurements per child.

**Analysis**

The datasets were cleaned and weights converted into Z scores relative to the WHO growth standard. Data beyond 2 years were excluded, when numbers were low and bias was likely. The measurements at birth and 6–8 weeks, 3, 6, 9, 12, 18 and 24 months were summarised by age, sex and cohort (table 1). The mean Z score, SD, skewness and kurtosis were modelled as functions of age for each dataset by sex, using Generalized Additive Models for Location, Scale and Shape (GAMLSS) as implemented in the GAMLSS package in R. Multiple GAMLSS models were fitted using different hyperparameters and distribution families. In the final models, based on the Box-Cox power exponential family (BCPE), the mean Z score was allowed to vary with age in each of the datasets, whereas the SD, skewness and kurtosis were constrained to be constant as exploratory analyses showed this made little difference to the model fit as determined using the Bayesian Information Criterion. The skewness adjustment ensured that the mean and median were effectively the same. The BCPE requires values to be positive, so all Z scores had 10 added to them prior to analysis, and 10 was then subtracted from the mean curves. The

---

**Table 1  Number of observations at target ages by dataset**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Birth</td>
<td>M</td>
<td>573</td>
<td>142</td>
<td>1589</td>
<td>484</td>
<td>929</td>
</tr>
<tr>
<td></td>
<td></td>
<td>F</td>
<td>519</td>
<td>122</td>
<td>1582</td>
<td>477</td>
<td>889</td>
</tr>
<tr>
<td>7</td>
<td>5–9 weeks</td>
<td>M</td>
<td>452</td>
<td>135</td>
<td>1201</td>
<td>388</td>
<td>1202</td>
</tr>
<tr>
<td></td>
<td></td>
<td>F</td>
<td>410</td>
<td>116</td>
<td>1217</td>
<td>379</td>
<td>1145</td>
</tr>
<tr>
<td>13</td>
<td>2–4 months</td>
<td>M</td>
<td>573</td>
<td>137</td>
<td>1357</td>
<td>392</td>
<td>1220</td>
</tr>
<tr>
<td></td>
<td></td>
<td>F</td>
<td>520</td>
<td>118</td>
<td>1365</td>
<td>389</td>
<td>1189</td>
</tr>
<tr>
<td>26</td>
<td>5–7 months</td>
<td>M</td>
<td>570</td>
<td>135</td>
<td>1186</td>
<td>240</td>
<td>1192</td>
</tr>
<tr>
<td></td>
<td></td>
<td>F</td>
<td>516</td>
<td>117</td>
<td>1212</td>
<td>247</td>
<td>1174</td>
</tr>
<tr>
<td>39</td>
<td>8–10 months</td>
<td>M</td>
<td>569</td>
<td>134</td>
<td>898</td>
<td>222</td>
<td>993</td>
</tr>
<tr>
<td></td>
<td></td>
<td>F</td>
<td>515</td>
<td>116</td>
<td>893</td>
<td>233</td>
<td>962</td>
</tr>
<tr>
<td>52</td>
<td>11–14 months</td>
<td>M</td>
<td>501</td>
<td>131</td>
<td>1068</td>
<td>308</td>
<td>1161</td>
</tr>
<tr>
<td></td>
<td></td>
<td>F</td>
<td>476</td>
<td>116</td>
<td>1109</td>
<td>300</td>
<td>1134</td>
</tr>
<tr>
<td>78</td>
<td>15–21 months</td>
<td>M</td>
<td>–</td>
<td>125</td>
<td>642</td>
<td>61</td>
<td>1132</td>
</tr>
<tr>
<td></td>
<td></td>
<td>F</td>
<td>–</td>
<td>116</td>
<td>670</td>
<td>66</td>
<td>1098</td>
</tr>
<tr>
<td>104</td>
<td>22–30 months</td>
<td>M</td>
<td>–</td>
<td>124</td>
<td>122</td>
<td>–</td>
<td>1076</td>
</tr>
<tr>
<td></td>
<td></td>
<td>F</td>
<td>–</td>
<td>115</td>
<td>115</td>
<td>–</td>
<td>1068</td>
</tr>
</tbody>
</table>

CIGS, Cambridge Infant Growth Study; GDS, Growth and Development Study; GMS, Gateshead Millennium Study.
Widdowson also showed the steepest rise in Z score after the
in the earlier cohorts (table 3). The earliest dataset
and 0.5 SDs from birth to 7 weeks, the fall tending to be greater
sively shallower over time. Mean Z score fell by between 0.2
8 weeks. Boys and girls followed broadly the same pattern.

GAMLSS models were plotted by cohort on separate charts by
median. A mean Z score of 0, SD 1, skewness 1 and kurtosis (tau)
sex, along with the constant SD (sigma), skewness (nu) and kur-

There is no standard definition of what constitutes a good or
fit to a growth reference. In this study fit was defined in
terms of mean weight Z score, measured in fractions of a centile
channel width relative to the WHO median, where one channel
width=0.67 SDs. An excellent fit was defined as a difference of
no more than ¼ of a channel width (0.17 SD) and a poor
fit as greater than a channel width (0.67 SD).

RESULTS

Mean birthweight Z scores in the five datasets were all positive
and close to zero (table 2). They were progressively higher in
later years, particularly in the boys, rising from 0.03 in 1959 to
0.37 in 2003. For all cohorts the fit was adequate in early
infancy, mostly staying within half a channel width of the
median, but by 1 year boys and girls in Widdowson and boys in
Tampere fitted poorly, more than a channel width above the
median.

The smoothed curves of mean Z score versus age from the
GAMLSS models were plotted by cohort on separate charts by
sex, along with the constant SD (sigma), skewness (nu) and kur-
tosis (tau) (figure 1). The curves provide a visual assessment of
the fit to the standard by cohort. All datasets showed consistent
differences relative to the standard. The mean Z scores all
tended to fall in the early weeks, with the four UK datasets
ending up below zero. This steep fall was followed by a slightly
less steep rise, creating a ‘trough’ in each curve at around
8 weeks. Boys and girls followed broadly the same pattern.

The depth of the trough, like birthweight, became progres-
sively shallower over time. Mean Z score fell by between 0.2
and 0.5 SDs from birth to 7 weeks, the fall tending to be greater
in the earlier cohorts (table 3). The earliest dataset
(Widdowson) also showed the steepest rise in Z score after the

models for each cohort were plotted as mean Z score versus age
in boys and girls, along with the constant SD, skewness and kur-
tosis, to compare how well the cohorts fitted the WHO 2006
standard. A mean Z score of 0, SD 1, skewness 1 and kurtosis 2
indicates a perfect fit, whereas a mean Z score above 0 is
heavier than the WHO standard, and below is lighter.

There is no standard definition of what constitutes a good or
poor fit to a growth reference. In this study fit was defined in

The strengths of this study are that the five datasets represent
weight gain before and during the obesity epidemic. The data-
sets are of high quality and all but CIGS are representative of
their populations, with high recruitment rates. The GAMLSS
modelling approach makes maximum use of the available data
and controls for potential bias. A weakness is that the most con-
temporary study is from Finland, not the UK, where growth pat-
terns may differ systematically.

Overall, most cohorts were within a channel width of the
WHO median and many showed an excellent fit in the early
weeks, though less so beyond this point. It seems unlikely that
these later differences could be simply genetic, since two of the
six WHO datasets were of predominantly Northern European
origin and they showed no consistent differences in stature." The
lack of variation by era argues against obesity being the
explanation.

The increase in birthweight over time is consistent with
research in other high-income countries. This is generally
thought to reflect less maternal cigarette smoking and more
maternal obesity and gestational diabetes. It has been sug-

gested that the higher birthweight seen for all these cohorts
compared with the WHO standard may be explained by prior
maternal undernutrition in some of the cohorts used to develop
the WHO standard. However, a recent study suggested that

**DISCUSSION**

**Table 2** Mean (SD) weight Z scores at target ages by dataset

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>M</td>
<td>0.03 (1.11)</td>
<td>0.06 (0.97)</td>
<td>0.12 (1.01)</td>
<td>0.18 (1.05)</td>
<td>0.37 (1.08)</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>0.10 (1.01)</td>
<td>0.30 (0.89)</td>
<td>0.08 (1.07)</td>
<td>0.17 (1.11)</td>
<td>0.40 (1.1)</td>
</tr>
<tr>
<td>7</td>
<td>M</td>
<td>−0.40 (1.02)</td>
<td>−0.28 (0.99)</td>
<td>−0.25 (0.95)</td>
<td>−0.16 (0.96)</td>
<td>0.14 (1.05)</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>−0.30 (0.83)</td>
<td>−0.19 (0.83)</td>
<td>−0.30 (0.91)</td>
<td>−0.17 (0.9)</td>
<td>0.12 (0.95)</td>
</tr>
<tr>
<td>13</td>
<td>M</td>
<td>−0.33 (1.07)</td>
<td>−0.34 (0.97)</td>
<td>−0.21 (0.98)</td>
<td>−0.16 (0.97)</td>
<td>0.17 (1.03)</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>−0.17 (0.83)</td>
<td>−0.22 (0.86)</td>
<td>−0.29 (0.9)</td>
<td>−0.18 (0.9)</td>
<td>0.12 (0.91)</td>
</tr>
<tr>
<td>26</td>
<td>M</td>
<td>0.28 (0.99)</td>
<td>−0.09 (0.98)</td>
<td>0.15 (1.01)</td>
<td>0.30 (0.99)</td>
<td>0.46 (1.07)</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>0.42 (0.82)</td>
<td>0.05 (0.87)</td>
<td>0.05 (0.91)</td>
<td>0.21 (0.93)</td>
<td>0.38 (0.93)</td>
</tr>
<tr>
<td>39</td>
<td>M</td>
<td>0.62 (1)</td>
<td>0.05 (0.97)</td>
<td>0.38 (1.05)</td>
<td>0.56 (1.01)</td>
<td>0.56 (1.09)</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>0.67 (0.8)</td>
<td>0.21 (0.86)</td>
<td>0.23 (0.93)</td>
<td>0.43 (0.92)</td>
<td>0.48 (0.96)</td>
</tr>
<tr>
<td>52</td>
<td>M</td>
<td>0.69 (0.97)</td>
<td>0.16 (0.97)</td>
<td>0.53 (1.01)</td>
<td>0.55 (1.06)</td>
<td>0.60 (1.01)</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>0.76 (0.81)</td>
<td>0.26 (0.83)</td>
<td>0.38 (0.89)</td>
<td>0.51 (0.91)</td>
<td>0.52 (0.91)</td>
</tr>
<tr>
<td>78</td>
<td>M</td>
<td>–</td>
<td>0.20 (0.94)</td>
<td>0.46 (1)</td>
<td>0.44 (0.97)</td>
<td>0.62 (0.96)</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>–</td>
<td>0.23 (0.76)</td>
<td>0.27 (0.95)</td>
<td>0.41 (1.01)</td>
<td>0.56 (0.89)</td>
</tr>
<tr>
<td>104</td>
<td>M</td>
<td>–</td>
<td>0.04 (0.96)</td>
<td>0.38 (1.12)</td>
<td>–</td>
<td>0.52 (0.96)</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>–</td>
<td>0.18 (0.71)</td>
<td>0.12 (1.12)</td>
<td>–</td>
<td>0.53 (0.91)</td>
</tr>
</tbody>
</table>

Fit: excellent (<0.17 SD), poor (>0.67 SD).
CIGS, Cambridge Infant Growth Study; GDS, Growth and Development Study; GMS, Gateshead Millennium Study.
birth weights across diverse countries, including some of those included in the WHO Multicentre Growth Reference Study (MGRS) cohorts, were not significantly different by country. Though born heavier, the infants here initially lost weight relative to the WHO standard and then regained it, causing a trough to appear in the Z score growth curve. This has been described in other cohorts, and was influential in the UK rejecting the use of the WHO standard at birth. This pattern is the mirror image of the way breastfed infants used to grow on charts based mainly on bottle-fed infants and though seen in all datasets, its depth decreases over time. We do not have information about breastfeeding rates in the routine cohorts but most will have been lower than the WHO sample. However, breastfeeding initiation rates have risen in the UK over the past 50 years from 36% in 1970 to over 70% in the early 21st century (UK) while when the Tampere cohort were born, 93% of Finnish infants were initially breastfed. The CIGS dataset, a more selective research sample, showed the shallowest trough and had higher rates of initial breastfeeding (75%) and breastfeeding to 6 months (48%). This pattern by era may also reflect changes in formula milk composition which has increasingly mirrored the nutritional content of breast milk.

Table 3  Change in mean (SD) weight Z scores between target ages by dataset

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Change in SDS</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td>M</td>
<td>F</td>
<td>M</td>
<td>F</td>
</tr>
<tr>
<td>0–7</td>
<td>−0.43</td>
<td>−0.34</td>
<td>−0.37</td>
<td>−0.34</td>
<td>−0.23</td>
</tr>
<tr>
<td>7–52</td>
<td>1.09</td>
<td>0.44</td>
<td>0.78</td>
<td>0.71</td>
<td>0.46</td>
</tr>
</tbody>
</table>

CIGS, Cambridge Infant Growth Study; GDS, Growth and Development Study; GMS, Gateshead Millennium Study.
In contrast, the later rise relative to the standard did not show the same trend over time, suggesting that it may be unrelated to breastfeeding. While observational studies have found an association between use of breast milk substitutes and faster weight gain later in infancy, this effect was not seen in the Belarus trial of breastfeeding promotion. The sampling process used to construct the WHO standard could in principle mean that the MGRS cohorts were lighter than the general population. Children who were either outliers or not breastfed to 12 months were excluded from the sample and it has been shown that larger children are more likely to cease breastfeeding early. However, the WHO MGRS group did not find any systematic difference in size between those included and excluded (personal communication De Onis, email communication, 2006).

Measurement error might also play a role, since the MGRS cohorts were measured using the same research protocols, where all clothes were removed or adjusted for. The present cohorts mainly used routine data collected by health professionals; UK and Finnish guidelines recommend weighing infants naked up to the age of 2 years, but parents of older infants may be reluctant to comply, leading to higher weights at later ages. The CIGS cohort where all infants were weighed naked to a research protocol tracks closest to the WHO standard beyond 6 months. However, another study that compared routinely collected measurements with research measurements of the same children found little difference in weight.

CONCLUSIONS
The overall fit to the WHO weight standard of these cohorts ranged from excellent to adequate, so that use of the WHO standard is unlikely to introduce major bias in the assessment of individual children. Some of the more subtle variation in fit is likely to reflect variations in the levels of breast and formula feeding, as well as the composition of formula milk, in different eras. However, the lack of a consistent trend for weight gain after the initial weeks suggests that the higher average weight gain seen in North European cohorts cannot simply be explained by increasing rates of obesity in later childhood and adulthood.

Data sharing statement This paper uses secondary data analysis from five studies. To the authors’ knowledge none are currently publicly available.

REFERENCES


Fit to WHO weight standard of European infants over time

Daniel Levin, Louise Marryat, Tim J Cole, John McColl, Ulla Harjunmaa, Per Ashorn and Charlotte Wright

Arch Dis Child 2016 101: 455-460 originally published online February 16, 2016
doi: 10.1136/archdischild-2015-309594

Updated information and services can be found at:
http://adc.bmj.com/content/101/5/455

These include:

References
This article cites 32 articles, 16 of which you can access for free at:
http://adc.bmj.com/content/101/5/455#BIBL

Open Access
This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See:
http://creativecommons.org/licenses/by/4.0/

Email alerting service
Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Topic Collections
Articles on similar topics can be found in the following collections
- Open access (181)
- Infant nutrition (including breastfeeding) (405)
- Childhood nutrition (708)
- Reproductive medicine (938)

Notes

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/