Developing regional weight-for-age growth references for malaria-endemic countries to optimize age-based dosing of antimalarials

Daniel J Hayes, Stef van Buuren, Feiko O ter Kuile, D Mikis Stasinopoulos, Robert A Rigby & Dianne J Terlouw

Objective To derive regional weight-for-age growth references to help optimize age-based dosing of antimalarials in Africa, the Americas, South-East Asia and the Western Pacific.

Methods A weight-for-age database was constructed from pre-existing population-based anthropometric data obtained from household surveys and research groups. It contained data collected between 1995 and 2012 on 1,263,119 individuals (909,368 female, 353,751 male) older than 14 days and younger than 50 years in 64 malaria-endemic countries. Regional growth references were generated using a generalized additive model for location, scale and shape by combining data with varying distributions from a range of sources. Countries were weighted by their population at risk of malaria to enable references to be used in optimizing the dosing of antimalarials.

Findings Large differences in weight-for-age distributions existed between the regions and between the regions and global growth standards. For example, the average adult male from the Americas weighed 68.1 kg – 6.0 kg more than males in South-East Asia and the Western Pacific (average: 62.1 kg). For adult women, the difference was over 10.4 kg; the average was 60.4 kg in the Americas and 50.0 kg in South-East Asia and the Western Pacific.

Conclusion There were substantial variations in weight-for-age growth curves between malaria-endemic areas. The growth reference charts derived here can be used to guide the evidence-based optimization of age-based dosing regimens for antimalarials and other drugs often prescribed by age.

Introduction

Manufacturers of antimalarials recommend that dosing should be based on body weight but, in many low- and middle-income countries, the dose is frequently based on age, which is used as a proxy for body weight because these drugs are often sold over the counter or prescribed in settings without weighing facilities. Unfortunately, the effectiveness and safety of this approach has rarely been assessed and it has been shown that the use of inadequate, age-based dosing regimens is responsible for a considerable proportion of treatment failures.1–2 Since age-based dosing results in much more variability in drug intake than weight-based dosing, over- and underdosing inevitably occur. Substantial improvements could be made by optimizing age-based dosing regimens but this idea has received little attention from manufacturers or policy-makers. Accurate knowledge of the weight-for-age distribution of the population at risk of malaria is vital for establishing the optimal dosing regimen.

Previously, the optimal age-based dosing regimen and drug ratio of a fixed-dose combination of the antimalarials artesunate and amodiaquine for use in sub-Saharan Africa were predicted using weight-for-age data from 88,054 individuals in several African countries.3 A recent analysis of the efficacy of artesunate and amodiaquine using pooled data on 5,410 patients from 24 studies showed that administered doses of the fixed combination were significantly better with both weight- and age-based regimens than when non-fixed combinations were used. The fixed combination also showed a lower risk of recrudescence.4

Currently weight-for-age reference data are lacking for most middle- and low-income countries. The existing global growth standards are the World Health Organization (WHO) Child Growth Standards for children aged 0 to 59 months5 – based on the Multicentre Growth Reference Study6 – and the 2007 WHO growth reference for school-aged children and adolescents aged 5 to 19 years.7 Optimal growth can be assessed against these standards but they are inadequate for establishing optimal dosing of antimalarials as they do not describe how children actually grow at a particular time and place. Since growth varies between regions, regional or country-specific reference data would enable dosing regimens to be tailored to the population affected.

Over the past two decades, a wealth of population-representative anthropometric data has become publicly available from low- and middle-income countries. These data are largely based on national household surveys that monitor standard socioeconomic and health indicators for children younger than 5 years and women of reproductive age.5,6 Research surveys and monitoring and evaluation activities can also provide data on locally representative samples of school-age children, adolescents and adult males. Conventional growth curve modelling methods, such as the LMS method,8,9 and the more recent generalized additive model for location, scale and shape10 assume a single data source and distribution. We developed an extension of the generalized additive model using heterogeneous data from representative population samples that reflect different underlying population growth distributions.11

Here we present the first application of this extended generalized additive model. We compiled individual-level,
weight-for-age data from population-representative data sources in countries where malaria is endemic, principally in three areas: (i) the WHO African Region; (ii) the WHO Region of the Americas; and (iii) the WHO South-East Asia and Western Pacific Regions combined. In addition, malaria-endemic areas in the WHO Eastern Mediterranean Region were initially considered but insufficient data were available. We constructed weight-for-age growth references for the three regions that could be used to optimize dosing regimens for antimalarials and other drugs prescribed by age.

Methods

We compiled databases of individual-level, anthropometric data from publicly available sources collected between 1995 and 2012 on individuals of both sexes and all ages living in malaria-endemic countries. We also contacted researchers to obtain data on age groups underrepresented in publicly available sources, particularly school-aged children aged 5–15 years and adult males aged 16–50 years.

Malaria-endemic countries were defined according to the World malaria report 2011. Eligible data sets had to contain individual-level data; include data on the country, the year of sampling and individuals’ age, sex and body weight; and have been obtained using random sampling strategies that ensured samples were potentially representative of the age and sex group of interest in the population. Studies of subgroups selected according to nutritional, economic or health criteria were excluded.

We used four principal data sources for anthropometric survey data: (i) Demographic and Health Surveys (DHS); (ii) United Nations Children’s Fund Multiple Indicator Cluster Surveys; (iii) data on adults in low- and middle-income countries that were collected through multistage, cluster-randomized, household surveys using the WHO STEPwise approach to the surveillance of risk factors for major chronic diseases; and, to address gaps in specific countries; and (iv) trial and survey data on population-representative samples from individual studies. Both DHS and Multiple Indicator Cluster Surveys used multistage, random sampling to obtain nationally representative samples and collected data on anthropometric measurements in children aged 0–59 months and in women of childbearing age (i.e. 15–49 years).

We collected information on the data source; the country where data were collected; and each individual’s sex; age, to the nearest day, month or year; and body weight, to the nearest hundred grams, where available. Records with missing observations or obvious errors (e.g. weight under 0 kg or age under 0 months) were excluded. In studies where age was documented in full years, half a year was added to average ages over the year. Data from eligible data sets were merged to create crude weight-for-age databases for Africa, the Americas, and South-East Asia and the Western Pacific. Individuals with weight-for-age Z scores that appeared improbable when compared with recent Centers for Disease Control and Prevention (CDC) growth charts were identified. As the observed spread in weight-for-age Z scores between countries and studies was large, the definition of an improbable value was different for different regions: it was: less than −11 standard deviations (SD) or more than +8 SD in Africa, less than −7 SD or more than +5 SD in the Americas, less than −10 SD or more than +5 SD in South-East Asia and the Western Pacific, and less than −6 SD or more than +4 SD in the Eastern Mediterranean. These cut-off values were conservative and reflected biological differences in growth between regions and the fact that median weights in our regions were considerably lower than those in the CDC reference set, which was based on a population from the United States of America. As a final step, only individuals older than 14 days and younger than 50 years were included: the lower limit was set to avoid modelling postpartum weight loss in the first 2 weeks of life and the upper limit, because data from older age groups were scarce.

Modelling strategy

Details of our modelling method have been published previously. In brief, country-specific, weight-for-age reference curves were derived from individual-level, population-representative data using a generalized additive model for location, scale and shape. Subsequently, country curves were combined to create regional, weight-for-age, reference curves. To increase the applicability of these regional reference curves for optimizing antimalarial doses, individual country curves were weighted according to the size of the population at risk of malaria caused by Plasmodium falciparum or P. vivax. Our model (GAMLSS package version 4.2.6 for the statistical package R version 2.15.2, London Metropolitan University, London, United Kingdom of Great Britain and Northern Ireland), which is a generalization of the LMS model, allowed the mean and spread of the weight-for-age distribution to vary between countries. We were able to generate growth charts for the full age range by extrapolating information from neighbouring countries with comparable growth curves, even for countries where information on certain age–sex categories was lacking.

Box 1 provides an overview of the main steps in our modelling strategy. The full data set was split into two parts: 70% for model building and 30% for validation. Subsequently, country- and sex-specific, smoothed, weight-for-age curves were generated. The final models were selected by determining the optimal number of degrees of freedom for the regional curves. Different model distributions were compared to establish which provided the best fit and we examined whether the addition of spline functions would improve the fit. We then used regional mean square error graphs to assess the deviation between various modelled country combinations and global values to determine whether country-level, weight-for-age distributions in each region were sufficiently homogeneous to allow joint modelling of the countries or whether countries needed to be split into groups with similar patterns to model their distributions separately before merging them. The countries in Africa and countries in South-East Asia and the Western Pacific were split into two country groups each, whereas countries in the Americas were modelled together (Appendix A, available from: http://archive.lstmed.ac.uk/4566/). Thereafter, country curves for each region were merged into a single regional weight-for-age reference curve using finite mixture distributions weighted for the size of each country's
Box 1. Modelling strategy for generating weight-for-age reference curves for malaria-endemic regions

- Clean and prepare individual data sets for joint modelling.
- Split the full data set into two parts: 70% for model building and 30% for validation.
- Set up separate databases for each sex.
- Calculate weight-for-age Z scores for all individuals and exclude improbable values.
- Establish country modelling groups by examining the divergence of the weight-for-age distributions of individual countries using regional mean square error graphs – countries in each of the following regions were divided into two modelling groups each: Africa, and South-East Asia and the Western Pacific.
- Generate weight-for-age curves for each country.
- Identify the type of weight-for-age distribution.
- Determine whether splines improve the fit of the models.
- Identify the optimal number of degrees of freedom.
- Apply the Schwarz Bayesian criterion as a numerical guide to determine the optimal fit.
- Select final model parameters by comparing country-level worm plots and regional mean square error graphs over a range of numbers of degrees of freedom for all variables.16
- Validate the country models by using the validation database to calculate the percentage of individuals in each country with weight-for-age values that corresponded to specific modelled centiles.
- Weight countries according to the population at risk of malaria using Malaria Atlas Project data.17
- Calculate weight-for-age reference curves for the three regions by combining country curves using a finite mixture distribution.

Table 1. Data used to derive weight-for-age reference curves for malaria-endemic countries in the WHO Regions of Africa, the Americas, South-East Asia and Western Pacific, 1995–2012

<table>
<thead>
<tr>
<th>Data set and source</th>
<th>No. of countries</th>
<th>No. of studies</th>
<th>No. of individuals</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Initial database</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Demographic and Health Surveys</td>
<td>56</td>
<td>111</td>
<td>1,266,209</td>
</tr>
<tr>
<td>Multiple Indicator Cluster Surveys</td>
<td>45</td>
<td>45</td>
<td>246,596</td>
</tr>
<tr>
<td>Household surveys using the WHO STEPwise approach</td>
<td>10</td>
<td>10</td>
<td>29,315</td>
</tr>
<tr>
<td>Other</td>
<td>26</td>
<td>95</td>
<td>733,201</td>
</tr>
<tr>
<td><strong>Database after cleaning</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Demographic and Health Surveys</td>
<td>53</td>
<td>106</td>
<td>870,430</td>
</tr>
<tr>
<td>Multiple Indicator Cluster Surveys</td>
<td>41</td>
<td>41</td>
<td>234,188</td>
</tr>
<tr>
<td>Household surveys using the WHO STEPwise approach</td>
<td>9</td>
<td>9</td>
<td>23,802</td>
</tr>
<tr>
<td>Other</td>
<td>25</td>
<td>48</td>
<td>203,516</td>
</tr>
<tr>
<td><strong>Final database</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Africa</td>
<td>38</td>
<td>140</td>
<td>681,022</td>
</tr>
<tr>
<td>Americas</td>
<td>13</td>
<td>29</td>
<td>302,279</td>
</tr>
<tr>
<td>South-East Asia and the Western Pacific</td>
<td>13</td>
<td>23</td>
<td>279,818</td>
</tr>
<tr>
<td>Total</td>
<td>64</td>
<td>192</td>
<td>1,263,119</td>
</tr>
</tbody>
</table>

WHO: World Health Organization.

1 WHO: World Health Organization.
2 Cleaning involved excluding: data with missing or invalid entries for country, sex, body weight or age; duplicate data; treatment data; data not representative of the population; and data from areas where malaria was not endemic.
3 For the final data set, only individuals older than 14 days and younger than 50 years were included and individuals with improbable weight-for-age Z scores were excluded. In addition, data from the WHO Eastern Mediterranean Region were excluded.

Results

An overview of the data sets selected for modelling is shown in Table 1. We considered data on 2,275,321 individuals from 218 data sources and 77 countries. After limiting data to that collected between 1995 and 2012 and excluding data with missing or invalid entries for country, sex, body weight or age, duplicate data and data from countries for which information on the risk of malaria was missing, the database comprised data on 1,331,936 individuals from 207 sources and 71 countries – 41.5% of the original data were excluded. Improbable weight-for-age Z scores were identified, leading to the removal of 0.20% of records in Africa, 0.11% in the Americas, 0.11% in South-East Asia and the Western Pacific, and 0.87% in the Eastern Mediterranean – no more than 1.3% was removed for any one country. The final database included data on 1,263,119 individuals: 909,368 women and 353,751 men. Appendix B (available from: http://archive.istmed.ac.uk/4566/) lists the data sources included in the final database.

As shown in Fig. 1, our data covered 64 of the 102 (62.7%) malaria-endemic countries for which malaria risk data were available. More than 80% of the data came from standardized health surveys, whereas the remaining data were obtained through literature and web searches or from public or private institutions or individual researchers. Sources included demographic survei-
The availability of data varied substantially by age and sex between countries. Overall, 44.2% (558,218 individuals) of the data were for children younger than 5 years and 47.3% (596,883 individuals) were for women of childbearing age (i.e. aged 15–49 years). Only 2.5% (31,021 individuals) were for adult males (i.e. aged 18 years or older) and 3.8% (47,581 individuals) were for school-aged children (i.e. aged 6–12 years). Several non-DHS-type data sets showed signs of rounding and heaping of weights and ages. Data covering the full age range for both sexes were available for eight countries in Africa, two in South-East Asia and the Western Pacific, and one in the Americas. Fig. 2 shows an example of a weight-for-age scatter plot (plots for other regions and both sexes are in Appendix C, available from: http://archive.lstmed.ac.uk/4566/) and Fig. 3 shows a smoothed, country-level curve of weight-for-age medians (curves for other regions and both sexes are in Appendix D, available from: http://archive.lstmed.ac.uk/4566/).

Appendix A lists the number of degrees of freedom that provided the best model fit for the reference curves in each region as derived using the Box–Cox power exponential model with cubic or penalized splines. Smoothed, regional, weight-for-age references were obtained for both sexes by combining country-level distributions weighted for the size of the population in malaria-endemic areas. Fig. 4 shows weight-for-age reference curves for different centiles for females in South-East Asia and the Western Pacific (curves for other regions and both sexes are in Appendix E, available from: http://archive.lstmed.ac.uk/4566/). Comprehensive tables containing Box–Cox power exponential model estimates, weight-for-age Z scores and all the centiles needed to reproduce the regional references can be downloaded from the WorldWide Antimalarial Resistance Network.18

Model validation

Country-level, weight-for-age references were validated using data from the initial data set reserved for validation. Appendix F (available from: http://archive.lstmed.ac.uk/4566/) shows the percentage of individuals in the validation set that fell below our calculated 3rd, 10th,
50th, 90th and 97th centiles for both sexes in each individual country in the three regions. Most modelled centiles were well on target: 96.9% of centiles in the validation set fell within 3 percentage points of the modelled weight-for-age. Of the centile points tested, 2.0% lay between 3 and 5 percentage points from the modelled weight-for-age and 1.1% lay more than 5 percentage points distant. Deviations beyond 3 percentage points were more frequent in countries for which data were relatively scarce, such as Belize, the Philippines, the Solomon Islands and Suriname.

**Comparison between references**

Modelled weight-for-age distributions differed considerably between the three regions (Fig. 5). Overall, body weight in South-East Asia and the Western Pacific was lower than in Africa or in the Americas. The median weight of an adult male aged 18 years or older in Africa and the Americas was 64.2 kg and 68.1 kg, respectively, whereas the median was 62.1 kg in South-East Asia and the Western Pacific. The corresponding median weights in females were 57.4 kg in Africa, 60.4 kg in the Americas and 50.0 kg in South-East Asia and the Western Pacific. Fig. 6 shows the percentage difference between median weight-for-age values in different regions and corresponding median values in our modelled global reference. Differences were most marked during the onset of the growth spurt in early adolescence. The modelled global reference curves are similar to curves for South-East Asia and the Western Pacific because the population at risk of malaria in this region was large.

Fig. 5 and Fig. 6 also show that median weight-for-age values in the three regions are much lower for all ages than in recent CDC growth charts – Fig. 5 shows that they are, at times, similar to values 2 SD below the medians in these charts. The difference in median weight between our modelled global reference and the growth chart median was around 20% for 1-year-olds and almost 70% for adolescents aged 13–14 years (Fig. 6). The peak difference in median weight between continents occurred during adolescence, which reflects the late onset of puberty in South-East Asia and the Western Pacific relative to Africa and America.

**Discussion**

We developed regional weight-for-age growth references representative of populations in malaria-endemic areas of Africa, the Americas and South-East Asia and the Western Pacific using existing anthropometric data. To our knowledge, these are the first region-specific estimates of the weight achieved by age in these populations. Although our reference values were tailored to optimize antimalarial regimens, they are relevant to age-based therapies used for other diseases. Our modelling method was specifically designed such that data from a wide range of sources could be pooled to create robust, regionally representative references. The results demonstrate that large differences in weight-for-age distributions existed between the three regions and between these regions and
Fig. 3. Median weight-for-age curves for females in malaria-endemic countries in the WHO South-East Asia and Western Pacific Regions, by country, 1995–2012

Fig. 4. Weight-for-age reference curves for females in malaria-endemic countries in the WHO South-East Asia and Western Pacific Regions by selected centiles, 1995–2012
optimal growth curves developed by WHO, which indicates that age-based dosing should shift from global to region-based regimens.

In all three regions, our models achieved good fits with the original data, which confirms that our generalized additive model for location, scale and shape extension method is a robust way of establishing growth references using mixed-source data in situations where multicentre growth reference studies are not feasible. The relative homogeneity of countries within the regions justifies modelling pooled country data sets in each region. Validation using independent data showed that the fitted country-level references adequately modelled the empirical distributions in each country. The large majority of our data came from DHS, which are representative of national or subnational populations and have well-established designs and quality assurance methods. Extensive data were available for most of the age range of interest (i.e. 0–5 years and 15–49 years), which helped in modelling the age range for which fewer data were available (i.e. 5–14 years). Data for adolescents and adult males were missing for many countries, however, they were available across the full age spectrum for at least one country in each region. Since our two-step modelling approach was designed to use data from adjacent countries and age groups, where these showed similar growth distributions, we were able to model the growth curves for adolescents and adult males for all three regions. In the Eastern Mediterranean no data were available for older children or adult males. Several non-DHS-type data sets showed heaping that resulted from rounding of crude weight and age values. As heaped entries made up a very small proportion of the observations, however, the effect was limited and did not influence the final models.

To increase the representativeness of our curves, we did not use data sets from before 1995: 167 of 192 (87%) sets were from surveys conducted in 2000 or later. Although obesity is becoming more common in several low- and middle-income countries, only a wealthier subpopulation living in urban areas is likely to be affected, not the population in the rural areas most at risk of malaria. However, growth curves should be updated periodically (e.g. every 5 to 10 years) using the latest data from representative surveys. Our multisource modelling method provides a dynamic framework that enables new data to be incorporated when they become available and makes it possible to include non-malaria-endemic countries in the three regions modelled. The growth charts we generated for individual countries as part of the modelling process could serve as powerful public health tools to support decision-making at a national level. The logical next step would be to further validate these charts for countries where data were limited. Country-specific growth references could be improved further by extrapolating female data to derive male growth curves.

Our method provides a way of deriving regional growth references by collating weight-for-age data available for populations. Furthermore, our method facilitates the transition
Malnourishment results in growth references. The method could also help monitor nutrition and optimize age-based dosing of other drugs.

Acknowledgements
We thank all individuals and research groups who provided data (Appendix G, available from: http://archive.lstm.ac.uk/4566/), David Lalloo and Eva Maria Staelhi-Hodel.

Funding: Medical Research Council (MRC)/Liverpool School of Tropical Medicine grant GO0501401 and MRC grant G1100522.

Competing interests: None declared.
Regional weight-for-age growth references

Daniel J Hayes et al.

Abstract

To develop regional age-specific weight references to optimize antimalarial drug dosing based on age in malaria endemic countries.

Objectif

Obtenir les références de croissance pour le rapport poids/âge entre les régions et entre les régions et les normes mondiales de croissance. Par exemple, l'homme adulte moyen de les Amériques pesait 68,1 kg – soit 6,0 kg de plus que les hommes vivant dans le Pacifique occidental et l'Asie du Sud-Est (moyenne: 62,1 kg). Pour les femmes adultes, la différence dépassait les 10,4 kg: la moyenne était de 60,4 kg dans les Amériques et de 50,0 kg dans le Pacifique occidental et l'Asie du Sud-Est.

Conclusion

Il existait des variations substantielles dans les courbes de croissance pour le rapport poids/âge entre les zones où le paludisme est endémique. Les tableaux de référence de croissance qui sont obtenus ici peuvent être utilisés pour guider l'optimisation factuelle des schémas posologiques des médicaments antipaludiques en fonction de l'âge et des autres médicaments qui sont souvent prescrits par âge.

Résumé

Développer des références régionales de croissance pour le rapport poids/âge dans les pays où le paludisme est endémique, afin d'optimiser la posologie des médicaments antipaludiques en fonction de l'âge
показателей веса и роста как между исследуемыми регионами, так и между данными регионами и мировыми стандартами роста. Например, вес среднестатистического взрослого мужчины в Америке составлял 68,1 кг — что на 6,0 кг больше веса мужчин в Юго-Восточной Азии и Западной части Тихого океана (в среднем 62,1 кг). Для взрослых женщин эта разница составляла более 10,4 кг: средний вес женщин равнялся 60,4 кг в Америке и 50,0 кг в Юго-Восточной Азии и Западной части Тихого.

Вывод
Были обнаружены существенные различия в кривых роста соответствующих веса к возрасту между эндемичными по малярии районами. Полученные в ходе исследования таблицы справочных показателей роста могут быть использованы в качестве ориентировочных при проведении оптимизации на основе фактических данных схем дозирования противомалярийных и других лекарственных средств, часто назначаемых по возрасту.