Acta Pædiatrica, 2006; Suppl 450: 66–75

Taylor & Francis
Taylor & Francis Group

Assessment of sex differences and heterogeneity in motor milestone attainment among populations in the WHO Multicentre Growth Reference Study

WHO MULTICENTRE GROWTH REFERENCE STUDY GROUP[1,2]

[1]*Department of Nutrition, World Health Organization, Geneva, Switzerland, and* [2]*Members of the WHO Multicentre Growth Reference Study Group (listed at the end of the first paper in this supplement)*

Abstract

Aim: To assess the heterogeneity of gross motor milestone achievement ages between the sexes and among study sites participating in the WHO Multicentre Growth Reference Study (MGRS). *Methods:* Six gross motor milestones (sitting without support, hands-and-knees crawling, standing with assistance, walking with assistance, standing alone, and walking alone) were assessed longitudinally in five of the six MGRS sites, namely Ghana, India, Norway, Oman and the USA. Testing was started at 4 mo of age and performed monthly until 12 mo, and bimonthly thereafter until all milestones were achieved or the child reached 24 mo of age. Four approaches were used to assess heterogeneity of the ages of milestone achievement on the basis of sex or study site. *Results:* No significant, consistent differences in milestone achievement ages were detected between boys and girls, nor were any site–sex interactions noted. However, some differences among sites were observed. The contribution of inter-site heterogeneity to the total variance was $<5\%$ for those milestones with the least heterogeneous ages of achievement (hands-and-knees crawling, standing alone, and walking alone) and nearly 15% for those with the most heterogeneous ages of achievement (sitting without support, standing with assistance, and walking with assistance).

Conclusion: Inter-site differences, most likely due to culture-specific care behaviours, reflect normal development among healthy populations across the wide range of cultures and environments included in the MGRS. These analyses support the appropriateness of pooling data from all sites and for both sexes for the purpose of developing an international standard for gross motor development.

Key Words: *Gross motor milestones, longitudinal, motor skills, standards, young child development*

Introduction

The WHO Multicentre Growth Reference Study (MGRS) was designed to provide a description of the physical growth and gross motor development in healthy infants and children throughout the world. Previous efforts to develop growth references relied on data collected from infants and young children "free from disease" who were representative of defined geographical areas. When appropriately carried out, such studies provide accurate snapshots of how children grow and/or develop in a particular time and place. The MGRS, however, adopted a prescriptive approach designed to describe how children *should* grow independently of time and place. In so doing, it defined health not only as the absence of disease but also as the adoption of healthy practices known to promote health, e.g. breastfeeding. The rationale, design and protocol for the MGRS have been described in detail elsewhere [1,2].

The second unique feature of the MGRS is that it included children from many of the world's major regions: Brazil (South America), Ghana (Africa), India (Asia), Norway (Europe), Oman (the Middle East) and the USA (North America). This design feature tested the assertion that growth in infancy and early childhood is very similar among diverse ethnic groups when conditions that favour growth are met [1]. The MGRS also offered an opportunity to assess the heterogeneity/similarity in gross motor development across distinct cultures and environments.

Undoubtedly, MGRS participants from diverse sites differed genetically; however, it is unlikely that functions and traits such as motor development

Correspondence: Mercedes de Onis, Study Coordinator, Department of Nutrition, World Health Organization, 20 Avenue Appia, 1211 Geneva 27, Switzerland. Tel: +41 22 791 3320. Fax: +41 22 791 4156. E-mail: deonism@who.int

ISSN 0803-5326 print/ISSN 1651-2227 online
DOI: 10.1080/08035320500495530

and linear growth, which reflect the coordinated expression of multiple genes, differ substantially and systematically among large populations living in healthy environments. At the population level, it is likely that environmental disparities such as those seen in developing countries influence phenotypic expressions of multigenic functions and traits to a greater extent than genetic differences do [3].

The literature provides only a limited basis on which to directly evaluate how these views relate to motor development. A number of studies have considered relationships between general nutritional or specific nutrient status [4–9], feeding mode in early infancy [10,11], and specific disease states or conditions [12,13] and motor development. Some have examined differences in motor development among diverse cultural or ethnic groups in healthy and unhealthy states [14–18]. These interests are not new. For example, Garcia-Coll [19] reviewed early papers that evaluated potential aetiologies of the putative motoric precocity of African American infants and infants of African descent in developing and developed countries [19–23]. Clearly, there are significant difficulties associated with isolating biological from caretaker socio-economic and attitudinal/behavioural influences. Complexities such as these thus make it difficult to interpret results of evaluations of the role that ethnicity and culture play in motor development [19,24].

Although the literature includes a discussion of differences in motor development between boys and girls [16,18,25,26], findings are inconsistent in that either no differences are found between boys and girls, or boys are observed to be either more delayed or at risk of being delayed when faced with various forms of stress. Apparently, no study has evaluated potential interactions among sex, ethnicity and cultural background when assessing motor development in young children.

The aim of this paper is to assess the heterogeneity between the sexes and among MGRS study sites of gross motor milestone achievement ages. Analyses are carried out to evaluate the need for distinct standards for boys and girls and the appropriateness of pooling observations from all MGRS sites that performed motor development assessments.

Methods

General study design

The rationale, planning, design and methods of the MGRS, including its motor development component and site-specific protocol implementation, have been described in detail elsewhere [1,2,27].

Six distinct gross motor milestones were assessed: sitting without support, hands-and-knees crawling, standing with assistance, walking with assistance, standing alone, and walking alone. These were selected because they are considered universal, fundamental to the acquisition of self-sufficient locomotion, and simple to test and evaluate. These milestones were assessed longitudinally beginning at 4 mo of age on all children enrolled in the longitudinal sample in five of the six MGRS sites, namely Ghana, India, Norway, Oman and the USA. Motor development was not assessed in Brazil because most of that site's longitudinal sample was older than 4 mo when motor development was added to the MGRS protocol.

Using standardized testing procedures and criteria, study staff performed monthly assessments until 12 mo of age and bimonthly assessments thereafter until all milestones were achieved or the child reached 24 mo of age. No fixed milestone sequence was assumed and all milestones were assessed at each visit. Training and standardization procedures and data collection protocols, described in detail elsewhere [27,28], were similar among sites.

Sample used for analyses

Analyses of differences between the sexes or among sites in age of motor milestone achievement were based on the same sample of children included in assessments of inter-site heterogeneity in linear growth [29]. In the five study sites where motor development was assessed, 1433 children were enrolled in the MGRS longitudinal component. Because of missing data, 149 (10%) of these children were not included in the assessment of inter-site heterogeneity for linear growth. Of the children ($n = 1284$) included in the linear growth assessment, 75 (5%) did not participate in the MGRS motor development assessment component.

Variable numbers of motor milestone assessments by trained MGRS personnel were available for individual children in the remaining sample ($n = 1209$, 85%). This was mainly the result of late initiation of this MGRS component at the Norwegian and Ghanaian sites due to funding constraints, which meant that some children were too old to participate fully in motor assessments.

Statistical analyses

Estimation of ages of motor milestone achievement. The MGRS design [2] did not permit the determination of exact ages of milestone achievement because subjects were not supervised daily by trained staff. "True" ages of milestone achievement were linked to intervals between visits by staff documenting the first observed

achievements of specific milestones and the most recent previous visit. Specific ages of achievement within those designated intervals were assigned randomly based on the assumption that achievement ages were distributed uniformly between scheduled visits. Detailed descriptions of the uses of fieldworker observations and caretaker reports of achievement ages are described in a companion paper in this supplement [30].

Evaluation of heterogeneity of milestone achievement ages between the sexes and among the MGRS sites. Two model-based approaches were used to characterize inter-site and inter-sex heterogeneity of the ages of milestone achievement.

A within-subject design ANOVA was used to assess proportional contributions of sex and site, both as main effects, to the total observed variation in ages of achievement of motor milestones and to evaluate site-sex interactions [31].

Another model-based approach applied a three-level variance components model (level 1: milestone indicator; level 2: individual child; and level 3: site). This model treated milestone achievement ages as successive occasions, assumed that achievement ages equalled fixed effects, and allowed for random perturbation on the normal scale [32]. To account for inter-level heterogeneity, a random effect was assigned to each clustering level. The percentages of the total variance attributable to each clustering level were calculated as fractions of the total variance [32,33]. Log-likelihood ratio was used to test the significance of sources of heterogeneity [32].

We also evaluated the magnitude of differences in ages of achievement of specific milestones between the sexes and among sites by calculating differences between the pooled mean age of achievement and the means for either sex or single sites as fractions of the pooled mean's standard deviation, i.e.

$$\frac{Y_A - Y}{SD} = Diff$$

where Y_A is the mean for *site A* or *sex A*, Y is the pooled mean, and SD is the standard deviation of the respective age of achievement corresponding to the pooled sample.

Site-specific and all-site average differences (in days) between boys' and girls' ages of achievement for each milestone were also calculated, and two-sample t-tests were performed to assess site- and milestone-specific differences in motor milestone achievement ages between boys and girls.

Lastly, the impact of inter-site heterogeneity was assessed further by evaluating the impact of excluding individual sites on percentile estimates. Differences were calculated between the 1st, 50th and 99th percentiles corresponding to "all-site" pooled values and the values calculated when single sites were individually omitted. Normalized differences were expressed as fractions of the standard deviations of the all-site pooled means.

Statistical significance was assigned to comparisons with p-values <0.05.

Results

Statistically significant differences in milestone achievement ages were not detected between boys and girls, nor were significant site-sex interactions noted (Table I) when a within-subject design ANOVA was applied. Figure 1 summarizes site-specific and overall differences in the ages of motor milestone achievement between boys and girls.

Two-sample t-tests assessing site- and motor milestone-specific differences between boys' and girls' ages of achievement detected statistically significant differences in five of 30 comparisons (Table II), namely sitting without support in India, walking with assistance in the USA, standing alone in Oman, and walking alone in Ghana and Oman. For all sites, statistically significant differences in the ages of achievement between boys and girls were detected for sitting without support (mean difference <5 d earlier for girls) and standing alone (mean difference of approximately 7 d earlier for girls).

Table I. Analysis of variance comparing the effect of sex, site and their interaction on milestone achievement ages.

Source of variation	Partial sum of squares	Degrees of freedom	p-value (prob >F)	Proportion of variance (%)
Among subjects:				
Site	1 119 723.3	4	0.0000	2.61
Sex	8626.0	1	0.2756	0.02
Interaction (site, sex)	50 649.9	4	0.1374	0.12
Residual (inter-subject)	8 686 429.2	1,198		20.22
Within subjects:				
Milestone	26 262 996.5	5	0.0000	61.12
Residual (intra-subject)	6 101 140.7	5,771		14.20
Total	42 970 018	6,983		100.00

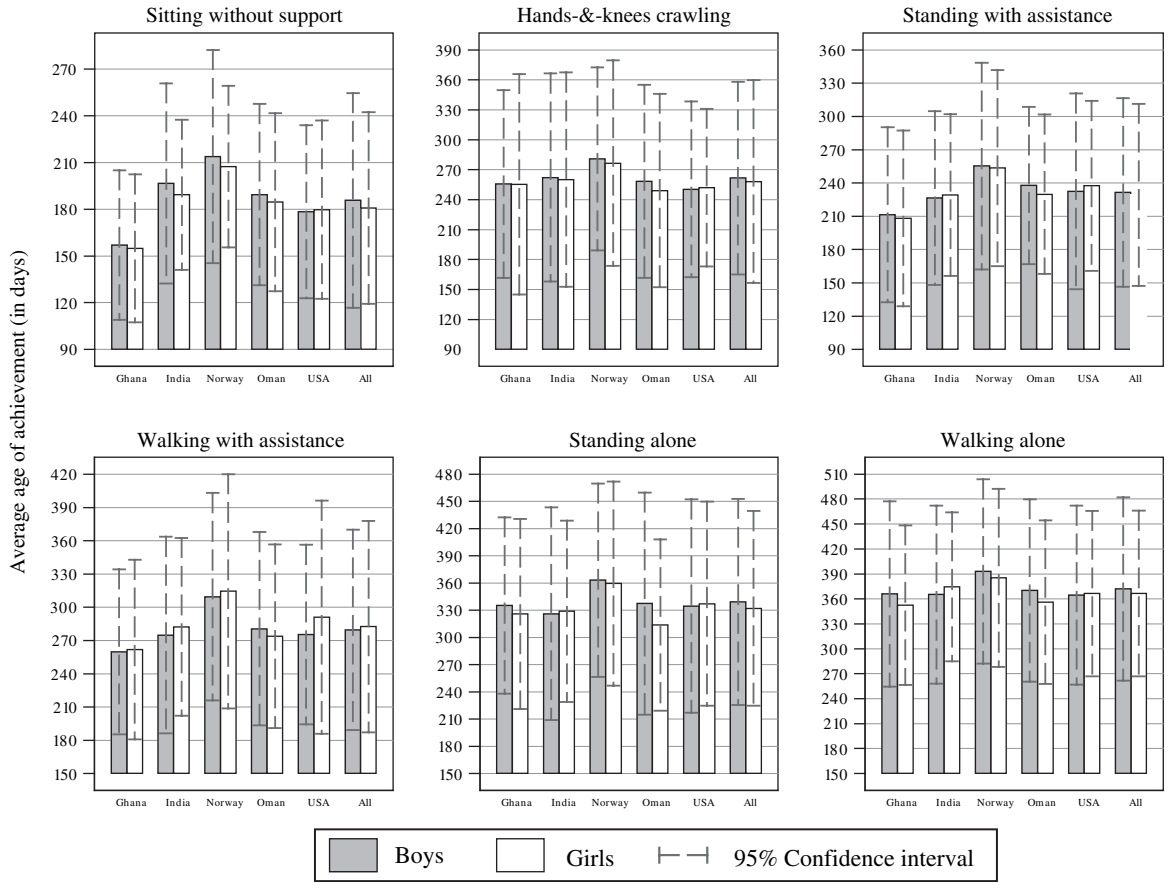

Figure 1. Average ages of gross motor milestone achievement in boys and girls.

Sitting without support exhibited the statistically most significant difference ($p = 0.0125$) in ages of achievement between boys and girls when all sites were pooled. Figure 2 illustrates the cumulative frequencies of the ages of achievement of sitting without support for boys and girls separately.

Small, though statistically significant, differences were observed among sites (sites accounted for 2.6% of the observed variance in Table I). Table III characterizes heterogeneity, by milestone, in the ages of milestone achievement. Ages of achievement for sitting without support demonstrated the greatest heterogeneity among sites. The least heterogeneity was observed for hands-and-knees crawling, standing alone and walking alone. *P*-values of log-likelihood ratio testing the significance of variance components due to site heterogeneity were <0.05. With the exception of standing alone ($p = 0.0298$), no evidence of heterogeneity due to sex, and no interaction of site and sex, were observed.

Estimates of the proportion of the total variance contributed by inter-site heterogeneity and inter-individual differences are summarized in Table IV. Inter-site heterogeneity contributed the least to the total variance (8.3%). Table IV also summarizes the contributions of inter-site heterogeneity to total variance when milestones with the greatest and

Table II. *P*-values of the two-sample *t*-tests on the equality of means between boys and girls.

Site	Sitting without support	Hands-and-knees crawling	Standing with assistance	Walking with assistance	Standing alone	Walking alone
Ghana	0.4665	0.9614	0.4885	0.6831	0.1377	0.0376*
India	0.0423*	0.7579	0.5608	0.1582	0.6988	0.1304
Norway	0.1730	0.5437	0.7861	0.4570	0.6073	0.2865
Oman	0.1781	0.1303	0.0798	0.2089	0.0008*	0.0371*
USA	0.7591	0.7860	0.4326	0.0348*	0.7718	0.8135
Total	0.0125*	0.2254	0.3900	0.3184	0.0297*	0.0654

*Statistically significant ($p < 0.05$).

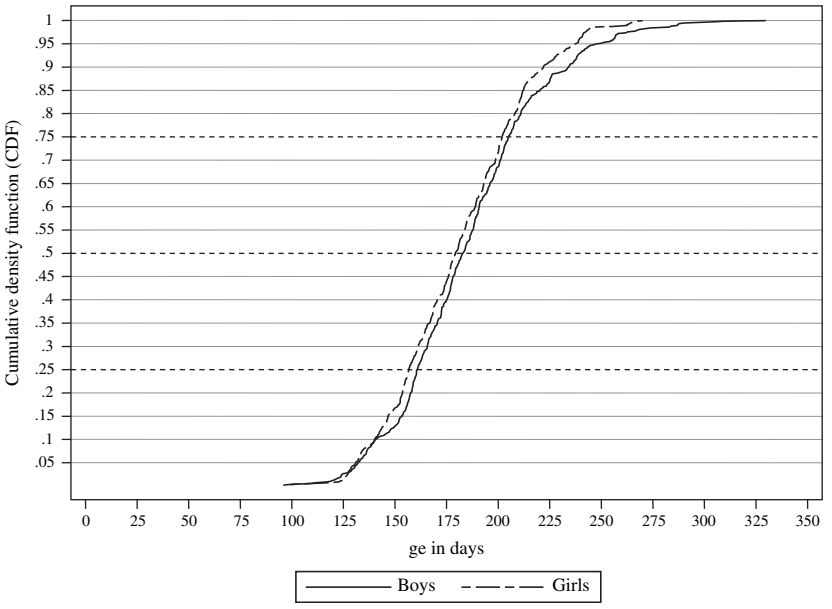

Figure 2. Cumulative frequency of motor achievement of sitting without support for boys and girls.

least heterogeneity were grouped. The contribution of inter-site heterogeneity to the total variance was $<5\%$ for those milestones with the least heterogeneous ages of achievement (hands-and-knees crawling, standing alone, and walking alone) and nearly 15% for those with the most heterogeneous ages of achievement (sitting without support, standing with assistance, and walking with assistance). *P*-values of log-likelihood ratios testing the significance of variance components due to site heterogeneity were <0.05. No evidence of heterogeneity due to sex or significant interaction of site and sex was observed.

Site-specific mean achievement ages and pooled means are presented in Table V. Normalized differences (expressed as fractions of the standard deviation of the pooled means) between site-specific means and the pooled mean varied by milestone. The Ghanaian sample exhibited the earliest mean ages of achievement for sitting without support (-0.82), standing with assistance (-0.49), walking with assistance (-0.43) and walking alone (-0.19). Normalized differences for all other sites with mean ages of achievement below the all-site pooled mean ranged from -0.17 to -0.05.

The Norwegian sample exhibited the latest mean ages of achievement for all six milestones (Table V). Normalized differences for all other sites with mean ages of achievement greater than the all-site pooled mean varied from 0.01 to 0.29.

Table VI summarizes the impact of eliminating single sites on the mean, 1st, 50th and 99th age of achievement percentiles. The impact of site elimination was assessed by comparing the "single-site elimination" values with "all-site" pooled values. Excluding the Ghanaian site increased the remaining site pooled mean (and corresponding percentiles) for sitting without support, standing with assistance,

Table III. Variance components two-level model comparing site heterogeneity by milestone.

Milestone	Variance component[a]	Estimate	Standard error (estimate)	*p*-value	Proportion of variance (%)
Sitting without support	Var(Site)	438.4	279.5	<0.000	34.8
	Var(Error)	823.1	33.6		65.2
Hands-and-knees crawling	Var(Site)	87.1	61.7	<0.000	3.5
	Var(Error)	2382.1	99.1		96.5
Standing with assistance	Var(Site)	255.8	166.1	<0.000	13.9
	Var(Error)	1584.8	64.6		86.1
Walking with assistance	Var(Site)	289.5	188.5	<0.000	12.8
	Var(Error)	1976.8	80.8		87.2
Standing alone	Var(Site)	177.2	120.4	<0.000	5.5
	Var(Error)	3042.3	125.1		94.5
Walking alone	Var(Site)	123.1	85.5	<0.000	4.3
	Var(Error)	2776.4	114.4		95.7

[a]"Site" as a random effect.

walking with assistance and walking alone by 9, 7, 6 and 3 d, respectively. Excluding Norway decreased the remaining site pooled mean for all six milestones by 5, 4, 5, 7, 6 and 5 d, respectively. As absolute values, these differences represent less than 0.3 of the pooled mean's SD for all estimated differences. Of the 30 "single-site exclusion" means calculated for all milestones, 23 differed from the pooled mean by ≤ 0.1 of the all-site pooled mean's SD; six were between 0.1 and 0.2, and one was between 0.2 and 0.3.

Discussion

These findings support the conclusion that MGRS gross motor development data from female and male infants and toddlers should be pooled for the purpose of constructing standards. The statistical insignificance of sex as a source of variability in the ages of milestone achievement that is documented in Tables I, III and IV is underscored by Figure 2.

This view is justified despite sporadic statistically significant differences in the ages of motor milestone achievement between boys and girls when two-sample *t*-tests were applied (Table II). These differences were small, i.e. 7 d or less, and inconsistent. Also, they should be interpreted cautiously given that the study's large sample size and the large number of two-sample *t*-tests performed increase the possibility of alpha errors. As reported in other studies [25,26], girls in the MGRS tended to achieve milestones at earlier ages than did boys. The tendency of girls to achieve motor milestones earlier than boys observed in Figure 1 is of interest from a developmental perspective; however, the magnitude of observed differences is too small to justify sex-specific norms.

The absence of any site–sex interaction is also reassuring. Its absence discounts the possibility that boys and girls were treated differentially in diverse sites in a manner that operated across sites to obscure sex-based differences. The paucity of other information evaluating differences in gross motor development between male and female infants and toddlers raised in diverse cultural settings and environments makes this finding particularly valuable to the construction of an international standard. These findings also support the view that any disparities between boys and girls in gross motor development likely reflect dissimilarities in care practices and/or other factors, which is to say that it is unlikely they are due to physiological sex-based differences.

These analyses found statistically significant inter-site differences in the ages of motor milestone achievement. This finding is generally consistent with another WHO collaborative study designed to develop and standardize culturally appropriate scales of psychosocial development [18]. That study included a wide array of developmental assessments. Although specific tests of inter-site differences were not included in the cited reference, tabulated information documents homogeneity in ages of achievement among some milestones but not among others. These findings suggest that environmental diversity may have accounted for the lack of homogeneity across all measures, which is consistent with observations made by others. For example, Lima et al. [25] reported that environments influence mental and motor development to a much greater degree than do biological factors (e.g. birthweight).

Analyses summarized in Table I indicate that sites contributed $<3\%$ of the variability observed in the MGRS. This estimate merits close examination. The variability and error introduced by the random point determination of ages of milestone achievement and the likelihood of uneven susceptibility of different milestones to caretaker influences (discussed further below) may have decreased the proportional contribution of inter-site differences. The most important challenge presented by statistically significant inter-site differences and considerations of the determinants of variability is assessing their implications for the purpose of constructing an international standard. Three aspects of the analyses addressed this point. The first assessed the magnitude of differences among

Table IV. Variance components three-level model comparing site heterogeneity by milestones combined.

Milestones grouped	Variance component [a]	Estimate	Standard error (estimate)	*p*-value	Proportion of variance (%)
All six milestones	Var(Site)	192.4	125.0	<0.000	8.3
	Var(Child)	1067.3	50.9	<0.000	46.1
	Var(Error)	1057.6	19.4		45.6
Sitting without support, standing with assistance, walking with assistance	Var(Site)	248.7	159.9	<0.000	14.5
	Var(Child)	690.7	39.5	<0.000	40.1
	Var(Error)	781.6	22.5		45.4
Hands-and-knees crawling, standing alone, walking alone	Var(Site)	129.6	87.5	<0.000	4.5
	Var(Child)	1701.9	85.1	<0.000	59.1
	Var(Error)	1046.6	31.0		36.4

[a]"Site" as a random effect.

Table V. Site-specific and "all-site" achievement ages (in days) by milestone.

	n	Mean	SD	Diff. in SD		*n*	Mean	SD	Diff. in SD
Sitting without support					Hands-and-knees crawling				
Pooled estimate	1139	183.3	33.4	0.00	Pooled estimate	1128	260.0	50.4	0.00
Estimate for Ghana	280	156.0	24.1	−0.82	Estimate for Ghana	261	255.7	51.7	−0.09
Estimate for India	262	193.1	29.0	0.29	Estimate for India	244	261.1	53.3	0.02
Estimate for Norway	173	210.8	30.8	0.82	Estimate for Norway	203	278.8	48.8	0.37
Estimate for Oman	258	187.1	29.2	0.12	Estimate for Oman	255	253.9	49.1	−0.12
Estimate for USA	166	179.1	28.3	−0.12	Estimate for USA	165	251.3	41.9	−0.17
Standing with assistance					Walking with assistance				
Pooled estimate	1169	230.5	42.6	0.00	Pooled estimate	1185	281.1	47.3	0.00
Estimate for Ghana	280	209.8	39.9	−0.49	Estimate for Ghana	278	260.9	39.3	−0.43
Estimate for India	262	227.8	38.3	−0.06	Estimate for India	262	278.6	42.9	−0.05
Estimate for Norway	203	254.5	45.7	0.56	Estimate for Norway	224	311.9	50.1	0.65
Estimate for Oman	258	234.0	36.2	0.08	Estimate for Oman	255	277.4	43.1	−0.08
Estimate for USA	166	235.0	41.4	0.11	Estimate for USA	166	283.3	47.8	0.05
Standing alone					Walking alone				
Pooled estimate	1182	335.6	56.4	0.00	Pooled estimate	1182	369.3	53.6	0.00
Estimate for Ghana	268	330.5	51.2	−0.09	Estimate for Ghana	266	359.2	52.8	−0.19
Estimate for India	262	327.4	55.2	−0.14	Estimate for India	261	369.7	50.1	0.01
Estimate for Norway	231	361.3	55.1	0.46	Estimate for Norway	236	389.3	55.1	0.37
Estimate for Oman	255	325.8	56.6	−0.17	Estimate for Oman	255	363.3	53.1	−0.11
Estimate for USA	166	335.9	57.7	0.01	Estimate for USA	164	365.4	52.0	−0.07

sites. As noted in the results section, the largest deviations from all-site pooled values were observed for Ghana and Norway. Those deviations were large in several instances, but neither Ghana nor Norway consistently accounted for the largest deviations (Table V).

Other analyses examined the consequences of specific single-site elimination on the resulting pooled means and selected percentiles. The greatest impact was observed when either Ghana or Norway was excluded from the sample. However, the exclusion of either country did not result consistently in the largest deviations from all-site pooled values. Also, as summarized in Table VI, the exclusion of any single site seldom resulted in normalized differences greater than 0.2 SD between corresponding means and the 1st, 50th and 99th centile values. Normalized differences most often were below 0.1 SD.

The contributions of inter-site differences to the total variability of specific milestones were also examined. Among the statistically significant sources of variation, sites contributed least to the variability in ages of achievement for hands-and-knees crawling (3.5%), standing alone (5.5%) and walking alone (4.3%). The most marked contribution to total variability by inter-site differences was observed for sitting without support (35%). Inter-site contributions to the total variability were intermediate in magnitude for the milestones standing with assistance (13.9%) and walking with assistance (12.8%).

Among the inferences that may be drawn from these differences is that developmental domains governing milestone achievement are influenced significantly by environmental and/or genetic factors specific to individual sites. Theories of motor development and skill acquisition and of genetic controls of development [34–37] make it unlikely that genetic factors linked to ethnicity determine the ability to sit without support to a greater extent than they do hands-and-knees crawling. The involvement of multiple gene networks seems unavoidable in the orchestration of anatomical, cognitive and other changes linked to development [38]. Thus, environmental influences appear to provide the more parsimonious explanation for observed differences. The two most relevant potential environmental influences relate to distinct gestational and/or perinatal conditions among participants and/or childcare practices in the various sites. It seems unlikely that unspecified gestational and/or perinatal site-specific conditions carry over only to the "earliest" motor milestone that was examined, but such possibilities cannot be discounted based on data collected by this study.

Although neither genetic nor environmental influences can be discounted completely as explanations for observed inter-site differences, inconsistencies within and among sites (e.g. children in Ghana did not always demonstrate the earliest ages of achievement for all milestones) and field observations suggest that childcare practices likely explain observed inter-site differences. As indicated earlier, inter-site differences were greatest between Ghana and Norway. Field reports indicate that Ghanaian caretakers commonly engaged in practices consistent with the training of infants so as to accelerate their achievement of motor milestones.

Table VI. Comparisons of achievement ages (days) by milestones when all sites are pooled and when single sites are excluded.

	n	Mean	SD	Diff. in SD	P1	Diff. in SD	P50	Diff. in SD	P99	Diff. in SD
Sitting without support										
Pooled estimate	1139	183.3	33.4	0.00	121.2	0.00	181.0	0.00	270.0	0.00
Excluding Ghana	859	192.2	31.1	0.27	127.3	0.18	190.2	0.28	282.9	0.39
Excluding India	877	180.3	34.1	−0.09	117.1	−0.12	177.4	−0.11	270.9	0.03
Excluding Norway	966	178.3	31.4	−0.15	118.9	−0.07	176.8	−0.12	265.4	−0.14
Excluding Oman	881	182.2	34.5	−0.03	117.1	−0.12	179.6	−0.04	268.5	−0.04
Excluding USA	973	184.0	34.2	0.02	122.8	0.05	181.6	0.02	274.6	0.14
Hands-and-knees crawling										
Pooled estimate	1128	260.0	50.4	0.00	169.4	0.00	254.2	0.00	410.4	0.00
Excluding Ghana	867	261.3	50.0	0.03	170.0	0.01	255.6	0.03	409.9	−0.01
Excluding India	884	259.7	49.6	−0.01	167.7	−0.03	254.4	0.00	415.2	0.10
Excluding Norway	925	255.8	49.9	−0.08	165.4	−0.08	251.1	−0.06	405.1	−0.11
Excluding Oman	873	261.7	50.7	0.04	167.7	−0.03	255.5	0.03	415.2	0.10
Excluding USA	963	261.5	51.6	0.03	170.0	0.01	255.3	0.02	417.1	0.13
Standing with assistance										
Pooled estimate	1169	230.5	42.6	0.00	153.1	0.00	227.0	0.00	351.5	0.00
Excluding Ghana	889	237.0	41.3	0.15	156.0	0.07	233.8	0.16	357.2	0.13
Excluding India	907	231.3	43.7	0.02	153.1	0.00	228.6	0.04	353.6	0.05
Excluding Norway	966	225.5	40.1	−0.12	150.2	−0.07	224.0	−0.07	340.9	−0.25
Excluding Oman	911	229.5	44.2	−0.02	150.2	−0.07	225.5	−0.04	353.6	0.05
Excluding USA	1003	229.8	42.7	−0.02	153.8	0.02	226.5	−0.01	351.5	0.00
Walking with assistance										
Pooled estimate	1185	281.1	47.3	0.00	190.6	0.00	275.4	0.00	423.7	0.00
Excluding Ghana	907	287.3	47.8	0.13	195.0	0.09	281.8	0.14	426.0	0.05
Excluding India	923	281.8	48.5	0.02	190.6	0.00	276.1	0.01	424.6	0.02
Excluding Norway	961	273.9	43.6	−0.15	189.8	−0.02	269.6	−0.12	406.1	−0.37
Excluding Oman	930	282.1	48.4	0.02	190.7	0.00	275.5	0.00	424.6	0.02
Excluding USA	1019	280.8	47.2	−0.01	190.6	0.00	275.1	−0.01	420.6	−0.06
Standing alone										
Pooled estimate	1182	335.6	56.4	0.00	230.7	0.00	329.9	0.00	491.0	0.00
Excluding Ghana	914	337.1	57.8	0.03	230.7	0.00	331.2	0.02	491.0	0.00
Excluding India	920	337.9	56.6	0.04	233.9	0.06	333.2	0.06	491.0	0.00
Excluding Norway	951	329.3	54.9	−0.11	221.6	−0.16	323.7	−0.11	486.0	−0.09
Excluding Oman	927	338.3	56.1	0.05	230.7	0.00	333.2	0.06	487.7	−0.06
Excluding USA	1016	335.5	56.2	0.00	232.3	0.03	329.7	0.00	491.0	0.00
Walking alone										
Pooled estimate	1182	369.3	53.6	0.00	256.8	0.00	361.2	0.00	517.0	0.00
Excluding Ghana	916	372.2	53.5	0.05	264.7	0.15	363.5	0.04	515.0	−0.04
Excluding India	921	369.2	54.6	0.00	256.7	0.00	360.1	−0.02	521.0	0.07
Excluding Norway	946	364.3	52.1	−0.09	255.8	−0.02	357.5	−0.07	513.6	−0.06
Excluding Oman	927	370.9	53.7	0.03	256.8	0.00	363.8	0.05	515.0	−0.04
Excluding USA	1018	369.9	53.9	0.01	257.1	0.01	361.9	0.01	517.0	0.00

For example, Ghanaian mothers often propped infants in a variety of ways to assist the infant's assumption of an upright sitting position. Norwegians, on the other hand, were encouraged by paediatric care norms not to push children to perform but to rely on a child's spontaneous interest and development, e.g. allowing infants to achieve an upright sitting position without assistance or prompting. The greater homogeneity in ages of achievement for milestones that require the most coordinated movements and control, namely hands-and-knees crawling and standing and walking alone, thus may be the least amenable to trainer "interference". However, this explanation merits further investigation.

Although the origins of inter-site heterogeneity in the ages of milestone achievement and differences in the degree of heterogeneity in the ages of achievement among the six milestones remain unclear, the implications of these analyses for the purposes of the MGRS appear straightforward. The ranges of observed ages of achievement amply document the variability of normal development in diverse cultural and environmental settings. Thus, given the health and environmental advantages inherent in the MGRS sample, pooling observations from all five sites appears to be the most appropriate manner to reflect

the range of normal development. This and other considerations led to the formulation of "windows of achievement" for specific milestones [30] that reflect the range of ages of achievement of motor milestones observed in the MGRS population. For reasons described in a companion paper in this supplement [30], these "windows" were estimated conservatively (as bounded by the 1st to 99th percentile age interval for individual milestone achievement).

Lastly, consideration is given to the relative contributions of inter-site and inter-individual differences to the total variability in ages of milestone achievement. The detailed evaluations of the roles of inter-site and inter-individual differences summarized in Tables III and IV are particularly informative. Clearly, the heterogeneity in ages of milestone achievement differs markedly among milestones (Table III). We suggest that milestones with the most homogeneous ages of achievement are likely to provide the most robust assessments of inherent inter-site differences, i.e. those that are least influenced by caretaker behaviours. Partitioning of variability for milestones with the most homogenous ages of achievement (hands-and-knees crawling, standing alone, and walking alone) attributes approximately 4% of the total variability to site differences and approximately 60% to inter-individual differences. The remaining 36% is ascribed to other sources of variation and random error, a proportion likely to be inflated by the random point method of determining ages of achievement and the inability to partition out a reasonable estimate of intra-individual variability. The 15-fold difference in the proportional contributions of inter-site and inter-individual differences are consistent with estimates of human genetic variability across and within populations. Population genomic analyses suggest that 85 to 90% of genetic variation resides within populations, whereas approximately 10 to 15% resides among populations [38]. The likely multigenic control of motor development suggests that variability between and within populations should be distributed similarly.

In summary, since these analyses found only small and sporadic differences in ages of achievement of gross motor milestones due to sex, we conclude they are of no practical relevance to the construction of gross motor development standards. Similarly, no significant site–sex interactions were observed. Significant differences among sites, however, were observed. Inter-site differences most likely reflect factors related to culture-specific care behaviours, but the aetiology of those differences cannot be discerned adequately from these analyses. Most importantly, however, these differences reflect the range of normal development among healthy populations across the relatively wide range of cultures and environments included in the MGRS, and they provide a useful basis for assessing motor development in populations. Lastly, the relative contributions of between- and within-site variability to the total variability across all six milestones are consistent with the relative contributions of those sources of variability to the total variability in child length discussed in a companion paper in this supplement [29]. These analyses support the appropriateness of pooling data from all sites for the purposes of developing an international standard for the six motor development milestones assessed by the MGRS.

Acknowledgements

This paper was prepared by Cutberto Garza, Mercedes de Onis, Reynaldo Martorell, Kathryn G. Dewey and Maureen Black on behalf of the WHO Multicentre Growth Reference Study Group. The statistical analysis was conducted by Amani Siyam.

References

[1] Garza C, de Onis M, for the WHO Multicentre Growth Reference Study Group. Rationale for developing a new international growth reference. Food Nutr Bull 2004;25 Suppl 1:S5–14.

[2] de Onis M, Garza C, Victora CG, Onyango AW, Frongillo EA, Martines J, for the WHO Multicentre Growth Reference Study Group. The WHO Multicentre Growth Reference Study: Planning, study design and methodology. Food Nutr Bull 2004;25 Suppl 1:S15–26.

[3] Cooper RS, Kaufman JS, Ward R. Race and genomics. N Engl J Med 2003;348:1166–70.

[4] Kariger PK, Stoltzfus RJ, Olney D, Sazawal S, Black R, Tielsch JM, et al. Iron deficiency and physical growth predict attainment of walking but not crawling in poorly nourished Zanzibari infants. J Nutr 2005;135:814–9.

[5] Black MM, Baqui AH, Zaman K, Ake Persson L, El Arifeen S, Le K, et al. Iron and zinc supplementation promote motor development and exploratory behavior among Bangladeshi infants. Am J Clin Nutr 2004;80:903–10.

[6] Kuklina EV, Ramakrishnan U, Stein AD, Barnhart HH, Martorell R. Growth and diet quality are associated with the attainment of walking in rural Guatemalan infants. J Nutr 2004;134:3296–300.

[7] Lozoff B, De Andraca I, Castillo M, Smith JB, Walter T, Pino P. Behavioral and developmental effects of preventing iron-deficiency anemia in healthy full-term infants. Pediatrics 2003; 112:978.

[8] Black MM. The evidence linking zinc deficiency with children's cognitive and motor functioning. J Nutr 2003;133: 1473S–6S.

[9] Jahari AB, Saco-Pollitt C, Husaini MA, Pollitt E. Effects of an energy and micronutrient supplement on motor development and motor activity in undernourished children in Indonesia. Eur J Clin Nutr 2000;54 Suppl 2:S60–8.

[10] Dewey KG, Cohen RJ, Brown KH, Rivera LL. Effects of exclusive breastfeeding for four versus six months on maternal nutritional status and infant motor development: results of two randomized trials in Honduras. J Nutr 2001;131:262–7.

[11] Lucas A, Fewtrell MS, Morley R, Singhal A, Abbott RA, Isaacs E, et al. Randomized trial of nutrient-enriched formula versus standard formula for postdischarge preterm infants. Pediatrics 2001;108:703–11.

[12] Carrel AL, Moerchen V, Myers SE, Bekx MT, Whitman BY, Allen DB. Growth hormone improves mobility and body composition in infants and toddlers with Prader-Willi syndrome. J Pediatr 2004;145:744–9.
[13] Chase C, Ware J, Hittelman J, Blasini I, Smith R, Llorente A, et al. Early cognitive and motor development among infants born to women infected with human immunodeficiency virus. Pediatrics 2000;106:E25.
[14] Campbell SK, Hedeker D. Validity of the Test of Infant Motor Performance for discriminating among infants with varying risk for poor motor outcome. J Pediatr 2001;139:546–51.
[15] Treuth MS, Sherwood NE, Butte NF, McClanahan B, Obarzanek E, Zhou A, et al. Validity and reliability of activity measures in African-American girls for GEMS. Med Sci Sports Exerc 2003;35:532–9.
[16] Allen MC, Alexander GR. Gross motor milestones in preterm infants: correction for degree of prematurity. J Pediatr 1990; 116:955–9.
[17] Garn SM, Petzold AS, Ridella SA, Johnston M. Effect of smoking during pregnancy on Apgar and Bayley scores. Lancet 1980;2:912–3.
[18] Lansdown RG, Goldstein H, Shah PM, Orley JH, Di G, Kaul KK, et al. Culturally appropriate measures for monitoring child development at family and community level: a WHO collaborative study. Bull World Health Organ 1996;74:283–90.
[19] Garcia Coll CT. Developmental outcome of minority infants: a process-oriented look into our beginnings. Child Dev 1990; 61:270–89.
[20] Pasamanick B. A comparative study of the behavioral development of Negro infants. J Genet Psychol 1946;69:3–44.
[21] Williams JR, Scott RB. Growth and development of Negro infants: IV. Motor development and its relationship to child rearing practices in two groups of Negro infants. Child Dev 1953;24:103–21.
[22] Geber M, Dean RF. The state of development of newborn African children. Lancet 1957;272:1216–9.
[23] Super CM. Cross-cultural research on infancy. In: Triandis HC, Heros A, editors. Handbook of cross-cultural psychology: Developmental psychology, vol. 4. Boston: Allyn & Bacon; 1981. p. 17–53.
[24] Pachter LM, Dworkin PH. Maternal expectations about normal child development in 4 cultural groups. Arch Pediatr Adolesc Med 1997;151:1144–50.
[25] Lima MC, Eickmann SH, Lima AC, Guerra MQ, Lira PI, Huttly SR, et al. Determinants of mental and motor development at 12 months in a low income population: a cohort study in northeast Brazil. Acta Paediatr 2004;93:969–75.
[26] To T, Guttmann A, Dick PT, Rosenfield JD, Parkin PC, Cao H, et al. What factors are associated with poor developmental attainment in young Canadian children. Can J Public Health 2004;95:258–63.
[27] Wijnhoven TM, de Onis M, Onyango AW, Wang T, Bjoerneboe GE, Bhandari N, et al. for the WHO Multicentre Growth Reference Study Group. Assessment of gross motor development in the WHO Multicentre Growth Reference Study. Food Nutr Bull 2004;25 Suppl 1:S37–45.
[28] WHO Multicentre Growth Reference Study Group. Reliability of motor development data in the WHO Multicentre Growth Reference Study. Acta Paediatr Suppl 2006;450:47–55.
[29] WHO Multicentre Growth Reference Study Group. Assessment of differences in linear growth among populations in the WHO Multicentre Growth Reference Study. Acta Paediatr Suppl 2006;450:56–65.
[30] WHO Multicentre Growth Reference Study Group. WHO Motor Development Study: Windows of achievement for six gross motor development milestones. Acta Paediatr Suppl 2006;450:86–95.
[31] Kirk RE. Experimental design. 2nd ed. Belmont, CA: Brooks/Cole; 1992.
[32] Goldstein H. Multilevel statistical models. 2nd ed. Kendall's Library of Statistics 3. London: Arnold Publications; 1995.
[33] Johnson NL, Kotz S. Continuous univariate distributions. In: The Houghton Mifflin Series in Statistics, vol. 1. New York: John Wiley and Sons; 1970.
[34] Lockman JJ, Thelen E. Developmental biodynamics: brain, body, behavior connections. Child Dev 1993;64:953–9.
[35] Thelen E. Motor development. A new synthesis. Am Psychol 1995;50:79–95.
[36] Gibson G, Wagner G. Canalization in evolutionary genetics: a stabilizing theory? Bioessays 2000;22:372–80.
[37] Davidson EH, Rast JP, Oliveri P, Ransick A, Calestani C, Yuh CH, et al. A genomic regulatory network for development. Science 2002;295:1669–78.
[38] Jorde LB, Wooding SP. Genetic variation, classification and 'race'. Nat Genet 2004;36 Suppl 11:S28–33.