Original Article

Diagnostic value of maternal anthropometric measurements for predicting low birth weight in developing countries: a meta-analysis

Eita Goto MD, MM, MPH

Nagoya Medical Science Research Institute, Meitou-ku, Nagoya, Japan

Objective: Weighing scales are often lacking at home birth in developing countries. Therefore, simple, reliable, and inexpensive methods for detecting low birth weight especially before birth would be useful. This study was performed to evaluate the diagnostic value of maternal anthropometric measurements for predicting low birth weight. **Methods**: Bivariate diagnostic meta-analysis was conducted to construct hierarchical summary receiver operating characteristic curves. All English language studies included in the meta-analysis enrolled apparently healthy pregnant women and provided the data necessary to construct two-by-two tables (i.e., true positive, false positive, false negative, and true negative values). Ten data bases, including PubMed, were searched to identify these studies. **Results**: A sufficient number of studies involving 309,419 women paired with their newborns in Africa, Asia, Europe, Latin America, the Middle East, and Oceania, included data on maternal height, weight, arm circumference, body mass index, and weight gain during pregnancy (n=85, 80, 23, 51, and 16, respectively) to provide generalizable findings. However, sensitivity of 0.46 (95% confidence interval (CI)=0.35-0.56) to 0.63 (95% CI=0.54-0.71), specificity of 0.55 (95% CI=0.42-0.67) to 0.71 (95% CI=0.61-0.80), and diagnostic odds ratios of 2 (95% CI=2-2) to 4 (95% CI=3-5) were not sufficiently high for primary screening. The generalizability of abdominal circumference data could not be guaranteed due to the limited sample (one article). **Conclusions**: Maternal anthropometric measurements are unsuitable for predicting low birth weight.

Key Words: anthropometry, low birth weight, meta-analysis, pregnant women, sensitivity and specificity

INTRODUCTION

Low birth weight (<2500 g) contributes to infant mortality.[1] Weighing scales and professional birth attendants may be unavailable at home delivery in outlying areas largely of developing countries.[2-4] Therefore, simple, reliable, and inexpensive methods for detecting low birth weight especially before birth would be useful. At present, the best anthropometric indicator of low birth weight among newborns delivered by apparently healthy pregnant women in developing countries is newborn chest circumference,[3,5] but maternal anthropometric measurements may be preferable.

In 1995, the World Health Organization (WHO) published a meta-analysis suggesting the predictive capability of maternal anthropometric measurements based on their positive associations with low birth weight, i.e., odds ratio (>1).[6] However, sensitivity (0.30-0.77) and specificity (0.71-0.85) in other studies did not necessarily confirm their usefulness in primarily screening for low birthweight.[7,8] No meta-analysis has yet determined whether the diagnostic value of maternal anthropometric measurements is sufficiently high based on sensitivity, specificity, likelihood ratios (LRs), and diagnostic odds ratio (DOR), as well as area under the curve (AUC).[6,9-13]

Here, bivariate diagnostic meta-analysis was performed and hierarchical summary receiver operating characteristic (HSROC) curves were drawn[14-16] to summarize the data of studies the quality of which was assessed using the Quality Assessment of Diagnostic Accuracy Studies (QUADAS),[17] with the aim to evaluate the diagnostic value of maternal anthropometric measurements for predicting low birth weight in developing countries.

METHODS

Primary outcome and selection criteria

The primary outcomes were sensitivity and specificity for maternal height; weight; head; chest, (mid-upper) arm, abdominal, and calf circumferences; biceps, triceps, and total skinfold thicknesses; body mass index (BMI); and weight gain during pregnancy to predict low birth weight. However, DORs and AUCs for these measurements were also included. This meta-analysis included all English language studies in apparently healthy pregnant women that provided the data necessary to construct two-by-two tables (i.e., true positive, false positive, false negative, and true negative values).

Corresponding Author: Dr Eita Goto, Nagoya Medical Science Research Institute 1-118 Kamenoi, Meitou-ku, Nagoya 465-0094, Japan.
Tel: +81-52-702-0941; Fax: +81-52-702-0975
Email: egoto1@nifty.com
Manuscript received 01 June 2014. Initial review completed 18 July 2014. Revision accepted 18 July 2014.
doi: 10.6133/apjcn.2015.24.2.05

Search strategy, article collection, and data extraction

PubMed (MEDLINE) was queried by scanning titles and abstracts using the Falck-Ytter filter[18] adding key terms of the above-mentioned maternal anthropometric measurements and ("low birth weight" OR "low birthweight") to identify the articles reporting one or more studies that provided true positive, false positive, false negative, and true negative results (October 2013). No limitation was set on the publication period. Occasionally, one article could consist of multiple studies due to the use of multiple samples, measurements, thresholds, and/or periods of measurement. In addition, the aims of the studies were not limited to evaluating this prediction as long as they produced extractable results. The articles that were unavailable online were collected from the Library of the Japan Medical Association. The full texts of the articles were retrieved, partly to calculate the missing true positive, false positive, false negative, and/or true negative values from some of the existing data, e.g., number of participants, prevalence of low birth weight, and the true positive, false positive, false negative, and/or true negative values that were not missing. The articles judged to be unrelated, not to provide all data for two-by-two tables, or to be reviews were excluded; the remaining articles and their reported studies became potentially eligible. The "related citations" listed by clicking the "See all" tabs on the PubMed screens displaying potentially eligible articles and the reviews were scanned to identify additional potentially eligible articles. Similarly, the bibliographies of the potentially eligible articles and the reviews were checked. Searches in EMBASE, CINAHL, PsycINFO, Wiley InterScience, ProQuest Health and Medical Complete™, ProQuest Dissertations & Theses Database, the entire Cochrane Library (e.g., CENTRAL), Google Scholar, and Scopus identified additional articles that were subjected to the same process. Before statistical analysis, the conflicts between all of the originally available and missing but calculated data were checked. The potentially eligible articles and studies left after excluding those with conflicting data were finally included in the meta-analysis. All of the true positive, false positive, false negative, and true negative values in all finally included studies were extracted and used as input for the meta-analysis. The literature search and data collection were performed by a single observer but were periodically repeated.

Study quality assessment

QUADAS, comprising 14 items,[17] was used (Table 1). The QUADAS score (0-14) was defined as the sum of allocated numbers, with a value of "1" assigned for a "yes" response to each item, otherwise a value of "0" was assigned. Subgroup analysis compared the results between QUADAS $\geq$10 vs <10. The reason for using this cut-off point is described in "Quality assessment" of a previously published article.[5]

Statistical analysis

Stata/MP 13.1 (StataCorp LP, College Station, TX, USA) was used for statistical analysis. The potential outliers were identified using spike plots of the Cook's distance to show the most influential data points and using scatter plots of the standardized residuals of healthy (x-axis) and diseased (y-axis) samples for each study. The potential outliers with flaws that the remaining studies did not have in study design were considered to be true outliers and were therefore omitted from subsequent analysis.

Bivariate meta-analysis provided pooled sensitivity and specificity, LRs, and DORs.[16,19] The HSROC curve was also constructed showing the AUC and the 95% confidence contour and 95% prediction contour, i.e., a given probability (e.g., 95%) of including true sensitivity and specificity of a future study.[14-16] The following guidelines were used to interpret AUC for low, i.e., 0.5≤AUC≤0.7, moderate, i.e., 0.7≤AUC≤0.9, and high, i.e., 0.9≤AUC≤1.0, diagnostic accuracy[20] and to interpret positive and negative LRs for small (and rarely important), i.e., LRs =1-2 and 0.5-1, small (but sometimes important), i.e., LRs=2-5 and 0.2-0.5, moderate, i.e., LRs=5-10 and 0.1-0.2, and conclusive, i.e., LRs >10 or <0.1, informational value.[21]

Heterogeneity was assessed using an I^2 statistic of 0%-100%.[22] Attempts were made to reduce heterogeneity from I^2>50% to I^2≤50% by eliminating potential confounders, i.e., changing study regions or responses to each QUADAS item (sensitivity analysis). Sensitivity and specificity and DORs were pooled separately by subdividing into Africa, Asia, Europe, Latin America, the

Table 1. Quality Assessment of Diagnostic Accuracy Studies (QUADAS) tool (14 items)

1	Was the spectrum of patients representative of the patients who will receive the test in practice?
2	Were selection criteria clearly described?
3	Is the reference standard likely to correctly classify the target condition?
4	Is the time period between the reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?
5	Did the whole sample or a random selection of the sample, receive verification using a reference standard of diagnosis?
6	Did patients receive the same reference standard regardless of the index test result?
7	Was the reference standard independent of the index test (i.e., the index test did not form part of the reference standard)?
8	Was the execution of the index test described in sufficient detail to permit its replication?
9	Was the execution of the reference standard test described in sufficient detail to permit its replication?
10	Were the index test results interpreted without knowledge of the results of the reference standards?
11	Were the reference standard results interpreted without knowledge of the results of the reference standards?
12	Were the same clinical data available when test results were interpreted as would be available when the test is used in practice?
13	Were uninterpretable/ intermediate test results reported?
14	Were withdrawals from the study explained?

Middle East, and Oceania; QUADAS scores of ≥10 and <10; and pre- and post-delivery measurements (subgroup analysis). Bivariate meta-regression analysis was also conducted to evaluate the association of sensitivity and specificity with covariates, i.e., study characteristics including Africa, Asia, Europe, Latin America, the Middle East, or Oceania vs other regions, QUADAS scores (≥10 vs <10) and responses to QUADAS items, periods of measurement (i.e., post-delivery vs others), and the three major sources of bias. The three major sources of bias were: 1) giving the same reference tests to all participants, irrespective of the index test outcomes; 2) recruiting the clinical population (cohort study) rather than the diseased population plus the control group (case-control study); and 3) collecting data prospectively rather than retrospectively.[23,24] Publication bias was assessed using Deeks' funnel plot asymmetry tests.[25] Welch's *t* test was used to determine the statistical significance of differences in DOR by changing anthropometric measurements, i.e., maternal vs newborn anthropometric measurements,[5] under the assumption that their log-transformation could approximate normal distributions.[16,19] Efforts to propose a cut-off point for each anthropometric measurement were made according to the Youden index, i.e., the point on the HSROC curve with the greatest distance to a straight line drawn at an angle of 45° from the origin.[26]

RESULTS

Among 546 articles collected for full text retrieval, 121 were potentially eligible (Figure 1). After excluding eight articles with conflicts between the originally available and the missing but calculated data and 42 articles from developed countries, 71 articles were selected for final analysis (Table 2). From these 71 articles, 259 studies involving 309,419 women paired with their newborns were finally included in the analysis, while a study by Chumnijarakij et al[27] regarding weight was omitted as a true outlier. Height was evaluated in 85 studies, weight in 80, arm circumference in 23, abdominal circumference in 4, BMI in 51, and weight gain in 16 (Tables 2 and 3). The studies extracted from two articles by Bisai et al were not considered duplicates (Table 2) because of post-delivery weight in one article and pre-delivery weight in the other. One study evaluating biceps skinfold thickness could not be subjected to meta-analysis.[28] None of the included studies evaluated other maternal anthropometric measurements. The studies were conducted in various regions of Africa, Asia, Europe, Latin America, the Middle East, and Oceania (Table 2). Only 33 studies aimed to evaluate the diagnostic accuracy of maternal anthropometric measurements to predict low birth weight. Other studies mainly aimed to identify the risk of low birth weight from maternal biological and/or socioeconomic characteristics. Based on QUADAS, however, overall study quality was relatively well controlled, as Figure 2 shows narrow white vs wide black and gray spaces.

With the exception of a study by Chumnijarakij et al[27] regarding weight, each potential outlier was not omitted as a true outlier. No omission was because each potential outlier was one of a series of studies among which cut-off points varied and some of which could not be omitted in case others should be analyzed. There was no evidence of a high diagnostic value of maternal anthropometric measurements for predicting low birth weight (Table 3 and Figure 3). For height, weight, arm circumference, BMI and weight gain, the pooled sensitivity of 0.46 (95% confidence interval (CI)=0.35-0.56) to 0.63 (95% CI=0.54-0.71), specificity of 0.55 (95% CI=0.42-0.67) to 0.71 (95% CI=0.61-0.80), and DORs of 2 (95% CI=2-2) to 4 (95% CI=3-5) were not high among the total population. Further, these measurements had low diagnostic accuracy and small (and rarely important) or at best small (but sometimes important) informational value, based on the AUC and LR interpretation guidelines.[20,21] Another concern regarding the practical use of maternal anthropometric measurements was larger 95% confidence and prediction contours. Abdominal circumference alone had moderate diagnostic accuracy. However, abdominal circumference also showed insufficient levels of pooled sensitivity (0.66, 95% CI=0.54-0.77), specificity (0.77, 95%

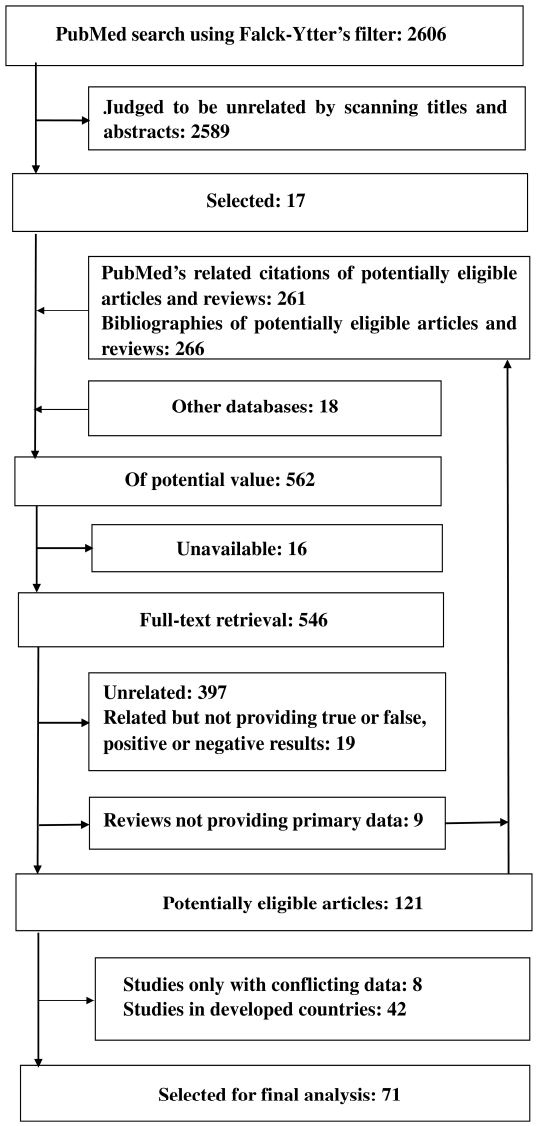

Figure 1. Flow diagram of article selection process. As many as 71 articles finally selected for material height (*n*=40), weight (*n*=31), arm circumference (*n*=12), body mass index (*n*=23), and weight gain (*n*=9), unlike abdominal circumference (*n*=1), may reflect generalizability to guarantee external validity (i.e., generalizability to or across different populations and settings), thus supporting the soundness of the conclusions.

Table 2. Characteristics of studies identified from 10 online databases, reference harvesting and PubMed

Author	Source	Country	Study design	Anthropometry (number of participating women)
Amin	Indian J Pediatr. 1993;60:269-74.	India	Case-control	Height (102); BMI (102)
Andersson	Trop Med Int Health. 1997;2:1080-7.	Central African Rep.	Cohort	Weight (1477)
Azimul and Matin	J Dhaka Med Coll. 2009;18:64-9 and 83-7.	Bangladesh	Case-control	Height (573); weight (583); BMI (573)
Awoleke	Arch Gynecol Obstet. 2012;285:1-6.	Nigeria	Case-control	Height (419)
Badshah	BMC Public Health. 2008;8:197.	Pakistan	Cohort	Height (1006); weight (1009); BMI (998)
Belizán†	Bull Pan Am Health Organ. 1989;23:414-23.	Argentina	Cohort	Height (120)
Bisai [A]‡	Environment Concerns and Perspective. New Delhi: APH Publishing Corporation; 2007. pp. 65-80.	India	Cohort	Weight (139)
Bisai [B]† and ‡	Ann Hum Biol. 2007;34:91-101.	India	Cohort	Weight (295)
Bisai	Coll Antropol. 2009;3:725-8.	India	Cohort	Weight (233)
Bondevik	Acta Obstet Gynecol Scand. 2001;80:402-8.	Nepal	Case-control	Height (1366); BMI (1236)
Chan	J Obstet Gynaecol Res. 2009;35:307-14.	China	Cohort	Height (13606)
Chhabra	Asia Pac J Public Health. 2004;16:95-8.	India	Cohort	Height (450); weight (450)
Chumnijarakij	J Med Assoc Thai. 1992;75:445-52.	Thailand	Case-control	Height (6095); weight (6095); BMI (6095)
Das	Bangladesh Med Res Counc Bull. 1997;23:10-5.	Bangladesh	Cohort	Height (150); weight (150); arm circumference (150)
Dasgupta	Indian J Public Health. 2004;48:218-0.	India	Cohort	Height (343); weight (315)
Dhar†	Bangladesh Med Res Counc Bull. 2008;34:64-6.	Bangladesh	Cohort	Height (316); weight (316); arm circumference (316)
Dinh	Ann Trop Paediatr. 1996;16:327-33.	Vietnam	Cohort	BMI (1375)
Ferreira	Indian J Community Med. 1991;16:106-9.	India	Cohort	BMI (105)
Ganesh Kumar	Indian J Pediatr. 2010;77:87-9.	India	Case-control	Height (450); weight (450)
Gazali	Paediatr Indones. 1987;27:1-9.	Indonesia	Cohort	Height (1067)
Gebremariam	East Afr Med J. 2005;82:554-8.	Ethiopia	Cohort	Height (588)
Ghosh	Indian Pediatr. 1977;14:107-14.	India	Cohort	Height (3625)
Hirve†	Indian Pediatr. 1994;31:1221-5.	India	Cohort	Weight (1922)
Hosain	J Trop Pediatr. 2006;52:87-91.	Bangladesh	Cohort	Height (350); weight (350); arm circumference (350); BMI (350); weight gain (350)
Isaranurug	J Med Assoc Thai. 2007;90:2559-64.	Thailand	Cohort	Height (3118); weight gain (2221)
Jafari	Public Health. 2010;124:153-8.	Iran	Cohort	Height (4510); weight (3621); BMI (3621)
Jain	J Obstet Gynaecol India. 2012;62:429-31.	India	Cohort	BMI (300)
Janjua	Public Health Nutr. 2009;12:789-98.	Pakistan	Cohort	Arm circumference (540); BMI (540)
Jariyapitaksakul	J Med Assoc Thai. 2013;96:259-65.	Thailand	Case-control§	Weight gain (1152)
Kamalados	Indian J Pediatr. 1992;59:299-304.	India	Cohort	Height (268); weight (268)

BMI: body mass index; RCT: randomized controlled trial; QUADAS: quality assessment of diagnostic studies.

†Belizán's study on weight, Bisai's study on weight only with the cut-off point of 47 kg, Dhar's study on height only with the cut-off point of 156 cm, Karim's studies on weight only with the cut-off point of 45 kg, on arm circumference and on BMI only with the cut-off point of 22.5 kg/m^2, Lawyoin's studies (1992) on arm circumference, Mohany's studies (2006) on arm circumference only with the cut-off point of 22 cm and on BMI only with the cut-off points of 19 and 19.5 kg/m^2, and Walraven's study (1995) on height and arm circumference are excluded because of disparities between calculated values of true and false positive and true and false negative results and other existing data in the articles. Hirve's study on height is excluded because of not determining one cut-off point. Zerfas' study on weight gain only with the cut-off point of 16% (8 kg) is excluded because it was conducted in the US.

‡Not regarded as a duplication, because maternal weight was measured at post-delivery in Bisai [A] and at pre-delivery in Bisai [B].

§Case group involved small-for-gestational-age rather than low birth weight.

¶Lechtig's studies evaluated weight gain during pregnancy (≥16th vs <16th percentile) and weight gain as a percentage of weight for height (≥90% vs <90%) and Mangklabruks' studies evaluated weight gain during, the 2nd trimester, 3rd trimester, and all trimesters, respectively.

††Height in urban areas and rural areas, respectively.

Table 2. Characteristics of studies identified from 10 online databases, reference harvesting and PubMed (cont.)

Author	Source	Country	Study design	Anthropometry (number of participating women)
Karim[†]	Ann Hum Biol. 1997;24:387-401.	Bangladesh	Cohort	Weight (247); BMI (247)
Ko	J Nurs Res. 2002;10:83-9.	China	Cohort	Height (620)
Klufio	P N G Med J. 1997;40:136-45.	Papua New Guinea	Case-control	Weight gain (666)
Lawyoin[†]	Afr J Med Sci. 1992;21:33-9.	Nigeria	Cohort	Height (452)
Lawyoin	East Afr Med J. 1993;70:746-8.	Nigeria	Cohort	Weight (913)
Lechtig	J Trop Pediatr. 1988;34:34-41.	Guatemala	Cohort	Arm circumference (445); weight gain (7061,71)[¶]
Maddah	Eur J Clin Nutr. 2005;59:1208-12.	Iran	Cohort	BMI (704); weight gain (1916)
Makki	Ann Saudi Med. 2002;22:333-5.	Yemen	Cohort	Height (2190); weight (2142); arm circumference (2222)
Malik	Indian J Pediatr. 1997;64:373-7.	India	Cohort	Height (984); weight (984)
Mangklabruks	J Med Assoc Thai. 2012;95:358-65.	Thailand	Cohort	Height (2182); weight (2183); arm circumference (2172); BMI (2182); weight gain (2050, 2102, 1491)[¶]
Mobasheri	J Med Sci. 2006;7:905-8.	Iran	Cohort	BMI (315)
Mohanty	J Trop Pediatr. 2000;46:363-4.	India	Cohort	Abdominal circumference (151)
Mohanty[†]	J Trop Pediatr. 2006;52:24-9.	India	Cohort	Height (395); weight (395); arm circumference (395) BMI (395)
Mo-Suwan	J Med Assoc Thai. 1989; 72 Suppl 1:52-6.	Thailand	Cohort	Height (292); arm circumference (252)
Nahar	Am J Clin Nutr. 2000;72:1010-7.	Bangladesh	Cohort	Height (660); weight (660)
Neyzi	Hum Biol. 1987;59:387-8.	Turkey	Cohort	Height (1191); weight (1153) Arm circumference (1236); BMI (1133)
Niyogi	J Indian Med Assoc. 1963;40:64-8.	India	Cohort	Height (121); weight (122)
Oduntan	Trop Geogr Med.1976;28:220-3.	Nigeria	Cohort	Height (541, 571)[††]
Oni	East Afr Med J. 1986;63:121-30.	Nigeria	Cohort	Height (205)
Panahadeh	Iran J Med Sci. 2007;32:36-9.	Iran	Cohort	BMI (480)
Rowshan	Bangladesh Med Res Counc Bull. 1978;4:1-9.	Bangladesh	Cohort	Weight (100)
Roy	J Fam Welf. 2009;55:79-83.	India	Cohort	Height (462); weight gain (470)
Saereeporncharenkul	J Med Assoc Thai. 2011;94 Suppl 2:S52-8.	Thailand	Cohort	BMI (3715)
Sahu	J Obstet Gynaecol Res. 2007;33:655-9.	India	Cohort	BMI (380)
Saigal	Indian Pediatr. 1969;6:773-82.	India	Cohort	Height (890)
Sebayang	Trop Med Int Health. 2012;17:938-50.	Indonesia	RCT	Height (11 600); arm circumference (11, 600)
Sharma	Internet J Health. 2009;9: doi: 10.5580/10f1.	India	Cohort	Weight (193)
Silva	Paediatr Perinat Epidemiol. 2001;15:257-64.	Brazil	Cohort	Height (2225)

BMI: body mass index; RCT: randomized controlled trial; QUADAS: quality assessment of diagnostic studies.

[†]Belizán's study on weight, Bisai's study on weight only with the cut-off point of 47 kg, Dhar's study on height only with the cut-off point of 156 cm, Karim's studies on weight only with the cut-off point of 45 kg, on arm circumference and on BMI only with the cut-off point of 22.5 kg/m^2, Lawyoin's studies (1992) on arm circumference, Mohany's studies (2006) on arm circumference only with the cut-off point of 22 cm and on BMI only with the cut-off points of 19 and 19.5 kg/m^2, and Walraven's study (1995) on height and arm circumference are excluded because of disparities between calculated values of true and false positive and true and false negative results and other existing data in the articles. Hirve's study on height is excluded because of not determining one cut-off point. Zerfas' study on weight gain only with the cut-off point of 16% (8 kg) is excluded because it was conducted in the US.

[‡]Not regarded as a duplication, because maternal weight was measured at post-delivery in Bisai [A] and at pre-delivery in Bisai [B].

[§]Case group involved small-for-gestational-age rather than low birth weight.

[¶]Lechtig's studies evaluated weight gain during pregnancy (≥16th vs <16th percentile) and weight gain as a percentage of weight for height (≥90% vs <90%) and Mangklabruks' studies evaluated weight gain during, the 2nd trimester, 3rd trimester, and all trimesters, respectively.

[††]Height in urban areas and rural areas, respectively.

Table 2. Characteristics of studies identified from 10 online databases, reference harvesting and PubMed (cont.)

Author	Source	Country	Study design	Anthropometry (number of participating women)
Singh	MJAFI. 2009;65:10-2.	India	Case-control	BMI (340)
Siza	Tanzanian J Health Res. 2008;10:1-8.	Tanzania	Cohort	BMI (2515)
Taha	Paediatr Perinat Epidemiol. 1995;9:185-200.	Sudan	Case-control	Weight (1003)
Tin	Malays J Reprod Health. 1994;12:32-7.	Myanmar	Cohort	Height (2613)
Tyagi	Indian Pediatr. 1985;22:507-14.	India	Cohort	Height (341)
Walraven†	Br J Obstet Gynaecol. 1995;102:525-9.	Tanzania	Cohort	Weight (1509)
Walraven	Trop Med Int Health. 1997;2:558-67.	Tanzania	Cohort	Height (2142); weight (2142)
Wannous	East Mediterr Health J. 2001;7:966-74.	Syria	Cohort	Weight (862)
Wong	J Am Diet Assoc. 2000;100:791-6.	China	Cohort	Weight gain (745)
Wongcharoenkiat	J Med Assoc Thai. 2006; 89 Suppl 4:S65-9.	Thailand	Cohort	Height (660)
Yadav	Med J Malaysia. 2013;68:44-7.	Malaysia	Cohort	BMI (666)
Zerfas†	Maternal Nutrition and Pregnancy. Washington DC: Pan American Health Organization; 1991. pp. 138-51.	Guatemala	Unclear	Arm circumference (445); weight gain (71)

BMI: body mass index; RCT: randomized controlled trial; QUADAS: quality assessment of diagnostic studies.

†Belizán's study on weight, Bisai's study on weight only with the cut-off point of 47 kg, Dhar's study on height only with the cut-off point of 156 cm, Karim's studies on weight only with the cut-off point of 45 kg, on arm circumference and on BMI only with the cut-off point of 22.5 kg/m^2, Lawyoin's studies (1992) on arm circumference, Mohany's studies (2006) on arm circumference only with the cut-off point of 22 cm and on BMI only with the cut-off points of 19 and 19.5 kg/m^2, and Walraven's study (1995) on height and arm circumference are excluded because of disparities between calculated values of true and false positive and true and false negative results and other existing data in the articles. Hirve's study on height is excluded because of not determining one cut-off point. Zerfas' study on weight gain only with the cut-off point of 16% (8 kg) is excluded because it was conducted in the US.

‡Not regarded as a duplication, because maternal weight was measured at post-delivery in Bisai [A] and at pre-delivery in Bisai [B].

§Case group involved small-for-gestational-age rather than low birth weight.

¶Lechtig's studies evaluated weight gain during pregnancy (≥16th vs <16th percentile) and weight gain as a percentage of weight for height (≥90% vs <90%) and Mangklabruks' studies evaluated weight gain during, the 2nd trimester, 3rd trimester, and all trimesters, respectively.

††Height in urban areas and rural areas, respectively.

Table 3. Results of meta-analysis

Measurement	Population (number of studies)	Number of participating women	SENS 95% CI	SPEC 95% CI	PLR 95% CI	NLR 95% CI	DOR 95% CI	I^2 (%) 95% CI
Height	Total (n=85)	137,420	0.46 (0.35, 0.56)	0.71 (0.61, 0.80)	1.6 (1.4, 1.8)	0.77 (0.71, 0.84)	2 (2, 2)	100 (100, 100)
	Africa (n=19)	10,933	0.45 (0.19, 0.73)	0.72 (0.43, 0.90)	1.6 (1.2, 2.2)	0.77 (0.62, 0.95)	2 (2, 3)	100 (100, 100)
	Asia (n=58)	109,591	0.42 (0.32, 0.53)	0.73 (0.62, 0.81)	1.5 (1.3, 1.8)	0.80 (0.73, 0.86)	2 (2, 2)	100 (100, 100)
	The Middle East (n=6)	13,462	0.83 (0.48, 0.96)	0.40 (0.11, 0.78)	1.4 (0.9, 2.1)	0.42 (0.24, 0.74)	3 (2, 3)	100 (100, 100)
	QUADAS ≥10 (n=6)	9,728	0.62 (0.31, 0.86)	0.59 (0.28, 0.84)	1.5 (1.0, 2.3)	0.64 (0.40, 1.03)	2 (1, 5)	100 (99, 100)
	QUADAS <10 (n=79)	127,692	0.44 (0.33, 0.55)	0.72 (0.62, 0.81)	1.6 (1.4, 1.8)	0.78 (0.72, 0.85)	2 (2, 2)	100 (100, 100)
Weight	Total (n=80)	58,316	0.63 (0.54, 0.71)	0.69 (0.60, 0.76)	2.0 (1.7, 2.4)	0.54 (0.46, 0.63)	4 (3, 5)	100 (100, 100)
	Africa (n=16)	20,176	0.71 (0.45, 0.88)	0.75 (0.51, 0.90)	2.9 (1.5, 5.5)	0.38 (0.20, 0.73)	8 (3, 19)	100 (100, 100)
	Asia (n=56)	22,783	0.60 (0.52, 0.69)	0.66 (0.57, 0.74)	1.8 (1.6, 2.0)	0.60 (0.53, 0.67)	3 (3, 3)	100 (100, 100)
	The Middle East (n=8)	15,357	0.58 (0.23, 0.86)	0.66 (0.30, 0.90)	1.7 (1.2, 2.4)	0.64 (0.44, 0.94)	3 (2, 3)	100 (100, 100)
	Post-delivery (n=34)	19,158	0.59 (0.45, 0.72)	0.68 (0.54, 0.79)	1.9 (1.5, 2.2)	0.60 (0.50, 0.72)	3 (2, 4)	100 (100, 100)
	Pre-delivery (n=20)	13,072	0.62 (0.51, 0.72)	0.75 (0.64, 0.84)	2.5 (1.6, 3.8)	0.51 (0.37, 0.69)	5 (2, 10)	100 (99, 100)
	QUADAS ≥10 (n=20)	12,791	0.62 (0.49, 0.73)	0.65 (0.52, 0.75)	1.7 (1.4, 2.1)	0.60 (0.49, 0.72)	3 (2, 4)	100 (100, 100)
	QUADAS <10 (n=60)	45,525	0.64 (0.52, 0.74)	0.70 (0.59, 0.79)	2.1 (1.7, 2.6)	0.52 (0.43, 0.64)	4 (3, 5)	100 (100, 100)
Arm circumference	Total (n=23)	28,070	0.58 (0.42, 0.74)	0.61 (0.45, 0.75)	1.5 (1.3, 1.7)	0.68 (0.57, 0.82)	2 (2, 3)	100 (100, 100)
Abdominal circumference	Total (n=4)	604	0.66 (0.54, 0.77)	0.77 (0.63, 0.87)	2.9 (1.9, 4.5)	0.43 (0.34, 0.55)	7 (4, 11)	92 (85, 99)
	Asia (n=4)	604	0.66 (0.54, 0.77)	0.77 (0.63, 0.87)	2.9 (1.9, 4.5)	0.43 (0.34, 0.55)	7 (4, 11)	92 (85, 99)
BMI	Total (n=51)	60,812	0.61 (0.47, 0.73)	0.55 (0.42, 0.67)	1.4 (1.2, 1.5)	0.71 (0.61, 0.83)	2 (2, 2)	100 (100, 100)
	Asia (n=35)	37,451	0.57 (0.42, 0.71)	0.60 (0.46, 0.73)	1.4 (1.2, 1.7)	0.72 (0.60, 0.86)	2 (1, 3)	100 (100, 100)
	The Middle East (n=12)	13,301	0.64 (0.37, 0.84)	0.48 (0.25, 0.71)	1.2 (1.1, 1.4)	0.75 (0.59, 0.96)	2 (1, 2)	100 (100, 100)
Weight gain	Total (n=16)	24,197	0.62 (0.47, 0.74)	0.60 (0.47, 0.73)	1.6 (1.3, 1.8)	0.63 (0.52, 0.77)	2 (2, 3)	100 (100, 100)
	Asia (n=10)	12,496	0.57 (0.41, 0.72)	0.66 (0.53, 0.77)	1.7 (1.4, 2.0)	0.65 (0.51, 0.81)	3 (2, 4)	100 (100, 100)

SENS: sensitivity; SPEC: specificity; PLR: positive likelihood ratio; NLR: negative likelihood ratio; DOR: diagnostic odds ratio; CI: confidence interval; QUADAS: quality assessment of diagnostic accuracy studies; BMI: body mass index.

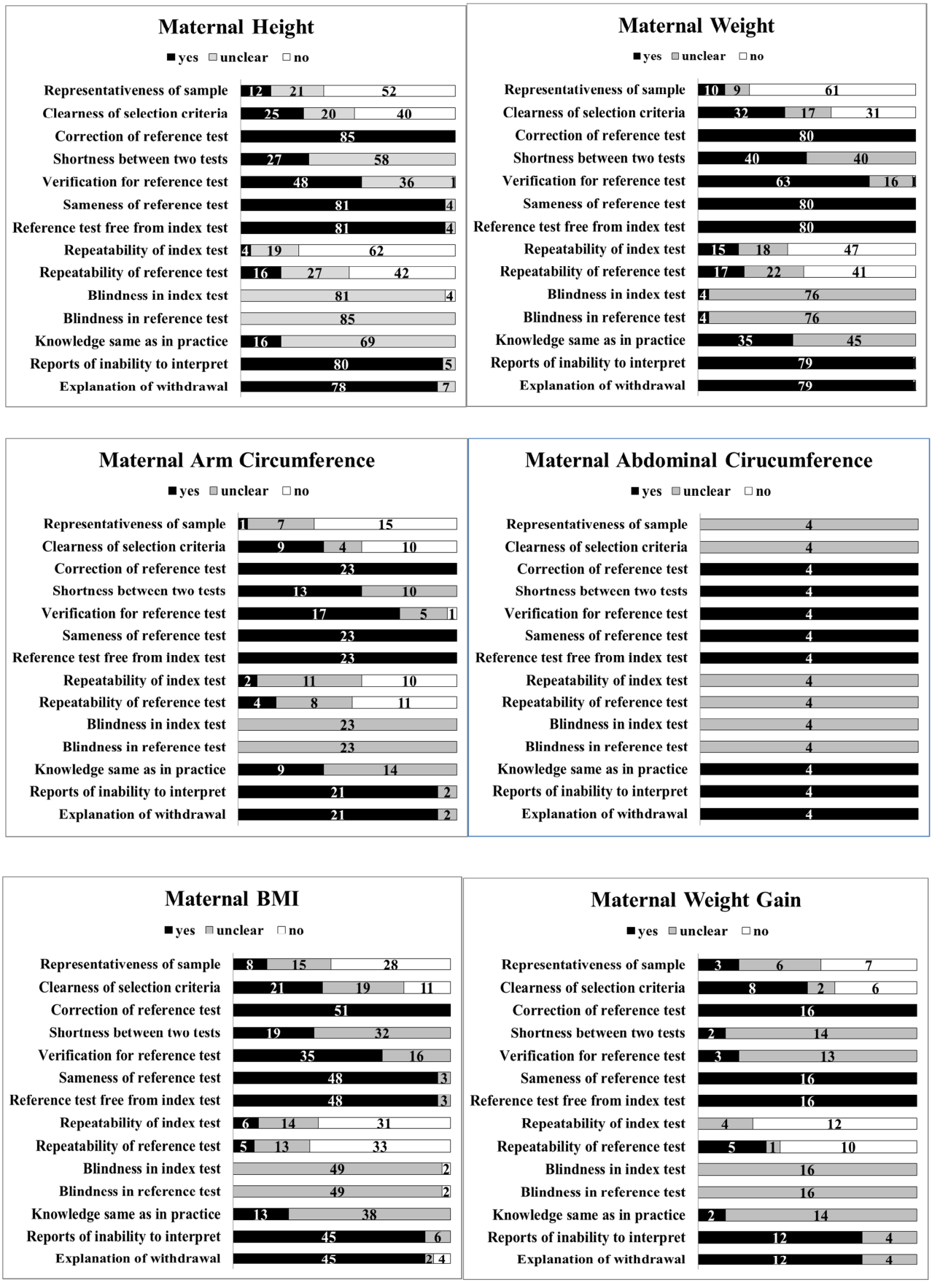

Figure 2. Results of study quality assessment using the Quality Assessment of Diagnostic Accuracy Studies (QUADAS) tool comprised of 14 question items. Relatively well-controlled study quality, as shown by narrow white vs wide black and grey spaces, may minimize risks of bias to guarantee internal validity (i.e., studies that were designed with sufficient rigour to minimize bias), thus supporting the soundness of the conclusions.

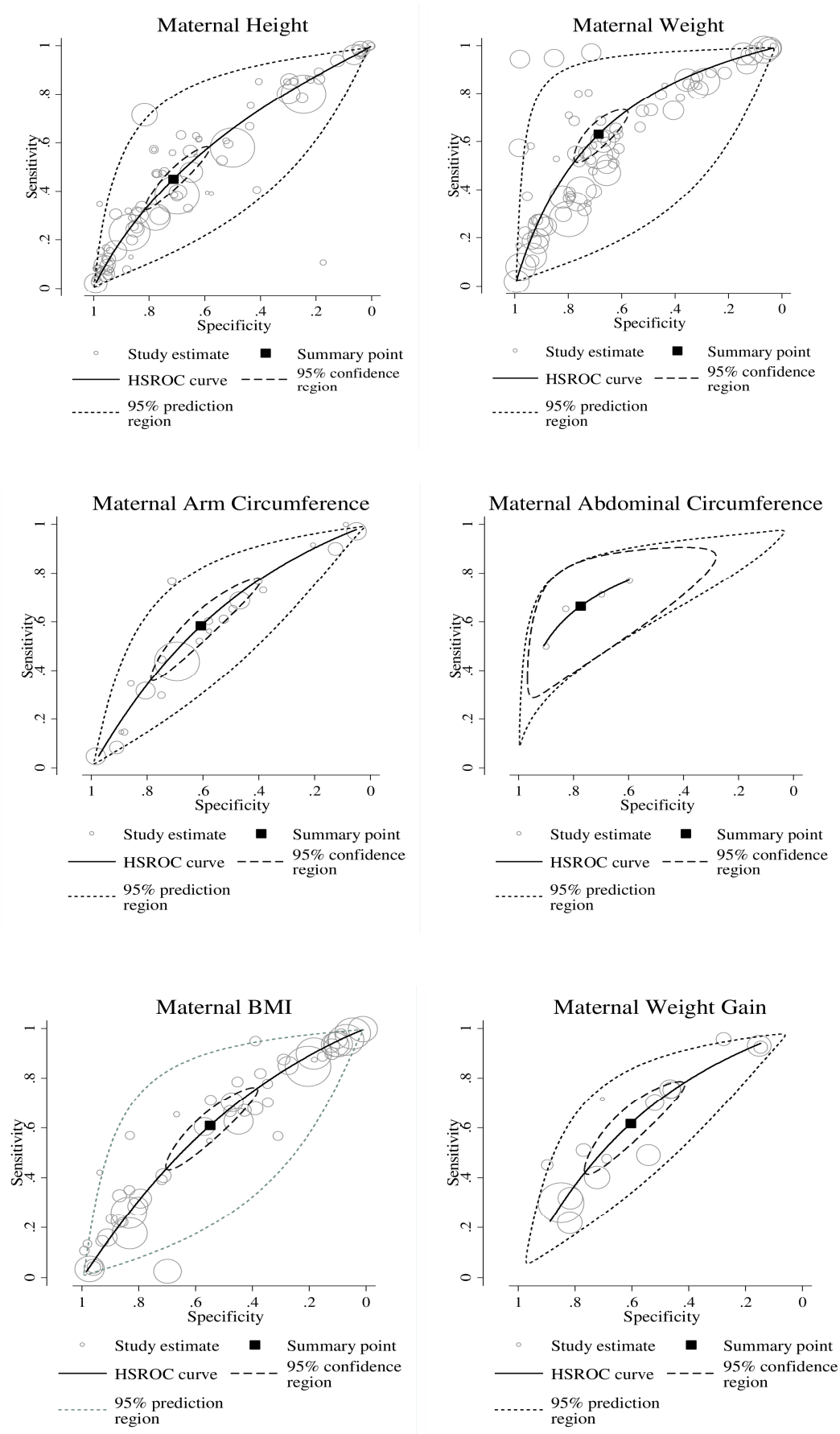

Figure 3. Hierarchical summary receiver operating characteristic curves provided no evidence of suitability of maternal anthropometric measurements for predicting low birth weight. Low sensitivity and specificity were shown for maternal height, weight, arm circumference, body mass index (BMI), and weight gain during pregnancy based on relatively large numbers of studies (n = 85, 80, 23, 51, and 16, respectively). Slightly higher but still insufficient sensitivity and specificity were shown for maternal abdominal circumference, which was not generalizable because there was only one data source.

CI=0.63-0.87), and DOR (7, 95% CI=4-11) and larger 95% confidence and prediction contours. The above findings were also true in the results of subgroup analysis (Table 3).

Eliminating potential confounders, i.e., changing study regions or responses to QUADAS items, could not reduce heterogeneity to $I^2 \leq 50\%$. Depending on the response to the 4th QUADAS item (i.e., short period from index test to reference test) and between Oceania vs other regions, the meta-regression showed statistically significant differences only in pooled specificity for weight (p=0.03) and in pooled sensitivity and specificity for weight gain (p=0.03 and 0.02, respectively), respectively. The differences in pooled sensitivity and specificity were not statistically significant according to any other feasible meta-regression analysis. There was also publication bias for weight and BMI (p=0.05 and 0.04, respectively), but no publication bias for any of the other maternal anthropometric measurements (p=0.18-0.39) except abdominal circumference for which assessment was not feasible due to the extraction of the data from only one article. The Youden index on the HSROC curve for each maternal anthropometric measurement could not provide a proposed cut-off point because the radius of curvature of the HSROC curve was too large and there were too few studies around the point likely to correspond to the index (Figure 3).

DISCUSSION

There was no solid evidence to suggest the suitability of maternal anthropometric measurements for predicting low birth weight. Rather, maternal height, weight, arm circumference, BMI, and weight gain showed poor low birth weight predictive capability during pregnancy because of low pooled sensitivity and specificity and DORs (Table 3 and Figure 3). The relatively good study quality may indicate the minimization of bias (Table 1 and Figure 2), and there are actually no threats to three major sources of bias-internal validity (i.e., studies that were designed with sufficient rigour to minimize bias). The findings for height, weight, arm circumference, BMI, and weight gain among the total population are also generalizable due to the large numbers of participants (n=24,197-137,420) extracted from the 16-85 studies of 71 articles (Tables 2 and 3 and Figure 1) - external validity (i.e., generalizability to or across different populations and settings). Internal and external validity supported the soundness of the conclusion. In addition, similar levels of pooled estimates were actually shown among the different study regions, between QUADAS ≥10 vs <10, and between post- vs pre-delivery measurements within the available data (Table 3). The publication bias detected for weight and BMI implied the probability of a poorer outcome in reality than that estimated in this meta-analysis. Abdominal circumference was shown to have slightly higher but still insufficient levels of pooled sensitivity and specificity and DOR (Table 3 and Figure 3). With regard to abdominal circumference, the smaller number of participants (n=604) in one article (only four studies) may not ensure the generalizability of the findings even among the total population.

The three major sources of bias in diagnostic meta-analyses[23,24] may not seriously affect the results. Nearly all of the included studies (n=252) used the same reference tests (Figure 2), and most of the included studies (n=202) were cohort studies (Table 2). A relatively small number of the included studies (n=69) involved prospective data collection. Changing the controls of these three major sources of bias did not reach statistical significance with regard to the differences in pooled sensitivity (p=0.13-0.92) or specificity (p=0.26-0.99), based on all obtainable p values in meta-regression analysis. Within the available data, changing the QUADAS score from ≥10 to <10 also did not reach statistical significance with regard to the differences in pooled sensitivity (p=0.42-0.84) or specificity (p=0.20-0.73) or cause the categories for diagnostic accuracy as a function of AUC[20] to differ, despite negligible alteration, if any, of the categories for informational value as a function of LRs[21] for weight (Table 3).

According to meta-regression analysis, one of the QUADAS items (i.e., the period between index and reference tests) would be a confounder for weight (see Results). This is reasonable because maternal weight may vary before delivery. However, comparison of the responses to the QUADAS item was not made between "yes" vs "no" but between "yes" vs "unclear" (Figure 2). In addition, statistically significant differences in sensitivity and specificity for weight gain (p=0.03 and 0.02, respectively) between Oceania (n=1) vs other regions (n=15) would be due mainly to a threshold effect rather than a region difference because of high sensitivity, i.e., 0.93 (95% CI=0.75-1.00) vs 0.58 (95% CI=0.45-0.72) and low specificity, i.e., 0.15 (95% CI=0.00-0.38) vs 0.64 (95% CI=0.52-0.75) in Oceania relative to other regions but almost the same DORs, i.e., 2 (95% CI=1-3) vs 2 (95% CI=2-3) between Oceania and other regions.

Newborn chest or arm circumference at birth may exceed maternal anthropometric measurements in predicting low birth weight in developing countries,[2] because of its high diagnostic accuracy ($0.9 \leq AUC \leq 1.0$)[20] and moderate informational value (positive and negative LRs=5-10 and 0.1-0.2, respectively).[5,21] Further, there were statistically significant differences (p<0.01) in DOR between newborn chest or arm circumference vs any maternal anthropometric measurement subjected to meta-analysis.

A significant strength of the present meta-analysis lies in its methodology, i.e., formulating review questions, defining inclusion and exclusion criteria, locating and selecting studies, assessing study quality, extracting data, and analyzing, presenting, and interpreting results.[28] With the exception of the absence of contact with the authors to obtain raw data or the lack of multiple observers to search the literature and assess study quality, the methodology closely or generally agrees with the proposed guidelines.[29-31] Moreover, the statistical analysis was distinguished by: 1) bivariate meta-analysis retaining the trade-off relationship between specificity and sensitivity instead of separate meta-analyses and allowing for between-study heterogeneity,[16] and 2) Deeks' funnel plot asymmetry test, which exceeds conventional tests in assessment of publication bias in diagnostic meta-analysis.[25]

Another strength of this meta-analysis was the inclusion of 259 studies and 309,419 women by employing various search engines and databases as well as checking

the references and PubMed-related citations of the potentially eligible articles as well as six systematic reviews.[6, 9-13] True positive, false positive, false negative, and true negative results were extracted wherever possible from the studies not clearly presenting two-by-two tables or aiming to evaluate diagnostic accuracy. Interestingly, the DORs of all the commonly evaluated maternal anthropometric measurements, except abdominal circumference, were consistent between this meta-analysis (1-4) and that performed previously by the WHO (1.2-2.5).[6] Sufficient sample sizes and good study quality support the generalizable conclusions for height, weight, arm circumference, BMI, and weight gain among the total population.

Another strength of this meta-analysis was the lack of serious influence of bias on the estimates. Although the studies included were not limited based on the QUADAS scores, overall their quality was relatively good (see Results, Table 1, and Figure 2). As mentioned above, the differences in pooled selectivity and specificity did not reach statistical significance depending on the three major sources of bias in diagnostic meta-analysis, and two of the three sources (i.e., the same reference test irrespective of the results of the index test and cohort rather than case-control study) were nearly always or mostly controlled. There was no publication bias for height, arm circumference, or weight gain. Publication bias for weight or BMI did not alter the interpretation of the estimates. Importantly, there were no statistically significant differences in pooled sensitivity (p=0.42-0.84) or in pooled specificity (p=0.20-0.73) between QUADAS ≥10 vs <10 within the available data.

One weakness of this meta-analysis was the application of pooled estimates from various populations to subgroups, such as males or females, singletons or non-singletons, and pre-term or full term infants. Further studies exclusively enrolling each subgroup are warranted to determine the subgroup-specific pooled estimates. Further, many of the studies considered here were not performed in participants' homes lacking weighing scales and healthcare facilities lacking foetal ultrasound equipment. However, meta-regression analysis of each maternal anthropometric measurement, except abdominal circumference in which meta-regression was not feasible as the data were extracted from only one article, did not show a statistically significant differences in pooled sensitivity (p=0.16-0.93) or pooled specificity (p=0.40-0.87) depending on the response to the first QUADAS item assessing the representativeness of participants.

Another weakness was the inclusion of BMI measured before as well as after delivery, although those taken only before delivery may be the most ideal indicators. For weight, the data could be separately pooled at the pre-delivery or post-delivery stage. Pooled sensitivity and specificity for weight did not show statistically significant differences between pre-delivery vs post-delivery (p=0.44 and 0.08, respectively). However, whether BMI was measured before or after delivery was not sufficiently clearly described for the data to be subjected to subgroup and meta-regression analyses. On the other hand, it is unreasonable that weight or BMI taken before delivery would yield better diagnostic performance than a post-delivery measurement for predicting low birth weight.

The final weakness of this meta-analysis was the unavailability of 16 articles (Figure 1) and the selection and review of studies by only a single person. However, the number of articles to be finally included among unavailable articles was estimated to be zero or one by multiplying the number of unavailable articles by the proportion of articles reporting finally included studies relative to the fully retrieved articles. For height, arm circumference, and weight gain, there was no evidence of publication bias in the results, although heterogeneity may affect funnel plot asymmetry.[25,32] For weight or BMI, publication bias did not seriously affect the interpretation of the estimates.

In summary, maternal height, weight, arm circumference, BMI, and weight gain do not show high sensitivity or specificity for predicting low birth weight. The findings for abdominal circumference could not be generalized because of the small sample size and limited population. Policymakers and clinical or public health practitioners are not encouraged to use maternal anthropometric measurements for predicting low birth weight.

ACKNOWLEDGEMENTS

The author would like to acknowledge the personnel in the Medical Library of the Japan Medical Association (Tokyo, Japan), who contributed to article retrieval, particularly by ordering the articles from other Japanese libraries, the U.S. National Library of Medicine, and the British Medical Library, as necessary.

AUTHOR DISCLOSURES

The author declares no potential conflicts of interest with respect to the authorship and/or publication of this article. The author received no financial support for the research and/or authorship of this article.

REFERENCES

1. United Nations Children's Fund and World Health Organization. II. Causes and consequences of low birthweight. In: Wardlaw T, Blanc A, Zupan J, Åhnman E, editors. Low Birthweight: Country, Regional and Global Estimates. New York: UNICEF Editorial and Publication Section; 2004. pp. 2-3.
2. Das S, Ghosh M, Mitra S, Chatterjee R, Bhattacharyya S, De AK, Chakraborty S. Developing nomogram to estimate birth weight from head circumference and mid-upper arm circumference. J Trop Pediatr. 2012;58:307-10. doi: 10.1093/tropej/fmr080.
3. Elizabeth NL, Christopher OG, Patrick K. Determining anthropometric surrogate measures for identifying low birth weight babies in Uganda: a hospital-based cross sectional study. BMC Pediar. 2013;13:54. doi: 10.1186/1471-243113-54.
4. Organisation for Economic Co-operation and Development. The DAC List of ODA 2011 Recipients. 2011/01/01 [cited 2014/5/23]; Available from: http://www.oecd.org/dac/stats/DAC%20List%20 used%20for%202011%20flows.pdf.
5. Goto E. Meta-analysis: identification of low birthweight by other anthropometric measurements at birth in developing countries. J Epidemiol. 2011;21:354-62. doi: 10.2188/jea.JE20100182.
6. World Health Organization. Maternal anthropometry and pregnancy outcome. A WHO collaborative study. Bull World Health Organ. 1995;73(suppl):1-98.
7. Bisai S, Mahalanabis D, Sen A, Bose K, Datta N. Maternal

early second trimester pregnancy weight in relation to birth outcome among Bengalee Hindus of Kolkata, India. Ann Hum Biol. 2007;34:91-101. doi: 10.1080/03014460601080728.
8. Zerfus A. Methodological issues in the measurement of maternal arm circumference. In: Krasovec K, Anderson MA, editors. Maternal nutrition and pregnancy outcomes: anthropometric assessment. Washington DC: Pan American Health Organization; 1991. pp. 131-58.
9. Han Z, Mulla S, Beyene J, McDonald SG; Knowledge Synthesis Group. Maternal underweight and the risk of preterm birth and low birth weight: a systematic review and meta-analysis. Int J Epidemiol. 2011;40:65-101. doi: 10.1093/ije/dyq195.
10. Kramer MS. Determinants of low birth weight: methodological assessment and meta-analysis. Bull World Health Organ. 1987;65:663-737.
11. McDonald SD, Han Z, Mulla S, Beyene J; Knowledge Synthesis Group. Overweight and obesity in mothers and risk of preterm birth and low birth weight infants: systematic review and meta-analysis. BMJ. 2010;341:c3428. doi: 10.1136/bmj.c3428.
12. Viswanathan M, Siega-Riz AM, Moos MK, Deierlein A, Mumford S, Knaack J, Thieda P, Lux LJ, Lohr KN. Outcomes of maternal weight gain. Evid Rep Technol Assess (Full Rep). 2008;168:1-223.
13. Siega-Riz AM, Viswanathan M, Moos MK, Deierlein A, Mumford S, Knaack J, Thieda P, Lux LJ, Lohr KN. A systematic review of outcomes of maternal weight gain according to the Institute of Medicine recommendations: birthweight, fetal growth, and postpartum weight retention. Am J Obstet Gynecol. 2009;201:339.e1-14. doi: 10.1016/j.ajog.2009.07.002.
14. Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. J Clin Epidemiol. 2005;58:982-90. doi: 10.1016/j.jclinepi.2005.02.022.
15. Arends LR, Hamza TH, van Houwelingen JC, Hijienbrok-Kal MH, Hunink MG, Stinen T. Bivariate random effects meta-analysis of ROC curves. Med Decis Making. 2008;28:621-38. doi: 10.1177/0272989X08319957.
16. Harbord RM, Whiting P. Metandi: meta-analysis of diagnostic accuracy using hierarchical logistic regression. In: Sterne JAC, editor. Meta-analysis in Stata: an updated collection from the Stata Journal. Texas: Stata Press; 2009. pp.181-99.
17. Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kelijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. BMC Med Res Methodol. 2003;3:25.
18. Falck–Ytter Y, Motschall E. New search filter for diagnostic studies: Ovid and Pubmed versions not the same. BMJ. 2004;328:1040. doi: 10.1136/bmj.38068.557998.EE.
19. Harbord RM. Metandi -- meta-analysis of diagnostic accuracy. 2013/02/19 [cited 2014/5/23]; Available from: http://fmwww.bc.edu/repec/bocode/m/metandi.html.
20. Swets JA. Measuring the accuracy of diagnostic systems. Science. 1988;240:1285-93. doi: 10.1126/science.3287615.
21. Jaeschke R, Guyatt GH, Sackett DL. Users' guides to medical literature. III. How to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients? The Evidence-Based Medicine Working Group. JAMA. 1994;271:703-7. doi: 10.1001/jama.271.9.703.
22. Higgins JP, Thompson SG, Deeks JJ, Altman DJ. Measuring inconsistency in meta-analyses. BMJ. 2003;327:557-60. doi: 10.1136/bmj.327.7414.557.
23. Ljimer JG, Mol BM, Heisterkap S, Bonsel GJ, Prins MH, van der Meulen JH, Bossuyt PM. Empirical evidence of design-related bias in studies of diagnostic tests. JAMA. 1999;282:1061-6. doi: 10.1001/jama.282.11.1061.
24. Rutjes AW, Reitsma JB, Di Nisio M, Smidt N, van Rijn JC, Bossuyt PM. Evidence of bias and variation in diagnostic accuracy studies. CMAJ. 2006;174:469-76. doi: 10.1503/cmaj.050090.
25. Deeks JJ, Macaskill P, Irwig L. The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic accuracy was assessed. J Clin Epidemiol. 2005;58:882-93. doi: 10.1016/j.jclinepi.2005.01.016.
26. Akobeng AK. Understanding diagnostic tests 3: receiver operating characteristic curves. Acta Paediatr. 2007;96:644-7. doi: 10.1111/j.1651-2227.2006.00178.x.
27. Chumnijarakij T, Nuchprayoon T, Chitinand S, Onthuman Y, Quamkul N, Dusitsin N et al. Maternal risk factors for low birth weight newborn in Thailand. J Med Assoc Thai. 1992;75:445-52.
28. Janjua NZ, Delzell E, Larson RR, Meleth S, Kristensen S, Kabagambe E, Sathiakumar N. Determinants of low birth weight in urban Pakistan. Public Health Nutr. 2009;12:789-98. doi: 10.1017/S1368980008002942.
29. Egger M, Smith GD, Altman DG. Systematic reviews in healthcare: meta-analysis in context. 2nd ed. London: BMJ; 2001.
30. Centre for Reviews and Dissemination. 2.2 Diagnostic tests. In: Systematic reviews: CRD's guidance for undertaking reviews in health care. York: Centre for Reviews and Dissemination; 2009. 2009/01/01 [cited 2014/5/23]; Available from: http://www.york.ac.uk/inst/crd/index_guidance.htm.
31. Diagnostic Test Accuracy Working Group. Handbook for DTA reviews. The Cochrane Collaboration; 2013. 2013/12/13 [cited 2014/5/24]; Available from: http://srdta.cochrane.org/handbook-dta- reviews.
32. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ. 1997;315:629-34. doi: 10.1136/bmj.315.7109.629.

Original Article

Diagnostic value of maternal anthropometric measurements for predicting low birth weight in developing countries: a meta-analysis

Eita Goto MD, MM, MPH

Nagoya Medical Science Research Institute, Meitou-ku, Nagoya, Japan

发展中国家孕妇人体测量指标预测低出生体重的诊断价值：一个荟萃分析

目的：发展中国家在家出生的婴儿往往没有进行称重。因此，尤其是在出生前检测低出生体重的简单、可靠和廉价的方法是很有用的。本研究为评估产妇人体测量指标预测婴儿低出生体重的诊断价值。**方法**：采用二元诊断荟萃分析，以构建多层次综合的受试者工作特征曲线。所有英语报道的看上去健康的孕妇并提供了构建 2×2 表（即真阳性、假阳性、假阴性和真阴性值）必要资料的研究均纳入本荟萃分析。在 PubMed 等 10 个数据库中进行检索确定这些研究。**结果**：检索到了包括来自非洲、亚洲、欧洲、拉丁美洲、中东和大洋洲的 309,419 名妇女及其新生儿的大量研究，其提供的可推广的研究结果包括母亲的身高、体重、上臂围、体质指数和怀孕期间体重增加（n 分别为 85、80、23、51 和 16）。然而， 0.46（95% CI=0.35-0.56）到 0.63 （95% CI=0.54-0.71）的灵敏度， 0.55（95% CI=0.42-0.67）到 0.71（95% CI=0.61-0.80）的特异度和 2（95% CI=2-2）到 4（95% CI=3-5）的诊断比值比对初步筛查是不够高的。由于样本有限（1 篇文章），腹围数据无法推广。**结论**：产妇人体测量指标不适合预测低出生体重。

关键词：人体测量学、低出生体重、荟萃分析、孕妇、灵敏度和特异度