

RESEARCH

Open Access



# Influence of different methods for calculating gestational age at birth on prematurity and small for gestational age proportions: a systematic review with meta-analysis

Gabriela Luiza Nogueira Vitral<sup>1,2\*</sup>, Roberta Maia de Castro Romanelli<sup>1</sup>, Tiago Alves Leonel<sup>1</sup>, Juliano de Souza Gaspar<sup>1</sup>, Regina Amélia Lopes Pessoa de Aguiar<sup>1</sup> and Zilma Silveira Nogueira Reis<sup>1</sup>

## Abstract

**Background** Recognizing premature newborns and small-for-gestational-age (SGA) is essential for providing care and supporting public policies. This systematic review aims to identify the influence of the last menstrual period (LMP) compared to ultrasonography (USG) before 24 weeks of gestation references on prematurity and SGA proportions at birth.

**Methods** Systematic review with meta-analysis followed the recommendations of the PRISMA Statement. PubMed, BVS, LILACS, Scopus-Elsevier, Embase-Elsevier, and Web-of-Science were searched (10–30-2022). The research question was: (P) newborns, (E) USG for estimating GA, (C) LMP for estimating GA, and (O) prematurity and SGA rates for both methods. Independent reviewers screened the articles and extracted the absolute number of preterm and SGA infants, reference standards, design, countries, and bias. Prematurity was birth before 37 weeks of gestation, and SGA was the birth weight below the p10 on the growth curve. The quality of the studies was assessed using the New-Castle-Ottawa Scale. The difference between proportions estimated the size effect in a meta-analysis of prevalence.

**Results** Among the 642 articles, 20 were included for data extraction and synthesis. The prematurity proportions ranged from 1.8 to 33.6% by USG and varied from 3.4 to 16.5% by the LMP. The pooled risk difference of prematurity proportions revealed an overestimation of the preterm birth of 2% in favor of LMP, with low certainty: 0.02 (95%CI: 0.01 to 0.03);  $I^2$  97%). Subgroup analysis of USG biometry (eight articles) showed homogeneity for a null risk difference between prematurity proportions when crown-rump length was the reference: 0.00 (95%CI: -0.001 to 0.000;  $I^2$ : 0%); for biparietal diameter, risk difference was 0.00 (95%CI: -0.001 to 0.000;  $I^2$ : 41%). Only one report showed the SGA proportions of 32% by the USG and 38% by the LMP.

**Conclusions** LMP-based GA, compared to a USG reference, has little or no effect on prematurity proportions considering the high heterogeneity among studies. Few data (one study) remained unclear the influence of such references on SGA proportions. Results reinforced the importance of qualified GA to mitigate the impact on perinatal statistics.

\*Correspondence:

Gabriela Luiza Nogueira Vitral  
gabrielavitral@gmail.com

Full list of author information is available at the end of the article



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

**Trial registration** Registration number PROSPERO: CRD42020184646.

**Keywords** Gestational age, Prematurity, Ultrasound, Last menstrual period

## Background

A qualified gestational age (GA) is essential to support healthcare decisions and to guarantee reliable perinatal health indicators for planning public policies [1]. The difference between the date of birth and the last menstrual period (LMP) is a useful method for GA calculation. However, the evaluation of GA by ultrasonography (USG) in early pregnancy with the crown-rump length (CRL) assessment is currently the most accurate method for dating pregnancy [2]. Clinical dating by LMP is fairly easy information to obtain GA in a birth scenario, however not always accurate being subject to memory bias, irregular menstrual cycles, and breastfeeding, among others [3]. Even easy to access, birth weight, as a single data, is not enough to discriminate preterm from term or to identify SGA newborns, although it is a risk marker for newborns [4]. The birth weight, gender, and GA allow the identification of the small for GA (SGA) newborn, according to weight below the 10th percentile for the standard expected in the growth curve [5].

Prematurity and nutritional classification at birth depend on the GA, which is essential to evaluate neonatal risks since they are associated with the chance of adverse neonatal outcomes [6]. Prematurity is the leading cause of death in children under five years old and is responsible for one million neonatal deaths annually [7, 8]. Recognizing the premature newborn, SGA, or both conditions at birth is important for the care provided at birth, as it helps timely interventions, indicating actions for greater effectiveness in care [9, 10]. In addition, the absence of a reliable gestational chronology negatively impacts the correct use of fetal and neonatal growth curves and leads to inaccuracy in epidemiological information associated with birth conditions [11, 12]. Affecting perinatal statistics, the scarcity of healthcare funding is still one determinant of the low access to prenatal USG in low-income countries (LMIC). This contributes to poor prenatal care coverage and makes estimating GA reliability even more difficult [1, 13]. The evaluation of GA by USG in early pregnancy with the crown-rump length (CRL) assessment is currently the most accurate method for redating pregnancy, but it requires high-cost equipment, early prenatal care, and a specialized health professional [2].

It is still unknown how much LMP and USG dating affect the rates of prematurity and SGA rates, based on studies with a high level of scientific evidence. Each method of gestational calculation can interfere differently with these rates. Previous publications point to the

possibility that the LMP overestimates the rate of premature and post-term births, whereas the more advanced obstetric USG performed in GA could underestimate such dating [14, 15]. In this context of uncertainties, this systematic review aims to identify the influence between LMP compared with USG before 24 weeks of gestational references on prematurity and SGA rates at birth.

## Methods

The research protocol followed the recommendations of the PRISMA Statement [16] and was registered in the International Prospective Registry of Systematic Reviews under PROSPERO number: CRD42020184646. This systematic review included all publications available in bibliographic bases: PubMed (MEDLINE), Scopus-Elsevier, Embase-Elsevier, BVS/LILACS, and Web of Science, until the date of October 2022. The research question considered "PECO" structuring: (P) newborns, (E) ultrasound for estimating GA, (C) LMP for estimating GA and (O) prematurity and SGA rates for both methods. The complete search strategy, adopting specific descriptors linked to Boolean operators, was "("gestational age" OR "pregnancy dating") AND ("last menstrual period" OR "menstrual date") AND (ultrasound OR ultrasonography OR "diagnostic imaging" OR ultrasonic) AND ((premature OR Preterm OR "immaturity at birth" OR prematurity) OR (small for gestational age" OR SGA OR "low birth weight" OR "intrauterine growth restriction" OR "small birth size"))". The search strategy was applied for descriptors present in the title, abstract, and keywords.

## Study eligibility criteria

The inclusion criteria were: [1] to have the GA at birth calculated by LMP and obstetric USG up to 24 weeks of gestation; [2] to have information about the data source for calculating the GA; [3] to have specific information about the standard fetal growth curve used to diagnose SGA. Primary studies, cohorts, epidemiological analyses, and database studies were also considered. There were no restrictions on language and date of publication. The exclusion criteria were: [1] GA at birth estimated by combining the two methods, [2] study with subgroups of newborns presenting specific diseases or conditions, such as carriers of congenital anomalies or growth anomalies, [3] population or sample composed only of premature or only SGA.

The preterm birth proportions was defined as births before 37 weeks of gestation divided by the total number

of births. SGA was birth weight below the 10th percentile for gestational age and gender based on the growth chart.

### Study appraisal and synthesis methods

We used the State of the Art software through Systematic Review (StArt) [17] to import articles and to support identifying duplicates, exclusions, and inclusions. Independent reviewers screened the title and abstract and performed the full reading with extracted the absolute number of preterm infants and SGA, reference standards, design, and countries. This research had two pairs of independent reviewers for each step, and a third reviewer was the judge in case of disagreements. According to the registered protocol, the absolute and relative values of the number of preterm infants and SGA were extracted, in addition to the characteristics of the primary studies.

### Evaluation of the quality of studies

Two independent reviewers assessed the methodological quality of the studies using the Newcastle–Ottawa Scale (NOS) [18] adjusted for the context of this review, detailed in Additional file 1. In cohort studies, 10 stars were possible, four for sample selection, two for comparability between the two GA estimation techniques, and four for outcomes. For cross-sectional studies, nine stars were possible, in which the outcomes part was scored with a maximum of three stars, valuing the appropriate and clearly described statistical treatment. So, lower-quality articles obtained fewer stars.

### Statistical analysis

The difference between two proportions using the risk difference [19] compared preterms when using LMP-based GA and prenatal USG-based GA. When the information was found, the same approach was used for the newborn SGA proportions. The null difference between prematurity proportions was adjusted in the center on the forest-plot graphs, with 95% confidence intervals (95%CI), considering two decimal places, according to Revman default. In this way, a result of 0.00 does not necessarily means zero difference, but the third decimal place onwards might have been hidden. Values below zero corresponded to a greater number of newborns being classified as preterm in favor of USG-based GA, while values above zero indicated greater proportions of preterm birth classification in favor of LMP-based GA. The random-effects model was adopted to mitigate high heterogeneity. The heterogeneity among studies was calculated using  $\chi^2$  and  $I^2$  for inconsistencies among proportions.

### Subgroup and meta-regression analysis

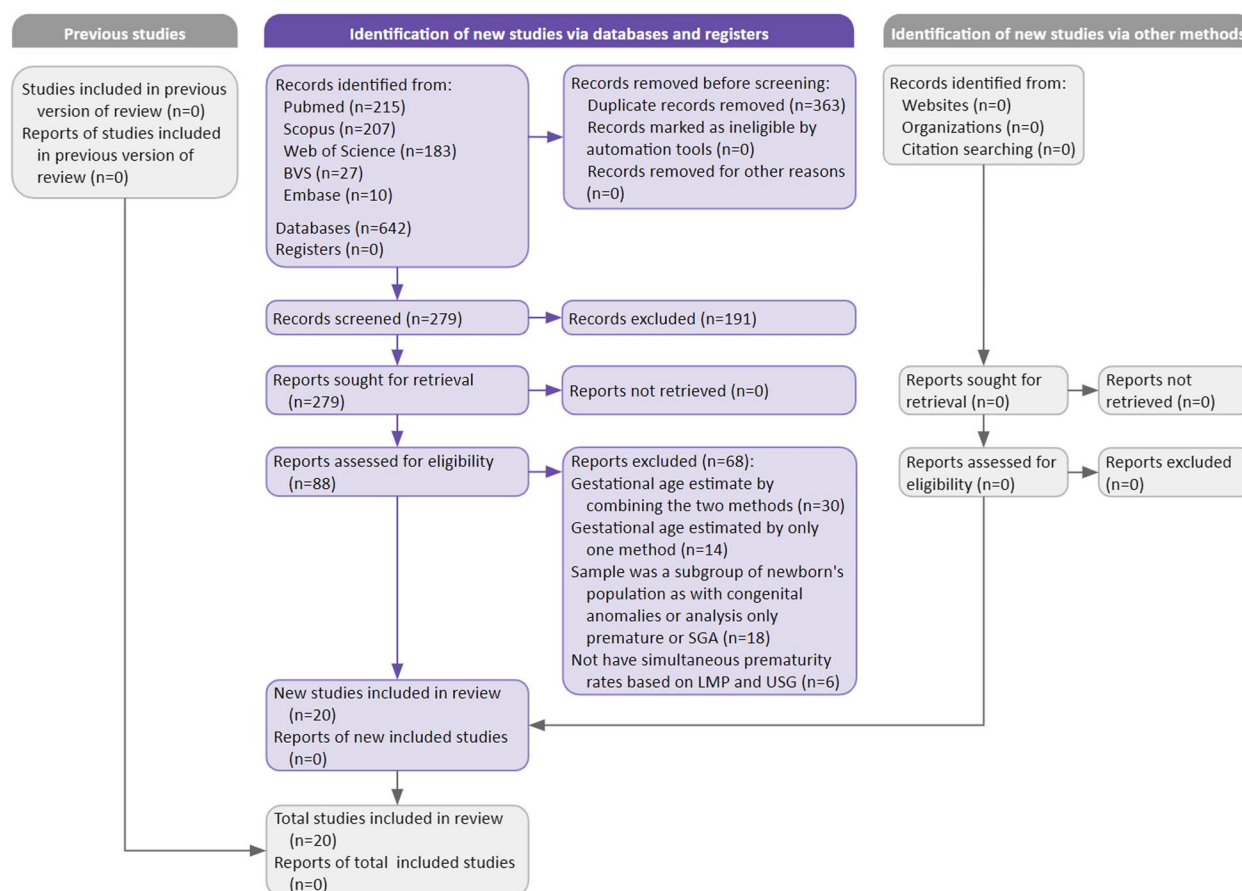
Subgroups were analyzed by study design: cohort or cross-sectional, income countries' economies: LMIC or high-income countries (HIC), and antenatal USG measurements: crown rump length (CRL) or biparietal diameter (BPD) biometrics. The Review Manager software (RevMan 5.4.1) was used for the meta-analysis.

### Results

The selection procedure of articles is shown in Fig. 1. In total, 642 articles were found, 215 PubMed, 207 Scopus, 183 Web of science, 27 BVS/Lilacs, and 10 Embase. Finally, 20 articles met the criteria for data extraction and synthesis according to the selection process depicted. Only one [15] among the 20 articles evaluated the SGA proportions estimate considering the 2 methods (LMP and USG).

We detailed the general characterization of the articles in Table 1. Among the 20 articles included in the meta-analysis, five presented a cross-sectional study design, and 15 were cohort studies. The year of publication ranged from 1995 to 2022, and the sample ranged from 171 to 165,908 newborns. Regarding the income countries of study location, ten were carried out in LMIC, such as Bangladesh, Colombia, Guatemala, Brazil, India, and Zambia, and ten in HIC, such as the USA, England, Denmark, and the Netherlands. Regarding the target population, eleven were carried out with mothers from the general population, two in rural areas, and seven were associated with other screening programs conducted widely in the studied population, such as the study of the Alpha-Fetoprotein Screening Program (XAPF) [20]. One report comparing pregnancy outcomes between women living with HIV and HIV-negative had data extraction only from the control group [21].

Table 2 provides information on the data sources for estimating GA at birth. The moment of the pregnancy when assessment with the obstetric USG occurred and the source of this data. It covered the information source of the LMP, the proportions of prematurity, and SGA newborns by the two methods of estimate analyzed. Fourteen articles reported the USG report in the medical record, and the USG was performed by the researchers in six. Information about LMP was collected through maternal interviews in ten articles, and ten were searched in medical records. The prematurity proportion by USG ranged from 3.4 to 16.5% and from 1.8 to 33.6% by LMP. Only one study reported comparisons between the SGA proportions, which was 32% (95%CI: 25–39) by USG and 38% LMP (95%CI: 31–46) [15].



**Fig. 1** Study selection process. SGA: Small for gestational age, LMP: Date of last menstrual period, USG: Ultrasonography

### Quality analysis of included articles

Table 3 shows the evaluation of the quality of the articles and, according to the NOS Scale. Regarding the cohort studies, three of them reached 9/10 stars [21, 24, 27]. The other studies are cross-sectional, with a maximum score of 7/9 stars in only two articles [32, 37]. The quality of the results refers to the independent or blind assessment of the prematurity proportions by the two references, the clarity in obtaining the date of birth, and the connection to the data of the beginning of pregnancy. In addition, it relates to the number of follow-up losses in the cohort and cross-sectional studies. The bias in the comparability between the proportions obtained by USG in relation to that obtained by LMP was based on the control of the reliability of the LMP. For this, we considered the resources used to certify this information, such as regular cycles, absence of abortion, and close birth influences on the female cycles, and statistical analysis. In database studies, efforts to qualify information on LMP and USG were better scored. Six studies presented a high risk of bias [25, 28, 30, 31, 33, 35].

Among the 15 cohort studies, the ones with the highest quality scores were those that stood out for their exceptional control in the prospective collection of LMP data, with control of preconception cycles [21, 24] and those that used prospective records or a large representation of the population of pregnant women [22, 27]. Among the five cross-sectional studies, the criterion with the highest risk of bias was the Outcome-item, either due to excessive data loss or a specific statistical approach to compare the prematurity proportions obtained by the two benchmarks. Part of the studies mainly focused on the analysis by grouping birth weight [20, 29] or prematurity subgroups [23, 25, 32].

### Risk difference between two proportions

Twenty-two proportions of prematurity were extracted from the 20 articles by each method, Fig. 2. The difference between preterm birth proportions was combined, resulting in 0.02 (95%CI: 0.01 to 0.03) in favor of LMP. It means that LMP-based GA overestimated the preterm birth proportions by 2%. However, the  $I^2$  value was 97%,  $p < 0.001$ , indicating high heterogeneity among



**Table 1** General characterization of the articles eligible for the systematic review

Author	Year	Study design	N	Study location	Period	Target population
AMANHI Study group [22]	2022	Cohort	9,974	South-Asia and sub-Saharan Africa	Jul 2012 – Sept 2016	AMANHI participants group
Dietz, P. M., et al. [20]	2007	Cross-sectional	165,908	California, USA	2002	XAPF study participants
Gardosi, J.; Francis, A. [23]	2000	Transversal	21,069	Birmingham, England	1988—1995	General population
Gernand, A. D., et al. [24]	2016	Cohort	353	Bangladesh	Fev 2009—Mar 2010	Rural area
Gonzalez, L. G., et al. [25]	2015	Cross-sectional	344	Manizales, Colombia	Sept 2012	General population
Henriksen, T. B., et al. [26]	1995	Cohort	3,606	Denmark	1989—1991	General population
Hoffman, C. S., et al. [27]	2008	Cohort	1,867	North Carolina and Texas, USA	2000—2004	RFTS Study Participants
Malaba, T. R., et al. [21]	2021	Cohort	2,507	Cape Town, South Africa	Apr 2015 – Oct 2016	General Population
Medeiros, M. N. L., et al. [28]	2015	Cohort	2,847	Ribeirão Preto and São Luis, Brazil	Mar 2010—Out 2011	General Population
Mongelli, M.; Gardosi, J. [29]	1997	Cross-sectional	34,249	Nottingham, England	-	General Population
Naslund Thagaard, I., et al. [30]	2016	Cohort	8,551	Copenhagen, Denmark	2006—2012	General Population
Neufeld, L. M., et al. [31]	2006	Cohort	171	Rural Guatemala	Aug 1996—Jun 1999	Rural Guatemala
Nguyen, T. H., et al. [32]	1999	Cohort	17,450	Copenhagen, Denmark	Jan 1986—Dec 1996	General Population
Pereira, A.P., et al. [33]	2013	Cohort	1,483	Rio de Janeiro, Brazil	Dez 2007—Nov 2008	General Population
Price, J. T., et al. [14]	2019	Cohort	692	Lusaka, Zambia	Aug 2015—Aug 2017	ZAPPS Study Participants
Reuss, M. L., et al. [34]	1995	Cohort	367	Pennsylvania and New York, USA	Jan 1987—Jun 1989	General Population
Savitz, D. A., et al. [35]	2002	Cohort	3,655	North Carolina, USA	Out 1995—Mai 2001	PIN Study Participants
van Oppenraaij, R. H. F., et al. [36]	2015	Cross-sectional	24,665	Rotterdam, Netherlands	Jan 2000—Dec 2009	General Population
Vijayram, R. et al. [37]	2021	Cohort	1,721	India	Mai 2015 –Nov 2017	GARBH-Ini study participants (rural and semi urban)
Weinstein, J. R., et al. [15]	2018	Cohort	188	Guatemala	Mar 2013—Fev 2015	SPHB study participants

N number of participants in the study, XAPF Alpha-Fetoprotein Screening Program, RFTS Right from the Start, ZAPPS Zambia Prematurity Prevention Study, PIN Pregnancy Infection and Nutrition, SPHB Safe Pregnancy Healthy Baby, GARBH-Ini Interdisciplinary Group for Advanced Research on Birth outcomes—India Initiative

the studies. In five studies, confidence intervals did not cross the null effect. These findings were robust since a sensitivity analysis removing the duplicity of proportions of prematurity [30, 37] showed a risk difference of 0.02 (95%CI: 0.01 to 0.03),  $I^2$  98%,  $p < 0.001$ , Additional file 2.

Comparisons between SGA proportions were measured in one study. There was a non-significant effect of LMP or USG methods on the SGA proportion at birth, Fig. 3 since the risk difference was 0.06 (95%CI: -0.04 to 0.16), crossing the null value.

In the subgroup analysis by study design Fig. 4, the pooled risk difference was 0.02 (95%CI: 0.01 to 0.03),  $I^2$  97%,  $p < 0.001$ , with LMP-based GA overestimation preterm birth by 2%. Across the 15 cohorts, however, risk difference showed no effect between LMP or USG references on the preterm birth proportions, with substantial heterogeneity: 0.03 (95%CI: 0.00 to 0.05). Likewise, in the cross-sectional study subgroup,  $I^2$  94%. These findings indicated that the study design may have unaffected the risk difference of prematurity since the total group and subgroups had high values of heterogeneity.

In the subgroup analysis by the country income of study Fig. 5, the risk difference of prematurity was 0.02 (95%CI: 0.01 to 0.03),  $I^2$  97%,  $p < 0.001$  in the studies conducted in LMIC. The findings indicated considerable heterogeneity among the studies, as well as in the subgroup of studies conducted in HIC in which  $I^2$  87%,  $p < 0.001$ . The risk difference was not significant, 0.04 (95%CI: -0.00 to 0.09),  $I^2$  97%, despite the higher prevalence of prematurity by LMP in countries of LMIC.

The USG captured CRL in five studies and BPD in four studies. In the subgroup analysis by antenatal USG measure Fig. 6, the risk difference for prematurity proportions was null calculated by CRL and null calculated by BPD when USG was used. Besides, we highlighted a fall in the heterogeneity among the studies for both methods of measure, considering  $I^2$  0% ( $p = 0.50$ ) in the CRL measure and  $I^2$  41% ( $p = 0.17$ ) in the BDP measure. It revealed the importance of segment of body measure by USG to explain the high heterogeneity among studies partially.

Thus, considering 20 reports, there was a low (2% in favor of LMP), however, with a high uncertainty risk

**Table 2** Characterization of the studied methods LMP and US of the systematic review

Author	Time of USG assessment	Source of USG data	Source of LMP data	LMP	USG	Prematurity rates		SGA rates	
						Difference	LMP	USG	Difference
AMANHI [22]	8 to 19 weeks	Researcher exam	Maternal interview	21.6%*	9.8%*	11.8%	-	-	-
Dietz, P.M. et al. [20]	< 20 weeks	Medical record	Medical record	8.7%	7.9%	0.8%	-	-	-
Gardosi, J. and Frandis, A. [23]	14 to 21 weeks	Medical record	Medical record	6.3%	7%	-0.7%	-	-	-
Gernand, A. D. et al. [24]	< 15 weeks	Researcher exam	Maternal interview	10.8%	8.7%	2.1%	-	-	-
Gonzalez, L. G. et al. [25]	56 to 97 days	Medical record	Medical record	5.2%	4.3%	0.9%	-	-	-
Henriksen, T. B. et al. [26]	< 20 weeks	Medical record	Medical record	4.0%	3.4%	0.6%	-	-	-
Hoffman, C. S., et al. [27]	4 to 17 weeks	Researcher exam	Maternal interview	10%	8.9%	1.1%	-	-	-
Malaba, T. R. et al. [21]	< 24 weeks	Researcher exam	Maternal interview	33.6%	14.0%	19.6%	-	-	-
Medeiros, M. N. L., et al. [28]	< 20 weeks	Medical record	Maternal interview	9.7%	8.5%	1.2%	-	-	-
Mongelli, M.; Gardosi, J. [29]	< 24 weeks	Medical record	Medical record	6.2%	6.6%	-0.4%	-	-	-
Naslund Thagaard, I., et al. (a) [30]	CRL between 45 and 84 mm	Medical record	Medical record	4.1%	4.5%	-0.4%	-	-	-
Naslund Thagaard, I., et al. (b) [30]	BPD first trimester	Medical record	-	4.1%	4.6%	-0.5%	-	-	-
Neufeld, L. M., et al. [31]	15 to 24 weeks	Researcher exam	Maternal interview	1.8%	4.1%	-2.3%	-	-	-
Nguyen, T. H., et al. [32]	12 to 22 weeks	Medical record	Medical record	4.0%	4.5%	-0.5%	-	-	-
Pereira, A. P., et al. [33]	7 to 20 weeks	Medical record	Maternal interview	17.7%*	13.2%	4.5%	-	-	-
Price, J. T., et al. [14]	< 20 weeks	Medical record	Medical record	20.2%	11.4%	8.8%	-	-	-

**Table 2** (continued)

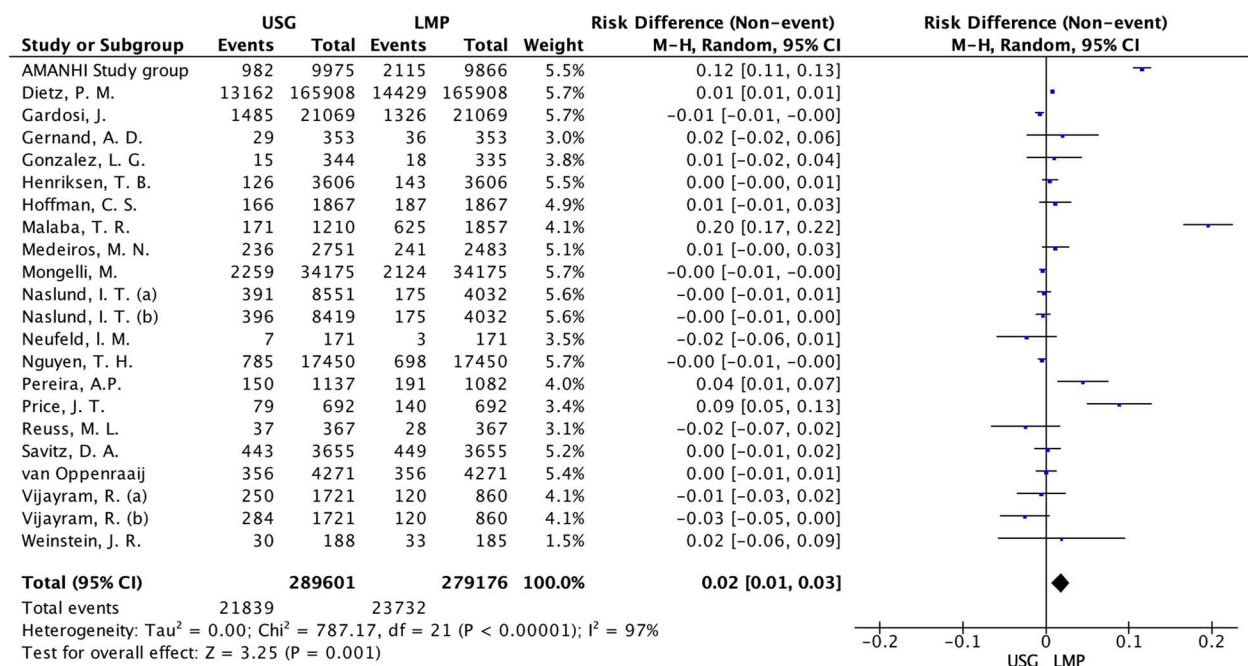
Author	Time of USG assessment	Source of USG data	Source of LMP data	Prematurity rates			SGA rates	
				LMP	USG	Difference	LMP	USG
Reuss, M. L., et al.[34]	< 18 weeks	Medical record	Maternal interview	7.6%	10.1%	-2.5%	-	-
Savitz, D. A., et al.[35]	< 18 weeks	Medical record	Medical record	12.2%	12.1%	0.1%	-	-
van Oppenraaij, R. H. F., et al.[36]	8 weeks 0 day to 11 weeks 6 days	Medical record	Medical record	8.3%	8.3%	0%	-	-
Vijayram, R. et al. (a) [37]	< 20 weeks Hadlock	Researcher exam	Maternal interview	14.0%	14.5%	-0.5%	-	-
Vijayram, R. et al. (b) [37]	< 20 week Robinson-Fleming	Researcher exam	-	14.0%	16.5%	-2.5%	-	-
Weinstein, J. R., et al. [15]	< 24 weeks	Medical record	Maternal interview	17.8%	16.0%	1.8%	38%	32%
								6%

LMP Last menstrual period, SGA Small for gestational age, USG Ultrasonography, Naslund (a) CRL measurement, Naslund (b) BPD measurement in the first trimester, Vijayram (a) US Hadlock, Vijayram (b) US Robinson-Fleming formula

\* Average of countries

**Table 3** Quality evaluation of the studies selected in the systematic review by the Newcastle–Ottawa Scale

Study	Type	Selection	Comparability	Outcome	Total
AMANHI [22]	Prospective cohort	★★★	★	★★★	7/10
Gernand, A. D.; et al. [24]	Prospective cohort	★★★★	★	★★★★	9/10
Henriksen, T. B. et al. [26]	Prospective cohort	★★★	★	★★	6/10
Hoffman, C. S., et al. [27]	Prospective cohort	★★★★	★★	★★★	9/10
Malaba, T. R., et al. [21]	Prospective cohort	★★★★	★	★★★★	9/10
Medeiros, M. N. L., et al. [28]	Prospective cohort	★★	-	★★	4/10
Naslund T., et al. [30]	Prospective cohort	★	★	★★	4/10
Neufeld, L. M. et al. [31]	Prospective cohort	★★★★	-	-	4/10
Pereira, A.P., et al. [33]	Prospective cohort	★★★	-	★	4/10
Price, J. T., et al. [14]	Prospective cohort	★★★★	★	★	6/10
Reuss, M. L., et al. [34]	Prospective cohort	★★★	★	★★★	7/10
Savitz, D. A., et al. [35]	Prospective cohort	★★★★	-	-	4/10
Van Oppenraaij, R. H. F., et al. [36]	Retrospective cohort	★★	★★	★	5/10
Weinstein, J. R., et al. [15]	Prospective cohort	★★★	★	★★	6/10
Vijayram, R. et al. [37]	Prospective cohort	★★★	★	★★★	7/10
Dietz, P. M., et al. [20]	Cross-sectional	★★★	★	★	5/9
Gardosi, J. & Francis, A. [23]	Cross-sectional	★★★★	★	★	6/9
Gonzalez, L. G., et al. [25]	Cross-sectional	★	-	★	2/9
Mongelli, M. & Gardosi, J. [29]	Cross-sectional	★★★★	★	★	6/9
Nguyen, T. H., et al. [32]	Cross-sectional	★★★	★★	★★	7/9

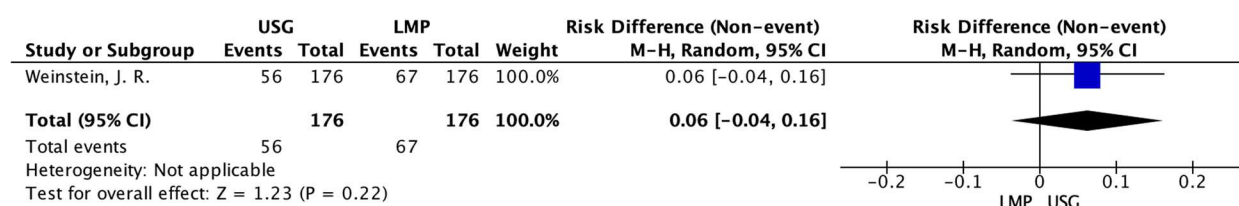
**Fig. 2** Forest plot of premature proportions by Last menstrual period and USG before 24 weeks. LMP: Last menstrual period; USG: Ultrasonography

difference in prematurity proportions between methods of reference. However, the CRL assessment in early pregnancy as a reference for GA resulted in similar prematurity proportions compared to the LMP reference.

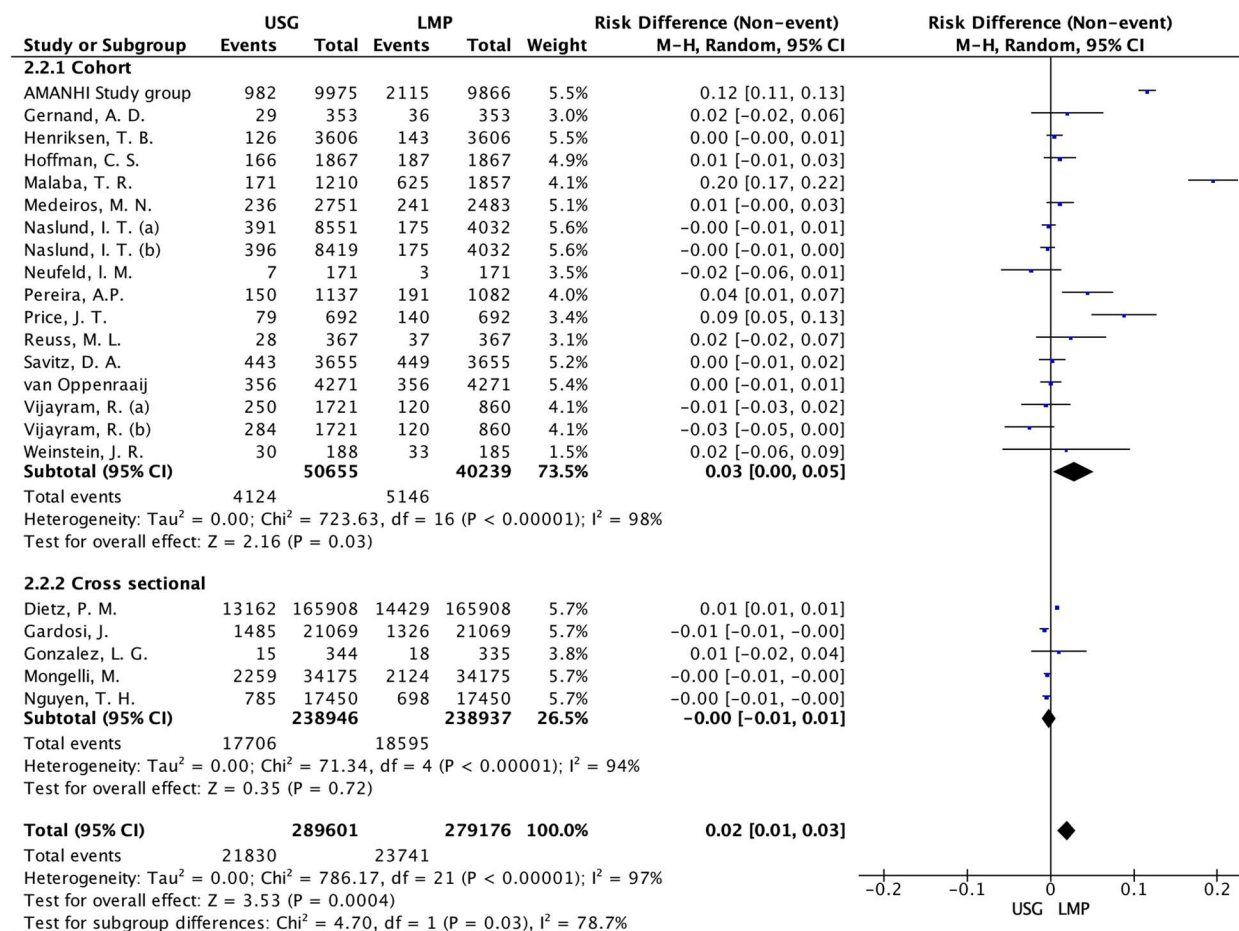
## Discussion

This review compared the proportions of prematurity and SGA between LMP and USG before 24 weeks as a reference for estimating GA. The occurrence of preterm





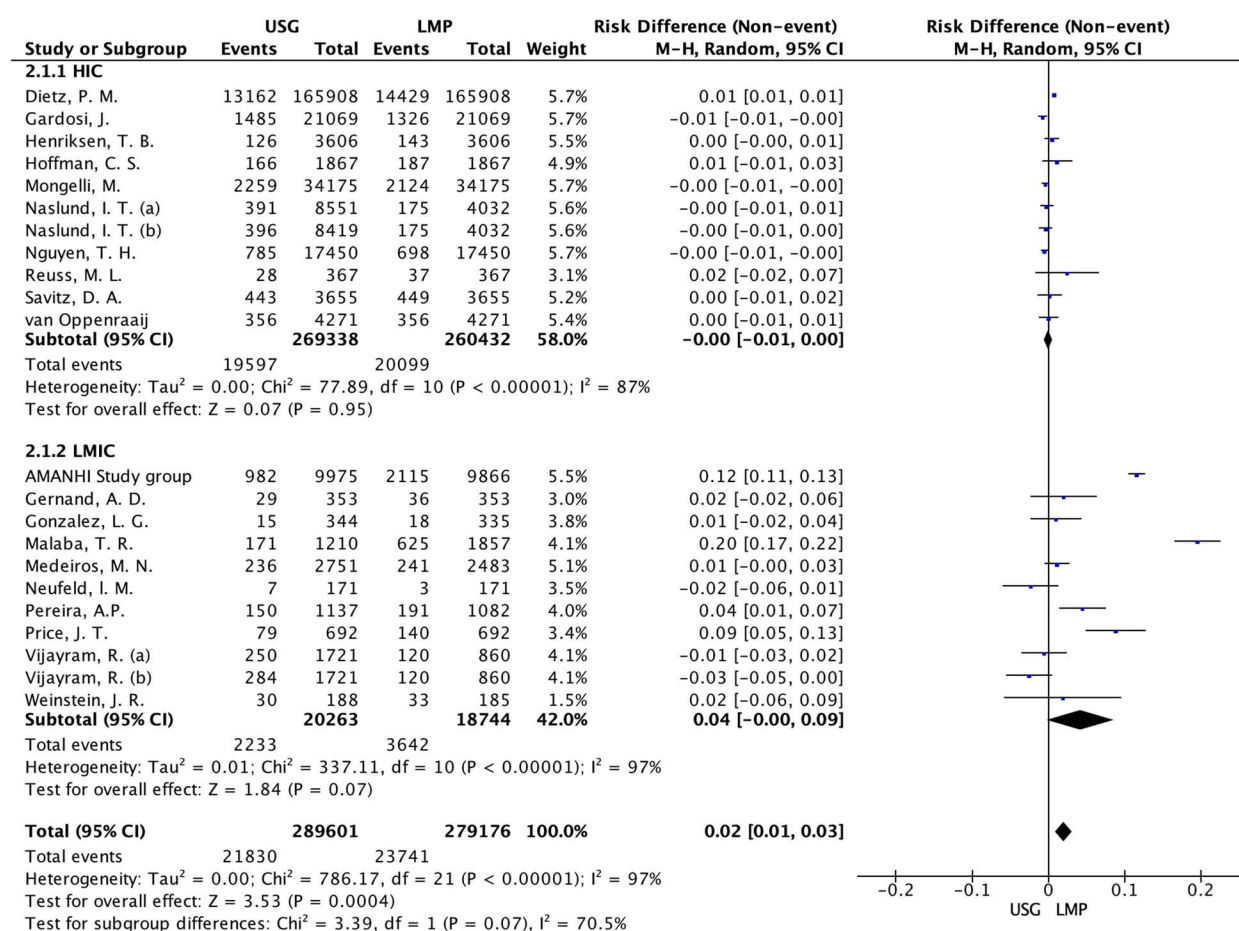
**Fig. 3** Forest plot of small-for-gestational-age proportions by last menstrual period and ultrasonography before 24 weeks. LMP: Last menstrual period; USG: Ultrasonography



**Fig. 4** Forest plot of subgroup by study design: Cohort or cross-sectional. LMP: Last menstrual period; USG: Ultrasonography. Naslund (a): CRL measurement; Naslund (b): BPD measurement in the first trimester. Vijayram (a): US Hadlock; Vijayram (b): US Robinson-Fleming formula

among infants varied significantly in the reports, with prematurity proportions ranging from 1.8% to 20.2% by LMP [14, 31]. It means different complexities of health assistance from varied countries and diverse ultrasonography approaches. In an attempt to reduce such differences, the meta-analysis considered the random effects model, sensitivity, and subgroup analysis. Five studies showed a significant risk difference for prematurity in favor of the GA calculated by the LMP [14, 20–22,

33], and no studies by the USG-based GA. In the all set of studies, a significant risk difference of 2% (1% to 3% 95%CI) was found between methods in favor of higher prematurity proportions when using LMP. Our interpretation regarding this outcome relies on the tendency for underestimation of GA based on LMP in preterm newborns and the overestimation in the term and post-term compared to GA based on USG. The overlapped GA distribution comparisons between LMP and USG references



**Fig. 5** Forest plot of subgroup by country of study location: Low-middle Income countries or high-income countries. LMP: Last menstrual period; USG: Ultrasonography; HIC: high-income country; LMIC: Low-middle income countries. Naslund (a): CRL measurement; Naslund (b): BPD measurement in the first trimester. Vijayram (a): US Hadlock; Vijayram (b): US Robinson-Fleming formula

to support this possibility were represented in previous reports [22, 38]. We speculate such differences unbalance the proportions in the direction of higher prematurity. Even with high heterogeneity and moderate risk of bias, we raised concern about the overestimation of preterm rates in birth scenarios where GA is mainly calculated based on LMP. Regarding the comparison of SGA proportions, the lack of studies kept unclear the influence of such references for GA calculation on the indicator.

### Investigation of subgroup analyses

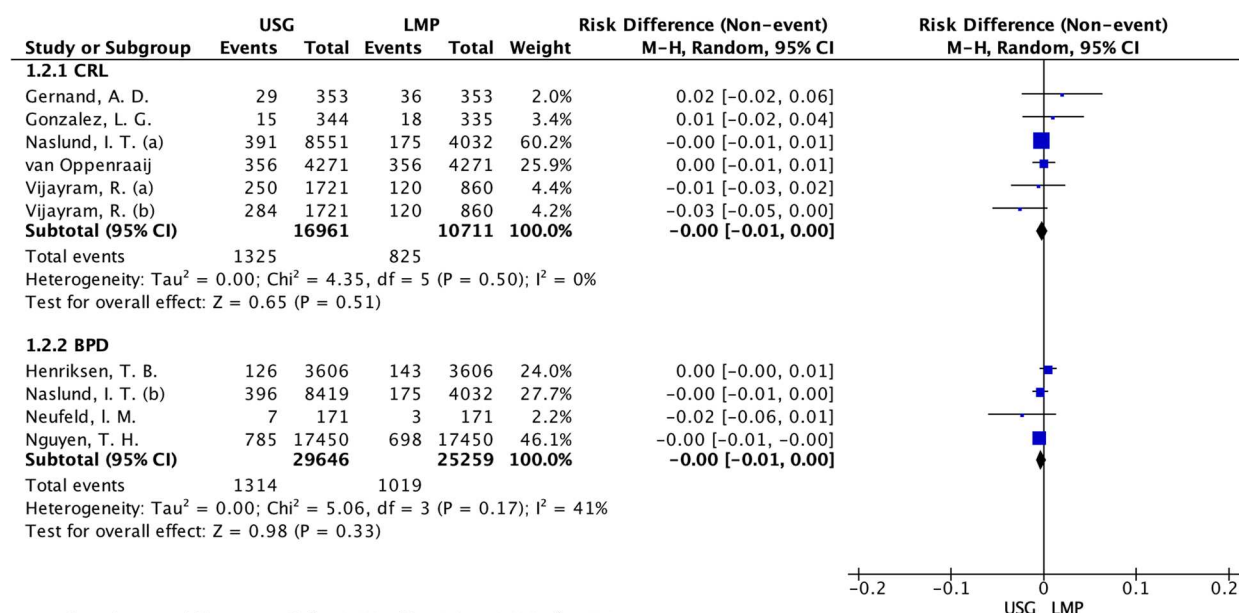
#### Income

The subgroup analysis supported that the difference between proportions of prematurity is practically null in the scenario of HIC since the pooled result had no evidence of prematurity in favor of any side. Early access to prenatal care and higher quality of assistance may be associated with this finding [39]. Studies in LMIC countries had risk differences between methods in favor of higher prematurity proportions when using LMP-based

GA. However, it presented a 95% CI reaching the null value. Among the ten studies in this group, four showed evidence of significant risk differences in favor of LMP [14, 21, 22, 33], and two studies found differences in favor of USG, but not significant [31, 37]. Most of the studies in this subgroup presented high proportions of prematurity as compared to those in the HIC group. One possible interpretation is that the qualification of clinical information about menstrual cycles was more valuable in studies of LMIC, while the early access to prenatal care, including USG, may be limited [6].

### Study design

Regarding the study's design, the group of articles with a cohort approach had a similar and high  $I^2$  value compared to the cross-sectional group of articles (98% vs. 94%, respectively). Thus, the study design did not explain the high heterogeneity among the studies. With cohort design, 15 articles gathering 50,655 newborns, a difference between prematurity proportions was in favor of



Test for subgroup differences:  $\chi^2 = 0.03$ ,  $df = 1$  ( $P = 0.86$ ),  $I^2 = 0\%$

**Fig. 6** Forest plot of subgroup by antenatal USG measure: Crown-rump length or biparietal diameter. LMP: Last menstrual period; USG: Ultrasonography; CRL: Crown-rump length; BPD: Biparietal diameter. Naslund (a): CRL measurement; Naslund (b): BPD measurement in the first trimester. Vijayram (a): US Hadlock; Vijayram (b): US Robinson-Fleming formula

LMP when compared to USG, with high uncertainty. Otherwise, in the cross-sectional study subgroup, five articles with 238,946 newborns, there was no evidence of a difference in favor of any side. Prospective cohorts of pregnant women occurred in scenarios with a high prematurity proportion in LMICs [14, 21, 22, 33], all articles in favor of LMP-based GA overestimating prematurity proportions. Also, attention is drawn to a large number of cases that Dietz et al. (2007) studied in a cross-sectional analysis of 165,908 women participating in the XAPF screening program in California, USA [20], also with a risk difference in favor of LMP-based. The other studies showed a 95% CI for the risk difference crossing the null effect, i.e., without evidence of impact on the proportions of prematurity.

#### Antenatal USG measure

Stratified analysis of USG biometry revealed high homogeneity among six-paired risk differences of five studies when CRL was the reference for GA estimate at birth [24, 25, 30, 36, 37]. Even though such pooled risk difference leaned to the USG-based GA side, there was no significant evidence that prematurity proportions based on CRL would differ from LMP reference. The CRL assessment in early pregnancy resulted in similar prematurity proportions compared to the LMP reference. The clinical application of this outcome is aligned with prior reports when the CRL is considered the most accurate method for GA redating [2]. Similarly, the risk difference

of prematurity when DPB is available <24 weeks was the reference for GA leaned to the USG side, however, with high uncertainty. Further comparative studies are still needed to confirm this trend. This last finding corroborates the USG before 24 weeks gestation as an acceptable standard the World Health Organization recommended to improve prenatal care [40]. For a proper interpretation, the inclusion criterion for USG <24 weeks may have influenced the high heterogeneity among the set of studies since the fall of heterogeneity was clear using USG biometry subgroup analysis. CRL measurement is the best parameter for calculating the GA [2]. However, we consider that using first-trimester USG with CRL measurement would have limited the inclusion of many studies, especially those carried out in LMIC.

#### Quality of the studies

The main bias domain in the systematic review was the procedure to access the reference used in the GA calculation. It is worth mentioning that collecting information about the menstrual cycle varied among the studies. This information was obtained prospectively in rural Bangladesh [24], Guatemala [31], and others [14, 21] through the woman's self-report or from a medical record and database, most of them without mention of procedures for data qualification. Accordingly, the discrepancy in the proportion of prematurity found in the same country is noteworthy: in the study by Weinstein et al., 17.8% and Neufeld et al., 1.8% based on the LMP [15, 31]. Even

though these studies analyze different birth scenarios, the figures are still far from Guatemala's prematurity rate, estimated at 10% [41]. The quality of the prematurity rate is directly linked to the accuracy of the GA estimate. Incompleteness, lost, and underreported data are more frequent limitations in LMICs [40]. Several factors are recognized for affecting the quality of data in this indicator, such as the collection protocols used by health professionals, early access to prenatal care and USG, and the value of this information for women, caregivers, and governments, among others [42].

### **Strength and limitations of the review**

The main contribution was to provide unprecedented comparison proportions of prematurity and SGA by two references recommended for calculating GA. However, the results found present limitations regarding the variability of research designs, different USG biometrics, and contexts of birth scenarios among the reports, which would possibly explain the heterogeneity observed in the set of all articles, at least in part. The considerable heterogeneity in the meta-analysis was a limiting factor in interpreting the results for clinical practice. Future controlled studies should mitigate methodological weakness in blinding the GA between methods. Another point to highlight was the duplication of prematurity proportions by different USG approaches, Hadlock vs. Robinson-Fleming; and CRL vs. DBP. However, the sensitivity analysis remained similar to the total 22 prematurity proportions, considering just one per study.

Despite the limitations, this review is helpful as a basis for studies involving estimates of GA and its relationship to the prematurity and SGA proportions. Transparent criteria in the study group selection, appropriate statistical treatment, and a clearly described and detailed methodology are essential for a study to confirm the differences among the rates of prematurity by the two reference methods. Another point was the lack of studies comparing the SGA proportions obtained by different reference calculations for GA.

The correct determination of GA can affect the results of pregnancy [43], improve decision-making in childbirth and neonatal care [44], optimizing health costs [45]. Investigating the influence of different references for prematurity and SGA rates is relevant for any delivery scenarios and public policies and research [13, 46, 47]. A valid prematurity rate is unknown in many places due to the lack of qualified data, especially in LMIC [47]. This systematic review reinforced the importance of early prenatal care with qualified LMP and USG access to adjust the due date of birth. The lack of studies kept unclear the influence of LMP and USG references on SGA proportions. Insights for future primary research are to compare

the rates paying attention to the fetal biometric measures for pregnancy dating and considering the diversity of high-cost healthcare technologies access.

### **Conclusions**

Meta-analysis showed that LMP compared to the USG before 24 weeks of gestational references for GA calculation, has little or no effect on prematurity proportions at birth, considering the high heterogeneity among studies. The CRL by USG assessment in early pregnancy resulted in similar prematurity proportions compared to the LMP reference. The lack of studies kept unclear the influence of such references on SGA proportions. Results corroborated the importance of qualified GA to mitigate the impact on perinatal statistics.

### **Abbreviations**

BPD	Biparietal diameter
CRL	Crown-rump length
GA	Gestational age
HIC	High-Income Countries
LMIC	Low-Middle Income Countries
LMP	Date of last menstrual period
NOS	New Castle-Ottawa scale
SGA	Small for gestational age
USG	Ultrasonography

### **Supplementary Information**

The online version contains supplementary material available at <https://doi.org/10.1186/s12884-023-05411-0>.

**Additional file 1.** NewCastle-Ottawa scale adjusted for the context of the review

**Additional file 2.** Forest plot of premature proportions by last menstrual period and USG before 24 weeks without duplicated proportions

### **Acknowledgements**

Fundação Oswaldo Cruz Rio de Janeiro, Brazil (grant VPPIS-002-FEX-20), and Grand Challenges Canada's programs, Government of Canada Toronto, Canada (grant number R-ST-POC-1807-13515). ZSNR is a researcher with a grant from the Conselho Nacional de Pesquisa (CNPq 305837/2021-4), Brazil.

### **Authors' contributions**

The authors declare that they have a substantial contribution to the conception and design of the study, the acquisition, analysis, and interpretation of data; All of them participated in drafting the work and revised it critically for important intellectual content; They approved the final draft version; They agree for all aspects of the study in ensuring that questions related to the accuracy or integrity of any part are appropriately investigated and resolved. The authors read and approved the final manuscript.

### **Funding**

This study is funded by Grand Challenges Exploration from the Bill & Melinda Gates Foundation (grant OPP1128907, contract). The funders played no role in the study design, data collection, analysis, decision to publish, or manuscript preparation.

### **Availability of data and materials**

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.



## Declarations

### Ethics approval and consent to participate

Not applicable.

### Consent for publication

Not applicable.

### Competing interests

Authors declare a patent deposit on behalf of the Universidade Federal de Minas Gerais and Fundação de Amparo a Pesquisa de Minas Gerais, Brazil, <http://www.fapemig.br/en/>. The inventors were GLNV, RMCR and ZSNR: BR1020200209825 (CTIT-PN1094).

### Author details

<sup>1</sup>Faculdade de Medicina, Universidade Federal de Minas Gerais, 30.130.100, Belo Horizonte, Avenida Professor Alfredo Balena, 190, Sala 601, Brazil. <sup>2</sup>Faculdade de Ciências Médicas de Minas Gerais, Alameda Ezequiel Dias, 275, Belo Horizonte 30130-110, Brazil.

Received: 3 March 2022 Accepted: 27 January 2023

Published online: 11 February 2023

## References

- Sharma LK, Bindal J, Shrivastava VA, Sharma M, Choorakuttil RM, Nirmalan PK. Discordant dating of pregnancy by LMP and ultrasound and its implications in perinatal statistics. *Indian J Radiol.* 2020;30(1):27–31.
- ACOG. Committee Opinion No 700 Methods for Estimating the Due Date. *Obstet Gynecol.* 2017;129(5):e150–4.
- Nguyen TH, Larsen T, Engholm G, Møller H. Increased adverse pregnancy outcomes with unreliable last menstruation. *Obstet Gynecol.* 2000;95(61):867–73.
- Damhuis SE, Ganzevoort W, Gordijn SJ. Abnormal Fetal Growth: Small for Gestational Age Fetal Growth Restriction Large for Gestational Age Definitions and Epidemiology. *Obstet Gynecol Clin.* 2021;48(2):267–79.
- Gordijn SJ, Beune IM, Thilaganathan B, Papageorgiou A, Baschat AA, Baker PN, et al. Consensus definition of fetal growth restriction: a Delphi procedure. *Ultrasound Obstet Gynecol.* 2016;48(3):333–9.
- Travers CP, Carlo WA. How to Save 1 Million Lives in a Year in Low- and Middle-Income Countries. *Neonatal.* 2017;111(4):431–6.
- Organization WH. Atlas of EHealth Country Profiles: The Use of EHealth in Support of Universal Health Coverage: Based on the Findings of the Third Global Survey on EHealth 2015. Volume 3. World Health Organization; 2016. 392 p.
- Darmstadt GL, Bhutta ZA, Cousens S, Adam T, Walker N, de Bernis L. Evidence-based cost-effective interventions: how many newborn babies can we save? *Lancet.* 2005;365(9463):977–88.
- Howson CP, Kinney MV, McDougall L, Lawn JE. the Born Too Soon Preterm Birth Action Group Born Too Soon Preterm birth matters. *Reprod Health.* 2013;10(1):1.
- Lees CC, Stampalija T, Baschat A, Costa da S F, Ferrazzi E, Figueras F, et al. ISUOG Practice Guidelines: diagnosis and management of small-for-gestational-age fetus and fetal growth restriction. *Ultrasound Obstet Gynecol Off J Int Soc Ultrasound Obstet Gynecol.* 2020;56(2):298–312.
- Villar J, Papageorgiou AT, Pang R, Ohuma EO, Ismail LC, Barros FC, et al. The likeness of fetal growth and newborn size across non-isolated populations in the INTERGROWTH-21st Project: the Fetal Growth Longitudinal Study and Newborn Cross-Sectional Study. *Lancet Diabetes Endocrinol.* 2014;2(10):781–92.
- Villar J, Ismail LC, Victora CG, Ohuma EO, Bertino E, Altman DG, et al. International standards for newborn weight, length, and head circumference by gestational age and sex: the Newborn Cross-Sectional Study of the INTERGROWTH-21st Project. *Lancet.* 2014;384(9946):857–68.
- Karl S, Suen CSNLW, Unger HW, Ome-Kaius M, Mola G, White L, et al. Pre-term or Not - An Evaluation of Estimates of Gestational Age in a Cohort of Women from Rural Papua New Guinea. *PLOS ONE.* 2015;10(5):e0124286.
- Price JT, Winston J, Vwalika B, Cole SR, Stoner MCD, Lubeya MK, et al. Quantifying bias between reported last menstrual period and ultrasonography estimates of gestational age in Lusaka. *Zambia Int J Gynecol Obstet.* 2019;144(1):9–15.
- Weinstein JR, Thompson LM, DíazArtiga A, Bryan JP, Arriaga WE, Omer SB, et al. Determining gestational age and preterm birth in rural Guatemala: A comparison of methods. *Plos One.* 2018;13(3):e0193666 (Ryckman KK, organizador).
- Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ.* 2021;372:71.
- Zamboni A, Thommazo AD, Hernandez ECM, Fabbri SCPP. StArt uma ferramenta computacional de apoio à revisão sistemática. In Proc.: Congresso Brasileiro de Software (CBSOFT'10), Salvador Brazil. 2010. p. 91–96.
- Wells GA, Shea B, O'Connell D, Peterson J, Welch V, Losos M, et al. The Newcastle–Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. *Ottawa Health Research Institute;* 2014.
- Barker TH, Migliavaca CB, Stein C, Colpani V, Falavigna M, Aromataris E, et al. Conducting proportional meta-analysis in different types of systematic reviews: a guide for synthesizers of evidence. *BMC Med Res Methodol.* 2021;21(1):189.
- Dietz PM, England LJ, Callaghan WM, Pearl M, Wier ML, Kharrazi M. A comparison of LMP-based and ultrasound-based estimates of gestational age using linked California livebirth and prenatal screening records. *Paediatr Perinat Epidemiol.* 2007;21(s2):62–71.
- Malaba TR, Mukonda E, Matjila M, Madlala HP, Myer L, Newell ML, et al. Pregnancy outcomes in women living with HIV and HIV-negative women in South Africa: Cohort analysis based on bias-corrected gestational age. *Paediatr Perinat Epidemiol.* 2021;36(4):525–35.
- AMANHI. Population-based rates, risk factors and consequences of preterm births in South-Asia and sub-Saharan Africa: A multi-country prospective cohort study. *J Glob Health.* 2022;12:04011.
- Gardosi J, Francis A. Early pregnancy predictors of preterm birth: the role of a prolonged menstruation-conception interval. *BJOG Int J Obstet Gynaecol.* 2000;107(2):228–37.
- Gernand AD, Paul RR, Ullah B, Taher MA, Witter FR, Wu L, et al. A home calendar and recall method of last menstrual period for estimating gestational age in rural Bangladesh: a validation study. *J Health Popul Nutr.* 2016;35(1):34.
- González LG, Serna MS, González VV, Arenas OAV. Comparación de la confiabilidad de la ecografía obstétrica del primer trimestre y la fecha de la última menstruación en la predicción de la fecha de parto vaginal espontáneo en pacientes admitidas en un centro de atención de I nivel en Manizales (Colombia) 2012–2013. *Arch Med.* 2015;15(1):107–13.
- Henriksen TB, Wilcox AJ, Hedegaard M, Secher NJ. Bias in Studies of Pre-term and Postterm Delivery Due to Ultrasound Assessment of Gestational Age. *Epidemiol.* 1995;6(5):533–7.
- Hoffman CS, Messer LC, Mendola P, Savitz DA, Herring AH, Hartmann KE. Comparison of gestational age at birth based on last menstrual period and ultrasound during the first trimester. *Paediatr Perinat Epidemiol.* 2008;22(6):587–96.
- Medeiros MNL, Cavalcante NCN, Mesquita FJA, Batista RLF, Simões VMF, Cavalli R de C, et al. Validity of pre and post-term birth rates based on the date of last menstrual period compared to early obstetric ultrasonography. *Cad Saude Publica.* 2015;31:885–90.
- Mongelli M, Gardosi J. Birth weight, prematurity and accuracy of gestational age. *Int J Gynecol Obstet.* 1997;56(3):251–6.
- NäslundThagaard I, Krebs L, Lausten-Thomsen U, Larsen SO, Holm JC, Christiansen M, et al. Dating of Pregnancy in First versus Second Trimester in Relation to Post-Term Birth Rate: A Cohort Study. *PLOS ONE.* 2016;11(1):e0147109.
- Neufeld LM, Haas JD, Grajeda R, Martorell R. Last menstrual period provides the best estimate of gestation length for women in rural Guatemala. *Paediatr Perinat Epidemiol.* 2006;20(4):290–8.
- Nguyen TH, Larsen T, Engholm G, Møller H. Evaluation of ultrasound-estimated date of delivery in 17 450 spontaneous singleton births: do we need to modify Naegele's rule? *Ultrasound Obstet Gynecol.* 1999;14(1):23–8.
- Pereira APE, Dias MAB, Bastos MH. da Gama N granado N, Leal M do C. Determining gestational age for public health care users in Brazil: comparison of methods and algorithm creation. *BMC Res Notes.* 2013;6(1):60.



34. Reuss ML, Hatch MC, Susser M. Early ultrasound dating of pregnancy: Selection and measurement biases. *J Clin Epidemiol*. 1995;48(5):667–74.
35. Savitz DA, Terry JW, Dole N, Thorp JM, Siega-Riz AM, Herring AH. Comparison of pregnancy dating by last menstrual period, ultrasound scanning, and their combination. *Am J Obstet Gynecol*. 2002;187(6):1660–6.
36. van Oppenraaij R, Eilers P, Willemsen S, van Dunné F, Exalto N, Steegers E. Determinants of number-specific recall error of last menstrual period: a retrospective cohort study. *BJOG Int J Obstet Gynaecol*. 2015;122(6):835–41.
37. Vijayram R, Damaraju N, Xavier A, Desiraju BK, Thiruvengadam R, Misra S, et al. Comparison of first trimester dating methods for gestational age estimation and their implication on preterm birth classification in a North Indian cohort. *BMC Pregnancy Childbirth*. 2021;21(1):343.
38. Rittenhouse KJ, Vwalika B, Keil A, Winston J, Stoner M, Price JT, et al. Improving preterm newborn identification in low-resource settings with machine learning. *Plos One*. 2019;14(2):e0198919.
39. Arsenaault C, Jordan K, Lee D, Dinsa G, Manzi F, Marchant T, et al. Equity in antenatal care quality: an analysis of 91 national household surveys. *Lancet Glob Health*. 2018;6(11):e1186–95.
40. Lee AC, Blencowe H, Lawn JE. Small babies, big numbers: global estimates of preterm birth. *Lancet Glob Health*. 2019;7(1):e2–3.
41. Save the Children Federation. Healthy Newborn Network. Guatemala Data [Internet]. 2022 [citado 17 de novembro de 2022]. Disponível em: <https://www.healthynewbornnetwork.org/country/guatemala/>
42. Scott K, Gupta S, Williams E, Arthur M, Somayajulu UV, Noguchi L. “I can guess the month . . . but beyond that, I can’t tell” an exploratory qualitative study of health care provider perspectives on gestational age estimation in Rajasthan, India. *BMC Pregnancy Childbirth*. 2020;20(1):529.
43. Thorsell M, Kaijser M, Almström H, Andolf E. Expected day of delivery from ultrasound dating versus last menstrual period—obstetric outcome when dates mismatch. *BJOG Int J Obstet Gynaecol*. 2008;115(5):585–9.
44. Reis ZSN, Gaspar JDS, Vital GLN, Abrantes VB, De-Souza IMF, Moreira MTS, et al. Quality of Pregnancy Dating and Obstetric Interventions During Labor: Retrospective Database Analysis. *JMIR Pediatr Parent*. 2020;3(1):e14109.
45. Kuo DZ, Lyle RE, Casey PH, Stille CJ. Care System Redesign for Preterm Children After Discharge From the NICU. *Pediatrics*. 2017;139(4):e20162969.
46. Lee AC, Panchal P, Folger L, Whelan H, Whelan R, Rosner B, et al. Diagnostic Accuracy of Neonatal Assessment for Gestational Age Determination: A Systematic Review. *Pediatrics*. 2017;140(6):e20171423.
47. Walani SR. Global burden of preterm birth. *Int J Gynecol Obstet*. 2020;150(1):31–3.

## Publisher’s Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

