



SYSTEMATIC REVIEW

Body composition in preterm infants: a systematic review on measurement methods

Dana F. J. Yumani¹✉, Dide de Jongh^{2,3}, Johannes C. F. Ket⁴, Harrie N. Lafeber^{1,5} and Mirjam M. van Weissenbruch¹

© The Author(s), under exclusive licence to the International Pediatric Research Foundation, Inc 2022

BACKGROUND: There are several methods to measure body composition in preterm infants. Yet, there is no agreement on which method should be preferred.

METHODS: PubMed, Embase.com, Wiley/Cochrane Library, and Google Scholar were searched for studies that reported on the predictive value or validity of body composition measurements in preterms, up to 6 months corrected age.

RESULTS: Nineteen out of 1884 identified studies were included. Predictive equations based on weight and length indices, body area circumferences, skinfold thickness, bioelectrical impedance, and ultrasound did not show agreement with body composition measured with air displacement plethysmography (ADP), dual-energy x-ray absorptiometry (DXA), magnetic resonance imaging (MRI), or isotope dilution. ADP agreed well with fat mass density measured by isotope dilution (bias -0.002 g/ml, limits of agreement ± 0.012 g/ml, $n = 14$). Fat mass percentage measured with ADP did not agree well with fat mass percentage measured by isotope dilution (limits of agreement up to $\pm 5.8\%$) and the bias between measurements was up to 2.2%. DXA, MRI, and isotope dilution were not compared to another reference method in preterms.

CONCLUSIONS: DXA, ADP, and isotope dilution methods are considered trustworthy validated techniques. Nevertheless, this review showed that these methods may not yield comparable results.

Pediatric Research (2023) 93:1120–1140; <https://doi.org/10.1038/s41390-022-02262-x>

IMPACT:

- Based on validation studies that were conducted in a limited number of study subjects, weight and length indices, body area circumferences, skinfold thickness, bioelectrical impedance, and ultrasound seem to be a poor representation of body composition in preterm infants.
- DXA, ADP, and isotope dilution methods are considered trustworthy and validated techniques. Nevertheless, these methods may not yield comparable results.

INTRODUCTION

It is becoming more evident that preterm birth increases the incidence of risk factors for cardiovascular disease in early as well as later life. Preterm birth is associated with a higher percentage of body fat, higher blood pressure, and an increased risk of dysglycemia from infancy into adulthood.^{1–6} Furthermore, preterm birth has been associated with an increased risk of ischemic heart disease in adulthood.⁷ Therefore, it is of utmost importance to identify proper screening tools to identify risk factors for cardiovascular disease at an early age and implement preventive measures accordingly.

As body fat percentage and fat mass index have been shown to positively correlate with the occurrence of metabolic syndrome components,^{8,9} monitoring body composition during early life could help to implement timely preventive measures. To gain more insight into which methods should be used to monitor body

composition in early life in preterm infants, this systematic review will assess validation studies in preterm infants from birth up to 6 months corrected age.

Several methods to assess body composition are available—ranging from inexpensive, bedside techniques, such as skinfold thickness (SFT) measurement, to expensive and bulky equipment, like air displacement plethysmography (ADP).^{10,11} Currently, ADP and dual energy x-ray absorptiometry (DXA) are frequently used in research settings and are considered as reliable methods.¹² Nevertheless, these methods are not widely implemented in clinical practice. Furthermore, there is no consensus on which method should preferentially be used to assess body composition in preterm infants. This systematic review aims to determine the validity of different methods used to measure body composition in preterm infants and to show whether validated methods yield comparable results.

¹Department of Pediatrics, VU University Medical Center, Amsterdam UMC, Amsterdam, The Netherlands. ²Faculty of Science, VU University, Amsterdam, The Netherlands.

³Faculty of Medical Ethics and Internal Medicine, Erasmus MC, Rotterdam, The Netherlands. ⁴Medical Library, Vrije Universiteit, Amsterdam, The Netherlands. ⁵Deceased: Harrie N. Lafeber. ✉email: d.yumani@amsterdamumc.nl

Received: 18 May 2021 Revised: 20 June 2022 Accepted: 24 July 2022

Published online: 22 August 2022

Table 1. Search strategies.

Search strategy for PubMed (29 September 2020)		
Search	Query	Results
#4	#3 NOT (animals[mh] NOT humans[mh])	337
#3	#1 AND #2	339
#2	"Body Composition"[Mesh] OR fatty tissue*[ti] OR adipos*[ti] OR body composition*[ti] OR fat mass*[ti] OR fat percentage*[ti] OR lean mass[ti] OR lean body mass[ti] OR fat free mass[ti] OR body fat[ti] OR ffm[ti] OR fat free body mass[ti] OR lean weight*[ti] OR fatty tissue*[ot] OR adipos*[ot] OR body composition*[ot] OR fat mass*[ot] OR fat percentage*[ot] OR lean mass[ot] OR lean body mass[ot] OR fat free mass[ot] OR body fat[ot] OR ffm[ot] OR fat free body mass[ot] OR lean weight*[ot]	100,890
#1	"Infant, Premature"[Mesh] OR "Premature Birth"[Mesh] OR "Intensive Care, Neonatal"[Mesh] OR "Intensive Care Units, Neonatal"[Mesh] OR nicu*[tiab]	86,314
Search strategy for Embase.com (29 September 2020)		
Search	Query	Results
#5	#4 NOT ("conference abstract"/it OR "conference paper"/it OR "editorial"/it OR "letter"/it OR "note"/it)	589
#4	#3 NOT ([animals]/lim NOT [humans]/lim)	803
#3	#1 AND #2	911
#2	"immature and premature labor"/de OR "premature labor"/exp OR "prematurity"/exp OR prematur*:ti,ab,kw OR preterm*:ti,ab,kw OR "newborn intensive care"/exp OR nicu*:ti,ab,kw	341,349
#1	"body composition"/mj/exp OR "adipose tissue"/mj/exp OR ((fatty NEAR/3 tissue*):ti,kw) OR adipos*:ti,kw OR "body composition*":ti,kw OR "fat mass*":ti,kw OR "fat percentage*":ti,kw OR "lean mass*":ti,kw OR "lean body mass*":ti,kw OR "fat free mass*":ti,kw OR "body fat*":ti,kw OR ffm:ti,kw OR "fat free body mass*":ti,kw OR "lean weight*":ti,kw	120,114
Search strategy for Wiley/Cochrane Library (29 September 2020)		
Search	Query	Results
#1	(prematu* OR preterm* OR (neonatal NEXT intensive NEXT care) OR nicu*):ab,ti,kw	29,187
#2	((fatty NEAR/3 tissue*) OR adipos* OR (body NEXT composition*) OR (fat NEXT mass*) OR (fat NEXT percentage*) OR (lean NEXT mass) OR (lean NEXT body NEXT mass) OR (fat NEXT free NEXT mass) OR (body NEXT fat) OR ffm OR (fat NEXT free NEXT body NEXT mass) OR (lean NEXT weight*)):ti,ab,kw	24,093
#3	#1 AND #2	280
Cochrane Reviews: 17; CENTRAL: 263.		
Google Scholar query		
	(~premature ~preterm "neonatal intensive care" nicu) AND ("fatty tissue" adipose "body composition" "fat mass" "fat percentage" "lean mass" "lean body mass" "fat free mass" "body fat" "fat free body mass" "lean weight")	+/-193,000
The first 949 hits were retrieved and screened using Publish or Perish ⁶⁵		

METHODOLOGY

This systematic review primarily assessed the accuracy of various methods used to measure or estimate body composition in infants born preterm.

A systematic literature review was conducted according to the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) statement (www.prisma-statement.org). The review was registered and the protocol was published on PROSPERO International prospective register of systematic reviews under ID CRD42018107821. Searches were performed in PubMed, Embase.com, and Wiley/Cochrane Library from inception (1809) up to 29 September 2020 by D.F.J.Y. and J.C.F.K. The search included keywords and free-text terms for "premature" and "body composition" or "adiposity" or "lean." Animal studies, conference papers and editorial letters were excluded. No language or publication date restrictions were applied. The full search strategies can be found in Table 1. In addition, the reference lists of relevant articles and Google Scholar were checked. Where needed, authors were contacted for clarification or additional information.

Study eligibility criteria

Studies were included if they reported on body composition measurement in infants born before 37 weeks of gestation. The body composition measurement had to take place between birth and 6 months corrected age. Studies needed to evaluate methods which measure or estimate fat (-free) mass in (kilo)grams or

percentage. Studies measuring or estimating total body water (TBW) were also included. See Table 2 for a description of the included methods.^{12,13} In addition to validation studies, randomized controlled trials, cohort studies, and epidemiologic studies were included if these studies reported the accuracy or predictive values of body composition measurements.

Definitions

A method was deemed validated if the method showed good statistical agreement to a reference method. Currently, there is no golden standard for the measurement of body composition. Hence, studies of all possible reference methods were accepted.

Good agreement was defined as a maximum allowed difference of 10% of the mean value of the body composition parameter in the study population. For example, if the mean of the fat free mass in a study population was 2000 g, then the bias ± limits of agreement had to be smaller than ±200 g.

For studies where agreement was not assessed, effect size of the different methods was determined by the *r*-squared value. A value below 0.5 was considered a poor predictive value, 0.5–0.7 as a moderate predictive value, and >0.7 was considered as good predictive value.¹⁴

Data extraction

Two reviewers (D.F.J.Y. and D.d.J.) separately screened the studies, initially based on title and abstract, followed by full-text review of the relevant studies.

Table 2. In vivo techniques for measuring body composition in preterm born infants.^{12,13}

Technique	Method	Application
Body proportionality calculations	Ponderal index (kg/m ³), Body mass index (BMI) (kg/m ²), Benn index (weight/length ^P), waist, mid-arm, and mid-thigh circumference, mid-arm/head circumference ratio	These indices are generally included in regression models to formulate predictive equations that estimate fat (free) mass (g) or fat (free) mass percentage
Skinfold thickness measurements	Skinfold thickness is measured with a caliper to estimate subcutaneous fat	The absolute value of the skinfold at a particular site or the sum of skinfolds measured at different sites are generally included in regression models to formulate predictive equations that estimate fat (free) mass (g) or fat (free) mass percentage
Isotope dilution	Isotope distribution is measured in blood or urine	Total body water is calculated from the isotope distribution. Fat free mass (g) is calculated from total body water
Bioelectrical impedance analysis/bioelectrical impedance spectroscopy	The flow of electrical currents through tissue is measured	Total body water is calculated from the resistance to the flow of electrical currents running to body tissue (Total body water = $\rho L^2/Z$; ρ = volume resistivity in Ω -cm, L = length in cm, Z = impedance in Ω)
Ultrasound	Measures muscle area	Estimates subcutaneous fat by subtracting the muscle cross-sectional area from the total cross-sectional area
Magnetic resonance imaging	Measures adipose tissue volume based on tissue-specific energy absorption	Measures fat mass (g)
Dual-energy x-ray absorptiometry	Measures fat mass and fat free mass based on tissue-specific x-ray attenuation	Measures fat (free) mass (g)
Air displacement plethysmography	Measures body volume based on air displacement in a test chamber	Calculates fat mass (g) and fat mass percentage through age- and sex-specific density coefficients

Data extraction was performed by these two reviewers. The data collected included the method and reference method, details on the study setting, methods, and results. In case of any discrepancies between the two reviewers, the two reviewers came to agreement through discussion.

Reviewers were not blinded for authors or journal details. Where needed, authors were contacted for clarification or additional information.

Risk of bias (quality) assessment

Two reviewers (D.F.J.Y. and D.d.J.) primarily assessed bias using the Critical Appraisal Skills Program (CASP) checklist.¹⁵ In case of any discrepancies between the two reviewers, the two reviewers discussed and asked the expert opinion of the two other reviewers until an agreement was made. The synthesis was based on the final decision made under agreement of all reviewers. The quality of individual studies was assessed with CASP checklists. In addition, the Oxford Centre for Evidence-based Medicine's Levels of Evidence was used to grade the level of evidence of each manuscript.¹⁶

Strategy for data synthesis

A narrative synthesis was primarily done by two researchers (D.F.J.Y. and M.M.v.W.) and was reviewed by D.d.J., J.C.F.K., and H.N.L. before finalization.

RESULTS

Out of the 1884 identified records, 48 full-text articles were assessed for eligibility and 19 were included in this synthesis (Fig. 1). Nine studies ($n = 1539$) reported about the predictive value or validity of body proportionality measures. Five studies ($n = 319$) investigated the validity of bioelectrical impedance analysis (BIA), three studies ($n = 90$) investigated the validity of SFT, two studies ($n = 24$) investigated the validity of ADP, one study ($n = 63$) investigated the predictive value of ultrasound, and one study ($n = 15$) investigated the validity of MRI. There were no human studies that reported about the validation of DXA and

isotope dilution studies in preterm infants (Table 3). Body composition measurements were performed at various postnatal ages, ranging from 24 h postpartum to 4 months corrected age.

Body proportionality measures

Table 3 shows our findings on the predictive value and validity of body proportionality measures. Weight and length indices had a moderate to good predictive value for fat-free mass (in grams). On the contrary, the predictive value of weight and length indices for fat mass was poor to moderate and fat mass percentage was poorly predicted by weight/length indices.^{17–21}

Larcade et al. and Simon et al. assessed predictive equations with clinical parameters, such as caloric and macronutrient intake, and z-scores for weight, length and head circumference.^{22,23} They found that fat-free mass (g) could be predicted by the amount of human milk feeding, respiratory support, antenatal corticosteroid use, growth parameters and sex. A newly modeled equation by Larcade et al. showed good agreement for fat free mass (g). However, Larcade and colleagues did not assess fat mass percentage, while Simon and colleagues could only explain 24% of the variance in fat mass percentage with their predictive model^{22,23} (see Table 3).

Daly-Wolfe and colleagues reported that mid-arm circumference had a moderate predictive value for fat mass percentage measured by ADP²⁴ (see Table 3). On the other hand, Koo et al. measured chest, abdomen, mid-thigh, and mid-arm circumference. After including sex, race, gestation, weight and length in a predictive equation for fat mass (g) and fat free mass (g), these body and limb circumference did not explain any additional part of the variance in fat (free) mass (g).¹⁷ Of note, these findings were based on the entire study population that included both term and preterm infants (respectively, $n = 68$ and $n = 52$) who were large, appropriate as well as small for gestational age. Nevertheless, Pereira-Da-Silva and colleagues investigated a group of exclusively preterm infants and reported mid-arm circumference to have a poor predictive value arm fat area (mm²) measured by MRI²⁵ (see Table 3).

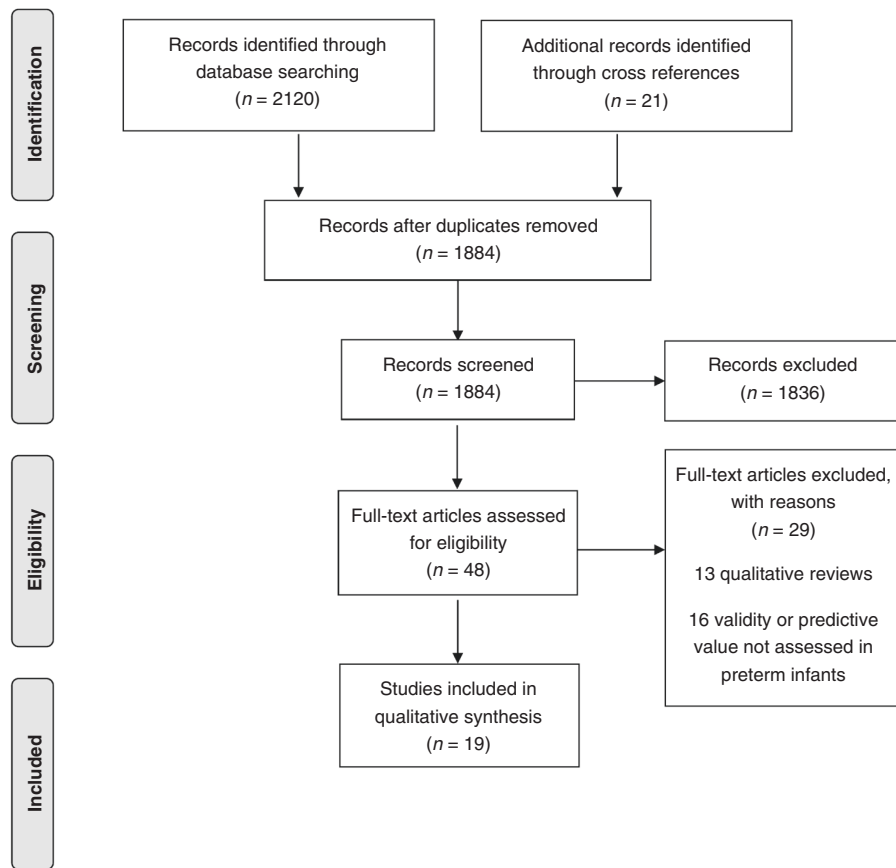


Fig. 1 PRISMA Flow diagram of included studies. Flow diagram.

SFT measurements

Table 4 describes the predictive value of SFT. Schmelzle et al. showed that the sum of SFT measured at four sites had a good predictive value for fat mass (g) in a study population that included both preterm as well as term infants.²⁶ However, only 10 out of 104 infants in this study population were born preterm.

Koo and colleagues also assessed the predictive value of SFT in a mixed population of term and preterm infants (respectively, $n = 68$ and $n = 52$). They reported that SFT, when added to weight and length, explained an additional 13% of the variance in fat free mass (g).¹⁷ Thus, SFT had a poor predictive value for fat free mass (g).

Schmelzle et al. as well as Koo et al. did not assess the predictive value of SFT for fat (free) mass percentage. In contrast, Daly-Wolfe et al. did and reported a poor predictive value of SFT for fat mass percentage (see Table 4).²⁴

Bioelectrical impedance measurements

Table 5 describes the predictive value and validity of bioelectrical impedance measurements. The impedance index (height^2 in $\text{cm}^2/\text{impedance}$ in Ω) measured with BIA adds little to the variance in fat-free mass already explained by weight²⁷ (see Table 5). Indeed Raghavan et al. reported that the least bias was obtained when weight alone was used to estimate TBW.²⁸ Table 5 shows that models used to estimate body composition based on the impedance index alone showed poor agreement. In contrast, the predictive equation by Dung et al. based on weight and the impedance index showed good agreement²⁷ (see Table 5).

Table 6 describes the predictive value of ultrasound measurements. Ultrasound measurements of muscle and fat mass showed high reliability (intraclass correlation coefficient 0.874–0.975; technical error of measurements 0.251–0.628 mm), but had a

poor predictive value for fat mass percentage measured by ADP²⁹ (see Table 6).

Table 7 describes the predictive value of MRI. Dyke and colleagues assessed the accuracy of body composition measured with rapid whole-body MRI. Repeated scans showed good agreement of fat mass percentage (95% limits of agreement 1.3%).³⁰ However, body composition measurement was not compared to a reference method (see Table 7).

Table 8 describes the validity of ADP. Compared to isotope dilution, ADP showed good agreement when measuring fat free mass density.³¹ Forsum et al. and Roggero et al. demonstrated that there was a small bias in the measurement of fat mass percentage.^{31,32} Nevertheless, the limits of agreement were relatively wide and thus there was poor agreement between fat mass percentage measured with ADP compared to fat mass percentage measured by isotope dilution (see Table 8).

DISCUSSION

Numerous studies have addressed how to measure body composition in preterm infants.^{17–35} Indeed there is an urgent need to monitor body composition in this population given the less favorable body composition in infancy and childhood as well as the increased risk of adverse cardiometabolic outcomes in later life.^{1–7} Nevertheless, till date, there is no consensus on which method should preferentially be used to assess body composition in preterm infants.

Reference methods

The studies included in this review used ADP, DXA, isotope dilution, and MRI as reference methods. In our opinion, ADP, DXA, and isotope dilution are acceptable reference methods.

Table 3. Body proportionality calculations to estimate body composition in preterm infants.

Author, year, country	Predictor for body composition	Reference method	Study setting		Time of assessment	Results	Quality assessment
			Subjects	No. of preterms			
Koo, 2004, United states of America	Circumferences of head, chest, abdomen, mid-arm, and mid-thigh Mid-arm and mid-thigh muscle and fat areas Total body weight Length	DXA (Hologic QDR 1000/W densitometer, pediatric software version 5.64P)	120 preterm and term infants	52 (GA 30–36 weeks)	3 (+2.8) days postnatal age	Results include preterm as well as term infants Weight and length accounted for >97% of the variance of fat free mass (g) in AGA and SGA infants and 46% of the variance in LGA infants. Mid-arm-to-head circumference ratio and mid-arm/thigh muscle areas accounted for, respectively, 17 and 6% of the variance in fat free mass (g). Chest, abdomen, mid-arm, and mid-thigh circumference did not explain any additional part of the variance in fat (free) mass (g) after including gender, race, gestation, weight and length in the predictive equation	Level of evidence 4 Strengths and limitations Assessment of agreement with reference method – Assessment of intraobserver and interobserver variability + Repeated measurements + Coefficients of variance assessed – Sensitivity analysis – Bootstrapping analysis – Cross-validation group – External validation – Large study population – Exclusively preterm infants –
Daly-Wolfe, 2015, United States of America	Mid-arm, mid-thigh, abdominal circumference	ADP	56 preterm and term infants	28 (GA 33.3 ± 0.9 weeks)	Discharge home	Mid-arm circumference was a covariate for fat mass percentage and accounted for 60.4% of the variance	Level of evidence 2 Strengths and limitations Assessment of agreement with reference method – Assessment of intraobserver and interobserver variability – Repeated measurements + Coefficients of variance assessed – Sensitivity analysis – Bootstrapping analysis – Cross-validation group – External validation – Large study population – Exclusively preterm infants –

Table 3. continued

Author, year, country	Predictor for body composition	Reference method	Study setting	Results	Quality assessment
Pereira-Da-Silva, 2009, Portugal	Mid-arm circumference Upper arm muscle and fat areas (calculated from mid-arm circumference and triceps skinfold thickness)	Regional adiposity measured by MRI	Subjects: 30 preterm infants No. of preterms: 30 (GA 30.7 ± 1.9 weeks) Time of assessment: 35 ± 1 weeks PMA	Mid-arm circumference predicted 41% of the variance in mid-arm circumference measured by MRI Arm fat and muscle area predicted, respectively, 40 and 7% of the variance in the area measured by MRI Mid-arm circumference predicted 4.8% of the variance in arm fat are (mm ²) measured by MRI	Level of evidence 2 Strengths and limitations Assessment of agreement with reference method – Assessment of intraobserver and interobserver variability + Repeated measurements + Coefficients of variance assessed + Sensitivity analysis – Bootstrapping analysis – Cross-validation group – External validation – Large study population – Exclusively preterm infants +
Kiger, 2016, United States of America	BMI Weight-for-length Ponderal index Body surface area Benn index	ADP	Subjects: 239 preterm infants No. of preterms: 239 (GA 28.1 [27.1–30.1] weeks) Time of assessment: PMA < 50: median 37 (IQR 34–39) weeks PMA > 50: median 54 (IQR 52–65) weeks	51% of the variance in fat mass percentage could be explained by weight-length indices before 50 weeks PMA and 16% after 50 weeks PMA The best fit regression model most closely matched BMI	Level of evidence 4 Strengths and limitations Assessment of agreement with reference method – Assessment of intraobserver and interobserver variability – Repeated measurements – Coefficients of variance assessed – Sensitivity analysis – Bootstrapping analysis – Cross-validation group – External validation – Large study population + Exclusively preterm infants +

Table 3. continued

Author, year, country	Predictor for body composition	Reference method	Study setting		Time of assessment	Results	Quality assessment
			Subjects	No. of preterms			
Larcade, 2017, France	Weight (z-score) Length (z-score) Head circumference (z-score)	ADP	155 preterm infants	155 (GA 29.2 ± 2.2 weeks)	Hospital discharge (35–41 weeks PMA)	Bias fat free mass –179 g (old predictive equation including weight z-scores and clinical parameters) Limits of agreement old equation ±249 g, i.e., poor agreement (mean fat free mass study population 2500 g) Bias fat free mass –12 g (new predictive equation, including weight z-scores and taking into account new feeding regimen) Limits of agreement new equation ±210 g, i.e., good agreement (mean fat free mass population 2500 g)	Level of evidence 2 Strengths and limitations Assessment of agreement with reference method + Assessment of intraobserver and interobserver variability – Repeated measurements – Coefficients of variance assessed + Sensitivity analysis + Bootstrapping analysis + Cross-validation group – External validation – Large study population + Exclusively preterm infants +
Simon, 2014, France	Weight (z-score) Length (z-score) Head circumference (z-score)	ADP	141 preterms	141 (GA 31.5 ± 2.5 weeks)	Last week before discharge. Mean PMA 37.3 ± 0.7 weeks	82% of the variance in fat free mass (g) was explained by predictive equation including clinical factors and change in weight z-score 24% of variance in fat mass percentage was explained by clinical factors and change in weight z-score	Level of evidence 4 Strengths and limitations Assessment of agreement with reference method – Assessment of intraobserver and interobserver variability – Repeated measurements + Coefficients of variance assessed – Sensitivity analysis – Bootstrapping analysis – Cross-validation group – External validation – Large study population + Exclusively preterm infants +

Table 3. continued

Author, year, country	Predictor for body composition	Reference method	Study setting	Time of assessment	Results	Quality assessment
Liotto, 2018, Italy	Weight Length Weight-for-length BMI Ponderal index	ADP	Subjects 1239 preterm and term infants	No. of preterms 585 (GA 32.4 ± 3.1 weeks)	Term equivalent age or 3 days postnatal age	Level of evidence 4 Strengths and limitations Assessment of agreement with reference method + Assessment of intraobserver and interobserver variability – Repeated measurements – Coefficients of variance assessed – Sensitivity analysis – Bootstrapping analysis – Cross-validation group – External validation – Large study population + Exclusively preterm infants –
Ramel, 2017, United States of America	Weight (z-score) Length (z-score) Head circumference (z-score) Weight-for-length Ponderal index BMI	ADP	218 preterm infants	218 (GA 33.8 ± 1.9 weeks)	1–3 days postnatal age	Level of evidence 4 Strengths and limitations Assessment of agreement with reference method – Assessment of intraobserver and interobserver variability – Repeated measurements – Coefficients of variance assessed – Sensitivity analysis – Bootstrapping analysis – Cross-validation group – External validation – Large study population + Exclusively preterm infants +

Table 3. continued

Author, year, country	Predictor for body composition	Reference method	Study setting	Time of assessment	Results	Quality assessment
Villar, 2017, United Kingdom	Weight-for-length BMI Ponderal index	ADP	Subjects 1019 preterm and term infants	No. of preterms 91 (GA 36.0 ± 0.7 weeks)	ADP within 96 h of birth, anthropometry within 12 h of birth	Level of evidence 4 Strengths and limitations Assessment of agreement with reference method — Assessment of intraobserver and interobserver variability — Repeated measurements — Coefficients of variance assessed — Sensitivity analysis — Bootstrapping analysis — Cross-validation group — External validation — Large study population — Exclusively preterm infants —

Large study population > 100 participants.

Levels of evidence based on the Oxford Centre for Evidence-based Medicine's Levels of Evidence.¹⁶

Level 1: Validating cohort study with good reference standards; or clinical decision rule tested within one clinical center.

Level 2: Exploratory cohort study with good reference standards; clinical decision rule after derivation or validated only on split-sample or databases.

Level 3: Non-consecutive study; or without consistently applied reference standards.

Level 4: Case-control study, poor or non-independent reference standard.

Level 5: Expert opinion without explicit critical appraisal, or based on physiology, bench research or "first principles."

ADP air displacement plethysmography, AGA appropriate for gestational age, BMI body mass index, DXA dual-energy x-ray absorptiometry, GA gestational age, LGA large for gestational age, MRI magnetic resonance imaging, PMA postmenstrual age, SFT skinfold thickness, SGA small for gestational age.

Table 4. Skinfold thickness measurements to estimate body composition in preterm infants.

Author, year, country	Predictor for body composition	Reference method	Study setting		Time of assessment	Results	Quality assessment
			Subjects	No. of preterms			
Koo, 2004, United states of America	Triceps, subscapular, suprailiac, and mid-thigh SFT on the left side	DXA (Hologic QDR 1000/W densitometer, pediatric software version 5.64P)	120 preterm and term infants	52 (GA 30–36 weeks)	3 (+2.8) days postnatal age	Results include preterm as well as term infants When added to weight and length, SFT explained an additional 13% of the variance in fat free mass (g)	Level of evidence 4 Strengths and limitations Assessment of agreement with reference method – Assessment of intraobserver and interobserver variability + Repeated measurements + Coefficients of variance assessed – Sensitivity analysis – Bootstrapping analysis – Cross validation group – External validation – Large study population – Exclusively preterm infants –
Schmelzle, 2002, Germany	Triceps, biceps, suprailiac, and subscapular SFTs	DXA (Hologic QDR 1500, infant whole-body software version 5.67)	104 preterm and term infants	10 (GA 34–36 weeks)	1–10 days postnatal age, and at 2 and 4 months of age	Results include preterm as well as term infants Newly modeled predictive equation based on SFT explained 94% of the variance in fat mass (g). The mean error of the predictive equation for the estimation of fat mass (g) was ± 75 , ± 170 , ± 300 , and ± 380 g for infants with an FM ≤ 500 , 501–1000, 1001–2000, or > 2000 g, respectively Limits of agreement not reported Previously modeled predictive equations could not be validated	Level of evidence 2 Strengths and limitations Assessment of agreement with reference method + Assessment of intraobserver and interobserver variability – Repeated measurements + Coefficients of variance assessed – Sensitivity analysis – Bootstrapping analysis + Cross-validation group + External validation – Large study population – Exclusively preterm infants –

Table 4. continued

Author, year, country	Predictor for body composition	Reference method	Study setting		Results	Quality assessment
			Subjects	No. of preterms		
Daly-Wolfe, 2015, United States of America	Triceps, biceps, suprailiac, and subscapular SFTs	ADP	56 preterm and term infants	28 (GA 33.3 ± 0.9 weeks)	Suprailiac SFT was a covariate for fat mass percentage and together with mid-arm circumference accounted for 49% of the variance	Level of evidence 2 Strengths and limitations Assessment of agreement with reference method — Assessment of intraobserver and interobserver variability — Repeated measurements + Coefficients of variance assessed — Sensitivity analysis — Bootstrapping analysis — External validation — Large study population — Exclusively preterm infants —

Large study population >100 participants.

Levels of evidence based on the Oxford Centre for Evidence-based Medicine's Levels of Evidence.¹⁶

Level 1: Validating cohort study with good reference standards; or clinical decision rule tested within one clinical center.

Level 2: Exploratory cohort study with good reference standards; clinical decision rule after derivation, or validated only on split-sample or databases.

Level 3: Non-consecutive study; or without consistently applied reference standards.

Level 4: Case-control study, poor or non-independent reference standard.

Level 5: Expert opinion without explicit critical appraisal, or based on physiology, bench research, or "first principles."

ADP air displacement plethysmography, DXA dual-energy x-ray absorptiometry, GA gestational age, SFT skinfold thickness.

Table 5. Bioelectrical impedance to estimate body composition in preterm infants.

Author, year, country	Predictor for body composition	Reference method	Study setting		Time of assessment	Results	Quality assessment
			Subjects	No. of preterms			
Dung, 2007, Germany	Impedance index derived from monofrequency BIA	DXA (Hologic QDR 1500, infant software version 5.67)	118 preterm infants	118 (GA 30.1 ± 3.1 weeks)	38.6 ± 3.8 weeks PMA	Weight and impedance index together explained 94% of the variance in fat free mass (g) Precision in prediction gained by impedance index was max 3.03% Bias fat free mass: -40 g (estimated with weight and impedance index) Limits of agreement ±20 g, i.e., good agreement (mean fat free mass study population 2260 g)	Level of evidence 2 Strengths and limitations Assessment of agreement with reference method + Assessment of intraobserver and interobserver variability – Repeated measurements – Coefficients of variance assessed – Sensitivity analysis – Bootstrapping analysis + External validation – Large study population + Exclusively preterm infants +
Raghavan, 1998, United States of America	Impedance index derived from monofrequency BIA	Isotope dilution	42 preterm infants	42 (GA 24–34 weeks)	Day 1–7 postnatally	Weight, length and the impedance index together explained 90% of the variance in TBW. Bias TBW 25.7 g (estimated with impedance alone) Limits of agreement TBW ± 95 g, i.e., poor agreement (mean TBW study population 794 g)	Level of evidence 4 Strengths and limitations Assessment of agreement with reference method + Assessment of intraobserver and interobserver variability – Repeated measurements + Coefficients of variance assessed + Sensitivity analysis – Bootstrapping analysis – Cross-validation group + External validation – Large study population – Exclusively preterm infants –

Table 5. continued

Author, year, country	Predictor for body composition	Reference method	Study setting		Results	Quality assessment
			Subjects	No. of preterms		
Kushner, 1992, Unites States of America and Peru	Impedance index derived from monofrequency BIA	Isotope dilution	133 preterm infants, prepubertal children and adults	32 (GA 33 ± 1.6 weeks)	Day 1 and day 4–7 postnatally	<p>Level of evidence 2</p> <p>Strengths and limitations</p> <p>Assessment of agreement with reference method –</p> <p>Assessment of intraobserver and interobserver variability –</p> <p>Repeated measurements +</p> <p>Coefficients of variance assessed +</p> <p>Sensitivity analysis –</p> <p>Bootstrapping analysis –</p> <p>Cross-validation group +</p> <p>External validation –</p> <p>Large study population –</p> <p>Exclusively preterm infants –</p>
Tang, 1997, United Kingdom	Impedance index derived from monofrequency BIA	Isotope dilution	28 preterm and term infants	Unknown GA median 30.5, range 24–38 weeks	Day 1–7 postnatally	<p>Level of evidence 4</p> <p>Strengths and limitations</p> <p>Assessment of agreement with reference method –</p> <p>Assessment of intraobserver and interobserver variability –</p> <p>Repeated measurements –</p> <p>Coefficients of variance assessed –</p> <p>Sensitivity analysis –</p> <p>Bootstrapping analysis –</p> <p>External validation –</p>

Table 5. continued

Author, year, country	Predictor for body composition	Reference method	Study setting	No. of preterms	Time of assessment	Results	Quality assessment
Collins, 2013, Australia	Bioelectrical impedance spectroscopy	Isotope dilution	Subjects 99 preterm infants	99 (GA 33 ± 2 weeks)	Mean PMA 35 weeks	Bias TBW 10 ml Bias extracellular water 40 ml Limits of agreement TBW ± 650 ml, i.e., poor agreement (mean TBW study population 1760 ml) Limits of agreement extracellular water ± 360 ml, i.e., poor agreement (mean extracellular water study population 1025 ml)	Level of evidence 2 Strengths and limitations Assessment of agreement with reference method + Assessment of intraobserver and interobserver variability – Repeated measurements + Coefficients of variance assessed – Sensitivity analysis – Bootstrapping analysis – Cross-validation group + External validation – Large study population – Exclusively preterm infants +

Large study population > 100 participants.

Levels of evidence based on the Oxford Centre for Evidence-based Medicine's Levels of Evidence.¹⁶

Level 1: Validating cohort study with good reference standards; or clinical decision rule tested within one clinical center.

Level 2: Exploratory cohort study with good reference standards; clinical decision rule after derivation, or validated only on split-sample or databases.

Level 3: Non-consecutive study; or without consistently applied reference standards.

Level 4: Case-control study, poor or non-independent reference standard.

Level 5: Expert opinion without explicit critical appraisal, or based on physiology, bench research, or "first principles."

BIA bioelectrical impedance, CI confidence interval, DXA dual-energy x-ray absorptiometry, GA gestational age, PMA postmenstrual age, SFT skinfold thickness, TBW total body water.

Table 6. Ultrasound to estimate body composition in preterm infants.

Author, year, country	Measurement method	Reference method	Study setting		Results	Quality assessment
			Subjects	No. of preterms		
Nagel, 2020, United States of America	Ultrasound images of the biceps (brachii and brachialis), abdomen (rectus abdominis), and quadriceps (rectus femoris and vastus intermedius) using a portable B-mode US device	ADP	63 preterm infants	63 (GA 32.0 ± 2.2 weeks)	PMA 35.1 + 1.2 weeks Biceps adipose tissue and total sum of adipose tissue thickness explained, respectively, 17% and 16% of fat mass percentage	Level of evidence 2 <u>Strengths and limitations</u> Assessment of agreement with reference method – Assessment of intraobserver and interobserver variability + Repeated measurements + Coefficients of variance assessed + Sensitivity analysis + Bootstrapping analysis – Cross-validation group – External validation – Large study population – Exclusively preterm infants +

Large study population >100 participants.

Levels of evidence based on the Oxford Centre for Evidence-based Medicine's Levels of Evidence.¹⁶

Level 1: Validating cohort study with good reference standards; or clinical decision rule tested within one clinical center.

Level 2: Exploratory cohort study with good reference standards; clinical decision rule after derivation, or validated only on split-sample or databases.

Level 3: Non-consecutive study; or without consistently applied reference standards.

Level 4: Case-control study, poor or non-independent reference standard.

Level 5: Expert opinion without explicit critical appraisal, or based on physiology, bench research, or "first principles."

ADP air displacement plethysmography, GA gestational age, PMA postmenstrual age.

Table 7. Magnetic resonance imaging to estimate body composition in preterm infants.

Author, year, country	Measurement method	Reference method	Study setting		Results	Quality assessment
			Subjects	No. of preterms		
Dyke, 2018, United States of America	Rapid whole-body imaging	No reference method for fat (free) mass. A weighing scale used as reference for weight	25 preterm and term infants	15 (GA 25.9 ± 1.3 weeks)	Term (equivalent) age	Level of evidence 2 Strengths and limitations Assessment of agreement with reference method – Assessment of intraobserver and interobserver variability – Repeated measurements + Coefficients of variance assessed – Sensitivity analysis – Bootstrapping analysis – Cross-validation group – External validation – Large study population – Exclusively preterm infants –

Large study population > 100 participants.

Levels of evidence based on the Oxford Centre for Evidence-based Medicine's Levels of Evidence.¹⁶

Level 1: Validating cohort study with good reference standards; or clinical decision rule tested within one clinical center.

Level 2: Exploratory cohort study with good reference standards; clinical decision rule after derivation, or validated only on split-sample or databases.

Level 3: Non-consecutive study; or without consistently applied reference standards.

Level 4: Case-control study, poor or non-independent reference standard.

Level 5: Expert opinion without explicit critical appraisal, or based on physiology, bench research, or "first principles."

GA gestational age, MRI magnetic resonance imaging.

Table 8. Air displacement plethysmography to estimate body composition in preterm infants.

Author, year, country	Measurement method	Reference method	Study setting		Results	Quality assessment
			Subjects	No. of preterms		
Forsum, 2016, Sweden	ADP	Isotope dilution	14 preterm infants	14 (GA 34.1 ± 1.3 weeks)	Bias fat mass percentage -1% Limits of agreement ±5.8%, i.e., poor agreement (study population mean fat mass percentage 3.2%) Bias fat free mass density g/ml -0.002 g/ml Limits of agreement ±0.012 g/ml, i.e., good agreement (study population mean fat free mass density 1.06 g/ml)	Level of evidence 2 Strengths and limitations Assessment of agreement with reference method + Assessment of intraobserver and interobserver variability - Repeated measurements + Coefficients of variance assessed - Sensitivity analysis - Bootstrapping analysis - Cross-validation group - External validation - Large study population - Exclusively preterm infants +
Roggero, 2012, Italy	ADP	Isotope dilution	79 preterm and term infants	70 (GA 24-36 weeks)	<i>Precision of ADP was assessed in 57 preterm infants</i> Bias for fat mass percentage 0.15% between repeated ADP measurements Limits of agreement ±2.2%, i.e., poor agreement between repeated ADP measurements (study population mean fat mass percentage 9.2%) <i>Accuracy was assessed in subgroup of 10 preterm infants</i> Bias fat mass percentage 0.32% Limits of agreement ±3.1%, i.e., poor agreement (study population mean fat mass percentage 6.0%) <i>Interdevice reliability was assessed in a subgroup of 12 preterm and term infants</i> Bias fat mass percentage 0.42% Limits of agreement ±2.29%, i.e., poor agreement (study population mean fat mass percentage 8.8%)	Level of evidence 2 Strengths and limitations Assessment of agreement with reference method + Assessment of intraobserver and interobserver variability - Repeated measurements + Coefficients of variance assessed + Sensitivity analysis - Bootstrapping analysis - Cross-validation group - External validation - Large study population - Exclusively preterm infants -

Large study population > 100 participants.

Levels of evidence based on the Oxford Centre for Evidence-based Medicine's Levels of Evidence.¹⁶

Level 1: Validating cohort study with good reference standards; or clinical decision rule tested within one clinical center.

Level 2: Exploratory cohort study with good reference standards; clinical decision rule after derivation, or validated only on split-sample or databases.

Level 3: Non-consecutive study; or without consistently applied reference standards.

Level 4: Case-control study, poor or non-independent reference standard.

Level 5: Expert opinion without explicit critical appraisal, or based on physiology, bench research, or "first principles."

ADP air displacement plethysmography, GA gestational age.

ADP, DXA, and isotope dilution have been validated against chemical carcass analysis in piglets.^{36–40} The body composition of piglets are considered to be comparable to the body composition of preterm infants.³⁷ Therefore, in practice, these methods are accepted as accurate measures. Nevertheless, we should take into consideration that there is variation within and between these methods. For example, different types of software are used to analyze DXA. It has been reported that pediatric and infant software rely on different assumptions and yield varying results.⁴¹ Furthermore, as discussed below (under the accuracy of ADP), the statistical agreement between ADP and isotope dilution may be interpreted as poor.^{31,32} Therefore, we believe that some reservation is needed when comparing different reference methods.

In contrast to ADP, DXA, and isotope dilution, MRI has not been validated against chemical carcass analysis in subjects comparable to the neonatal population, but it has been found accurate in adult human cadaver and animal studies.^{42,43} Over the years, MRI has been increasingly used to measure body composition in the neonatal population as well.^{30,44–46} However, it is yet to be widely implemented and we find it preliminary to use MRI as a reference method in comparative studies with preterm infants.

In our opinion, due to the lack of a gold standard and the difference between reference methods, some reservation is needed when drawing conclusions on the validation studies included in this review.

Assessment of validity

To assess whether two methods agree, Bland–Altman analyses are an accepted and widely implemented method.⁴⁷ Agreement should be based on a maximum difference between the two methods that is clinically acceptable.^{47–50} Nevertheless, this so-called predefined clinical agreement limit was omitted in all the studies included in this review. Hence, the interpretation of these studies is limited.

Accuracy of body proportionality measures

Studies conducted so far have mainly assessed the predictive value of body proportionality measures.^{17,18,20,21} Two studies assessed the agreement between ADP and a predictive equation including weight and clinical parameters.^{19,22} Liotto and colleagues found poor agreement between ADP and their predictive equation which estimated fat (free) mass adjusted by length (g/cm).¹⁹ However, they included both preterm and term infants in their analysis, which makes it difficult to extrapolate their findings to only preterm infants—the target population of this review. Larcade et al.²² investigated a study population of exclusively preterm infants and could not validate the predictive equation for fat free mass (g) made by Simon and colleagues.²³ A difference in nutritional care and ensuing better growth in Larcade's population may have been the cause of an underestimation of fat free mass by the previously modeled equation.

In our opinion, it is difficult to develop a predictive equation that can be validated externally. Just as Larcade et al. found nutritional practices to influence the prediction model, changes in neonatal care over time and across neonatal intensive care units (NICUs) influence the predictive equations and limit their universal application. Moreover, investigators have used mixed study populations which include small, appropriate as well as large for gestational age infants.¹⁷ Meanwhile, Koo et al. demonstrated that associations between weight/length indices and body composition differ for those born large for gestational age, which makes the use of a mixed study population inappropriate.¹⁷

In addition, it is important to note that predictive equations generally found that a large proportion of the variance in fat (free) mass (g) could be explained by weight or BMI. This logically follows the fact that fat mass (g) and fat free mass (g) together make up total body weight. However, fat mass percentage was

poorly explained by weight or length indices. Meanwhile, in our opinion, fat (free) mass percentage, may be a more relevant parameter when it comes to comparing the body composition of an individual or groups because it takes the subject's weight into account.

All in all, we conclude that the predictive equations based on weight and length indices currently cannot be implemented in clinical practice, because of the lack of external validation and a poor predictive value for fat (free) mass percentage.

Daly-Wolfe and colleagues found that mid-arm circumference had a moderate predictive value for fat mass percentage measured with ADP.²⁴ Koo et al., on the other hand, found that mid-arm circumference together with chest, abdomen, and mid-thigh circumference, added <5% to the variance in fat mass percentage already explained by weight and length. In their study, however, fat mass was determined by DXA. Pereira-Da-Silva and colleagues compared upper arm anthropometry to regional fat mass measured with MRI and found it to be an inaccurate predictor of regional body composition.²⁵ In our view, currently there is inconclusive evidence on the predictive value of body area circumferences and more research is needed to assess the potential of mid-arm circumference as a predictor of whole body or regional fat mass.

Accuracy of skinfold measurements

Several studies assess SFT in preterm infants.^{24,51–56} For example, Daly-Wolfe and colleagues investigated the predictive value of SFT and found that, together with the mid-arm circumference, SFT explained 49% of the variance in fat mass percentage.²⁴ However, only one study included in the review assessed the validity of SFT in preterm infants.²⁶ Unfortunately, only 10 late preterm infants were included in this study and analysis included term infants as well, so no robust conclusions can be drawn from their findings.²⁶ Recently, we found poor agreement between SFT and body composition in a study with exclusively preterm infants.⁵⁶

Moreover, SFT is influenced by the fluid status and there is a high interobserver variability.¹³ In addition, it could be deemed controversial to use SFT calipers in extremely preterm infants in light of their vulnerable skin. Therefore, there is insufficient evidence to support SFT as an clinically useful measure of fat mass in preterm infants at this time.

Accuracy of bioelectrical impedance

Both BIA, as well as bioelectrical impedance spectroscopy, have a poor predictive value for TBW measured with isotope dilution analyses^{28,35,57} and body composition measured with DXA.²⁷ BIA did not seem to provide an additive predictive effect for fat-free mass or TBW, compared to body weight alone.^{27,28} Though both Kushner and colleagues as well as Tang and colleagues found the impedance index (cm^2/Ω) to significantly improve the prediction of TBW, the majority of variance in TBW measured with isotope dilution was still explained by weight.^{33,57} Kushner and colleagues concluded that the impedance index (cm^2/Ω) explained 99% of the variance in TBW. However, in their subgroup of preterm infants there was a significant bias: a higher variance was found for higher values of the impedance index. This bias was only eliminated by the addition of weight to the prediction equation³³—implying that bioelectrical impedance on its own is not an adequate predictor of body composition in preterm infants.

Accuracy of ultrasound

Ahmad et al. previously demonstrated that ultrasound measurements correlate with fat mass in preterm infants.⁵⁸ Depending on the site of ultrasound measurement the intraobserver variability was reported to be up to 14.7% in preterm infants.⁵⁹ To our knowledge, the interobserver variability has not been investigated in preterm infants but has been reported to show high interobserver agreement (0.89–0.95) in 1- and 2-year-olds.⁶⁰

Nevertheless, we only found one study which assessed the predictive value of ultrasound measurement for body composition in preterm infants and they found a poor predictive value.²⁹ Others did report on ultrasound measurements as a means to estimate body composition in preterm infants, but they did not assess the predictive value or validity.^{58,59} Hence, in our opinion, there seems to be insufficient evidence for the use of ultrasound as a reference method for body composition. However, since ultrasound measurements showed high reliability, it may be of interest to investigate whether other body sites are a better representation of body composition.

Accuracy of MRI

Though several authors suggested the use of MRI to measure body composition in preterm infants,^{44,61} only one study was found that assessed the validity of MRI for the assessment of body composition in a small study population.³⁰ Despite a sound assessment of repeatability, the actual fat mass (g) measurement was not compared with other techniques. Therefore, it should be concluded that more studies are necessary to draw conclusions on the use of MRI in determining fat mass in preterm infants.

Accuracy of ADP

Roggero and colleagues found a small bias when comparing ADP to isotope dilution.³² However, the limits of agreement were relatively wide, resulting in poor agreement between the two methods. Of note, the accuracy of ADP was only assessed in a small subgroup of 10 preterm infants. In line with their findings Forsum et al. also found a small bias when comparing ADP with deuterium dilution in 14 preterm infants.³¹ Fat mass percentage, however, had relatively wide limits of agreement and thus poor agreement. Fat free mass density (g/ml), however, agreed well. Precision was studied by Roggero in a larger group of 57 preterm infants and also showed a small bias for fat mass percentage between repeated ADP measurements.³² Despite wide limits of agreement the authors concluded that ADP shows good agreement with isotope dilution for fat mass percentage as well as fat free mass density.^{31,32} Likewise, carcass analyses showed small bias, with relatively wide limits of agreement.³⁶ Nevertheless, it is generally accepted that ADP is a reliable method in infants. Taking into account the relatively small study populations and small bias, it is to be expected that larger studies would yield better agreement. Therefore, we conclude that ADP is a reliable method to assess fat mass in preterm infants.

Accuracy of isotope dilution

Isotope dilution is a well-established method for the measurement of TBW from which the fat free mass can be derived.⁶² As a result there were no studies validating isotope dilution against another reference method, such as DXA, in preterm infants. Hartnoll and colleagues, however, did show similar results as late nineteenth century cadaver studies, even though no comparative analysis was done.⁶³ A conclusion thus cannot be drawn based on human studies in preterm infants but is deemed reliable based on carcass analysis of piglets.³⁷

Accuracy of DXA

There are no comparative studies with preterm infants where DXA was compared with other methods, such as isotope dilution. DXA has been validated in piglets^{38,40} and in practice is accepted as an accurate measure. Nevertheless, in human as well as in animal studies DXA has been reported to overestimate fat mass, especially in lower weights.^{40,64} Moreover, different software algorithms yield varying results in body composition.⁴¹

It would be insightful to investigate the agreement between isotope dilution, ADP, and DXA in preterm infants. This would help to give us guidance on which method should be preferentially used to assess body composition in preterm infants.

Patient-friendliness, ease of use, and costs

Methods are chosen based on local experience and available resources. In a clinical setting, there is a preference for quick, easy-to-use, but accurate methods, which could be used at the bedside. In contrast, in a research setting there is more room for less flexible methods.

Body proportionality measures are quick, low-cost, and minimally invasive methods. They are ideal in the intensive care setting as well as in outpatient department for follow-up. Unfortunately, studies so far have not confirmed these methods have sufficient accuracy. Likewise, SFT and bioelectrical impedance techniques are easy bedside methods, but robust evidence supporting their use is lacking. Moreover, there is a high interobserver variability, and in extremely preterm infants SFT measurements should be taken with caution to prevent injury of their vulnerable skin. Nevertheless, when used with caution SFT is a safe, non-invasive method.¹⁰

DXA, ADP, and isotope dilution are accurate methods but have some practical downsides. For DXA, infants need to be clinically stable and free from respiratory support and monitoring, making DXA more appropriate from term age onwards. Taking into account that movement is not allowed, it could be used in infants up to 6 months corrected age who can be swaddled or nursed to sleep during the procedure.

ADP could be used in the NICU if the infant is clinically stable and not on respiratory support. Nonetheless, the machinery is bulky and recalibration is needed after movement, making it less suitable for such use.

On the contrary, isotope dilution can be used in small infants who are not stable yet, making it very suitable for use in the NICU—were it not that it has a relatively high workload. In addition, oral administration of the isotope solution is challenging in older infants.

MRI is yet to become a widely used method to measure fat mass in preterm infants, but seems promising with high precision. Furthermore, MRI is a safe, radiation-free method. Nevertheless, infants on respiratory support and monitoring cannot easily undergo an MRI procedure. All in all, at this time there is no prospect of an accurate, easy, low-cost point-of-care instrument that could be used during the NICU stay or the follow-up period in a clinical setting. In a research setting, ADP may be the most practical, yet reliable, method to use in infants up to 6 months corrected age.

Limitations

This review included all potential reference methods, which made it challenging to come to a concise conclusion. Furthermore, it is important to note that, overall, validation studies were conducted in a limited number of study subjects with a wide range of gestational ages and varying postnatal ages at the time of assessment. Hence, it was even more difficult to draw definitive conclusions on the assessment of body composition at different gestational and postnatal ages. Moreover, the reference methods used in the various studies have not been validated in humans or were only validated in a small number of subjects. This lack of a solid golden standard further undermines any conclusions drawn from these studies. In addition, only formally published data was included, leaving potential publication bias unassessed. There is a need for larger cross-sectional studies comparing these instruments at different time points as well as longitudinal studies investigating the accuracy of the use of the instruments over time.

CONCLUSIONS

Monitoring body composition remains important in the light of the increased cardiometabolic disease risk in adults born prematurely.^{1–7} Therefore, the quest for not only accurate but also practical methods to assess body composition should continue.

This review reaffirmed that weight and length indices, body area circumferences, SFT, BIA, and ultrasound do not adequately reflect body composition. MRI looks promising for the use in preterm infants but has not been validated for the measurement of body composition. On the other hand, DXA, ADP, and isotope dilution methods are considered trustworthy and validated techniques. Nevertheless, this review showed that these methods may not yield comparable results. Therefore, caution should be taken when comparing body composition measured with different methods. Moreover, to facilitate future studies and support clinical practice it would be valuable for researchers and physicians to come to an agreement on which reference should preferentially be used to measure body composition in preterm infants.

DATA AVAILABILITY

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

REFERENCES

- Kerkhof, G. F., Breukhoven, P. E., Leunissen, R. W., Willemsen, R. H. & Hokken-Koelega, A. C. Does preterm birth influence cardiovascular risk in early adulthood? *J. Pediatr.* **161**, 390.e1–396.e1 (2012).
- Sipola-Leppanen, M. et al. Cardiometabolic risk factors in young adults who were born preterm. *Am. J. Epidemiol.* **181**, 861–873 (2015).
- Morrison, K. M. et al. Cardiometabolic health in adults born premature with extremely low birth weight. *Pediatrics* **138**, e20160515 (2016).
- Johnson, M. J., Wootton, S. A., Leaf, A. A. & Jackson, A. A. Preterm birth and body composition at term equivalent age: a systematic review and meta-analysis. *Pediatrics* **130**, e640–e649 (2012).
- Stigson, L. et al. Bone and fat mass in relation to postnatal levels of insulin-like growth factors in prematurely born children at 4 y of age. *Pediatr. Res.* **75**, 544–550 (2014).
- Darlow, B. A., Martin, J. & Horwood, L. J. Metabolic syndrome in very low birth weight young adults and controls: the New Zealand 1986 VLBW Study. *J. Pediatr.* **206**, 128–133 e125 (2019).
- Crump, C. et al. Association of preterm birth with risk of ischemic heart disease in adulthood. *JAMA Pediatr.* **173**, 736–743 (2019).
- Liu, P., Ma, F., Lou, H. & Liu, Y. The utility of fat mass index vs. body mass index and percentage of body fat in the screening of metabolic syndrome. *BMC Public Health* **13**, 629 (2013).
- Ramirez-Velez, R. et al. Percentage of body fat and fat mass index as a screening tool for metabolic syndrome prediction in Colombian University students. *Nutrients* **9**, 1009 (2017).
- Strydom, K., Van Niekerk, E. & Dhansay, M. A. Factors affecting body composition in preterm infants: assessment techniques and nutritional interventions. *Pediatr. Neonatol.* **60**, 121–128 (2019).
- Mazahery, H., von Hurst, P. R., McKinlay, C. J. D., Cormack, B. E. & Conlon, C. A. Air displacement plethysmography (pea pod) in full-term and pre-term infants: a comprehensive review of accuracy, reproducibility, and practical challenges. *Matern. Health Neonatol. Perinatol.* **4**, 12 (2018).
- Andrews, E. T., Beattie, R. M. & Johnson, M. J. Measuring body composition in the preterm infant: evidence base and practicalities. *Clin. Nutr.* **38**, 2521–2530 (2019).
- Demerath, E. W. & Fields, D. A. Body composition assessment in the infant. *Am. J. Hum. Biol.* **26**, 291–304 (2014).
- Moore, D. S., Notz, W. I. & Flinger, M. A. *The Basic Practice of Statistics* (W. H. Freeman and Company, 2013).
- Critical Appraisal Skills Programme. CASP checklist. <https://casp-uk.net/casp-tools-checklists/> (2018).
- Philips, B. et al. Oxford Centre for Evidence-Based Medicine: Levels of Evidence (March 2009). <https://www.cebm.ox.ac.uk/resources/levels-of-evidence/oxford-centre-for-evidence-based-medicine-levels-of-evidence-march-2009> (2009).
- Koo, W. W., Walters, J. C. & Hockman, E. M. Body composition in neonates: relationship between measured and derived anthropometry with dual-energy x-ray absorptiometry measurements. *Pediatr. Res.* **56**, 694–700 (2004).
- Kiger, J. R., Taylor, S. N., Wagner, C. L., Finch, C. & Katikaneni, L. Preterm infant body composition cannot be accurately determined by weight and length. *J. Neonatal Perinat. Med.* **9**, 285–290 (2016).
- Liotto, N. et al. Can basic characteristics estimate body composition in early infancy? *J. Pediatr. Gastroenterol. Nutr.* **66**, e76–e80 (2018).
- Ramel, S. E., Zhang, L., Misra, S., Anderson, C. G. & Demerath, E. W. Do anthropometric measures accurately reflect body composition in preterm infants? *Pediatr. Obes.* **12**, 72–77 (2017).
- Villar, J. et al. Body composition at birth and its relationship with neonatal anthropometric ratios: the newborn body composition study of the INTERGROWTH-21st Project. *Pediatr. Res.* **82**, 305–316 (2017).
- Larcade, J. et al. Estimation of fat-free mass at discharge in preterm infants fed with optimized feeding regimen. *J. Pediatr. Gastroenterol. Nutr.* **64**, 115–118 (2017).
- Simon, L. et al. Determinants of body composition in preterm infants at the time of hospital discharge. *Am. J. Clin. Nutr.* **100**, 98–104 (2014).
- Daly-Wolfe, K. M., Jordan, K. C., Slater, H., Beachy, J. C. & Moyer-Mileur, L. J. Mid-arm circumference is a reliable method to estimate adiposity in preterm and term infants. *Pediatr. Res.* **78**, 336–341 (2015).
- Pereira-Da-Silva, L., Abecasis, F., Virella, D. & Videira-Amaral, J. M. Upper arm anthropometry is not a valid predictor of regional body composition in preterm infants. *Neonatology* **95**, 74–79 (2009).
- Schmelzle, H. R. & Fusch, C. Body fat in neonates and young infants: validation of skinfold thickness versus dual-energy x-ray absorptiometry. *Am. J. Clin. Nutr.* **76**, 1096–1100 (2002).
- Dung, N. Q., Fusch, G., Armbrust, S., Jochum, F. & Fusch, C. Body composition of preterm infants measured during the first months of life: bioelectrical impedance provides insignificant additional information compared to anthropometry alone. *Eur. J. Pediatr.* **166**, 215–222 (2007).
- Raghavan, C. V. et al. Estimation of total body water in very-low-birth-weight infants by using anthropometry with and without bioelectrical impedance and H₂[(18)O]. *Am. J. Clin. Nutr.* **68**, 668–674 (1998).
- Nagel, E. et al. Can ultrasound measures of muscle and adipose tissue thickness predict body composition of premature infants in the neonatal intensive care unit? *J. Parenter. Enteral Nutr.* **45**, 323–330 (2020).
- Dyke, J. P., Garfinkel, A. C., Groves, A. M. & Kovanlikaya, A. High-resolution rapid neonatal whole-body composition using 3.0 Tesla chemical shift magnetic resonance imaging. *Pediatr. Res.* **83**, 638–644 (2018).
- Forsum, E., Olhager, E. & Tornqvist, C. An evaluation of the pea pod system for assessing body composition of moderately premature infants. *Nutrients* **8**, 238 (2016).
- Roggero, P. et al. Evaluation of air-displacement plethysmography for body composition assessment in preterm infants. *Pediatr. Res.* **72**, 316–320 (2012).
- Kushner, R. F., Schoeller, D. A., Fjeld, C. R. & Danford, L. Is the impedance index (Ht²/R) significant in predicting total body water? *Am. J. Clin. Nutr.* **56**, 835–839 (1992).
- Tang, W., Ridout, D. & Modi, N. Influence of respiratory distress syndrome on body composition after preterm birth. *Arch. Dis. Child. Fetal Neonatal Ed.* **77**, F28–F31 (1997).
- Collins, C. T. et al. Prediction of body water compartments in preterm infants by bioelectrical impedance spectroscopy. *Eur. J. Clin. Nutr.* **67**, S47–S53 (2013).
- Fronchas-Chauty, A., Louveau, I., Le Huerou-Luron, I., Roze, J. C. & Darmaun, D. Air-displacement plethysmography for determining body composition in neonates: validation using live piglets. *Pediatr. Res.* **72**, 26–31 (2012).
- Rudolph, B. C., Stahly, T. S. & Cromwell, G. L. Estimation of body composition of neonatal pigs via deuterium oxide dilution: validation of technique. *J. Anim. Sci.* **66**, 53–61 (1988).
- Brunton, J. A., Bayley, H. S. & Atkinson, S. A. Validation and application of dual-energy x-ray absorptiometry to measure bone mass and body composition in small infants. *Am. J. Clin. Nutr.* **58**, 839–845 (1993).
- Picaud, J. C., Rigo, J., Nyamugabo, K., Milet, J. & Senterre, J. Evaluation of dual-energy x-ray absorptiometry for body-composition assessment in piglets and term human neonates. *Am. J. Clin. Nutr.* **63**, 157–163 (1996).
- Rigo, J. et al. Reference values of body composition obtained by dual energy x-ray absorptiometry in preterm and term neonates. *J. Pediatr. Gastroenterol. Nutr.* **27**, 184–190 (1998).
- Picaud, J. C. et al. Software and scan acquisition technique-related discrepancies in bone mineral assessment using dual-energy x-ray absorptiometry in neonates. *Acta Paediatr.* **91**, 1189–1193 (2002).
- Fowler, P. A., Fuller, M. F., Glasbey, C. A., Cameron, G. G. & Foster, M. A. Validation of the in vivo measurement of adipose tissue by magnetic resonance imaging of lean and obese pigs. *Am. J. Clin. Nutr.* **56**, 7–13 (1992).
- Abate, N., Burns, D., Peshock, R. M., Garg, A. & Grundy, S. M. Estimation of adipose tissue mass by magnetic resonance imaging: validation against dissection in human cadavers. *J. Lipid Res.* **35**, 1490–1496 (1994).
- Uthaya, S., Bell, J. & Modi, N. Adipose tissue magnetic resonance imaging in the newborn. *Horm. Res.* **62**, 143–148 (2004).
- Harrington, T. A. et al. Fast and reproducible method for the direct quantitation of adipose tissue in newborn infants. *Lipids* **37**, 95–100 (2002).

46. Stokes, T. A. et al. The clinical utility of anthropometric measures to assess adiposity in a cohort of prematurely born infants: correlations with MRI fat quantification. *J. Neonatal Perinat. Med.* **10**, 133–138 (2017).
47. Lu, M. J. et al. Sample size for assessing agreement between two methods of measurement by Bland-Altman method. *Int. J. Biostat.* <https://doi.org/10.1515/ijb-2015-0039> (2016).
48. Bland, J. M. & Altman, D. G. Comparing methods of measurement: why plotting difference against standard method is misleading. *Lancet* **346**, 1085–1087 (1995).
49. Bland, J. M. & Altman, D. G. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* **1**, 307–310 (1986).
50. Giavarina, D. Understanding Bland Altman analysis. *Biochem. Med.* **25**, 141–151 (2015).
51. Bustamante, S. A., Jacobs, P. & Gaines, J. A. Body weight, static and dynamic skinfold thickness in small premature infants during the first month of life. *Early Hum. Dev.* **8**, 217–224 (1983).
52. Dauncey, M. J., Gandy, G. & Gairdner, D. Assessment of total body fat in infancy from skinfold thickness measurements. *Arch. Dis. Child.* **52**, 223–227 (1977).
53. Olhager, E. & Forsum, E. Assessment of total body fat using the skinfold technique in full-term and preterm infants. *Acta Paediatr.* **95**, 21–28 (2006).
54. Petersen, S., Gotfredsen, A. & Ursin Knudsen, F. Lean body mass in small for gestational age and appropriate for gestational age infants. *J. Pediatr.* **113**, 886–889 (1988).
55. Sheng, H. P., Muthappa, P. B., Wong, W. W. & Schanler, R. J. Pitfalls of body fat assessments in premature infants by anthropometry. *Biol. Neonate* **64**, 279–286 (1993).
56. Yumani, D. F. J., de Jongh, D., Lafeber, H. N. & van Weissenbruch, M. M. A comparative study using dual-energy x-ray absorptiometry, air displacement plethysmography, and skinfolds to assess fat mass in preterms at term equivalent age. *Eur. J. Pediatr.* **180**, 919–927 (2021).
57. Tang, W., Ridout, D. & Modi, N. Assessment of total body water using bioelectrical impedance analysis in neonates receiving intensive care. *Arch. Dis. Child. Fetal Neonatal Ed.* **77**, F123–F126 (1997).
58. Ahmad, I. et al. Body composition and its components in preterm and term newborns: a cross-sectional, multimodal investigation. *Am. J. Hum. Biol.* **22**, 69–75 (2010).
59. McLeod, G. et al. Feasibility of using ultrasound to measure preterm body composition and to assess macronutrient influences on tissue accretion rates. *Early Hum. Dev.* **89**, 577–582 (2013).
60. Holzhauer, S. et al. Sonographic assessment of abdominal fat distribution in infancy. *Eur. J. Epidemiol.* **24**, 521–529 (2009).
61. Ward, L. C., Poston, L., Godfrey, K. M. & Koletzko, B. Assessing early growth and adiposity: report from an Early Nutrition Academy Workshop. *Ann. Nutr. Metab.* **63**, 120–130 (2013).
62. International Atomic Energy Agency. *Body Composition Assessment from Birth to Two Years of Age* (IAEA, 2013).
63. Hartnoll, G., Bétrémieux, P. & Modi, N. Body water content of extremely preterm infants at birth. *Arch. Dis. Child. Fetal Neonatal Ed.* **83**, F56–F59 (2000).
64. Brunton, J. A., Weiler, H. A. & Atkinson, S. A. Improvement in the accuracy of dual energy x-ray absorptiometry for whole body and regional analysis of body composition: validation using piglets and methodologic considerations in infants. *Pediatr. Res.* **41**, 590–596 (1997).
65. Harzing, A. W. Publish or perish. <https://harzing.com/resources/publish-or-perish> (2007).

AUTHOR CONTRIBUTIONS

D.F.J.Y.: conceptualization and design, acquisition, analysis and interpretation of data, drafting and revision of the manuscript, final approval of the manuscript. D.d.J.: acquisition, analysis and interpretation of data, drafting and revision of the manuscript, final approval of the manuscript. J.C.F.K.: acquisition and analysis of data, revision of the manuscript, final approval of the manuscript. H.N.L.: conceptualization and design, acquisition, analysis and interpretation of data, revision of the manuscript. M.M.v.W.: conceptualization and design, acquisition, analysis and interpretation of data, revision of the manuscript, final approval of the manuscript.

COMPETING INTERESTS

The authors declare no competing interests.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

Patient consent was not required.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41390-022-02262-x>.

Correspondence and requests for materials should be addressed to Dana F. J. Yumani.

Reprints and permission information is available at <http://www.nature.com/reprints>

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

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.