Review Article

Clinical relevance and validity of tools to predict infant, childhood and adulthood obesity: a systematic review

Oliver J Canfell[1,2,*], Robyn Littlewood[1,2,3], Olivia RL Wright[1] and Jacqueline L Walker[1]

[1]School of Human Movement and Nutrition Sciences, Faculty of Health and Behavioural Sciences, The University of Queensland, Brisbane, QLD 4072, Australia: [2]Lady Cilento Children's Hospital, Children's Health Queensland Hospital and Health Service, Brisbane, Australia: [3]Queensland Child and Youth Clinical Network, Office of Strategy Management, Clinical Excellence Division, Brisbane, Australia

Abstract

Objective: To determine the global availability of a multicomponent tool predicting overweight/obesity in infancy, childhood, adolescence or adulthood; and to compare their predictive validity and clinical relevance.

Design/Setting: The PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines were followed. The databases PubMed, EMBASE, CINAHL, Web of Science and PsycINFO were searched. Additional articles were identified via reference lists of included articles. Risk of bias was assessed using the Academy of Nutrition and Dietetics' Quality Criteria Checklist. The National Health and Medical Research Council's Levels of Evidence hierarchy was used to assess quality of evidence. Predictive performance was evaluated using the ABCD framework.

Subjects: Eligible studies: tool could be administered at any life stage; quantified the risk of overweight/obesity onset; used more than one predictor variable; and reported appropriate prediction statistical outcomes.

Results: Of the initial 4490 articles identified, twelve articles (describing twelve tools) were included. Most tools aimed to predict overweight and/or obesity within childhood (age 2–12 years). Predictive accuracy of tools was consistently adequate; however, the predictive validity of most tools was questioned secondary to poor methodology and statistical reporting. Globally, five tools were developed for dissemination into clinical practice, but no tools were tested within a clinical setting.

Conclusions: To our knowledge, a clinically relevant and highly predictive overweight/obesity prediction tool is yet to be developed. Clinicians can, however, act now to identify the strongest predictors of future overweight/ obesity. Further research is necessary to optimise the predictive strength and clinical applicability of such a tool.

Despite numerous past and ongoing public health prevention and intervention initiatives, childhood and adulthood overweight and obesity (overweight/obesity) remains one of the most significant global public health challenges[1]. The estimated annual global direct economic impact and worldwide investment to mitigate obesity is currently third only to smoking and armed conflict[2]. Within Australia, childhood overweight/obesity occurs in approximately one in four (27·6 %) children aged 5–17 years, while overweight/ obesity is present in 63·4 % of adults[3].

Early identification of overweight/obesity within infants and children is largely contingent on the screening behaviours of primary care clinicians (PCC)[4]. Currently in the USA and Australia, BMI percentile screening for overweight/obesity in children is recommended within primary care (age ≥6 years and age 2–18 years, respectively), its role being the detection of current overweight/obesity[5,6]. Yet, once overweight/ obesity has developed in children and adolescents, the long-term success of intervention strategies is modest[7]. Even with global recognition and recommendation for the use of BMI percentile screening within paediatric primary care, it is performed infrequently and inconsistently[4,8].

Corresponding author: Email oliver.canfell@uqconnect.edu.au

To maximise success in preventive health with currently available tools, the use of BMI as a predictor of future overweight/obesity has been considered[9]. However, BMI has limited clinical efficacy in predicting future overweight/obesity; a large proportion of obese adults do not identify as obese in childhood[9]. Therefore, the utility of BMI alone as an adequate tool to assess the risk of future overweight/obesity and potential co-morbidities is questionable.

Investigation into other anthropometric predictors of childhood and adulthood obesity has resulted in strong evidence for maternal pre-pregnancy weight status[10,11], low[12] and high birth weight[13,14] and rapid infant weight gain[15] influencing obesity outcomes in childhood and adolescence. Additionally, non-anthropometric predictors have been identified, such as maternal smoking during pregnancy[16], breast-feeding duration[17,18] and parental education[19,20]. To maximise the accuracy of overweight/obesity prediction, a combination of child anthropometric, maternal, environmental and socio-economic predictors is likely necessary.

Paediatric PCC are well-positioned to effect change for children and their families; they have the ability to influence weight-related behaviours themselves, or initiate referrals to appropriate health professionals or health services for further intervention if necessary[4]. The ability of PCC to discriminate low-risk and high-risk infants for future overweight/obesity may increase the preventive reach of the health-care system, as early interventions for high-risk infants can be prioritised before overweight/obesity develops.

The aim of the present review, therefore, was to assess the global availability of a prediction tool estimating the risk of developing overweight/obesity at any stage throughout infancy, childhood, adolescence or adulthood, and to compare their predictive validity and relevance to current clinical settings.

For the purposes of the review, all included questionnaires, equations, models, indices, evaluations or tools are collectively referred to as 'tools'.

Methodology

The protocol for the present systematic review was published in PROSPERO prior to commencement (CRD42017071048; available at https://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42017071048).

Search strategy and identification of included articles

The databases PubMed, EMBASE, CINAHL, Web of Science and PsycINFO were searched for articles published since individual database inception until September 2017. Key words were identified via literature scoping, group discussion and consultation with a research librarian. Identified key words were translated into a PICO (Population, Intervention, Comparison and Outcome)[21] format and used as the template in searching for all relevant publications within each database (see online supplementary material, Table S1). The final strategy involved a combination of key words and controlled vocabulary, limited by abstract/title to achieve precise results. Two reviewers (O.J.C. and J.L.W.) independently screened all articles based on title and abstract. Full-text articles were then screened by two reviewers (O.J.C. and J.L.W.) based on the specified inclusion and exclusion criteria, with any discrepancies discussed and resolved by a third reviewer (R.L.). The reference lists of all included articles were also checked to identify additional articles of relevance. The PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses)[22] guidelines were followed in preparing the review (see online supplementary material, Table S2 for the PRISMA checklist and Fig. 1 for the PRISMA flow diagram).

Inclusion and exclusion criteria

Articles were included if they described the development of a prediction tool designed to predict overweight/obesity. Articles were required to meet the following inclusion criteria:

1. Tool was designed to be administered at any stage throughout gestation, infancy, childhood, adolescence and adulthood; and
2. Tool determined the risk of overweight/obesity onset at any stage throughout infancy, childhood, adolescence or adulthood, irrespective of study design; and
3. Tool employed a multicomponent approach, either through:
   a. The use of more than one predictive variable in the final tool; or
   b. If using only one predictive variable, the use of more than one time-point measurement of the same variable; and
4. Article described the risk output, which was accepted to be either of the following:
   a. A percentage likelihood of overweight/obesity development; or
   b. A numerical or qualitative score output specific to the tool; and
5. Article reported statistical analyses pertinent to prediction tool performance: calibration, discrimination and decision-curve analysis, i.e. calibration-in-the-large, calibration slope, goodness-of-fit (P value), receiver operating characteristic (ROC) curve analysis (area under the ROC curve (AUC)), sensitivity, specificity and the tool's decision threshold. Positive predictive value (PPV) and negative predictive value (NPV) were also considered acceptable, as were odds ratio and/or relative risk analyses.
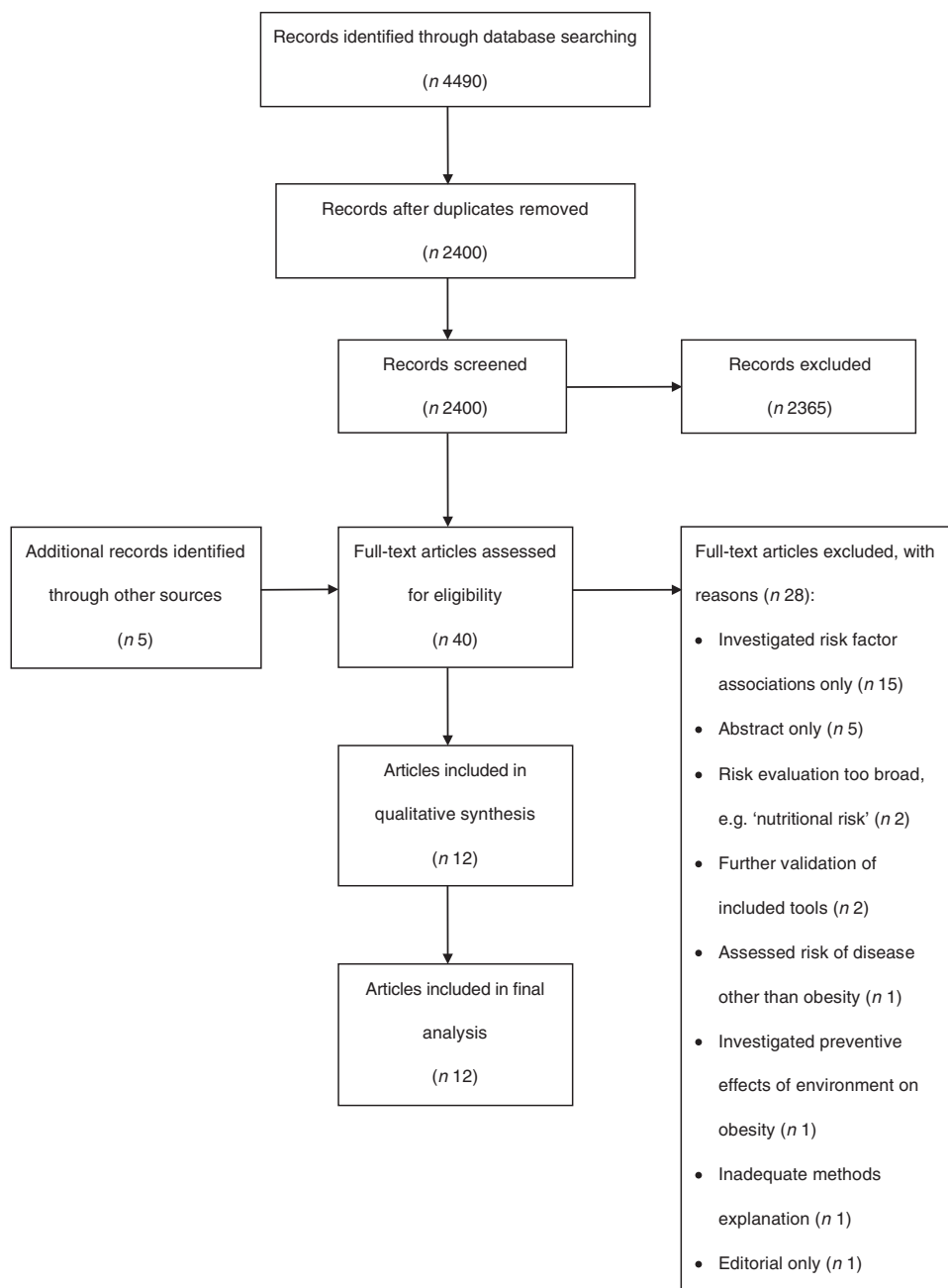
Fig. 1 PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram of search strategy results for the present systematic review of tools to predict infant, childhood and adulthood obesity

Exclusion criteria were any articles not published in English, those not published as full-text articles and any cohorts investigating the outcome within unwell children.

Data extraction and synthesis

The following characteristics of each article were extracted for qualitative synthesis: year of publication, country, study type, cohort description, outcome measure, outcome age and diagnostic criteria. The predictive variables of each tool were also extracted and, for ease of reporting, were classified into one of the groups: anthropometric; socio-economic/sociodemographic; and clinical. Clinical predictive variables were classified as those representing a clinical factor unrelated to anthropometric or socio-economic/sociodemographic elements. The following results of each article were extracted for quantitative synthesis: calibration outcome (e.g. calibration-in-the-large; calibration slope; and/or goodness-of-fit (P value)), AUC (%), sensitivity (%), specificity (%), PPV (%), NPV (%) and the decision threshold of the tool.

Prediction tool performance

The optimal predictive accuracy of a tool depends on the individual outcome being predicted, the costs of misclassification and the clinical pathway each individual will enter after risk assessment[23]. Therefore, there is no known standard definition for statistical predictive outcomes that justify a tool's use in a clinical setting. Instead, prediction tool validity, and thus performance, may be estimated according to the Alpha, Beta, c-statistic and Decision-curve analysis (ABCD) framework, developed by Steyerberg and Vergouwe[24], which considers the following indicators to assess overall prediction tool performance: calibration (calibration-in-the-large; calibration slope); discrimination (AUC); and decision-curve analysis (sensitivity; specificity). Indicators pertinent to the ABCD framework were extracted for each study and compared qualitatively. For the purposes of the present review, strength of discriminative accuracy was determined as: inadequate if AUC < 70%; adequate if AUC = 70–80 %; and excellent if AUC > 80%, as indicated by Lee et al.[25].

Clinical relevance

The clinical relevance or usefulness of a prediction tool can be defined as its ability to generate a decision threshold, or a 'cut-off' point, to classify patients as low-risk or high-risk, to contribute to and enhance clinician decision making[24]; however, the strength of a prediction tool cannot be determined solely by statistical metrics[26] – practical factors also influence a tool's immediate clinical relevance. To our knowledge, there are no available criteria to objectively assess a prediction tool's immediate clinical applicability. Therefore, the authors defined factors relating to clinical relevance as the following:

1. Author mention, within the article, of tool translatability to current practice.
2. Ease of clinician accessing predictor variable information.
3. Defined action and/or referral pathway for individuals identified as 'high-risk'.

Quality assessment and levels of evidence

Full-text articles included in the review were assessed for quality using the Academy of Nutrition and Dietetics' Quality Criteria Checklist for Primary Research, located within the Evidence Analysis Manual (see online supplementary material, Fig. S1)[27]. Articles were designated a positive (+), neutral (Ø) or negative (–) quality rating, according to questions relating to each individual article's relevance, validity and risk of bias (see Table 1).

Each included article was then assessed against the National Health and Medical Research Council's Levels of Evidence hierarchy[28] (see online supplementary material, Fig. S2). The hierarchy classifies evidence into descending levels of strength (Level I, II, III-1, III-2, III-3, IV)

Table 1 Quality rating and level of evidence results for studies included in the present systematic review of tools to predict infant, childhood and adulthood obesity

Study	Quality rating*	Level of evidence†
Classen and Hokayem (2005)[30]	Ø	Level II
de Kroon et al. (2011)[31]	Ø	Level II
Druet et al. (2012)[40]	Ø	Level III-2
Manios et al. (2013)[32]	Ø	Level II
Morandi et al. (2012)[33]	Ø	Level II
Pei et al. (2013)[34]	Ø	Level II
Potter and Ulijaszek (2013)[35]	Ø	Level II
Santorelli et al. (2013)[36]	Ø	Level II
Seyednasrollah et al. (2017)[39]	Ø	Level III-1
Steur et al. (2011)[37]	Ø	Level II
Timpka et al. (2007)[29]	+	Level II
Weng et al. (2013)[38]	Ø	Level II

*Academy of Nutrition and Dietetics' Quality Criteria Checklist for Primary Research[27]: positive (+); neutral (Ø); negative (–).
†National Health and Medical Research Council's Levels of Evidence hierarchy[28]: Level I; II; III-1; III-2; III-3; IV.

based on robustness of the study design for a specific research question (see Table 1). This informs a study's overall level of evidence quality.

Two reviewers (O.J.C. and J.L.W.) independently completed the quality checklist and Levels of Evidence hierarchy for each article and any inconsistencies were discussed with a third reviewer (R.L.) until an agreement was made.

Results

A total of 4490 articles were identified from the initial search. Following removal of duplicates, title and abstract screening was performed on 2400 articles, with forty subsequently full-text screened for eligibility. Of these, twelve articles describing the development of twelve tools were selected for inclusion in the review (see Fig. 1).

Quality rating and level of evidence

Table 1 presents a summary of all quality and level of evidence outcomes for each article. No articles were excluded due to poor quality. Eleven of the twelve included articles received a neutral rating (Ø), with only one article receiving a positive quality rating (+)[29]. Ten included articles were classified as Level II evidence[29–38]. One article was classified as Level III-1 evidence[39], while the study by Druet et al.[40] was classified as Level III-2 evidence.

Description of study cohorts

Table 2 presents the characteristics of included studies. The majority of articles utilised established prospective birth cohort samples (n 9)[29,31,33–38,40] to gather baseline and outcome data, then employed a retrospective approach to develop their tool. The size of cohorts used

varied considerably, but was typically large: >10 000 participants (n 3)[35,38,40]; 5000–10 000 (n 1)[29]; 2000–5000 (n 4)[30,32–34]; 1000–2000 (n 3)[36,37,39]; and <1000 (n 1)[31]. As a supplementary component to their meta-analysis, Druet et al.[40] employed a multivariate logistic regression approach within a sub-sample of cohorts (ALSPAC, SEYCHELLES and CPP; see Table 2) to develop their prediction tool.

Nine of the twelve included articles performed validation of their tool using internal or external methods, or both (see Table 2). Of these, five articles randomised a varying percentage of participants from their development cohort into a validation cohort[29,34,38–40]. Two articles performed validation externally using a different cohort[33,36]. Internal validation using bootstrapping methods was completed for three articles[31,36,37]. One article performed both internal validation using bootstrapping methods (1000 repetitions) and validation with an external cohort[36]. Three articles[30,32,35] did not perform internal or external validation.

Prediction tool design

All tools were designed to be administered post-gestation. Risk estimation within two tools could be performed solely by the administering PCC[29,31], with the remaining ten tools requiring variable input from both the administering PCC and mother[30,32–40].

Outcome age

Outcome age is the age at which the tool predicts the likelihood of overweight/obesity onset (see Table 2). No tools were designed to predict overweight/obesity development in infancy and young childhood (≤2 years). The outcome age of one tool[36] was designated at 2 years; however, the outcome was only 'risk for childhood obesity' at a later, unspecified age. Nine of the included tools estimated the risk of overweight/obesity in childhood (age 2–12 years) across singular ages or age ranges: 3 years[38]; 3–6 years[39]; 7 years[33,40]; ≤8 years[30]; 8 years[37]; 9–12 years[39]; 9–13 years[32]; ≥9 years[30]; 10 years[29,34]; and 11 years[40]. Only two tools estimated risk of overweight/obesity development in adolescence (age 13–18 years), at 15–18 years[39] and 16 years[33]. Similarly, two tools measured risk of overweight/obesity development in adulthood (age ≥18 years), at 23 years[31,35], 33 years[35] and 42 years[35].

Outcome measure

Obesity was assessed as the single outcome in four tools[29,32,39,40] (see Table 2). The combined outcome of overweight and obesity was assessed in seven tools[30,31,33–35,37,38]. One tool investigated the outcome 'risk for childhood obesity'[36], defined as BMI > 91st percentile at 2 years of age in conjunction with a >1 centile band weight Z-score gain between birth and 2 years of age.

Risk output

Three articles adapted the output of their screening tools to present an estimated percentage likelihood risk of future obesity[31–33]. A score output unique to the specific tool was developed within four articles, presented as either numerical or qualitative: Manios et al.[32] (numerical score unique to tool translated into percentage risk of future obesity); Weng et al.[38] (numerical); Santorelli et al.[36] (qualitative – high, medium, low); and Potter and Ulijaszek[35] (qualitative – low-risk, fixed-risk, acute-risk, high-risk). Within six included tools[29,30,34,37,39,40], risk estimation was performed via a statistical equation without a score output.

Predictive variables

The mean number of predictive variables in all included tools was 4·92 (SD 3·06; range 1–13). Substantial variability existed between tool variables that largely depended on the type of baseline and outcome data available within each cohort. For a description of predictive variables included within each tool, refer to Table 3.

Anthropometric

All tools included at least one anthropometric variable, either from the child only (n 3)[29,31,34] or both the child and parent (n 9)[30,32,33,35–40]. The most common included variable was a maternal measure of weight (n 9), either pre-pregnancy (n 2)[32,38] or current (n 7)[30,33,35–37,39,40]. The most common child anthropometric variable included was birth weight (n 8), classified as kilograms[30,33–35,37], SD-score[40], quintiles[38] or Z-scores[36].

Socio-economic/sociodemographic

Seven tools included at least one variable of a socio-economic nature[30,32–34,37–39], while a sociodemographic variable was included within five tools[30,32,36,37,40]. Education level of either the mother[30,32] or both parents[34] was the most common socio-economic variable, as was a measure of income, either per household member[30] or as a family[34,39]. In terms of demographics, five articles included sex as a tool variable, all associating female gender with a greater risk of overweight/obesity development[30,32,36,37,40].

Clinical

Smoking status was the most commonly included clinical predictive variable (n 5), defined as maternal smoking during pregnancy (n 4)[32–34,38], maternal smoking (n 1)[33] or smoking in the parental house (n 1)[37]. Breast-feeding history was included as a variable within two tools, categorised as either 'ever breast-fed in first year' (n 1)[38] or 'breast-fed partially or wholly' (n 1)[35].

Two tools[33,39] attempted to enhance the predictive accuracy of anthropometric and socio-economic variables

Table 2 Characteristics of studies included in the present systematic review of tools to predict infant, childhood and adulthood obesity

Study	Country	Study type	Cohort description (study title and participants*)		Outcome measure	Outcome age (years)	Diagnostic criteria
			Development	Validation			
Classen and Hokayem (2005)[30]	USA	Prospective cohort	NLSY79 n 4980	–	Overweight and obesity	≤8 and ≥9	CDC
de Kroon et al. (2011)[31]	Netherlands	Prospective birth cohort	Terneuzen Birth Cohort n 762	Bootstrapping (internal) 1000 repetitions	Overweight and obesity	23	BMI (Dutch)
Druet et al. (2012)[40]	UK	Meta-analysis	ALSPAC (n 1976) MRC HCS (n 2981) MRC NSHD (n 3341) HAGUENAU (n 11 812) SWEDES (n 248) EUROAM (n 653) CPP (n 19 355) PBPP (n 300) SEYCHELLES (n 5406) Total: n 47 661	ALSPAC (n 1976) CPP (n 19 355) SEYCHELLES (n 5406) Sample 1: n 8236 Sample 2: n 8236	Obesity	7 and 11	IOTF
Manios et al. (2013)[32]	Greece	Cross-sectional	Healthy Growth Study n 2294	–	Obesity	9–13	IOTF
Morandi et al. (2012)[33]	France	Prospective birth cohort	NFBC1986 n 4032	1. Veneto Retrospective cohort n 1503 2. Project Viva Prospective cohort n 1032	Overweight and obesity	7–16	IOTF
Pei et al. (2013)[34]	Germany	Prospective birth cohort	GINIplus and LISAplus Total: n 2272 Training data set: n 1515	GINIplus and LISAplus Validation data set: n 757	Overweight and obesity	10	WHO
Potter and Ulijaszek (2013)[35]	UK	Prospective birth cohort	NCDS n 11 752	–	Overweight and obesity	23, 33 and 42	WHO
Santorelli et al. (2013)[36]	UK	Prospective birth cohort	Born in Bradford (BiB) n 1868	Bootstrapping (internal) 1000 repetitions ALSPAC Prospective birth cohort n 880 (Equation 2) n 867 (Equation 3)	Risk for childhood obesity	2	UK90
Seyednasrollah et al. (2017)[39]	Finland	Prospective cohort	Cardiovascular Risk in Young Finns n 1625	Cardiovascular Risk in Young Finns n 637	Obesity	3–6, 9–12 and 15–18	–
Steur et al. (2011)[37]	Netherlands	Prospective birth cohort	PIAMA n 1687	Bootstrapping (internal) 200 repetitions	Overweight and obesity	8	IOTF
Timpka et al. (2007)[29]	Sweden	Prospective birth cohort	Children are important, Swedish only Total: n 5105 Derivation cohort: n 2234	Children are important, Swedish only Validation cohort: n 2235	Obesity	10	IOTF
Weng et al. (2013)[38]	UK	Prospective birth cohort	Millennium Cohort Study n 13 513	Millennium Cohort Study n 1715	Overweight and obesity	3	IOTF

NLSY79, National Longitudinal Survey of Youth; ALSPAC, Avon Longitudinal Study of Parents and Children; MRC HCS, Medical Research Council Hertfordshire Cohort Study; MRC NSHD, Medical Research Council National Survey of Health and Development; HAGUENAU, Haguenau Case–Control Study; SWEDES, Stockholm Weight Development Study; EUROAM, Formula-fed European Americans Cohort Study; CPP, Collaborative Perinatal Project; PBPP, Philadelphia Blood Pressure Project; SEYCHELLES, Seychelles Cohort; NFBC1986, Northern Finland Birth Cohort; GINIplus, German INfant Nutritional Intervention plus environmental and genetic influences on allergy development; LISAplus, Influences of LIfestyle-Related Factors on the Immune System and the Development of Allergies in Childhood plus Air Pollution and Genetics; NCDS, National Child Development Study; PIAMA, Prevention and Incidence of Asthma and Mite Allergy; CDC, Centers for Disease Control and Prevention; IOTF, International Obesity Task Force.
* Total number (n) of participants in each study indicates those who participated in the final development of the tool, rather than the initial cohort size.

by including genetic variables associated with obesity. Seyednasrollah et al.[39] derived genetic variables from a genome-wide association study, using significant SNP (n 19) to create a weighted genetic risk score (WGRS19) that was included within the final predictive tool. Morandi et al.[33] selected thirty-nine SNP related to obesity, creating a cumulative genotype score by summing the number of risk alleles.

Table 3 Description of predictive variables within each tool included in the present systematic review of tools to predict infant, childhood and adulthood obesity

Study	Total no. of variables	Anthropometric	Socio-economic/sociodemographic	Clinical
			Predictive variables	
Classen and Hokayem (2005)[30]	13	Maternal BMI; birth weight (kg)	Full-time v. at-home mothers; mother's marital status; number of household members; income per household member; health insurance status; maternal educational level; ethnicity; gender	Order of birth among siblings; mother's age at time of birth; indications of child and youth depression
de Kroon et al. (2011)[31]	1	BMI SD-score at ages 2, 4 and 6 years	–	–
Druet et al. (2012)[40]	4	Weight SD-score gain (0–1 year); birth weight SD-score; maternal BMI	Gender	–
Manios et al. (2013)[32]	4	Maternal pre-pregnancy BMI; weight change (0–6 months)	Gender; maternal educational level	Maternal smoking during pregnancy
Morandi et al. (2012)[33]	6	Parental BMI; birth weight (kg); maternal GWG	Number of household members; maternal occupation	Smoking habits (maternal smoking or gestational smoking)
Pei et al. (2013)[34]	5	Birth weight (kg); BMI (age 60–64 months)	Parental educational level; family income	Maternal smoking during pregnancy
Potter and Ulijaszek (2013)[35]	4	Birth weight (kg); parental BMI (at 11 years old)	–	Breast-feeding partially or wholly; physical activity (at 11 years old)
Santorelli et al. (2013)[36]	5	Birth weight Z-score; weight Z-score change (birth to screening age); maternal BMI	Female sex	Age (determines which equation is used)
Seyednasrollah et al. (2017)[39]	4	BMI SD-score; maternal BMI	Family income	WGRS97 or WGRS19 (genetic model only)
Steur et al. (2011)[37]	6	Paternal BMI; maternal BMI; birth weight (kg)	Female sex	Smoking in parental house; hospital delivery
Timpka et al. (2007)[29]	1	Boys: obesity at 4- and 5-years check-ups; girls: BMI above 20 kg/m^2 at 5-year check-up	–	–
Weng et al. (2013)[38]	6	Birth weight (quintile); weight change (0–1 year); maternal pre-pregnancy BMI; paternal BMI	–	Maternal smoking during pregnancy; ever breast-fed infant in first year

SD-score, standard deviation score; GWG, gestational weight gain; WGRS, weighted genetic risk score.

Prediction tool performance

Table 4 details the predictive statistical outcomes of each included tool. Only two studies[33,37] provided all three metrics (calibration, discrimination and decision-curve analysis) pertinent to assessing predictive performance, as described by the ABCD framework[24]. Discrimination was more frequently reported, with nine out of twelve studies measuring AUC[30–33,36–40]. For decision-curve analysis, ten out of twelve studies reported sensitivity and specificity values[29–34,36–38,40]. Eight of twelve studies included PPV and/or NPV statistical outcomes[29,31,33,34,36–38,40]. Potter and Ulijaszek[35] did not perform any statistical analyses related to predictive accuracy, instead calculating odds ratios and relative risks for the development of adulthood overweight/ obesity within each risk group (low-risk; fixed-risk; acute-risk; high-risk). Due to the quantity of calculations and their heterogeneity to the other included articles, the results are not presented here.

Calibration

The two studies[33,37] that reported calibration outcomes used the Hosmer–Lemeshow test for goodness-of-fit. Similar values were reported for Morandi et al.[33] and Steur et al.[37], at $P = 0.33$ and $P = 0.3$, respectively. This indicates that their models' predictions of weight outcome were consistently correct[41]. No studies reported the calibration outcomes of calibration-in-the-large or calibration slope as per the ABCD framework[24].

Discrimination

Only three tools achieved an excellent[25] AUC outcome or higher: Santorelli et al.[36], at 85·8 % (Equation 1), 86·1 % (Equation 2) and 91·1 % (Equation 3); Morandi et al.[33] (obesity prediction tool only), at 85 %; and de Kroon et al.[31] (boys aged 2–6 years only, predicting to age 23 years), at 83 %. As an example, in practice this means the tool of Morandi et al.[33] can accurately predict childhood obesity, at age 7–16 years, 85 % of the time. The highest discriminative accuracy outcome achieved for five of nine tools reporting AUC was adequate[30,37–40], with only Manios et al.[32] demonstrating inadequate discriminative accuracy (63·8 %) for their tool predicting obesity at age 9–13 years.

In tools that assessed discriminative accuracy outcomes for both obesity and overweight and obesity, the outcome of obesity was associated with stronger predictive

Table 4 Predictive statistical outcomes, according to the ABCD framework*, of each tool included in the present systematic review of tools to predict infant, childhood and adulthood obesity

Study	Calibration	Discrimination	Decision-curve analysis				
	Goodness-of-fit (*P* value)	AUC (%)†	Sensitivity (%)	Specificity (%)	Decision threshold‡	PPV (%)	NPV (%)
Classen and Hokayem (2005)[30]	–	Obesity; overweight: 70; 65·7 (basic child) 73; 68·4 (full child) 72·9; 68·5 (full youth)	58·9 (basic child) 60·6 (full child) 62·9 (full youth)	31·4 (basic child) 31·7 (full child) 34·4 (full youth)	15·70 %	–	–
de Kroon *et al.* (2011)[31]§	–	Boys¶: 79 (2–4 years) 83 (2–6 years) Girls¶: 76 (2–4 years) 79 (2–6 years)	28 (2–4 years) 38 (4–6 years) 36 (2–6 years)	92 (2–4 years) 93 (4–6 years) 93 (2–6 years)	50 %	58 (2–4 years) 66 (4–6 years) 67 (2–6 years)	–
Druet *et al.* (2012)[40]§	–	77	70·1	71·2	≥30 %	6·9	98·7
Manios *et al.* (2013)[32]	–	63·8 (95 % CI 60·2, 67·5)	54	65	≥5	–	–
Morandi *et al.* (2012)[33]§	0·33	Obesity: 85 Overweight/obesity: 75	Obesity: 79 Overweight/obesity: 63	Obesity: 75·5 Overweight/obesity: 78	75th percentile	Obesity: 4 Overweight/obesity: 21	Obesity: 99·5 Overweight/obesity: 96
Pei *et al.* (2013)[34]§	–	–	37·1	95·1	–	72·5	86·1
Potter and Ulijaszek (2013)[35]	–	–	–	–	–	–	–
Santorelli *et al.* (2013)[36]§,‖	–	85·8 (Equation 1) 86·1 (Equation 2) 91·1 (Equation 3)	71·1 (Equation 1) 69·4 (Equation 2) 77·0 (Equation 3)	84·5 (Equation 1) 84·2 (Equation 2) 85·2 (Equation 3)	≥20 %	28·8 (Equation 1) 27·5 (Equation 2) 32·2 (Equation 3)	97·1 (Equation 1) 97·0 (Equation 2) 97·6 (Equation 3)
Seyednasrollah *et al.* (2017)[39]§	–	WGRS19 & clinical factors: 76·9 WGRS94 & clinical factors: 74·9 Clinical factors: 74·7	–	–	–	–	–
Steur *et al.* (2011)[37]§	0·3	75·4	67·2	74·9	≥15 %	30·2	93·4
Timpka *et al.* (2007)[29]§	–	–	21 (boys) 41 (girls)	100 (boys) 100 (girls)	–	80 (boys) 82 (girls)	–
Weng *et al.* (2013)[38]§	–	75·5	76·9	66·5	≥25	37	89

AUC, area under the receiver operating characteristic curve; PPV, positive predictive value; NPV, negative predictive value; WGRS, weighted genetic risk score.
*If internal or external tool validation was performed, only these quantitative outcomes have been reported, rather than also including the quantitative outcomes of the tool's development phase. If validation was not performed, the quantitative values in Table 4 are outcomes of the tool's development phase only.
†Strength of discriminative accuracy was determined as: inadequate if AUC < 70 %; adequate if AUC = 70–80 %; and excellent if AUC > 80 %.
‡Where predictive statistical analyses had been performed using multiple decision thresholds for each tool, quantitative data were extracted from the median cut-off point only and rounded up where appropriate. Where the author has not designated a decision threshold, this cell has been left blank.
§Tool validation completed.
‖Tool is administered at ages: Equation 1 at 6 (sd 1·5) months; Equation 2 at 9 (sd 1·5) months; Equation 3 at 12 (sd 1·5) months.
¶Tool is administered at the upper age threshold.

performance. Morandi *et al.*[33] achieved an excellent AUC value (85 %) for their obesity prediction tool, but an adequate AUC value for their overweight and obesity tool (75 %). Classen and Hokayem[30] demonstrated adequate discriminative accuracy for their obesity prediction tool, but inadequate discriminative accuracy for their overweight tool. The excellent AUC value of de Kroon *et al.*'s[31] model (83 %) (boys aged 2–6 years, predicting to age 23 years) could not be replicated among their other prediction models, with all others achieving adequate AUC values.

Decision-curve analysis
Morandi *et al.*[33], Santorelli *et al.*[36] and Weng *et al.*[38] reported the highest sensitivity values of all tools at 79 % (obesity tool), 77 % (Equation 3) and 76·9 %, respectively. Again, using Morandi *et al.*[33] as an example, their tool would predict childhood obesity, at age 7–16 years,

correctly in 79 % of children initially assessed at baseline. The corresponding specificity for each of these tools was similarly strong for Morandi *et al.*[33] (75·5 %) and Santorelli *et al.*[36] (85·2 %), while not as strong for Weng *et al.*[38] (66·5 %). In the tool of Morandi *et al.*[33], 75·5 % of children assessed at baseline would be correctly predicted to not develop childhood obesity. Of the ten studies reporting decision-curve analysis, two did not report a corresponding decision threshold related to the tool's sensitivity and specificity values[29,34], meaning it was unclear how the sensitivity and specificity values would be prioritised to inform clinical decisions.

Clinical relevance
The immediate translatability of the developed prediction tool to clinical practice was mentioned within five

studies[31–33,36,37]. Within three of these[31,32,37], a risk graph or chart was developed for clinical use to provide an output of estimated risk. Morandi et al.[33] developed a spreadsheet-based risk calculator with an output of percentage estimated risk. Santorelli et al.[36] integrated their prediction tool within an app, 'Healthy Infant Weight?'; however, at the time of review, this app was not available for download.* Those studies that did not mention tool translatability to clinical practice developed a statistical prediction equation only[29,30,34,35,38–40].

Predictor variable data collection was achieved by combined self-report and medical record access for five prediction tools[32–34,38,40]. Methods of only self-report and medical record access were appropriate within two prediction tools[35,37] and one prediction tool[36], respectively. Two prediction tools required only anthropometric measurement of the individual[29,31]. One prediction tool required the collection of a genetic variable in the form of obesity-related SNP[39]. The tool of Classen and Hokayem[30] included an indication of child depression and therefore required a clinical assessment component.

Only one prediction tool[36] provided a component beyond risk prediction, with the addition of relevant educational material and resource links for those identified as high-risk. No tool described a relevant referral pathway, to either a tertiary or primary care health service, or a community-based preventive or treatment programme, for individuals identified as high-risk.

Discussion

The current systematic review provides a comprehensive summary of known available tools designed to predict overweight/obesity via an estimation of risk. Twelve tools were developed within specific populations, with nine[29,31,33,34,36–40] of these having been validated either using internal or external methods, or a combination of both. The prediction tools developed and validated by Morandi et al.[33] and Steur et al.[37] were assessed as the strongest overall, secondary to: strong congruency with the ABCD framework for assessing prediction tool validity; adequate–excellent discriminative accuracy; internal[33,37] and external[33] tool validation; and relevance to current clinical practice.

Prediction tool performance

Assessing the validity of a prediction tool does not solely rest with its applicability to multiple settings. Robust reporting of outcomes pertinent to predictive performance is crucial in the development of any prediction tool. Following the ABCD framework[24] for assessing prediction

tool validity, only two studies[33,37] were identified as having reported all outcomes necessary: calibration; discrimination; and decision-curve analysis. Despite this, the calibration metric described within these two studies has been demonstrated as inappropriate: the Hosmer–Lemeshow goodness-of-fit test is not recommended as P value outputs can vary by sample size and are thus deemed imprecise[24]. Instead, calibration-in-the-large and assessment of the calibration slope is recommended[24]. Poor methodological considerations and inadequate reporting of all measures necessary to assess prediction tool performance are consistent across multiple diseases and settings, not just overweight/obesity, as identified in a recent systematic review[42].

Discriminative accuracy (AUC %) was the most commonly reported statistical outcome relevant to predictive performance. Excellent discriminative accuracy is expected in tools which predict overweight/obesity as a proximal outcome. Equation 3, a prediction tool of Santorelli et al.[36] designed to be administered at 12 (SD 1·5) months to predict the risk of childhood obesity at 2 years of age, achieved the highest AUC value at 91·1%. Notably, only the outcome of 'risk for childhood obesity' is associated with this prediction tool, rather than the actual outcome of obesity. No follow-up anthropometric data were collected at a later time point in childhood and no age range was specified for 'childhood obesity'. Similarly, in their prediction of persistent childhood obesity (i.e. obesity at ages 7 and 16 years), Morandi et al.[33] produced an excellent AUC outcome (85%), which is also expected considering the high likelihood of childhood obesity tracking into adolescence[14].

In its mere classification of individuals who develop overweight/obesity compared with those who do not, discriminative accuracy is not a true indicator of clinical usefulness. When determining the overall goal of the tool, consequences secondary to false negative and false positive predictions also need to be considered[43]. Designation of the decision threshold (cut-off point) is a subjective decision at the discretion of the end user, informed by the priority assigned to high or low sensitivity and specificity values. Consideration needs to be made for the tool's overall impact on health service delivery; the decision threshold should be assigned based on characteristics of the health system in which the predictive tool is to be embedded[24]. For example, Timpka et al.[29] prioritised strong specificity in order to avoid unnecessary interventions, reducing health system burden and cost. Other end users may prioritise sensitivity to maximise the number of children identified who will go on to develop overweight/obesity and thus be highly eligible for preventive intervention. This approach, however, may be feasible only for health systems with access to an evidence-based preventive intervention.

Considering the low prevalence of weight discussion and overweight/obesity diagnosis in a paediatric clinical setting[44], the availability of a structured, user-friendly and

---

* At the time of submitting the revised manuscript, the app 'Healthy Infant Weight?' was not available for download.

outcomes-based platform to a clinician may also simply encourage the discussion of weight with families and children. This action in itself improves clinical service delivery and may lead to appropriate intervention. Prior to the development of a clinically relevant prediction tool, clinicians can also begin to identify the strongest predictors for overweight/obesity as part of routine clinical practice. Predictors most commonly included in the tools of the present review, such as maternal overweight/obesity, high birth weight, low parental education and maternal smoking during pregnancy, can be assessed by the clinician to inform an overall, subjective risk assessment to influence preventive clinical decisions.

Two articles[33,39] attempted to enhance their prediction tool with the addition of genetic risk factors derived from a genome-wide association study. The addition of a cumulative genetic score based on thirty-nine obesity predisposing SNP did not improve discriminative accuracy of Morandi et al.'s[33] tool compared with using traditional risk factors alone and was excluded as a variable in the final tool. Conversely, Seyednasrollah et al.[39] reported that inclusion of a weighted genetic risk score based on nineteen SNP slightly improved the predictive accuracy of the tool compared with clinical factors alone (AUC 76·9 % v. 74·7 %) in both the development and validation cohorts for children aged 3–6 years. Globally, only limited added clinical value has been reported for the use of genetic factors in accurately predicting adulthood obesity[45]. Potential exists for genetic factors to add strength to prediction tools; however, the clinical utility of such tools is questionable, considering the timely process involved with obtaining, securing and analysing genetic data.

Clinical relevance

For the successful dissemination of a health service innovation within a clinical setting, a combination of quantitative and qualitative development, implementation and evaluation processes is necessary[46]. Five prediction tools[31–33,36,37] included within the review had been modified with the intention of integration into clinical practice, via a risk graph/chart, spreadsheet-based calculator or app. Despite this, there had been no known attempt by the authors to progress to widespread clinical uptake within the health-care sector. It remained unclear how the prediction tools may integrate with population electronic health systems, PCC practice management software or the Internet. While these tools could theoretically be used within current practice, issues remain, including: perceived usability; clinician awareness and uptake; clinician access to required child and maternal anthropometric data; potential need to upskill existing PCC and train new PCC; the effect on clinician behaviour change, including the tool's influence on clinical decision

making; as well as actual overweight/obesity outcomes for all children exposed to the tool.

In developing a clinically valid overweight/obesity prediction tool, the most appropriate outcome age for preventive success needs to be determined. Prediction during adolescence may be a suitable medium between childhood and adulthood. Within the analysis of their large-sampled longitudinal cohort (n 17 638), Potter and Ulijaszek[35] identified that a large proportion of measured overweight/obese adults did not present as overweight/ obese in childhood (aged 11 years). This is consistent with findings of a recent systematic review and meta-analysis[9], demonstrating that childhood weight status is a poor predictor of adulthood weight status: 70 % of obese adults were not obese as children. There is considerable evidence, however, supporting the association between adolescent overweight/obesity and adulthood overweight/obesity[47–50]. Additionally, adolescent overweight/obesity is strongly and consistently associated with adulthood morbidity, including CVD[51,52] and type 2 diabetes mellitus[51,53]. The outcome of adolescent overweight/obesity may be proximal enough to ensure adequate predictive accuracy, yet distal enough to encompass prediction of adulthood obesity and chronic disease in certain cases.

Prediction tool development needs to consider the optimal time point for tool administration. Tools that require the input of predictor variables in late childhood, such as those by de Kroon et al.[31] (age 6 years) and Potter and Ulijaszek[35] (age 11 years), arguably provide an estimation of risk that is too late to elicit a preventive benefit. Modifiable predictors within infancy, such as prolonged breast-feeding exclusivity and duration[17,54,55], as well as appropriate timing of introduction to solids (approximately 6 months of age)[56,57], have been demonstrated as protective against future overweight/ obesity. Additionally, the ability of a prediction tool to be administered at multiple time points is highly desirable, such as the tool of Santorelli et al.[36]. The variability of health service visits in the early years is high between individuals. It is unrealistic to expect all infants or young children to be assessed for overweight/obesity risk at exactly the same time point[36]. A prediction tool validated only for administration at one time point may miss high-risk individuals who would have benefited from preventive intervention.

The concurrent availability of a cost-effective, evidence-based preventive intervention and/or referral pathway for individuals identified as high-risk is critical for a clinically embedded obesity prediction tool. Santorelli et al.[36] developed the only prediction tool to incorporate a component beyond risk prediction, with the inclusion of generic, evidence-based recommendations for healthy eating, physical activity and positive parenting behaviours within their app. To embed an electronic tool within clinical practice creates capacity to integrate individualised

recommendations based on risk output. Potential exists for further risk-determined customisation, perhaps providing a referral pathway for the clinician to action, a recommendation to emphasise positive early feeding habits or the encouragement of frequent anthropometric monitoring by the clinician.

Strengths and limitations of included articles

A number of limitations were seen as trends across all included articles. All studies were conducted in high-income countries. Multiple articles were restricted in their choice of tool variables. Depending on the cohort used to inform the prediction tool, only certain baseline variables and outcome measures were accessible, potentially limiting the accuracy and scope of the tool. None of the included articles measured body composition as indicators of overweight/obesity outcomes. In addition, there was considerable heterogeneity in each study's justification for including its predictive variables, with most only extracting the strongest variables observed within a single cohort[29,32,33,36,39] and some providing no justification[30,31]. A strength of all included articles was the large sample sizes commonly used within both the development and validation phases.

Limitations of review

The principal limitation of the present review was the inability to perform a quantitative synthesis of the tools' predictive strength. The heterogeneity observed in each tool's outcome measure, diagnostic criteria used, outcome age and chosen methods of statistical analysis suited a qualitative approach. While no articles were excluded secondary to a low quality rating, only one included article achieved a high quality rating, potentially subjecting the review to bias. Additionally, the review focused solely on studies detailing the development and validation of prediction tools and thus excluded any follow-up validation performed.

Conclusions

The current systematic review is the first of its kind aiming to determine the availability, clinical relevance and validity of tools to predict overweight/obesity development across the life course. Our findings indicate that further research should prioritise the following: optimising the predictive strength of currently available, population-specific tools; improving the immediate clinical applicability of existing prediction tools; and the methodologically robust development of new prediction tools in novel countries and settings. Following, investigation into the use of a validated prediction tool concurrently with an overweight/obesity preventive intervention for subjects identified as high-risk could be considered.

Clinicians can act now to identify strong predictors, in infancy, for future overweight/obesity, such as maternal overweight/obesity and high birth weight. The successful integration of an overweight/obesity prediction tool within clinical practice may be a valid front-line preventive strategy to assist in decreasing overweight/obesity prevalence.

Supplementary material

To view supplementary material for this article, please visit https://doi.org/10.1017/S1368980018001684

References

1. World Health Organization (2015) *Global Status Report on Noncommunicable Diseases 2014*. Geneva: WHO.
2. Dobbs R, Sawers C, Thompson F *et al.* (2014) *Overcoming Obesity: An Initial Economic Analysis*. New York: McKinsey Global Institute.
3. Australian Bureau of Statistics (2017) *National Health Survey: First Results, 2014–15*. Canberra, ACT: ABS; available at http://www.abs.gov.au/ausstats/abs@.nsf/Lookup/by%20 Subject/4364.0.55.001 ~ 2014-15 ~ Main%20Features ~ Over weight%20and%20obesity ~ 22
4. Klein JD, Sesselberg TS, Johnson MS *et al.* (2010) Adoption of body mass index guidelines for screening and counseling in pediatric practice. *Pediatrics* **125**, 2008–2985.
5. Grossman DC, Bibbins-Domingo K, Curry SJ *et al.* (2017) Screening for obesity in children and adolescents: US Preventive Services Task Force recommendation statement. *JAMA* **317**, 2417–2426.
6. National Health and Medical Research Council (2013) *Clinical Practice Guidelines for the Management of Overweight and Obesity in Adults, Adolescents and Children in Australia*. Melbourne, VIC: NHMRC.
7. Hassink SG (2017) Mobilizing during infancy to prevent severe obesity. *J Pediatr* **183**, 6–7.
8. Barlow SE, Bobra SR, Elliott MB *et al.* (2007) Recognition of childhood overweight during health supervision visits: does BMI help pediatricians? *Obesity (Silver Spring)* **15**, 225–232.

9. Simmonds M, Burch J, Llewellyn A *et al.* (2015) The use of measures of obesity in childhood for predicting obesity and the development of obesity-related diseases in adulthood: a systematic review and meta-analysis. *Health Technol Assess* issue 43, 1–336.

10. Zalbahar N, Najman J, McIntyre H *et al.* (2017) Parental pre-pregnancy obesity and the risk of offspring weight and body mass index change from childhood to adulthood. *Clin Obes* **7**, 206–215.

11. Rath S, Marsh J, Newnham J *et al.* (2016) Parental pre-pregnancy BMI is a dominant early-life risk factor influencing BMI of offspring in adulthood. *Obes Sci Pract* **2**, 48–57.

12. Hales C & Barker D (2001) The thrifty phenotype hypothesis. *Br Med Bull* **60**, 5–20.

13. Yu Z, Han S, Zhu G *et al.* (2011) Birth weight and subsequent risk of obesity: a systematic review and meta-analysis. *Obes Rev* **12**, 525–542.

14. Evensen E, Emaus N, Kokkvoll A *et al.* (2017) The relation between birthweight, childhood body mass index, and overweight and obesity in late adolescence: a longitudinal cohort study from Norway, The Tromsø Study, Fit Futures. *BMJ Open* **7**, e015576.

15. Ong KK & Loos RJ (2006) Rapid infancy weight gain and subsequent obesity: systematic reviews and hopeful suggestions. *Acta Paediatr* **95**, 904–908.

16. Ino T (2010) Maternal smoking during pregnancy and offspring obesity: meta-analysis. *Pediatr Int* **52**, 94–99.

17. Horta BL, Loret de Mola C & Victora CG (2015) Long-term consequences of breastfeeding on cholesterol, obesity, systolic blood pressure and type 2 diabetes: a systematic review and meta-analysis. *Acta Paediatr* **104**, 30–37.

18. Scott JA, Ng SY & Cobiac L (2012) The relationship between breastfeeding and weight status in a national sample of Australian children and adolescents. *BMC Public Health* **12**, 107.

19. Birbilis M, Moschonis G, Mougios V *et al.* (2012) Obesity in adolescence is associated with perinatal risk factors, parental BMI and sociodemographic characteristics. *Eur J Clin Nutr* **67**, 115–121.

20. Parrino C, Vinciguerra F, La Spina N *et al.* (2016) Influence of early-life and parental factors on childhood overweight and obesity. *J Endocrinol Invest* **39**, 1315–1321.

21. Aslam S & Emmanuel P (2010) Formulating a researchable question: a critical step for facilitating good clinical research. *Indian J Sex Transm Dis* **31**, 47–50.

22. Moher D, Liberati A, Tetzlaff J *et al.* (2009) Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med* **6**, e1000097.

23. Eusebi P (2013) Diagnostic accuracy measures. *Cerebrovasc Dis* **36**, 267–272.

24. Steyerberg EW & Vergouwe Y (2014) Towards better clinical prediction models: seven steps for development and an ABCD for validation. *Eur Heart J* **35**, 1925–1931.

25. Lee Y-H, Bang H & Kim DJ (2016) How to establish clinical prediction models. *Endocrinol Metab (Seoul)* **31**, 38–44.

26. Dent THS, Wright CF, Stephan BCM *et al.* (2012) Risk prediction models: a framework for assessment. *Public Health Genomics* **15**, 98–105.

27. Academy of Nutrition and Dietetics (2016) *Evidence Analysis Manual: Steps in the Academy Evidence Analysis Process*. Chicago, IL: Academy of Nutrition and Dietetics.

28. National Health and Medical Research Council (2009) *NHMRC Additional Levels of Evidence and Grades for Recommendation for Developers of Guidelines*. Canberra, ACT: NHMRC.

29. Timpka T, Angbratt M, Bolme P *et al.* (2007) A high-precision protocol for identification of preschool children at risk for persisting obesity. *PLoS One* **2**, e535.

30. Classen T & Hokayem C (2005) Childhood influences on youth obesity. *Econ Hum Biol* **3**, 165–187.

31. de Kroon ML, Renders CM, van Wouwe JP *et al.* (2011) Identifying young children without overweight at high risk for adult overweight: the Terneuzen Birth Cohort. *Int J Pediatr Obes* **6**, e187–e195.

32. Manios Y, Birbilis M, Moschonis G *et al.* (2013) Childhood Obesity Risk Evaluation based on perinatal factors and family sociodemographic characteristics: CORE index. *Eur J Pediatr* **172**, 551–555.

33. Morandi A, Meyre D, Lobbens S *et al.* (2012) Estimation of newborn risk for child or adolescent obesity: lessons from longitudinal birth cohorts. *PLoS One* **7**, e49919.

34. Pei Z, Flexeder C, Fuertes E *et al.* (2013) Early life risk factors of being overweight at 10 years of age: results of the German birth cohorts GINIplus and LISAplus. *Eur J Clin Nutr* **67**, 855–862.

35. Potter CM & Ulijaszek SJ (2013) Predicting adult obesity from measures in earlier life. *J Epidemiol Community Health* **67**, 1032–1037.

36. Santorelli G, Petherick ES, Wright J *et al.* (2013) Developing prediction equations and a mobile phone application to identify infants at risk of obesity. *PLoS One* **8**, e71183.

37. Steur M, Smit HA, Schipper CM *et al.* (2011) Predicting the risk of newborn children to become overweight later in childhood: the PIAMA birth cohort study. *Int J Pediatr Obes* **6**, e170–e178.

38. Weng SF, Redsell SA, Nathan D *et al.* (2013) Estimating overweight risk in childhood from predictors during infancy. *Pediatrics* **132**, e414–e421.

39. Seyednasrollah F, Makela J, Pitkanen N *et al.* (2017) Prediction of adulthood obesity using genetic and childhood clinical risk factors in the Cardiovascular Risk in Young Finns study. *Circ Cardiovasc Genet* **10**, e001554.

40. Druet C, Stettler N, Sharp S *et al.* (2012) Prediction of childhood obesity by infancy weight gain: an individual-level meta-analysis. *Paediatr Perinat Epidemiol* **26**, 19–26.

41. Hosmer DW, Lemeshow S & Sturdivant RX (2013) *Applied Logistic Regression*, 3rd ed. Hoboken, NJ: Wiley.

42. Bouwmeester W, Zuithoff NPA, Mallett S *et al.* (2012) Reporting and methods in clinical prediction research: a systematic review. *PLoS Med* **9**, e1001221.

43. Mallett S, Halligan S, Thompson M *et al.* (2012) Interpreting diagnostic accuracy studies for patient care. *BMJ* **345**, e3999.

44. Campbell M, Bryson H & Wake M (2010) The burden of childhood obesity in Australian paediatric practice: children attending paediatricians survey. *J Paediatr Child Health* **46**, 1–2.

45. Juonala M, Juhola J, Magnussen CG *et al.* (2011) Childhood environmental and genetic predictors of adulthood obesity: the Cardiovascular Risk in Young Finns study. *J Clin Endocrinol Metab* **96**, E1542–E1549.

46. Damschroder LJ, Aron DC, Keith RE *et al.* (2009) Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. *Implement Sci* **4**, 50.

47. Guo SS, Huang C, Maynard LM *et al.* (2000) Body mass index during childhood, adolescence and young adulthood in relation to adult overweight and adiposity: the Fels Longitudinal Study. *Int J Obes Relat Metab Disord* **24**, 1628–1635.

48. Kvaavik E, Tell GS & Klepp KI (2003) Predictors and tracking of body mass index from adolescence into adulthood: follow-up of 18 to 20 years in the Oslo Youth Study. *Arch Pediatr Adolesc Med* **157**, 1212–1218.

49. The NS, Suchindran C, North KE *et al.* (2010) Association of adolescent obesity with risk of severe obesity in adulthood. *JAMA* **304**, 2042–2047.

50. Wang LY, Chyen D, Lee S *et al.* (2008) The association between body mass index in adolescence and obesity in adulthood. *J Adolesc Health* **42**, 512–518.

51. Llewellyn A, Simmonds M, Owen CG *et al.* (2016) Childhood obesity as a predictor of morbidity in adulthood: a systematic review and meta-analysis. *Obes Rev* **17**, 56–67.

52. Park MH, Falconer C, Viner RM *et al.* (2012) The impact of childhood obesity on morbidity and mortality in adulthood: a systematic review. *Obes Rev* **13**, 985–1000.

53. Narayan KMV, Boyle JP, Thompson TJ *et al.* (2007) Effect of BMI on lifetime risk for diabetes in the US. *Diabetes Care* **30**, 1562–1566.

54. Owen CG, Martin RM, Whincup P *et al.* (2005) Effect of infant feeding on the risk of obesity across the life course: a quantitative review of published evidence. *Pediatrics* **115**, 1367–1377.

55. Yan J, Liu L, Zhu Y *et al.* (2014) The association between breastfeeding and childhood obesity: a meta-analysis. *BMC Public Health* **14**, 1267.

56. Lin SL, Leung GM, Lam TH *et al.* (2013) Timing of solid food introduction and obesity: Hong Kong's 'children of 1997' birth cohort. *Pediatrics* **131**, e1459–e1467.

57. Wang J, Wu Y, Xiong G *et al.* (2016) Introduction of complementary feeding before 4 months of age increases the risk of childhood overweight or obesity: a meta-analysis of prospective cohort studies. *Nutr Res* **36**, 759–770.