Validation, optimal threshold determination, and clinical utility of the Infant Risk of Overweight Checklist (IROC) for early prevention of child overweight

Authors

Professor Sarah A Redsell*, PhD, BSc (Hons), RGN, RHV, CPsychol
Professor of Public Health, Faculty of Health, Social Care and Education, Anglia Ruskin University, Young Street Site, East Road Campus, Cambridge, CB1 1PT, UK
Email address: Sarah.Redsell@anglia.ac.uk

Dr Stephen Weng, PhD, MPH, BA (Hons)
Research Fellow, School of Medicine, Division of Primary Care, University of Nottingham, Room 1307, The Tower, University Park, Nottingham, NG7 2RD, UK
Email address: Stephen.Weng@nottingham.ac.uk

Dr Judy A Swift, PhD, MSc, BSc (Hons)
Associate Professor in Human Nutrition, University Of Nottingham, Sutton Bonington Campus, LE12 5RD, UK
Email address: Judy.Swift@nottingham.ac.uk

Dr Dilip Nathan, BMedSci (Hons), BMBS, MRCP (UK), FRCPCH (UK), MMedSci
Consultant Paediatrician, Nottingham University Hospitals NHS Trust, Queen's Medical Centre, Nottingham, NG7 2UH, UK
Email address: dilip.nathan@nhs.net

Professor Cris Glazebrook, PhD, BSc (Hons), RGN, CPsychol
Head of Division of Psychiatry and Applied Psychology & Professor of Health Psychology, School of Medicine, Room B12, B Floor, Institute of Mental Health, University of Nottingham Innovation Park, Triumph Road, Nottingham, NG7 2TU, UK
Email address: Cris.Glazebrook@nottingham.ac.uk

*Corresponding Author

Abstract

Background

Previous research has demonstrated the predictive validity of the Infant Risk of Overweight Checklist (IROC). This study further establishes the predictive accuracy of the IROC using data from the Avon Longitudinal Study of Parents and Children (ALSPAC) and examines the optimal threshold for determining high risk of childhood overweight.

Methods

Using the IROC algorithm, we calculated the risk of being overweight, based on International Obesity Task Force (IOTF) criteria, in the first year of life for 980 children in the ALSPAC cohort at 5 years. Discrimination was assessed by the area under the receiver operating curve (AUC c-statistic). Net reclassification index (NRI) was calculated for risk thresholds ranging from 2.5% to 30% which determine cut-offs for identifying infants at risk of becoming overweight.

Results

At five years of age, 12.3% of boys and 19.6% of girls were categorised overweight. Discrimination (AUC c-statistic) ranged from 0.67 (95% CI 0.62 – 0.72) when risk scores were calculated directly to 0.93 (95% CI 0.88 – 0.98) when the algorithm was recalibrated and missing values of the risk factor algorithm were imputed. The NRI showed there were positive gains in reclassification using risk thresholds from 5% to 20%, with the maximum NRI being at 10%.

Conclusions

This study confirms the IROC has moderately good validity for assessing overweight risk in infants and offers an optimal threshold for determining high risk. The IROC algorithm has been imbedded into a computer programme for Proactive Assessment of Obesity Risk during Infancy (ProAsk) which facilitates early overweight prevention through communication of risk to parents.

Abbreviations:

IROC – Infant Risk of Overweight Checklist
ProAsk – Proactive Assessment of Obesity Risk during Infancy
ALSPAC – Avon Longitudinal Study of Parents and Children
IOTF – International Obesity Task Force
AUC – Area under receiver operating curve
PPV – Positive predictive value
NPV – Negative predictive value
CiF – Children in Focus
MCS – Millennium Cohort Study

Introduction

The risk factors for being overweight and obese in childhood are identifiable antenatally and/or during early infancy[1] and some of these, such as rapid weight gain, are potentially modifiable. Interventions which enable parents to address modifiable risk factors could contribute to early prevention of childhood overweight and obesity and therefore identifying infants at greatest risk is important. Current guidance[2-4] around the identification and assessment of overweight and obesity risk in infants and young children is contingent on a passive approach with little explicit guidance for health professionals. Health policy from the UK National Institute for Health and Care Excellence (NICE) and US Institute of Medicine (IoM) states that primary prevention and evidence-based interventions are important, but there is little consensus on whether childhood overweight risk can be identified by health professionals and, if so, how it should be communicated.

The Infant Risk of Overweight Checklist (IROC)[7] was developed from a comprehensive systematic review[1] of early-life risk factors associated with being overweight in childhood, and validated using data from the UK Millennium Cohort Study (MCS)[8]. The MCS is a prospective birth cohort in the United Kingdom which consists of data from longitudinal interviews with parents. The study analysis was restricted to 13,513 children who had completed anthropometric data at 3 years. The sample was divided into two cohorts: 80% was randomly selected to a derivation cohort for the development of the risk algorithm, and the remaining 20% was used to validate the algorithm. Stepwise logistic regression was used to determine the best predictor model for overweight at three years[8]. Seven predictors were found to be significantly associated with overweight at 3 years in a mutually adjusted predictor model: gender, birth weight, age-adjusted weight gain, maternal pre-pregnancy Body Mass Index (BMI), paternal BMI, maternal smoking in pregnancy and breast feeding status during the first year. The model yielded a moderately good ability to predict whether the infant would be overweight at three years of age (Area Under the Receiver Operator Curve ((R^2 = 0.92, AUC) = 0.72).

Additionally, the IROC algorithm was found to have a sensitivity (true-positive rate) of 70% (corresponding Positive Predictive Value [PPV] = 38%) and a specificity (true-negative rate) of 68% (corresponding Negative Predictive Value [NPV] = 87%)[8]. However, the IROC has only been validated in the MCS birth cohort and can therefore only be generalised to similar populations. Recruitment to the MCS was purposively sampled to ensure adequate representation of children from all four UK countries, deprived areas, and areas with high concentrations of Black and Asian families[8]. It is therefore important to validate the IROC in a second population-based birth cohort. The Avon Longitudinal Study of Parents and Children (ALSPAC) is a UK based birth cohort which contains a higher proportion of families with White British ethnicity and lower levels of deprivation compared to the MCS Cohort[9]. The aim of this study is to confirm the predictive accuracy of the IROC algorithm using data from the ALSPAC. The previous validation[8] used rapid weight gain over the first 12 months of life as a key predictor of child overweight. The present study evaluates whether earlier assessment of rapid weight gain, at 4 months, has acceptable predictive value as arguably earlier intervention will be more effective, and allow parents to make informed choices about weaning. A further aim is to establish the optimal threshold for determining high childhood overweight risk in infants, and to consider the operationalisation of the IROC algorithm for prevention of childhood obesity in a clinical setting.

Methods

Participants
The data for this study were a 10% sample (1,432 families) of births, known as the Children in Focus (CiF) cohort, from ALSPAC, a prospective birth cohort which recruited 14,541 pregnant women in Avon, UK from 1991-1992. Participants in the CiF cohort attended clinics at the University of Bristol at various time intervals between four months to 61 months of age (5 years). Follow up data for child's weight at 5 years is available for 980 families. Details on the selection, enrolment and cohort profile are published

elsewhere. Ethical approval was obtained from the ALSPAC Ethics and Law Committee and the Local Research Ethics Committees.

Infant Risk of Obesity Checklist (IROC)

The IROC algorithm was based on seven predictors which can be easily assessed in the early months of life: gender, infant birth weight, infant rapid weight gain, maternal pre-pregnancy BMI, paternal BMI, maternal smoking in pregnancy, and breast feeding in the first year (**Table 1**). The variables of gender (male/female), maternal smoking in pregnancy (no/yes), ever breastfed (no/yes) and infant rapid weight gain (no/yes) were dichotomised. Rapid weight gain was classified as weight gain > 0.67 standard deviation (SD)[11-13], from birth to the four month assessment by trained interviewers. Infant birth weight, obtained from the birth certificates or health records, was categorised into quintiles: < 2.93 kg; 2.93 to < 3.24 kg; 3.24 to < 3.49 kg; 3.49 to < 3.81 kg; ≥ 3.81 kg. Self-reported maternal pre-pregnancy BMI and paternal BMI were categorised into clinically recognised categories of <18.5 kg/m^2 for underweight, 18.5-< 25 kg/m^2 normal weight, 25-<30 kg/m^2 overweight or ≥30 kg/m^2 for obese.

INSERT TABLE 1

Outcome measure
The primary clinical outcome measure for the validation study was overweight in childhood defined by the International Obesity Task Force (IOTF)[14] sex and age-specific cut-offs which correspond to an adult BMI ≥ 25 kg/m^2. In addition, we compared IOTF overweight criteria[14] assessed at five-years to UK national guidelines[4] using the UK 1990 growth reference[15], which defines clinical overweight based on the 91st centiles, respectively. Both criteria were used to ensure the IROC can also be interpreted in a broader context.

Statistical Analysis

Data from all children in the ALSPAC CiF cohort on the seven risk factors utilised in the IROC algorithm were extracted to calculate risk and predict actual outcome of being overweight at 5 years. A total of four models were developed and analysed: (1) a *clinical model* which uses the original algorithm and assigns null values to missing data on risk factors, (2) a *recalibrated model* which uses multivariate logistic regression to generate and apply a recalibrated algorithm to better reflect the demographics of the ALSAPC CiF cohort, (3) an *imputed model* using multiple imputation[16] where ten copies of the existing dataset were generated to predict missing risk factor from multivariate models (assuming that the data were "missing-at-random"), (4) a *recalibrated imputed model* applying the recalibrated algorithm to the imputed dataset where there are no missing risk factor data. Discrimination, a measure of predictive accuracy defined by the ability to distinguish a case from a non-case, for each of the four models was assessed by the area under the receiver operating curve (AUC *c*-statistic). This calculates the probability that the IROC predictive risk score is higher for overweight children than for those who are healthy weight or underweight. The AUC can range from 0.5 (algorithm predicts based on complete random chance) to 1.0 (algorithm has perfect predictive accuracy). Models with an AUC of greater than 0.70 for predicting clinical outcomes are seen to be valid for use in practice, with consideration given to the balance of benefits of prevention against the harms of treatment[17]. To generate 95% confidence intervals around AUC *c*-statistic, standard errors were bootstrapped using a jack-knife procedure[18]. To test the effect of various risk thresholds for identifying "high-risk", the net reclassification index (NRI) was calculated for risk thresholds ranging from 2.5% to 30% for the primary clinical model[19]. The NRI in this study can be interpreted as the net change in percentage of children correctly re-classified from an initial baseline risk threshold of 2.5%[19]. This is calculated as: *NRI = (net increase of classification for cases/total number of cases) + (net decrease of classification of non-cases/total number of non-cases) x 100.* All analyses were conducted using STATA 13 MP; imputation using Multiple Imputation Chained Equations (MICE) *mim* module[16], calibration using multivariate logistic

regression (*logit* command) of IROC risk factors against the overweight outcome, discrimination analysis generating AUC and associated confidence intervals by *somersd* module[18], and net reclassification analysis using *diagt* module[20].

Results

Study population

At five years of age, 12.3% of boys and 19.6% of girls were categorised overweight by IOTF standards (Table 2). The proportion of children who were overweight increased as birth weight increased with the exception of children with birth weights less than 2.93 kg. Twenty-four percent of children who experienced rapid weight gain in the first four months were considered overweight at five years whereas only 13.3% of children who had not experienced rapid weight gain were considered overweight. There was an overall directly proportional relationship between maternal and paternal BMI and childhood overweight apart from the lowest BMI category < 18.5 possibly due to the low numbers of parents in that category. Mothers who smoked in pregnancy also had a high proportion of overweight children compared to those who did not smoke, while breast feeding resulted in a lower proportion of children who were overweight compared with not breast feeding. Missing values ranged from 0.70% (n = 7) for birth weight to 38.9% (n = 381) for paternal BMI. Additional analysis on the relationship between IROC risk factors and overweight status at 5 years is shown as a multivariate logistic model provided in Supplemental Table 1.

INSERT TABLE 2

External validation

The primary clinical model (Model 1) which directly applied the original IROC algorithm to the ALSPAC cohort of children resulted in a moderate discrimination using both IOTF (AUC *c*-statistic 0.67, 95% CI 0.62 to 0.72) and UK 1990 (AUC *c*-statistic 0.65, 95% CI 0.60 - 0.71) overweight criteria (Table 2). In Model 2, the algorithm was simply re-

7

calibrated reflecting the ALSPAC population characteristics. Compared to Model 1, recalibration (Model 2) increased the discrimination by an average of 3% and 2% based on IOTF and UK 1990 overweight criteria, respectively. Compared to Model 1, multiple imputation (Model 3) increased the discrimination by an average of 12% and 8% based on IOTF and UK 1990 overweight criteria, respectively. Moreover, if both recalibration and multiple imputation were used in combination (Model 4), an average of 26% and 25% increase in discrimination could be achieved, based on IOTF and UK 1990 overweight criteria, respectively. Further analysis on the IROC algorithm performance compared to single risk factors can be found in Supplemental Figure 1.

INSERT TABLE 3

Risk threshold analysis

The risk threshold for determining "high-risk" increased as the number of cases of children who were correctly identified decreased (sensitivity: true positive rate), while the number of non-cases correctly identified (specificity: true negative rate) increased (Table 4).

INSERT TABLE 4

The net reclassification index (NRI) showed there were positive gains in reclassification anywhere from 5% to 20%, with the maximum NRI being at 10%. This can be interpreted as the optimal point of balance between sensitivity and specificity. At this 10% threshold using IOTF defined overweight, 53% (n = 81) of all overweight children were correctly identified while 71% (n = 587) of all healthy weight children were correctly ruled-out, resulting in a maximum NRI of 21. Using UK 1990 defined overweight at a 19% threshold, 52% (n = 60) of all overweight children were correctly identified while 70% (n = 603) of all healthy weight children were correctly ruled out, resulting in a maximum NRI of 19.

Discussion

The results confirms that the IROC algorithm is a valid measure of risk, predicting overweight in childhood up to 5 years of age. The IROC algorithm incorporates seven factors which are routinely recorded or can easily be assessed in clinical practice. Whilst some risk factors in the ALSPAC were missing or undocumented, the algorithm remains clinically valid by assuming the average population risk for those particular missing risk factors. Even with missing values and no recalibration of the original algorithm to reflect target population, the IROC can achieve moderate prediction accuracy. With a simple recalibration this accuracy improves. With complete and accurate information for all the predictor variables, extremely high levels of predictive accuracy could be achieved. This study has also determined that the optimal risk threshold for determining "high risk" based on a balance between sensitivity and specificity is 10% using ALSPAC, a large UK based longitudinal population study. Although several prediction algorithms[21-24] exist, the significant advantage of the IROC is that it has proved valid in a heterogeneous UK population-based cohort, using parameters identified in a systematic review of the available literature[1].

Clinical implications

Studies have shown that UK healthcare professionals are often unsure whether and how to intervene with infants who gain weight rapidly. Similar studies in the US have found that clinicians only diagnose overweight or obesity in 1.1% to 31% of all overweight children, leading to suboptimal levels advice given and failure to refer to appropriate interventions. Whilst it is possible that 'at risk' infants could be identified solely by rapid weight gain/growth[27-29] during the first few months of life, this will not account for interactions between, for example, birth weight and weight gain[30] or the protective effects of breast feeding[1]. This could increase the number of false positive cases resulting in unnecessary intervention. The IROC algorithm has the potential to increase identification of infants at risk, whilst the high specificity and corresponding NPV of the algorithm indicates a relatively low chance of engaging parents unnecessarily. IROC is

9

designed to assess overweight in infants as young as four months, thus offering a valuable opportunity for effective early intervention. Counselling on weight status of pre-school children results in parents making positive lifestyle changes. There is also some evidence from randomised controlled trials of interventions delivered during infancy, typically including parental support, nutrition modification, healthy eating, breast feeding and weaning advice, that it is possible to reduce obesity risk during infancy[33-38].

ProAsk tool

Confirmation of the validity of the IROC and an optimal cut-off threshold for overweight risk enables potential implementation in a clinical setting. The IROC algorithm (Model 1) has been embedded within an interactive digital programme named ProAsk (Proactive Assessment of Obesity Risk during Infancy) for use on a hand-held tablet device. ProAsk is designed for clinical use with parents and prompts them to enter information on the seven risk factors used by the IROC algorithm. If information on any of the risk factors is unavailable the programme defaults to the average population risk of the missing risk factor. A percentage risk is calculated using the original algorithm[16] and, based upon the 10% risk threshold determined by the current study, gives text-based feedback of risk of overweight compared to other infants. Rather than using the terminology "high risk" as is appropriate in this technical report, "above average" risk is employed. ProAsk prompts the health professional to share this information with the parent with supporting text to aid appropriate, non-stigmatising language. The ProAsk programme then directs the healthcare professionals to a therapeutic wheel based on a recent systematic review[38], which promotes evidence-based behaviour change strategies in four areas: active play, milk and solid foods, sleeping and soothing, and infant feeding cues. Parents are encouraged to identify one area and to co-produce solution-focused yet evidence-based strategies for behaviour change with health professionals. ProAsk draws on the extended Health Belief model of behaviour change to increase parents' perceived susceptibility to the risk of child obesity and promote self-efficacy for behaviour change[39]. Importantly

the use of a digital tool supports the healthcare professional to communicate infant obesity risk in a personalised, non-stigmatising manner.

There is currently very little research guidance about the most appropriate way for health professionals to assess and communicate infant risk of overweight to parents in an effective, non-judgemental and sensitive way. Interactive digital technology has the potential to deliver personalised information about health risks in a way which increases understanding of risk and facilitates appropriate behaviour change without increasing patient anxiety[40]. Evidence from qualitative research suggests that general practitioners appreciate the options offered via tailored digital technology for health promotion in primary care[41] and dietitians have identified a need for digital resources to support their communication with children and parents around the sensitive topic of child obesity[42]. Although the results of a large systematic review suggest that health information technology can promote and support patient centred care[43], the potential of digital technology to enhance communication in child focused clinical settings has been largely neglected[42].

The MCS and ALSPAC databases both experienced higher attrition rates from study participants from lower socio-economic backgrounds leading to missing data. There is no current research evidence which demonstrates the best way to collect data on obesity risk factors in clinical practice. A multi-faceted approach is likely to be needed such as face-to-face data collection during a clinical consultation, via a questionnaire sent to parents. supplemented through access to medical records[44]. For instance, risk factors such as parental BMI could be obtained from consenting participants through their general practice computer records[44]. A study funded by the UK Medical Research Council Public Health Intervention Development Scheme (MRC-PHIND PH01/14-15) is exploring the feasibility of health visitors (public health nurses), using ProAsk to identify infants at increased risk of obesity, and to communicate that risk in conjunction with brief targeted health advice. The ProAsk study is being conducted in areas of high deprivation to

determine the feasibility of collecting this data clinically as well as the acceptability of the approach.

Limitations

There were several limitations of the study. The most substantial contributor towards missing data was paternal BMI due to lack of availability, or single-parent status. Although the sensitivity analysis performed in this study show the effects of missing values may be marginal, the missing values have the potential to generate greater variability in the observed overweight outcomes. This process allows assessment of all children despite missing data. Imputation techniques allow for the efficient analysis of all children in the cohort despite having missing data. This method of accounting for missing data is a robust and widely accepted[45] technique to improve the efficiency of the analysis and validity of the results. Additionally, the overweight outcomes assessed were limited to BMI and a relatively short follow-up duration (5 years of age). While there is a strong association between childhood overweight and adult obesity[46], it still unknown whether identifying infants at risk of overweight would extend to adulthood limited by the 5-year follow-up duration.

Conclusion

This study has confirmed that the IROC algorithm has moderately good validity for assessing overweight risk in infants as early as four months. It has also determined the optimal threshold for determining "high risk" infants, enabling services to prioritise those at greatest risk. Interactive digital technology (ProAsk) has the potential to facilitate the identification of "high risk" infants in clinical settings, to support the sensitive communication of this information to parents, and to empower parents to make evidence-based solution-focused behavioural changes.

Acknowledgments

Competing interests

The authors have no conflicts of interest to disclose

Author contributions

SAR and SFW wrote the first draft of the manuscript and the subsequent authors (SAR, SFW, JAS, DN, CG) contributed towards the study design, data analysis, revision, and preparation of the final manuscript. All five study authors have met authorship criteria set by International Committee of Medical Journal Editors (ICMJE).

References