Clinical Prediction of Iron Deficiency at Age 2 Years: A National Cross-sectional Study in France

Claire Guivarch, Anne-Sylvia Sacri, Corinne Levy, Alain Bocquet, Nathanaël Lapidus, Serge Hercberg, Pascale Hebel, Anne Chevé, Christian Copin, Morched Zouari, et al.

Clinical prediction of iron deficiency at 2 years-old: a national cross-sectional study in France

Authors' full name and academic degrees

Claire Guivarch, MSc [1]; Anne-Sylvia Sacri, MD, PhD [1,2]; Corinne Levy, MD [3,4,5]; Alain Bocquet, MD [3,6]; Nathanaël Lapidus, MD, PhD [7]; Serge Hercberg, MD, PhD [8,9]; Pascale Hebel, PhD [10]; Anne Chevé, MD [11]; Christian Copin, MD [12]; Morched Zouari, MD [13]; Laurent Gouya, MD, PhD [14]; Mariane de Montalembert, MD, PhD [1,2]; Jérémie F. Cohen, MD, PhD [1,2]; Martin Chalumeau MD, PhD [1,2]

Authors' affiliations

[1] Centre of Research in Epidemiology and Statistics, Obstetrical, Perinatal and Pediatric Epidemiology Research team (EPOPé), INSERM UMR 1153, Université de Paris, 75004 Paris, France

[2] Department of General Pediatrics and Pediatric Infectious Diseases, AP-HP, Necker-Enfants malades hospital, Université de Paris, 75015 Paris, France

[3] Association Française de Pédiatrie Ambulatoire (AFPA), Talence, France

[4] Association Clinique et Thérapeutique Infantile du Val-de-Marne (ACTIV), Saint-Maur des Fossés, France

[5] Université Paris Est, IMRB- GRC GEMINI, 94000 Créteil, France

[6] Comité de nutrition de la Société Française de Pédiatrie

[7] Sorbonne University, INSERM, Institut Pierre Louis d'Epidémiologie et de Santé Publique IPLESP, Public Health Department, Saint-Antoine Hospital, Assistance Publique-Hôpitaux de Paris, Paris, France

[8] Centre of Research in Epidemiology and Statistics, Nutritional Epidemiology Research Team (EREN), INSERM UMR 1153, Université de Paris, Sorbonne Paris Cité, France

[9] Département de Santé Publique, AP-HP, Hôpital Avicenne, Bobigny

[10] Centre de Recherche pour l'Etude et l'Observation des Conditions de vie (CREDOC), Paris, France

[11] AFPA, Brest, France.

[12] AFPA, Provins, France

[13] AFPA, Etampes, France

[14] Centre de recherche sur l'inflammation, INSERM UMR 1149, Université Paris Diderot; ERL CNRS 8252, Faculté de Médecine site Bichat, Paris, France

Corresponding author: Claire Guivarch

Maternité de Port-Royal, 53 Avenue de l'Observatoire, 75014 Paris

Phone: +33 1 42 34 55 80

E-mail: claire.guivarch@inserm.fr

Short title

Clinical prediction of iron deficiency at 2 years-old

Conflict of interest statement

The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. The authors have no patents, products in development or marketed products to declare.

Funding sources

This work was supported by the Secteur Français des Aliments de l'Enfance and the French Ministry of Health "DGOS PHRC régional 2014 no. AOR14053". The sponsor was Assistance Publique – Hôpitaux de Paris (Délégation à la Recherche Clinique et à l'Innovation [DRCI]).

Clinical Trial Registration

Iron deficiency (ID) in infants, NCT02484274.

Deidentified individual participant data (including data dictionaries) will be made available, in addition to study protocols, the statistical analysis plan, and the informed consent form. The data will be made available upon publication to researchers who provide a methodologically sound proposal for use in achieving the goals of the approved proposal. Proposals should be submitted to martin.chalumeau@inserm.fr

Abbreviations

AAP: American Academy of Pediatrics

AUROC: Area under the receiver operating characteristic curve

CI: Confidence interval

ID: Iron deficiency

SF: Serum ferritin

TRIPOD: Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis

WG: Weeks of gestation

Keywords

Screening; prediction tools; nationwide cohort study

Abstract

Objective(s): The American Academy of Pediatrics (AAP) recommends iron deficiency (ID) screening in at-risk infants, based on several clinical criteria which diagnostic accuracy has never been evaluated. We aimed to assess this diagnostic accuracy and to develop prediction tools for ID in 2-years-old infants.

Study design: In a national cross-sectional study conducted in primary care pediatricians' practices throughout France, 2-years-old infants were consecutively included (2016-2017). Multivariable logistic regression modeling and bootstrapping were used to develop several clinical models to predict ID (serum ferritin <12 μg/L). These models used the best criteria and combinations among AAP criteria adapted to the European context (n=10), then all potential predictors (n=19). One model was then simplified into a simple prediction tool.

Results: Among 568 included infants, 38 had ID (6.7%). In univariable analyses, no significant association with ID was observed for 8 of the 10 adapted AAP criteria. Three criteria (both parents born outside the European Union, low weight at one year old, and weaning to cow's milk without supplemental iron) were retained in the "AAP model", which AUROC, sensitivity, and specificity were 0.62 (95% CI 0.58-0.67), 30% (22-39%) and 95% (92-97%), respectively. Four criteria were retained in a newly derived "simple prediction tool" (≥one criterion among the three previous plus duration of iron-rich formula consumption <12 months), which AUROC, sensitivity, and specificity were 0.72 (0.65-0.79), 63% (47-80%) and 81% (70-91%), respectively.

Conclusion(s): All prediction tools achieved acceptable diagnostic accuracy. The newly derived simple prediction tool offered potential ease of use.

INTRODUCTION

Iron deficiency (ID) is the most prevalent micronutrient deficiency worldwide, affecting an estimated two billion people (1). Infants under two years are at risk of developing ID, because of the high iron requirements needed for their rapid growth. ID may be associated with short- and long-term adverse neurocognitive outcomes when it occurs in infants (2-4). Medical societies and public health authorities have developed primary prevention programs aiming at tackling ID in infants by recommending optimized iron intakes (5-7). Despite these preventive strategies, ID is estimated frequent among infants in industrialized countries (6-8), with prevalence oscillating between 3% and 33% in Europe and being much higher in deprived populations (5, 9, 10). This partial failure of preventive strategies raises the question of the relevance of adding validated screening strategies to reduce the consequences of ID.

There is no consensus between learned societies on the relevance of ID screening compared to no screening (6, 11). The US Preventive Services Task Force (11), and the European Society for Pediatric Gastroenterology, Hepatology, and Nutrition (5) do not recommend screening for ID, based on the current insufficient evidence. On the contrary, the American Academy of Pediatrics (AAP) recommends universal screening for anemia at one-year-old and targeted screening when infants are considered at high risk of developing ID (6) (**Appendix 1**). The Canadian Paediatric Society (CPS) recommends risk assessment in the first 2 years and targeted screening for ID in high-risk infants (12) (**Appendix 1**). In these recommendations, there is no consensus on the relevance of universal screening of ID compared to targeted screening based on clinical predictors of ID (6, 11, 12). One of the most used reference standards for detecting ID is a low serum ferritin level (5-7, 13), but its measurement requires an invasive blood sampling and would be costly if measured for an entire population. Clinical signs of ID, such as pallor of the skin and fatigue, are neither sensitive nor specific (14). Predictors of ID are well known, including low socio-economic

status, poor perinatal iron stock and imbalance between requirements and dietary intake (1, 5-7). A reason for such variability in guidance at national and international levels could be the lack of formal assessment of the diagnostic accuracy of the proposed criteria (11), taken individually or combined in prediction tools. As a consequence, iron status is not routinely examined in infancy (15), and primary healthcare providers are left with various potential screening strategies ranging from no screening to universal screening.

Our objectives were to assess the diagnostic accuracy of AAP criteria and to develop and validate new prediction tools for detecting ID in 2-years-old infants, using data from a national cross-sectional study in France.

MATERIAL AND METHODS

Study design and setting

The present study is a planned ancillary analysis of a French cross-sectional observational study aiming at evaluating the effectiveness of a national ID prevention strategy based on using iron-fortified formula at weaning (16). The study protocol was approved by local ethics and administrative authorities (CPP IDF III no. 3295) and registered (ClinicalTrials.gov identifier: NCT02484274). Signed consent of at least one parent was obtained before inclusion. This study is reported according to the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD) statement (17). We used Stata/SE 13.1 (StataCorp, College Station, Texas) for data analysis and R 3.6.2 software for internal validation. Sample size calculation was performed for the primary objective of the study (16), and no a posteriori calculation was performed for this ancillary analysis.

Study participants

From January 2016 to December 2017, 58 primary care pediatricians throughout 16 French regions were asked to include 10 consecutive patients (**Appendix 2**), aged from 22 to 26 months, living in France, with health insurance coverage (16). Children were excluded if they were known to have any chronic disease that may affect iron metabolism (e.g., celiac disease; complete list in **Appendix 3**).

Clinical and laboratory data collection

Socio-demographic data, medical history, and past and current dietary habits were collected by the investigator at inclusion (16). A fasting blood sample was taken in a nearby laboratory. Serum ferritin and CRP measurements were centralized in a unique laboratory (CERBA laboratory, Saint-Ouen l'Aumône, France) (16). In case of fever or any other condition that could alter the biological iron status (e.g., bronchiolitis, gastroenteritis), the blood sampling was postponed to 15 days after discontinuation of symptoms. Children for whom no blood sample was taken or who had a CRP level ≥ 10 mg/L were secondarily excluded (16) (flow chart in **Appendix 4**). Then, sensitivity analyses were performed by excluding infants who had a CRP level ≥ 5 mg/L.

Iron deficiency and potential predictors definition

ID was defined by a serum ferritin level of <12 µg/L, as in several previous studies (5, 6, 18, 19), and measured blinded to all clinical data.

AAP recommendations are reproduced verbatim in **Appendix 1** (6). Some AAP criteria required an adaptation to the European context (e.g., "children of Mexican American descent"), or clarification (e.g., "poor growth"). Some were not relevant to the study context (e.g., prevalence of lead poisoning which is almost null among French children (20)). Thus,

we studied the 10 following modified and dichotomized AAP criteria: parents' country of birth (both outside the European Union [EU] or not), parents' education level (both did not attend university or not), parents' professional status (both unemployed or not), family having a health coverage fully funded publicly, preterm birth (<37 or ≥ 37 weeks of gestation), small birth weight (<2500 or ≥ 2500 g), low weights for age at one and two year(s) old (WHO weight for age z-score <-2 SD), "exclusive" breastfeeding ("no other type of milk" but other liquids could be ingested) beyond four months without supplemental iron, and weaning to cow's milk without supplemental iron (**Appendix 1**).

We also used 9 other potential predictors of ID that were previously suggested in the literature, related to socio-economic status, perinatal and medical history, and nutrition (5-7): single motherhood (yes or no), rank among siblings (first/second child or ≥ third child), mother's ID during pregnancy (yes or no), iron supplementation since birth (yes or no), and durations of consumption of infant formula (<4 or ≥ 4 months), iron-rich formula [<12 or ≥ 12 months; a global variable that sums the durations of consumption of "follow-on formula" (recommended at age 6-12 months) and "young children formula" (recommended after 12 months of age)], milk protein hydrolysates (<6 or ≥ 6 months), cow's milk (<12 or ≥ 12 months) and iron-fortified cereals (<12 or ≥ 12 months) (**Appendix 1**).

We chose to a priori dichotomize all potential predictors of ID to take into account the results of a qualitative survey regarding clinical applicability of prediction tools conducted among a panel of ten primary care pediatricians. The panel stated that any clinical prediction tool aiming at a day-to-day use without computation should be based on less than 5 yes/no criteria. Dichotomization was performed after multiple imputations (see below). Sensitivity analyses were performed with different thresholds to test the robustness of our results. A target zone for acceptability of prediction models was also defined during this panel survey (see below).

Statistical analyses

We first described the general characteristics of study participants. Unadjusted associations between ID and each of the 19 dichotomized potential predictors were assessed by odds ratios (OR) with corresponding 95% confidence intervals (95% CI). Then, a backward stepwise selection of the 10 dichotomized adapted AAP criteria was performed in 5,000 bootstrap samples (*b*=200 bootstrap samples per multiple imputations dataset, *m* = 25, see below), using a *P*-value of 0.157 for removal, as suggested (21). The final model retained AAP criteria that were selected in at least 60% of the bootstrap models and is referred to as "*parsimonious AAP model*". One predictor (i.e., weaning to cow's milk without supplemental iron) that was not selected with this procedure, due to its low proportion in this study, was finally forced into prediction models because of its strong association with ID in univariable analysis and previous literature (22). Additional analyses were performed to study the diagnostic accuracy of AAP criteria according to how they were combined: (i) in a "*complete AAP model*" where all 10 dichotomized AAP were forced or (ii) in a "*simple additive AAP model*" where each of the 10 AAP criteria was assigned one point.

To develop new prediction models, all dichotomized candidate predictors (n=19; including the 10 adapted AAP criteria) were included in a multivariable logistic regression model, regardless of their association with ID in univariable analysis (23). The same above-described selection process of predictors was applied and provided a "*new model*". This "*new model*'' was further simplified into a "*simple prediction tool*" (binary tool: no criterion vs. ≥ 1) to fit with applicability constraints suggested by the panel of primary care pediatricians as described above. To assess all models' diagnostic accuracy, we used the area under the receiver operating characteristics curve (AUROC), sensitivities, specificities, and accuracies (correctly classified percentage). Internal validation was performed by bootstrap (*b*=500 per

multiple imputation dataset) (24). Calibration was studied graphically by plotting observed proportions of ID by deciles of predicted probabilities of ID (25).

Finally, to interpret the diagnostic accuracy of both individual predictors and prediction models, we used the target zone provided by the pediatricians panel, i.e., a specificity of at least 75%. This target zone was defined regarding the current situation in some European countries including France where no ID screening strategy is recommended, for a theoretical sensitivity of 0% and a theoretical specificity of 100%. Thus, clinical prediction models were dichotomized on the threshold providing a specificity of at least 75%, then sensitivity, specificity, accuracy, positive predictive value, negative predictive value, positive likelihood ratio (LR+) and negative likelihood ratio (LR-) were estimated. For the "*simple prediction tool*", theoretic positive predictive value and (1- negative predictive value) were estimated in settings with different prevalences of ID.

The number of missing data ranged from 0% to 12% per candidate predictor variable (**Appendix 5**). We used multiple imputations with chained equations (m = 25) with predictive mean matching for continuous variables and logistic regression for categorical variables to generate values for missing data. Statistical analyses were performed separately in each imputed dataset; estimated parameters and corresponding variances were pooled using Rubin's rules (23, 26, 27).

RESULTS

Participants

As previously described (16), the mean age of the 568 infants included in the analysis was 24 months (SD, ±0.6), 49% were girls, 5.2% had both parents born outside the EU, 2.8% lived in

a single-motherhood family, 3.5% had both parents unemployed, 7.6% had health coverage fully funded publicly, and 38 (6.7%; 95% CI, 4.6-8.8%) had ID. Analyzed infants did not differ from excluded infants, except for some dietary variables (**Appendix 6**).

Performance of AAP criteria alone or combined

In univariable analyses, no significant association with ID was observed for 8 of the 10 adapted AAP criteria: parents' educational level, parents' professional status, having or not a health coverage fully funded publicly, birth term, birth weight, low weights at one and two years old, and "exclusive" breastfeeding beyond four months without iron supplementation. ID was significantly associated with two of the 10 adapted AAP criteria: both parents born outside the EU (OR 4.1; 95% CI, 1.6-10.7) and weaning to cow's milk without supplemental iron (OR 80.2; 95% CI, 9.1-706.0) (**Appendix 7**).

Two AAP predictors were selected in at least 60% of the backward stepwise selection models: both parents born outside the EU and low weight at one year old (**Appendix 8.A**). The logistic regression "*parsimonious AAP model*" was built with these two predictor variables plus weaning to cow's milk without supplemental iron that was forced into the model. Calibration plots of the "*parsimonious AAP model*" showed high agreement between predicted probabilities of ID and observed outcomes (**Figure 1**). After bootstrap internal validation, this model had an optimism-corrected AUROC of 0.62 (95% CI, 0.58 to 0.67). After dichotomization around an individual risk threshold of 0.18 (the first one providing a specificity >75%), its sensitivity was 30% (95% CI, 22-39%), its specificity 95% (95% CI, 92-97%) and its accuracy 90% (95% CI, 88-93%). Other prediction tools based on the 10 adapted AAP criteria (i.e., a "*simple additive AAP model*" and the "*complete AAP model*"), had similar diagnostic accuracy (**Appendix 9**). Estimations of the positive predictive value,

the negative predictive value, the LR+, the LR- and the 1/LR- of each developed prediction tools are reported in **Appendix 10.**

New prediction models

In univariable analyses, significant positive associations with ID were found for 3 of the 9 non-AAP criteria: mother's ID during pregnancy (OR 2.4; 95% CI, 1.2-4.7), duration of consumption of "iron-rich formula" <12 months (OR 7.2; 95% CI, 3.6-14.2), and duration of consumption of cow's milk ≥ 12 months (OR 4.0; 95% CI, 1.8-9.0) (**Appendix 7**).

Three among the 19 potential predictors were selected in at least 60% backward stepwise selection models performed on the bootstrap samples: both parents born outside the EU, low weight at one year old, and duration of consumption of "iron-rich formula" <12 months (**Appendix 8.B**). The "*new model*" was built with these three predictor variables plus weaning to cow's milk without supplemental iron. In logistic regression analysis, significant positive associations with ID were found with having both parents born outside the EU (adjusted OR 3.5; 95% CI, 1.8-10.3), low weight at one year (adjusted OR 5.9; 95% CI, 1.0-34.0), duration of consumption of "iron-rich formula" <12 months (adjusted OR 5.2; 95% CI, 2.5-11.0), and weaning to cow's milk without supplemental iron (adjusted OR 25.6; 95% CI, 2.7-239.2). Calibration plots of the "*new model*" showed high agreement between predicted probabilities of ID and observed outcomes (**Figure 1**). After bootstrap internal validation, this model had an optimism-corrected AUROC of 0.73 (95% CI, 0.66-0.80). After dichotomization around an individual risk threshold of 0.10, its sensitivity was 62% (95% CI, 49-78%), its specificity 81% (95% CI, 77-86%) and its accuracy 80% (95% CI, 76-84%). The "*new model*" was then simplified by dichotomization for a threshold of at least one criterion, providing the "*simple prediction tool*" (no criterion vs ≥ one criterion among the four above criteria), that had, after bootstrap internal validation, an optimism-corrected AUROC of 0.72

(95% CI, 0.65-0.79%). Its sensitivity was 63% (95% CI, 47-80%), its specificity 81% (95% CI, 70-91%) and its accuracy 80% (95% CI, 70-89%) (**Table 1**). For each developed prediction tool, estimations of the positive predictive value, the negative predictive value, the LR+, the LR- and the 1/LR- are reported in **Appendix 10**. For the "*simple prediction tool*", estimations of the positive predictive value and the (1 – negative predictive value) in settings with different prevalences of ID, but with the same sensitivity and specificity are reported in **Appendix 11**.

To test the robustness of these results, we first modified the thresholds used to dichotomize some potential predictors (durations of consumption of infant formula and iron-rich formula -**Appendix 12**-) and then excluded infants who had CRP level ≥ 5 mg/L, which led to a sample of 530 infants (**Appendix 13**). These sensitivity analyses yielded similar results to those of our main analyses.

DISCUSSION

Main results

This study is the first attempt to formally validate the AAP criteria for targeted screening ID in at-risk infants. We found that three AAP criteria were strong predictors of ID. Various combinations of AAP criteria provided clinically acceptable diagnostic accuracy, with a sensitivity varying from 30% (95% CI, 22-39%) to 41% (95% CI, 26-56%) for a specificity of at least 75%. We developed two new clinical prediction tools for ID that provided similar accuracy than tools based on AAP criteria. Notably, a simple tool in which high risk of ID was defined by the presence of at least one out of four simple criteria had an acceptable

AUROC of 0.72 (0.65 – 0.79) (28) and provided a sensitivity of 63% (95% CI, 47-80%) for a specificity of 81% (95% CI, 70-91%).

Strengths and limitations

Our study has several strengths. Measurements of serum ferritin levels were standardized and centralized, as recommended by the World Health Organization (29). Biological and clinical data were measured in a blinded way (16). We used recommended statistical methods for developing and validating clinical prediction tools, such as multiple imputations and bootstrap (21, 24), however we cannot exclude residual predictor selection bias.

There were several limitations. First, the interpretation and adaptation of some of the criteria of the AAP is arguable. For example, "poor growth" definitions in the literature (30) and other thresholds used to dichotomize the potential predictors are not consensual, leading to a potential non-differential misclassification bias. We applied AAP criteria to our study population of 2-years-old infants, while AAP seems to target preferably 1-year-old infants, which could explain why no association was found between ID and well-known risks factors of ID, such as history of low birth weight (31). We did not explore some risk factors of ID in the present study (e.g., special health care needs) (6, 31), because we found more relevant to target healthy infants seen in general medical practice. Second, to reach the pre-specified target zone for specificity of at least 75%, all prediction models finally had quite low sensitivities (between 30% and 63%) but good accuracies (between 78% and 90%). In the actual context of lack of international consensus about ID screening and as ID consequences are not fatal, these results were deemed acceptable from clinical and public health points of view. These conclusions may differ depending on the national epidemiology of ID and the local strategy to prevent its consequences. Third, the low number of infants with ID in our study (n=38, 6.7%) and the dichotomization of serum ferritin (<12 µg/L) and of potential

predictors may have been responsible for low statistical power (32). Indeed, in a study performed on the same dataset, history of prematurity was significantly associated with serum ferritin used as a continuous variable (16). Finally, the lowest prevalence of health coverage fully funded publicly (7.6%) than expected in the general French population points to a potential selection bias (33). An explanation could be that participants were recruited by pediatricians, while in France only 20% of infants are regularly followed by a pediatrician at 2 years of age (34). Children followed by pediatricians may have more favorable socio-economic characteristics and may be more likely to comply with nutritional recommendations, regardless of socio-economic level, than those followed by general practitioners. Thus, our results should be validated in a group of infants with a higher proportion of low socio-economic status families.

Our findings may not be generalizable to other high-income settings that have a higher prevalence of ID in infants. External validation of our screening tools is thus warranted (35, 36). The fact that dietary habits and ID prevention policies are different across countries also suggests the need for external validation studies. Our screening tools may only apply to countries where populations are used to consuming iron-rich formulas, such as France where respectively 65% of one year old and 43% of 2-years-old infant consume it regularly (37). Few data on iron-rich formulas consumption in toddlers is available in other high-income countries (37).

Implications

Primary healthcare providers were, until now, without clear recommendations regarding ID screening in infants, with guidance ranging from universal screening (6, 12) to no screening at all (5, 11). An intermediate solution would be to perform targeted screening based on validated combinations of ID predictors. The AAP suggested both universal screening and

targeted screening based on criteria which diagnostic accuracies have never been evaluated (6). In this first external validation study, two simple prediction tools had similar diagnostic accuracies than combinations of the AAP criteria. Thus, primary healthcare providers in contexts similar to the French one who wish to screen for ID in at-risk young children could rely on one of these new tools. The simple prediction tool, in which high risk of ID was defined by the presence of at least 1 criterion out of four (both parents born outside the EU, low weight at one year, duration of consumption of "iron-rich formula" <12 months and weaning to cow's milk without supplemental iron), offered potential ease of use, with the aim of limiting as much as possible additional costs and loss of time during the consultation. Nevertheless, impact analyses are necessary to assess the acceptability, usefulness and ease of use of our screening tools in primary healthcare providers' day-to-day practice (21, 38). Even if internal bootstrap validation of the prediction tools shown that the optimism bias in our study was low, these encouraging results need to be further validated in other age groups (closer to one year old) and in other settings to evaluate their transportability in different nutritional contexts and populations.

Acknowledgments

The authors thank URC-CIC Paris Descartes Necker Cochin (project manager: Nelly Briand and clinical research associates: Camille Boucheny, Charlène Ambvouli and Emilie Le Roux) for the implementation and monitoring of the study. The authors thank ACTIV network clinical research associates: Claire Prieur, Deborah Kern. They thank the participant families and the pediatricians who recruited them: Jean-Marie André, Pierre Bakhache, Sophie Barrois, Bénédicte Blanc, Marie-Emmanuelle Bouloumié, Fabienne Cahn-Sellem, Marie Helene Cavert, Anne Cheve, Christian Copin, Muriel Cornic, François Corrard, Armelle Courboin, Véronique Dagrenat, Caroline De Backer, Pascal Dechelette, Jean-François Delobbe, Michele Dudreuilh, Corinne Dumas, Christine Duvant-Fay, Annie Elbez, Chantal Gayot, Nathalie Gelbert, Piotr Gembara, Patricia Giraud-Escoffier, Geneviève Granat, Cécile Guiheneuf, Frédéric Hassid, Kai Kassmann, Sylvie Labarge, Sophie Langlais, Hélène Le Scornet, Dominique Lemarié, Claudia Litzler, Christine Magendie, Delphine Menet, Bénédicte Nold, Jean- Claude Oilleau, France Paratte, Brigitte Pinçant, Massimiliano Rallo, Sophie Ravilly, Olivier Romain, Bechara Saade, Catherine Salinier, Claire Sarreau, Ruth-Caroline Schweitzer, Anne Sengier-Njinou, Philippe Simon, David Somerville, Joachim Stoller, Marie-Pierre Streicher, Franck Thollot, Catherine Tuberg Romain, Jean-Christophe Vigreux, Pascale Vurpillat-Almanza, Andreas Werner, Alain Wollner, Morched Zouari.

References

1. Zimmermann MB, Hurrell RF. Nutritional iron deficiency. Lancet. 2007;370:511-20.

2. Lozoff B, Jimenez E, Hagen J, Mollen E, Wolf AW. Poorer behavioral and developmental outcome more than 10 years after treatment for iron deficiency in infancy. Pediatrics. 2000;105:E51.

3. Algarin C, Nelson CA, Peirano P, Westerlund A, Reyes S, Lozoff B. Iron-deficiency anemia in infancy and poorer cognitive inhibitory control at age 10 years. Dev Med Child Neurol. 2013;55:453-8.

4. Lozoff B, Jimenez E, Wolf AW. Long-term developmental outcome of infants with iron deficiency. N Engl J Med. 1991;325:687-94.

5. Domellöf M, Braegger C, Campoy C, Colomb V, Decsi T, Fewtrell M, et al. Iron requirements of infants and toddlers. J Pediatr Gastroenterol Nutr. 2014;58:119-29.

6. Baker RD, Greer FR. Diagnosis and prevention of iron deficiency and iron-deficiency anemia in infants and young children (0-3 years of age). Pediatrics. 2010;126:1040-50.

7. EFSA. European Food Safety Authority. Panel on Dietetic Products, Nutrition and Allergies (NDA). Scientific Opinion on nutrient requirements and dietary intakes of infants and young children in the European Union. European Food Security Authority (EFSA) Journal [serial online]. 2013;11:3408-511.

8. van der Merwe LF, Eussen SR. Iron status of young children in Europe. Am J Clin Nutr. 2017;106:1663s-71s.

9. Sacri AS, Hercberg S, Gouya L, Levy C, Bocquet A, Blondel B, et al. Very low prevalence of iron deficiency among young French children: A national cross-sectional hospital-based survey. Matern Child Nutr. 2017;14:e12460.

10. Eussen S, Alles M, Uijterschout L, Brus F, van der Horst-Graat J. Iron intake and status of children aged 6-36 months in Europe: a systematic review. Ann Nutr Metab. 2015;66:80-92.

11. Siu AL. Screening for iron deficiency anemia in young children: USPSTF recommendation statement. Pediatrics. 2015;136:746-52.

12. Unger SL, Fenton TR, Jetty R, Critch JN, O'Connor D L. Iron requirements in the first 2 years of life. Paediatr Child Health. 2019;24(8):555-6.

13. Joint World Health Organization/Centers for Disease Control and Prevention technical consultation on the assessment of iron status at the population level. Assessing the iron status of populations. Geneva, Switzerland: World Health Organization (WHO) Press; 2007 [Available from: http://www.who.int/nutrition/publications/micronutrients/anaemia_iron_deficiency/9789241596107/en/].

14. Lopez A, Cacoub P, Macdougall IC, Peyrin-Biroulet L. Iron deficiency anaemia. Lancet. 2016;387:907-16.

15. Qasem WA, Friel JK. An Overview of Iron in Term Breast-Fed Infants. Clin Med Insights Pediatr. 2015;9:79-84.

16. Sacri AS, Bocquet A, de Montalembert M, Hercberg S, Gouya L, Blondel B, et al. Young children formula consumption and iron deficiency at 24 months in the general population: A national-level study. Clin Nutr. 2020.

17. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. BJOG. 2015;122:434-43.

18. Cook JD, Lipschitz DA, Miles LE, Finch CA. Serum ferritin as a measure of iron stores in normal subjects. Am J Clin Nutr. 1974;27:681-7.

19. World Health Organization. Nutritional anaemias: tools for effective prevention and control Geneva, Switzerland: World Health Organization (WHO) Press; 2017 [Available from: https://www.who.int/publications/i/item/9789241513067].

20. Etchevers A, Bretin P, Lecoffre C, Bidondo ML, Le Strat Y, Glorennec P, et al. Blood lead levels and risk factors in young children in France, 2008-2009. Int J Hyg Environ Health. 2014;217:528-37.

21. Steyerberg EW. Clinical prediction models: a pratical approach to development, validation, and updating. 1st, editor. New York, NY: Springer; 2009. 500 p.

22. Leung AK, Sauve RS. Whole cow's milk in infancy. Paediatr Child Health. 2003;8:419-21.

23. Moons KGM, Kengne AP, Woodward M, Royston P, Vergouwe Y, Altman DG, et al. Risk prediction models: I. Development, internal validation, and assessing the incremental value of a new (bio)marker. Heart. 2012;98:683.

24. Harrell Jr FE. Regression modeling strategies: with applications to linear models, logistic and ordinal regression, and survival analysis: Springer; 2015.

25. Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. Epidemiology. 2010;21:128-38.

26. White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. Stat Med. 2011;30:377-99.

27. Rubin DB. Multiple imputation for nonresponse in surveys: John Wiley & Sons; 2004.

28. Mandrekar JN. Receiver operating characteristic curve in diagnostic test assessment. J Thorac Oncol. 2010;5:1315-6.

29. World Health Organization. Iron deficiency anaemia. Assessment, prevention and control. A guide for programme managers Geneva, Switzerland: World Health Organization (WHO) Press; 2001 [Available from: http://www.who.int/nutrition/publications/en/ida_assessment_prevention_control.pdf].

30. Scherdel P, Dunkel L, van Dommelen P, Goulet O, Salaun JF, Brauner R, et al. Growth monitoring as an early detection tool: a systematic review. Lancet Diabetes Endocrinol. 2016;4:447-56.

31. Burke RM, Leon JS, Suchdev PS. Identification, prevention and treatment of iron deficiency during the first 1000 days. Nutrients. 2014;6:4093-114.

32. Altman DG, Royston P. The cost of dichotomising continuous variables. BMJ. 2006;332(7549):1080.

33. Blondel B, Coulm B, Bonnet C, Goffinet F, Le Ray C. Trends in perinatal health in metropolitan France from 1995 to 2016: Results from the French National Perinatal Surveys. J Gynecol Obstet Hum Reprod. 2017;46:701-13.

34. Bocquet A, Chalumeau M, Bollotte D, Escano G, Langue J, Virey B. Comparison of prescriptions by pediatricians and general practitioners: a population-based study in Franche-Comte from the database of Regional Health Insurance Fund. Arch Pediatr. 2005;12:1688-96.

35. Steyerberg EW, Harrell FE, Jr. Prediction models need appropriate internal, internal-external, and external validation. J Clin Epidemiol. 2016;69:245-7.

36. Moons KG, Kengne AP, Grobbee DE, Royston P, Vergouwe Y, Altman DG, et al. Risk prediction models: II. External validation, model updating, and impact assessment. Heart. 2012;98:691.

37. Sacri AS, de Lauzon–Guillain B, Dufourg MN, Bois C, Charles MA, Chalumeau M. Iron-fortified formula use in young children and association with socioeconomic factors in the French nationwide ELFE cohort. Acta Paediatr. 2019;108:1285-94.

38. Kappen TH, van Klei WA, van Wolfswinkel L, Kalkman CJ, Vergouwe Y, Moons KGM. Evaluating the impact of prediction models: lessons learned, challenges, and recommendations. Diagn Progn Res. 2018;2:11.

TABLES

Table 1. Diagnostic accuracy of clinical prediction tools under evaluation (n=568)

Figure 1. Calibration plots of predicted probabilities of ID from the "*parsimonious AAP model*" (A) and the "*new model*" (B) and observed outcome

Table 1. Diagnostic accuracy of clinical prediction tools under evaluation (n=568) [a]

Tool	Number of predictor variables	AUROC: Estimate, (95% CI) Validation [b], (95% CI)	Positivity threshold [c]	Sensitivity: Estimate, % (95% CI) Validation [b], % (95% CI)	Specificity: Estimate, % (95% CI) Validation [b], % (95% CI)	Accuracy [d]: Estimate, % (95% CI) Validation [b], % (95% CI)
"*Parsimonious AAP model*"	3	0.63 (0.56 – 0.71) 0.62 (0.58 – 0.67)	Predicted probability of ID ≥ 0.18	32 (17 – 46) 30 (22 – 39)	94 (92 – 96) 95 (92 – 97)	90 (88 – 93) 90 (88 – 93)
"*New model*"	4	0.74 (0.66 – 0.83) 0.73 (0.66 – 0.80)	Predicted probability of ID ≥ 0.10	63 (48 – 79) 62 (49 – 78)	81 (77 – 86) 81 (77 – 86)	80 (76 – 83) 80 (76 – 84)
"Simple prediction tool"	4	0.72 (0.64 – 0.80) 0.72 (0.65 – 0.79)	≥ 1 criterion	63 (48 – 79) 63 (47 – 80)	81 (77 – 84) 81 (70 – 91)	80 (76 – 84) 80 (70 – 89)

"*Parsimonious AAP model*" included the 3 AAP adapted criteria retained in at least 60% of the bootstrap models, i.e. both parents born outside the EU, low weight at one year old and weaning to cow's milk without supplemental iron (score range between 0 and 1, obtained as the individual risk prediction from a logistic regression model which included these 3 criteria).
"*New model"* included the 4 criteria retained in at least 60% of the bootstrap models, i.e. both parents born outside the EU, low weight at one year old, duration of consumption of "iron-rich formula" <12 months and weaning to cow's milk without supplemental iron (score range between 0 and 1, obtained as the individual risk prediction from a logistic regression model which included these 4 criteria)
"Simple prediction tool" included the 4 criteria mentioned in the "*New Model*" (binary score: no criterion vs at least 1 criterion)

[a] After performing statistical analyses separately in each imputed dataset (m=25); estimated parameters were pooled using Rubin's rules

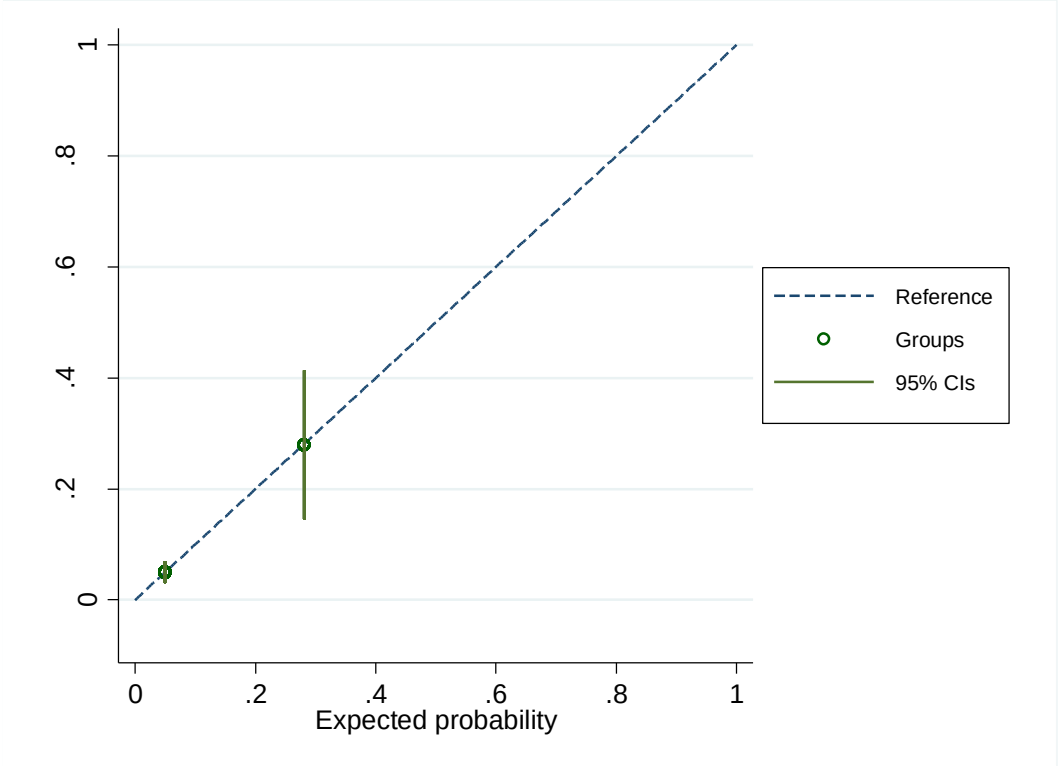
[b] Internal validation was performed by bootstrap (*b*=500 per multiple imputation dataset)

[c] The threshold was set to obtain a specificity of at least 75% (see explanations in the text)

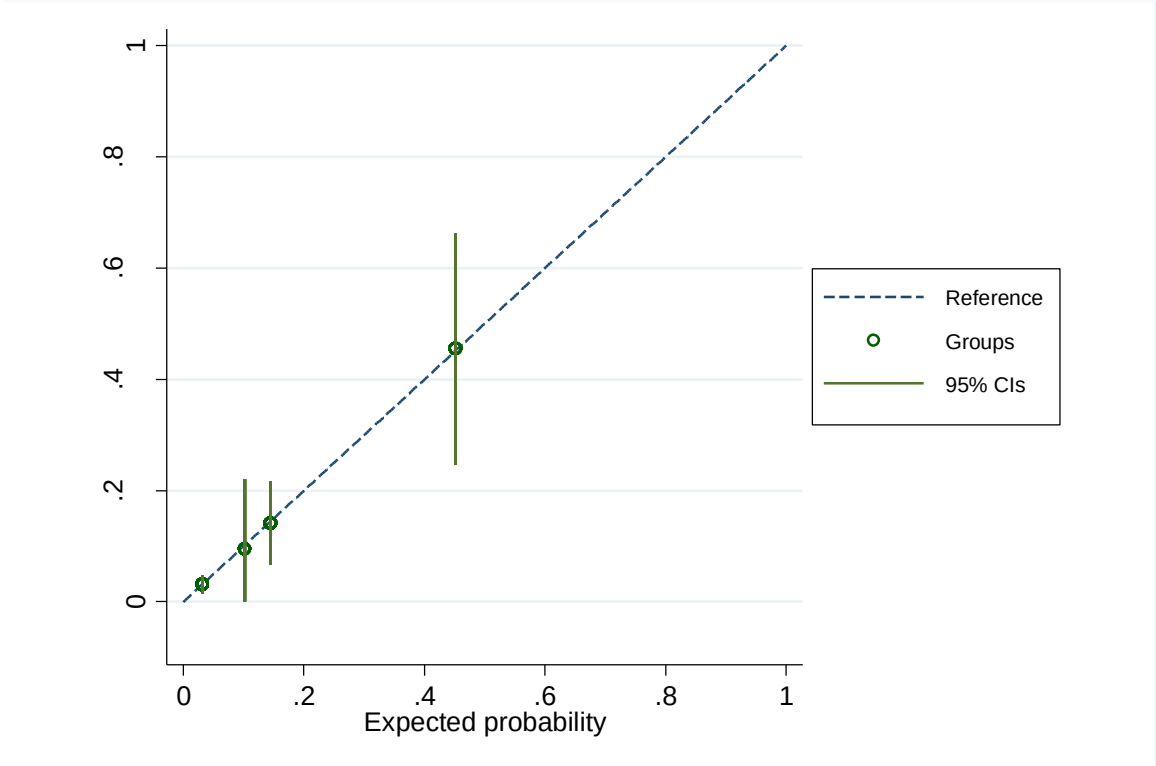
[d] Accuracy: correctly classified percentage

Figure 1. Calibration plots of predicted probabilities of ID from the "*parsimonious AAP model*" (A) and the "*new model*" (B) and observed outcome

A.



B.



Circles represent mean predicted probabilities versus observed proportions in subgroups defined by deciles of the predicted ID probabilities from the “complete” model and realized on the first imputed dataset. Vertical bars are 95% confidence intervals. Dashed diagonal line represents perfect calibration. Several infants have the same probability of having ID: eight (A) and six (B) “missing” deciles overlap with the first one.

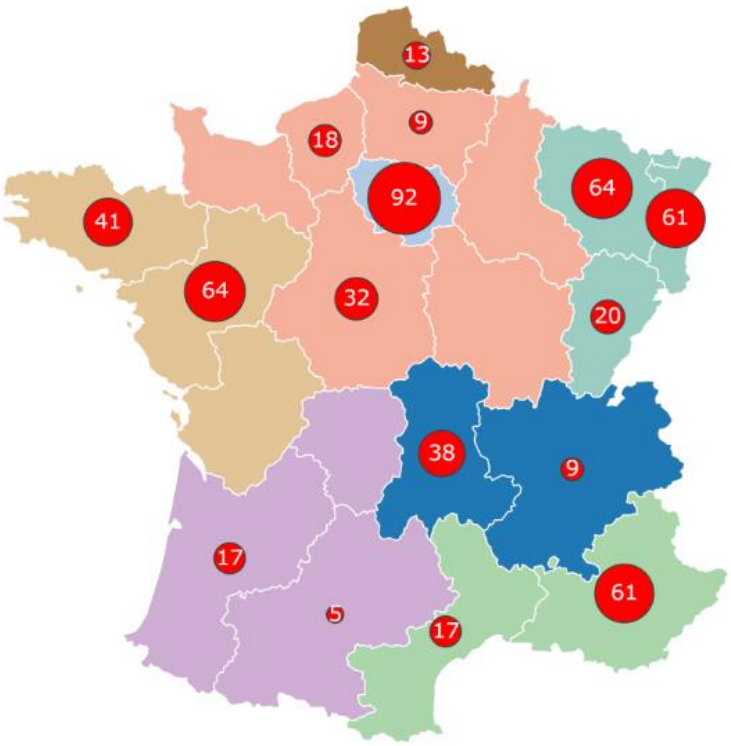
Appendix 1. Verbatim of the AAP and CPS recommendations for ID screening, AAP criteria adaptations needed for the present study, and list of other potential predictors of ID included in analyses

The verbatim of the recommendations of the AAP is as follow: "*Universal screening for anemia should be performed at approximately 12 months of age with determination of Hb concentration and an assessment of risk factors associated with ID/IDA. These risk factors would include low socio-economic status (especially children of Mexican American descent), a history of prematurity or low birth weight, exposure to lead, exclusive breastfeeding beyond 4 months of age without supplemental iron, and weaning to whole milk or complementary foods that do not include iron-fortified cereals or foods naturally rich in iron. Additional risk factors are the feeding problems, poor growth, and inadequate nutrition typically seen in infants with special health care needs*" (6).

AAP criterion (verbatim)	**Adaptation needed**
History of prematurity or low birth weight	No adaptation needed
Exclusive breastfeeding beyond four months without supplemental iron	No adaptation needed
Weaning to whole milk or complementary foods that do not include iron-fortified cereals or foods naturally rich in iron	No adaptation needed for "weaning to whole milk". Weaning "to complementary foods that do not include iron-fortified cereals or foods naturally rich in iron" could not be assessed in the present study due to lack of information on foods poor in iron used at weaning.
Low socio-economic status (especially children of Mexican American descent)	Needed an adaptation to the European context. We decided to study low socio-economic status with 4 potential risk factors: - parent's country of birth - parent's education level - parent's professional status - health coverage fully funded publicly
Poor growth	Needed clarification. We decided to define "poor growth" with the WHO weight for age z-score at one and two years old
Exposure to lead	Not relevant in the French context (see text). We did not consider this variable in the present study
Feeding problems Inadequate nutrition typically seen in infants with special health care needs	Not relevant in the study context that targeted apparently healthy infants. Infants with feeding problems or with special health care needs were excluded of the present study. We did not consider these variables in the present study.
CPS criteria (verbatim)	**Discussion**

Children living with chronic illness	Not relevant in the study context that targeted apparently healthy infants. Infants with chronic illness were excluded of the present analyses. We did not consider this variable in the present study.
Low socio-economic status	Common with AAP criteria
Suboptimal intake of iron-rich foods	We consider this variable by studying different nutritional variables (including those from AAP recommendations and other potential predictors from literature - see below).
Prolonged bottle feeding	This criterion was assessed by studying duration of consumption of infant formula, iron-rich formula (see below), milk protein hydrolysates, cow's milk iron-fortified infant cereals.
Potential predictors from literature included in the present study	**Additional information**
Single motherhood	(yes/no)
Rank among siblings	(<3 or ≥ 3rd child)
Mother's ID during pregnancy	(yes/no)
Iron supplementation since birth	(yes/no)
Infant formula: duration of consumption	(<4 or ≥ 4 months) – Infant formula is recommended for non-breastfed infants from birth to 6 months of age. In France, 4 months is the mean age of weaning from breastfeeding and the start of solid food.
"Iron-rich formula": duration of consumption	(<12 or ≥ 12 months) – Global variable that sums durations of consumption of "follow-on formula" (recommended from 6 to 10-12 months of age) and "young children formula" (recommended after 12 months of age).
Milk protein hydrolysates: duration of consumption	(<6 or ≥ 6 months)
Cow's milk: duration of consumption	(<12 or ≥ 12 months)
Iron-fortified infant cereals: duration of consumption	(<12 or ≥ 12 months)

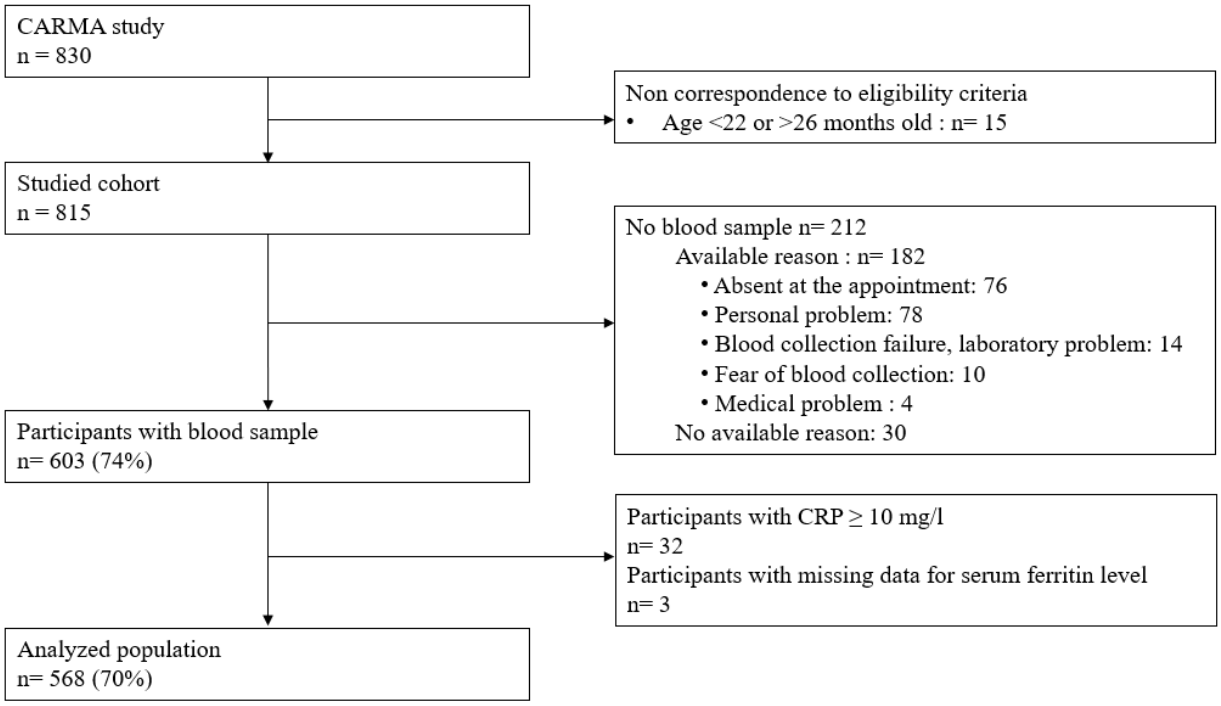
Appendix 2. Geographical distribution of children included, by French continental region



Appendix 3. Full list of the study exclusion criteria

CARMA study exclusion criteria
• Transfusion(s) since birth • Celiac disease • Inflammatory bowel disease • Cystic fibrosis • Other enteropathy (except allergy to cow's milk protein) • Enteral nutrition for more than 15 days within the last six months • Chronic hemolytic diseases (e.g. sickle cell disease) • Chronic kidney disease • Hemophilia • Hemochromatosis • Malignant diseases • Lead poisoning

Appendix 4. Flowchart of study participants



CRP: C-reactive protein

Appendix 5. Multiple imputations of missing data

Patterns of missing values for candidate predictors of ID (n=568)

Complete case analysis would have led to excluding 152/568 patients (26.8%)

Candidate predictors of ID	Number of missing values	Percentage
Sociodemographic characteristics		
Single motherhood	1	0.18
Both parents born outside the EU	5	0.88
Both parents who did not attend university	4	0.70
Both parents unemployed	4	0.70
Perinatal and medical history		
Mother's ID during pregnancy	32	5.63
Iron supplementation since birth	1	0.18
Weight z-score at one year	13	2.29
Feeding since birth		
Exclusive breastfeeding > 4 months of age without supplemental iron	2	0.35
Infant formula: duration of consumption (months)	2	0.35
Iron-rich formula: duration of consumption (months) [a]	18	3.17
Milk protein hydrolysates: duration of consumption (months)	1	0.18
Cow's milk: duration of consumption (months)	43	7.57
Iron-fortified infant cereals: duration of consumption (months)	69	12.15

[a] Global variable that sums durations of consumption of "follow on formula" and "young children formula"

Characteristics of patients with at least one missing value and complete cases

Candidate predictors of ID	Patients with at least one missing value (n=152)	Complete cases (n=416)	P[a]
ID [a]	15/152 (9.9)	23/416 (5.5)	0.09
Socio-demographic characteristics			
Single motherhood [a]	10/151 (6.6)	6/416 (1.4)	0.002
Both parents born outside the EU [a]	8/147 (5.4)	21/416 (5.1)	0.853
Both parents who did not attend university [a]	26/148 (17.6)	66/416 (15.9)	0.630
Both parents unemployed [a]	0/148 (0.0)	19/416 (4.6)	0.008
Health coverage fully funded publicly [a]	15/152 (9.9)	28/416 (6.7)	0.211
Perinatal and medical history			
Rank among siblings (≥ 3rd child) [a]	29/152 (19.1)	50/416 (12.0)	0.031
Mother's ID during pregnancy [a]	51/120 (42.5)	142/416 (34.1)	0.093
Iron supplementation since birth [a]	16/151 (10.6)	41/416 (9.9)	0.796
Term of birth (GW) [b]	39.3 (1.7)	39.1 (1.8)	0.239
Weight at birth (g) [b]	3257.7 (528.4)	3257.2 (519.4)	0.992
Weight z-score at one year [b]	0.03 (0.99)	0.04 (0.87)	0.889
Weight z-score at two years [b]	0.06 (0.91)	0.04 (0.89)	0.818
Feeding since birth			
Exclusive breastfeeding >4 months of age without supplemental iron [a]	40/150 (26.7)	83/416 (20.0)	0.106
Weaning to cow's milk without supplemental iron [a]	2/152 (1.3)	4/416 (1.0)	0.661
Infant formula: duration of consumption (months) [b]	4.0 (2.9)	4.2 (2.7)	0.383
Iron-rich formula: duration of consumption (months) [b, c]	16.4 (5.4)	16.1 (5.5)	0.596
Milk protein hydrolysates: duration of consumption (months) [b]	0.8 (3.4)	0.9 (3.7)	0.780
Cow's milk: duration of consumption (months) [b]	2.6 (4.9)	2.7 (2.3)	0.821
Iron-fortified infant cereals: duration of consumption (months) [b]	8.3 (8.7)	11.1 (8.3)	0.006

Data are no./N of patients (%) or mean (SD).

[a] Chi-square or Fisher's exact test

[b] Student's *t* test

[c] Global variable that sums durations of consumption of "follow-on formula" and "young children formula"

GW: Weeks of gestation

Multiple imputations of missing data

We used multiple imputations with chained equations (MICE) because of missing values in almost all clinical potential predictors of ID we identified (26). MICE was performed in STATA 13/SE (*m*=25). Analyses were repeated in each dataset and estimates of interest were combined using Rubin's rules (27).

Convergence of imputation diagnostics

Convergence of the MICE algorithm was checked by graphically investigating the trends in the means and standard deviations of the imputed values over 100 iterations. The trace plots did not show apparent trends in the summaries of the imputed values, so the number of burn-in iterations was set to 25. The distribution of imputed and observed values were compared graphically. Distributional plots did not show significant departure between imputed values and observed values. We concluded that the fit of the imputation model was good.

Appendix 6. Characteristics of excluded and included patients

Characteristics	Excluded (n=247)	Included (n=568)	P
Sex (boy) [a]	129/247 (52%)	291/568 (51%)	0.79
Age [b]	24.26 (0.59)	24.25 (0.60)	0.79
Socio-demographic characteristics			
Single motherhood [a]	10/247 (4%)	16/568 (3%)	0.36
Number of parents born outside the EU (one or both) [a]	32/246 (13%)	84/568 (15%)	0.51
Number of parents who did not go to the university (one or both) [a]	117/243 (48%)	268/568 (47%)	0.80
Number of parents unemployed (one or both) [a]	68/242 (28%)	140/568 (25%)	0.30
Health coverage fully funded publicly [a]	22/247 (9%)	43/568 (8%)	0.52
Perinatal and medical history			
Rank among siblings (≥ 3rd child) [a]	35/247 (14%)	79/568 (14%)	0.92
Mother's ID during pregnancy [a]	96/237 (41%)	206/568 (36%)	0.26
Iron supplementation since birth [a]	19/246 (8%)	58/568 (10%)	0.27
Term of birth (GW) [b]	39.30 (1.68)	39.11 (1.74)	0.17
Weight at birth (g) [b]	3282.15 (489.03)	3257.38 (521.40)	0.53
Feeding since birth			
Exclusive breastfeeding >4 months of age without supplemental iron [a]	55/245 (22%)	125/568 (22%)	0.89
Weaning to cow's milk without supplemental iron [a]	4/247 (2%)	6/568 (1%)	0.50
Duration of breastfeeding (months) [b]	5.13 (6.23)	4.58 (5.47)	0.21
Infant formula: duration of consumption (months) [b]	4.32 (2.87)	4.20 (2.74)	0.58
Iron-rich formula: duration of consumption (months) [b, c]	14.52 (6.22)	16.03 (5.52)	0.004
Milk protein hydrolysates: duration of consumption (months) [b]	0.34 (2.05)	0.84 (3.58)	0.04
Cow's milk: duration of consumption (months) [b]	1.60 (3.97)	3.00 (4.75)	<0.001
Iron-fortified infant cereals: duration of consumption (months) [b]	7.46 (7.87)	11.01 (8.31)	<0.001

Data are no./N of patients (%) or mean (SD).

[a] Chi-square or Fisher's exact test

[b] Student's *t* test. The results presented in this table were obtained on the first imputed dataset;

[c] Global variable that sums durations of consumption of "follow-on formula" and "young children formula"

GW: Weeks of gestation

Appendix 7. Characteristics of participants (n=568), univariable associations with iron deficiency (serum ferritin, SF <12 µG/L) and diagnostic accuracy

Characteristics		Iron deficiency (SF <12 µg/l)	Diagnostic accuracy	
	Proportion [a] (%)	Unadjusted OR (95% CI)	Sensitivity, % (95% CI)	Specificity, % (95% CI)
Sociodemographic characteristics				
Single motherhood				
No	97.2	Ref.		
Yes	2.8	2.04 (0.45 – 9.31)	5.3 (0.0 – 12.4)	97.3 (96.0 – 98.7)
* Both parents born outside the EU				
No	94.8	Ref.		
Yes	5.2	4.07 (1.55 – 10.69)	15.8 (4.2 – 27.4)	95.6 (93.8 – 97.3)
* Both parents who did not attend university				
No	83.4	Ref.		
Yes	16.6	1.62 (0.74 – 3.54)	23.7 (10.2 – 37.2)	83.9 (80.8 – 87.0)
* Both parents unemployed				
No	96.5	Ref.		
Yes	3.5	0.72 (0.09 – 5.55)	2.6 (0.0 – 7.7)	96.4 (94.8 – 98.0)
* Having a health coverage fully funded publicly [b]				
No	92.4	Ref.		
Yes	7.6	1.48 (0.50 – 4.39)	10.5 (0.8 – 20.3)	92.6 (90.4 – 94.9)
Perinatal and medical history				
Rank among siblings				
1st or 2nd child	86.1	Ref.		
≥ 3rd child	13.9	2.04 (0.93 – 4.49)	23.7 (10.2 – 37.2)	86.8 (83.9 – 89.7)
Mother's ID during pregnancy				
No	63.8	Ref.		
Yes	36.2	2.36 (1.19 – 4.68)	55.8 (39.6 – 71.9)	65.2 (61.0 – 69.3)
Iron supplementation since birth				
Yes	10.1	Ref.		
No	89.9	2.10 (0.49 – 8.94)	94.7 (87.6 – 100)	10.4 (7.8 – 13.0)
* Term of birth (WG)				
≥ 37	92.3	Ref.		
<37	7.7	1.02 (0.30 – 3.47)	7.9 (0.0 – 16.5)	92.3 (90.0 – 94.5)
* Birth weight (g)				
≥ 2500	91.9	Ref.		
<2500	8.1	1.37 (0.46 – 4.04)	10.5 (0.1 – 20.3)	92.1 (89.8 – 94.4)
* Low weight at one year old (weight for age z-score <-2 SD)				
No	98.5	Ref.		
Yes	1.5	4.67 (0.91 – 24.00)	5.3 (0.0 – 12.4)	98.8 (97.9 – 99.8)
* Low weight at two years old (weight for age z-score <-2 SD)				
No	98.8	Ref.		
Yes	1.2	2.36 (0.28 – 20.12)	2.6 (0.0 – 7.7)	98.9 (98.0 – 99.8)

Appendix 7 (*continued*)

Characteristics		Iron deficiency (SF <12 µg/l)	Diagnostic accuracy	
	Proportion [a] (%)	Unadjusted OR (95% CI)	Sensitivity, % (95% CI)	Specificity, % (95% CI)
Feeding since birth				
* "Exclusive" breastfeeding beyond 4 months without iron supplementation				
No	78.1	Ref.		
Yes	21.9	1.95 (0.97 – 3.93)	34.2 (19.1 – 49.3)	78.9 (75.5 – 82.4)
Infant formula: duration of consumption				
≥ 4 months	53.5	Ref.		
<4 months	46.5	1.30 (0.67 – 2.52)	52.6 (36.8 – 68.5)	54.0 (49.7 – 58.2)
"Iron-rich formula": duration of consumption [c]				
≥ 12 months	82.5	Ref.		
<12 months	17.5	7.17 (3.61 -14.23)	55.4 (39.5 – 71.2)	85.2 (82.2 – 88.3)
Milk protein hydrolysates: duration of consumption				
≥ 6 months	5.5	Ref.		
<6 months	94.5	1.03 (0.24 – 4.47)	5.3 (0.0 – 12.4)	94.5 (92.6 – 96.5)
Cow's milk: duration of consumption				
<12 months	90.0	Ref.		
≥ 12 months	10.0	4.02 (1.79 – 9.03)	28.0 (12.9 – 42.9)	91.2 (88.6 – 93.8)
Iron-fortified infant cereals: duration of consumption				
≥ 12 months	52.7	Ref		
<12 months	47.3	1.95 (0.95 – 4.01)	62.6 (44.6 – 78.9)	53.8 (49.3 – 58.3)
* Weaning to cow's milk without supplemental iron				
No	98.9	Ref.		
Yes	1.1	80.15 (9.10 – 705.96)	13.2 (0.2 – 23.9)	99.8 (99.4 – 100)

[a] After performing statistical analyses separately in each imputed dataset (m=25); estimated parameters were pooled using Rubin's rules. It explains why we presented proportions instead of number of participants

[*] Criteria recommended by AAP are marked with an asterisk to differentiate them from those suggested by the literature

[b] Health coverage fully funded publicly; an indicator strongly linked to poverty in France

[c] Global variable that sums durations of consumption of "follow-on formula" (recommended at age 6-12 months) and "young children formula" (recommended after 12 months of age)

WG: weeks of gestation

Appendix 8. Selection of predictor variables in 200 bootstrap backwards stepwise selection procedures across 25 imputed datasets (N=568)

A. Selection among all AAP potential predictor variables (binary) adapted to the European context: implementation of the *"parsimonious AAP model"*

	Number of times each predictor was selected in the 200 bootstraps by the 25 imputed datasets																										
Candidate predictor	**Imputed dataset**																										
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**Total (%)**	**In final model**
Both parents born outside the EU	169	173	172	174	184	168	168	168	171	184	168	177	163	174	172	173	180	173	163	176	163	159	177	167	167	4283 (85.7)	**Yes**
Both parents unemployed	34	30	31	26	33	38	26	30	42	33	30	33	27	38	42	31	33	31	33	31	32	32	26	29	26	797 (15.9)	No
Both parents who did not attend university	76	80	69	73	76	67	76	74	84	67	72	70	64	68	72	74	63	64	68	67	74	81	64	84	73	1800 (36.0)	No
Health coverage fully funded publicly [a]	53	49	43	49	52	35	49	48	53	50	49	45	49	58	36	51	46	38	51	54	46	45	44	49	45	1187 (23.7)	No
Prematurity [b]	38	27	34	32	38	38	34	36	39	40	33	32	38	36	35	33	38	36	38	36	28	28	40	40	38	885 (17.7)	No
Low birth weight [c]	29	21	35	19	30	24	17	33	26	28	33	24	20	29	26	27	28	26	28	24	24	27	20	29	28	655 (13.1)	No
Low weight at one year old [d]	134	124	126	122	123	115	111	114	111	124	114	116	115	127	132	123	116	114	125	132	123	120	123	133	117	3024 (60.5)	**Yes**
Low weight at two years old [d]	31	32	35	27	36	35	41	38	40	37	43	25	41	33	38	30	22	31	35	34	36	41	29	34	29	853 (17.1)	No
Breastfeeding > 4 months	36	34	38	46	35	37	32	43	35	39	40	30	43	32	34	35	35	45	36	41	36	33	33	40	42	930 (18.6)	No
Weaning to cow's milk without supplemental iron	71	79	83	72	83	81	64	71	89	76	72	73	58	72	70	90	78	68	73	81	74	76	77	72	73	1876 (37.5)	**Yes** [e]

[a] Health coverage publicly, an indicator strongly linked to poverty in France

[b] Prematurity, defined as <37 weeks of gestation in this study

[c] Low birth weight, defined as <2500 g

[d] Low weight at one and two year(s) old, defined as weight for age z-score <-2

[e] Weaning to cow's milk without supplemental iron was forced in the model

B. Selection among all binary predictor variables: implementation of the "*new model*".

	Number of times each predictor was selected in the 200 bootstraps by the 25 imputed datasets																										
Candidate predictor	**Imputed dataset**																										
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**Total (%)**	**In final model**
Single motherhood	84	71	84	83	86	80	84	93	74	66	62	53	87	70	78	72	83	77	90	68	91	83	71	74	65	1929 (38.6)	No
Both parents born outside the EU	164	160	155	160	152	149	158	153	150	156	152	156	159	154	151	158	155	155	152	168	159	151	155	144	158	3884 (77.7)	**Yes**
Both parents unemployed	35	33	24	14	28	24	21	20	27	19	25	26	39	18	26	20	24	25	27	35	22	21	24	27	27	631 (12.6)	No
Both parents who did not attend university	45	46	53	43	44	42	49	38	57	58	48	38	38	41	43	50	41	50	46	55	41	47	57	39	38	1147 (22.9)	No
Health coverage fully funded publicly [a]	95	94	107	88	87	102	107	104	107	113	102	76	83	91	83	94	99	89	94	94	97	79	97	80	76	2338 (46.8)	No
Rank among siblings: ≥ 3rd child	60	71	60	67	62	78	58	59	52	59	79	71	64	64	59	64	63	63	66	57	68	80	63	65	65	1617 (32.3)	No
Mother's ID during pregnancy	86	84	93	79	80	77	113	66	127	98	71	110	91	84	100	108	75	109	92	88	93	84	68	106	118	2300 (46.0)	No
Iron supplementation since birth	27	23	23	24	38	18	26	18	20	15	19	28	22	18	21	24	14	16	24	18	21	21	16	23	25	542 (10.8)	No
Prematurity [b]	37	40	39	41	50	34	37	33	38	32	36	47	48	42	41	41	46	42	38	39	38	41	35	40	43	998 (20.0)	No
Low birth weight [c]	46	39	32	30	38	34	42	39	34	33	46	34	41	24	32	39	29	37	40	40	47	28	44	46	30	924 (18.5)	No
Low weight at one year old [d]	124	129	134	126	135	124	127	150	125	120	113	115	126	121	123	109	127	121	131	125	130	123	138	116	120	3132 (62.6)	**Yes**
Low weight at two years old [d]	30	46	28	40	22	34	54	36	34	40	37	28	29	29	37	38	38	49	27	25	29	46	38	53	44	911 (18.2)	No
Breastfeeding > 4 months	51	57	38	52	46	51	56	41	41	51	59	60	57	59	61	55	52	62	52	54	60	56	56	57	48	1332 (26.6)	No
Consumption of "infant formula" <4 months	69	45	59	71	68	51	52	59	63	60	58	71	58	62	58	55	57	66	65	67	59	63	65	64	69	1534 (30.7)	No
Consumption of "iron-rich formula" <12 months [e]	199	196	195	199	192	197	192	195	197	198	198	198	199	198	198	194	199	195	199	199	198	197	199	196	199	4926 (98.5)	**Yes**
Consumption of milk protein hydrolysates <6 months	48	88	56	72	49	63	45	55	61	75	76	60	71	64	82	58	83	63	75	83	62	59	63	62	67	1640 (32.8)	No
Consumption of cow's milk ≥ 12 months	61	49	108	51	81	60	123	89	85	71	54	43	38	48	42	68	42	48	46	40	55	39	46	62	43	1492 (29.8)	No
Consumption of cereals <12 months	73	61	52	44	37	58	60	37	65	85	40	42	50	55	39	49	57	45	45	35	37	46	47	39	35	1233 (24.7)	No
Weaning to cow's milk without supplemental iron	80	74	67	69	76	75	74	69	66	79	80	79	83	61	70	71	69	72	80	74	67	60	63	71	75	1804 (36.1)	**Yes** [f]

[a] Health coverage publicly, an indicator strongly linked to poverty in France

[b] Prematurity, defined as <37 weeks of gestation in this study

[c] Low birth weight, defined as <2500 g

[d] Low weight at one and two year(s) old, defined as weight for age z-score <-2

[e] Global variable that sums durations of consumption of "follow on formula" and "young children formula"

[f] Weaning to cow's milk without supplemental iron was forced in the model

Appendix 9. Sensitivity analyses: diagnostic accuracy of AAP criteria using additive or complete combination (N=568) [a]

Tool	Number of predictor variables	AUROC: Estimate, (95% CI) Validation [b], (95% CI)	Positivity threshold [c]	Sensitivity: Estimate, % (95% CI) Validation [b], % (95% CI)	Specificity: Estimate, % (95% CI) Validation [b], % (95% CI)	Accuracy [d]: Estimate, % (95% CI) Validation [b], % (95% CI)
"Simple additive AAP model"	10	0.62 (0.53 – 0.72) 0.58 (0.52 – 0.65)	≥ 2 criteria	34 (19 – 49) 34 (21 – 47)	83 (80 – 86) 83 (78 – 88)	80 (76 – 83) 80 (75 – 85)
"Complete AAP model"	10	0.66 (0.57 – 0.76) 0.63 (0.54 – 0.71)	Predicted probability of ID ≥ 0.05	53 (37 – 69) 41 (26 – 56)	77 (73 – 80) 81 (72 – 90)	75 (72 – 79) 78 (66 – 90)

"*Simple additive AAP model*", combination of the 10 adapted AAP criteria (score range between 0 and 10, assigning one point for each AAP adapted criterion).
"*Complete AAP model*" included all of the 10 AAP adapted criteria (score range between 0 and 1, obtained as the individual risk prediction from a logistic regression model where all AAP criteria were forced to enter the model).

[a] After performing statistical analyses separately in each imputed dataset (m=25); estimated parameters were pooled using Rubin's rules

[b] Internal validation was performed by bootstrap (*b*=500 per multiple imputation dataset)

[c] The threshold was set to obtain a specificity of at least 75%

[d] Accuracy: correctly classified percentage

Appendix 10. Additional diagnostic accuracy of clinical prediction tools under evaluation (n=568)

A. Positive predictive value and negative predictive value of clinical prediction tools under evaluation

Tool	**Number of predictor variables**	**Positivity threshold** [a]	**PPV** [b]**, % (95% CI)**	**NPV** [c]**, % (95% CI)**
"Parsimonious AAP model"	3	Predicted probability of ID ≥ 0.18	28 (15 – 42)	95 (93 – 97)
"New model"	4	Predicted probability of ID ≥ 0.10	19 (12 – 26)	97 (95 – 98)
"Simple prediction tool"	4	≥ 1 criterion	19 (12 – 26)	97 (95 – 98)
"Simple additive AAP model"	10	≥ 2 criteria	12 (6 – 19)	95 (93 – 97)
"Complete AAP model"	10	Predicted probability of ID ≥ 0.05	14 (8 – 20)	96 (94 – 98)

"*Parsimonious AAP model*" included the 3 AAP adapted criteria retained in at least 60% of the bootstrap models, i.e. both parents born outside the EU, low weight at one year old and weaning to cow's milk without supplemental iron (score range between 0 and 1, obtained as the individual risk prediction from a logistic regression model which included these 3 criteria).

"*New model*" included the 4 criteria retained in at least 60% of the bootstrap models, i.e. both parents born outside the EU, low weight at one year old, duration of consumption of "iron-rich formula" <12 months and weaning to cow's milk without supplemental iron (score range between 0 and 1, obtained as the individual risk prediction from a logistic regression model which included these 4 criteria)

"*Simple prediction tool*" included the 4 criteria mentioned in the "*New Model*" (binary score: no criterion vs at least 1 criterion)

"*Simple additive AAP model*", combination of the 10 adapted AAP criteria (score range between 0 and 10, assigning one point for each AAP adapted criterion).

"*Complete AAP model*" included all of the 10 AAP adapted criteria (score range between 0 and 1, obtained as the individual risk prediction from a logistic regression model where all AAP criteria were forced to enter the model).

[a] The threshold was set to obtain a specificity of at least 75%

[b] PPV: positive predictive value

[c] NPV: negative predictive value

B. Positive likelihood ratio and negative likelihood ratio of clinical prediction tools under evaluation

Tool	Number of predictor variables	Positivity threshold [a]	LR+ [b] (95% CI)	LR- [c] (95% CI)	1/LR-
"Parsimonious AAP model"	3	Predicted probability of ID ≥ 0.18	6.00 (3.26 – 11.06)	0.74 (0.60 – 0.91)	1.35
"New model"	4	Predicted probability of ID ≥ 0.10	3.27 (2.41 – 4.43)	0.47 (0.31 – 0.70)	2.13
"Simple prediction tool"	4	≥ 1 criterion	3.31 (2.45 – 4.47)	0.46 (0.30 – 0.69)	2.17
"Simple additive AAP model"	10	≥ 2 criteria	2.00 (1.23 – 3.23)	0.80 (0.63 – 1.00)	1.25
"Complete AAP model"	10	Predicted probability of ID ≥ 0.05	2.16 (1.42 – 3.29)	0.73 (0.56 – 0.95)	1.37

"*Parsimonious AAP model*" included the 3 AAP adapted criteria retained in at least 60% of the bootstrap models, i.e. both parents born outside the EU, low weight at one year old and weaning to cow's milk without supplemental iron (score range between 0 and 1, obtained as the individual risk prediction from a logistic regression model which included these 3 criteria).

"*New model*" included the 4 criteria retained in at least 60% of the bootstrap models, i.e. both parents born outside the EU, low weight at one year old, duration of consumption of "iron-rich formula" <12 months and weaning to cow's milk without supplemental iron (score range between 0 and 1, obtained as the individual risk prediction from a logistic regression model which included these 4 criteria)

"*Simple prediction tool*" included the 4 criteria mentioned in the "*New Model*" (binary score: no criterion vs at least 1 criterion)

"*Simple additive AAP model*", combination of the 10 adapted AAP criteria (score range between 0 and 10, assigning one point for each AAP adapted criterion).

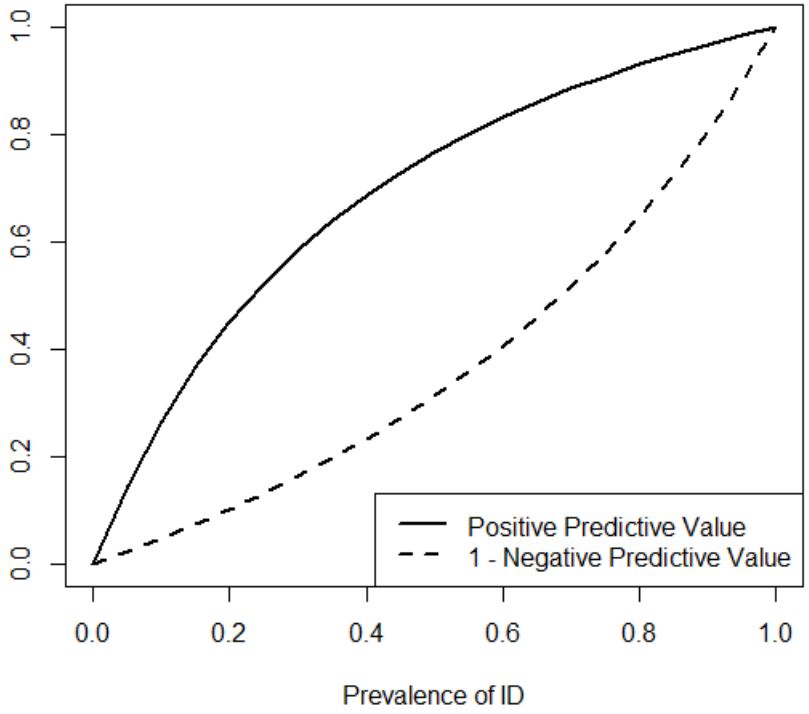
"*Complete AAP model*" included all of the 10 AAP adapted criteria (score range between 0 and 1, obtained as the individual risk prediction from a logistic regression model where all AAP criteria were forced to enter the model).

[a] The threshold was set to obtain a specificity of at least 75%

[b] LR+: positive likelihood ratio

[c] LR-: negative likelihood ratio

Appendix 11. *Simple prediction tool*: estimations of the predictive positive value and of the (1 – negative predictive value) for settings with different prevalences of ID, a sensitivity of 63% and a specificity of 81%



Appendix 12. Sensitivity analyses: using different thresholds to dichotomize durations of consumption of infant formula and iron rich formula

Appendix 12.1. Characteristics of participants (n=568), univariable associations with ID (serum ferritin, SF <12 µg/L) and diagnostic accuracy

Characteristics		Iron deficiency (SF <12 µg/l)	Diagnostic accuracy	
	Proportion [a] (%)	Unadjusted OR (95% CI)	Sensitivity, % (95% CI)	Specificity, % (95% CI)
Infant formula: duration of consumption				
≥ 3 months	64.1	Ref.		
<3 months	35.9	1.48 (0.76 – 2.88)	44.7 (28.9 – 60.5)	64.7 (60.6 – 68.8)
Infant formula: duration of consumption				
≥ 4 months	53.5	Ref.		
<4 months	46.5	1.30 (0.67 – 2.52)	52.6 (36.8 – 68.5)	54.0 (49.7 – 58.2)
Infant formula: duration of consumption				
≥ 5 months	45.1	Ref.		
<5 months	54.9	1.14 (0.58 – 2.21)	58.0 (42.2 – 73.6)	45.3 (41.0 – 49.5)
"Iron-rich formula": duration of consumption [b]				
≥ 11 months	83.1	Ref.		
<11 months	16.9	7.50 (3.78 – 14.89)	55.3 (39.4 – 71.1)	85.9 (82.9 – 88.9)
"Iron-rich formula": duration of consumption [b]				
≥ 12 months	82.5	Ref.		
<12 months	17.5	7.17 (3.61 – 14.23)	55.4 (39.5 – 71.2)	85.2 (82.2 – 88.3)
"Iron-rich formula": duration of consumption [b]				
≥ 13 months	76.0	Ref.		
<13 months	24.0	6.40 (3.20 – 12.81)	63.3 (47.9 – 78.6)	78.8 (75.3 – 82.3)

[a] After performing statistical analyses separately in each imputed dataset (m=25); estimated parameters were pooled using Rubin's rules. It explains why we presented proportions instead of number of participants

[b] Global variable that sums durations of consumption of "follow-on formula" (recommended between age 6-12 months) and "young children formula" (recommended after 12 months of age)

Appendix 12.2. Selection of predictor variables in 200 bootstrap backwards stepwise selection procedures across 25 imputed datasets (N=568), among all binary predictor variables: implementation of the "*new model*"

A. Infant formula – duration of consumption: dichotomized at 3 months

	Number of times each predictor was selected in the 200 bootstraps by the 25 imputed datasets																										
Candidate predictor	**Imputed dataset**																										
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**Total (%)**	**In final model**
Single motherhood	79	72	79	81	83	84	81	88	75	63	71	70	66	65	75	87	82	67	80	93	82	70	66	76	70	1905 (38.1%)	No
Both parents born outside the EU	162	161	154	166	146	153	150	157	145	161	147	153	148	160	157	150	155	151	156	165	156	148	162	155	152	3870 (77.4%)	**Yes**
Both parents unemployed	30	24	19	36	20	29	18	22	24	21	23	13	23	25	19	18	32	26	28	22	23	27	26	13	33	594 (11.9%)	No
Both parents who did not attend university	56	60	59	46	40	53	43	47	50	33	43	43	38	42	57	43	46	47	45	52	50	43	41	39	50	1166 (23.3%)	No
Health coverage fully funded publicly [a]	112	87	114	97	104	91	118	96	104	112	71	96	91	85	92	91	97	82	92	116	98	85	90	106	85	2412 (48.2%)	No
Rank among siblings: ≥3rd child	66	73	72	65	67	66	65	81	71	61	76	64	71	69	57	64	72	68	74	61	69	74	70	74	63	1713 (34.3%)	No
Mother's ID during pregnancy	107	78	97	82	96	82	101	60	117	98	81	108	89	79	96	123	74	104	103	98	92	83	75	109	102	2334 (46.7%)	No
Iron supplementation since birth	16	22	20	23	18	17	23	20	18	25	14	13	17	24	26	17	15	16	22	25	27	19	22	25	20	504 (10.1%)	No
Prematurity [b]	34	36	37	34	30	40	35	38	49	28	37	49	42	31	39	41	36	30	29	34	41	32	39	42	32	915 (18.3%)	No
Low birth weight [c]	38	41	37	31	45	43	39	40	39	48	34	36	47	33	31	35	32	29	33	45	36	50	38	40	36	956 (19.1%)	No
Low weight at one year old [d]	128	130	126	142	129	122	117	137	113	132	122	126	117	115	127	105	129	104	137	131	136	121	126	114	142	3128 (62.6%)	**Yes**
Low weight at two years old [d]	31	30	42	43	44	48	56	26	52	37	40	42	35	34	45	40	47	53	30	44	36	46	27	42	37	1007 (20.1%)	No
Breastfeeding >4 months	48	43	46	40	43	39	41	41	51	33	51	48	58	48	53	57	44	44	37	54	56	42	52	58	41	1168 (23.4%)	No
Consumption of "infant formula" <3 months	83	67	74	70	62	73	75	69	73	37	68	74	73	70	72	77	81	81	72	76	77	79	73	88	61	1805 (36.1%)	No
Consumption of "iron-rich formula" <12 months [e]	196	198	196	198	200	196	185	194	198	194	199	196	198	200	200	199	199	197	198	199	198	195	196	197	200	4926 (98.5%)	**Yes**
Consumption of milk protein hydrolysates <6 months	47	62	49	67	60	56	40	53	50	57	74	57	64	65	64	58	62	54	74	58	68	52	65	58	62	1476 (29.5%)	No
Consumption of cow's milk <12 months	56	51	98	47	72	60	122	84	77	78	44	49	38	45	57	63	48	41	36	45	45	55	53	63	50	1477 (29.5%)	No

Candidate predictor	Number of times each predictor was selected in the 200 bootstraps by the 25 imputed datasets																										
	Imputed dataset																										
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	Total (%)	In final model
Consumption of cereals <12 months	87	67	49	37	32	62	34	48	63	91	38	44	54	54	47	34	49	41	37	43	26	52	35	37	43	1204 (24.1%)	No
Weaning to cow's milk without supplemental iron	81	61	68	70	72	74	73	68	70	81	70	65	72	76	71	68	77	70	73	79	74	78	68	72	71	1802 (36.0%)	**Yes** [f]

[a] Health coverage publicly, an indicator strongly linked to poverty in France

[b] Prematurity, defined as <37 weeks of gestation in this study

[c] Low birth weight, defined as <2500 g

[d] Low weight at one and two year(s) old, defined as weight for age z-score <-2

[e] Global variable that sums durations of consumption of "follow on formula" and "young children formula"

[f] Weaning to cow's milk without supplemental iron was forced in the model

B. Infant formula – duration of consumption: dichotomized at 5 months

Candidate predictor	Number of times each predictor was selected in the 200 bootstraps by the 25 imputed datasets																										
	Imputed dataset																										
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**Total (%)**	**In final model**
Single motherhood	82	55	84	71	82	79	85	90	73	58	65	74	65	68	77	70	73	63	70	72	92	73	81	66	59	1827 (36.5%)	No
Both parents born outside the EU	162	155	151	157	154	162	149	152	164	158	153	148	163	169	149	146	163	147	163	157	156	154	163	139	148	3882 (77.6%)	**Yes**
Both parents unemployed	24	34	17	23	25	27	19	28	17	16	27	32	31	37	23	23	23	22	31	15	28	24	25	23	26	620 (12.4%)	No
Both parents who did not attend university	46	49	49	57	38	49	45	41	37	37	43	42	45	43	42	42	45	34	56	53	34	46	50	36	50	1109 (22.2%)	No
Health coverage fully funded publicly [a]	109	88	108	92	83	94	120	83	117	107	90	92	90	90	98	109	80	91	107	105	90	97	83	80	85	2388 (47.8%)	No
Rank among siblings: ≥3rd child	55	68	51	60	78	71	55	60	63	34	67	68	62	73	69	73	77	69	71	70	65	73	64	63	58	1617 (32.3%)	No
Mother's ID during pregnancy	103	87	84	83	98	69	107	63	117	98	75	110	81	68	97	111	67	101	75	89	81	76	62	93	104	2199 (44.0%)	No
Iron supplementation since birth	16	26	21	30	23	20	18	31	24	21	21	18	15	18	13	24	22	14	15	18	17	10	28	22	24	509 (10.2%)	No
Prematurity [b]	44	34	33	46	46	45	43	40	44	35	35	44	38	36	37	42	39	37	41	31	42	39	34	40	35	980 (19.6%)	No
Low birth weight [c]	40	35	38	43	3	36	43	35	39	35	38	37	39	25	39	36	31	29	32	34	33	39	29	37	31	856 (17.1%)	No
Low weight at one year old [d]	128	125	131	123	118	127	111	128	121	136	117	113	129	131	124	100	130	105	126	128	123	118	138	115	129	3074 (61.5%)	**Yes**
Low weight at two years old [d]	30	29	35	23	31	36	44	28	34	37	32	28	22	30	26	35	39	42	35	24	31	38	33	50	22	814 (16.2%)	No
Breastfeeding >4 months	51	58	44	58	49	59	59	35	57	55	68	50	65	54	65	54	48	60	47	51	69	48	54	53	54	1365 (27.3%)	No
Consumption of "infant formula" <5 months	55	77	55	71	77	71	75	64	69	60	32	81	53	76	66	70	73	71	64	71	53	66	71	65	79	1665 (33.3%)	No
Consumption of "iron-rich formula" <12 months [e]	197	200	198	199	199	198	195	196	196	194	196	198	199	200	200	199	200	199	199	197	199	197	196	197	198	4946 (98.9%)	**Yes**
Consumption of milk protein hydrolysates <6 months	57	64	55	77	61	70	42	56	62	58	71	59	88	70	76	56	37	57	65	67	79	72	59	63	88	1609 (32.2%)	No
Consumption of cow's milk <12 months	59	36	105	49	80	46	125	87	78	67	43	47	56	47	58	66	41	48	47	49	48	41	42	52	45	1462 (29.2%)	No
Consumption of cereals <12 months	70	62	45	41	42	65	55	43	67	80	34	35	51	46	49	54	53	48	44	41	35	44	46	39	38	1277 (25.5%)	No
Weaning to cow's milk without supplemental iron	71	77	73	67	70	77	69	74	67	75	70	79	69	78	78	71	59	71	78	78	76	62	71	68	57	1785 (35.7%)	**Yes** [f]

[a] Health coverage publicly, an indicator strongly linked to poverty in France

[b] Prematurity, defined as <37 weeks of gestation in this study

[c] Low birth weight, defined as <2500 g

[d] Low weight at one and two year(s) old, defined as weight for age z-score <-2

[e] Global variable that sums durations of consumption of “follow on formula” and “young children formula”

[f] Weaning to cow’s milk without supplemental iron was forced in the model

Appendix 12.3. Selection of predictor variables in 200 bootstrap backwards stepwise selection procedures across 25 imputed datasets (N=568), among all binary predictor variables: implementation of the "*new model*"

A. Iron rich formula – duration of consumption: dichotomized at 11 months

	Number of times each predictor was selected in the 200 bootstraps by the 25 imputed datasets																										
Candidate predictor	**Imputed dataset**																										
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**Total (%)**	**In final model**
Single motherhood	79	68	88	67	77	81	85	90	75	66	75	73	63	61	74	77	83	74	85	65	85	71	77	65	68	1872 (37.4%)	No
Both parents born outside the EU	150	153	144	157	157	157	142	150	153	167	150	153	158	154	159	161	159	156	155	154	160	161	162	145	149	3866 (77.3%)	**Yes**
Both parents unemployed	20	22	19	23	24	21	15	20	15	17	21	27	34	26	23	25	20	26	25	24	21	19	26	22	30	565 (11.3%)	No
Both parents who did not attend university	39	52	55	42	44	53	47	43	46	55	37	43	41	36	45	58	43	45	48	57	53	51	52	41	45	1171 (23.4%)	No
Health coverage fully funded publicly [a]	100	92	101	99	94	92	108	84	107	112	97	74	71	82	82	104	90	97	112	94	100	92	104	85	77	2350 (47.0%)	No
Rank among siblings: ≥3rd child	68	85	68	63	68	60	57	57	54	63	77	73	69	63	69	59	79	67	83	66	81	73	67	66	65	1700 (34.0%)	No
Mother's ID during pregnancy	95	88	99	81	96	91	11	71	117	94	87	111	88	83	96	112	74	107	112	94	101	88	65	103	114	2278 (45.6%)	No
Iron supplementation since birth	17	21	23	16	20	16	23	28	24	23	21	113	23	17	22	18	27	15	22	15	20	20	24	17	16	601 (12.0%)	No
Prematurity [b]	47	43	42	33	34	36	44	37	27	38	36	41	37	38	45	34	38	39	36	37	34	39	31	29	39	934 (18.7%)	No
Low birth weight [c]	29	32	33	41	34	39	45	33	45	32	37	33	32	36	32	36	32	34	35	40	44	44	40	34	42	914 (18.3%)	No
Low weight at one year old [d]	126	129	127	134	126	107	117	148	137	133	125	125	123	137	113	91	120	106	139	125	128	119	121	128	127	3111 (62.2%)	**Yes**
Low weight at two years old [d]	32	39	33	28	29	30	44	32	42	37	32	38	31	30	38	56	41	49	36	30	27	43	25	48	43	913 (18.3%)	No
Breastfeeding >4 months	55	62	44	54	55	63	50	51	46	65	61	63	61	74	55	57	28	58	54	67	67	66	60	43	61	1420 (28.4%)	No
Consumption of "infant formula" <4 months	64	65	63	50	50	63	68	55	60	54	55	65	62	62	62	63	47	71	62	56	61	57	44	65	60	1484 (29.6%)	No
Consumption of "iron-rich formula" <11 months [e]	194	197	193	200	196	197	196	197	190	195	198	197	199	199	197	199	197	198	200	199	200	200	197	199	199	4933 (98.7%)	**Yes**
Consumption of milk protein hydrolysates <6 months	52	96	56	75	67	69	45	61	45	61	71	70	76	78	61	63	79	80	81	74	90	81	75	69	78	1753 (35.1%)	No
Consumption of cow's milk <12 months	71	53	95	49	65	52	123	83	115	67	55	48	41	41	49	55	55	41	41	52	44	43	48	49	59	1494 (29.9%)	No

Candidate predictor	Number of times each predictor was selected in the 200 bootstraps by the 25 imputed datasets																										
	Imputed dataset																										
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**Total (%)**	**In final model**
Consumption of cereals <12 months	74	49	61	32	35	63	62	32	37	80	45	48	58	51	53	44	41	50	34	37	31	44	35	35	42	1173 (23.5%)	No
Weaning to cow's milk without supplemental iron	77	74	74	70	71	66	77	64	76	73	84	67	76	65	65	76	63	78	78	68	73	69	78	74	69	1805 (36.1%)	**Yes** [f]

[a] Health coverage publicly, an indicator strongly linked to poverty in France

[b] Prematurity, defined as <37 weeks of gestation in this study

[c] Low birth weight, defined as <2500 g

[d] Low weight at one and two year(s) old, defined as weight for age z-score <-2

[e] Global variable that sums durations of consumption of "follow on formula" and "young children formula"

[f] Weaning to cow's milk without supplemental iron was forced in the model

B. Iron rich formula – duration of consumption: dichotomized at 13 months

Candidate predictor	Number of times each predictor was selected in the 200 bootstraps by the 25 imputed datasets																										
	Imputed dataset																										
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**Total (%)**	**In final model**
Single motherhood	96	80	102	83	102	98	101	99	101	83	81	97	94	75	93	98	82	92	97	88	86	80	75	81	85	2249 (45.0%)	No
Both parents born outside the EU	147	139	147	139	143	169	145	140	140	141	132	144	149	158	137	138	147	144	138	141	145	152	145	137	142	3599 (72.0%)	**Yes**
Both parents unemployed	17	26	16	21	16	31	10	30	13	14	27	21	24	34	29	21	24	29	18	28	30	25	25	22	27	578 (11.6%)	No
Both parents who did not attend university	48	51	49	43	42	52	48	45	48	40	46	50	46	49	47	49	44	42	35	48	43	38	50	45	43	1141 (22.8%)	No
Health coverage fully funded publicly [a]	73	68	82	66	78	80	88	77	86	92	75	69	66	64	69	82	60	73	61	81	72	65	80	71	58	1836 (36.7%)	No
Rank among siblings: ≥3rd child	66	63	54	61	69	77	56	66	69	60	69	65	65	67	46	69	72	63	71	53	64	66	64	58	74	1607 (32.1%)	No
Mother's ID during pregnancy	96	58	72	80	81	60	107	57	98	93	67	86	65	51	84	109	71	97	76	74	86	71	57	89	102	1987 (39.7%)	No
Iron supplementation since birth	19	17	15	13	29	17	21	23	25	26	15	15	18	17	17	17	20	20	14	18	20	21	19	21	23	480 (9.6%)	No
Prematurity [b]	60	56	48	53	60	45	56	58	58	61	50	63	62	53	40	55	68	51	51	54	44	49	55	52	47	1349 (27.0%)	No
Low birth weight [c]	40	27	43	33	37	41	58	36	36	39	32	27	41	29	29	34	30	36	32	28	30	48	28	35	37	886 (17.7%)	No
Low weight at one year old [d]	129	133	124	136	141	141	126	124	127	134	124	133	136	142	143	127	142	113	129	132	138	113	131	130	132	3280 (65.6%)	**Yes**
Low weight at two years old [d]	26	23	22	3	23	33	37	17	35	24	24	23	26	27	35	37	19	29	21	13	27	36	21	33	21	635 (12.7%)	No
Breastfeeding >4 months	56	54	48	57	62	52	53	47	45	51	53	50	46	52	58	55	45	58	45	42	43	62	47	48	47	1276 (25.5%)	No
Consumption of "infant formula" <4 months	62	69	60	66	60	55	53	59	55	65	53	54	49	62	64	61	63	49	60	50	48	60	31	59	62	1429 (28.6%)	No
Consumption of "iron-rich formula" <13 months [e]	194	200	194	198	197	197	191	196	198	194	198	200	200	200	195	196	200	199	199	200	198	199	196	194	200	4933 (98.7%)	**Yes**
Consumption of milk protein hydrolysates <6 months	46	62	35	41	58	55	39	37	38	54	58	62	54	66	40	44	65	47	51	44	51	36	51	47	52	1233 (24.7%)	No
Consumption of cow's milk <12 months	65	42	109	55	82	51	129	88	112	84	33	42	42	42	52	80	28	37	49	29	45	62	37	65	35	1495 (29.9%)	No
Consumption of cereals <12 months	73	64	65	37	33	79	67	47	75	84	48	48	48	55	56	34	60	61	54	52	42	59	52	36	44	1373 (27.5%)	No
Weaning to cow's milk without supplemental iron	69	71	81	66	73	73	62	70	67	69	70	68	71	84	85	73	73	76	62	70	69	82	72	74	73	1803 (36.1%)	**Yes** [f]

[a] Health coverage publicly, an indicator strongly linked to poverty in France

[b] Prematurity, defined as <37 weeks of gestation in this study

[c] Low birth weight, defined as <2500 g

[d] Low weight at one and two year(s) old, defined as weight for age z-score <-2

[e] Global variable that sums durations of consumption of “follow on formula” and “young children formula”

[f] Weaning to cow’s milk without supplemental iron was forced in the model

Appendix 12.4. Sensitivity analyses: additional diagnostic accuracy of clinical prediction tools under evaluation (n=568)

A. Sensitivity, specificity, positive predictive value and negative predictive value

Tool	Number of predictor variables	AUROC (95% CI)	Positivity threshold [b]	Sensitivity, % (95% CI)	Specificity, % (95% CI)	Accuracy [c], % (95% CI)	PPV [d], % (95% CI)	NPV [e], % (95% CI)
Iron rich formula – duration of consumption: cut off value of 11 months								
"*New model*"	4	0.74 (0.66 – 0.83)	Predicted probability of ID ≥ 0.10	63 (48 – 78)	81 (78 – 85)	80 (77 – 83)	20 (13 – 27)	97 (95 – 98)
"Simple prediction tool"	4	0.72 (0.64 – 0.80)	≥ 1 criterion	63 (48 – 78)	81 (78 – 85)	80 (77 – 83)	20 (13 – 27)	97 (95 – 98)
Iron rich formula – duration of consumption: cut off value of 13 months								
"New model"	4	0.76 (0.67 – 0.84)	Predicted probability of ID ≥ 0.12	66 (51 – 81)	78 (74 – 81)	77 (74 – 81)	18 (11 – 24)	97 (95 – 99)
"Simple prediction tool"	4	0.73 (0.66 – 0.81)	≥ 1 criterion	71 (57 – 86)	75 (71 – 79)	75 (71 – 78)	17 (11 – 23)	97 (96 – 99)

"*New model*" included the 4 criteria retained in at least 60% of the bootstrap models, i.e., both parents born outside the EU, low weight at one year old, duration of consumption of "iron-rich formula" <11 months and weaning to cow's milk without supplemental iron (score range between 0 and 1, obtained as the individual risk prediction from a logistic regression model which included these 4 criteria)

"*Simple prediction tool*" included the 4 criteria mentioned in the "*New Model*" (binary score: no criterion vs at least 1 criterion)

[a] After performing statistical analyses separately in each imputed dataset (m=25); estimated parameters were pooled using Rubin's rules

[b] The threshold was set to obtain a specificity of at least 75%

[c] Accuracy: correctly classified percentage

[d] PPV: positive predictive value

[e] NPV: negative predictive value

B. Positive likelihood ratio and negative likelihood ratio

Tool	Number of predictor variables	AUROC (95% CI)	Positivity threshold [a]	LR+ [b] (95% CI)	LR- [c] (95% CI)	1/LR-
Iron rich formula – duration of consumption: cut off value of 11 months						
"New model"	4	0.74 (0.66 – 0.83)	Predicted probability of ID ≥ 0.10	3.31 (2.45 – 4.47)	0.46 (0.30 – 0.69)	2.17
"Simple prediction tool"	4	0.72 (0.64 – 0.80)	≥ 1 criterion	3.31 (2.45 – 4.47)	0.46 (0.30 – 0.69)	2.17
Iron rich formula – duration of consumption: cut off value of 13 months						
"New model"	4	0.76 (0.67 – 0.84)	Predicted probability of ID ≥ 0.12	3.00 (2.27 – 3.97)	0.44 (0.28 – 0.68)	2.27
"Simple prediction tool"	4	0.73 (0.66 – 0.81)	≥ 1 criterion	2.84 (2.21 – 3.65)	0.39 (0.23 – 0.64)	2.5

"*New model*" included the 4 criteria retained in at least 60% of the bootstrap models, i.e., both parents born outside the EU, low weight at one year old, duration of consumption of "iron-rich formula" <11 months and weaning to cow's milk without supplemental iron (score range between 0 and 1, obtained as the individual risk prediction from a logistic regression model which included these 4 criteria)

"*Simple prediction tool*" included the 4 criteria mentioned in the "*New Model*" (binary score: no criterion vs at least 1 criterion)

[a] The threshold was set to obtain a specificity of at least 75%

[b] LR+: positive likelihood ratio

[c] LR-: negative likelihood ratio

Appendix 13. Sensitivity analyses: diagnostic accuracy of clinical prediction tools under evaluation (n=530), after excluding infants with a CRP > 5 mg/L [a]

A. Sensitivity, specificity, accuracy, positive predictive value and negative predictive value

Tool	**Number of predictor variables**	**AUROC (95% CI)**	**Positivity threshold [b]**	**Sensitivity, % (95% CI)**	**Specificity, % (95% CI)**	**Accuracy [c], % (95% CI)**	**PPV [d], % (95% CI)**	**NPV [e], % (95% CI)**
"Parsimonious AAP model"	3	0.63 (0.55 – 0.71)	Predicted probability of ID ≥ 0.18	32 (17 – 46)	94 (92 – 96)	89 (86 – 92)	29 (15 – 43)	95 (93 – 97)
"New model"	4	0.74 (0.66 – 0.83)	Predicted probability of ID ≥ 0.11	63 (48 – 79)	81 (77 – 84)	80 (76 – 83)	20 (13 – 28)	97 (95 – 98)
"Simple prediction tool"	4	0.72 (0.64 – 0.80)	≥ 1 criterion	63 (48 – 79)	81 (77 – 84)	80 (76 – 83)	20 (13 – 28)	97 (95 – 98)
"Simple additive AAP model"	10	0.62 (0.52 – 0.71)	≥ 2 criteria	34 (19 – 49)	82 (78 – 85)	78 (75 – 82)	13 (6 – 19)	94 (92 – 96)
"Complete AAP model"	10	0.66 (0.57 – 0.76)	Predicted probability of ID ≥ 0.05	50 (34 – 66)	75 (71 – 79)	73 (70 – 77)	13 (8 – 19)	95 (93 – 97)

"*Parsimonious AAP model*" included the 3 AAP adapted criteria retained in at least 60% of the bootstrap models, i.e. both parents born outside the EU, low weight at one year old and weaning to cow's milk without supplemental iron (score range between 0 and 1, obtained as the individual risk prediction from a logistic regression model which included these 3 criteria).

"*New model*" included the 4 criteria retained in at least 60% of the bootstrap models, i.e. both parents born outside the EU, low weight at one year old, duration of consumption of "iron-rich formula" <12 months and weaning to cow's milk without supplemental iron (score range between 0 and 1, obtained as the individual risk prediction from a logistic regression model which included these 4 criteria)

"*Simple prediction tool*" included the 4 criteria mentioned in the "*New Model*" (binary score: no criterion vs at least 1 criterion)

"*Simple additive AAP model*", combination of the 10 adapted AAP criteria (score range between 0 and 10, assigning one point for each AAP adapted criterion).

"*Complete AAP model*" included all of the 10 AAP adapted criteria (score range between 0 and 1, obtained as the individual risk prediction from a logistic regression model where all AAP criteria were forced to enter the model).

[a] After performing statistical analyses separately in each imputed dataset (m=25); estimated parameters were pooled using Rubin's rules

[b] The threshold was set to obtain a specificity of at least 75%

[c] Accuracy: correctly classified percentage

[d] PPV: positive predictive value

[e] NPV: negative predictive value

B. Positive likelihood ratio and negative likelihood ratio

Tool	Number of predictor variables	AUROC (95% CI)	Positivity threshold [a]	LR+ [b] (95% CI)	LR- [c] (95% CI)	1/LR-
"Parsimonious AAP model"	3	0.63 (0.55 – 0.71)	Predicted probability of ID ≥ 0.18	5.35 (3.02 – 9.48)	0.72 (0.58 – 0.90)	1.39
"New model"	4	0.74 (0.66 – 0.83)	Predicted probability of ID ≥ 0.11	3.31 (2.45 – 4.37)	0.46 (0.30 – 0.69)	2.17
"Simple prediction tool"	4	0.72 (0.64 – 0.80)	≥ 1 criterion	3.31 (2.45 – 4.37)	0.46 (0.30 – 0.69)	2.17
"Simple additive AAP model"	10	0.62 (0.52 – 0.71)	≥ 2 criteria	1.89 (1.17 – 3.05)	0.81 (0.64 – 1.00)	1.25
"Complete AAP model"	10	0.66 (0.57 – 0.76)	Predicted probability of ID ≥ 0.05	2.00 (1.41 – 2.84)	0.67 (0.48 – 0.92)	1.49

"*Parsimonious AAP model*" included the 3 AAP adapted criteria retained in at least 60% of the bootstrap models, i.e. both parents born outside the EU, low weight at one year old and weaning to cow's milk without supplemental iron (score range between 0 and 1, obtained as the individual risk prediction from a logistic regression model which included these 3 criteria).

"*New model*" included the 4 criteria retained in at least 60% of the bootstrap models, i.e. both parents born outside the EU, low weight at one year old, duration of consumption of "iron-rich formula" <12 months and weaning to cow's milk without supplemental iron (score range between 0 and 1, obtained as the individual risk prediction from a logistic regression model which included these 4 criteria)

"*Simple prediction tool*" included the 4 criteria mentioned in the "*New Model*" (binary score: no criterion vs at least 1 criterion)

"*Simple additive AAP model*", combination of the 10 adapted AAP criteria (score range between 0 and 10, assigning one point for each AAP adapted criterion).

"*Complete AAP model*" included all of the 10 AAP adapted criteria (score range between 0 and 1, obtained as the individual risk prediction from a logistic regression model where all AAP criteria were forced to enter the model).

[a] The threshold was set to obtain a specificity of at least 75%

[b] LR+: positive likelihood ratio

[c] LR-: negative likelihood ratio