SCIENTIFIC OPINION

EFSA Journal

ADOPTED: 3 July 2019

doi: 10.2903/j.efsa.2019.5780

Appropriate age range for introduction of complementary feeding into an infant's diet

EFSA Panel on Nutrition, Novel Foods and Food Allergens (NDA), Jacqueline Castenmiller, Stefaan de Henauw, Karen-Ildico Hirsch-Ernst, John Kearney, Helle Katrine Knutsen, Alexandre Maciuk, Inge Mangelsdorf, Harry J McArdle, Androniki Naska, Carmen Pelaez, Kristina Pentieva, Alfonso Siani, Frank Thies, Sophia Tsabouri, Marco Vinceti, Jean-Louis Bresson, Mary Fewtrell, Mathilde Kersting, Hildegard Przyrembel, Céline Dumas, Ariane Titz and Dominique Turck

Abstract

Following a request from the European Commission, the Panel on Nutrition, Novel Foods and Food Allergens (NDA) revised its 2009 Opinion on the appropriate age for introduction of complementary feeding of infants. This age has been evaluated considering the effects on health outcomes, nutritional aspects and infant development, and depends on the individual's characteristics and development. As long as foods have an age-appropriate texture, are nutritionally appropriate and prepared following good hygiene practices, there is no convincing evidence that at any age investigated in the included studies (< 1 to < 6 months), the introduction of complementary foods (CFs) is associated with adverse health effects or benefits (except for infants at risk of iron depletion). For nutritional reasons, the majority of infants need CFs from around 6 months of age. Infants at risk of iron depletion (exclusively breastfed infants born to mothers with low iron status, or with early umbilical cord clamping (< 1 min after birth), or born preterm, or born small-for-gestational age or with high growth velocity) may benefit from earlier introduction of CFs that are a source of iron. The earliest developmental skills relevant for consuming pureed CFs can be observed between 3 and 4 months of age. Skills for consuming finger foods can be observed in some infants at 4 months, but more commonly at 5–7 months. The fact that an infant may be ready from a neurodevelopmental perspective to progress to a more diversified diet before 6 months of age does not imply that there is a need to introduce CFs. There is no reason to postpone the introduction of potentially allergenic foods (egg, cereals, fish and peanut) to a later age than that of other CFs as far as the risk of developing atopic diseases is concerned. Regarding the risk of coeliac disease, gluten can be introduced with other CFs.

Keywords: complementary food, introduction, timing, infant, health outcome, development, systematic review

Requestor: European Commission

Question number: EFSA-Q-2016-00482

Correspondence: nda@efsa.europa.eu

Panel members: Jacqueline Castenmiller, Stefaan de Henauw, Karen-Ildico Hirsch-Ernst, John Kearney, Helle Katrine Knutsen, Alexandre Maciuk, Inge Mangelsdorf, Harry J McArdle, Androniki Naska, Carmen Pelaez, Kristina Pentieva, Alfonso Siani, Frank Thies, Sophia Tsabouri, Dominique Turck and Marco Vinceti.

Acknowledgements: The Panel wishes to thank the following EFSA statutory staff members (Declarations of interest available upon request) for the support provided to this scientific output: Federica Barrucci, Andrea Bau, Laura Ciccolallo, Agnès de Sesmaisons-Lecarré, Lucia Fabiani, Leonard Matijević, Irene Muñoz Guajardo, and Ermolaos Ververis. The Panel also wishes to thank the following EFSA non-statutory staff members* for the support provided to this scientific output: Mathias Amundsen, Claudia Denami, Pauline Lachevre, Federico Morreale, Charlotte Salgaard Nielsen and Qinqing Sun.

Interests and EFSA funding: All authors of this Scientific Opinion have submitted Declarations of Interest (DoI) which have been evaluated in accordance with EFSA's policy on Independence and Competing Interest Management (http://www.efsa.europa.eu/en/corporate/pub/policyonindependence, https://www.efsa.europa.eu/en/corporate/pub/independencepolicy17). The DoIs of members of the NDA Panel and of the Working Group are available in the 'DoI database' on EFSA's website (https://ess.efsa.europa.eu/doi/doiweb/doisearch/). In accordance with Article 21 of the Decision of the Executive Director on Competing Interest Management a waiver was granted to an expert of the Working Group until end of October 2018. Pursuant to Article 21(6) of the aforementioned Decision, the concerned expert was allowed to take part in the discussion and in the drafting phase of the scientific output but was not allowed to take up the role of rapporteur within that time frame. Any competing interests are recorded in the respective minutes of the meetings of the NDA Panel (http://www.efsa.europa.eu/en/events/advanced-search) and of the Working Group (https://www.efsa.europa.eu/sites/default/files/wgs/nutrition/ndainfantnutrition.pdf). EFSA statutory staff members are also subject to DoI submission and screening requirements and are assessed in line with the above policy under the EU Staff Regulations and its implementing rules. DoIs of the Heads of Department and Heads of Unit responsible for managing EFSA's scientific and operational services are also available on EFSA's website. Allowances and indemnities received by experts for their involvement in EFSA's work are outlined in the EFSA experts' compensation guide (http://www.efsa.europa.eu/sites/default/files/Experts_compensation_guide.pdf).
EFSA is fully funded by the general budget of the European Union or of Member States participating to its governance bodies and receives no fees or contributions from private sources.

Suggested citation: EFSA NDA Panel (EFSA Panel on Nutrition, Novel Foods and Food Allergens), Castenmiller J, de Henauw S, Hirsch-Ernst K-I, Kearney J, Knutsen HK, Maciuk A, Mangelsdorf I, McArdle HJ, Naska A, Pelaez C, Pentieva K, Siani A, Thies F, Tsabouri S, Vinceti M, Bresson J-L, Fewtrell M, Kersting M, Przyrembel H, Dumas C, Titz A and Turck D, 2019. Scientific Opinion on the appropriate age range for introduction of complementary feeding into an infant's diet. EFSA Journal 2019;17(9):5780, 241 pp. https://doi.org/10.2903/j.efsa.2019.5780

ISSN: 1831-4732

The EFSA Journal is a publication of the European Food Safety Authority, an agency of the European Union.

* Non-statutory staff members (i.e. trainees, visiting scientists and interim staff) are not covered by the EU staff regulations and are therefore not subject to DoIs submission.

Summary

Following a request from the European Commission, the Panel on Nutrition, Novel Foods and Food Allergens (NDA Panel) revised its Scientific Opinion of 2009 on the appropriate age for introduction of complementary feeding of infants.

This request arises in the context of the information regarding the use of processed cereal-based foods and baby foods. This information is required for a future delegated act of the European Commission on these foods foreseen in Regulation (EU) No 609/2013 on food intended for infants and young children. This Regulation revises the legal framework set out in Directive 2009/39/EC on foodstuffs intended for particular nutritional uses and the specific Directives adopted under this framework, including Directive 2006/125/EC on processed cereal-based foods and baby foods for infants and young children. This Directive required the mandatory indication of a statement on the appropriate age from which processed cereal-based foods and baby foods may be used, that shall be not less than four months for any products.

The Panel specified upfront in a protocol the strategy and methodology to collect and evaluate scientific data on possible relationships between the timing of introduction of complementary foods (CFs) and a number of (health) outcomes. This protocol was released for public consultation and published, alongside a report on how comments received during the public consultation were taken into account in the final protocol. A draft of this Scientific Opinion was also released for public consultation and revised according to the comments received, where appropriate. The comments that were received were addressed in detail in a technical report that is published together with this Scientific Opinion.

The Panel considers that exclusive breastfeeding is nutritionally appropriate up to 6 months of age for the majority of healthy infants born at term from healthy well-nourished mothers.

The purpose of this Scientific Opinion is to assess the scientific evidence in relation to whether there are:

1) any developmental factors relevant for the introduction of CFs,
2) any adverse health effects associated with the introduction of CFs before 6 months of age, and
3) any benefits associated with the introduction of CFs before 6 months of age.

Out of the scope of this Scientific Opinion are:

- public health recommendations for the introduction of CFs; this task is outside the remit of the European Food Safety Authority (EFSA) but it is the role of public health authorities in Member States;
- the effects of the duration of exclusive breastfeeding on the selected health outcomes, as the assessment is performed irrespective of whether infants were initially exclusively breastfed or formula fed;
- the health benefits of breastfeeding itself (for the infant/child and the mother);
- the effects on health outcomes of introduction of CFs solely after 6 months of age, as there is a nutritional requirement for CFs for the majority of exclusively breastfed infants by 6 months onwards;
- the effects of the amount, order of introduction, variety, composition and texture of CFs;
- the role of aspects, such as social interactions and the cultural context, on the appropriate age of introduction of CFs;
- risks related to, e.g. chemical or microbiological contaminants or pesticides.

The definition of CFs differs in different publications. In the context of this Scientific Opinion, complementary feeding is defined as the period when CFs are given together with either breast milk or formula or both. CFs in this Scientific Opinion comprise foods other than breast milk, formula, water or vitamins that are given to infants and can be beverages, spoon-fed pureed foods, spoon-fed lumpy foods or finger foods, either prepared at home or produced commercially. This definition is in line with that used by some other bodies, such as the European Society for Paediatric Gastroenterology, Hepatology and Nutrition (ESPGHAN), the UK Scientific Advisory Committee on Nutrition (SACN), the United States Department of Agriculture (USDA) and the American Academy of Pediatrics (AAP) but differs from the one that has been used by, for example the World Health Organization (WHO), which included formula in the definition of CFs.

In the interpretation of the Terms of Reference, the choice has been made by the Panel to limit the assessment to health effects associated with the timing of introduction of CFs or specific foods before

the age of 6 months. This led to the exclusion of studies that had been considered by other bodies in their assessments done in different contexts than this Scientific Opinion. This is, for example, the case for some studies that investigated the introduction of some allergenic foods, such as fish, egg and peanut, or of gluten after 6 months of age.

The appropriate age of introduction of CFs is influenced not only by nutritional considerations, but also by effects on health outcomes and by infant development. Considering the influence of various factors, the Panel considers that it is likely that there is an appropriate age range rather than a single appropriate age for the introduction of CFs.

The Panel undertook a systematic literature search of intervention and observational studies for the assessment of the association between the timing of introduction of CFs and health outcomes, while an extensive literature search was carried out specifically for developmental determinants of the introduction of CFs. The Panel also appraised the risk of bias (RoB) of the studies included from the systematic search, thus classifying them as low, intermediate or high RoB (Tiers 1, 2 or 3).

Studies considered pertinent for this assessment were those in infants and children, generally healthy at the time of introduction of CFs, either born at term or preterm. The study groups had to be alike in terms of the type of milk feeding (breast milk or formula or mixed, with no additional behavioural interventions), i.e. the study groups had to differ only in the timing of the introduction of CFs. The selected papers were studies in which at least one group was introduced to CFs before 6 months of age. Studies on a specific CF item or food group were also considered for certain health outcomes (e.g. gluten in relation to the risk of coeliac disease). The list of outcomes to be evaluated was defined in the protocol, based on the previous EFSA Scientific Opinion of 2009, and expanded when evidence was available. Endpoints for which only one study was available were not included. In the systematic review, the Panel has assessed 283 studies that reported on the relationship between the timing of introduction of CFs (or specific foods for some outcomes) in relation to (1) body weight and growth, including body mass index (BMI), risk of developing overweight and obesity, as well as body composition, (2) risk of developing atopic diseases or symptoms of atopic diseases, such as asthma-like symptoms, eczema, allergic rhinitis and symptomatic food allergy, (3) risk of developing coeliac disease and type 1 diabetes mellitus, (4) blood pressure, (5) infections, (6) sleep, (7) infant and child development, (8) nutrient status (i.e. iron) and (9) food preferences and eating behaviours later in life. For these outcomes, whenever enough data were available, forest plots were created, and pooled estimates were calculated from the individual studies, with associated 95% confidence and prediction intervals, using random effects meta-analyses. Evidence was discussed separately for infants born at term and those born preterm.

Developmental skills relevant for the progression from a liquid to a diversified diet

For the assessment of the oral–motor developmental readiness of infants to receive CFs, the Panel conducted an extensive literature search to retrieve studies, review papers and text books that provided information on when certain milestones indicative of the oral–motor readiness to receive CFs are reached in the normally developing term infant.

One determinant of the appropriate age range of introduction of CFs is the infant's anatomical, physiological and oral–motor readiness to receive foods other than breast milk or formula. Gastrointestinal and renal functions are not limiting factors with respect to the timing of introduction of CFs once the infant has the necessary neuromotor skills and has developed an apparent interest in non-milk foods and feeding. The changes that are required for progressing from a liquid to a semi-solid and solid diet are: (1) anatomical changes in the oral cavity, (2) the disappearance or diminishing of reflexes present at birth that coordinate suckling, swallowing and respiration, and protect the infant from aspiration and choking (i.e. the extrusion reflex of the tongue), in favour of more voluntary movements and (3) the development of gross motor skills (head and trunk control to allow an improved movement of the jaw) and fine motor skills (lip, tongue and jaw movements).

The age range at which infants attain these developmental milestones shows considerable variation within and between populations, presumably reflecting the infant's innate developmental trajectory combined with the opportunities and experiences provided by the carer.

The earliest gross motor skills indicative of developmental readiness for spoon-feeding of pureed foods (i.e. holding the head in midline when in supine position and to control its head well when pulled to sitting or at aided sitting) can be observed between 3 and 4 months of age. At this age, it can be assumed that the rooting and the extrusion reflexes may have also diminished in some infants. The gross motor skill indicative of developmental readiness for self-feeding finger foods (i.e. sitting without support) can be observed in some infants at 4 months, but more commonly between 5 and 7 months

of age. In preterm infants, the necessary developmental milestones for feeding are also reached around the same age range (post-term), depending on the severity of illness experienced during the neonatal period, the degree of prematurity and any sequelae.

Nutritional need for the introduction of CFs

Most infants do not need CFs for nutritional reasons up to around 6 months of age, with the exception of some infants at risk of iron depletion who may benefit from earlier introduction of CFs that are a source of iron. From the systematic review, the Panel concludes that there is high confidence in the evidence that the introduction of CFs at 4 months of age compared with 6 months of age reduces the risk of iron depletion at 6 months of age in exclusively breastfed infants at risk of iron depletion. However, the effect on iron depletion is not an effect of introducing CFs *per se*, but an effect of introducing CFs that are a source of iron. Infants that may benefit from an early introduction of CFs that are a source of iron are exclusively breastfed infants born to mothers with a low iron status, or with early umbilical cord clamping (< 1 min after birth), or born preterm, or born small-for-gestational age, or with a high growth velocity.

Adverse health effects or benefits associated with the introduction of CFs before 6 months of age

There is no convincing evidence for adverse health effects of introducing CFs at any of the ages investigated in the included studies. In the studies rated as Tiers 1 and 2, the definition of 'early introduction of CFs' ranged from < 1 month to < 6 months. In most instances, < 3 or < 4 months of age was investigated as 'early introduction' without precise information on the earliest age at which infants in the study were introduced to CFs. The Panel applied a weight of evidence approach to derive its conclusions and grade the confidence in the evidence.

The Panel concludes (high level of confidence) (1) that there was no effect of introduction of CFs at 3–4 months of age, compared with 6 months of age, on body weight, body length, head circumference, BMI and body composition; (2) that there is no effect of the introduction of gluten at 4 months of age compared with 6 months of age on the risk of developing coeliac disease; and (3) that there is no evidence for an effect or an association between the timing of introduction of CFs in mixed fed populations and iron status at 10–12 months of age.

The Panel concludes (moderate level of confidence) that there is no evidence for an association between the timing of introduction of CFs and body weight (between < 2 and < 6 months vs thereafter), body length (between 2–3 and < 6 months vs thereafter), BMI (between ≤ 2 and ≤ 5 months vs thereafter), body composition (< 4 months vs ≥ 4 to > 6 months) and coeliac disease (for gluten, between ≤ 3 and ≤ 4 months vs thereafter). The Panel also concludes (moderate level of confidence) that there is no evidence for an effect or an association between the timing of introduction of CFs and overweight (between ≤ 2 and < 4 months vs > 2 to > 6 months), obesity (between < 1 and < 4 months vs ≥ 3 to ≥ 6 months), atopic diseases (at 3–4 vs 6 months), asthma-like symptoms (at 3–4 vs 6 months for CFs, < 3.75–5.5 months vs thereafter for cereals and < 5.25 to ≤ 6 months vs >5.25 to 8.5 months for fish), eczema (between < 3 and ≤ 6 months vs thereafter), allergic rhinitis (at 3–4 vs 6 months), symptomatic food allergy (at 3–4 vs 6 months), type 1 diabetes mellitus (gluten and CFs, between < 3 and < 5 months vs thereafter), blood pressure (between < 3 and < 5 months vs thereafter) and infections in general (between 3–4 months and < 6 months vs at 6 and > 6 months).

The Panel considers that the confidence level in the evidence was low to very low for a number of outcomes related to atopic diseases (and introduction of specific foods) as well as for gastrointestinal and lower respiratory tract infections, sleep, and infant and child development.

For some outcomes, the evidence was inconsistent and therefore the confidence in the evidence was not graded (i.e. timing of introduction of peanut and peanut allergy, upper respiratory tract infections, and food preferences and eating habits (introduction of CFs and fruit and vegetables)).

Even though there is no convincing evidence for a harmful effect of CF introduction at any age that was studied on the selected health outcomes, the Panel emphasises that foods given to infants should be presented in an age-appropriate texture (to prevent aspiration and choking), are nutritionally appropriate and are prepared according to good hygiene practices. Also, the fact that, based on the available evidence, CFs could be introduced before 6 months of age does not imply that this is necessary or desirable.

In the following the main findings are summarised:

- **Specific allergenic foods**

In relation to the introduction of allergenic foods (egg, cereals, fish and peanut) into an infant's diet, the Panel concludes that allergenic foods can be introduced in the same way as other CFs once the infant has the necessary neuromotor skills and has developed an apparent interest in non-milk foods and feeding. There is no evidence to support postponing the introduction of potentially allergenic foods to a later age than the introduction of other CFs.

  - **Hen's egg and egg allergy**

With respect to egg introduction, the data pointed towards a favourable effect of its introduction between around 3–4 months compared with 6 months of age on the risk of developing egg allergy. However, the confidence in the evidence is low to moderate and is, therefore, insufficient to support introducing egg at around 3–4 months of age in all infants for the prevention of egg allergy. In the available studies, no serious adverse reactions occurred with consumption of cooked egg, while anaphylactic reactions were observed when the intervention consisted of pasteurised raw egg powder. As far as the risk of allergy is concerned, cooked egg can be introduced into the diet of infants when other CFs are introduced.

  - **Peanut and peanut allergy**

There is evidence that peanut introduction during the first year of life (either at 4–10 months or at 4–6 months) compared with peanut avoidance up to 5 years of age reduces the risk of developing peanut allergy. However, the evidence is insufficient to conclude whether, when comparing infants introduced to peanut $\leq$ 6 months of age with those introduced $>$ 6 months (but still within the first year of life, which is the subject of this mandate), a similar effect occurs. As the evidence was inconsistent, no level of confidence was assigned.

- **Overweight and obesity**

There is no evidence that the timing of introduction of CFs is associated with higher risk of developing overweight and obesity (moderate confidence in the evidence). This finding is supported by the results on body weight, BMI and fat mass (moderate to high confidence in the evidence, depending on the outcome).

- **Coeliac disease and type 1 diabetes mellitus**

If gluten is introduced, there is no evidence for beneficial or adverse health effects of gluten introduction $<$ 6 months of age compared with thereafter with respect to the risk of developing coeliac disease or type 1 diabetes mellitus, nor is there evidence that (any) continued breastfeeding could modify the effect of gluten introduction at that age (moderate to high level of confidence in the evidence, depending on the age of introduction of CFs investigated). As far as the risk of developing coeliac disease or type 1 diabetes mellitus is concerned, gluten can be introduced to an infant's diet when other CFs are introduced. Time to onset of coeliac disease or type 1 diabetes mellitus in relation to the timing of introduction of CFs was not considered.

- **Infections**

When hygiene conditions are satisfactory,[1] there is no evidence that the introduction of CFs $<$ 6 months of age compared with thereafter is associated with an increased risk of (1) gastrointestinal infections (low level of confidence in the evidence), (2) lower respiratory tract infections (moderate level of confidence in the evidence) or (3) infections in general (moderate level of confidence in the evidence). The evidence for upper respiratory tract infections is inconsistent and insufficient to draw conclusions.

- **Sleep-related endpoints**

Even though the statistical analyses of the effect of the age of introduction of CFs on sleep-related endpoints was significant (low level of confidence), the Panel considered that the size of the effect was not biologically relevant.

[1] Studies in low-income and lower-middle-income countries that were conducted in poor hygiene conditions were excluded.

- **Preterm infants**

The available evidence on preterm infants is limited and comprised only one study in the main line of evidence. From this study, there is no evidence for an effect of introduction of CFs at 4 months post-term compared with 6 months post-term on body weight, body length and head circumference (low level of confidence in the evidence).

Conclusions

The appropriate age range of introduction of CFs has been evaluated taking into account effects on health outcomes, nutritional aspects and infant development.

The available data do not allow the determination of a single age for the introduction of CFs for infants living in Europe. The appropriate age range depends on the individual's characteristics and development, even more so if the infant was born preterm.

As long as the foods are given in an age-appropriate texture, are nutritionally appropriate and prepared according to good hygiene practices, there is no convincing evidence that the introduction of CFs is associated with either adverse or beneficial health effects (except for infants at risk of iron depletion) at any age investigated in the included studies (< 1 month to < 6 months for earlier introduction).

For nutritional reasons, the majority of infants need CFs from around 6 months of age. For preterm infants, this refers to post-term age. Infants at risk of iron depletion (exclusively breastfed infants born to mothers with low iron status, or with early umbilical cord clamping (< 1 min after birth), or born preterm, or born small-for-gestational age or with high growth velocity) may benefit from introduction of CFs that are a source of iron before 6 months of age.

The earliest developmental skills relevant for the consumption of spoon-fed pureed CFs can be observed between 3 and 4 months of age. Skills necessary for consuming self-fed finger foods can be observed in some infants at 4 months, but more commonly between 5 and 7 months of age. For preterm infants, this refers to post-term age.

The fact that an infant may be ready from a neurodevelopmental point of view to progress from a liquid to a more diversified diet before 6 months of age does not imply that there is a need to introduce CFs.

There is no reason to postpone the introduction of potentially allergenic foods (egg, cereals, fish and peanut) to a later age than that of other CFs as far as the risk of developing atopic diseases is concerned. Regarding the risk of coeliac disease, gluten can be introduced with other CFs.

Table of contents

1. Introduction

1.1. Background and Terms of Reference as provided by the requestor

1.1.1. Background

Directive 2009/39/EC[2] of the European Parliament and of the Council on foodstuffs intended for particular nutritional uses lays down general compositional and information requirements of such foods that are specially designed to meet the particular nutritional requirements of the persons to whom they are intended, including those 'of infants and young children in good health'.

Directive 2006/125/EC[3] has established compositional and labelling requirements for processed cereal-based foods and baby foods for infants and young children which are defined in the legislation as "foodstuffs for particular nutritional use fulfilling the particular requirements of infants and young children in good health (...) and are intended for the use by infants while they are being weaned, and by young children as a supplement to their diet and/or for their progressive adaptation to ordinary food".

The Directive defines 'infants' as "children under the age of 12 months" and 'young children' as "children aged between one and three years".

In particular, Article 8(1)(a) of Directive 2006/125/EC requires the mandatory indication of a statement as to the appropriate age from which processed cereal-based food and baby food may be used. According to this provision the stated age shall be not less than four months for any products. The product, if its use is recommended from four months, may indicate that it is suitable from that age unless independent persons having qualifications in medicine, nutrition or pharmacy, or other professionals responsible for maternal and child care, advise otherwise. This requirement is in line with EFSA's scientific opinion on the appropriate age for introduction of complementary feeding of infants.

Regulation (EU) No 609/2013[4] of the European Parliament and of the Council on food intended for infants and young children, food for special medical purposes, and total diet replacement for weight control revises the legal framework applicable to foods for particular nutritional uses as set out in Directive 2009/39/EC and the specific Directives adopted under this framework, including Directive 2006/125/EC.

The Regulation includes in its scope processed cereal-based food and baby food, maintains the definitions as laid down in Directive 2006/125/EC for them. With respect to labelling, presentation and advertising Article 9(5) of the Regulation generally requires amongst others that the food governed by this legislation "shall provide information for the appropriate use of such food".

In addition to the general requirements of Regulation (EU) No 609/2013 the Commission is required to lay down by the means of delegated act specific compositional and information requirements for processed cereal-based food and baby food, taking into account relevant technical and new scientific evidence and knowledge available.

In the context of the information to be provided regarding the use of processed cereal-based and baby food, questions have been raised on the appropriate age for introduction of complementary feeding of infants.

Taking into account the abovementioned, it is considered necessary, at this stage to request EFSA to update the conclusions of its scientific opinion on the appropriate age for introduction of complementary feeding of infants in light of more recent scientific evidence.

1.1.2. Terms of reference

In accordance with Article 29(1)(a) of Regulation (EC) No 178/2002, the European Commission asks EFSA to:

- Update EFSA's scientific opinion on the appropriate age for introduction of complementary feeding of infants in light of more recent scientific evidence and knowledge available.

[2] Directive 2009/39/EC of the European Parliament and of the Council of 6 May 2009 on foodstuffs intended for particular nutritional uses (recast), OJ L 124, 20.5.2009, p. 21–29.

[3] Commission Directive 2006/125/EC of 5 December 2006 on processed cereal-based foods and baby foods for infants and young children, OJ L 339, 6.12.2006, p. 16–35.

[4] Regulation (EU) No 609/2013 of the European Parliament and of the Council of 12 June 2013 on food intended for infants and young children, food for special medical purposes, and total diet replacement for weight control and repealing Council Directive 92/52/EEC, Commission Directives 96/8/EC, 1999/21/EC, 2006/125/EC and 2006/141/EC, Directive 2009/39/EC of the European Parliament and of the Council and Commission Regulations (EC) No 41/2009 and (EC) No 953/2009; OJ L 181, 29.6.2013, p. 35–56.

1.2. Previous assessments

In its previous Scientific Opinion (EFSA NDA Panel, 2009), the Panel concluded that 'the introduction of complementary food into the diet of healthy term infants in the European Union (EU) between the age of 4 and 6 months is safe and does not pose a risk of adverse health effects'. The Panel also concluded that 'available data on the risk of coeliac disease and type 1 diabetes mellitus (T1DM) support also the timing of the introduction of gluten-containing food (preferably while still breastfeeding) not later than 6 months of age'. These conclusions were based on data from high-income countries, and primarily on observational data in exclusively breastfed infants, healthy and born at term. The list of endpoints, discussed narratively in the Scientific Opinion in relation to exclusive breastfeeding and/or age of introduction of complementary foods (CFs), were nutrient requirement, growth, neurodevelopment, digestive system, renal function, development of food preferences, and risk of obesity, type 2 diabetes mellitus, atopic diseases, coeliac disease, T1DM, infectious morbidity and caries.

The Panel was also aware of the following position statements or reports. In the UK, the Scientific Advisory Committee on Nutrition (SACN) and the Committee on Toxicity of Chemicals in food, Consumer Products and the Environment (COT) published statements on health benefits and risks of introduction of peanut and hen's egg into the infant diet before 6 months and on the timing of introduction of gluten into the infant diet (SACN-COT, 2011, 2018). Their main conclusions were that the 'evidence that the introduction of hen's egg before 6 months might be beneficial was limited'. The committees concluded as well that 'there were insufficient data to demonstrate that the introduction of peanut or hen's egg into the infant diet between four and six months of age reduced the risk of developing food allergy to any greater extent than introduction from around six months'. The committees also concluded that 'currently available evidence on the timing of introduction of gluten into the infant diet and subsequent risk of coeliac disease and [type 1 diabetes mellitus T1DM] is insufficient to support recommendations about the appropriate timing of introduction of gluten into the infant diet beyond 3 completed months of age, for either the general population or high-risk sub-populations'. They also considered that the evidence was insufficient to support the introduction of gluten into the infant's diet not later than 6 completed months of age with the objective of reducing the risk of developing coeliac disease and T1DM.

The Panel was also aware that the SACN report on feeding in the first year of life covers aspects of infant feeding other than complementary feeding, such as the adequate duration of breastfeeding (SACN, 2018). Its main conclusions in relation to the timing of introduction of complementary foods (CFs) were that '(a) observed relationships between the timing of introduction of complementary foods and obesity were in most prospective studies attributed to rapid early weight gain rather than early introduction of complementary foods, (b) there is insufficient evidence to demonstrate that introduction of peanut, hen's egg, gluten or fish before 6 months of age reduces the risk of developing food allergy as compared to the introduction at around 6 months of age, (c) there is high quality evidence that the timing of introduction of gluten is not related to the risk of developing coeliac disease, (d) there is low quality evidence that fish introduction before 6 to 12 months of age [i.e. from evidence covering different ages of introduction between < 6 and 12 months of age] is associated with a reduced risk of developing allergic rhinitis and sensitisation, (e) there is no "critical window" for introducing complementary foods that is related to later food acceptance'.

The Panel took note of the position papers of the European Society for Paediatric Gastroenterology, Hepatology and Nutrition (ESPGHAN) on complementary feeding and on gluten introduction and risk of coeliac disease (Szajewska et al., 2016; Fewtrell et al., 2017). Regarding specifically the introduction of CFs, their main conclusions were that 'complementary foods (solids and liquids other than breast milk or infant formula) should not be introduced before 4 months but should not be delayed beyond 6 months'. Regarding the age of introduction of allergenic foods, their main conclusions were that 'allergenic foods may be introduced when complementary food is commenced any time after 4 months'. In addition, ESPGHAN considered that 'infants at high risk of peanut allergy [...] should have peanut introduced between 4 and 11 months, following evaluation by an appropriate trained specialist' and 'gluten may be introduced between 4 and 12 months'. ESPGHAN indicated that 'although breastfeeding should be promoted for its other well-established health benefits, neither any breastfeeding nor breastfeeding during gluten introduction has been shown to reduce the risk of coeliac disease'.

The World Health Organization (WHO) report on Feeding and Nutrition of Infants and Young Children (WHO Regional Office for Europe, 2003) concluded, based on a narrative description of the evidence, that 'complementary foods should be introduced at about 6 months of age. Some infants may need complementary foods earlier, but not before 4 months of age'.

The United States Department of Agriculture (USDA) and the Department of Health and Human Services launched the Pregnancy and Birth to 24 months project, which involved conducting a series of systematic reviews about the timing of introduction of complementary feeding in healthy term infants. They concluded that there was moderate evidence that there was no relationship between the introduction of CFs at 4–5 months compared with 6 months and weight, length, overweight and obesity, and body composition. However, limited evidence was found that introducing CFs before 4 months compared with later could increase the odds of overweight and obesity (English et al., 2019a). For outcomes on atopic diseases, Obbagy et al. (2019a) reported that there was moderate evidence for no association between the age of CF introduction and the risk of developing food allergy, atopic dermatitis, or childhood asthma. Limited to strong evidence (depending on the specific food studied) suggested that the risk of food allergy and atopic dermatitis did not increase by introducing allergenic foods after 4 months of age but within the first year of life, although it may prevent peanut and egg allergy. For bone health and developmental milestones, only three articles were available (English et al., 2019a; Obbagy et al., 2019b). Hence, the authors concluded that insufficient evidence was available to draw conclusions on the relationships, or to grade the confidence in the evidence. For micronutrient status, Obbagy et al. (2019c) found moderate evidence that introducing CFs at 4 months of age compared with 6 months does not affect iron status, derived from evidence generated in high-income countries.

The Panel also took note of a recent report (Greer et al., 2019) of the American Academy of Pediatrics (AAP). This report concluded that 'there is no evidence that delaying the introduction of allergenic foods, including peanut, egg, and fish, beyond 4 to 6 months prevents atopic disease'. It also concluded that 'there is now evidence that the early introduction of infant-safe forms of peanuts reduces the risk of peanut allergies. Data are less clear for timing of introduction of egg'.

The National Institute of Allergy and Infectious Diseases in the United States provided guidelines on early introduction of peanut into the diet of infants who were at three risk levels (Togias et al., 2017). To reduce the risk of peanut allergy, it was recommended to introduce peanut-containing foods from 4 to 6 months of age into the diet of infants with severe eczema, egg allergy or both. Moreover, it was suggested to introduce peanut-containing foods around 6 months of age into the diet of infants with mild-to-moderate eczema, and freely into the diet of infants without eczema or any food allergy.

1.3. Definitions

Complementary feeding means the period when CFs are given together with either breast milk or formula or both (EFSA NDA Panel, 2009). This definition is in line with the terms of reference received from the European Commission and is also in line with the definition used by other bodies (e.g. ESPGHAN (Fewtrell et al., 2017), SACN (SACN, 2018), USDA (Obbagy et al., 2019b) or the AAP (AAP, 2014). It differs from the definition of WHO which defined 'complementary feeding' as 'the process starting when breast milk alone is no longer sufficient to meet the nutritional requirements of infants, and therefore other foods and liquids are needed, along with breast milk'.[5] The Panel understands that these 'other foods' in this last definition may also comprise formula.

CFs in this Scientific Opinion comprises, therefore, all liquid, semisolid and solid foods other than breast milk, formula, water or vitamins that are given to infants. CFs can be beverages, spoon-fed pureed foods, spoon-fed lumpy foods or finger foods (EFSA NDA Panel, 2009), depending on the age of the infant. They can be either prepared at home or produced commercially.

Weaning in this Scientific Opinion means the time period of gradual reduction of frequency and volume of breast milk or formula which starts with the first introduction of CFs and gradually leads to a dietary pattern customary in the infant's family during the second year of life (EFSA NDA Panel, 2009).

Breastfeeding may be exclusive, predominant, full, mixed or partial. Exclusive breastfeeding means that no other food or liquid is given besides breast milk and medicines or vitamin drops. It is predominant if, in addition to breast milk, the infant receives 'non-milk liquids' (i.e. other than breast milk or formula) like water or energy-free 'teas'. Exclusive and predominant breastfeeding together are called full breastfeeding. Mixed breastfeeding means that, in addition to breast milk, the infant receives

[5] http://www.who.int/elena/titles/complementary_feeding/en/

formula. Partial breastfeeding is breastfeeding together with CFs (EFSA NDA Panel, 2009). The Panel notes that different definitions may be found in the literature.

Appropriate, according to the Oxford English Dictionary, means suitable for a given circumstance.

The Panel notes that, from a scientific point of view, the assessment of the appropriate age range of introduction of CFs (which is the subject of this mandate) is not an assessment of the optimal duration of exclusive breastfeeding.

1.4. Need for complementary foods for infants

The following Section summarises the knowledge that is available on the nutritional adequacy of exclusive breastfeeding in the first months of life in healthy infants born at term from healthy well-nourished mothers.

1.4.1. Nutritional adequacy of exclusive breastfeeding

Breast milk composition changes with gestational and post-natal age, from the start to the end of a feed, and follows a diurnal pattern.

1.4.1.1. Energy and protein

Energy content of breast milk is fairly stable over the first year of life (Nommsen et al., 1991; Nielsen et al., 2011; Gidrewicz and Fenton, 2014). It is sufficient to meet the energy requirements of exclusively breastfed infants during the first six months of life (Butte et al., 2002; Nielsen et al., 2011). This consideration is based on (1) the comparison of energy intakes from breast milk (using age-specific volume intakes corrected for insensible water losses[6]) to data on total energy expenditure and energy deposition related to growth and accretion of fat and protein (Butte et al., 2002) and (2) data on adequate growth of infants exclusively breastfed up to 6 months of age (Nielsen et al., 2011).

Measured content of true protein of term breast milk was observed to decrease in the first few weeks of life (Gidrewicz and Fenton, 2014) and to be fairly stable thereafter up to 12 months of age (Nommsen et al., 1991). The protein content of breast milk fulfils the protein requirements of infants, as derived from factorial estimates of requirements for maintenance and deposition (EFSA NDA Panel, 2013). In addition, weight and length gain of exclusively breastfed healthy term infants who received a protein supplement from 4 to 6 months of age was similar to a control group exclusively breastfed for 6 months in a randomised controlled trial (RCT), despite a 20% higher protein intake (Dewey et al., 1996).

The Panel considers that the energy and the protein contents of breast milk are sufficient to cover the nutritional needs of infants up to 6 months of age.

1.4.1.2. Minerals, vitamins and fatty acids

The iron concentration of breast milk decreases with the duration of lactation, and is unaffected by maternal iron status and diet (EFSA NDA Panel, 2015a). The healthy term infant of a well-nourished mother is born with a store of iron (body content about 75 mg/kg body weight), which can be increased by about 30–35 mg through delayed clamping of the umbilical cord (i.e. > 2 min after birth). According to the review by Chaparro (2008), this store is sufficient to supply the iron needed for the formation of haemoglobin (Hb) and myoglobin concomitant with growth until about 6 months of age in most fully breastfed infants (EFSA NDA Panel, 2015a).

However, some infants who are at risk of iron depletion, e.g. infants born to mothers with a low iron status, infants with early umbilical cord clamping (< 1 min after birth), infants born preterm, infants born small-for-gestational age (SGA) and infants with a high growth velocity, may need additional iron before 6 months of age. This was investigated in three RCTs (Dewey et al., 1998; Dewey et al., 2004; Jonsdottir et al., 2012), performed in healthy term exclusively breastfed infants, both SGA and appropriate-for gestational age (AGA), at some degree at risk of iron depletion. A meta-analysis of these trials done by EFSA (Appendix A.48) showed that the risk of iron depletion (serum ferritin (SF) concentrations < 12 μg/L) at 6 months of age was statistically significantly lower when CFs were introduced at 4 months of age (Section 15.3). It should be emphasised that iron depletion is a risk factor for iron-deficiency anaemia which is associated with deleterious effects (e.g. delayed attention, poor recognition memory, long-lasting poor cognitive and behavioural performance) (Geng et al., 2015; Lynch et al., 2018).

[6] Losses via transepidermal diffusion and via the respiratory tract.

Zinc concentrations in breast milk sharply decline over the early months of lactation and are not associated with maternal zinc status, her dietary zinc intake or zinc supplementation (EFSA NDA Panel, 2014b). However, there are no reports describing zinc deficiency in term breastfed infants up to 6 months of age in well-nourished populations. Zinc concentration in breast milk is considered to be adequate for the majority of healthy term breastfed infants up to six months of life (EFSA NDA Panel, 2013) and thus is not a determinant for the need to introduce CFs.

There is a general agreement that breast milk does not contain sufficient vitamin D to prevent rickets in the breastfed infant. The vitamin D content of breast milk is, however, not a determinant for the need to introduce CFs, because infants in the EU are routinely supplemented with vitamin D (daily supplement of 10 μg to all infants is recommended by ESPGHAN (Braegger et al., 2013)).

The vitamin A concentration in breast milk is dependent on the maternal vitamin A status and decreases with prolonged lactation (EFSA NDA Panel, 2015b). There is no indication that vitamin A insufficiency occurs in exclusively breastfed infants in well-nourished populations (Butte et al., 2002), in which the vitamin A content of breast milk is thus not a determinant for the need to introduce CFs.

Breast milk has a low phylloquinone content, which can increase the risk of vitamin K deficiency bleeding. Administration of phylloquinone at a pharmacological dose is usual practice for prevention of haemorrhagic disease in newborn infants (EFSA NDA Panel, 2013, 2017) and phylloquinone content of breast milk is thus not a determinant for the need to introduce CFs.

Concentrations of most B vitamins, iodine and selenium and certain fatty acids, for example docosahexaenoic acid (DHA) in breast milk are directly influenced by current maternal intake and are, in well-nourished populations, not determinants for the need to introduce CFs. However, there are case reports of infants from mothers with undetected pernicious anaemia or adhering to a strict vegan diet without taking supplements that show that clinical symptoms of cobalamin deficiency may occur in exclusively breastfed infants (Dror and Allen, 2008; EFSA NDA Panel, 2015c).

The Panel concludes that the micronutrient and fatty acid contents of breast milk are not determinants for the need to introduce CFs. However, the Panel considers that the iron status of the infants may be a determinant for the need to introduce CFs.

1.4.1.3. Growth of exclusively breastfed infants

Compared to formula fed infants, infants breastfed for at least 12 months grow more rapidly in the first 2–3 months and less rapidly (particularly in weight) from 3 to 12 months of age (Dewey, 1998). The growth pattern of breastfed infants is generally considered a healthier growth pattern. Indeed, many studies have shown that a high growth velocity during infancy is associated with an increased risk of non-communicable diseases such as obesity and cardiovascular diseases later in life (Singhal, 2017).

In a systematic review, Kramer and Kakuma (2012) did not find any differences in measures of growth of infants exclusively breastfed for 6 months compared with shorter durations of exclusive breastfeeding. In addition, the RCT by Jonsdottir et al. (2012) compared the effects on growth of exclusive breastfeeding for 6 months, with exclusive breastfeeding for 4 months followed by complementary feeding in addition to breast milk. Infants in both groups grew at the same rate between 4 and 6 months of age. In a follow-up study, there were no differences in anthropometric outcomes between both groups up to 29–38 months of age (Jonsdottir et al., 2014).

Several longitudinal or cross-sectional studies that assessed growth of exclusively breastfed infants for more than 6 months of age are available in low-income settings (Sidhu et al., 1981; Khan, 1984; Kumari et al., 1985; Rao and Kanade, 1992) and high-income settings (French, 1967; Ahn and MacLean, 1980; Salmenpera et al., 1985). Most of them showed a decline in the rate of weight and/or length gain after the age of 6 months (French, 1967; Sidhu et al., 1981; Khan, 1984; Kumari et al., 1985; Rao and Kanade, 1992). However, many studies have methodological limitations (e.g. small number of infants, lack of adjustment for confounding factors, high attrition rate) and/or were performed in low-income settings, thereby preventing firm conclusions being drawn on the adequacy of exclusive breastfeeding for more than 6 months in infants living in Europe.

The Panel concludes that exclusive breastfeeding for a duration of 6 months allows a normal growth pattern in most healthy term infants.

1.4.2. Nutritional adequacy of exclusive breastfeeding: overall conclusions

The Panel concludes that exclusive breastfeeding is nutritionally adequate up to 6 months for the majority of healthy infants born at term from healthy well-nourished mothers. However, some infants

at risk of iron depletion may benefit from the introduction of CFs that are a source of iron, before 6 months of age in addition to breastfeeding (see Sections 1.4.1.2 and 15.3).

1.5. Interpretation of the Terms of Reference

The appropriate age of introduction of CFs is influenced not only by nutritional considerations, but also by effects on health outcomes and by infant development. Aspects, such as social interactions and the cultural context, may also play a role but are not within the remit of the mandate. Considering the influence of various factors, the Panel considers it likely that there is an appropriate age range rather than a single appropriate age for the introduction of CFs. Taking into consideration the conclusions from Section 1.4.2 and the considerations above, EFSA interprets this mandate as follows:

To evaluate the appropriate age range for introduction of CFs to healthy infants, by answering the following questions:

1) Are there any developmental factors relevant for the introduction of complementary foods (CFs);
2) Is there evidence (based on a systematic literature review, Section 4 and following) to indicate that there would be (an) adverse (health) effect(s) for the child to have CFs introduced before the age of 6 months (selection of the age limit of 6 months based on conclusions of Section 1.4.2)?
3) Is there evidence (based on a systematic literature review, Section 4 and following) to indicate that there would be (a) benefit(s) for the child to have CFs introduced before the age of 6 months (selection of the age limit of 6 months based on conclusions of Section 1.4.2)?

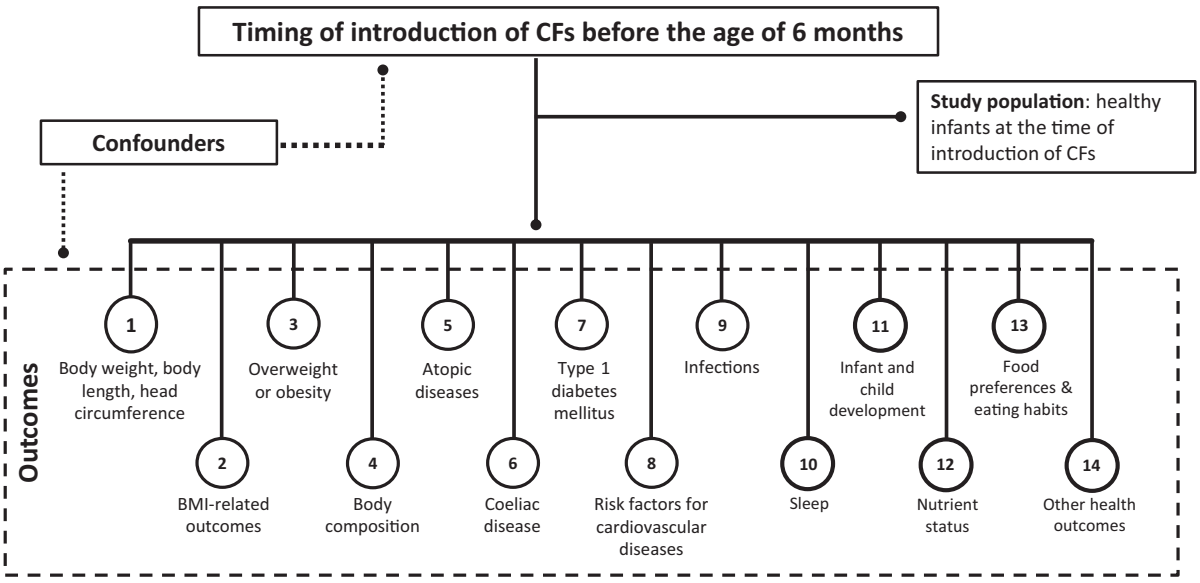
Out of the scope of this mandate are:

- public health recommendations for the introduction of CFs; this task is outside the remit of EFSA but it is the role of public health authorities in Member States;
- the effects of the duration of exclusive breastfeeding on the selected health outcomes, as the assessment is performed irrespective of whether infants were initially exclusively breastfed or formula fed;
- the health benefits of breastfeeding itself (for the infant/child and the mother);
- the effects on health outcomes of introduction of CFs solely after 6 months of age, as there is a nutritional requirement for CFs for the majority of exclusively breastfed infants from around 6 months onwards;
- the effects of the amount, order of introduction, variety, composition and texture of CFs;
- the role of aspects, such as social interactions and the cultural context, on the appropriate age of introduction of CFs;
- risks related to e.g. chemical or microbiological contaminants or pesticides.

1.6. General considerations on the outcomes assessed

(Health) outcomes that were considered in the systematic literature review (Section 4 and following) were identified *a priori,* in particular based on the Panel's previous Scientific Opinion (EFSA NDA Panel, 2009), and listed in a protocol for this assessment (EFSA, 2017b). The conceptual framework for this assessment is outlined in Figure 1.

Each outcome covered several endpoints (e.g. weight-for-age and weight-for-length). Compared to the protocol, a dedicated Section on BMI was created (Section 5), additional outcomes were considered when relevant studies were identified e.g. sleep (in a dedicated Section) or juvenile arthritis. The risk of type 2 diabetes mellitus was not discussed as no relevant data were identified on this outcome.



BMI: body mass index; CF: complementary food.

Figure 1: Conceptual framework for the systematic review on the appropriate age range of introduction of complementary foods (CFs) into an infant's diet

No limit on the length of follow-up between timing of introduction of CFs and the age at outcome assessment was applied during the literature selection, with the exception of the following:

- studies on growth in which the endpoint was measured at 6 months of age only, which were excluded (see Section 4.2 for reasons);
- studies on infections with an age at outcome assessment beyond 1 year of age (see Section 12.2 for reasons);
- studies on nutrient status with an age at outcome assessment beyond 1 year of age (see Section 15.2 for reasons).
- studies investigating outcomes at time points for which a relationship with the timing of introduction of CFs is unlikely considering the influence of the background diet on the outcome (e.g. kidney function at 6 years of age).

No exclusion criterion was applied in relation to the method of measurement of the outcome during the literature selection. The reliability of the different methods was considered in the appraisal of the risk of bias (RoB) (Appendix B). One exception was applied to a study that measured F2-isoprostane concentrations in spot urine samples (and not in 24-hour urine) as a marker of oxidative damage to lipids (Frederiksen et al., 2015). Spot urine samples are not considered an appropriate sampling unit for this outcome (EFSA NDA Panel, 2018).

Studies which reported on the attainment of individual developmental milestones in months were not considered in the systematic review (see Section 14.2 for reasons). However, they are discussed in the Section on the extensive literature search (Section 3).

The Panel notes that the studies selected for this assessment were heterogeneous with respect to the length of follow-up and the way in which the (health) outcomes were assessed.

2. Data and methodologies

A protocol was developed for this systematic review. It was subjected to public consultation (from 16 February to 23 March 2017) and amended as appropriate. The final version of the protocol described the methodology for data retrieval, study appraisal, data extraction and possible synthesis (EFSA, 2017b). It was published alongside a technical report on how the comments received during the public consultation were taken into account in the final protocol (EFSA, 2017a). Protocol amendments are listed in the following sections and Section 2.3. The EFSA guidance on the 'Application of systematic review methodology to food and feed safety assessments to support decision making' was applied for this assessment (EFSA, 2010).

2.1. Data

For all the (health) outcomes mentioned in Section 1.6, data selection and methodology followed the approach of a systematic literature review. For developmental readiness of term infants, in particular motor developmental milestones (called 'neuromuscular development' in the protocol), an extensive literature review was undertaken (as meta-analyses were not envisaged). The differences in the various steps between these two approaches (systematic or extensive) are explained in the following sections. For developmental readiness of preterm infants (Section 18.1), data came from a narrative review (in the following not further addressed).

2.1.1. Eligibility criteria for the systematic literature search

2.1.1.1. Inclusion

Study populations and exposures considered pertinent

Papers that were selected were only those investigating infants (i.e. aged 0 to < 1 year), children or adults, males and females, who were generally healthy at the time when they were introduced to CFs as infants and were either born at term or preterm (i.e. born at less than 37 weeks of gestation). These were considered pertinent study populations by the Panel for this assessment.

The study groups of the selected papers had to be alike in terms of the type of milk feeding (breast milk or formula[7] or mixed, with no additional behavioural interventions), i.e. the study groups had to differ only in the timing of introduction of CFs. In order to be included in this review, at least one study group had to have been introduced to CFs before 6 months of age (protocol amendment 2).

Introduction of CFs thus occurred with different types of milk feeding in the included studies, which compared:

- groups of exclusively breastfed infants introduced to CFs at different time points up to 6 months of age;
- groups of exclusively formula fed infants introduced to CFs at different time points up to 6 months of age;
- groups of infants receiving various types of background milk feeding (i.e. breast milk, formula, mixed) and introduced to CFs at different time points up to 6 months of age.

Introduction of a specific CF item or food group, irrespective of the introduction of other CFs, was also considered as providing potentially relevant information for some of the outcomes discussed in this assessment and mentioned in Section 1.6. Thus, studies which compared the early (before 6 months of age) vs later introduction of a specific CF item or food group were included if investigating the following outcomes:

- Atopic diseases: The specific foods considered were cereals (in particular wheat), egg, fish (as defined in the papers, i.e. generally undefined), peanut, soy (not in the form of formula), which are among the major food allergens relevant in children (EFSA NDA Panel, 2014a);
- Coeliac disease and T1DM: The specific food (item) considered was gluten and gluten-containing foods, as coeliac disease is triggered by the ingestion of gluten, found in wheat, barley and rye. For T1DM, gluten was considered relevant as the previous assessment of the Panel included specific conclusions on T1DM and gluten;
- Eating behaviours/food preferences: The specific foods considered were fruit and vegetables.

The studies were included irrespective of:

- the income of the population in the country in which the study was done, except for the outcome 'infections' as mentioned above;
- the age of assessment of the exposure, i.e. timing of introduction of CFs. This was, however, considered in the appraisal of the RoB (Section 2.2.2).

[7] The terms 'background milk feeding' are used in the following sections, even though the Panel is aware that formula is not 'milk' from a legal perspective.

Study designs and publication types considered pertinent

Articles were included if describing investigations based on the following study designs in humans:[8]

- intervention (experimental) studies;
- longitudinal prospective observational cohort studies;
- nested case–control studies with prospective data collection;
- letters to the editor, in a limited number of cases, i.e. if they provided sufficiently detailed information for assessment of the RoB and for data analysis (protocol amendment 4);
- retrospective studies[9] were included to assess the totality of the evidence in the context of a weight of evidence approach. The weight of evidence approach was not described in the protocol but was deemed necessary for transparent evidence integration (protocol amendment 8).

2.1.1.2. Exclusion

Study populations and exposures not considered pertinent

Human studies were not considered pertinent if they:

- focused on the duration of breastfeeding only or on the comparison of breastfeeding with formula feeding: e.g. studies that compared breastfeeding vs formula feeding independently of CF introduction, studies that compared the introduction of CFs at the same age in breastfed versus formula fed infants, or studies that investigated the nutritional content of breast milk or formula, the duration or promotion of any breastfeeding or the duration of exclusive breastfeeding without reporting on the timing of introduction of CFs;
- had an unclear definition of CFs, or defined CFs as including formula (Section 1.3), investigated the timing of introduction before 6 months of a specific food item/group not listed above (e.g. cow's milk for all outcomes, as the Panel considered that the effect of formula based on intact cow's milk protein and dairy products could not be disentangled);
- investigated the introduction of CFs (in general or specific foods) at ages only after 6 months (see above and protocol amendment 2);
- investigated texture (e.g. lumpy food introduction) or food diversity or preparation methods (e.g. home-cooked vs commercial baby foods) or composition of CFs or weaning methods (e.g. baby-led weaning);
- investigated growth or iron status in populations with high prevalence of undernutrition, wasting and/or stunting, in populations under clinical care or with diseases/disorders/medication use known to affect nutritional status (e.g. malaria and iron status);
- investigated the outcome 'infections' in low-income and lower-middle-income countries in settings with poor hygiene conditions (i.e. situations in which it is difficult to disentangle the relative effect of co-exposures on the incidence of respiratory and gastrointestinal infections from the effect of the timing of introduction of CFs on these outcomes; see Section 12.2 for reasons); low-income and lower-middle-income countries were identified according to the World Bank criteria, comparing the year in which the studies were conducted with the historical data of the World Bank per country.[10]

Study design and publication types not considered pertinent

The following design and publication types were not considered pertinent:

- *in vitro* studies;
- animal studies;
- case-only studies (i.e. on a relevant (health) outcome but composed of cases only, e.g. time to onset of coeliac disease or T1DM);
- publication types not providing sufficiently detailed information for assessment of the RoB and for data analysis or synthesis e.g. editorials or abstracts;
- narrative reviews;

[8] Randomised controlled trials (RCTs) may provide evidence of causality. Observational studies after adjustments for confounders can indicate the presence of an association. Associations shown in an observational study should also be interpreted in the light of possible reverse causality (Sections 4.2, 5.2, 6.2 and 8.2).

[9] i.e. case–control studies, sibling case–control studies, cross-sectional studies, cross-sectional analyses of otherwise prospective studies, retrospective cohorts and a prospective cohort in which exposure was assessed after the outcome.

[10] https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups

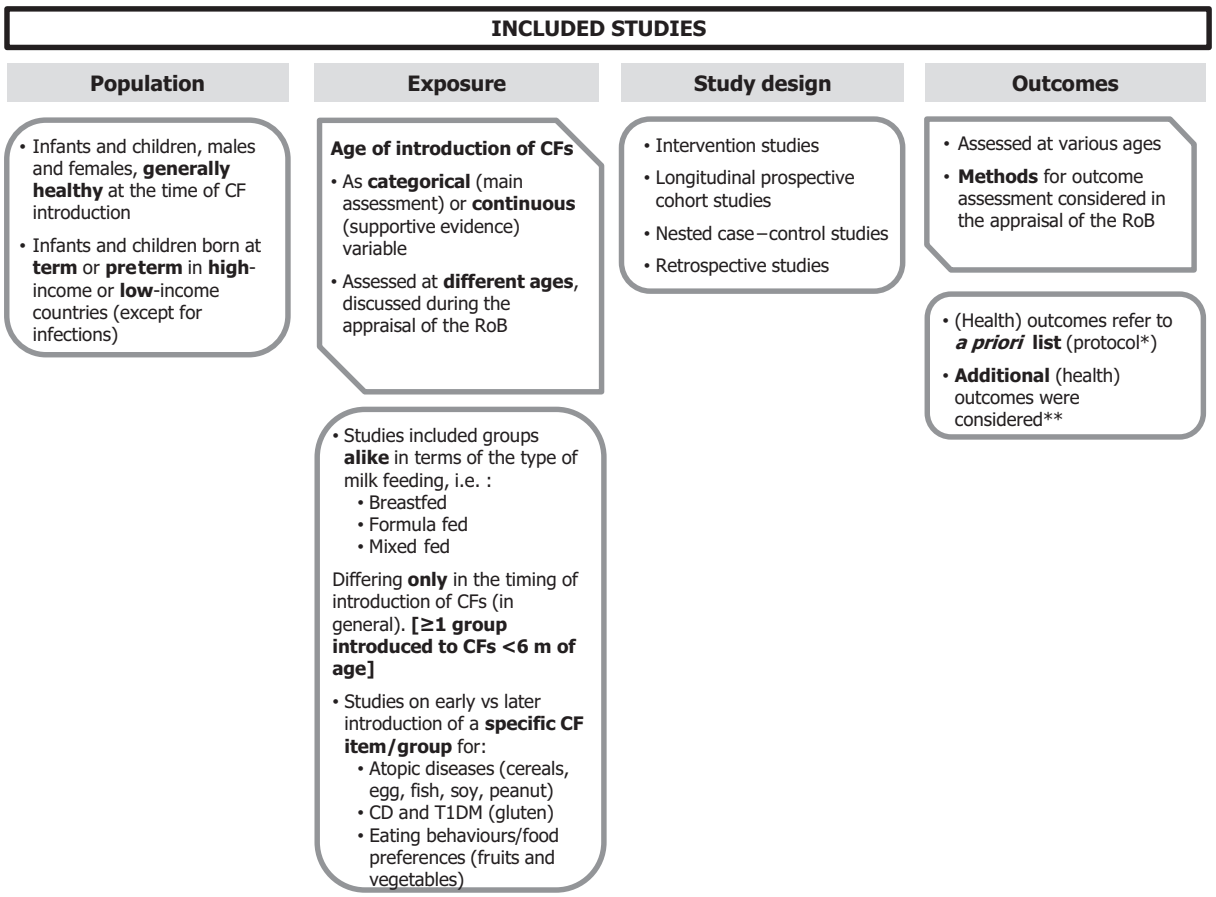
- systematic reviews with or without meta-analyses, and grey literature (i.e. conference abstracts, posters, dissertations, scientific reports). These were excluded from the assessment as such, and used only for hand search for peer-reviewed studies in their list of references;
- evidence-based guidelines comprising evidence-based and practice-based recommendations. Although a specific search and a quality assessment of evidence-based guidelines were required from an external contractor in the protocol (EFSA, 2017b), they were finally not used for this assessment (protocol amendment 4), in view of the large body of evidence coming from the peer-reviewed articles. However, some of these guidelines are mentioned in Section 1.2.

Additional exclusion criteria (protocol amendment 3):

Additional exclusions, not stated in the protocol (EFSA, 2017b), occurred at the 2nd step of the full-text screening (Section 2.1.1.2). The Panel estimated that the possible bias introduced by deciding on the exclusion of the following studies based on the knowledge of the evidence (and not *a priori* before study retrieval) was limited:

- Studies on growth in which the endpoint was measured in the first 6 months of life only (and not after) (see Section 4.2 for reasons);
- Studies on infections with an age at outcome assessment after 1 year of age and that did not cover the period during which CFs were introduced (see Section 12.2 for reasons);
- Studies investigating outcomes at time points for which a relationship with the timing of introduction of CFs is unlikely considering the influence of the background diet on the outcome (e.g. kidney function at 6 years of age);
- Studies on nutrient status with an age at outcome assessment after 1 year of age, e.g. Hb concentrations at 6 years (see Section 15.2 for reasons);
- Studies on nutrient status focussing on nutrients either non-critical for the European population of infants and young children or more influenced by other factors than the timing of introduction of CFs (see Section 15.2 for reasons);
- Studies on nutrient status reporting only on mean blood concentrations of biomarkers with no consideration of the proportion of subjects below a certain cut-off for nutrient sufficiency (see Section 15.2 for reasons);
- Studies on neurodevelopmental milestones reported in months or weeks only (see Section 14.2 for reasons);
- Studies in children reporting on bone mineral content (BMC) measurements not adjusted for bone area (see Section 7.2 for reasons);
- Studies on sensitisation to aeroallergens (see Section 8.2);
- Studies with inappropriate statistical analysis so that the results cannot be interpreted (e.g. matched (nested) case–control studies in which the matching factor was related to the exposure, but the matching was not taken into account in the analysis);
- Studies with undefined units of measurement.

Studies were excluded at the level of title or abstract screening, at the level of the first step of the full-text screening or at the second step of the full-text screening (Section 2.1.1.2, based on the criteria of protocol amendments 2 and 3). Annex C provides a list of 230 excluded references with the reasons for exclusion at step 2 of the full-text screening. These 230 references are composed of 221 references that were excluded overall and 9 references that were excluded from the assessment of certain outcomes, but included otherwise (Heinig et al., 1993; Cohen et al., 1994; Bainbridge et al., 1996; Mehta et al., 1998; Wilson et al., 1998; Kalanda et al., 2006; Hetzner et al., 2009; Jonsdottir et al., 2012; Nopporlertwong and Tantibhaedhyangkul, 2016).



CD: coeliac disease; CF: complementary food; m: months; RoB: risk of bias; T1DM: type 1 diabetes mellitus.
* EFSA (2017b); **For the complete list of outcomes, please see Figure 1 (Section 1.6).

Figure 2: Characteristics of the included human studies (body of evidence) from the systematic literature search

2.1.2. Eligibility criteria for the extensive literature search (developmental readiness)

2.1.2.1. Inclusion

Study populations considered pertinent and endpoints related to developmental readiness of term infants

Age of achievement of motor development milestones in (generally healthy) infants in relation to the introduction of CFs before 6 months of age was considered by the Panel as the relevant topic for this search.

The Panel was in particular interested in:

- when the extrusion reflex disappears,
- when the child is able to transport foods with the tongue to the back of the mouth,
- when the child gains some head control or postural control,
- when the child is able to sit with some support.

Publication types

- studies (whatever the design) described in peer-reviewed articles;
- reviews (either narrative or systematic);
- reports or books, when accessible.

2.1.2.2. Exclusion

The following exclusion criteria were applied.

Study populations not considered pertinent

- studies on subjects with a disease/disability, with no results from a healthy control group.

Study design and publication types

- *in-vitro* studies;
- animal studies;
- publication types not providing sufficiently detailed information, e.g. commentaries.

2.1.3. Considerations on the included data

The Panel notes that the studies selected were heterogeneous with respect to the length of follow-up, the methods and criteria used for the assessment of (health) outcomes, the study design, the way in which the exposure was assessed (i.e. the timing of introduction of CFs), the classification into exposure groups, the study settings and the populations investigated (Figure 2).

Heterogeneity is discussed per outcome/endpoint from Section 4 onwards.

2.2. Methodologies

Six literature searches were undertaken:

- four of them were systematic literature searches (see below);
- one was an additional quality check by EFSA based on artificial intelligence (see below);
- one was an extensive literature search by EFSA (on developmental readiness of term infants).

The general methodological approach regarding the systematic review (for the outcomes described in Section 1.6) was presented in broad terms in the protocol (EFSA, 2017b). The practical steps are described in the following sections and summarised in Figure 3. Some of them were not applied for the extensive review (developmental readiness of term infants) and this will be explained in the individual steps described below. Step f of the weighing and grading of the confidence in the evidence was not initially described in the protocol (protocol amendment 8).

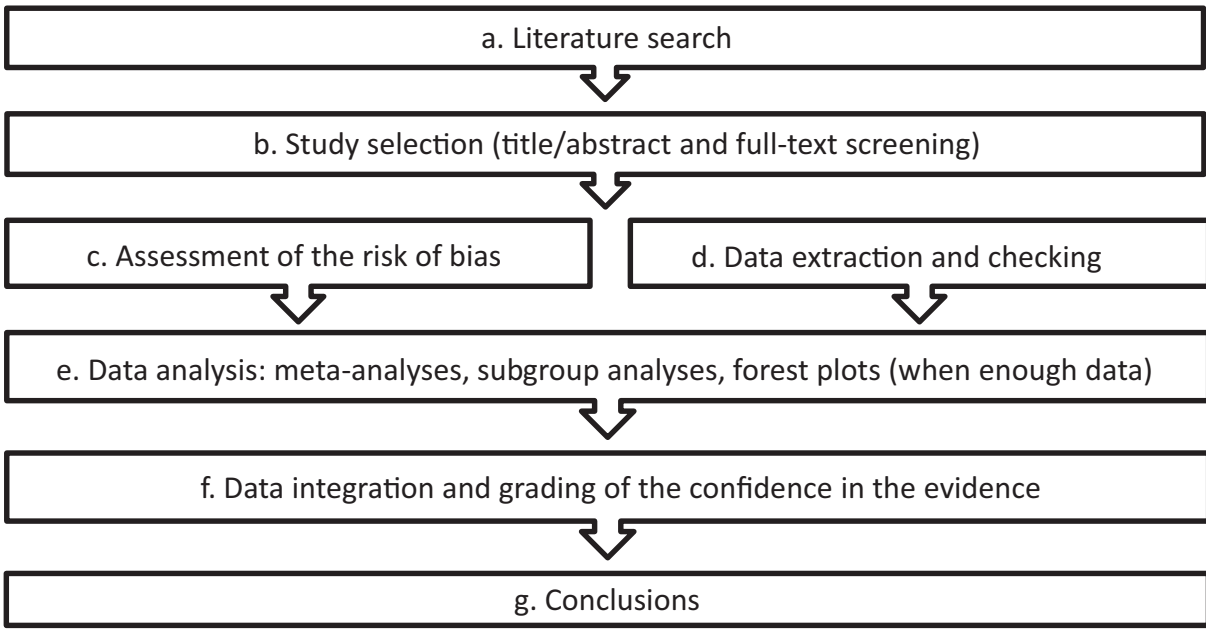


Figure 3: Methodological steps followed for the systematic review

2.2.1. Literature searches and study selection

2.2.1.1. Literature searches and study selection for the systematic review

Sources of information and publication date for published articles

Three databases were screened for the systematic literature searches, i.e. PubMed, Web of Science Core Collection and the Cochrane Library were searched for articles published since 1990. Data published before 1990 were obtained from hand searching in the reference lists of systematic reviews, grey literature and of the included primary studies. Thus, publication dates of the included studies ranged between 1973 and 2018.

Sources of information and publication date for grey literature (used for hand search, Section 2.1.1.2)

- five databases in addition to Google were used: the National Technical Information Service (NTIS[11]), the System for Information on Grey Literature in Europe,[12] CAB Abstracts, Open Access Theses and Dissertations,[13] the US National Guideline Clearinghouse;
- published since 2011 (conference abstracts or posters or dissertations) or most up-to-date versions (scientific reports and evidence-based guidelines).

Language

For the systematic searches, no language limits were applied. Studies described in articles not published in English were screened/extracted/appraised either based on the information provided by an EFSA staff member proficient in that language or based on the information provided by on-line translation tools. Eight studies in a language other than English, i.e. Chinese (Huang et al., 2013; Zheng et al., 2016), German (Forster et al., 1990), Japanese (Takahashi et al., 1999), Portuguese (Gomes et al., 2010) and Spanish (Bascunan Gamboa et al., 2012; Cu et al., 2015; Sandoval Jurado et al., 2016) were included.

Search strings

- for the first search, the search strings were created by an external contractor and are presented in the protocol and the report of the contractor (EFSA, 2017b; Pallas Health Research and Consultancy, 2019);
- for the other systematic searches, they were created by the information specialist of EFSA and are presented in Appendix D.

Dates and objectives of each of the searches

a) Initial literature search by the external contractor and quality check by EFSA

- For peer-reviewed articles, an external contractor conducted the initial search in May 2017, specifically on the 5th for Web of Science Core Collection) and 8th (PubMed and Cochrane Library) (Pallas Health Research and Consultancy, 2019);
- For grey literature, the contractor conducted the search in June/July 2017 (Pallas Health Research and Consultancy, 2019).

The number of papers (on prospective or retrospective studies) that were finally included by EFSA from this search is given in Table 1.

The following steps were undertaken by EFSA, after the initial literature search by the contractor:

- full-text screening step 2 (see below) and further exclusion, based on the criteria listed in Section 2.1.1.2;
- appraisal of the internal validity of the included studies, data extraction, presentation and synthesis (Sections 2.2.2);
- retrieval of relevant retrospective studies (protocol amendment 4) initially excluded by the contractor in line with the protocol). This retrieval was done by searching through the list of excluded papers provided by the contractor;

[11] www.ntis.gov
[12] www.opengrey.eu
[13] https://oatd.org/

- additional quality check: the 7,280 references excluded by the contractor were screened again by EFSA using a tool based on machine learning (artificial intelligence), i.e. 'ShinyR tool[14] for the automation of systematic review' that is available online in Zenodo[15] or in the web platform R4EU - Open Analytics.[16] This led to the identification of 1,037 references, which were screened first based on their title and abstract, and then on their full texts by EFSA staff members using single screening (i.e. not duplicate screening). The number of papers re-included is given in Table 1.

b) Complementary search

The information specialist of EFSA developed the search strings (Appendix D.1), and this complementary search by EFSA on 16 October 2017 retrieved:

- studies that included terms related to exclusive breastfeeding in the abstract (as such could have been considered not relevant), but that discussed complementary feeding in the full text;
- studies that were missing from the initial search (e.g. papers on timing of introduction of CFs and outcomes assessed in 'pre-school children').

The number of papers (on prospective or retrospective studies) included by EFSA from this search is given in Table 1.

c) Upgrade of searches a and b

Both the initial and the complementary searches ('a' and 'b') were updated and upgraded by EFSA on 2 October 2018 (Appendix D.2), to retrieve papers published since, respectively, May and October 2017. The search for grey literature was not updated.

Refined search strings (compared to those used in the initial search by the contractor) were developed by the information specialist of EFSA (protocol amendment 1):

- the search strings for countries were removed;
- some relevant terms for the population were added;
- the previous restriction on some study designs was removed (e.g. cross-sectional studies).

The results (number of hits) presented in Appendix D.2 included almost all those from the search of the contractor as no time limit was applied to the search. Duplicates were removed before the start of the screening process.

The upgraded searches were updated on 10 May 2019 to retrieve RCTs (protocol amendment 1) published since October 2018. Again, no time limit was applied, and duplicates were removed before screening. Search strings were those already used in the upgrade of searches a and b and are given in Appendix D.2.

The number of papers (on prospective or retrospective studies) that were included by EFSA from the upgraded and the updated searches is given in Table 1.

d) Hand search

EFSA staff hand-searched through the bibliography of:

- the studies included from all the searches described above,
- the systematic reviews (those performed by USDA (English et al., 2019b, English et al., 2019a; Obbagy et al., 2019c; Obbagy et al., 2019a; Obbagy et al., 2019b) and published shortly before the launch of the public consultation on this Scientific Opinion were searched during the public consultation and relevant papers were added thereafter),
- the theses found through the search of grey literature undertaken by the external contractor.

The number of additional papers that were included via hand-search is given in Table 1.

[14] Topic was used as feature space and the best prediction was obtained using the following ensemble: RF rose, NN rose, svm Linear rose, svm Poly rose, svm Poly smote, svm Linear smote, NN smote, svm Radial smote and GBM smote.

[15] EFSA. (2018, June 26). Shiny R tool for the automation of systematic reviews (Version v3). Zenodo. https://doi.org/10.5281/zenodo.1299654

[16] https://shiny-efsa.openanalytics.eu/

Study selection

For all searches ('a' to 'c' mentioned above), the study selection process was based on title and abstract and full-text screenings.

- **For the initial systematic search** by the external contractor ('a' mentioned above):
  - the study selection process (title/abstract screening and full-text screening) is described in a report (Pallas Health Research and Consultancy, 2019);
  - the outcome was provided as EndNote® databases to EFSA; a second step of full-text screening was applied by EFSA based on additional exclusion criteria (see below);
  - an additional quality check by EFSA was performed using an artificial intelligence tool (see above).
- **For the other systematic searches** ('b' and 'c' mentioned above):
  - the screening of the title and abstract was done in duplicate by EFSA staff members;
  - a full-text screening in two steps was undertaken:
    - the first step was based on the initial inclusion/exclusion criteria listed *a priori* in the protocol (EFSA, 2017b); it was done in duplicate and led to the exclusion of the studies irrelevant for this assessment;
    - the second step was based on the additional exclusion criteria described in Section 2.1.1.2 (protocol amendment 3); it was done by single screening (i.e. not in duplicate) and led to the further exclusion of papers (Annex C).

All systematic searches undertaken by EFSA ('b' to 'c' mentioned above) were screened in DistillerSR (Evidence Partners, Ottawa, Canada) and possible conflicts during the screening were discussed and resolved by EFSA staff.

2.2.1.2. Literature searches and study selection for the extensive review

Sources of information and publication date for published articles

Two literature databases were used for the extensive literature search, i.e. PubMed and Web of Science, without limiting the search with respect to publication dates.

Language

For the extensive literature search, only papers in English were selected.

Search strings

For the extensive literature search, search strings were created by the information specialist of EFSA and are presented in Appendix D.3, with the number of hits.

Dates and objectives of each of the searches

Specifically, for the aspects related to neuromotor developmental readiness of term infants, EFSA undertook an extensive literature search on 6 February 2019 in PubMed and Web of Science, for primary research studies and narrative or systematic reviews.

This led to the inclusion of 15 papers discussed in Section 3.3 (see final body of evidence further below and Table 1). These papers did not go through the steps described in the following sections, i.e. appraisal of the RoB, data extraction, data synthesis or grading the confidence in the evidence, as no meta-analysis was envisaged. EFSA staff also hand-searched through the bibliography of the included papers. The number of additional papers that were included via hand-search is given in Table 1.

Study selection

The study selection process was based on title and abstract and full-text screenings. The screening of title and abstract was done in duplicate by several EFSA staff members, and the full-text screening was done by a single EFSA staff member.

References were screened in DistillerSR (Evidence Partners, Ottawa, Canada) and possible conflicts during the title and abstract screening were discussed and resolved by EFSA staff.

2.2.1.3. Final body of evidence

The overall number of hits for the different steps of all these searches is provided as Table 1.

The total number of papers included in the systematic review on the relationship between the timing of introduction of CFs and health outcomes is 283.

The 201 papers on 131 prospective studies[17] included:

- 21 papers on 13 RCTs;
- 169 papers on 107 prospective cohort studies, of which:
    - 131 referred to 72 individual cohort studies with a specified name;
    - 38 referred to 35 individual cohort studies without a specified name;
- 9 papers on 9 nested case–control studies;
- 2 papers on 2 pooled analyses of prospective studies.

The 82 papers on 79 retrospective studies included:

- 12 papers on cross-sectional baseline analyses of 9 otherwise prospective studies (RCTs or prospective cohort studies);
- 29 papers on cross-sectional studies;
- 37 papers on case-control studies;
- 3 papers on retrospective cohort studies;
- 1 paper on a prospective cohort study in which the timing of introduction of CFs was assessed after the outcome.

This number is higher than that initially predicted in the protocol (EFSA, 2017b).

The total number of papers included from the extensive literature search in this assessment on motor development of term infants was 15. Another 8 publications were added by hand search and 2 papers were used that were originally retrieved through the systematic search. Thus, the total number of references discussed in relation to motor development was 25.

The total number of papers discussed in this Scientific Opinion is therefore 308.

[17] In some cases, several papers on the same study were available that investigated different outcomes and ages at which the outcome was assessed.

Table 1: Literature search flow

	Contractor search	Shiny R quality check in contractor excluded	Hand search in contractor excluded	Complementary search by EFSA	Upgrade contractor search by EFSA	Upgrade complementary search by EFSA	Hand search by EFSA	TOTAL	Developmental readiness
Initial hits	7,280	1,037	n/a	4,681	4,249 + 661[(b)]	1,877 + 446[(b)]	n/a	19,194[(a)]	1,412
Included full-text screening step 1	352	107	n/a	477	271 + 32[(b)]	115 + 22[(b)]	n/a	1,376	84
Included in full-text screening step 2	162	56	29	102	140 + 0[(b)]	21 + 0[(b)]	87	597	27
Duplicates	n/a	−2	n/a	−50	−41	−2	n/a	−95	n/a
Excluded full-text screening step 2 (without duplicates)	−44	−24	−10	−26	−65	−12	−40	−221	−12
Included papers on prospective studies	111	14	7	20	22	3	24	201	15 + 8 hand search + 2 from SR
Included papers on retrospective studies	7	18	12	6	12	4	23	82	n/a

n/a: not applicable; SR: systematic review.
(a): Excludes the 1,037 references re-screened using Shiny R;
(b): number of publications retrieved in the update of the search (limited to RCTs) performed during the time period when the draft Scientific Opinion was subjected to public consultation.

2.2.2. Assessment of the internal validity of studies included in the systematic review

Purpose and software

The appraisal of the included studies consisted of the assessment of their internal validity, i.e. their RoB. This was documented in the web-based systematic review software Distiller SR (Evidence Partners, Ottawa, Canada).

Study designs for which this step applied

The studies with the designs initially included based on the protocol (EFSA, 2017b) were those that went through the appraisal step: intervention studies (mainly RCTs), prospective cohort studies, pooled analyses of prospective studies and nested case–control studies.

The retrospective studies initially not included in the protocol but finally considered in this assessment (Section 2.2.1; protocol amendment 4) were not appraised. They were considered as being, by design, of high RoB (Tier 3).

Assessment at outcome level

For a study investigating several outcomes in relation to the age of introduction of CFs (e.g. symptomatic food allergy and weight), each outcome of the study was appraised individually, possibly leading to different assessments of the RoB.

Tool used and rating scale

The appraisal was based on the tool proposed by the US National Toxicology Program (NTP) Office of Health Assessment and Translation (OHAT) for conducting a literature-based health assessment (NTP, 2015), as mentioned in the protocol (EFSA, 2017b). The original set of questions proposed in the tool by NTP (2015) was reduced to those deemed most appropriate to the present assessment, as envisaged in the OHAT handbook (i.e. questions on reporting bias and on whether selection of study participants resulted in appropriate comparison groups in observational studies were dropped). The questions were answered on a four-level rating scale (low, probably low, probably high and high RoB). The protocol stipulated that judgements to the RoB questions should be combined into an overall RoB judgement (Tier of RoB), using an algorithm. The algorithm used for this assessment is described below.

Criteria to answer the individual questions and to combine them into an overall RoB judgement

The outline of the criteria used to answer each question is included in Appendix C. The rating for each question per study and outcome is presented in Annex B for RCTs and observational studies.

Four key questions were identified and used to conclude on the final RoB Tier (Table 2). Regarding the question on exposure to complementary feeding (detection bias) (EFSA, 2017b), the experts considered it relevant to formulate a specific question on compliance for intervention studies (Table 2; question 1).

The remaining questions (on concealed allocation and on blinding for RCTs, on other risks of bias for RCTs and prospective observational studies) that were considered in the protocol for the appraisal step were rated for completeness but did not influence the final overall allocation of a study to a RoB Tier.

In order for a study to be classified as Tier 1 (low RoB related to the outcome of interest), the publication must have been rated as 'definitely low' or 'probably low' RoB for all key questions. For a study to be classified as Tier 2 (intermediate RoB), one of the key questions, and for Tier 3 (high RoB), two of the key questions must have been rated as 'definitely high' or 'probably high' RoB.

Table 2: Four key questions which answers were combined into an overall judgement of the risk of bias (Tier of RoB)

Number	Key questions
1	– Observational studies: Can we be confident in the exposure characterisation (i.e. the assessment of the timing of introduction of CFs)? – Intervention studies: Can we be confident in the way compliance was assessed (i.e. can we be confident that complementary feeding was started/not started during the assigned time period)?
2	– Can we be confident in the outcome assessment?
3	– Observational studies: Did the study design or analysis account for important confounding variables? – Intervention studies: Was the study adequately randomised?
4	– Were outcome data incomplete due to attrition or exclusion from analysis?

Reviewers undertaking the appraisal

For the first half of the studies selected from the first two searches (Section 2.2.1.1), the appraisal was done in the full setting of the working group (WG) on Infant Nutrition, i.e. by each WG member present during the meetings. This resulted in an agreed rating of the individual RoB domains.

The appraisal for the remaining studies was done in parallel groups composed of half of the WG members, based on the experience gained.

Studies retrieved through the updated search (Section 2.2.1.2) were appraised by EFSA scientific staff members based on the same criteria established by the WG for the initial appraisal.

Insufficient information for appraisal

In case insufficient information was provided in a publication to allow an appropriate assessment of the RoB, the WG endeavoured to retrieve additional information:

- for example, additional information on the study methodology provided in other related publications or from original questionnaires, when publicly available.
- for RCTs or very large prospective studies for which information was missing on one or more items considered among the key questions, the authors of a limited number of papers (Brophy et al., 2009; Jonsdottir et al., 2012; Palmer et al., 2013; Vriezinga et al., 2014; Perkin et al., 2016) were contacted (protocol amendment 5).

If the information remained insufficient for an assessment, the 'probably high RoB' category was chosen by default.

2.2.3. Data extraction, presentation and synthesis in the systematic review

2.2.3.1. Data extraction

Data extraction was done in Microsoft Excel®. Data were extracted by one EFSA staff member and checked by a second EFSA staff member. The Microsoft Excel® files show all comparisons of age of introduction of CFs in a harmonised way, i.e. earlier introduction compared to later introduction.

Prospective (observational or intervention) studies

Most included studies considered the timing of introduction of CFs as categorical variable. In the following sections, studies in which the timing of introduction of CFs was used as a continuous variable in the analysis are identified as such in the text (the absence of such indication in the following sections means that the discussed study considered the timing of introduction as categorical).

The types of data extracted are listed in Table 3, and the detailed data are in Annex A (Microsoft Excel®).

- In cases where several models were reported in a paper, specifically an unadjusted model and several adjusted models with different sets of confounders, data from the fully adjusted model were extracted.
- In cases where data were reported in a paper for the 'full' study population as well as for subgroups, the data from the 'full' study population was extracted. An exception to this was if papers reported separately results for breastfed and formula fed infants: the data from such subgroups were extracted and used in the subgroup analyses and dedicated forest plots described in the following sections (Appendix A).

– For RCTs for which different types of analyses may be described (e.g. intention-to-treat (ITT), full analysis set (FAS), per protocol (PP)), the results of the most complete analyses (in most cases the FAS) were extracted. However, PP analyses may also be discussed in the following sections whenever needed.

Table 3: Type of data extracted and used for data presentation and synthesis

Type of data extracted	
Identification number of the comparison	Endpoint (e.g. attained BMI, WAZ)
Bibliography	Allergy to (e.g. egg, fish, peanut), if relevant
Inclusion (or not) in main analysis	Age at outcome assessment
Inclusion (or not) in subgroup analysis	Point estimate
Tier	Lower bound of the confidence interval (as reported in the paper or calculated by EFSA)
Study design	Upper bound of the confidence interval (as reported in the paper or calculated by EFSA)
Study name	Unit/Type (e.g. OR, RR)
Country (abbreviation)	Adjusted (yes/no)
At-risk group (yes/no)	Remarks
Heredity of allergy (yes/no, mixed population, unclear)	Statistical significance (significant/not significant)
Allergic symptoms at introduction of CFs (yes/no/unclear)	'Reverse causality addressed through' (e.g. sensitivity analysis)
Characteristics of the population (e.g. children with heredity of T1DM)	'Earlier introduction associated with' (in case of significant result)
Specific study group (e.g. breastfed, preterm)	Exposure assessment time point (e.g. multiple ≤ 6 months, as classified for appraisal, see Appendix B)
E1 (age of introduction of CFs when used as categorical variable, group 1)	Exposure assessment method (e.g. interview, questionnaire)
E2 (age of introduction of CFs when used as categorical variable, group 2)	Outcome assessment (e.g. parent's report of symptoms)
Age of introduction of CFs as a continuous variable (yes/no)	Reference data/cut-offs/method used (e.g. BMI ≥ P99 of CDC 2000)
N1 (number of subjects, group 1)	Food (e.g. CFs in general, egg, fish, gluten)
N2 (number of subjects, group 2)	Specific food (e.g. egg yolk if 'food' is egg)
Total N (total number of subjects of the comparison)	Amount (when available)
Section in opinion	Comparator (in RCTs)
Outcome (e.g. BMI, weight)	List of confounders (included, considered, not considered)

CDC 2000: growth charts by CDC released in 2000; BMI: body mass index; CDC: US Centers for Disease Control and Prevention; CF: complementary food; OR: odds ratio; RCT: randomised controlled trial; RR: risk ratio; T1DM: type 1 diabetes mellitus; WAZ: weight-for-age z-score.

Retrospective studies

The type of data extracted for retrospective studies was simplified compared to the list for prospective studies and is given in Table 4. The detailed data are in Annex A as Microsoft Excel® file.

Table 4: Type of data extracted and used for data presentation and synthesis

Type of data extracted	
Identification number of the comparison	Total N (total number of subjects of the comparison)
Bibliography	Section in opinion
Inclusion (or not) in main analysis	Outcome (e.g. overweight, blood pressure)
Study design	Endpoint (e.g. attained BMI, WAZ)
Study name	Allergy to (e.g. egg, fish, peanut), if relevant
Country (abbreviation)	Food (e.g. CFs in general, egg, fish, gluten)
At-risk group (yes/no)	Specific food (e.g. egg yolk if 'food' is egg)
Characteristics of the population (e.g. children with heredity of T1DM)	Age at outcome assessment
Specific study group (e.g. breastfed, preterm)	Point estimate
E1 (age of introduction of CFs when used as categorical variable, group 1)	Lower bound of the confidence interval (as reported in the paper or calculated by EFSA)
E2 (age of introduction of CFs when used as categorical variable, group 2)	Upper bound of the confidence interval (as reported in the paper or calculated by EFSA)
Age of introduction of CFs as a continuous variable (yes/no)	Unit/Type (e.g. OR)
N1 (number of subjects, group 1)	Adjusted (yes/no)
N2 (number of subjects, group 2)	Statistical significance (significant/not significant)

BMI: body mass index; CF: complementary food; OR: odds ratio; T1DM: type 1 diabetes mellitus; WAZ: weight-for-age z-score.

2.2.3.2. Data analysis and subgroup analyses and data presentation

Forest plots and estimates from meta-analyses

Data were visualised in forest plots whenever more than two studies were available for an endpoint. These forest plots are included in Appendix A.

In all forest plots representing RCTs and prospective observational studies (mostly prospective cohort studies), individual age comparisons from the included studies were organised in strata (subgroups) according to the following order:

- First, study design (i.e. separating RCTs from prospective observational studies);
- Second, RoB Tier: Tier 1 and Tier 2 studies were grouped together, separately from Tier 3 studies;
- Then, alphabetical order of the name of the first author.

Retrospective studies (Tier 3 by design) were represented in separate forest plots, in line with the approach outlined above to separate studies by their study design.

This allows an assessment if, for a given endpoint, the response is consistent or changes according to study design or RoB Tier.

In addition to the name of first author of the paper, the publication date, the study design and the RoB Tier, the forest plots display the following information:

- The point estimate for each age comparison with its 95% confidence interval (CI).
- The forest plots indicate whether each comparison was adjusted for the four to five confounders considered by the Panel as most relevant for each outcome (Appendices A and C). The forest plot also includes information if a comparison from an observational study was completely unadjusted for confounders. This is indicated by an 'N'.
- The forest plots also display the country as abbreviation, the age at outcome assessment, the age categories of introduction of CFs and additional information whenever needed (e.g. reference population on which z-scores are based).
- Whenever a single publication or several publications on the same cohort provided results for different ages at outcome assessment or different relevant populations, the results displayed in the forest plots were those which referred to the latest age at outcome assessment in the lowest RoB Tier and the most complete analysis set, unless the comparison was from an unadjusted analysis of an observational study. A similar approach was followed for the assessment of individual studies when no meta-analysis was possible.

– For studies on atopic-disease-related endpoints that reported on several interrelated endpoints (e.g. wheeze and asthma) or presented results for the same endpoints assessed in different ways (e.g. parents' report of a physician's diagnosis and parents' report of symptoms), the disease-related endpoint (e.g. asthma) and the most reliable outcome assessment (e.g. parents' report of a physician's diagnosis) (based on Appendix C) were included in the forest plots. A similar approach was followed for the assessment of individual studies when no meta-analysis was possible.

For comparisons not included in forest plots, the reader is still able to obtain information on the time course of the effect/association and on all endpoints assessed in a study in Annex A as Microsoft Excel® file.

Whenever possible, a pooled estimate of observed effect measures from individual studies (e.g. mean difference, OR etc.) was calculated using a random effects model carried out in the meta package of the R software (version R 3.5.0). Associated 95% CIs and prediction intervals were estimated for each stratum (as defined above) and represented in the forest plots.[18]

Heterogeneity index

The value of the I^2 together with its 95% CI is shown in the forest plots (Appendix A).[19]

The Panel notes that, when the number of comparisons/studies in meta-analyses is low, as is the case for several meta-analyses conducted by the Panel, the uncertainty associated with the I^2 estimate can be 'large'; therefore the I^2 has to be interpreted with caution (Ioannidis et al., 2007). With respect to the interpretation of I^2, the Panel followed the classification proposed by Higgins and Green (2011), also taken over by NTP (2015), i.e. 0–40% heterogeneity might not be important; 30–60% may represent moderate heterogeneity; 50–90% may represent substantial heterogeneity; 75–100% represents considerable heterogeneity.

For grading the confidence in the evidence (Section 2.2.3.3), the Panel focused on the estimate of I^2 (and not the 95% CI) as the most likely value of I^2.

Calculations, estimations and methodological approaches

A number of calculations and estimations were made by EFSA to produce the forest plots in case of missing summary statistics:

– When the articles did not report point estimates (e.g. OR) and associated 95% CIs, the point estimates and 95% CIs were either calculated based on the information reported numerically in the papers (e.g. number of subjects) or extracted by EFSA from graphs using an on-line tool.[20]
– Some studies reported point estimates without measures of spread. If an exact p-value was provided in these studies, 95% CIs were calculated from p-values. Otherwise, standard deviations (SDs) were imputed from other similar studies on the same endpoint. This was done for two papers on weight-for-age z-scores (WAZ) (Haschke and van't Hof, 2000; Gaffney et al., 2012), two papers on length(height)-for-age z-scores (L(H)AZ) (Dewey et al., 1999; de Beer et al., 2015) and two papers on BMI-for-age z-scores (BMIZ) (Haschke and van't Hof, 2000; Zheng et al., 2015).

The following methodological approaches were taken:

– The effect measures of the individual studies considered for a given endpoint were pooled and their 95% CIs estimated applying the Hartung and Knapp modification (Knapp and Hartung, 2003) to the DerSimonian and Laird approach (DerSimonian and Laird, 1986), a different approach to that originally described in the protocol (EFSA, 2017b) (protocol amendment 6). This method was used, as the DerSimonian and Laird approach without modification does not preserve the type 1 error rate[21] in situations where the number of comparisons/studies included in a meta-analysis is low and heterogeneity is high (Veroniki

[18] The pooled estimate of a random effects meta-analysis represents the estimated weighted average of the effect/associations observed in the different sub-populations and settings investigated. It should not be interpreted as an estimation of the 'true mean effect' (which is estimated by a fixed effect meta-analysis). The 95% CI provides information about the uncertainty around the weighted average. The prediction interval illustrates the range in which results of future studies in similar populations and settings could fall with a certain probability.
[19] I^2 is a measure of inconsistency in study results that cannot be attributed to sampling error.
[20] https://apps.automeris.io/wpd/
[21] The probability of rejecting the null hypothesis if it is true is > 5%.

et al., 2016; Jackson et al., 2017). This was a situation that was present for a number of meta-analyses performed by the Panel. The Hartung and Knapp modification was applied as it has been suggested that it may perform better in many situations and across types of outcomes (IntHout et al., 2014). However, it is not without criticism. Especially, it has been suggested to produce narrower 95% CIs than the DerSimonian and Laird approach in some instances, especially when τ^2 [22] is zero (Jackson et al., 2017). This is contrary to what is intended by the use of this modification.

- Therefore, sensitivity analyses were conducted, for all endpoints for which a random effects meta-analysis could be done, to check the relative performance of the DerSimonian and Laird approach with and without the Hartung and Knapp modification. In addition, the performance of another between-study variance estimator, proposed by Paule and Mandel (Paule and Mandel, 1989), again with and without the Hartung and Knapp modification was tested in the sensitivity analyses. When the Hartung and Knapp modification was not used, Wald-type CIs using a t-distribution with k-1 degrees of freedom were derived. The results are reported in Annex E. Indeed, in some instances applying the Hartung and Knapp modification led to narrower CIs as compared with not using this modification. This was the case for:
    - attained body weight (Appendix A.3): subgroups of 1) RCTs rated as Tiers 1 and 2 (3 studies, $\tau^2 = 0$, $I^2 = 0\%$ [95% CI 0 to 38%]) and 2) prospective cohort studies rated as Tiers 1 and 2 (4 studies, $\tau^2 = 0$, $I^2 = 0\%$ [95% CI 0 to 2%]);
    - attained body length/height (Appendix A.8): subgroups of 1) RCTs rated as Tiers 1 and 2 (3 studies, $\tau^2 = 0$, $I^2 = 0\%$ [95% CI 0 to 0%]) and 2) prospective cohort studies rated as Tiers 1 and 2 (3 studies, $\tau^2 = 0$, $I^2 = 0\%$ [95% CI 0 to 34%]);
    - attained body length by feeding mode (Appendix A.9): subgroup of formula fed infants (4 studies, $\tau^2 = 0$, $I^2 = 0\%$ [95% CI 0 to 25%]);
    - attained head circumference (HC) (Appendix A.10): subgroups of RCTs rated as Tiers 1 and 2 (3 studies, $\tau^2 = 0$, $I^2 = 0\%$ [95% CI 0 to 0%]);
    - odds of developing (at least) overweight (Appendix A.16): subgroup of prospective cohort studies rated as Tier 3 (10 studies; $\tau^2 = 0.01$, $I^2 = 61\%$ [95% CI 23 to 81%]);
    - asthma-like symptoms and fish – general population (Appendix A.24): subgroup of prospective cohort studies rated as Tiers 1 and 2 (3 studies, $\tau^2 = 0$, $I^2 = 0\%$ [95% CI 0 to 15%]);
    - eczema and fish – general population (Appendix A.31): subgroup of prospective cohort studies rated as Tiers 1 and 2 (2 studies, $\tau^2 = 0$, $I^2 = 0\%$);
    - risk of iron depletion in exclusively breastfed infants at 6 months of age (3 studies, $\tau^2 = 0$, $I^2 = 0\%$ [95% CI 0 to 4%]) (Appendix A.48).

  In none of the above-mentioned cases, the results with respect to their statistical significance changed according to the method applied. Considering that the Hartung and Knapp modification in these instances did not perform well, the results of the DerSimonian and Laird approach without the modification is indicated underneath the respective forest plots and used for reporting in the Scientific Opinion. In all other cases, the results of the Hartung and Knapp modification are reported. The Paule and Mandel approach gave similar results to the DerSimonian and Laird approach and they were therefore not considered further in the discussion of results. In some instances,[23] the DerSimonian and Laird approach showed statistically significant results, while when the Hartung and Knapp modification was applied, the findings became non-significant. However, considering the increased false positive rate of the unmodified DerSimonian and Laird approach, the results with the Hartung and Knapp modification were considered.
- Prediction intervals (95% level) were estimated following the DerSimonian and Laird approach and using a t-distribution with k-2 degrees of freedom, whenever more than 2 studies were available per subgroup.
- Some studies reported on more than two comparisons for the ages of introduction of CFs (e.g. a study comparing < 4 months vs >6 months, 4–5 months vs > 6 months, 5–6 months vs > 6 months). In these cases, the correlation among comparisons including a common

[22] Estimate of the between-study variance.

[23] (1) attained body length: prospective cohort studies in Tier 3; (2) symptomatic food allergy and CFs (general population): prospective cohort studies in Tier 3, (3) sensitisation and CFs (at-risk populations): prospective cohort studies Tiers 1 and 2, and (4) sensitisation and egg (at-risk populations): RCTs Tiers 1 and 2.

reference category was considered by combining the estimates to obtain a single comparison for each study (e.g. ≤ 6 months vs > 6 months) (Higgins and Green, 2011), whenever possible.[24] Sensitivity analysis showed that this had little impact on the estimate and the associated CIs of the meta-analysis. Detailed information for each comparison remains available in Annex A (Microsoft Excel®).

Sensitivity and subgroup analyses

The possibility of sensitivity and subgroup analyses was mentioned in the protocol with some examples (EFSA, 2017b). Sensitivity analyses which were conducted were as follows:

- applying alternative approaches to the estimation of CIs (see above);
- for some forest plots (Appendix A), to test the influence of a specific study on the pooled estimates and heterogeneity.

The protocol also stipulated that the following sensitivity analyses were to be conducted: (1) a 'sensitivity analysis to assess the impact of imputed summary data' and (2) a sensitivity analysis on the inclusion of studies with high participant attrition (or with other missing data)'.

As the number of studies for which imputation was made by EFSA was low, such a sensitivity analysis was not undertaken. In addition, most included studies had high attrition, and attrition/exclusion was considered among the key questions in the assessment of the RoB (Section 2.2.2 and Appendix C). Therefore, this impacted on the attribution of studies to Tiers of RoB (according to which data were pooled in meta-analyses, as explained above).

Enough studies were available for subgroup analyses regarding type of 'milk' feeding at the time of introduction of CFs (exclusively breastfed vs exclusively formula fed infants), for five endpoints: WAZ, attained body weight, BMIZ, L(H)AZ, attained body length. This was done on studies of RoB Tiers 1 and 2 only and was irrespectively of the study design.

Unplanned subgroup analyses were undertaken by EFSA:

- for atopic diseases, data from populations at-risk and the general population were analysed separately (protocol amendment 7): This was done to investigate whether different associations were observed for the general population and for the at-risk population (a specific subpopulation of the general population). If associations were indeed not different, the Panel considered that data generated in at-risk populations could be generalised to the whole population of infants living in Europe.[25] Therefore, if not stated otherwise in the text, the conclusions of the Panel on atopic-diseases apply to the whole population of infants living in Europe, which is the target of this mandate.
- for coeliac disease, data were analysed according to age of introduction of CFs: This was done to investigate whether there is a differential effect of introducing gluten < 4 months of age and between 4 and 6 months of age on the risk of coeliac disease as concluded by the Panel in its previous Scientific Opinion (EFSA NDA Panel, 2009).

Limitations of the meta-analyses

The Panel notes that for some meta-analyses, the number of studies that could be considered by subgroup/stratum was low. However, the Panel wishes to highlight that the pooled estimates were calculated with the objective to summarise the data and describe the direction of the effect or association (if any observed). The uncertainty that is associated with meta-analyses with a low number of studies, especially when heterogeneity is high, is expressed in the wide CIs around the point estimates. This uncertainty was addressed through the grading of the confidence in the evidence (Section 2.2.3.3). In addition, whenever the meta-analysis in a subgroup/stratum was based on two studies only, both the results of the meta-analysis and the individual studies are discussed.

[24] This was not possible when the reference category was the 'middle' category. In this case, the individual comparisons were included in the analysis.
[25] If no other limitations prevented the generalisation of results.

Publication bias

Publication bias[26] was assessed by EFSA in the body of evidence by generating funnel plots[27] whenever ≥ 10 comparisons/studies were available (Higgins and Green, 2011). Funnel plots are shown in Annex D. The reference line that was used in the funnel plots was the value of the pooled estimate of the random effects meta-analysis.

Whenever the funnel plots indicated asymmetry from visual inspection, the Egger's regression test (Egger et al., 1997; Sterne and Egger, 2005) and the trim-and-fill analysis (Duval and Tweedie, 2000) were used. In case of results suggesting asymmetry obtained from the aforementioned analyses (i.e. for odds of developing (at least) overweight, odds of developing obesity, BMIZ (Sections 5 and 6)), contours of statistical significance were overlaid on the funnel plot (Peters et al., 2008) (data not shown).[28]

2.2.3.3. Reporting, evidence integration and grading of the confidence in the evidence

Reporting of the ages of introduction of CFs

Ages at introduction of CFs reported in this Scientific Opinion should be interpreted bearing in mind the uncertainties that are associated with the ages that are described in the included papers. For RCTs, the ages reported in the Scientific Opinion are those when infants were randomised to start consuming CFs. However, variability is to be expected as to when infants were actually introduced to CFs or were able to consume the assigned CFs after randomisation. This may well span over some weeks. Therefore, reported ages for RCTs should not be interpreted as a single time point, but rather as a time span of one month. For example, introduction of CFs at 4 months should be read as introduction during the fifth month of life and an introduction of CFs at 3–4 months of life as an introduction during the fourth and fifth months of life. With respect to the reporting of observational studies, it should be noted that the uncertainty in the ages that are described in the included papers is usually higher than those reported for RCTs. This is owing to the mostly retrospective assessment of the timing of introduction of CFs and the lack of exact definition of the ages that are reported in the papers. Therefore, the Panel considered it valid to summarise the timing of introduction of CFs in individual studies, in particular RCTs, into an overarching age range of introduction of CFs in these studies (e.g. an introduction of CFs at 3–4 months and one at 4 months is summarised into an age of 3–4 months).

Preterm and term infants

The Panel decided to present and discuss in separate sections the results on infants born at term from the results on preterm infants, because of their differences in developmental steps and nutritional requirements (Section 18). Papers that did not specify if the included children were born at term or not or included both populations ('mixed populations') were grouped with papers on children born at term.

Main and supportive lines of evidence

In the following sections, for each outcome, the main line of evidence is discussed first. It consists of RCTs and prospective observational studies rated as Tiers 1 and 2. Within this line of evidence, endpoints for which forest plots could be created are discussed first. The conclusions by the Panel are based on the results of the meta-analyses and not the individual studies, unless only two studies were available in a subgroup/stratum (Section 2.2.3.2 above). Studies providing information on p-values without point estimates are discussed individually.

The supportive line of evidence consists of:

- RCTs and prospective observational studies rated as Tier 3;
- Retrospective studies (i.e. case–control studies, studies in which the exposure was assessed at the same time as the outcome or thereafter, and retrospective cohort studies; all Tier 3 because of the study design, as mentioned in previous sections);
- Studies in which the timing of introduction of CFs was used as a continuous variable in the analyses (whatever RoB Tier);

[26] i.e. the selective publication of positive findings.

[27] Funnel plots show the effect estimate from individual studies (plotted on the x-axis) against a measure of variability represented on the y-axis, such as the standard error. Smaller and less precise studies tend to scatter at the bottom of the 'inverted funnel'.

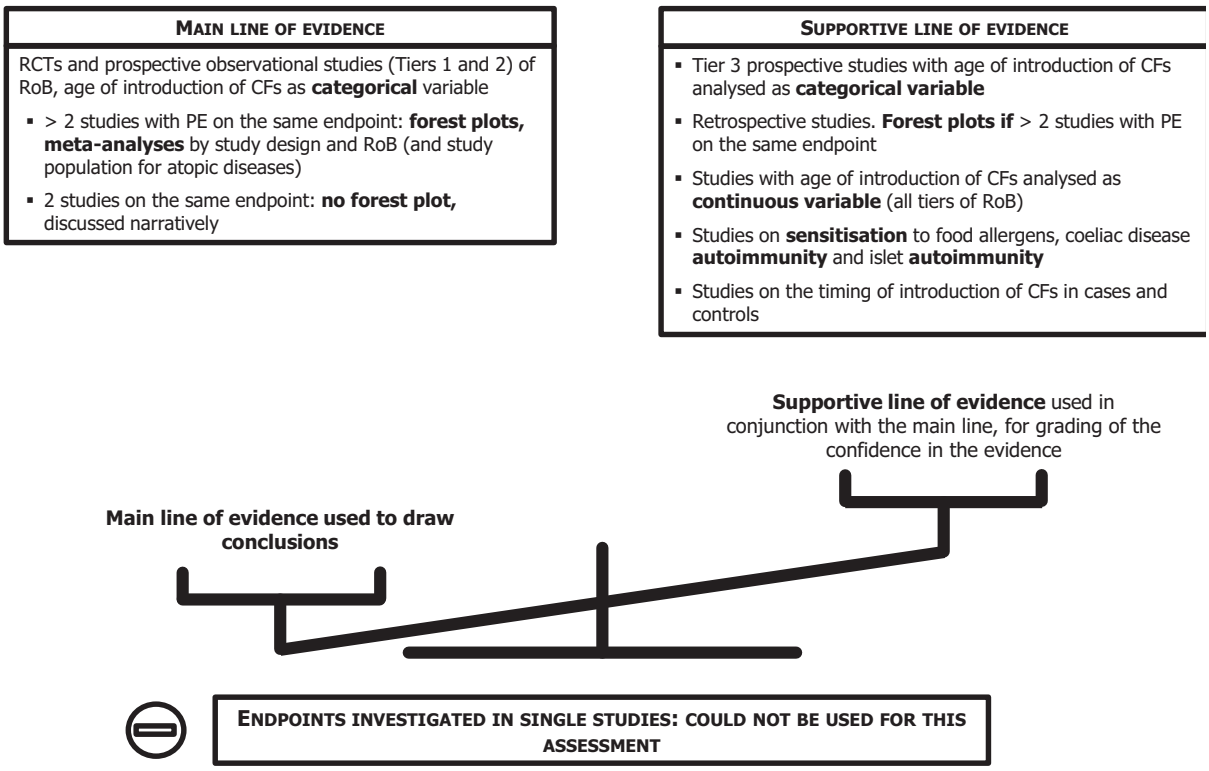
[28] If studies are missing in the area of non-statistical significance (i.e. $p \geq 0.05$), publication bias is a possible explanation for the asymmetry that is observed; if studies are missing in the area of statistical significance (i.e. $p < 0.05$) factors other than publication bias are the more likely explanation for asymmetry.

- Studies on sensitisation to food allergens (i.e. considered supportive compared with data on symptomatic food allergy), coeliac disease autoimmunity (i.e. considered supportive compared with data on coeliac disease) and islet autoimmunity (i.e. considered supportive compared with data on type 1 diabetes mellitus);
- Studies on the timing of introduction of CFs in cases and controls.

The Panel considers that the evidence from studies in the supportive line of evidence is insufficient by itself to draw conclusions on an appropriate age range of introduction of CFs. This is either because of the high RoB (Tier 3) or because they do not directly address the research question, i.e. studies in which the timing of introduction of CFs is used as a continuous variable in the analyses do not allow to conclude on an appropriate age range of introduction of CFs, and also studies on sensitisation or coeliac disease and islet autoimmunity do not allow to draw conclusions on the disease. Therefore, the evidence from studies in the supportive line of evidence is only used in conjunction with evidence from the main line of evidence (see below for the approach for grading the evidence).

The Panel considers that endpoints investigated in single studies (pertaining to the main or supportive lines of evidence) or only in studies in the supportive line of evidence could not be used to establish the appropriate age range of introduction of CFs. Thus, they are mentioned in the following sections but are not considered further by the Panel.

This hierarchy of the available evidence is described in Figure 4.



CF: complementary food; PE: point estimate; RCT: randomised controlled trial; RoB: risk of bias.

Figure 4: Hierarchy of the available evidence discussed in the systematic review

Evidence integration and level of confidence

The determination of the level of confidence followed an approach that was inspired by the approach proposed by OHAT (NTP, 2015). Evidence derived from RCTs was initially attributed a high confidence level (i.e. ++++), evidence derived from prospective observational studies was considered to provide a moderate confidence level (i.e. +++) and the evidence derived from retrospective studies was considered to have a low confidence level (i.e. ++). Whenever a study was most likely underpowered for an endpoint and did not show a statistically significant association or effect, the Panel excluded it from the integration of the evidence, because its findings were not considered reliable.

The initial level of confidence in the evidence could then be downgraded:

- for the RoB (i.e. prospective observational studies rated as Tier 3).
- for inconsistency in the findings; whenever a meta-analysis was available (Appendix A), substantial inconsistency was considered to be present when heterogeneity as identified by I^2 exceeded 75%[29] and could not be explained. Whenever I^2 exceeded 75% and, from visual inspection of the forest plot, this was most likely attributable to a single study, sensitivity analyses were performed (protocol amendment 7) by removing the study from the analysis to investigate its impact on I^2 and the results of the meta-analysis. This was done for (i) one prospective cohort study on BMIZ (Section 5.3), (ii) one retrospective study on obesity (Section 6.3) and (iii) one prospective cohort study on coeliac disease (Section 9.3).
- for serious imprecision (i.e. a wide 95% CI; to assess imprecision, results from the meta-analysis were prioritised over results from individual studies). For odds/risk/hazard ratios (OR, RR, HR), the Panel considered that the estimate was imprecise if the upper bound of the CI divided by the lower bound of the CI was higher than 10. For other kinds of measurements (e.g. BMIZ), the Panel considered an estimate imprecise when the CIs were wide and if the lower or the upper bound of the 95% CI was indicative of biological relevance of the finding.
- for limitations in the generalisability of the findings, i.e. (i) when lines of evidence consisted only of studies in exclusively breastfed infants or only of studies in exclusively formula fed infants (unless there was evidence that the background milk feeding was not an effect modifier); (ii) for atopic diseases, lines of evidence consisting only of studies performed in countries with a prevalence of the respective disease that is different from Europe and this difference could not be explained.
- for evidence of publication bias.
- if the main line of evidence contained only one study or, for endpoints related to atopic diseases, only one study per population group (i.e. lack of replication) and when (1) the supportive line of evidence was non-existent, (2) consisted of only one study or (3) provided inconsistent findings.

The confidence could also be upgraded when

- the effect or association in the line of evidence was large (the magnitude of the effect was defined as large when the RR or the OR exceeded 2 or was less than 0.5) and
- when an indication for a dose-response was available.

Aspects other than those listed above were also considered (e.g. discrepancies in the findings of the FAS analysis and the PP analysis), if they increased or decreased the confidence in a finding.

Confidence levels were truncated ++++ at the upper end and at + at the lower end.

Three conclusions were possible when findings were statistically non-significant or of no biological relevance:

1) 'no effect': for conclusions derived from RCTs with a high level of confidence in the evidence;
2) 'no evidence for an effect': for conclusions derived from RCTs with a very low, low or moderate level of confidence in the evidence;
3) 'no evidence for an association': for conclusions derived from prospective observational studies, irrespective of the level of confidence in the evidence.

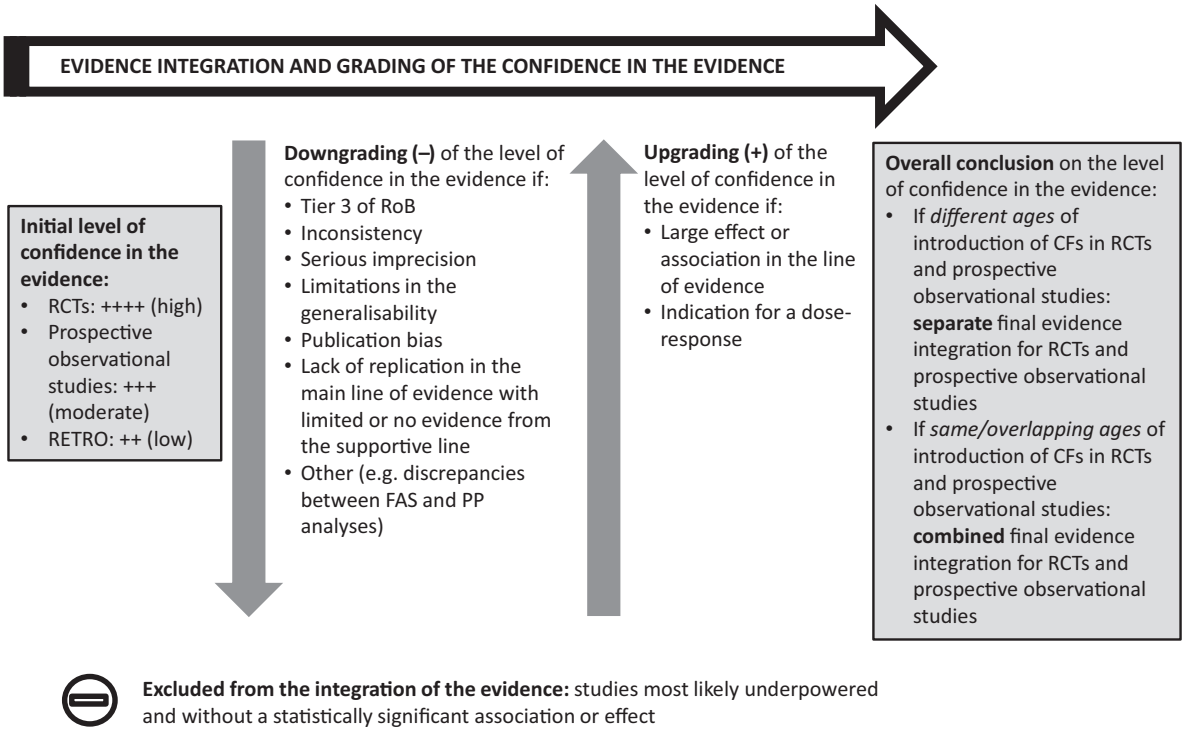
In case there was insufficient evidence to conclude if the timing of introduction of CFs was associated with an outcome, no level of confidence was derived.

For all outcomes considered in the following sections, a paragraph on grading the confidence in the evidence lists the main considerations in relation to imprecision, inconsistency, generalisability, and publication bias, and is followed by the overall conclusions of the Panel.

In the overall conclusions of the Panel, for most outcomes, RCTs were considered separately from prospective cohort studies, as the ages of introduction of CFs were different between these study designs. This allowed a higher confidence to be attributed to the conclusions derived in relation to the ages of introduction of CFs studied in RCTs, while the confidence in the evidence for the age ranges studied in prospective cohort studies was lower. The results from RCTs and prospective cohort studies were integrated only when the ages of introduction of CFs in both study designs overlapped. The overall conclusions of the Panel per outcome refer to the main line of evidence only (and not the supportive line).

[29] Lower bound for considering heterogeneity as considerable.

The approach for integrating the evidence and grading the confidence in the evidence is described in Figure 5.



CF: complementary food; FAS: full-analysis set; PP: per protocol; RCT: randomised controlled trial; RETRO: retrospective studies; RoB: risk of bias.

Figure 5: Approach for integrating the evidence and grading the confidence in the evidence

2.3. Protocol amendments

As previously discussed, the following amendments to the protocol (EFSA, 2017b) have been made:

1) Upgrade and update of the literature search: the search strings were upgraded for the updated search that was performed before the release for public consultation; the additional literature search before the final adoption of the Scientific Opinion was limited to RCTs.
2) Eligibility criteria: consideration of studies in which at least one group was introduced to CFs before 6 months of age (Sections 1.4 and 2.1.1.1), instead of studies involving infants not older than 12 months of age at introduction of CFs as mentioned in the protocol.
3) Eligibility criteria: additional exclusion criteria (Section 2.1.1.2).
4) Eligibility criteria: in a few cases, letters to the editor were included, if they provided sufficiently detailed information for assessment of the RoB and for data analysis or synthesis (instead of excluding all letters to the editor as mentioned in the protocol). Retrospective studies were also included. Evidence-based guidelines comprising evidence-based and practice-based recommendations collected and assessed by the external contractor, were finally not used for this assessment.
5) Missing data: gathering missing data for an appropriate assessment of the RoB was initially not mentioned in the protocol but was done in some cases described in Section 2.2.2 by contacting authors.
6) Data analysis: CIs of the pooled estimates were calculated based on the Hartung and Knapp modification to the DerSimonian and Laird approach (instead of using just the DerSimonian and Laird approach originally mentioned in the protocol) and sensitivity analyses were carried out.
7) Unplanned subgroup or sensitivity analyses: (i) for populations at-risk of different atopic diseases and for the general population; and (ii) for assessing the impact of one potentially influential study on the I^2 and the results of a meta-analysis (in case I^2 exceeded 75%).

8) Final step in the assessment: weighing the evidence and grading of the confidence in the evidence.

3. Assessment of the developmental readiness of the term infant to consume CFs

Developmental readiness can be defined as the physiological maturation necessary for an infant to metabolise 'non-milk foods' i.e. other than breast milk or formula, and the neurodevelopmental changes necessary for safe and effective progression from suckling to spoon- and self-feeding, including the infant's apparent emerging interest in non-milk foods and feeding. Developmental skills necessary to consume CFs will differ depending on the texture of the food. The skills needed for spoon-feeding of pureed CFs will appear earlier than the ones required for self-feeding and therefore, will be used to define the lower bound of the age range of developmental readiness.

3.1. Gastrointestinal function

The human gut is anatomically and functionally mature at birth in the healthy term infant, although the secretion and activity of gastric and pancreatic enzymes are not developed to adult levels (EFSA Scientific Committee, 2017). These functions mature at very different rates (EFSA Scientific Committee, 2017) and the ingested foods appear to play a part in triggering the maturation of gastric and pancreatic enzymes (WHO, 1989).

The Panel notes that gastrointestinal function is not a limiting factor with respect to the timing of introduction of CFs once the infant has the necessary neuromotor skills and has developed an apparent interest in non-milk foods and feeding.

3.2. Renal function

The renal control of water balance is not fully developed at birth. The rate of renal water excretion is influenced by the solute load to be excreted.[30] As renal concentrating capacity is limited in the neonatal period (Joppich et al., 1977), a high solute load could result in rapid and profound alteration in water balance.

The renal concentrating ability was reported to increase rapidly in healthy term infants within the first month of life, with average individual maximum values for osmolality in urine samples of 515 mOsm/L on day 3, 663 mOsm/L on day 6 and 896 mOsm/L in the first month of life. Thereafter, the increase is attenuated, with average individual maximum values for osmolality reached at 10–12 months of 1,118 mOsm/L and of 1,362 mOsm/L at 14–18 years of age (Polacek et al., 1965).

The Panel notes that renal function is not a limiting factor with respect to the timing of introduction of CFs once the infant has the necessary neuromotor skills and has developed an apparent interest in non-milk foods and feeding.

3.3. Neuromuscular coordination and neurodevelopment

At term, healthy infants are able to coordinate efficient suckling, swallowing and respiration (Bu'Lock et al., 1990; Morris and Klein, 2000) as a result of five feeding reflexes that develop prenatally: swallowing, sucking, gag, phasic bite and rooting. Changes in these reflexes over time combined with anatomical changes in the infant jaw and tongue facilitate the subsequent progression to solid foods (Stevenson and Allaire, 1991).

At birth, the tongue occupies most of the oral cavity (Bogaerts et al., 2012); the soft palate, the pharynx and the larynx are in close proximity. This protects airways from aspiration of liquids. However, it also limits the movements of the tongue with no room for up-down or lateral movements and chewing (Morris and Klein, 2000). During the first months of life, with head and neck growth, the oral cavity and upper pharynx enlarge (Stevenson and Allaire, 1991; Arvedson and Lefton-Greif, 1996) and free space for more refined tongue movements and for the infant to receive foods other than liquids (Morris and Klein, 2000). Initial suckling with peristaltic tongue movements decreases and is replaced by sucking with increasing voluntary up and down movements of the tongue. This is less automatic and requires more neurological control. These changes have been estimated to initiate at

[30] Renal solute load is derived from exogenous and endogenous sources. The former comes mainly from electrolyte intake, while the latter results from metabolism, particularly nitrogenous end-products related to protein metabolism.

around 3–4 months of age (Arvedson and Lefton-Greif, 1996; Morris and Klein, 2000). However, the Panel was unable to retrieve empirical data for this estimation.

While the swallowing reflex persists, other reflexes disappear or diminish. The gag reflex becomes less intense and is elicited over a smaller area of the tongue after about 6 months of age, whilst rooting disappears after about 3–4 months (Stevenson and Allaire, 1991). Another reflex that young infants' exhibit and that disappears with time is the tongue-thrust reflex, also called the extrusion reflex, when an object touches the infants' tongue or lip. As a reaction, the tongue moves forward and pushes any material, including food, that is placed on the infant's tongue outwards (Rogers and Arvedson, 2005). It has been estimated that this reflex diminishes and finally disappears between around 4 and 6 months of age (Sheppard and Mysak, 1984; Rogers and Arvedson, 2005). However, the Panel was unable to retrieve empirical data for this estimation.

Another aspect to consider is that, in order to efficiently accept spoon-fed foods, the infant has to be able to move the upper lip down to wipe the food from the spoon with the lips (instead of suckling it off the spoon) (Stevenson and Allaire, 1991; Ayano et al., 2000). It also must possess the necessary oral-motor functions that permit the tongue to receive food on its surface, form a bolus, lift it up and press it against the hard palate to transport it to the back of the mouth where the swallow reflex is triggered. This is a complex motion that requires oral structures to move independently instead of moving together as in suckling (Ayano et al., 2000; Bogaerts et al., 2012). A prerequisite for these skills to emerge is that the child has gained oral stability to control the jaw, the tongue and the lips. This develops alongside head and trunk stability and control (Ogg, 1975; Morris and Klein, 2000; Redstone and West, 2004).

The spectrum of head control ranges from basic head control when the infant is able to position the head in the body midline, to full head and trunk control that is present when the infant is able to sit by itself without any support. Intermediate measures of a developing, but not fully achieved, head and trunk control are, for example, sitting balance (e.g. sitting with some support) and the ability of the infant to bring the hands to the midline (Arvedson and Lefton-Greif, 1996).

The age at which infants attain different developmental milestones shows considerable variation within and between populations, presumably reflecting the infant's innate developmental trajectory combined with the opportunities and experiences provided by the caregiver (Lee and Galloway, 2012).

Also, feeding skills are acquired and consolidated over a period of time, so that the initial amount of food that is consumed by an infant when complementary feeding is started is small and increases over time with increasing feeding skills and repeated experiences. In an observational study in 39 healthy term infants who had spoon-fed pureed food introduced between 4 and 8 months of age, it took on average 5.7 (SD 2.1) weeks (range 2–10 weeks) for them to consolidate their feeding skills, regardless of the age at which CFs were first given, or whether the infant was breastfed or bottle-fed (van den Engel-Hoek et al., 2014).

3.3.1. Gross and fine motor skills relevant for spoon-feeding pureed CFs

The Panel considers the infant's ability of holding the head in midline when in supine position and to control its head well when pulled to sitting or at aided sitting to be the earliest gross motor skills indicative of an infant's developmental readiness to consume spoon-fed pureed CFs.

Table 5 gives an overview about the achievement of these and related milestones in the studies that were retrieved through the extensive literature search.

Table 5: Attainment of the gross motor developmental milestones indicative of an infant's developmental readiness to consume spoon-fed pureed CFs

Skill	Age	Result	N	Country	Study design	Author
Head in midline in supine position	3 m	60% could keep head in midline to a limited extent	8	SE	Cross-sectional	Hedberg et al. (2005)
Head in midline in supine position	3 and 4 m	Increased frequency of headline posture in midline	13	BR	Longitudinal[(a)]	Lima-Alvarez et al. (2014)
Head in midline in supine position	4 m	60% could keep the head adequately in the midline and another 20% could keep it in the midline to a limited extent	8	SE	Cross-sectional	Hedberg et al. (2005)
Head in midline in supine position	5 m	100% could keep the head adequately in the midline	8	SE	Cross-sectional	Hedberg et al. (2005)

Skill	Age	Result	N	Country	Study design	Author
Control of head movements in supine position	3 and 4 m	Increased proportion of midline-to-side and side-to-side movements	13	BR	Longitudinal	Lima-Alvarez et al. (2014)
Control of head movements in supine position	4 m	Increased peak velocity of head movements	13	BR	Longitudinal	Lima-Alvarez et al. (2014)
Head control when pulled to sitting	3–4 m	Mean age when milestone reached: 3.25 (SD 0.72) m	13,076	JP	Longitudinal	Yokoyama et al. (2011)
Head control when pulled to sitting	4 m	33% had a good head control	51	AU	Longitudinal	Pin et al. (2009)
Head balance at aided sitting	4 m	100% of infants had adequate head control at aided sitting	8	SE	Cross-sectional	Hedberg et al. (2005)

AU: Australia; BR: Brazil; JP: Japan; m: months; SE: Sweden.
(a): Studied at birth, 1, 2, 3 and 4 months of age.

Information on the development of fine motor skills relevant for spoon-feeding of pureed food was available from only one longitudinal study (Carruth and Skinner, 2002). In this study, infants were able to use the tongue to move food to the back of the mouth and swallow it at a mean age of 4.95 (SD 1.27) months with a range of 2.0–7.5 months. The Panel notes that the development of this skill is also indicative that the extrusion reflex had already disappeared at that age. In the same study, infants were able to keep food in their mouth without the need to be re-fed at a mean age of 5.72 (SD 1.58) months with a range of 0.5–10.5 months and to use the upper lip to remove food from the spoon at a mean age of 7.73 (SD 2.23) months with a range of 4.0–16.0 months.

The Panel notes that the earliest gross motor skills indicative of developmental readiness for spoon-feeding of pureed CFs (i.e. holding the head in midline when in supine position and to control its head well when pulled to sitting or at aided sitting) can be observed between 3 and 4 months of age. From the limited evidence that is available, fine motor skills indicative of developmental readiness for spoon-feeding of pureed foods and full disappearance of the extrusion reflex occur on average later.

3.3.2. Gross and fine motor skills relevant for self-feeding of finger foods

The Panel considers the infant's ability to sit alone is indicative of an infant having achieved the developmental readiness to consume self-fed finger foods.

Table 6 gives an overview about the achievement of this milestone in the studies that were retrieved through the extensive literature search.

Table 6: Attainment of the gross motor developmental milestone indicative of an infant's developmental readiness to consume self-fed finger foods

Skill	Result	N	Country	Study design	Author
Sitting in lap without support	Mean age: 5.54 m SD 2.08 m	98	US	Longitudinal	Carruth and Skinner (2002)
Sitting alone	Mean age: 5.4 m Range 3.8–9.2 m	816	GH, IN, NO, OM, US[(a)]	Longitudinal	WHO Multicentre Growth Reference Study Group (2006)
Sitting alone	Mean age: 5.6–6.0 m depending on the group of infants investigated	105	US	Longitudinal	Heinig et al. (1993)

Skill	Result	N	Country	Study design	Author
Sitting alone	Median age: 6 m Range: 4–9 m	189	CN	Retrospective	Wang et al. (2019)
Sitting alone	Median age: 6.3 m IQR: 6.0–7.2 m	542[(b)]	IT	Longitudinal	Agostoni et al. (2009)
Sitting alone	Mean age: 6.66 m SD 1.03 m	13,076	JP	Longitudinal	Yokoyama et al. (2011)
Sitting up from lying position	Mean age: 6.9 m SD 1.3 m	140[(c)]	HN	Longitudinal	Dewey et al. (2001)
Sitting up from lying position	Mean age: 7.8 m SD 1.6 m	108[(d)]	HN	Longitudinal	Dewey et al. (2001)

CN: China; GH: Ghana; HN: Honduras; IN: India; IQR: interquartile age; IT: Italy; JP: Japan; m: months; NO: Norway; OM: Oman, US: United States.
(a): Gross motor milestones were not assessed in the Brazilian study site.
(b): Infants in the control group of a randomised controlled trial.
(c): Appropriate and small-for-gestational age infants.
(d): Small-for-gestational age infants.

With respect to fine motor skills necessary for self-feeding of finger foods one study (Törölä et al., 2012) reported that in 11 term infants emerging chewing (i.e. lateral and diagonal movements) appeared at a median age of 5 (range 5–-8) months, while chewing (i.e. rotatory movements) occurred later: diagonal rotatory movements were observed at a median age of 7 (range 7–10) months and circulatory movements at a median age of 8 (range 7–11) months.

In the Gateshead Millennium Study, a population-based cohort study, 56% of infants (340 out of 602) were reported to be reaching for food before 6 months of age (Wright et al., 2011). During this time period, it is also expected that reaching movements become more organised and mature (von Hofsten, 1991). In addition, this may be interpreted as an apparent interest in food, even though the Panel acknowledges that this could be also an expression of interest in the environment.

The Panel notes that the gross motor skill indicative of developmental readiness for self-feeding finger foods (i.e. sitting without support) can be observed in some infants at 4 months, but more commonly between 5 and 7 months of age. From the limited evidence that is available, fine motor skills indicative of developmental readiness for self-feeding finger foods may occur at the same time or slightly later.

3.4. Developmental readiness of the term infant to receive CFs: conclusions

The Panel considers that the gastrointestinal and renal functions are not limiting factors with respect to the timing of introduction of CFs once the infant has the necessary neuromotor skills and has developed an apparent interest in such feeding.

The Panel further considers that there is a large biological variability when infants develop the necessary neuromotor skills for progressing from a liquid to a diet including pureed CFs and finger foods, depending on the individual infant. Furthermore, they are acquired and consolidated over a period of time with practice. From the neurodevelopmental data, it is not possible to define a precise age when introduction of CFs is appropriate.

The Panel concludes that the earliest gross motor skills indicative of developmental readiness for spoon-feeding of pureed CFs (i.e. holding the head in midline when in supine position and to control its head well when pulled to sitting or at aided sitting) can be observed between 3 and 4 months of age. At this age, it can be assumed that the rooting and the extrusion reflexes may have also diminished in some infants.

The Panel also concludes that the gross motor skill indicative of developmental readiness for self-feeding finger foods (i.e. sitting without support) can be observed in some infants at 4 months, but more commonly between 5 and 7 months of age.

4. Assessment of the data on body weight, body length/height and head circumference in individuals born at term or mixed populations

4.1. Body weight, body length/height and head circumference: final body of evidence

The 42 publications that were considered in the assessment in individuals born at term or mixed populations are given in Appendix B.1. These included two publications that were considered together (Kramer et al., 1985a,b) and one publication that covered four studies (Moschonis et al., 2017).

These publications reported on results from 42 studies:

- 5 RCTs (2 rated as Tier 1, 3 rated as Tier 2);
- 30 prospective cohort studies and 2 pooled analyses of prospective studies (3 rated as Tier 1, 15 rated as Tier 2 and 16 rated as Tier 3; 2 studies were allocated two different Tiers depending on the endpoint that was assessed);
- 6 retrospective studies (all Tier 3).

In these studies, 18 different endpoints related to body weight, body length and HC were investigated. Results of all the studies are given in Annex A as Microsoft Excel® file. In addition, for the main endpoints, results are summarised in the forest plots in Appendices A.1–A.10 of this Scientific Opinion.

With respect to the interpretation of the age at introduction of CFs as reported in the following, please refer to Section 2.2.3.3.

4.2. Body weight, body length/height and head circumference: endpoint and study selection

The first anthropometric measure to be impacted in the presence of nutritional imbalances is body weight, whilst body length/height and HC change at later stages. In the absence of evidence of an effect of the timing of introduction of CFs on body weight endpoints, body length/height and HC are unlikely to be altered. In addition, measurement errors for body length/height and HC are higher compared to body weight (Harrison et al., 2001). Therefore, the emphasis of the assessment is put on body weight endpoints.

The interpretation of the biological relevance of mean differences in anthropometric outcomes depends on the age at outcome assessment and the characteristics of the reference group to which the other groups are compared. For example, a 1 kg difference in weight at 12 months of age might be of relevance while it may be minor at, e.g. 10 years of age. This difference compared to a relatively underweight reference group will have a different meaning than if compared to a relatively overweight reference group. The advantage of the use of z-scores is that the age at outcome assessment (and gender) is already considered, which makes comparisons across measurement time points possible. However, the use of different reference populations between studies to transform absolute measurements into z-scores (e.g. WHO, US Centers for Disease Control and Prevention (CDC) or national growth standards) limits the comparability of results between these studies. Despite this, the Panel decided to give priority to results reported as z-scores.

Conditional body weight gain, expressed in z-scores, takes into account that, on average, children with a relatively higher or lower body weight at an initial time point will tend to have a body weight closer to the median at a subsequent time point (regression to the mean). It also allows more readily comparisons of outcomes of different studies. On the contrary, the interpretation of absolute body weight gain is hampered by the different time periods in which body weight gain is measured, and the use of different metrics (grams over the whole period, g/month). Absolute body weight gains expressed in grams per month also assume a linear growth rate of a child, which is, however, not observed biologically. This use of different metrics was also observed for absolute body length gain.

The biological relevance of differences in z-scores is judged compared to a difference of 0.5 z-scores (SCF, 2003), which is considered to be biologically relevant for anthropometric outcomes.[31] A difference of 3 g/day in weight gain over a 3- to 4-month period was suggested by the AAP (1988) to constitute a biologically relevant difference.

[31] Even though this value is an arbitrary value.

With respect to the minimum study duration, only studies that provided results beyond the age of 6 months were included in this assessment, considering that they provide more reliable estimates of associations that may persist beyond infancy.

Regarding possible reverse causality of observational studies, infants growing faster or are heavier may be introduced to CFs earlier. Thus, this may (partially) explain an association at later time points between age of introduction of CFs and anthropometric outcomes. This aspect has been addressed in the assessment of the RoB related to confounders (Section 2.2.2; key question 3), by considering whether a previous outcome measurement was taken into account as a covariate in the analysis.

4.3. Body weight: summary of the evidence

This section discusses firstly WAZ and attained body weight and the related subgroup analyses, secondly weight-for-length(height)-z-scores (WL(H)Z), thirdly endpoints for which the results are not shown in forest plots, i.e. conditional body weight gain, absolute body weight gain, rapid body weight gain (either because of the availability of only two studies with point estimates or because of lack of comparability of result metrics) and finally miscellaneous endpoints.

4.3.1. WAZ and attained body weight

Main line of evidence (12 studies)

For WAZ or attained body weight, the evidence derived from the five RCTs (Cohen et al., 1995a; Mehta et al., 1998; Dewey et al., 1999; Jonsdottir et al., 2014; Perkin et al., 2016) did not show an effect of the introduction of CFs at 3–4 months of age compared with introduction at 6 months on these endpoints assessed up to the age of 3 years (Appendices A.1 and A.3). This is true for the pooled estimate as well as for the results of the individual studies. Heterogeneity was not important (I^2 = 0% for both WAZ and attained body weight).

Seven prospective cohort studies were reported in eight papers (Forsyth et al., 1993; Wilson et al., 1998; Haschke and van't Hof, 2000; Grote et al., 2011; Imai et al., 2014; Noppornlertwong and Tantibhaedhyangkul, 2016; Eriksen et al., 2017). These investigated the timing of introduction of CFs at various ages, four of them investigated introduction at ages below 3 or 4 months vs later. They did not show a biologically relevant association of the age of introduction of CFs with WAZ and attained body weight assessed up to the age of 7 years (Appendices A.1 and A.3). This is true for the pooled estimate as well as for the results of the individual studies. Heterogeneity was substantial for WAZ (I^2 = 74%) and not important for attained body weight (I^2 = 0%).

Subgroup analyses were made on WAZ and attained body weight, in exclusively breastfed or formula fed infants (Section 2.2.3.2): there was no evidence for an association in any of these two groups, for these two endpoints (Appendices A.2 and A.4).

> The Panel notes, from the five RCTs and seven prospective cohort studies (Tiers 1 and 2) in the main line of evidence, that there is no evidence for an association between the timing of introduction of CFs and weight assessed up to the age of 7 years.

Supportive line of evidence (15 studies)

The reasoning behind the use of the data comprised in the supportive line of evidence is explained in Section 2.2.3.3.

- **Prospective cohort studies (9 studies, Tier 3)**

There was no evidence for an association from the meta-analysis of four studies (Heinig et al., 1993; Kalanda et al., 2006; Huh et al., 2011; Gaffney et al., 2012) that investigated the association between the introduction of CFs, mostly below 3 or 4 months of age vs thereafter (three out of four studies), on WAZ assessed up to 3 years of age. Heterogeneity was moderate to substantial (I^2 = 54%) (Appendix A.1).

The only studies for which the overall pooled estimate and associated 95% CI departed from the 'null' effect were the three prospective cohort studies that investigated attained body weight (Hodgson, 1978; Huh et al., 2011; Atkins et al., 2016). In the meta-analysis of five group comparisons from these three studies, there was an association between earlier introduction (< 1.5 to < 6 months) of CFs, compared with later introduction, and a higher attained body weight up to the age of 3 years

(mean difference 391 (95% CI 211 to 570) g) (Appendix A.3). All estimates were unadjusted and therefore it is likely that the association observed in the meta-analysis is overestimated. Heterogeneity was not important (I^2 = 0%). In addition, three studies (Warrington and Storey, 1988; WHO Working Group on Infant Growth, 1994; Morgan et al., 2004) did not report point estimates, but did not find statistically significant associations between the introduction of CFs and attained body weight or WAZ at 12, 18 and 24 months of age.

- **Retrospective studies (1 study, Tier 3)**

In one cross-sectional analysis of baseline data of a prospective cohort study (Sloan et al., 2008), a higher WAZ at 14 months of age with a borderline biological relevance was observed in infants introduced to CFs before 4 months of age compared with thereafter (adjusted mean difference 0.45 (95% CI 0.21 to 0.94) z-scores) (Annex A as Microsoft Excel® file).

- **Studies in which the timing of introduction of CFs was used as a continuous variable in the analysis, irrespective of the study design (5 studies)**

The prospective cohort studies described in Kramer et al. (1985a), Vail et al. (2015) and Butte et al. (2000) (Tiers 2 and 3) as well as a cross-sectional study (Zhu et al., 2015) and a retrospective cohort study (Klag et al., 2015) (both Tier 3) did not observe statistically significant associations between the timing of introduction of CFs and attained body weight at 2 years and WAZ at 1, 2 and 6 years (Annex A as Microsoft Excel® file).

The Panel notes that, in the supportive line of evidence, the meta-analysis of five group comparisons from three prospective cohort studies (Tier 3) as well as a retrospective study indicate a significant association between 'early' introduction of CFs and a higher body weight in childhood (in total four studies). However, the meta-analysis of the four prospective cohort studies (Tier 3) on WAZ and the results of the remaining individual studies (in total 10 studies) in the supportive line of evidence are consistent with the findings of the main line of evidence.

4.3.2. Other body weight-related endpoints

Main line of evidence

For WL(H)Z (4 studies) assessed up to 4 years of age, the RCT (Perkin et al., 2016) and the meta-analysis of the three prospective cohort studies (Grote et al., 2011; van Rossem et al., 2013; Eriksen et al., 2017) did not show an association with early introduction of CFs (in two studies: before 3 or 4 months of age). Heterogeneity was moderate to substantial (I^2=51%) (Appendix A.5).

For conditional body weight gain (2 studies) assessed up to 3 years of age, the two prospective cohort studies (Griffiths et al., 2009; de Beer et al., 2015) did not show a biologically relevant association between the age of introduction of CFs before 4 months compared with thereafter and the outcome (Annex A as Microsoft Excel® file).

For absolute body weight gain (5 studies) assessed in different time spans, results of the five prospective studies (including one RCT) (Cohen et al., 1995a; Simondon and Simondon, 1997; Imai et al., 2014; Mäkelä et al., 2014; Noppornlertwong and Tantibhaedhyangkul, 2016) are not directly comparable (as explained in Section 4.2). They are therefore not summarised graphically. There were no statistically significant findings comparing various time points of introduction of CFs (Annex A as Microsoft Excel® file).

For rapid body weight gain (2 studies), defined as change in z-score above 0.67, the prospective cohort study by Azad et al. (2018) showed higher odds of rapid body weight gain from 0 to 12 months, with introduction of CFs below 4 months compared with after 6 months (adjusted odds ratio (aOR) 1.43 (95% CI 1.01 to 2.01)) and also at 4–5 months compared with after 6 months (aOR 1.86 (95% CI 1.36 to 2.56)). On the contrary, Layte et al. (2014) showed no evidence for an association between rapid body weight gain from 0.75 to 3 years and introduction of CFs before 4 months of age compared with later (Annex A as Microsoft Excel® file).

For WAZ gain (1 study) between birth and 12 months, in the study by Azad et al. (2018), there was no biologically relevant association between the timing of introduction of CFs and WAZ gain (Annex A as Microsoft Excel® file).

The Panel notes that only one study (Tier 2) in the main line of evidence shows a relevant association between the timing of introduction of CFs before 6 months of age compared with thereafter and rapid weight gain in the first year of life, while the 12 other studies (Tiers 1 and 2) do not observe biologically relevant associations between the timing of introduction of CFs and the endpoints investigated.

Supportive line of evidence

The reasoning behind the use of the data comprised in the supportive line of evidence is explained in Section 2.2.3.3.

For WL(H)Z (3 studies, Tiers 2 and 3), the findings from two prospective studies (WHO Working Group on Infant Growth, 1994; Butte et al., 2000) and a cross-sectional study (Zhu et al., 2015) that analysed the timing of introduction of CFs as a continuous variable are consistent with the results of the main line of evidence (Annex A as Microsoft Excel® file).

For conditional body weight gain (2 studies, Tier 3), the findings from a prospective study (Wright et al., 2004) are consistent with the results of the main line of evidence, while in one cross-sectional analysis of baseline data of a prospective cohort study (Sloan et al., 2008), a higher conditional body weight gain between 2 and 14 months of age, with a borderline biological relevance, was observed in infants introduced to CFs before 4 months of age compared with thereafter (adjusted mean difference: 0.49 (95% CI 0.26 to 0.93) z-scores) (Annex A as Microsoft Excel® file).

For absolute body weight gain (6 studies, 5 Tier 3 and 1 Tier 2), three prospective studies (Heinig et al., 1993; Baker et al., 2004; Morgan et al., 2004) found earlier introduction of CFs (< 3 to < 6 months vs thereafter) to be associated both with higher and lower weight gain that was not considered by the Panel to be of biological relevance (ranging from mean differences of −200 g to 167 g for 15- and 12-month time periods, respectively). The prospective study that analysed the timing of introduction of CFs as a continuous variable (Haschke and van't Hof, 2000) (Tier 2), and a retrospective cohort study (Klag et al., 2015), are consistent with the main line of evidence, as well as one cross-sectional analysis of baseline data of a prospective cohort study (Kim and Peterson, 2008) in which the statistically significantly higher body weight gain between birth and 9 months (adjusted mean difference: 47 (95% CI 15 to 78) g) was not of biological relevance (Annex A as Microsoft Excel® file).

For rapid body weight gain (1 study, Tier 3), the results of a cross-sectional analysis of baseline data of an RCT (Mihrshahi et al., 2011) did not show an association between early introduction of CFs (< 3 vs ≥ 3 months) and rapid body weight gain from birth to 4–7 months of age (Annex A as Microsoft Excel® file).

For WAZ gain (1 study, Tier 3), Klag et al. (2015) in a retrospective cohort study did not observe a statistically significant association between the timing of introduction of CFs used as a continuous variable in the analysis and WAZ gain between birth and 12 months (Annex A as Microsoft Excel® file).

The Panel notes that the results in the supportive line of evidence are mostly consistent with those in the main line of evidence: 12 (10 Tier 3 and 2 Tier 2) out of 13 studies do not observe biologically relevant differences in the endpoints investigated, while one study (Tier 3) finds a borderline biologically relevant higher conditional weight gain between 2 and 14 months of age to be associated with introduction of CFs before 4 months of age compared with thereafter.

Endpoints investigated in single studies

Other endpoints investigated that were related to weight were: WL(H)Z-trajectories and WAZ-trajectories (Grote et al., 2011) (main line of evidence); proportion of children who had started CFs < 4 months of age in WAZ and WL(H)Z tertiles (Sit et al., 2001) (supportive line of evidence). These were assessed in single studies only. Therefore, they cannot be used to establish the appropriate age range of introduction of CFs (Section 2.2.3.3).

4.4. Body length/height: summary of the evidence

This section discusses first L(H)AZ and attained body length/height and the related subgroup analyses, then absolute body length gain for which no forest plot could be made (Section 2.2.3.2), and finally miscellaneous endpoints.

4.4.1. L(H)AZ and attained body length/height

Main line of evidence (11 studies)

For L(H)AZ or attained body length/height, the evidence derived from the five RCTs (Cohen et al., 1995a; Mehta et al., 1998; Dewey et al., 1999; Jonsdottir et al., 2014; Perkin et al., 2016) did not show an effect of the introduction of CFs at 3–4 months of age compared with the introduction at 6 months on these endpoints assessed up to around 3 years of age (Appendices A.6 and A.8). This is true for the pooled estimate as well as for the results of the individual studies. Heterogeneity was not important (I^2 = 0% for both L(H)AZ and attained body length).

The six prospective cohort studies, investigating various ages of introduction of CFs (in three studies: before 3 or 4 months of age), did not show biological relevant associations of earlier introduction of CFs with L(H)AZ or attained body length/height up to the age of 9 years ((Haschke and van't Hof, 2000; Grote et al., 2011; Imai et al., 2014; de Beer et al., 2015; Nopporlertwong and Tantibhaedhyangkul, 2016; Eriksen et al., 2017) as well as Moschonis et al. (2017) for the EDEN[32] study and the Avon Longitudinal Study of Parents and Children (ALSPAC)). This is true for each individual comparison and for the results of the meta-analyses. Heterogeneity was moderate to substantial for L(H)AZ (I^2 = 55%) and not important for attained body length/height (I^2 = 0%).

Subgroup analyses were made on L(H)AZ and attained body length/height, in exclusively breastfed or formula fed infants (Section 2.2.3.2): there was no evidence for an association in any of these two groups, for these two endpoints (Appendices A.7 and A.9).

> The Panel notes, from the five RCTs and the six prospective cohort studies (Tiers 1 and 2) in the main line of evidence, that there is no evidence for an association between the timing of introduction of CFs and body length/height assessed up to the age of 9 years.

Supportive line of evidence (10 studies)

The reasoning behind the use of the data comprised in the supportive line of evidence is explained in Section 2.2.3.3.

The meta-analyses of the four prospective cohort studies reported in three publications rated as Tier 3 on L(H)AZ (Heinig et al., 1993; Huh et al., 2011; Moschonis et al., 2017) and the meta-analysis on the two prospective cohort studies on attained body length/height (Huh et al., 2011; Atkins et al., 2016), which investigated various ages of introduction of CFs, were consistent with the findings in the main line of evidence (Appendices A.6 and A.8). Heterogeneity was not important (L(H)AZ I^2 = 0%, attained body length/height I^2 = 0%). Individually, the two studies on attained body length/height that were combined in the meta-analysis found that exclusively breastfed infants, but not formula fed infants, that were introduced to CFs < 4 months of age compared with thereafter (Huh et al., 2011) and infants introduced to CFs < 6 months of age compared with thereafter (Atkins et al., 2016) were taller at 3 years and 20 months of age, respectively. However, the estimates were unadjusted and therefore it is likely that the associations that were observed were overestimated.

In addition, Morgan et al. (2004) and WHO Working Group on Infant Growth (1994) that did not provide a point estimate reported that the timing of introduction of CFs were not associated with attained body length at 18 months and L(H)AZ at 9, 10, 11 and 12 months of age, respectively.

Also the prospective cohort studies by Vail et al. (2015) and Butte et al. (2000), and the cross-sectional study by Zhu et al. (2015) (all Tier 3) that analysed the timing of introduction of CFs as a continuous variable, found no evidence for an association (Annex A as Microsoft Excel® file).

> The Panel notes that the results of the ten studies (Tier 3) in the supportive line of evidence are consistent with those in the main line of evidence.

4.4.2. Absolute length gain

Main line of evidence (3 studies)

The results of the three individual prospective studies, including one RCT (Cohen et al., 1995a; Simondon and Simondon, 1997; Nopporlertwong and Tantibhaedhyangkul, 2016), are not directly

[32] Etude des déterminants pré et post natals précoces de la santé et de développement de l'enfant.

comparable (as explained in Section 4.2). Therefore, they are not summarised graphically in Appendix A. They did not find statistically significant associations between the timing of introduction of CFs (investigating various time points of introduction) and length gain assessed in different time spans up to 12 months of age (Annex A as Microsoft Excel® file).

> The Panel notes, from the RCT and two prospective cohort studies (all Tier 2) in the main line of evidence, that there is no evidence for an association between the timing of introduction of CFs and length gain assessed in different time spans up to the age of 12 months.

Supportive line of evidence (3 studies)

The prospective cohort study (Tier 3) by Heinig et al. (1993) was consistent with the findings in the main line of evidence (Annex A as Microsoft Excel® file).

Morgan et al. (2004) (Tier 3) reported a statistically significantly lower body length gain between 3 and 18 months of age associated with introduction of CFs < 3 months of age compared with thereafter (Annex A as Microsoft Excel® file). However, infants introduced to CFs < 3 months of age were longer at baseline than those introduced later, and the lower length gain led to a comparable body length in both groups of infants at 18 months of age. Therefore, the Panel considers this finding not to be of biological relevance.

Haschke and van't Hof (2000) (Tier 2) that analysed the timing of introduction of CFs as a continuous variable reported a statistically significantly lower body length gain between 1 and 24 months to be associated with earlier introduction of CFs (adjusted mean difference: −0.05 (95% CI −0.09 to −0.01) mm/month per month of earlier introduction of CFs) (Annex A as Microsoft Excel® file). Between 1 and 12 months, differences in body length gain were not statistically significant and the point estimate was on the other side of the 'null' line, which may indicate that the lower body length gain primarily occurred between 12 and 24 months of age and thus may not be a result that could be directly attributed to the timing of introduction of CFs.

> The Panel notes that the studies in the supportive line of evidence (2 Tier 3 and 1 Tier 2) show divergent results. However, these are considered by the Panel as being either of no biological relevance or unlikely to be a direct result of the timing of introduction of CFs. Therefore, the Panel considers the results of the studies in the supportive line of evidence to be consistent with the findings from the main line of evidence.

Endpoints investigated in single studies

Other investigated endpoints related to length were: conditional body length gain (de Beer et al., 2015) and L(H)AZ-trajectories (Grote et al., 2011) (main line of evidence) (Annex A as Microsoft Excel® file). These were assessed in single studies only. Therefore, they cannot be used to establish the appropriate age range of introduction of CF.

4.5. Head circumference: summary of the evidence

This section discusses the endpoints related to HC, mostly investigated in RCTs, i.e. attained HC, HC-for-age z-scores (HCZ), and finally, miscellaneous endpoints.

Main line of evidence (4 studies)

For HCZ, individual RCTs (Jonsdottir et al., 2014; Perkin et al., 2016) showed no statistically significant effect of the timing of introduction of CFs at 3–4 months of age compared with the introduction at ≥6 months on this endpoint assessed up to 3 years of age (Annex A as Microsoft Excel® file).

For attained HC, the evidence derived from the three RCTs (Mehta et al., 1998; Jonsdottir et al., 2014; Perkin et al., 2016) did not show an effect of the timing of introduction of CFs at 3–4 months of age compared with the introduction at 6 months on this endpoint assessed up to 3 years of age. This is true for the results of the meta-analysis and the individual studies (Appendix A.10). Heterogeneity was not important (I^2 = 0%). The results of the only prospective cohort study (Noppornlertwong and Tantibhaedhyangkul, 2016) that investigated introduction of CFs at 4–6 months of age vs > 6 months and attained HC at the age of 12 months, are consistent with the above.

For HC gain, one study (Noppornlertwong and Tantibhaedhyangkul, 2016) did not show statistically significant differences in HC gain between 4 and 12 months of age (Annex A as Microsoft Excel® file).

The Panel notes, from the three RCTs and the one prospective cohort study (Tiers 1 and 2) in the main line of evidence, that there is no evidence for an association between the timing of introduction of CFs and HC assessed up to 3 years of age.

Supportive line of evidence (1 study)

For HC gain, the prospective study by Morgan et al. (2004) (Tier 3) found a statistically significant lower HC gain between 3 and 18 months of age associated with introduction of CFs < 3 months of age compared with thereafter (Annex A as Microsoft Excel® file). However, infants introduced to CFs < 3 months of age had a higher HC at baseline than those introduced later, and the lower HC gain led to a comparable HC in both groups of infants at 18 months of age. Therefore, the Panel considers this finding not to be of biological relevance.

The Panel notes that the association observed in the study (Tier 3) in the supportive line of evidence is not of biological relevance. The Panel also notes that the result of the study in this line of evidence is consistent with the findings in the main line of evidence.

4.6. Body weight, body length/height and head circumference: conclusions and grading of the confidence in the evidence

Imprecision: The results of the meta-analyses did not indicate imprecision.

Inconsistency: The evidence is consistent across populations and the results of the studies in the supportive line of evidence (20 studies for body weight, 10 for body length/height and 1 for HC) were consistent with the main line. For all meta-analyses conducted in the main line of evidence, I^2 was below 75%.

Generalisability: For all outcomes, RCTs in exclusively breastfed and exclusively formula fed infants were available. Subgroup analyses (independent of study design) did not show different effects on WAZ, attained body weight, L(H)AZ or attained body length/height, of the timing of introduction of CFs in exclusively breastfed and in exclusively formula fed infants (Appendices A.2, A.4, A.7 and A.9). Therefore, the Panel considers that the evidence from RCTs can be generalised to the whole population of infants living in Europe. As a representative number of populations were studied in the prospective cohort studies, the Panel considers that their results can also be generalised.

Publication bias: From visual inspections of the funnel plots on WAZ, attained body weight and L(H)AZ, there was no convincing evidence for publication bias (Annexes D.1, D.2 and D.3). For the other endpoints, publication bias could not be evaluated, because of the insufficient number of studies.

The Panel concludes from the RCTs, that there is no effect of introduction of CFs at 3–4 months vs 6 months of age on body weight (5 RCTs), body length/height (5 RCTs) and HC (3 RCTs) assessed up to around 3 years of age (high level of confidence in the evidence).

The Panel concludes from prospective cohort studies (Tiers 1 and 2), covering a broader range of ages of introduction of CFs than the RCTs, that there is no evidence for an association between the age of introduction of CFs and body weight (12 studies) and body length/height (9 studies) (moderate level of confidence in the evidence). For the assessment of body weight, the ages of introduction of CFs ranged between < 2 months and < 6 months for early introduction and > 2 months and ≥ 6 months for later introduction. For the assessment of body length/height, early introduction ranged from 2–3 months to < 6 months, and later introduction from > 4 to ≥ 6 months. The latest age of outcome assessment was 7 years for body weight and 9 years for body length/height.

The prospective cohort study available on HC was integrated with the RCTs as the ages that were compared were already covered by the RCTs.

5. Assessment of the data on BMI and related endpoints in individuals born at term or mixed populations

5.1. BMI: final body of evidence

The 40 publications that were considered in the assessment of data on BMI in individuals born at term or mixed populations are given in Appendix B.2 (2 publications were considered together (Kramer et al., 1985a,b)).

These publications reported on results of 36 studies:

- 2 RCTs (Tier 1);
- 26 prospective cohort studies (5 rated as Tier 1, 12 rated as Tier 2 and 12 rated as Tier 3; three studies were attributed two different Tiers);
- 8 retrospective studies (Tier 3).

In line with the reasons given in Section 5.2 for the selection of papers that reported on different ages at assessment of an endpoint in the same study, the results provided by de Beer et al. (2015) were used for the Amsterdam Born Children and their Development (ABCD) study (instead of those provided by Sirkka et al. (2018)) and the results provided by Vogelezang et al. (2018) for the Generation R study (instead of those provided by Durmuş et al. (2014)).

In the included studies, nine different endpoints related to BMI were investigated. Results of all the studies are given in Annex A as Microsoft Excel® file, including the ones by Sirkka et al. (2018) and Durmuş et al. (2014). In addition, results are summarised in the forest plots in Appendices A.11–A.13 of this Scientific Opinion.

With respect to the interpretation of the age at introduction of CFs as reported in the following, please refer to Section 2.2.3.3.

5.2. BMI: endpoint and study selection

Previous considerations (Section 4.2) on advantages and limitations of using z-scores compared to absolute (attained) measurements in the context of endpoints related to body weight or body length, as well as previous considerations on biological relevance of differences in z-scores, are also true in the context of studies on BMI, discussed in the following sections.

BMI-related outcomes (i.e. a continuous outcome) are discussed separately from the dichotomised outcome of overweight or obesity. Even though the definition of overweight and obesity is based on BMI, the dichotomisation may lead to different findings compared with results obtained from an analysis of the outcome on a continuous scale (i.e. BMI). Therefore, results are not necessarily comparable.

Regarding possible reverse causality of observational studies, the same considerations and approach for the assessment of the RoB described above for weight endpoints (Section 4.2) were also relevant for BMI and related endpoints.

5.3. BMI: summary of the evidence

This section discusses first BMIZ and attained BMI for which forest plots could be made, then the subgroup analysis for BMIZ, and finally miscellaneous endpoints.

Main line of evidence (13 studies)

For BMIZ or attained BMI, the evidence derived from the two RCTs in exclusively breastfed infants (Jonsdottir et al., 2014; Perkin et al., 2016) did not show an effect of the timing of introduction of CFs at 3–4 months of age compared with the introduction at 6 months on these endpoints assessed up to 3 years of age, neither from the meta-analysis nor individually. Heterogeneity was not important (I^2 = 0% both for BMIZ and attained BMI) (Appendices A.11 and A.13).

For BMIZ, the nine prospective cohort studies (Burdette et al., 2006; Grote et al., 2011; Huh et al., 2011; de Beer et al., 2015; Fairley et al., 2015; Leary et al., 2015; Zheng et al., 2015; Azad et al., 2018; Vogelezang et al., 2018), the majority of which investigated the timing of introduction of CFs below the age of 3 or 4 months vs later, did not show a biologically relevant association between the age of introduction of CFs and BMIZ assessed up to 15 years of age. This is true for the result of the meta-analysis and for each individual comparison. Heterogeneity was not important to moderate (I^2 = 35%).

For attained BMI at the age of 2 years, neither the results of the two individual prospective studies nor the result of the meta-analysis (Grote et al., 2011; Wen et al., 2014) was statistically significant, comparing introduction < 3 and < 2 months with later, respectively. This was also true for the individual studies. Heterogeneity was not important ($I^2 = 0\%$). In addition, Agras et al. (1990) did not report a statistically significant association between the introduction of CF ≤ 5 months of age compared with thereafter on attained BMI at 6 years of age (results only presented as correlation coefficients, thus not included in the meta-analysis).

A subgroup analysis was performed for BMIZ, in exclusively breastfed or formula fed infants (Section 2.2.3.2): there was no evidence for an association in either of these two groups (Appendix A.12).

The Panel notes, from the two RCTs and 11 prospective cohort studies (Tiers 1 and 2) in the main line of evidence, that there is no evidence for an association between the timing of introduction of CFs and BMI assessed up to 10 years of age.

Supportive line of evidence (17 studies)

The reasoning behind the use of the data comprised in the supportive line of evidence is explained in Section 2.2.3.3.

- **Prospective cohort studies (7 studies, Tier 3)**

The meta-analysis of five studies (Wilson et al., 1998; Haschke and van't Hof, 2000; Iguacel et al., 2018; Schmidt Morgen et al., 2018) on BMIZ assessed up to 11 years of age did not show a statistically significant association with the timing of introduction of CFs before 3.5 or 4 months of age compared with thereafter (mean difference –0.06 (95% CI −0.37 to 0.25) z-scores) (Appendix A.11). However, heterogeneity was important ($I^2 = 95\%$). When the study by Haschke and van't Hof (2000), that showed results considerably different from the other studies in that Tier (the reasons for which cannot be explained), was removed in a sensitivity analysis, heterogeneity became non-important to moderate ($I^2 = 33\%$), the pooled point estimate shifted to the other side of the line of the 'null' effect and the 95% CI was reduced (i.e. mean difference 0.05 (95% CI −0.03 to 0.13) z-scores).

Equally, the meta-analysis of the four comparisons from the three studies on attained BMI assessed up to around 10 years of age (Veena et al., 2010; Huh et al., 2011; Imai et al., 2014) did not show a statistically significant association between the timing of introduction of CFs (in two studies < 4 months compared with later) and this endpoint (Appendix A.13). Heterogeneity was substantial ($I^2 = 63\%$).

- **Retrospective studies (3 studies, Tier 3)**

One cross-sectional study (Brambilla et al., 2016) and one prospective cohort study (in which the timing of introduction of CFs was assessed after the outcome) (Lin et al., 2013) did not find an association between the timing of introduction of CFs at various ages and BMIZ assessed up to 14 years of age.

Vafa et al. (2012) found a statistically significantly higher attained BMI in 7-year-old children introduced to CFs ≤ 4 months of age compared with thereafter (adjusted mean difference 0.88 (95% CI 0.26 to 1.50) kg/m^2).

- **Studies in which the timing of introduction of CFs was used as a continuous variable in the analysis, irrespective of the study design (7 studies)**

Three such studies on BMIZ (two prospective cohort studies (Schack-Nielsen et al., 2010; Vail et al., 2015) and one cross-sectional study (Zhu et al., 2015), all Tier 3) did not find biologically relevant associations between the timing of introduction of CFs and the outcome assessed up to 42 years of age (Annex A as Microsoft Excel® file).

The four studies that analysed the timing of introduction of CFs as a continuous variable, did not find a statistically significant association between the timing of introduction of CFs and attained BMI assessed up to 4 years of age (one prospective cohort study rated as Tier 1 (Lande et al., 2005), two prospective cohort studies rated as Tier 2 (Kramer et al., 1985a; Robinson et al., 2009) and one cross-sectional analysis of baseline data of a prospective cohort study (Tier 3) (Zive et al., 1992)) (Annex A as Microsoft Excel® file).

The Panel notes that, in the supportive line of evidence, the results of the meta-analyses of the prospective cohort studies (7 studies, Tier 3) as well as 9 of the 10 remaining individual studies (3 rated as Tiers 1 and 2; 6 Tier in 3) are consistent with the findings in the main line of evidence. Only one cross-sectional study (Tier 3) observed a higher attained BMI at 7 years to be associated with the introduction of CFs ≤ 4 months compared with later. Overall, the Panel considers that the results in the supportive line of evidence is consistent with those in the main line of evidence.

Endpoints investigated in single studies

Other investigated endpoints related to BMI were: BMIZ trajectories (Grote et al., 2011), BMI trajectory class membership (Garden et al., 2012), and 'high' BMI (Caleyachetty et al., 2013) (main line of evidence); % expected weight (Poskitt and Cole, 1978), waist circumference (Schack-Nielsen et al., 2010) and the Shukla index (Thorogood et al., 1979) (supportive line of evidence). These were assessed in single studies only. Therefore, they cannot be used to establish the appropriate age range of introduction of CFs (Section 2.2.3.3).

5.4. BMI: conclusions and grading of the confidence in the evidence

Imprecision: The results of the meta-analyses did not indicate imprecision.

Inconsistency: The evidence is consistent across populations and, overall, the supportive line of evidence (16 out of 17 studies) is consistent with the main line of evidence. For all meta-analyses conducted in the main line of evidence, I^2 was below 75%.

Generalisability: The study population of both RCTs consisted only of exclusively breastfed infants. Individually, these studies cannot be generalised to formula fed or mixed fed infants. However, considering that subgroup analyses on exclusively breastfed and exclusively formula fed infants did not show different effects of the timing of introduction of CFs on BMI in those infants (Appendix A.12) and that observational studies with a variety of different background milk feedings were consistent with the findings of the RCTs, the Panel considers that the results of these two RCTs can be generalised to the whole population of infants living in Europe. In the prospective cohort studies, a representative number of populations were studied. Therefore, the Panel considers that results from these studies can also be generalised.

Publication bias: Even though the funnel plot for BMIZ appeared to be asymmetrical (Annex D.4), none of the statistical methods (Egger's test, trim-and-fill analysis and contour plots) applied suggested asymmetry (data not shown). Therefore, the Panel considers that publications bias is unlikely. For attained BMI, publication bias could not be evaluated, because of the insufficient number of studies available.

The Panel concludes from the two RCTs (Tier 1) that there is no effect of introduction of CFs at 3–4 months of age compared with 6 months of age on BMI assessed up to 3 years of age (high level of confidence in the evidence).

The Panel concludes from 11 prospective cohort studies (Tiers 1 and 2) that there is no evidence for an association between the age of introduction of CFs and BMI (moderate level of confidence in the evidence). The ages of introduction ranged between ≤ 2 months and ≤ 5 months for early introduction and > 2 months and ≥ 6 months for later introduction. The latest age of outcome assessment was 10 years.

6. Assessment of the data on obesity and overweight in individuals born at term or mixed populations

6.1. Obesity and overweight: final body of evidence

The 55 publications that were considered in the assessment in individuals born at term or mixed populations are given in Appendix B.3. One publication covered four studies (Moschonis et al., 2017).

These papers reported on the results of 50 studies:

- 1 RCT (Tier 1);
- 29 prospective cohort studies (2 rated as Tier 1, 12 rated as Tier 2 and 15 rated as Tier 3);
- 20 retrospective studies (all Tier 3).

In line with the reasons given in Section 2.2.3.2 for the selection of papers that reported on different ages at assessment of an endpoint in the same study, the results provided by Massion et al. (2016) were used for the Millennium Cohort Study (MCS) (instead of those provided by Hawkins et al. (2009)) and the results provided by Moss and Yeaton (2014) for the Early Childhood Longitudinal Study, Birth Cohort (ECLS-B) (instead of those provided by Gibbs and Forste (2014), Flores and Lin (2013b) and Gooze et al. (2011)).

In the included studies, nine different endpoints related to obesity and overweight were investigated. Results of all the studies are given in Annex A as Microsoft Excel® file, including the ones by Hawkins et al. (2009), Gibbs and Forste (2014), Flores and Lin (2013b) and Gooze et al. (2011). In addition, results are summarised in the forest plots in Appendices A.14–A.17 of this Scientific Opinion.

With respect to the interpretation of the age at introduction of CFs as reported in the following, please refer to Section 2.2.3.3.

6.2. Obesity and overweight: endpoint and study selection

Different reference populations were used in the included studies, e.g. from the WHO, the CDC, the International Obesity Task Force (IOTF) or national growth standards, as well as different cut-offs (percentiles, z-scores) to define overweight and obesity. The Panel considers that this limits the comparability of results between studies. In the following sections, the disease outcome, i.e. odds/risk of developing obesity, is discussed first.

The studies that were considered by the Panel for the section on overweight included studies that investigated the odds of developing at least overweight (i.e. combining overweight and obese children) and studies that assessed the odds of developing overweight separately from the odds of developing obesity (i.e. separated overweight from obese children).

Regarding possible reverse causality of observational studies, the same considerations and approach for the assessment of the RoB described above for weight endpoints were also relevant for obesity and overweight outcomes (Section 4.2).

6.3. Obesity: summary of the evidence

Main line of evidence (6 studies)

The main line of evidence consists of six prospective cohort studies and no RCT (Reilly et al., 2005; Brophy et al., 2009; Neutzling et al., 2009; Huh et al., 2011; Layte et al., 2014; Zheng et al., 2015).

For these studies, which mainly investigated introduction of CFs at below 3 or 4 months of age, the result of the meta-analysis did not show a statistically significant association between the age of introduction of CFs and the odds of developing obesity up to 11 years of age. Heterogeneity was moderate (I^2 = 50%) (Appendix A.14).

> The Panel notes, from the six prospective cohort studies (Tiers 1 and 2) in the main line of evidence, that there is no evidence for an association between the timing of introduction of CFs and the odds of developing obesity up to 11 years of age.

Supportive line of evidence (10 studies)

The reasoning behind the use of the data comprised in the supportive line of evidence is explained in Section 2.2.3.3.

- **Prospective cohort studies (2 studies, Tier 3)**

The meta-analysis from the three comparisons of two studies (Moss and Yeaton, 2014; Barrera et al., 2016) did not show an association between the timing of introduction of CFs < 4 months of age compared with thereafter and the odds of developing obesity up to 6 years of age (Appendix A.14). Heterogeneity was not important (I^2 = 0%). Also, individually the results of these studies did not show an association between the timing of introduction of CFs and obesity.

- **Retrospective studies (4 studies, Tier 3)**

The meta-analysis of one case–control study (Zhou et al., 2011) and three cross-sectional studies (Birbilis et al., 2013; Vehapoglu et al., 2014; Sandoval Jurado et al., 2016) did not show an association between the timing of introduction of CFs (in three studies < 4 months of age compared with

thereafter or with > 6 months of age), and the odds of developing obesity up to 14 years of age (Appendix A.15). However, heterogeneity was important (I^2 = 81%). When the study by Zhou et al. (2011), that showed results that were considerably different from the other studies (which cannot be explained), was removed in a sensitivity analysis, heterogeneity became moderate (I^2 = 46%); the results remained non-statistically significant.

- **Studies in which the timing of introduction of CFs was used as a continuous variable in the analysis, irrespective of the study design (4 studies)**

The prospective cohort studies Mäkelä et al. (2014) (Tier 1) and Schack-Nielsen et al. (2010) (Tier 3) as well as the cross-sectional studies by Gillman et al. (2001) and Sinigaglia et al. (2016) (both Tier 3) did not observe statistically significant associations between the timing of introduction of CFs and the odds of developing obesity assessed up to 42 years of age.

- **Difference in the timing of introduction of CFs between cases and controls (2 studies)**

Two studies (both Tier 3) investigated the timing of introduction of CFs between obese and control subjects. One prospective cohort study (Flores and Lin, 2013a) found no statistically significant differences in the timing of introduction of CFs between 4-year-old children with severe obesity and their non-severely obese counterparts, while one case–control study (Zhou et al., 2011) found that 3-to 6-year-old obese children had significantly higher odds of having been introduced to CFs < 4 months of age than normal-weight controls (OR 6.58 (95% CI 2.71 to 15.93)). This analysis was unadjusted and therefore is likely to overestimate the association.

The Panel notes that, in the supportive line of evidence, results of the two meta-analyses of prospective cohort and retrospective studies are consistent with the findings in the main line of evidence (in total 6 studies). This is also true for the results of the four studies in which the timing of introduction of CFs was used as a continuous variable in the analysis. The results of the two studies investigating the difference in the introduction of CFs between cases and controls are inconsistent.

Endpoints investigated in single studies

One endpoint related to obesity was investigated in a single study in the supportive line of evidence only (i.e. %obese (Wolman, 1984) (Annex A as Microsoft Excel® file). Therefore, it cannot be used to establish the appropriate age range of introduction of CFs (Section 2.2.3.3).

6.4. Overweight: summary of the evidence

Main line of evidence (10 studies)

The RCT (Jonsdottir et al., 2014) that was available was relatively small in sample size, reflected in the wide 95% CI associated with the point estimate. No statistically significant effect of the timing of introduction of CFs (4 vs 6 months) on the odds of developing overweight up to 3 years of age was observed. However, this study was most likely underpowered for the outcome and its non-statistically significant findings were therefore not further used by the Panel for drawing conclusions (Appendix A.16).

For the 10 prospective cohort studies ((Neutzling et al., 2009; Rossiter and Evers, 2013; Durmuş et al., 2014; Wen et al., 2014; Fairley et al., 2015; Zheng et al., 2015; Massion et al., 2016; Azad et al., 2018) as well as Moschonis et al. (2017) for the studies EDEN and ALSPAC), the results of the meta-analysis did not show a statistically significant association between the timing of introduction of CFs (mostly < 4 vs ≥ 4 months) and the odds of developing overweight up to 13 years of age. Heterogeneity was substantial (I^2 = 66%).

The Panel notes, from the ten prospective cohort studies (Tiers 1 and 2) in the main line of evidence, that there is no evidence for an association between the timing of introduction of CFs and the odds of developing overweight up to 13 years of age.

Supportive line of evidence (30 studies)

The reasoning behind the use of the data comprised in the supportive line of evidence is explained in Section 2.2.3.3.

- **Prospective cohort studies (10 studies, Tier 3)**

The meta-analysis of 10 studies reported in 9 publications ((Abraham et al., 2012; Moss and Yeaton, 2014; Hollis et al., 2016; Aris et al., 2018; Bell S et al., 2018; Pluymen et al., 2018; Schmidt Morgen et al., 2018; Sirkka et al., 2018) as well as Moschonis et al. (2017) for the studies Greek EuroPrevall and Generation XXI) showed increased odds of developing overweight up to 17 years of age associated with earlier introduction of CFs (mostly < 4 months vs later) (OR 1.28 (95% CI 1.18 to 1.39)) (Appendix A.16). The 95% prediction interval crossed the null line (0.98–1.67). Heterogeneity was moderate to substantial (I^2 = 59%).

- **Retrospective studies (10 studies, Tier 3)**

From the meta-analysis of the 10 retrospective studies (Nascimento Simon et al., 2009; Jimenez-Cruz et al., 2010; Magalhaes et al., 2012; Birbilis et al., 2013; Lin et al., 2013; Rathnayake et al., 2013; Cu et al., 2015; Škledar and Milošević, 2015; Sun et al., 2016; Papoutsou et al., 2018), there was no evidence for an association between the timing of introduction of CFs (at various ages) and the odds of developing overweight up to 14 years of age (Appendix A.17). Heterogeneity was moderate to substantial (I^2 = 51%).

- **Studies in which the timing of introduction of CFs was used as a continuous variable in the analysis, irrespective of the study design (6 studies)**

The prospective cohort studies by Mäkelä et al. (2014) (Tier 1) and Schack-Nielsen et al. (2010) (Tier 3) as well as the cross-sectional studies by Butte (2009) and Gillman et al. (2001) (both Tier 3) did not find an association between the timing of introduction of CFs and the odds of developing overweight up to 42 years of age.

However, in the prospective cohort study by Seach et al. (2010) (Tier 3), a one month earlier introduction of CFs was associated with higher odds of developing overweight up to 10 years of age (aOR 1.11 (95% CI 1.03 to 1.19)). In the cross-sectional study by Hediger et al. (2001) (Tier 3), a one month earlier introduction of CFs was associated with a higher odds of developing overweight up to 3–5 years of age (aOR 1.0006 (95% CI 1.0003 to 1.001)). The Panel considers that the observed OR is unlikely to be of biological relevance.

- **Difference in the timing of introduction of CFs between cases and controls (4 studies)**

Four studies (all Tier 3) investigated the difference in the timing of introduction of CFs in overweight and normal weight subjects (Jiang et al., 2009; Gomes et al., 2010; Gungor et al., 2010; Flores and Lin, 2013b). The case–control study by Jiang et al. (2009) found that 1- to 3-year-old overweight children had statistically significantly higher odds of having been introduced of CFs < 4 months of age (aOR 1.76 (95% CI 1.15 to 3.64)) than controls. The retrospective cohort study by Gungor et al. (2010) found that 6- to 8-year-old overweight children had been introduced to CFs statistically significantly earlier than their controls (mean difference: −1.39 (95% CI −2.46 to −0.32) months). This analysis was unadjusted and therefore is likely to overestimate the association. The prospective cohort study (Flores and Lin, 2013b) and the other case–control study (Gomes et al., 2010) found no statistically significant differences in the timing of introduction of CFs between overweight cases and controls at 4 years and 2.2–6.8 years of age, respectively.

The Panel notes that the results within the supportive line of evidence (30 studies) are inconsistent.

Endpoints investigated in single studies

Another endpoint related to overweight was investigated in a single study only, i.e. %overweight (Burdette et al., 2006) (main line of evidence). Therefore, it cannot be used to establish the appropriate age range of introduction of CFs (Section 2.2.3.3).

6.5. Obesity and overweight: conclusions and grading of the confidence in the evidence

The RCT of the main line of evidence was most likely underpowered for this outcome and was therefore not considered further in the grading of the confidence in the evidence.

Imprecision: The results of the meta-analyses did not indicate imprecision.

Inconsistency: The evidence is consistent across populations and endpoints in the main line of evidence (12 studies in total (4 studies reported both on obesity and overweight)). For obesity, the results in the supportive line of evidence (10 studies) are consistent with the findings in the main line of evidence, while, for overweight, the results within the supportive line of evidence (30 studies) are inconsistent. As there was enough evidence in the main line of evidence and results were consistent across lines of evidence for the disease endpoint (i.e. obesity), the Panel did not downgrade the confidence in the evidence for the inconsistency observed in the supportive line of evidence for overweight. For all meta-analyses conducted in the main line of evidence, I^2 was below 75%.

Generalisability: A variety of populations and settings were covered in the available studies. Therefore, the Panel does not have concerns with respect to the generalisability of the findings to the whole population of infants living in Europe.

Publication bias: From visual inspections of the funnel plot on obesity, there was no convincing evidence for publication bias (Annex D.5). The funnel plot on overweight (Annex D.6) was asymmetrical as indicated by the Egger test and the trim-and-fill analysis, but the contour plot did not suggest that the asymmetry was due to publication bias (data not shown).

> The Panel concludes from the prospective cohort studies (Tiers 1 and 2) that there is no evidence for an association between the age of introduction of CFs and obesity (6 studies) or overweight (10 studies) (moderate level of confidence). The ages of introduction of CFs ranged between < 1 month and < 4 months for early introduction and > 2 months and ≥ 6 months for later introduction. The latest age of outcome assessment was 13 years.

7. Assessment of the data on body composition in individuals born at term or mixed populations

7.1. Body composition: final body of evidence

The 21 publications that were considered in the assessment in individuals born at term or mixed populations are given in Appendix B.4. These included two publications that were considered together (Kramer et al., 1985a,b) and one publication that reported on four prospective cohort studies (Moschonis et al., 2017).

These publications reported on results from 19 studies:

- 2 RCTs (1 rated as Tier 1, 1 rated as Tier 2);
- 13 prospective cohort studies and 1 pooled analysis of prospective studies (1 rated as Tier 1, 8 rated as Tier 2 and 6 rated as Tier 3; one study was assigned to two different Tiers, depending on how the outcome was measured);
- 3 retrospective studies (all Tier 3);

Results of all the studies are given in Annex A as Microsoft Excel® file. In addition, results are summarised in the forest plots in Appendices A.18 and A.20 of this Scientific Opinion.

In these studies, 25 different endpoints related to body composition were investigated.

With respect to the interpretation of the age at introduction of CFs as reported in the following, please refer to Section 2.2.3.3.

7.2. Body composition: endpoint and study selection

Body composition measurements performed by either dual-energy X-ray absorptiometry (DXA) or bioelectrical impedance analysis (BIA) (fat mass, fat-free mass, lean body mass and regional fat distribution) were preferred by the Panel over skinfold thickness (SFT) measurements expressed in millimetres, thus are described first in a separate subsection. The reliability of the outcome measurements was considered in the assessment of the RoB, i.e. DXA lower RoB than BIA.

BMC measurements not adjusted for bone area were not considered as an outcome for this assessment, owing to the lack of comparability in growing children that are of different size.

7.3. Fat mass: summary of the evidence

This section discusses first fat mass, then fat mass z-score and percentage of fat mass for which no forest plot could be made (Section 2.2.3.2), and finally miscellaneous endpoints.

Main line of evidence

For fat mass (2 studies), neither the RCT (Mehta et al., 1998) nor the prospective cohort study (Burdette et al., 2006) showed an association between early introduction of CFs (3–4 vs 6 months and < 4 vs ≥ 4 months of age) and this endpoint (Appendix A.18).

For fat mass z-scores (3 studies), the meta-analysis of three prospective cohort studies (Durmuş et al., 2014; de Beer et al., 2015; Leary et al., 2015) did not show statistically significant associations with the age of introduction of CFs ranging from ≤ 2 to < 5 months of age versus later. Heterogeneity was not important (I^2 = 0%) (Appendix A.19). The latest age at outcome assessment was 15 years.

For percentage of fat mass (2 studies), the RCT mentioned above (Mehta et al., 1998) also did not find a significant effect of the timing of introduction of CFs at 3–4 months, compared with 6 months of age, on this endpoint at 12 months of age. For the ALSPAC study when the outcome was assessed by DXA, Moschonis et al. (2017) did not report statistically significant associations between the introduction of CFs < 4 months of age and at 4–5 months of age, each compared with 5–6 months of age, and the percentage of fat mass at 13 years.

For high fat mass (1 study), the prospective cohort study by Burdette et al. (2006) did not find an association between the introduction of CFs < 4 months compared with thereafter and high fat mass (defined as the age and sex-specific 75th percentile of the cohort at 5 years of age).

> The Panel notes that the six studies in the main line of evidence (Tiers 1 and 2; some investigating several endpoints) showed consistently no association between the timing of introduction of CFs and fat mass assessed up to the age of 15 years.

Supportive line of evidence

The reasoning behind the use of the data comprised in the supportive line of evidence is explained in Section 2.2.3.3.

For fat mass (5 studies, reported in 2 papers), Moschonis et al. (2017) (Tier 3, covering four prospective cohort studies, including ALSPAC[33]), comparing < 4 and 4–5 vs 5–6 months of age (outcome assessed by BIA) (Appendix A.18), as well as one prospective cohort study that analysed the timing of introduction of CFs as a continuous variable (Robinson et al., 2009) (Tier 2), did not show an association between the timing of introduction of CFs and this endpoint.

For fat mass z-scores, no studies were available in the supportive line of evidence.

For percentage of fat mass (1 study), the prospective cohort study by Wilson et al. (1998) (Tier 3) found a higher percentage of fat mass at 7 years of age to be associated with introduction of CF < 3.5 months vs thereafter (adjusted mean difference of 2% points (95% CI 1.42 to 2.58)). The Panel considers that this difference is unlikely to be of biological relevance.

For high fat mass (1 study), the results of the retrospective cohort study by Magalhaes et al. (2012), comparing an introduction of CFs ≤ 3 with 4–6 months of age did not find a significant association between the timing of introduction of CFs and the outcome. The outcome in this study was assessed at 4–7 years of age and was defined as the age- and sex-specific 85th percentile of the cohort.

> The Panel notes that the results in the supportive line of evidence (seven studies, some investigating several endpoints) are consistent with those in the main line of evidence.

Endpoints investigated in single studies

Other investigated endpoints related to fat (or fat-free) mass were: lean mass z-scores (Leary et al., 2015), lean mass (Mehta et al., 1998), fat-free mass z-score (de Beer et al., 2015), android:

[33] In this study fat mass was measured both by DXA and BIA.

gynoid fat ratio z-score (Durmuş et al., 2014), preperitoneal abdominal fat area z-score (Durmuş et al., 2014), fat mass index z-score and fat-free mass index z-score (Vogelezang et al., 2018) (main line of evidence); and high fat from the android region (Magalhaes et al., 2012) (supportive line of evidence). These were assessed in single studies only. Therefore, they cannot be used to establish the appropriate age range of introduction of CFs (Section 2.2.3.3).

7.4. Skinfold thickness: summary of the evidence

This section discusses first SFT and the related forest plot, then miscellaneous endpoints.

For skinfold thickness, the included studies provided data on SFT measured in one site in the body or a combination of two to four sites (i.e. subscapular, triceps, subscapular + suprailiac, triceps + biceps, triceps + subscapular, triceps + subscapular + suprailiac, triceps + biceps + subscapular + suprailiac SFT, expressed in millimetres). As these measures cannot be directly compared, no meta-analysis was performed (Appendix A.20).

Main line of evidence (2 studies)

The RCT (Perkin et al., 2016) did not find statistically significant differences between the introduction of CFs at 3–4 months of age compared with 6 months of age on triceps and subscapular SFT at 3 years of age. The prospective cohort study (Durmuş et al., 2012) that investigated the association between various combinations of SFT measurements at 2 years of age and the timing of introduction of CFs before 4 months or at 4–5 months vs after 5 months of age, did not find a statistically significant association in most of the comparisons made.

> The Panel notes, from the RCT and the prospective cohort study (Tiers 1 and 2) in the main line of evidence, that there is no association between the timing of introduction of CFs and SFT assessed up to 3 years of age.

Supportive line of evidence (4 studies)

The reasoning behind the use of the data comprised in the supportive line of evidence is explained in Section 2.2.3.3.

This line is composed of the prospective cohort studies by Huh et al. (2011) (Tier 3) and by Kramer et al. (1985a) (in which the age of introduction of CFs was used as a continuous variable in the analysis, Tier 2), as well the cross-sectional study by Patterson et al. (1986) and the cross-sectional analysis of baseline data of a prospective cohort study by Zive et al. (1992) (both Tier 3). There was no evidence for an association between the timing of introduction of CFs and SFT measurements (assessed up to 4 years of age) in these studies.

> The Panel notes that results of the four studies in the supportive line are consistent with the findings of the main line of evidence.

Endpoints investigated in single studies

Other endpoints related to SFT were assessed in the line of supportive evidence in single studies only and were SFT gain (Morgan et al., 2004), %difference in SFT (Caleyachetty et al., 2013). These cannot be used to establish the appropriate age range of introduction of CFs and were not considered further.

7.5. Bone health: summary of the evidence

Endpoints related to bone health were: areal BMC (aBMC), bone mineral density (BMD) and bone area. These were assessed in a single study only (van den Hooven et al., 2016) (Tier 1). Therefore, they cannot be used to establish the appropriate age range of introduction of CFs.

7.6. Body composition: conclusions and grading of the confidence in the evidence

Imprecision: The results of the individual studies did not indicate imprecision.

Inconsistency: The evidence is consistent across populations and endpoints. The results of the studies in the supportive line of evidence (11 studies) are consistent with the findings of the main line of evidence.

Generalisability: The two available RCTs (one on fat mass and one on SFT) were conducted in exclusively breastfed and formula fed infants, and their findings were consistent. Therefore, the Panel does not have concerns with respect to the generalisability of the findings to the whole population of infants living in Europe. Even though the number of prospective cohort studies is limited, they were performed in three different countries and covered a sufficient number of populations. Therefore, the Panel considers that results from these studies can also be generalised.

Publication bias: Publication bias could not be evaluated, because of the insufficient number of studies available.

The Panel concludes from the two RCTs that there is no effect of introduction of CFs at 3–4 months of age compared with 6 months of age on fat mass or SFT (high level of confidence in the evidence).

The Panel concludes from prospective cohort studies (Tiers 1 and 2) that there is no evidence for an association between the age of introduction of CFs, covering a broader range of ages of introduction than the RCTs, on fat mass or SFT (6 studies) (moderate level of confidence in the evidence). The early introduction of CFs was defined in all of these studies as < 4 months of age and later introduction as ≥ 4 months to > 6 months. The latest age of outcome assessment was 13 years.

8. Assessment of the data on atopic diseases in individuals born at term or mixed populations

8.1. Atopic diseases: final body of evidence

The 92 publications that were considered in the assessment in individuals born at term or mixed populations are given in Appendix B.5. These included two publications that were considered together (Kajosaari, 1991, 1994).

These publications reported on results from 79 studies:

- 6 RCTs (5 rated as Tier 1, 1 rated as Tier 2 and 1 rated as Tier 3; 1 study was allocated two different Tiers depending on the outcome that was assessed);
- 45 prospective cohort studies, 5 nested case–control studies (one study was analysed as prospective cohort study and as a nested case–control study), 2 observational analyses of an RCT and 1 pooled analysis of prospective studies (7 rated as Tier 1, 22 rated as Tier 2 and 28 rated as Tier 3; one study was assigned to three Tiers and three studies to two different Tiers, depending on the outcome that was assessed);
- 21 retrospective studies (all Tier 3).

In these studies, six different outcomes (each possibly covering several endpoints) related to atopic-diseases were investigated. Results of all the studies are given in Annex A as Microsoft Excel® file. In addition, for the main endpoints, results are summarised in the forest plots in Appendices A.21–A.39 of this Scientific Opinion.

With respect to the interpretation of the age at introduction of CFs as reported in the following, please refer to Section 2.2.3.3.

8.2. Atopic diseases: endpoint and study selection

As previously explained (Section 2.1.1), the Panel investigated the association between atopic disease-related endpoints and the timing of introduction of CFs in general, as well as of egg, cereals (in particular wheat), fish (as defined in the papers, i.e. generally undefined), peanut and soy (not in the form of infant formula).

When assessing the timing of introduction of individual foods, the comparator can be either continued breast or formula feeding or CFs other than the one under investigation or mixed feeding regimens. This aspect will be discussed in each of the subsections on individual foods.

For food allergy, the Panel decided to draw its conclusions from the disease-related endpoint, i.e. symptomatic food allergy. Data on sensitisation to allergens are used only as supportive evidence to the results from the studies on symptomatic food allergy, as positive results are associated with a higher risk of allergy but alone are not predictive of the disease (Chokshi and Sicherer, 2016). Also, the Panel decided to consider only sensitisation to food allergens and not to aeroallergens (Section 2.1.1.2).

Information in the individual publications was often insufficient to ascertain whether the diagnostic criteria used for the diagnosis of asthma and atopic dermatitis were able to distinguish cases of these diseases from cases of wheeze or eczema due to other causes. Thus, the Panel clustered:

- under 'asthma-like symptoms', the endpoints 'wheeze', 'asthma' and associated endpoints as investigated in the individual studies;
- under 'eczema', the endpoints 'symptomatic eczema' and 'atopic dermatitis' as investigated in the individual studies.

An important consideration in the evaluation of the effect or association between the timing of introduction of CFs and an atopic-disease-related outcome is reverse causality which may be either due to the presence of an atopic family history on the one hand, and to the presence of allergic symptoms before the introduction of CFs on the other hand. In both cases, parents may decide to anticipate or postpone the introduction of CFs (depending on feeding recommendations given) while, at the same time, these children may already be at a higher risk of developing the disease, independent of the timing of introduction of CFs. This aspect was considered in the assessment of the RoB (Appendix B).

Populations considered as being at risk of the disease were those with a first-degree family history of the disease (i.e. presence of symptomatic allergy in at least one of the following: father, mother or siblings) or already showing atopic symptoms other than those related to the disease under investigation (e.g. children with eczema in a study investigating symptomatic food allergy), while the general population comprises at-risk and not-at-risk populations. At-risk populations and the general population were considered separately in this assessment, as potentially differential effects or associations could be observed in these two populations. However, the Panel notes that the above definition of at-risk infants is not comprehensive, as children without a first-degree family history of the disease and without the presence of atopic symptoms may also develop atopic diseases.

In line with the previously described approach of focussing on the most complete datasets for the step of the data extraction (Section 2.2.3.1), for atopic-disease-related endpoints, the most comprehensive population within the general population and within the at-risk population was used (e.g. results from the overall population were retained instead of results obtained in a subgroup of children without atopic symptoms before the introduction of CFs, if there was evidence that results in the overall population were not influenced by reverse causality). Also, the most reliable outcome assessment was kept for studies reporting results for several inter-related endpoints. For example, asthma diagnosed by a physician was retained in the assessment rather than the caregivers' reports of symptoms indicative of asthma.

Eczema and asthma-like symptoms were the most frequently investigated endpoint in prospective observational studies, while symptomatic food allergy was the most investigated endpoint in the RCTs included.

In line with Section 2.2.3.2, all conclusions on atopic-disease related endpoints refer to the general population of infants living in Europe.

8.3. Outcome cluster of atopic diseases: summary of the evidence

8.3.1. Timing of introduction of CFs in general

Main line of evidence

General population (1 study): In the RCT in exclusively breastfed infants (Perkin et al., 2016), no effect of introducing CF at 3–4 months compared with 6 months on the odds of developing an atopic disease up to 3 years of age was observed (Annex A as Microsoft Excel® file).

At-risk population (2 studies): The prospective cohort study (Pöysä et al., 1991) did not find an association between the introduction of CFs < 3 months of age, compared with later, on the odds of developing an atopic disease up to 9–10 years of age. However, there was a high imprecision in the estimate (Annex A as Microsoft Excel® file). Also, Sandini et al. (2011) reported non-significant findings, but without a point estimate, for CF introduction < 4 months of age compared with 4–6 months of age on the odds of developing an atopic disease up to 2 or 5 years of age.

The Panel notes for the general population that the RCT (Tier 1) available in the main line of evidence did not observe an effect of introducing CFs at 3–4 months of age compared with 6 months of age on the odds of developing an atopic disease up to 3 years of age in exclusively breastfed infants.

The Panel notes for the at-risk population that the two prospective cohort studies (Tier 2) in the main line of evidence, one with a high imprecision and one without point estimate, did not show an association between the timing of introduction of CFs and odds of developing an atopic disease up to 9–10 years of age.

Supportive line of evidence

The reasoning behind the use of the data comprised in the supportive line of evidence is explained in Section 2.2.3.3.

- **Prospective cohort studies (2 studies, Tier 3)**

General population (1 study): The prospective cohort study by Keijzers et al. (2018) (Tier 3) did not observe a statistically significant association between introduction of CFs < 6 months of age compared with later and the odds of developing an atopic disease up to 5 years of age.

At-risk population (1 study): The prospective cohort study in exclusively breastfed infants (Kajosaari, 1991, 1994) did not observe an association between the introduction of CFs at 3 months of age, compared with after 6 months, and the outcome assessed at 5 years of age (Annex A as Microsoft Excel® file).

- **Retrospective studies (2 studies, Tier 3)**

General population: The cross-sectional study by Hatakka et al. (2008) did not find a statistically significant association between introduction of CFs < 3 months vs thereafter and the odds of developing an atopic disease up to 1–6 years of age. One case–control study (Parihar et al., 1984) observed higher odds of atopic diseases assessed up to 2 years to be associated with the introduction of CFs < 3 months of age compared with thereafter (OR 7.37 (2.18 to 24.92)). This analysis was unadjusted and therefore is likely to overestimate the association (Annex A as Microsoft Excel® file).

- **Studies in which the timing of introduction of CFs was used as a continuous variable in the analysis, irrespective of the study design (2 studies)**

General population: The two studies that analysed the timing of introduction of CFs as a continuous variable (i.e. the prospective cohort study by Savilahti et al. (1987) (Tier 2) and the cross-sectional study by Forster et al. (1990) (Tier 3) did not find a relationship between the timing of introduction of CFs and the outcome assessed at 2 and 1.5 years, respectively (Annex A as Microsoft Excel® file).

- **Difference in the timing of introduction of CFs between cases and controls (1 study)**

General population: One case–control study (Yung et al., 2015) (Tier 3) did not find a statistically significant difference in the timing of introduction of CFs between on average 20-month-old cases with atopic diseases and controls (Annex A as Microsoft Excel® file).

The Panel notes for the general population that only one of the six studies in the supportive line of evidence found an association between the introduction of CFs < 3 months vs later and higher odds of developing an atopic disease in an unadjusted analysis, while the five others did not find an association. The Panel considers that the results of the supportive line of evidence in the general population are consistent with those the main line of evidence.

The Panel notes for the at-risk population that the result of the single study in the supportive line of evidence is consistent with the findings in the main line of evidence.

8.3.2. Timing of introduction of specific foods

Only single studies were available to assess the timing of introduction of specific foods, i.e. egg (Halpern et al., 1973) and cereals (Jonsson et al., 2017) (both supportive line of evidence) (Annex A as Microsoft Excel® file). Therefore, these cannot be used to establish the appropriate age range of introduction of CFs and were not considered further.

8.3.3. Outcome cluster of atopic diseases: conclusions and grading of the confidence in the evidence

Imprecision: Contrary to the RCT in the general population, the results of one of the two prospective cohort studies in at-risk populations showed a high imprecision. The other one did not provide a point estimate to allow a judgement to be made. Therefore, the Panel downgraded by one category the confidence in the evidence derived from the cohort studies in at-risk populations.

Inconsistency: The limited evidence that is available is consistent between the at-risk population and the general population and the results of the six out of seven studies in the supportive line of evidence were consistent with the main line of evidence.

Generalisability: The study population of the RCT available in the general population consisted only of breastfed infants. The Panel considers that the results of this study cannot be generalised to formula fed infants and thus to the whole population of infants living in Europe. Therefore, the Panel downgraded by one category the confidence in the evidence derived from the RCT. With respect to the prospective cohort studies, the Panel notes that even though only one single study in a small population was available in the main line of evidence in an at-risk population, the Panel did not have concerns with respect to generalisability of the findings to the whole population of infants living in Europe, considering that the evidence was consistent across populations, taking into account also the supportive line of evidence.

Publication bias: Publication bias could not be evaluated, because of the insufficient number of studies available.

Other: The Panel noted the limited evidence available for assessment in the main line of evidence (one study in the general population and one in an at-risk population). As six out of seven studies in the supportive line of evidence were consistent with the findings in the main line of evidence, the Panel decided not to downgrade the level of confidence in the main line of evidence for this low number of studies.

> The Panel concludes from the RCT (Tier 1) that there is no evidence for an effect of introduction of CFs at 3–4 months vs 6 months of age on the odds of developing an atopic disease up to 3 years of age (moderate confidence in the evidence).
>
> The Panel concludes from the two prospective cohort studies (Tier 2) that there is no evidence for an association between the introduction of CFs at < 3 months of age vs thereafter and at < 4 vs 4–6 months and the odds of developing an atopic disease up to 9–10 years of age (low confidence in the evidence).

8.4. Asthma-like symptoms: summary of the evidence

8.4.1. Timing of introduction of CFs in general

Main line of evidence

General population (5 studies): The RCT (Perkin et al., 2016) did not show an effect of introduction of CFs at 3–4 months compared with 6 months of age on the odds of developing asthma-like symptoms up to 3 years of age.

The result of the meta-analysis of the two prospective cohort studies that provided information that could be used for this analysis (Zutavern et al., 2004; Lossius et al., 2018) was not statistically significant. Heterogeneity was not important (I^2 = 27%). The imprecision around the pooled estimate was serious (Appendix A.21). However, also individually, these studies did not show an association between the timing of introduction of CFs (≤ 3 months and < 6 months vs thereafter, respectively) and the outcome. Equally, the studies by Wilson et al. (1998) and Nwaru et al. (2013a) that did not provide data in a form that could be incorporated into the meta-analysis did not observe an association between the timing of introduction of CFs and the outcome. The latest age at outcome assessment was 7 years.

At-risk populations (3 studies): The meta-analysis of the two prospective cohort studies (Marini et al., 1996; Mihrshahi et al., 2007) did not show an association between the introduction of CFs < 4 and < 3 months, respectively, compared with thereafter, and the outcome assessed up to 5 years of age. Heterogeneity was not important (I^2 = 0%) (Appendix A.22). The imprecision around the pooled estimate was serious. However, individually these two studies also did not observe an association. In

addition, Sandini et al. (2011) did not find an association between CF introduction < 4 months of age compared with 4–6 months and the odds of developing asthma-like symptoms up to 2 or 5 years of age, but did not provide a point estimate.

The Panel notes for the general population that, from one RCT and the four prospective cohort studies (Tiers 1 and 2) in the main line of evidence, there is no evidence for an association between the timing of introduction of CFs and the odds of developing asthma-like symptoms up to 7 years of age.

The Panel notes for the at-risk population that, from the three prospective cohort studies (Tier 2) in the main line of evidence, there is no evidence for an association between the timing of introduction of CFs and the odds of developing asthma-like symptoms up to 5 years of age.

Supportive line of evidence

The reasoning behind the use of the data comprised in the supportive line of evidence is explained in Section 2.2.3.3.

- **Prospective cohort studies (9 studies, Tier 3)**

General population (7 studies): The meta-analysis of five studies (Fergusson et al., 1983; Larsson et al., 2008; Snijders et al., 2008; Zutavern et al., 2008; Hetzner et al., 2009) comparing various time points of introduction of CFs in relation to the odds of developing asthma-like symptoms up to 9 years of age was not statistically significant. Heterogeneity was not important (I^2 = 32%) (Appendix A.21). In addition, Morgan et al. (2004) and Kurukulaaratchy et al. (2004), that did not provide a point estimate, reported non-significant findings in relation to the outcome at 1.5 and 10 years of age, respectively, comparing introduction of CFs < 3 with > 3 months of age.

At-risk populations (2 studies): The meta-analysis of the two prospective cohort studies reported in three publications (Van Asperen et al., 1984; Kajosaari, 1991, 1994) did not find an association between the timing of introduction of CFs (≤ 4 vs > 4 months and 3 vs > 6 months, respectively) and the outcome assessed up to 5 years of age. Heterogeneity was not important (I^2 = 0%) (Appendix A.22). The imprecision around the pooled estimate was serious. However, individually these two studies also did not observe an association.

- **Retrospective studies (1 study, Tier 3)**

General population: The case–control study by Karunasekera et al. (2001) did not find a relationship between the timing of introduction of CFs before and after 3 months of age and the odds of developing asthma-like symptoms up to 1–10 years of age (Annex A as Microsoft Excel® file).

- **Difference in the timing of introduction of CFs between cases and controls (1 study)**

General population: A nested case–control study (Hesselmar et al., 2010) (Tier 1) found no statistically significant difference between the timing of introduction of CFs in 1.5-year-old cases with asthma-like symptoms and controls (Annex A as Microsoft Excel® file).

The Panel notes for the general population that the results in the supportive line of evidence are consistent with those in the main line of evidence.

The Panel notes for the at-risk population that the results in the supportive line of evidence are consistent with those in the main line of evidence.

8.4.2. Timing of introduction of CFs in general and asthma-like symptoms: conclusions and grading of the confidence in the evidence

Imprecision: The imprecision in the results of the meta-analyses of the prospective cohort studies in the main line of evidence in the general as well as the at-risk populations was serious. Therefore, the Panel downgraded by one category the confidence in the evidence.

Inconsistency: The evidence is consistent across populations and the results of the studies in the supportive line of evidence (nine for the general population and two for at-risk populations) were

consistent with the main line. For all meta-analyses conducted in the main line of evidence, I^2 was below 75%.

Generalisability: The study population of the RCT consisted only of breastfed infants. The Panel considers that the results of this study cannot be generalised to formula fed infants and thus to the whole population of infants living in Europe. Therefore, the Panel downgraded by one category the confidence in the evidence derived from this RCT. With respect to prospective cohort studies, the Panel did not have any concerns with respect to their generalisability, as a representative number of populations were studied.

Publication bias: Publication bias could not be evaluated, because of the insufficient number of studies available.

The Panel concludes from the RCT (Tier 1) that there is no evidence for an effect of introduction of CFs at 3–4 months of age compared with 6 months of age on the odds of developing asthma-like symptoms up to 3 years of age (moderate level of confidence in the evidence).

The Panel concludes from the seven prospective cohort studies (Tiers 1 and 2) that there is no evidence for an association between the age of introduction of CFs, covering a range of ages from ≤ 3 months to < 6 months that was compared mostly with thereafter, and the odds of developing asthma-like symptoms up to 7 years of age (low level of confidence in the evidence).

8.4.3. Timing of introduction of egg

Main and supportive lines of evidence

General population (2 studies): Neither the prospective cohort study in the main line of evidence (Nwaru et al., 2013b) (Tier 2) nor the one in the supportive line of evidence (Tromp et al., 2011) (Tier 3) found an association between the timing of introduction of egg (comparing introduction at < 5 and ≤ 6 months to thereafter, respectively) and the odds of developing asthma-like symptoms up to 10 years of age (Annex A as Microsoft Excel® file).

At-risk populations (2 studies): Neither the RCT by Palmer et al. (2017) nor the prospective cohort study by Nwaru et al. (2013b) (both Tier 2; main line of evidence) showed an effect or association between the timing of introduction of egg (comparing introduction at 4–6.5 with ≥ 10 months and < 5 months with thereafter, respectively) and the odds of developing asthma-like symptoms up to 10 years of age (Annex A as Microsoft Excel® file).

The Panel notes for the general population that the prospective cohort study (Tier 2) in the main line of evidence did not show an association between the timing of introduction of egg and the odds of developing asthma-like symptoms up to 10 years of age. The study in the general population in the supportive line of evidence is consistent with this finding.

The Panel notes for the at-risk population that the results of one RCT and one prospective cohort study (Tier 2) in the main line of evidence did not show an effect or association between the timing of introduction of egg and the odds of developing asthma-like symptoms up to 10 years of age.

8.4.4. Timing of introduction of egg and asthma-like symptoms: conclusions and grading of the confidence in the evidence

Imprecision: There were no concerns with respect to imprecision of results.

Inconsistency: The evidence is consistent across populations and the results of the study in the supportive line of evidence is consistent with the main line.

Generalisability: The RCT was performed in Australia, a country that has an unexplained higher prevalence of allergy than Europe. Further the study used pasteurised raw egg powder as an intervention product, which is not the form that would be used when egg is introduced to infants. Therefore, the Panel downgraded the confidence level in the evidence twice for the RCT. With respect to prospective cohort studies, even though only one study in an at-risk population was available in the main line of evidence, the Panel did not have concerns with respect to generalisability of the findings to the whole population of infants living in Europe, considering the consistency across populations, taking into account also the study in the supportive line of evidence.

Publication bias: Publication bias could not be evaluated, because of the insufficient number of studies available.

Other: The confidence in the evidence was downgraded by one category because of the overall limited evidence that was available for the outcome in the main and the supportive lines of evidence.

> The Panel concludes, from one RCT and one prospective cohort study (Tier 2), that there is no evidence for an effect or association between the introduction of egg, at 4–6.5 months vs ≥ 10 months of age and < 5 months vs ≥ 5 months, and the odds of developing asthma-like symptoms up to 10 years of age (low level of confidence in the evidence).

8.4.5. Timing of introduction of cereals

Main line of evidence

General population (3 studies): From the meta-analysis of three prospective cohort studies (Zutavern et al., 2004; Nwaru et al., 2013a,b), there was no evidence for an association between the timing of introduction of cereals (comparing various time points between 3.75 and 5.5 months with thereafter) and the odds of developing asthma-like symptoms up to 10 years of age. Heterogeneity was moderate (I^2 = 48%) (Appendix A.23).

At-risk populations (1 study): In one study (Nwaru et al., 2013b), no association between the timing of introduction of cereals and the odds of developing asthma-like symptoms up to 10 years of age was observed, comparing an introduction of cereals < 3.75 months with thereafter (Annex A as Microsoft Excel® file).

> The Panel notes for the general population that, from the three prospective cohort studies (Tier 2) in the main line of evidence, there is no evidence for an association between the timing of introduction of cereals and the odds of developing asthma-like symptoms up to 10 years of age.
>
> The Panel notes for the at-risk population that, from the prospective cohort study (Tier 2) in the main line of evidence, there is no evidence for an association between the timing of introduction of cereals and the odds of developing asthma-like symptoms up to 10 years of age.

Supportive line of evidence

The reasoning behind the use of the data comprised in the supportive line of evidence is explained in Section 2.2.3.3.

General population (1 study): The prospective cohort study (Tromp et al., 2011) (Tier 3) did not find an association with the outcome assessed at 4 years of age comparing introduction of cereals ≤ 6 months with thereafter (Appendix A.23).

> The Panel notes for the general population that the result of the study in the supportive line of evidence is consistent with the findings in the main line of evidence.

8.4.6. Timing of introduction of cereals and asthma-like symptoms: conclusions and grading of the confidence in the evidence

Imprecision: There were no concerns with respect to imprecision of results.

Inconsistency: The evidence is consistent across populations and the results of the single study in the supportive line of evidence is consistent with the main line. For the meta-analysis conducted in the main line of evidence, I^2 was below 75%.

Generalisability: There were no concerns with respect to generalisability of the results of the prospective cohort studies.

Publication bias: Publication bias could not be evaluated, because of the insufficient number of studies available.

The Panel concludes from the three prospective cohort studies (Tier 2) that there is no evidence for an association between the age of introduction of cereals, covering a range of ages from < 3.75 months to ≤ 5.5 months vs thereafter, and the odds of developing asthma-like symptoms up to 10 years of age (moderate confidence in the evidence).

8.4.7. Timing of introduction of fish

Main line of evidence

General population (3 studies): From the meta-analysis of three prospective cohort studies (Zutavern et al., 2004; Virtanen et al., 2010; Nwaru et al., 2013b), there was no evidence for an association between the timing of introduction of fish (mostly comparing introduction < 5–6 months of age with > 5–6 months to > 8.5 months) and the odds of developing asthma-like symptoms up to 10 years of age. Heterogeneity was not important (I^2 = 0%) (Appendix A.24).

At-risk population (1 study): One study (Nwaru et al., 2013b) found no association between the timing of introduction of fish before and after 5.25 months of age and the development of asthma-like symptoms up to 10 years of age (Annex A as Microsoft Excel® file).

The Panel notes for the general population that, from the three prospective cohort studies (Tier 2) in the main line of evidence, there is no evidence for an association between the timing of introduction of fish and the odds of developing asthma-like symptoms up to 10 years of age.

The Panel notes for the at-risk population that, from the prospective cohort study (Tier 2) in the main line of evidence, there is no evidence for an association between the timing of introduction of fish and the odds of developing asthma-like symptoms up to 10 years of age.

Supportive line of evidence

The reasoning behind the use of the data comprised in the supportive line of evidence is explained in Section 2.2.3.3.

- **Prospective cohort studies (1 study, Tier 3)**

General population (1 study): The prospective cohort study by Kiefte-de Jong et al. (2012) found that introduction of fish before 6 months of age compared with introduction between 6 and 12 months of age was associated with higher odds of asthma-like symptoms at 4 years of age (aOR 1.53 (1.07 to 2.19)) (Appendix A.24).

The Panel notes for the general population that the result of the study in the supportive line of evidence is not consistent with those in the main line of evidence.

8.4.8. Timing of introduction of fish and asthma-like symptoms: conclusions and grading of the confidence in the evidence

Imprecision: There were no concerns with respect to imprecision of results.

Inconsistency: The evidence is consistent across populations in the main line of evidence. The result of the prospective cohort study in the supportive line of evidence is inconsistent with the main line. However, as there was enough evidence in the main line, the Panel did not downgrade the confidence in the evidence for this inconsistent finding. For the meta-analysis conducted in the main line of evidence, I^2 was below 75%.

Generalisability: There were no concerns with respect to generalisability of the results of the prospective cohort studies.

Publication bias: Publication bias could not be evaluated, because of the insufficient number of studies available.

The Panel concludes from the three prospective cohort studies (Tier 2) that there is no evidence for an association between the age of introduction of fish, covering a range of ages from < 5–6 months for earlier introduction and > 5–6 months to > 8.5 months for later introduction, and the odds of developing asthma-like symptoms up to 10 years of age (moderate confidence in the evidence).

8.4.9. Timing of introduction of soy and peanut

Only one study (Tromp et al., 2011) (supportive line of evidence) investigated the relationship between the timing of introduction of soy and peanut and the odds of developing asthma-like symptoms. Therefore, this study cannot be used to establish the appropriate age range of introduction of CFs and was not considered further.

8.5. Eczema: summary of the evidence

8.5.1. Timing of introduction of CFs in general

Main line of evidence

General population (5 studies): From the meta-analysis of five prospective cohort studies (Fergusson et al., 1981; Forsyth et al., 1993; Zutavern et al., 2004; Chuang et al., 2011; Roduit et al., 2012) (Appendix A.25), there was no evidence for an association between the timing of introduction of CFs, in most cases ≤ 3–4 months vs thereafter, and the odds of developing eczema up to 5.5 years of age. Heterogeneity was moderate (I^2 = 46%).

At-risk population (8 studies): The meta-analysis of the six prospective cohort studies (Fergusson et al., 1981; Ruiz et al., 1992; Marini et al., 1996; Schoetzau et al., 2002; Mihrshahi et al., 2007; Roduit et al., 2012) showed no association between the introduction of CFs (mostly before 3 or 4 months vs later) and the odds of developing eczema assessed up to 5 years of age. Heterogeneity was moderate to substantial (I^2 = 53%) (Appendix A.26). In addition, Sandini et al. (2011) and Moore et al. (1985), who did not provide point estimates, did not find an association between CF introduction < 4 months of age compared with 4–6 months and < 3 months of age compared with later, and the odds of developing eczema up to 2 or 5 years and 1 year of age, respectively.

The Panel notes for the general population that, from the five prospective cohort studies in the main line of evidence (Tiers 1 and 2), there is no evidence for an association between the timing of introduction of CFs and the odds of developing eczema up to 5.5 years of age.

The Panel notes for the at-risk population that, from the eight prospective cohort studies (Tiers 1 and 2) in the main line of evidence, there is no evidence for an association between the timing of introduction of CFs and the odds of developing eczema up to 5 years of age.

Supportive line of evidence

The reasoning behind the use of the data comprised in the supportive line of evidence is explained in Section 2.2.3.3.

- **Prospective cohort studies (14 studies, Tier 3)**

General population (11 studies): The results of these studies are heterogeneous. From the meta-analysis of nine studies (Hide and Guyer, 1981; Dunlop et al., 2006; Filipiak et al., 2007; Larsson et al., 2008; Snijders et al., 2008; Zutavern et al., 2008; Sariachvili et al., 2010; Huang et al., 2013; Taylor-Robinson et al., 2016), there was no evidence for an association between the timing of introduction of CFs (comparing various time points, but mostly below 3 or 4 months vs later) and the odds of developing eczema up to 9 years of age (Appendix A.25). Heterogeneity was substantial to considerable (I^2 = 75%) and cannot be explained. In addition, Morgan et al. (2004) and Nwaru et al. (2013a), who did not provide a point estimate, did not observe an association between the introduction of CFs before around 3 months of age compared with later and the odds of developing eczema up to 1.5 and 5 years of age, respectively (Annex A as Microsoft Excel® file).

At-risk populations (3 studies): The meta-analysis of three prospective studies (reported in 4 publications) (Van Asperen et al., 1984; Kajosaari, 1991, 1994; Ranucci et al., 2018) did not show an association between the timing of introduction of CFs (comparing various time points) and the

outcome investigated up to 5 years of age. Heterogeneity was not important (I^2 = 0%) (Appendix A.26).

- **Retrospective studies (6 studies, Tier 3)**

General population (5 studies): The results of these studies are also heterogeneous. The meta-analysis of the three case–control studies (Haileamlak et al., 2005; Sahakyan et al., 2006; Turati et al., 2016) together with the two cross-sectional studies (Zheng et al., 2016; Lee et al., 2017) did not show evidence for an association between the timing of introduction of CFs (comparing introduction < 4 months with later) and the odds of developing eczema up to 7 years of age (Appendix A.27). Heterogeneity was moderate to substantial (I^2 = 51%).

At-risk population (1 study): In the cross-sectional study by Suryati et al. (2006), no association between the timing of introduction of CFs (< vs ≥ 4 months) and the odds of developing eczema up to 1–5 years of age was observed (Annex A as Microsoft Excel® file).

- **Studies in which the timing of introduction of CFs was used as a continuous variable in the analysis, irrespective of the study design (2 studies)**

General population (2 studies): In the cross-sectional study by Takahashi et al. (1999) (Tier 3), no association was observed between the timing of introduction of CFs and the outcome at 1–2 years of age. Illi et al. (2004), in a prospective cohort study (Tier 3), did not find an association between the timing of introduction of CFs and the cumulative odds for developing eczema up to 2 years of age, but did not provide a point estimate. (Annex A as Microsoft Excel® file).

- **Difference in the timing of introduction of CFs between cases and controls (3 studies)**

General population (3 studies): Sariachvili et al. (2010) (Tier 3), in a nested case–control study, found a lower likelihood that 4-year-old cases with eczema had been introduced to CFs before 4 months of age (OR 0.60 (95% CI 0.43 to 0.84)). However, the analysis was unadjusted and therefore is likely to overestimate the association. In the nested case–control study by Hesselmar et al. (2010) (Tier 1) the timing of introduction of CFs was not statistically significantly different between 18-months-old cases and controls. Also, in the case–control study by Kramer and Moroz (1981) (Tier 3), no statistically significant differences in the timing of introduction of CFs were observed between 1- to 20-year-old cases and their controls (Annex A as Microsoft Excel® file).

> The Panel notes for the general population that the results in the supportive line of evidence are consistent with those in the main line of evidence.
>
> The Panel notes for the at-risk population that the results in the supportive line of evidence are consistent with those in the main line of evidence.

8.5.2. Timing of introduction of CFs in general and eczema: conclusions and grading of the confidence in the evidence

Imprecision: There was no imprecision associated with the results of the meta-analyses.

Inconsistency: The evidence is consistent across populations and the results of the studies in the supportive line of evidence (21 studies for the general population, 4 in at-risk populations), are consistent with the results of the studies in the main line of evidence. For all meta-analyses conducted in the main line of evidence, I^2 was below 75%.

Generalisability: There were no concerns with respect to the generalisability of the findings to the whole population of infants living in Europe.

Publication bias: From the visual inspection of the funnel plots of studies performed in the general population and at-risk populations (Annexes D.7 and D.8), there was no convincing evidence of asymmetry.

> The Panel concludes from the 11 prospective cohort studies (two in common in the general and at-risk populations; Tiers 1 and 2) that there is no evidence for an association between the age of introduction of CFs, covering a range of ages from < 3 months to ≤ 6 months vs thereafter and the odds of developing eczema up to 5.5 years of age (moderate confidence in the evidence).

8.5.3. Timing of introduction of egg

Main line of evidence

General population (2 studies): The two prospective cohort studies (Fergusson et al., 1990; Nwaru et al., 2013b) did not observe a relationship between the timing of introduction of egg < 5 months compared with later and the odds of developing eczema up to 10 years of age in the general population (Fergusson et al., 1990 did not provide a point estimate) (Appendix A.28 and Annex A as Microsoft Excel® file).

At-risk populations (5 studies): Neither the meta-analysis nor the results of the two RCTs considered individually (Palmer et al., 2017; Tan et al., 2017) (Tier 1) showed an effect of egg introduction at around 4–6 months of age, compared with an introduction at 8–10 months of age, on the odds of developing eczema up to 1 year of age. This is also true for the meta-analysis and for the results of the two individual prospective cohort studies (Ruiz et al., 1992; Nwaru et al., 2013b) (Tier 2) that investigated egg introduction ≤ 6 and < 5 months vs later, respectively. The latest age at outcome assessment was 10 years. The pooled estimates of the two meta-analyses were associated with a serious imprecision. Heterogeneity was substantial (I^2 = 69%) and moderate (I^2 = 37%), respectively (Appendix A.29).

In addition, the prospective cohort study by Fergusson et al. (1981) (Tier 2) (same study as Fergusson et al. (1990)), who did not provide a point estimate, showed no evidence for an association between egg introduction ≤ 4 months compared with later and the odds of developing eczema up to 2 years (Annex A as Microsoft Excel® file).

> The Panel notes for the general population that, from the two prospective cohort studies (Tier 2) in the main line of evidence, there is no evidence for an association between the timing of introduction of egg and the odds of developing eczema up to 10 years of age.
>
> The Panel notes for the at-risk population that, from the two RCTs and three prospective cohort studies (Tiers 1 and 2) in the main line of evidence, there is no evidence for an effect or association between the timing of introduction of egg and the odds of developing eczema up to 10 years of age.

Supportive line of evidence

The reasoning behind the use of the data comprised in the supportive line of evidence is explained in Section 2.2.3.3.

- **Prospective cohort studies (3 studies, Tier 3)**

General population (3 studies): The result of the meta-analysis of three prospective cohort studies (Zutavern et al., 2006; Filipiak et al., 2007; Elbert et al., 2017) comparing egg introduction before vs after 6 months of age was not statistically significant. Eczema was investigated up to 10 years of age. Heterogeneity was moderate (I^2 = 45%) (Appendix A.28).

At-risk populations (1 study): One prospective cohort study that investigated the association between egg introduction and eczema in the general population (Filipiak et al., 2007), provided also results on an at-risk population. It did not observe an association between egg introduction before and after 6 months of age and the odds of developing eczema up to 4 years of age (Annex A as Microsoft Excel® file).

- **Retrospective studies (2 studies, Tier 3)**

General population (1 study): In a cross-sectional analysis of baseline data of a prospective cohort study (Peters et al., 2015), no statistically significant differences were observed between various time points of egg introduction (i.e. < 4 vs 4–6, 4–6 vs 7–9 and 4–6 vs 10–12 months), except for those introduced to egg at 4–6 months of age compared with those introduced after 12 months of age. The odds of eczema were statistically significantly lower in the earlier group (aOR 0.5 (95% CI 0.33 to 0.74)) (Annex A as Microsoft Excel® file).

At-risk populations (1 study): In the cross-sectional study by Suryati et al. (2006), no association was observed between egg introduction < 4 months compared with ≥4 months of age and the odds of developing eczema up to 1–5 years of age. However, imprecision was serious in this study (Annex A as Microsoft Excel® file).

- **Studies in which the timing of introduction of egg was used as a continuous variable in the analysis, irrespective of the study design (1 study)**

General population (1 study): The cross-sectional study by Takahashi et al. (1999) (Tier 3) observed lower odds of eczema at 1–2 years to be associated with earlier introduction of egg (aOR for one month of earlier introduction 0.94 (95% CI 0.89 to 0.99)) (Annex A as Microsoft Excel® file).

> The Panel notes for the general population that, in the supportive line of evidence, two retrospective studies observed lower odds of eczema to be related to earlier introduction of egg, while in three prospective cohort studies no association was observed between the timing of introduction of egg and this outcome.
>
> The Panel notes for the at-risk population that the results in the supportive line of evidence are consistent with those in the main line of evidence.

8.5.4. Timing of introduction of egg and eczema: conclusions and grading of the confidence in the evidence

Imprecision: There was serious imprecision associated with the pooled estimate of the meta-analyses of the two RCTs and the two prospective cohort studies in at-risk populations. Therefore, the Panel downgraded by one category the confidence in the evidence in these lines of evidence.

Inconsistency: The findings were consistent across populations in the main lines of evidence. There were six studies in the supportive line of evidence (five in the general population, two in at-risk populations; one in common in both groups). While in the at-risk population the findings in the supportive line of evidence were consistent with the main line of evidence, the results of studies in the general population in the supportive line were inconsistent. However, as there was enough evidence available in the main line of evidence, the Panel did not downgrade the confidence in the evidence for this finding. For all meta-analyses conducted in the main line of evidence, I^2 was below 75%.

Generalisability: The two RCTs in at-risk populations were performed in Australia, a country that has an unexplained higher prevalence of allergy than Europe. Further, the study used pasteurised raw egg powder as an intervention product, which is not the form that would be used when egg is introduced to infants. Therefore, the Panel downgraded the confidence level in the evidence twice for these RCTs. With respect to prospective cohort studies, a representative number of populations were studied. Therefore, the Panel considers that their results can be generalised to the whole population of infants living in Europe.

Publication bias: Publication bias could not be evaluated, because of the insufficient number of studies available.

> The Panel concludes from the two RCTs and the three prospective cohort studies (Tiers 1 and 2) that there is no evidence for an effect or association between the timing of introduction of egg, covering a range of ages from ≤ 4 months to ≤ 6 months compared with thereafter and the odds of developing eczema up to 10 years of age (very low to low confidence in the evidence).

8.5.5. Timing of introduction of cereals

Main line of evidence

General population (3 studies): Neither the meta-analysis nor the results of the two prospective cohort studies considered individually (Zutavern et al., 2004; Nwaru et al., 2013b) showed an association between cereal introduction about ≤ 4 months of age compared with thereafter and the odds of developing eczema up to the age of 10 years. The pooled estimate of the meta-analysis was associated with serious imprecision. Heterogeneity was substantial (I^2 = 71%) (Appendix A.30). In addition, in the prospective cohort study by Fergusson et al. (1990) that did not provide a point estimate, no evidence for an association between cereal introduction ≤ 4 months compared with later and the odds of developing eczema up to 10 years was observed (Annex A as Microsoft Excel® file).

At-risk populations (2 studies): Two prospective cohort studies (Fergusson et al., 1981; Nwaru et al., 2013b) (Fergusson et al. (1981) report on the same study as Fergusson et al. (1990)) did not show an association between the timing of introduction of cereals (≤ 4 months of age vs thereafter) and the odds of developing eczema up to 10 years of age (Annex A as Microsoft Excel® file).

The Panel notes for the general population that, from the three prospective cohort studies (Tier 2) in the main line of evidence, there is no evidence for an association between the timing of introduction of cereals and the odds of developing eczema up to 10 years of age.

The Panel notes for the at-risk population that, from the two prospective cohort studies (Tier 2) in the main line of evidence, there is no evidence for an association between the timing of introduction of cereals and the odds of developing eczema up to 10 years of age.

Supportive line of evidence

The reasoning behind the use of the data comprised in the supportive line of evidence is explained in Section 2.2.3.3.

- **Prospective cohort studies (4 studies, Tier 3)**

General population (4 studies): The result of the meta-analysis of the three prospective cohort studies (Zutavern et al., 2006; Filipiak et al., 2007; Elbert et al., 2017), that investigated the association between cereal introduction ≤ 4 months and ≤ 6 months of age vs thereafter and the odds of developing eczema up to 10 years of age, was not statistically significant. Heterogeneity was not important (I^2 = 17%) (Appendix A.30). In addition, Nwaru et al. (2013a), that did not provide a point estimate, did not find an association between cereal introduction before 5 months vs thereafter on the odds of developing eczema up to 5 years of age (Annex A as Microsoft Excel® file).

At-risk population (1 study): The prospective cohort study that investigated the outcome in the general population, investigated it also in an at-risk population (Filipiak et al., 2007). This study did not find an association between the introduction of cereals ≤ 4 months of age vs thereafter and the odds of developing eczema up to 4 years (Annex A as Microsoft Excel® file).

The Panel notes for the general population that the results in the supportive line of evidence are consistent with those in the main line of evidence.

The Panel notes for the at-risk population that the results in the supportive line of evidence are consistent with those in the main line of evidence.

8.5.6. Timing of introduction of cereals and eczema: conclusions and grading of the confidence in the evidence

Imprecision: The imprecision associated with the results of the meta-analysis of the two studies in the general population was serious. Therefore, the Panel decided to downgrade by one category the confidence in the evidence.

Inconsistency: The evidence is consistent across populations. The findings in the supportive line of evidence (four studies in the general population, including one also in an at-risk population) were consistent with the findings in the main line of evidence. For the meta-analysis conducted in the main line of evidence, I^2 was below 75%.

Generalisability: A representative number of populations has been studied. Therefore, the Panel did not have any concerns with respect to the generalisability of the findings to the whole population of infants living in Europe.

Publication bias: Publication bias could not be evaluated, because of the insufficient number of studies available.

The Panel concludes from the three prospective cohort studies (Tiers 2) that there is no evidence for an association between the introduction of cereals ≤ 4 months compared with thereafter and the odds of developing eczema up to 10 years of age (low confidence in the evidence).

8.5.7. Timing of introduction of fish

Main line of evidence

General population (2 studies): The meta-analysis of two prospective cohort studies (Zutavern et al., 2004; Nwaru et al., 2013b) did not show an association between fish introduction ≤ 5–6 months of age compared with thereafter and the odds of developing eczema up to the age of 10 years. Heterogeneity was not important (I^2 = 0%) (Appendix A.31).

At-risk population (1 study): The prospective cohort study (Nwaru et al., 2013b) did not find an association between introduction of fish before 5.25 months of age compared with thereafter and the odds of developing eczema up to 5 years of age (Annex A as Microsoft Excel® file).

> The Panel notes for the general population, from the two prospective cohort studies (Tier 2) in the main line of evidence, that there is no evidence for an association between the timing of introduction of fish and the odds of developing eczema up to 10 years of age.
>
> The Panel notes for the at-risk population, from the prospective cohort study (Tier 2) in the main line of evidence, that there is no evidence for an association between the timing of introduction of fish and the odds of developing eczema up to 5 years of age.

Supportive line of evidence

The reasoning behind the use of the data comprised in the supportive line of evidence is explained in Section 2.2.3.3.

- **Prospective cohort studies (4 studies, Tier 3)**

General population (4 studies): The meta-analysis of three prospective cohort studies (Zutavern et al., 2006; Filipiak et al., 2007; Alm et al., 2009) did not show an association between fish introduction ≤6 months of age compared with thereafter and the odds of developing eczema up to the age of 4 years. Heterogeneity was not important (I^2 = 0%) (Appendix A.31). In addition, Nwaru et al. (2013a), that did not provide a point estimate, did not find an association between introduction of fish before 6 months of age compared with after 9 months of age and the odds of developing eczema up to 5 years of age (Annex A as Microsoft Excel® file).

> The Panel notes for the general population that the results in the supportive line of evidence are consistent with those in the main line of evidence.

8.5.8. Timing of introduction of fish and eczema: conclusions and grading of the confidence in the evidence

Imprecision: There were no concerns with respect to imprecision.

Inconsistency: The evidence is consistent across populations. The results of the supportive line of evidence in the general population (4 studies; no studies in at-risk populations) were consistent with the findings of the studies in the main line of evidence. For the meta-analysis conducted in the main line of evidence, I^2 was below 75%.

Generalisability: Both prospective cohort studies in the general or at-risk populations were conducted in the UK. Owing to the limited number of populations studied, generalisability is uncertain. Therefore, the Panel downgraded by one category the evidence.

Publication bias: Publication bias could not be evaluated, because of the insufficient number of studies available.

> The Panel concludes from the two prospective cohort studies (Tiers 2) that there is no evidence for an association between the introduction of fish at ≤ 5–6 months compared with later and the odds of developing eczema up to 10 years of age (low confidence in the evidence).

8.5.9. Timing of introduction of soy or peanut

Only one prospective cohort study reported in two publications (Tromp et al., 2011; Elbert et al., 2017) (supportive line of evidence) investigated the relationship between the timing of introduction of soy and peanut and the odds of developing eczema. Therefore, this study cannot be used to establish the appropriate age range of introduction of CFs.

8.6. Allergic rhinitis: summary of the evidence

8.6.1. Timing of introduction of CFs in general

Main line of evidence

General population (1 study): The RCT (Perkin et al., 2016) did not show an effect of introduction of CFs at 3–4 months of age compared with introduction at 6 months of age on the odds of developing allergic rhinitis up to 3 years of age in exclusively breastfed infants (Appendix A.32).

At-risk population (2 studies): The prospective cohort study rated as Tier 2 (Marini et al., 1996) did not show a relationship between the timing of introduction of CFs (≤ 4 months vs later) and the odds of developing allergic rhinitis up to 1–3 years of age. In addition, Sandini et al. (2011) (Tier 2) did not find an association between CF introduction < 4 months of age compared with 4–6 months and the odds of developing allergic rhinitis up to 2 or 5 years of age, but did not provide a point estimate. (Annex A as Microsoft Excel® file).

> The Panel notes for the general population, from the RCT (Tier 1) in the main line of evidence, that there is no evidence for an effect of the timing of introduction of CFs on the odds of developing allergic rhinitis up to 3 years of age.
>
> The Panel notes for the at-risk population, from the two prospective cohort studies (Tier 2) in the main line of evidence, that there is no evidence for an association between the timing of introduction of CFs and the odds of developing allergic rhinitis up to 5 years of age.

Supportive line of evidence

The reasoning behind the use of the data comprised in the supportive line of evidence is explained in Section 2.2.3.3.

- **Prospective cohort studies (6 studies, Tier 3)**

General population (5 studies): The meta-analysis of the four prospective cohort studies (Wright et al., 1994; Strachan et al., 1996; Larsson et al., 2008; Zutavern et al., 2008) that investigated various time points with respect to the timing of introduction of CFs (from ≤ 1 to ≤ 6 months vs thereafter) did not show an association between the timing of introduction of CFs and the odds of developing allergic rhinitis up to 16 years of age. Heterogeneity was not important (I^2 = 1%) (Appendix A.32). In addition, Nwaru et al. (2013a), that did not provide a point estimate, reported non-significant findings, comparing an introduction of CFs ≤ 4 months with thereafter, in relation to the odds of developing allergic rhinitis up to 5 years of age.

At-risk population (1 study): The prospective cohort study (Van Asperen et al., 1984) did not show an association between the timing of introduction of CFs (≤ 4 months vs later) and the odds of developing allergic rhinitis up to around 1.5 years of age (Annex A as Microsoft Excel® file).

> The Panel notes for the general population that the results in the supportive line of evidence are consistent with those in the main line of evidence.
>
> The Panel notes for the at-risk population that the results in the supportive line of evidence are consistent with those in the main line of evidence.

8.6.2. Timing of introduction of cereals or fish

Only one study reported in two publications (Virtanen et al., 2010; Nwaru et al., 2013a) (supportive line of evidence) investigated the relationship between the timing of introduction of cereals or fish and

the odds of developing allergic rhinitis. Therefore, this study cannot be used to establish the appropriate age range of introduction of CFs.[34]

8.6.3. Allergic rhinitis: conclusions and grading of the confidence in the evidence

Imprecision: There was no imprecision associated with the results of the studies. However, as the single prospective cohort study in the main line of evidence was small (n = 62) and the second one did not provide a point estimate, the Panel was still concerned about the precision of the result. Therefore, the confidence in the evidence was downgraded by one category.

Inconsistency: The evidence is consistent across populations and the supportive line of evidence (five studies in the general population, one in an at-risk population) was consistent with the main line of evidence.

Generalisability: The study population of the RCT consisted only of breastfed infants. The Panel considers that the results of this study cannot be generalised to formula fed infants and thus to the whole population of infants living in Europe. Therefore, the Panel downgraded by one category the confidence in the evidence derived from the RCT. With respect to the prospective cohort studies, the Panel did not have concerns with respect to the generalisability of its findings, considering that the supportive line of evidence in which a number of populations were studied was consistent with the findings in the main line of evidence.

Publication bias: Publication bias could not be evaluated, because of the insufficient number of studies available.

The Panel concludes from the RCT (Tier 1) that there is no evidence for an effect of the introduction of CFs at 3-4 months of age compared with 6 months of age on the odds of developing allergic rhinitis up to 3 years of age (moderate confidence in the evidence).

The Panel concludes from the two prospective cohort studies (Tier 2) that there is no evidence for an association between the introduction of CFs ≤ 4 months of age compared with thereafter or compared with 4–6 months of age and the odds of developing allergic rhinitis up to 5 years of age (low confidence in the evidence).

8.7. Symptomatic food allergy: summary of the evidence

8.7.1. Timing of introduction of CFs in general

Main line of evidence

General population (3 studies): The RCT (Perkin et al., 2016) did not show an effect of the timing of introduction of CFs in general and the risk of symptomatic food allergy in exclusively breastfed infants at 1 or 3 years of age in the FAS (RR 0.80 (95% CI 0.51 to 1.25) (Appendix A.33). In the PP analysis, introduction of CFs at 3–4 months of age compared with an introduction at 6 months of age was associated with a statistically significant reduction in the risk of developing symptomatic food allergy (RR 0.33 (95% CI 0.13 to 0.83). However, this could be mainly attributed to the effect of the early introduction of egg and peanut on symptomatic egg and peanut allergy, respectively, in this study (discussed in a separate Section) and not to the timing of introduction of CFs *per se*.

The meta-analysis of the nested case–control study (Grimshaw et al., 2013) and the prospective cohort study (Luccioli et al., 2014) did not show a significant association between the timing of introduction of CFs (≤ 3 and ≤ 4 months compared with thereafter) and the outcome assessed up to 6 years of age. The pooled estimate obtained from the meta-analysis was associated with a serious imprecision. Heterogeneity was considerable (I^2 = 78%) (Appendix A.32). However, the Panel considers that this could be explained by the different methods for assessing symptomatic food allergy (i.e. double-blind placebo-controlled food challenge vs parents' report of a doctor's diagnosis) and the different age at outcome assessment (i.e. 1 vs 6 years). Individually, one study (Grimshaw et al., 2013) found higher odds of developing symptomatic

[34] Only studies that investigated the timing of introduction of fish at at-least one time point before 6 months of age were pertinent for the present assessment, in line with the interpretation of the Terms of Reference, as explained previously (Section 2.1.1.1.). This led to the exclusion of studies related to the timing of introduction of fish that had been considered by other bodies (e.g. SACN) in their assessment undertaken in a different regulatory context from this opinion. Therefore, the assessment by the Panel is not necessarily comparable to assessments that have been performed by other bodies in this respect.

6 years). Individually, one study (Grimshaw et al., 2013) found higher odds of developing symptomatic food allergy in infants introduced to CFs earlier (aOR 4.08 (95% CI 1.47 to 11.34)), but the other larger study (Luccioli et al., 2014) did not find a statistically significant association.

At-risk population (2 studies): The prospective cohort study that investigated the outcome in the general population, also investigated it in an at-risk population (Luccioli et al., 2014). The study did not observe an association between the introduction of CF ≤ 3 months compared with thereafter and the change for developing symptomatic food allergy, assessed at 6 years of age. In addition, Sandini et al. (2011) did not find an association between CF introduction < 4 months of age compared with 4–6 months and the odds of developing symptomatic food allergy up to 2 or 5 years of age, but did not provide a point estimate. (Annex A as Microsoft Excel® file).

Supportive line of evidence

> The Panel notes for the general population, from one RCT and two prospective cohort studies (Tiers 1 and 2) in the main line of evidence, that there is no evidence for an effect or association between the timing of introduction of CFs and the odds of developing symptomatic food allergy up to 6 years of age.
>
> The Panel notes for the at-risk population, from the two prospective cohort studies (Tier 2) in the main line of evidence, that there is no evidence for an association between the timing of introduction of CFs and the odds of developing symptomatic food allergy up to 6 years of age.

The reasoning behind the use of the data comprised in the supportive line of evidence is explained in Section 2.2.3.3.

- **Prospective cohort studies (2 studies, Tier 3)**

General population (2 studies): The meta-analysis of two prospective cohort studies (Venter et al., 2009; Kim et al., 2011) that compared the introduction of CFs < 4 and < 6 months of age with thereafter did not show a statistically significant association between the timing of introduction of CFs and the odds of developing symptomatic food allergy investigated up to 3 years of age (Appendix A.33). The pooled estimate obtained from the meta-analysis was associated with serious imprecision. Heterogeneity was not important (I^2 = 0%). Individually, Venter et al. (2009) showed significantly lower odds of symptomatic food allergy to be associated with the introduction of CFs < 4 months of age (OR 0.51 (95% CI 0.28 to 0.92)). However, this analysis was unadjusted and therefore is likely to overestimate the association. Kim et al. (2011) did not find an association between the introduction of CFs < 6 months of age compared with thereafter on the odds of developing symptomatic food allergy up to 1 year of age.

- **Retrospective studies (2 studies, Tier 3)**

General population (1 study): The case–control study by Bascunan Gamboa et al. (2012) did not observe an association between the introduction of CFs < 6 months of age compared with later and the odds of developing symptomatic food allergy up to 6–24 months of age (Annex A as Microsoft Excel® file).

At-risk populations (1 study): One cross-sectional analysis of baseline data of a prospective cohort study (Koplin et al., 2010) did not find an association between the timing of introduction of CFs (< vs ≥ 4 months) and the odds of developing symptomatic egg allergy up to 1 year of age (Annex A as Microsoft Excel® file).

- **Studies in which the timing of introduction of CFs was used as a continuous variable in the analysis, irrespective of the study design (1 study, Tier 3)**

General population (1 study): The case–control study by Alkazemi et al. (2018) did not observe an association between the timing of introduction of CFs and the odds of developing symptomatic food allergy assessed up to 13 years of age (Annex A as Microsoft Excel® file).

- **Sensitisation to food allergens (10 studies)**

General population (6 studies): The results of the RCT (Perkin et al., 2016) (Tier 1) with respect to sensitisation to food allergens is consistent with the results on symptomatic food allergy. The meta-analysis of the four prospective cohort studies rated as Tiers 1 and 2 (Snijders et al., 2008; Zutavern et al., 2008; Joseph et al., 2011; Nwaru et al., 2013c) did not show statistically significant results (Appendix A.34). Heterogeneity was substantial to considerable (I^2 = 79%) and cannot be

explained. The results of the single prospective cohort study rated as Tier 3 (Venter et al., 2009) showed lower odds of sensitisation at 3 years to be associated with an introduction of CF < 4 months of age compared with thereafter (OR 0.33 (95% CI 0.11 to 0.86)). This analysis was unadjusted and therefore is likely to overestimate the association. Finally, with respect to the difference in the timing of introduction of CFs in cases sensitised to food allergens and controls, four studies that investigated this outcome did not find statistically significant differences (Kucukosmanoglu et al., 2008; Hesselmar et al., 2010; McGowan et al., 2015; Hua et al., 2017) (two Tier 1 and two Tier 3) (Annex A as Microsoft Excel® file).

At-risk populations (4 studies): The meta-analysis of four comparisons from three prospective cohort studies (Mihrshahi et al., 2007; Joseph et al., 2011; Nwaru et al., 2013c), that investigated the introduction of CFs before 3 or 4 months of age compared with thereafter, did not show a statistically significant association between the timing of introduction of CFs and sensitisation to food allergens. Heterogeneity was moderate to substantial (I^2 = 59%) (Appendix A.35). Equally, the case–control study by Sicherer et al. (2010) (Tier 3) in which the timing of introduction of CFs was used as a continuous variable in the analysis did not find an association between the timing of introduction of CFs and sensitisation to peanut protein at 3–15 months of age (Annex A as Microsoft Excel® file).

The Panel considers that, given that symptomatic food allergy was not investigated as an outcome in the prospective cohort studies, the findings of these studies with respect to sensitisation are difficult to interpret.

- **Difference in the timing of introduction of CFs in cases and controls**

General population (3 studies): One nested case–control study (McGowan et al., 2015) (Tier 2) reported that 5-year old cases with symptomatic food allergy were introduced to CFs statistically significantly earlier (median ages of introduction in cases and controls: 18 weeks vs 20 weeks, p = 0.04). The Panel notes that the difference in the timing of introduction of CFs between cases and controls is small and is unlikely to represent a true relationship between the timing of introduction of CFs and symptomatic food allergy. In addition, the analysis was unadjusted. Another nested case–control study (Hesselmar et al., 2010) (Tier 1) did not observe statistically significant differences in the timing of introduction of CFs between 1.5-year-old cases and controls. Also, in the case–control study by DesRoches et al. (2010) (Tier 3), the timing of introduction of CFs in 18-month-old peanut allergy cases compared with controls was not statistically significant (Annex A as Microsoft Excel® file).

At-risk populations (1 study): In a nested case–control study (McGowan et al., 2015) (Tier 2) in which cases with symptomatic food allergy and controls were selected from a population with heredity of atopic diseases, no statistically significant differences were observed in the timing of introduction of CFs between cases with symptomatic food allergy and controls (Annex A as Microsoft Excel® file).

> The Panel notes for the general population that the results of three out of four studies on symptomatic food allergy in the supportive line of evidence are consistent with the findings in the main line of evidence, as are the results related to the difference in the timing of introduction of CFs between cases and controls (three studies) and the results on sensitisation of the RCT. The results of the prospective cohort studies with respect to sensitisation cannot be interpreted in the absence of results on symptomatic food allergy in these studies.
>
> The Panel notes for the at-risk population that the results in the supportive line of evidence are consistent with those in the main line of evidence.

8.7.2. Timing of introduction of CFs in general and symptomatic food allergy: conclusions and grading of the confidence in the evidence

Imprecision: The imprecision associated with the results of the meta-analysis of the two observational studies on symptomatic food allergy in the general population was serious. Therefore, the Panel downgraded by one category the evidence.

Inconsistency: The evidence is consistent across populations and the results of the supportive line of evidence are overall consistent with those of the main line of evidence. For the meta-analysis conducted in the main line of evidence, I^2 was below 75%.

Generalisability: The study population of the RCT consisted only of breastfed infants. The Panel considers that the results of this study cannot be generalised to formula fed infants and thus to the whole population of infants living in Europe. Therefore, the Panel downgraded by one category the

confidence in the evidence derived from the RCT. There was no concern with respect to generalisability for the prospective cohort studies.

Publication bias: Publication bias could not be evaluated, because of the insufficient number of studies available.

> The Panel concludes from the RCT (Tier 1) that there is no evidence for an effect of introduction of CFs at 3–4 months of age compared with 6 months of age on the odds of developing symptomatic food allergy up to 3 years of age (moderate confidence in the evidence).
>
> The Panel concludes from the two prospective cohort studies (Tiers 1 and 2) that there is no evidence for an association between the age of introduction of CFs ≤ 3 and ≤ 4 months vs thereafter and the odds of developing symptomatic food allergy up to 6 years of age (low confidence in the evidence).

8.7.3. Timing of introduction of egg

Main line of evidence

General population (3 studies)

Two RCTs (Perkin et al., 2016; Bellach et al., 2017) and one prospective cohort study (Tham et al., 2017) were available in this line of evidence. The use of different comparators in the two RCTs precluded pooling of the results (Annex A as Microsoft Excel® file).

In the RCT by Perkin et al. (2016) (Tier 1) in exclusively breastfed infants conducted in the UK, introduction of egg at 3–4 months (intervention) was compared with continued exclusive breastfeeding and an introduction of egg at 6 months of age (control), in relation to the prevalence of egg allergy at 1 or 3 years of age as diagnosed by a double-blind placebo-controlled food challenge. In this study, infants were recruited from the general population. According to the protocol, egg was to be administered as boiled hen's egg in an amount of 2×2 g of egg protein per week, equivalent to around 30 g of egg without shell (equivalent to 1 (very) small egg).

Compliance with the protocol in the intervention group was defined as consumption of at least 3 g allergen protein/week for at least five weeks between 3 and 6 months of age. There was a considerable number of infants (56.9%) who did not reach the minimum targeted amount of consumption of cooked hen's egg, hence were excluded from the PP analysis in the intervention group. Adherence to the protocol in the control group was much higher and only 7.9% of the infants were excluded from the PP analysis. The Panel notes that the differential adherence rates to the protocol and subsequent exclusions from analysis in the intervention and control groups might have led to the violation of the principle of randomisation, and thus to a potentially biased result, in the PP analysis.

In the PP population, a statistically significantly lower risk of developing symptomatic egg allergy was found at 1 or 3 years of age in the intervention group compared with the control group (RR 0.25 (95% CI; 0.08 to 0.82)). This did not reach statistical significance in the FAS (RR 0.69 (95% CI; 0.40 to 1.18)). One possible explanation is that the significant results in the PP analysis were due to reverse causality, i.e. those infants who did not consume egg had developed or were developing egg allergy. It is also possible that those infants who did not consume egg in sufficient amounts were unable to handle the texture and their non-adherence was unrelated to the outcome. In this case, their inclusion in the FAS analysis would have diluted the overall findings. The authors of the study addressed this question by comparing the prevalence of egg allergy of the non-compliant infants in the intervention group with infants in the control group and did not find statistically significant differences between these two groups (6.0% vs 5.5%; $p = 0.79$). This increases the confidence that the findings in the PP analysis were not due to the exclusion of children with egg allergy who could not consume the food. However, overall the Panel considers that the confidence in the finding is reduced by the inconsistent results in the PP and FAS analyses of this RCT.

The Panel also notes that there was limited evidence for an inverse dose-response relationship when considering the amount of egg that was consumed by infants, and that the introduction of cooked egg at home did not result in any cases of anaphylaxis.

In the RCT by Bellach et al. (2017) conducted in Germany (Tier 1), the timing of introduction of egg between 4 and 6 months with regular consumption up to 12 months of age (intervention) was compared with egg avoidance (control), in relation to the prevalence of egg allergy at 12 months of age diagnosed by a double-blind placebo-controlled food challenge or an open food challenge. In this study, infants were recruited from the general population. Only infants who were not sensitised to

hen's egg (i.e. who had specific immunoglobulin E (sIgE) concentrations < 0.35 kU_A/L) were included. Egg was administered as a pasteurised raw egg white powder mixed with solid infant foods in an amount of 3 × 2.5 g of egg protein per week, equivalent to around 58 g of egg without shell (equivalent to 1 large egg).

The originally planned sample size of the trial to have 80% power to detect a 50% reduction (from 12% to 6%) in sensitisation to egg was 788 infants. Recruitment was stopped early when 383 infants had been included in the trial. The authors report that this decision was taken based on three reasons: 1) following an interim analysis performed by an independent statistician, 2) the high level of egg sensitisation and allergy in the infants screened for inclusion (5.3% were diagnosed with egg allergy of which 2/3 reacted with an anaphylactic reaction during challenge) and 3) the frequency of allergic symptoms that occurred during the course of the trial (7.1% (n = 13/184; 3 of which had egg allergy) in the intervention group and 0.5% (n = 1/199) in the placebo group). Results were presented in the FAS population. The PP analysis was not considered by the Panel as it excluded all infants who became allergic to egg during the intervention. This analysis was therefore not an informative analysis. No statistically significant differences between the intervention and the control groups were observed. However, as recruitment stopped early, the study was underpowered to detect a statistically significant effect. The point estimate in the FAS analysis indicated a higher risk to be associated with early introduction (RR 3.3 (95% CI; 0.35 to 31.32)).

The Panel notes that this study was designed to investigate the effect of egg introduction at 4–6 months of age compared with egg avoidance in infants who were not sensitised to egg at baseline. Therefore, it is not comparable to the other available evidence and the seemingly inconsistent findings may be explained by these factors. In addition, it was underpowered.

Allergic reactions to the study powder were reported in 7.1% of the intervention and in 0.5% of the control group. Two of the three children who were diagnosed with egg allergy in the intervention group had an anaphylactic reaction to the study powder at home.

In the population-based birth cohort study (Tham et al., 2017) (Tier 2) conducted in Singapore on 1,152 singleton infants of Chinese, Malay or Indian ethnicity, symptomatic food allergy was defined as a convincing history of IgE-mediated reaction to a food. Data were available at 12 months of age from 854 infants, at 18 months from 799 children and at 2 years from 796 children. The prevalence of egg allergy in the study population was 1.7% at 12 months, 1.1% at 18 months and 0.8% at 24 months. Only a few infants were introduced to egg before 6 months of age, i.e. 21, 19 and 19 of those assessed at the different time points mentioned above and none developed egg allergy. No OR could be calculated owing to the zero events in the group introduced to egg before 6 months of age. When EFSA used the Fisher's exact test on the data reported above, there was no statistically significant difference. However, considering the low event rate, the study was most likely underpowered to detect statistically significant differences. Therefore, the Panel notes that the non-statistically significant findings of this study may not be reliable.

At-risk populations (3 studies)

Three RCTs from two different research groups (Tier 1 (Palmer et al., 2013; Palmer et al., 2017), Tier 2 (Tan et al., 2017)) are included in this line of evidence.

The study populations consisted of infants with moderate-to-severe eczema (SCORAD ≥ 15) (Palmer et al., 2013), infants with atopic mothers (Palmer et al., 2017) and non-sensitised infants (skin prick test (SPT) wheal < 2 mm) with at least one first-degree relative with an atopic disease (Tan et al., 2017). In all trials, the introduction of egg between around 4–6 months was compared to an introduction of around 8–10 months. Foods other than egg were self-selected. Egg was administered as pasteurised whole egg powder (raw in Palmer et al. (2013) and Palmer et al. (2017); unspecified in Tan et al. (2017)). The amount was equivalent to around 48 g of egg without shell per week (equivalent to 1 medium egg per week (daily consumption of 0.9 g of egg protein in Palmer et al. (2013)) and to around 19–22 g of egg without shell per week (equivalent to half a medium egg) in Palmer et al. (2017) and Tan et al. (2017) (daily consumption of 350–400 mg of egg protein).

In both trials by Palmer et al. (2013) and Palmer et al. (2017), recruitment had to be stopped early because of funding constraints and, therefore, they were individually not sufficiently powered to detect an effect. The study by Tan et al. (2017) was powered for sensitisation. Symptomatic food allergy was a secondary outcome.

In the meta-analysis of all three trials (Appendix A.36), a statistically significantly lower risk of developing symptomatic egg allergy at 1 year of age was observed when comparing introduction of egg between 4 and 6 months of age with after 8–10 months of age (RR 0.69 (95% CI 0.51 to 0.93)).

Heterogeneity was not important (I^2 = 0%). The 95% prediction interval crossed the line of 'null' effect. However, the Panel notes the uncertainty around the estimation of a prediction interval when only three studies are available (Section 2.2.3.2).

Palmer et al. (2013) reported that 31% (n = 15/49) of the infants had a reaction to the egg powder used in the study, ten of those reacted at the first exposure, including one case of anaphylaxis. Palmer et al. (2017) reported that 6.1% (n = 25/407) had a confirmed allergic reaction to the egg powder used in the study with no case of anaphylaxis. In Tan et al. (2017), 4.4% of the infants had mild to moderate reactions to egg within one week of starting the intervention with no case of anaphylaxis. Palmer et al. (2017) also reported that 92% (n=60/65) of the infants who had a reaction to the pasteurised raw egg challenge tolerated baked or cooked egg in the diet.

The Panel notes for the general population that, in the main line of evidence, there is limited evidence from one RCT (Tier 1) conducted in Europe that the introduction of cooked egg at 3–4 months of age compared with 6 months may reduce the risk of symptomatic egg allergy at 3 years of age.

The Panel notes for the at-risk population that, in the main line of evidence, there is evidence from three RCTs (Tiers 1 and 2) that egg introduction between 4 and 6 months of age may be associated with a lower risk of developing symptomatic egg allergy at 1 year of age.

Supportive line of evidence

The reasoning behind the use of the data comprised in the supportive line of evidence is explained in Section 2.2.3.3.

- **Retrospective studies (1 study, Tier 3)**

General population (1 study): In a cross-sectional analysis of baseline data of a prospective cohort study performed in Australia (Koplin et al., 2012) diagnosis of egg allergy was based on an open food challenge. Children were also considered to be egg allergic and were not offered a food challenge if parents reported a definite reaction to egg, the children had a positive SPT and egg was avoided in the infants' diet. Children who tolerated one raw egg during the food challenge were given single servings of one whole raw egg for 7 days to exclude egg allergy. There was no statistically significant difference in the odds of egg allergy at 11–15 months between infants that were introduced to egg between 4–6 months of age and those introduced between 7–9 months of age as well as those introduced between 10–12 months of age. However, infants introduced to egg between 4-6 months of age had statistically significantly lower odds of egg allergy than those introduced to egg after 12 months of age (including those infants who had not yet been exposed to egg): aOR 0.23 (95% CI 0.15 to 0.35). The p-for-trend was statistically significant.

At-risk population (1 study): For the same study described above also results in an at-risk population (i.e. children with a SPT wheal size ≥ 1 mm) were available (Koplin et al., 2010). There was no statistically significant difference in the odds of egg allergy at 11–15 months of age between infants that were introduced to egg between 4–6 months of age and those introduced between 7–9 months of age (aOR: 0.77 (95% CI 0.48 to 1.25)). There was also no statistically significant difference in the 4- to 6-month-group compared with the 10- to 12-month group (aOR 0.63 (95% CI 0.38 to 1.00)). However, infants introduced to egg at 4–6 months compared with those introduced to egg after 12 months of age (including those infants who had not yet been exposed to egg) had statistically significantly lower odds of symptomatic food allergy (aOR 0.29 (95% CI 0.15 to 0.56)). The analysis was adjusted for allergic symptoms occurring before the introduction of egg. The p-for-trend was statistically significant (Annex A as Microsoft Excel® file).

- **Sensitisation to food allergens (7 studies)**

General population (4 studies): Four studies (Gabet et al., 2016; Perkin et al., 2016; Bellach et al., 2017; Tran et al., 2017) (Tiers 1 and 2), including two RCTs, investigated sensitisation to egg protein, except (Gabet et al., 2016) that investigated sensitisation to egg, cow's milk, wheat, fish, peanut, sesame, mustard, soy, shrimp, beef and kiwi together (Appendix A.37). The results for sensitisation of Perkin et al. (2016) and Bellach et al. (2017) are consistent with the findings on symptomatic food allergy. In the study by Perkin et al. (2016) again the PP analysis showed a statistically significantly lower risk of developing sensitisation in the intervention group compared with controls, while this was not the case for the FAS analysis. The two prospective cohort studies (Gabet

et al., 2016; Tran et al., 2017) (Tiers 1 and 2) did not find an association between the timing of introduction of egg (≤ 6 months vs thereafter) and sensitisation assessed up to 18 months of age. However, the Panel notes that their results are difficult to interpret in the absence of results on symptomatic food allergy in the same studies.

At-risk populations (3 studies): The result of the meta-analysis of the three RCTs (Palmer et al., 2013; Palmer et al., 2017; Tan et al., 2017) (Tiers 1 and 2), described in the main line of evidence (symptomatic food allergy) and that also investigated sensitisation, was not statistically significant (Appendix A.38). However, the study that was powered to detect an effect on sensitisation (Tan et al., 2017), showed statistically significantly reduced odds of sensitisation in the group that was introduced to egg at 4 months of age compared with the group introduced at > 8 months (OR 0.46 (95% CI 0.22 to 0.95)). The Panel notes that the other two studies (Palmer et al., 2013; Palmer et al., 2017) that were included in the meta-analysis were individually underpowered to detect significant findings. In addition, the higher uncertainty around the heterogeneity estimate led to a wider 95% CI than in the meta-analysis on symptomatic food allergy. Therefore, the Panel considers that it is difficult to interpret whether or not the findings of these studies in relation to sensitisation are consistent with their results on symptomatic food allergy.

The Panel notes for the general population that the results of the retrospective study are consistent with the findings in the main line of evidence. Results of the studies investigating sensitisation are consistent within the two RCTs that investigated both symptomatic egg allergy and sensitisation, and cannot be interpreted for the two prospective cohort studies in the absence of results on symptomatic food allergy in these studies.

The Panel notes for the at-risk population that the results of the retrospective study are consistent with the findings in the main line of evidence. The results of the studies investigating sensitisation cannot be interpreted.

8.7.4. Timing of introduction of egg and symptomatic food allergy: conclusions and grading of the confidence in the evidence

For the grading in the confidence of the evidence, the main and supportive lines of evidence were further subdivided into:

- Main-A: the RCT by Perkin et al. (2016) (Tier 1) conducted in the general population comparing egg introduction (cooked egg) at 3–4 months of age to continued exclusive breastfeeding and egg introduction at 6 months of age.
- Main-B: the RCT by Bellach et al. (2017) (Tier 1) conducted in the general population comparing egg introduction (pasteurised raw egg white powder) at 4–6 months of age with egg avoidance in infants not sensitised to egg at baseline.
- Main-C: the prospective cohort study by Tham et al. (2017) (Tier 2) in the general population comparing egg < 6 months of age with thereafter.
- Main-D: the three RCTs conducted in high-risk populations in Australia (Palmer et al., 2013; Palmer et al., 2017; Tan et al., 2017) comparing egg introduction (pasteurised (raw) egg powder) at 4–6 months with > 8–10 months.
- S-A and B: the cross-sectional analysis of baseline data of the HealthNuts study performed in Australia (Koplin et al., 2010; Koplin et al., 2012).

The results of the evaluation of inconsistency, generalisability, imprecision, magnitude of the effect, dose-response and 'other' are summarised in Table 7.

Publication bias: Publication bias could not be evaluated, because of the insufficient number of studies available.

Safety: In the studies, there were some anaphylactic reactions associated with the consumption of pasteurised raw egg powders as intervention products. In the trial in which cooked egg was given to infants, no such reactions were observed. The Panel considers that, as far as the odds/risk of allergy is concerned, cooked egg can be introduced into the diet of infants when other CFs are introduced.

The Panel concludes from four RCTs (Tiers 1 and 2) that introduction of egg at 3–4 months of age compared with 6 months of age may reduce the risk of developing egg allergy (low to moderate level of confidence).

Table 7: Grading of the confidence in the evidence for symptomatic food allergy and timing of introduction of egg[(f)]

Line of evidence and initial rating	Certainty assessment									Characteristics					No of subjects		Effect in the LoE	Certainty in the LoE	Certainty across LoE
	No studies	Design	RoB	Inconsistency	Generalisability	Imprecision	Magnitude	Dose response	Other	Comparator, population	Early CF	Late CF	Age	Analysis	Early	Late	Relative		
1 (main-A) **++++**	1	RCT	↔	o	↓	↓	↑	↑	↓[(a)]	EBF, EU, gen pop	3–4 m	6 m	1 or 3 y	FAS	21/569 (3.7%)	32/596 (5.4%)	RR **0.69** (0.40–1.18)	**+++**	**++** to **+++**[(g)]
														PP	3/214 (1.4%)	29/525 (5.5%)	RR **0.25** (0.08–0.82)		
2 (main-B) **++++**	1	RCT	↔	o	↓↓	↓↓	↑	o	↔	CF, EU, gen pop	4–6 m	No egg	1 y	FAS	3/142 (2.1%)	1/156 (0.6%)	RR **3.3** (0.35–31.3)	**+**[(d)]	
3 (main-C) **+++**	1	PC	↔	o	↓↓	↓	↔	o	↔	CF, SG, gen pop	< 6 m	> 6 m	2 y	n/a	0/19 (0%)	6/777 (0.8%)	n/a	**+**[(e)]	
4 (main-D) **++++**	3	RCT	↔	↔	↓↓	↔	↔	↔	↔	CF, AU, at-risk	4–6 m	> 8–10 m	1 y	FAS	48/542 (8.9%)	70/536 (13.1%)	RR **0.69** (0.51–0.93)	**++**	
4 (S-A) **+**	1	CS	↔	o	↓	↔	↔	o	↑[(b)]	CF, AU, at-risk	4–6 m	7–12 m	1 y	n/a	27/485 (5.6%)	147/1663 (8.8%)	OR **0.61**[(c)] (0.40-0.93)	**+**	**+**
4 (S-B) **+**	1	CS	↔	o	↓	↔	↔	o	↑[(b)]	CF, AU, gen pop	4–6 m	7–12 m	1 y	n/a	n/a	n/a	n/a	**+**	

AU: Australia; CS: cross-sectional; CF: complementary food; EBF: exclusively breastfed; FAS: full analysis set; gen pop: general population LoE: line of evidence; m: months; n/a: not applicable; OR: odds ratio; PP: per-protocol analysis; RCT: randomised controlled trial; RR: risk ratio; SG: Singapore; y: year(s); ↓: downgrade; ↑: upgrade; ↔: no concern/impact; o: not evaluable.

(a): Because of inconsistency between FAS and PP analysis.

(b): Significant p-for-trend across different age categories of introduction of egg.

(c): Unadjusted (calculated based on the raw data of events per group).

(d): This line of evidence is not considered to be inconsistent with the findings in the lines of evidence 1 and 3, as differential findings could be explained. In addition, the study was underpowered. Therefore, this line of evidence was not considered in the grading of the confidence of the evidence.

(e): The study in this line of evidence was most likely underpowered and therefore this line of evidence was not considered in the grading of the confidence of the evidence.

(f): Results for sensitisation were consistent with the results on symptomatic food allergy within each of the RCTs in the general population and could not be interpreted for two additional studies (Gabet et al., 2016; Tran et al., 2017) in the general population in the absence of results on symptomatic food allergy in these studies. Also, the results on sensitisation of the meta-analysis on the three RCTs considered in the line of evidence 4 in at-risk populations cannot be interpreted. Therefore, the studies on sensitisation were not used in the grading of the confidence in the evidence.

(g): Derived as a range from the certainty in the lines of evidence 4 (++) and 1 (+++).

8.7.5. Timing of introduction of cereals

Main line of evidence

General population (1 study): The RCT (Perkin et al., 2016) did not find a statistically significant effect of the timing of introduction of wheat on the risk of wheat allergy at 1 or 3 years of age comparing the introduction at 3-4 months of age with 6 months of age in exclusively breastfed infants (Annex A as Microsoft Excel® file).

No studies were available in at-risk populations.

> The Panel notes for the general population that, from the RCT (Tier 1) in the main line of evidence, there is no evidence for an effect of the timing of introduction of wheat in exclusively breastfed infants on the risk of developing wheat allergy up to 3 years of age.

Supportive line of evidence

The reasoning behind the use of the data comprised in the supportive line of evidence is explained in Section 2.2.3.3.

- **Prospective cohort studies (1 study, Tier 3)**

General population (1 study): The prospective cohort study (Poole et al., 2006) reported lower odds of wheat allergy at 4 years of age with introduction of wheat, rye, oats and barley ≤ 6 months of age compared with thereafter (aOR 0.26 (95% CI 0.08 to 0.85)) (Annex A as Microsoft Excel® file).

- **Retrospective studies (1 study, Tier 3)**

General population (1 study): The cross-sectional study (Kumar et al., 2010) found higher odds of wheat allergy at 0.2–21 years to be associated with introduction of wheat or rice before 6 months of age compared with thereafter (aOR 1.6 (95% CI 1.004 to 2.5)) (Annex A as Microsoft Excel® file).

- **Sensitisation to food allergens (2 studies)**

General population (2 studies): The RCT by Perkin et al. (2016) (Tier 1) reported a statistically significant lower risk of sensitisation to wheat protein to be associated with introduction of wheat at 3–4 months of age compared with 6 months of age at 1 year (RR 0.40 (95% CI 0.17 to 0.95)), but not at 3 years of age. The prospective cohort study by Nwaru et al. (2013c) (Tier 1) reported that introduction of wheat before 5 months of age compared with thereafter was not associated with sensitisation to wheat protein at 5 years of age (Annex A as Microsoft Excel® file). However, this finding in the last study is difficult to interpret in the absence of data on symptomatic food allergy in the same study.

At-risk populations (1 study): The same prospective cohort study (Nwaru et al., 2013c) (Tier 1) described above assessed sensitisation to wheat protein as an outcome in an at-risk population. It found lower odds of wheat protein sensitisation at 5 years of age to be associated with introduction of wheat < 5.1 months of age compared with > 6.6 months of age (aOR 0.76 (95% CI 0.58 to 0.99)) (Annex A as Microsoft Excel® file). However, this finding is difficult to interpret in the absence of data on symptomatic food allergy in the same study.

> The Panel notes for the general population that the results of the two studies on symptomatic wheat allergy in the supportive line of evidence are inconsistent. The Panel also notes that the results of the RCT with respect to sensitisation are inconsistent with those on symptomatic wheat allergy and that the results of the prospective cohort study in relation to sensitisation cannot be interpreted in the absence of data on symptomatic wheat allergy in the same study.

8.7.6. Timing of introduction of cereals and symptomatic food allergy: conclusions and grading of the confidence in the evidence

Imprecision: There were no concerns with respect to imprecision.

Inconsistency: Only one study was available in the main line of evidence. The findings of the two studies on symptomatic food allergy in the supportive line of evidence are inconsistent, as is the finding on sensitisation of the RCT that is not consistent with the results on symptomatic wheat allergy

of the same RCT. For the decision on the grading of the confidence in the evidence in relation to these inconsistent findings, see 'other'.

Generalisability: The study population of the RCT consisted of only breastfed infants. The Panel considers that the results of this study cannot be generalised to formula fed infants and thus to the whole population of infants living in Europe. Therefore, the Panel downgraded by one category the confidence in the evidence derived from this RCT.

Publication bias: Publication bias could not be evaluated, because of the insufficient number of studies available.

Other: The Panel downgraded by one category the evidence, because of the limited number of studies that were available in the main line of evidence that were not supported by the findings in the supportive line of evidence.

The Panel concludes from the RCT (Tier 1) that there is no evidence for an effect of introduction of cereals at 3–4 months of age compared with 6 months of age on the risk of developing wheat allergy, assessed up to 3 years of age (low confidence in the evidence).

8.7.7. Timing of introduction of fish

The timing of introduction of fish and symptomatic fish allergy was investigated in a single study (Perkin et al., 2016) (main line of evidence). Therefore, this study cannot be used to establish an appropriate age range of introduction of CFs. This means that the studies in the supportive line of evidence on sensitisation (Nwaru et al., 2013c; Gabet et al., 2016; Perkin et al., 2016) cannot not be used either.

8.7.8. Timing of introduction of peanut

Main line of evidence

General population (1 study): In the FAS analysis, Perkin et al. (2016) (Tier 1) did not observe an effect of introduction of peanut at 3–4 months of age compared with 6 months in exclusively breastfed infants on the risk of symptomatic peanut allergy, assessed up to 3 years of age (RR 0.49 (95% CI 0.20 to 1.19)) (Annex A as Microsoft Excel® file). In the PP analysis, it was observed that introduction of peanut at 3–4 months of age reduced the risk of developing symptomatic peanut allergy compared with an introduction at 6 months of age (0% vs 2.5%, p = 0.003; RR not calculable due to zero events in the early introduction group). There was limited evidence for an inverse dose-response relationship when taking into account the amount of peanut consumed.

At-risk populations (1 study): In the study by Du Toit et al. (2015), infants between 4 and 10 months with severe eczema or egg allergy or both were randomly assigned either to peanut consumption (that was started depending on the age of the infant at enrolment between 4 and 10 months of age) or to peanut avoidance up to 5 years of age. At 5 years of age, the early introduction group had statistically significantly reduced odds in developing peanut allergy (in the ITT analysis: OR 0.16 (95% CI 0.01 to 0.32; PP: 0.02 (0.002–0.12)). As such, this study did not meet the inclusion criteria set by the Panel for the systematic review, because the early introduction group covered a time span beyond the first six months of life.

However, in a letter of response to a publication by Greenhawt et al. (2017), Lawson et al. (2017) (Tier 2) (Annex A as Microsoft Excel® file) provided further data that were used by the Panel to evaluate whether the introduction of peanut before the age of 6 months of age was associated with the development of peanut allergy.

In infants who were introduced to peanut ≤ 6 months of age, the odds of developing peanut allergy up to the age of 5 years was significantly reduced compared with those who avoided peanut up to that age (OR 0.17 (95% CI 0.06 to 0.47)). When performing a comparison in the intervention arm of the trial between infants that were introduced to peanut ≤ 6 months and those introduced at 7–10 months of age, there was no statistically significant difference between these two groups in the odds of developing peanut allergy. However, it should be noted that both reanalyses were observational and not based on the original randomised group.

The Panel notes for the general population that, in the main line of evidence, there is limited evidence from one RCT (Tier 1) that the introduction of peanut at 3–4 months of age compared with introduction at 6 months of age may reduce the risk of developing peanut allergy.

The Panel notes for the at-risk population that in the main line of evidence there is limited evidence from one RCT that the introduction of peanut between 4 and 10 months or between 4 and 6 months compared with after 5 years reduces the risk of developing peanut allergy. However, this was not the case for introduction of peanut ≤ 6 months of age compared with 7–10 months.

Supportive line of evidence

- **Sensitisation to food allergens (1 study)**

The result for sensitisation in the study by Perkin et al. (2016) is consistent with the findings in relation to symptomatic peanut allergy (Annex A as Microsoft Excel® file).

8.7.9. Timing of introduction of peanut and symptomatic food allergy: conclusions

The Panel considers that there is evidence that the introduction of peanut between 4 and 10 months or between 4 and 6 months of age in at-risk infants compared with after 5 years reduces the risk of developing symptomatic peanut allergy. However, the evidence is insufficient to conclude whether a similar effect occurs when comparing infants introduced to peanut ≤ 6 months of age compared with > 6 months, but still within the first year of life, owing to the inconsistent evidence between the study in the general population and the study in an at-risk population. Therefore, no level of confidence was assigned.

8.8. Atopic diseases: conclusions

For egg, the Panel concludes that there is evidence that its introduction at 3–4 months of age compared with 6 months of age may reduce the risk of developing egg allergy (low to moderate confidence in the evidence). In the studies that investigated egg allergy, there were some anaphylactic reactions associated with the consumption of pasteurised raw egg powders as intervention products. In the trial in which cooked egg was given to infants, no such reactions were observed.

For peanut, there is evidence that the introduction of peanut between 4 and 10 months or between 4 and 6 months of age in at-risk infants compared with after 5 years reduces the risk of developing peanut allergy. However, the evidence is insufficient to conclude whether a similar effect occurs when comparing infants introduced to peanut ≤ 6 months of age with those introduced later within the first year of life (no level of confidence assigned).

For CFs in general, fish and cereals, there is no evidence for an association between the timing of their introduction and the odds for developing atopic diseases. The confidence in the evidence ranges from low to moderate, depending on the outcome, the food and the age range studied.

The Panel also concludes that, as far as the odds/risk of developing allergy is concerned, cooked egg, fish, peanut and cereals can be introduced to the diet of infants when other CFs are introduced; there is no need to postpone their introduction.

9. Assessment of the data on coeliac disease in individuals born at term or mixed populations

9.1. Coeliac disease: final body of evidence

The 17 publications that were considered in the assessment in individuals born at term or mixed populations are given in Appendix B.6.

These publications reported on results from 15 studies:

- 1 RCT (Tier 1);
- 7 prospective cohort studies, 2 nested case–control studies (one study was analysed both as a prospective cohort study and a nested case–control study; 6 rated as Tier 1, 1 rated as Tier 2 and 1 rated as Tier 3);
- 6 retrospective studies (all Tier 3).

In these studies, three different endpoints were investigated. All results of the studies are given in Annex A as Microsoft Excel® file. In addition, for the main endpoints, results are summarised in the forest plots in Appendices A.39–A.41 of this Scientific Opinion.

With respect to the interpretation of the age at introduction of CFs as reported in the following, please refer to Section 2.2.3.3.

9.2. Coeliac disease: endpoint and study selection

Studies were included if cases of coeliac disease were identified following the criteria established by the Guidelines of the European Society for Paediatric Gastroenterology, Hepatology, and Nutrition (ESPGHAN) (Husby et al., 2012) for the diagnosis of coeliac disease. In children that show symptoms indicative of coeliac disease and have high anti-tissue transglutaminase type 2 antibody (IgA-tTGA) titres (> 10 times the upper limit of normal), the diagnosis is based on the additional presence of an elevated titre of endomysial antibodies in a blood sample drawn at an occasion separate from the initial one and the presence of haplotypes in the human leucocyte antigen (HLA) region associated with the risk of developing coeliac disease (HLA-DQ2 or HLA-DQ8), with no need for a small bowel biopsy. Under all other circumstances, the diagnosis is confirmed by a small bowel biopsy.

Coeliac disease autoimmunity was defined in most of the studies as IgA-tTGA concentrations above a pre-defined cut-off in children not fulfilling the above-mentioned criteria for diagnosing coeliac disease based on IgA-tTGA concentrations. All studies were used in the assessment, irrespective of the cut-offs used. The upper limit of normal concentration of antibodies depends on the test kit that is used.

No distinction was made in the assessment between study populations at risk of disease or not, as there is a strong genetic predisposition to coeliac disease. Individuals having neither HLA-DQ2 nor HLA-DQ8 are unlikely to have or develop coeliac disease (Husby et al., 2012). Selecting the general population as study population rather than subjects with a positive HLA-DQ2 or HLA-DQ8 status will dilute the effect or association but will not lead to differential results. Therefore, results of all study populations were combined. Individuals that are positive for HLA-DQ2 or HLA-DQ8 (and therefore at risk of the disease) will only develop coeliac disease if they are exposed to gluten. The present assessment focussed on an evaluation of the risk of developing coeliac disease in infants introduced to gluten at different time points (with at least one time point < 6 months of age). Time to onset of the disease or the effect of complete gluten avoidance was not investigated.

The Panel decided to draw its conclusions from the disease-related endpoint, i.e. coeliac disease. Data on coeliac disease autoimmunity are used only as supportive evidence to the results from the studies on coeliac disease, as positive results are associated with a higher risk of developing coeliac disease but alone are not predictive of the disease (see above).

For this outcome, only studies were available that investigated the timing of introduction of gluten (at various ages between < 3 and < 6 months), but not of CFs in general.

9.3. Coeliac disease: summary of the evidence

Main line of evidence (5 studies)

The RCT by Vriezinga et al. (2014) (Appendix A.39) conducted in a population at risk of coeliac disease did not find a statistically significant effect of the introduction of gluten (200 mg/day of vital wheat gluten with lactose, equivalent to 100 mg of immunologically reactive gluten[35,36]) at 4 months vs 6 months of age, on the hazard of developing coeliac disease up to 3 years of age. The authors explained that the amount of gluten that was administered is sufficient to cause histologic lesions (i.e. villous atrophy) in patients with coeliac disease.

From the meta-analysis of four prospective cohort studies (Norris et al., 2005; Welander et al., 2010; Størdal et al., 2013; Andren Aronsson et al., 2015), there is no evidence for an association

[35] Vital gluten is a by-product of starch isolation obtained during wet milling, in which flour is separated into starch and proteins (including gluten).

[36] Translation of this amount into an amount of food is difficult owing to the different gluten content of flours, depending on the type of cereal, the variety and environmental factors that may influence the percentage of storage proteins of a grain (such as growing season or region) as well as the gluten that is added for technological reasons. For example, for durum wheat pasta, protein content usually varies between 10 and 15%, of which around 50% are gluten (Atwell and Finnie, 2016). Using these assumptions, 200 mg vital gluten are contained in 2–4 g of durum wheat pasta.

between the timing of introduction of gluten or gluten-containing foods and the hazard of coeliac disease studied up to 12 years of age (Appendix A.39). Heterogeneity was substantial (I^2 = 73%).

This heterogeneity was mainly caused by one study (Norris et al., 2005) which showed results that were substantially different from those of the other available studies, and this could not be explained. When Norris et al. (2005) was removed from the meta-analysis in a sensitivity analysis (data not shown), heterogeneity was reduced to 18% with no substantial change to the point estimate (HR 0.91 instead of 0.94) with a narrower 95% CI (still not statistically significant).

An unplanned subgroup analysis (data not shown) was performed to investigate whether introduction of gluten below 3 or 4 months of age compared with around 4–6 months of age would have a different effect than introduction around 4–6 months of age compared with thereafter. This was done following the conclusion of the Panel in the previous Scientific Opinion (EFSA NDA Panel, 2009) that 'introduction of gluten < 4 months might increase the risk of coeliac disease [...], whilst the introduction of gluten between 4 and 6 months while still breastfeeding might decrease the risk'.

- The meta-analysis comparing those who were introduced to gluten ≤ 3 or 4 months of age with those introduced around 4–6 months of age did not show an association between the timing of introduction of gluten and the outcome (HR 1.47 (95% CI 0.05 to 14.91)). Heterogeneity was substantial (I^2 = 83%) and imprecision around the point estimate was serious. When the study by Norris et al. (2005) was removed from the meta-analysis, heterogeneity reduced to 24%, the point estimate shifted to the other side of the line of the 'null' effect and the 95% CI was reduced (HR 0.85 (95% CI 0.37 to 1.96)).
- Equally, the meta-analysis comparing those who were introduced to gluten around 4–6 months of age with those introduced later did not show an association between the timing of introduction of gluten and the outcome (HR 0.85 (95% CI 0.45 to 1.62)). Heterogeneity was moderate to substantial (I^2 = 60%). When the study by Norris et al. (2005) was removed from the meta-analysis, heterogeneity reduced to 41%, without substantial effects on the point estimate and the 95% CI (HR 0.92 (95% CI 0.57 to 1.46)).

Two studies reported on breastfeeding at the time of gluten introduction before 6 months of age and the risk of developing coeliac disease. Based on data reported in Szajewska et al. (2015), the RCT by Vriezinga et al. (2014) did not find an effect of breastfeeding during gluten introduction at 4 months of age compared with 6 months of age (RR 1.31 (95% CI 0.77 to 2.23)) on the risk of developing coeliac disease (secondary observational analysis). Also, the prospective cohort study by Størdal et al. (2013) (Tier 1), including 45,156 infants in the analysis, did not observe an association between continued breastfeeding at the time of gluten introduction ≤ 6 months of age and the risk of developing coeliac disease.

The assessment of the effect of breastfeeding while introducing gluten over a wider age range (> 6 months) as investigated in the systematic review by Szajewska et al. (2015) and in the position paper by ESPGHAN (Szajewska et al., 2016) is not part of the current mandate and was not considered further.

> The Panel notes that, from the RCT and the meta-analysis of four prospective cohort studies (Tiers 1 and 2) in the main line of evidence, there is no evidence for an association between various timings of introduction of gluten or gluten-containing foods and the hazard of developing coeliac disease up to 12 years of age. There are also no differential effects of gluten introduction < 4 months of age and between 4 and 6 months of age, or gluten introduction < 6 months of age while still breastfeeding.

Supportive line of evidence

The reasoning behind the use of the data comprised in the supportive line of evidence is explained in Section 2.2.3.3.

- **Retrospective studies (4 studies, Tier 3)**

The meta-analysis of the four case–control studies (Auricchio et al., 1983; Greco et al., 1988; Peters et al., 2001; Ivarsson et al., 2002) did not find a statistically significant association between the various timings of introduction of gluten-containing foods (ranging from ≤ 2 to ≤ 4 months) and the odds of developing coeliac disease up to around 6 years of age (Appendix A.40). Heterogeneity was not important (I^2 = 26%).

- **Studies in which the timing of introduction of gluten was used as a continuous variable in the analysis, irrespective of the study design (2 studies)**

One prospective cohort study by Andren Aronsson et al. (2016) (Tier 1) and one case–control study (Myleus et al., 2012) (Tier 3) did not observe statistically significant associations between the timing of introduction of gluten-containing foods and the odds of developing coeliac disease up to 2 years of age (Annex A as Microsoft Excel® file).

- **Coeliac disease autoimmunity (5 studies)**

The findings of the RCT (Vriezinga et al., 2014) on coeliac disease autoimmunity were consistent with the findings on the disease endpoint (Appendix A.41).

From the meta-analysis of four prospective cohort studies (Norris et al., 2005; Jansen et al., 2014; Andren Aronsson et al., 2015; Chmiel et al., 2015) (Tiers 1 and 2), there is no evidence for an association between various timings of introduction of gluten-containing foods (ranging from < 3 months to ≤ 6 months compared with thereafter) and the hazard of developing coeliac disease autoimmunity up to around 9 years of age. Heterogeneity was moderate to substantial ($I^2 = 56\%$). Two of the studies also investigated coeliac disease and their findings on coeliac disease autoimmunity were consistent with those on the disease endpoint.

- **Difference in the timing of introduction of gluten in cases and controls (3 studies)**

Two nested case–control studies (Andren Aronsson et al., 2016; Savilahti et al., 2018) (Tiers 1 and 3, respectively) and one case–control study performed in cases with coeliac disease and their healthy siblings (Ascher et al., 1997) (Tier 3), did not find statistically significantly different timings of introduction of gluten or gluten-containing foods in coeliac disease cases and controls aged 2, 5 and 8 years, respectively (Annex A as Microsoft Excel® file).

> The Panel notes that the results in the supportive line of evidence are consistent with those in the main line of evidence.

9.4. Coeliac disease: conclusions and grading of the confidence in the evidence

Imprecision: There were no concerns with respect to imprecision.

Inconsistency: The results were consistent across populations and the results of the supportive line of evidence (six studies on coeliac disease, five on coeliac disease autoimmunity and three on the timing of introduction of gluten in cases and controls) were consistent with the results in the main line of evidence. For the meta-analysis conducted in the main line of evidence, I^2 was below 75%.

Generalisability: There were no concerns with respect to generalisability, as a variety of populations were studied.

Publication bias: Publication bias could not be evaluated, because of the insufficient number of studies available.

> The Panel concludes from the RCT (Tier 1) that there is no effect of the introduction of gluten at 4 months of age compared with 6 months of age and the hazard of developing coeliac disease up to 3 years of age (high confidence in the evidence).
>
> The Panel concludes from the four prospective cohort studies (Tiers 1 and 2) that there is no evidence for an association between the age of introduction of gluten ≤ 3 or 4 months of age compared with thereafter and the hazard of developing coeliac disease up to 12 years of age (moderate confidence in the evidence).
>
> In its previous Scientific Opinion, the Panel considered that introduction of gluten < 4 months of age might increase the risk of coeliac disease, whereas introduction between 4 and 6 months of age while still breastfeeding might decrease the risk of coeliac disease. With the data that have become available on coeliac disease since the publication of the last Scientific Opinion, these conclusions are no longer supported.