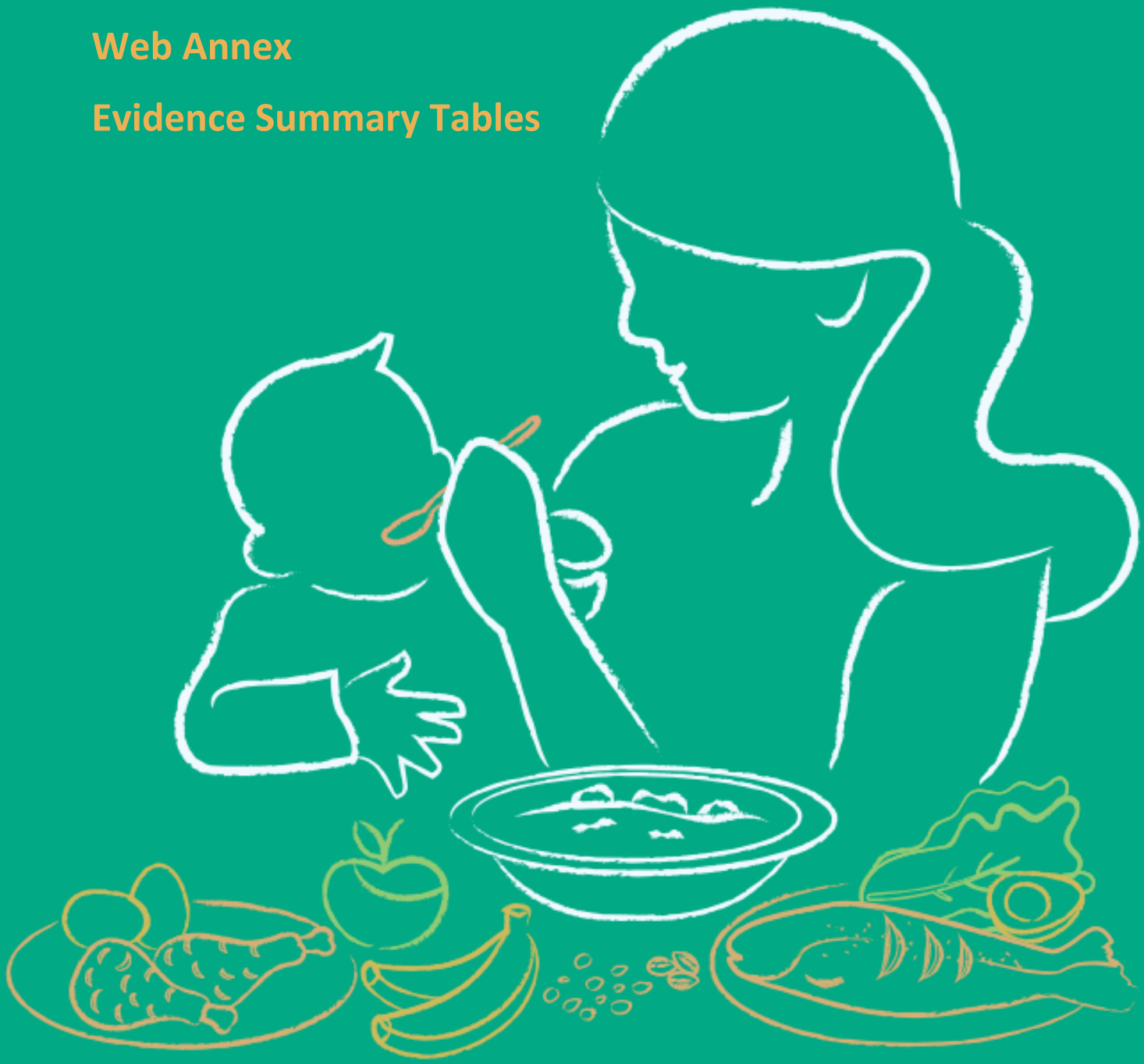
World Health Organization

WHO Guideline

for complementary feeding of infants and young children 6–23 months of age

Web Annex

Evidence Summary Tables



WHO Guideline for complementary feeding of infants and young **cihldren** 6-23 months of age

Web Annex

Evidence Summary Tables

World Health Organization

WHO Guideline for complementary feeding of infants and young children 6-23 months of age. Web Annex. Evidence Summary Tables

ISBN 978-92-4-008238-0 (electronic version)

Suggested citation. Web Annex. Evidence Summary Tables. In: WHO Guideline for complementary feeding of infants and young children 6-23 months of age. Geneva: World Health Organization; 2023. Licence: CC BY-NC-SA 3.0 IGO.

Cataloguing-in-Publication (CIP) data. CIP data are available at http://apps.who.int/iris.

Sales, rights and licensing. To purchase WHO publications, see https://www.who.int/publications/book-orders. To submit requests for commercial use and queries on rights and licensing, see https://www.who.int/copyright.

This publication forms part of the document entitled *WHO Guideline for complementary feeding of infants and young children 6-23 months of age*. It is being made publicly available for transparency purposes and information.

Contents

Continued breastfeeding

Certainty assessment							№ of patients	Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations		Relative (95% CI)	Absolute (95% CI)		
Child Development, Intelligence, Cognition and Behaviour											
Delayed Development											
2	Cohort Studies	serious[a]	not serious	very serious[b]	not serious	none	1872	1.15 [0.54, 2.43]		⊕◯◯◯ VERY LOW	CRITICAL
IQ											
6	Cohort Studies	very serious[c]	not serious	not serious	not serious	none	5079		MD 0.01 higher (0.06 lower to 0.08 higher)	⊕◯◯◯ VERY LOW	CRITICAL
Learning ability											
1	Cohort Studies	serious[a]	not serious	very serious[b]	not serious	none	413		MD 0.02 lower (0.27 lower to 0.23 higher)	⊕◯◯◯ VERY LOW	CRITICAL
Coding Wisc III											
1	Cohort Studies	serious[a]	not serious	very serious[b]	not serious	none	413		MD 0.14 higher (0.11 lower to 0.39 higher)	⊕◯◯◯ VERY LOW	CRITICAL
Visuospatial Problem Solving											
2	Cohort Studies	serious[a,d]	not serious	not serious	not serious	none	1403		**MD 0.17 lower (0.33 lower to 0.01 lower)**	⊕◯◯◯ VERY LOW	CRITICAL
Verbal Intelligence											
3	Cohort Studies	serious[a,d,e]	not serious	not serious	not serious	none	1936		MD 0.08 lower (0.18 lower to 0.01 higher)	⊕◯◯◯ VERY LOW	CRITICAL
Working Memory											
2	Cohort Studies	serious[a,e]	not serious	not serious	not serious	none	946		MD 0.05 lower (0.19 lower to 0.10 higher)	⊕◯◯◯ VERY LOW	CRITICAL
Reasoning											
3	Cohort Studies	serious[a,d,e]	not serious	not serious	not serious	none	1936		**MD 0.13 lower (0.22 lower to 0.03 lower)**	⊕◯◯◯ VERY LOW	CRITICAL
Processing Speed											
1	Cohort Studies	very serious[b,e]	not serious	not serious	not serious	none	533		MD 0.07 lower (0.24 lower to 0.11 higher)	⊕◯◯◯ VERY LOW	CRITICAL
Child Growth and anthropometry											
Underweight											
7	Cross Sectional	serious[f]	not serious	not serious	not serious	none	15441	**1.25 [1.08, 1.46]**		⊕◯◯◯ VERY LOW	CRITICAL
Stunting											
3	Cohort Studies	moderate[g]	not serious	serious[h]	not serious	none	1299	1.87 [0.95, 3.68]		⊕◯◯◯ VERY LOW	CRITICAL
Wasting											

2	Cohort Studies	serious[f]	not serious	not serious	not serious	none	342	**2.16 [1.18, 3.98]**		⊕◯◯◯ VERY LOW	CRITICAL
Weight [kg]											
1	Cohort Studies	serious[f]	not serious	serious[h]	not serious	none	191		MD 3.30 lower (7.39 lower to 0.79 higher)	⊕◯◯◯ VERY LOW	CRITICAL
WAZ											
1	Cohort Studies	moderate[g]	not serious	serious[h]	not serious	none	1799		**MD 0.25 higher (0.10 higher to 0.40 higher)**	⊕◯◯◯ VERY LOW	CRITICAL
Height [cm]											
2	Cohort Studies	serious[i,j]	not serious	not serious	not serious	none	600		MD 0.68 lower (2.46 lower to 1.11 higher)	⊕◯◯◯ VERY LOW	CRITICAL
HAZ											
2	Cross-sectional	moderate[j]	not serious	not serious	not serious	none	3050		MD 0.14 higher (0.07 lower to 0.36 higher)	⊕◯◯◯ VERY LOW	CRITICAL
WHZ											
2	Cohort Studies	moderate[j]	not serious	serious[h]	not serious	none	2148		MD 0.01 lower (0.33 lower to 0.31 higher)	⊕◯◯◯ VERY LOW	CRITICAL
BMI											
5	Cohort Studies	serious[f]	not serious	not serious	not serious	none	12356		**MD 0.10 lower (0.17 lower to 0.03 lower)**	⊕◯◯◯ VERY LOW	CRITICAL
BMIZ											
3	Cohort Studies	serious[f]	not serious	not serious	not serious	none	1750		MD 0.01 higher (0.14 lower to 0.16 higher)	⊕◯◯◯ VERY LOW	CRITICAL
Obesity and Overweight											
25	Cohort Studies	serious[f]	not serious	not serious	not serious	none	37715	0.94 [0.79, 1.12]		⊕◯◯◯ VERY LOW	CRITICAL
Waist Circumference (cm)											
1	Cross-sectional	moderate[j]	not serious	serious[h]	not serious	none	969		MD 0.07 higher (0.47 lower to 0.61 higher)	⊕◯◯◯ VERY LOW	CRITICAL
Skinfold thickness											
1	Cohort Studies	serious[f]	not serious	serious[h]	serious[k]	none	356		MD 0.10 lower (1.08 lower to 0.88 higher)	⊕◯◯◯ VERY LOW	CRITICAL
Child Morbidity											
Infections											
2	Cohort Studies	moderate[l]	not serious	not serious	not serious	none	46156	**1.08 [1.00, 1.16]**		⊕◯◯◯ VERY LOW	CRITICAL
GI Infections											
2	Cohort Studies	serious[m]	not serious	not serious	not serious	none	43018	1.00 [0.96, 1.04]		⊕◯◯◯ VERY LOW	CRITICAL
Acute Gastroenteritis											
1	Cohort Studies	serious[n]	not serious	serious[h]	serious[k]	none	270		**MD 2.23 lower (2.55 lower to 1.91 lower)**	⊕◯◯◯ VERY LOW	CRITICAL
Respiratory tract infections											
1	Cohort Studies	serious[n]	not serious	serious[h]	serious[k]	none	270		**MD 2.43 lower (3.99 lower to 0.87 lower)**	⊕◯◯◯ VERY LOW	CRITICAL
Acute Otitis Media											

1	Cohort Studies	serious[n]	not serious	serious[h]	serious[k]	none	270		**MD 1.40 lower (1.80 lower to 1.00 lower)**	⊕◯◯◯ VERY LOW	CRITICAL
Urinary System Infections											
1	Cohort Studies	serious[n]	not serious	serious[h]	serious[k]	none	270		MD 0.02 lower (0.25 lower to 0.21 higher)	⊕◯◯◯ VERY LOW	CRITICAL
Diabetes											
1	Cohort Studies	moderate[o]	not serious	not serious	not serious	none	2090	0.62 [0.29, 1.29]		⊕◯◯◯ VERY LOW	CRITICAL
Glucose (mg/dL)											
1	Cohort Studies	moderate[o]	not serious	serious[h]	not serious	none	969		MD 0.01 lower (0.10 lower to 0.08 higher)	⊕◯◯◯ VERY LOW	CRITICAL
Type 1 Diabetes											
1	Cohort Studies	moderate[o]	not serious	serious[h]	not serious	none	67803	0.98 [0.73, 1.32]		⊕◯◯◯ VERY LOW	CRITICAL
Systolic Sauce Pressure											
4	Cohort Studies	moderate[o]	not serious	serious[h]	not serious	none	4163		MD 0.65 lower (1.35 lower to 0.06 higher)	⊕◯◯◯ VERY LOW	CRITICAL
Diastolic Blood Pressure											
2	Cohort Studies	serious[n]	not serious	not serious	not serious	none	704		MD 1.60 lower (4.11 lower to 0.92 higher)	⊕◯◯◯ VERY LOW	CRITICAL
Prehypertension											
1	Cohort Studies	moderate[o]	not serious	not serious	not serious	none	2490		**MD 0.86 higher (0.69 higher to 1.07 higher)**	⊕◯◯◯ VERY LOW	CRITICAL
Carotid Intima Media Thickness											
2	Cohort Studies	serious[n]	not serious	serious[h]	serious[k]	none	806		MD 6.27 lower (16.54 lower to 4.01 higher)	⊕◯◯◯ VERY LOW	CRITICAL
Total Cholesterol											
2	Cohort Studies	moderate[o]	not serious	serious[h]	not serious	none	637		MD 2.14 higher (7.17 lower to 11.45 higher)	⊕◯◯◯ VERY LOW	CRITICAL
LDL											
2	Cohort Studies	moderate[o]	not serious	serious[h]	not serious	none	637		MD 2.38 higher (3.82 lower to 8.58 higher)	⊕◯◯◯ VERY LOW	CRITICAL
HDL											
2	Cohort Studies	moderate[o]	not serious	serious[h]	not serious	none	637		MD 0.02 higher (0.03 lower to 0.07 higher)	⊕◯◯◯ VERY LOW	CRITICAL
LDL/HDL ratio											
1	Cohort Studies	moderate[o]	not serious	serious[h]	not serious	none	409		MD 0.18 higher (0.04 lower to 0.40 higher)	⊕◯◯◯ VERY LOW	CRITICAL
Anaemia											
1	Cohort Studies	moderate[o]	not serious	serious[h]	not serious	none	139	**0.21 [0.05, 0.88]**		⊕◯◯◯ VERY LOW	CRITICAL
Iron Deficiency											
1	Cohort Studies	moderate[o]	not serious	serious[h]	not serious	none	171	**1.71 [1.05, 2.78]**		⊕◯◯◯ VERY LOW	CRITICAL
Child Bone Health, Oral Health and Allergies											
Bone Mineral Density											

2	Cohort Studies	serious[n]	not serious	not serious	not serious	none	985		MD 0.01 lower (0.01 lower to 0.00 lower)	⊕◯◯◯ VERY LOW	CRITICAL
Bone Mineral Content											
2	Cohort Studies	serious[n]	not serious	not serious	not serious	none	985		MD 12.27 lower (25.75 lower to 1.20 higher)	⊕◯◯◯ VERY LOW	CRITICAL
Dental caries											
20	Cohort Studies	moderate[o]	not serious	not serious	not serious	none	11286	**1.52 [1.24, 1.88]**		⊕◯◯◯ VERY LOW	CRITICAL
Asthma											
8	Cohort Studies	moderate[o]	not serious	serious[h]	not serious	none	42521	1.04 [0.95, 1.13]		⊕◯◯◯ VERY LOW	CRITICAL
Allergies											
2	Cohort Studies	moderate[o]	not serious	not serious	not serious	none	9824	0.88 [0.40, 1.95]		⊕◯◯◯ VERY LOW	CRITICAL
Eczema											
4	Cohort Studies	moderate[o]	not serious	not serious	not serious	none	13360	1.09 [0.94, 1.27]		⊕◯◯◯ VERY LOW	CRITICAL
Hay Fever											
1	Cohort Studies	moderate[o]	not serious	not serious	not serious	none	381	0.30 [0.07, 1.21]		⊕◯◯◯ VERY LOW	CRITICAL
Maternal Diabetes & CVD											
Diabetes											
5	Cohort Studies	moderate[o]	serious[p]	not serious	not serious	none	42987	0.93 [0.62, 1.39]		⊕◯◯◯ VERY LOW	CRITICAL
Blood Glucose (mg/dL)											
3	Cohort Studies	moderate[o]	not serious	not serious	not serious	none	5646		MD 0.35 higher (3.85 lower to 4.55 higher)	⊕◯◯◯ VERY LOW	CRITICAL
HbA1C											
1	Cohort Studies	moderate[o]	not serious	serious[h]	serious[q]	none	291		MD 0.01 lower (0.09 lower to 0.07 higher)	⊕◯◯◯ VERY LOW	CRITICAL
Fasting Insulin											
1	Cohort Studies	moderate[o]	not serious	serious[h]	serious[q]	none	89		MD 0.10 lower (2.03 lower to 1.83 higher)	⊕◯◯◯ VERY LOW	CRITICAL
Hypertension											
5	Cohort Studies	moderate[o]	not serious	not serious	not serious	none	262129	0.94 [0.83, 1.06]		⊕◯◯◯ VERY LOW	CRITICAL
Systolic Blood Pressure											
2	Cohort Studies	moderate[o]	not serious	serious[h]	not serious	none	5547		MD 2.57 higher (1.70 lower to 6.84 higher)	⊕◯◯◯ VERY LOW	CRITICAL
Diastolic Blood Pressure											
1	Cohort Studies	moderate[o]	not serious	serious[h]	not serious	none	5547		MD 0.46 higher (1.70 lower to 2.62 higher)	⊕◯◯◯ VERY LOW	CRITICAL
Stroke											
1	Cohort Studies	moderate[o]	not serious	serious[h]	not serious	none	239959	**1.02 [1.00, 1.05]**		⊕◯◯◯ VERY LOW	CRITICAL
Cardiovascular Mortality											

2	Cohort Studies	moderate[o]	not serious	not serious	not serious	none	272257	**0.97 [0.94, 1.00]**		⊕◯◯◯ VERY LOW	CRITICAL
All Cardiovascular Disease											
2	Cohort Studies	moderate[o]	not serious	not serious	not serious	none	227257	**0.97 [0.94, 1.00]**		⊕◯◯◯ VERY LOW	CRITICAL
Hyperlipidemia											
1	Cohort Studies	moderate[o]	not serious	serious[h]	not serious	none	3156	0.69 [0.39, 1.22]		⊕◯◯◯ VERY LOW	CRITICAL
Cholesterol (mG/dL)											
3	Cohort Studies	moderate[o]	not serious	not serious	not serious	none	5646		**MD 3.86 lower (6.26 lower to 1.45 lower)**	⊕◯◯◯ VERY LOW	CRITICAL
LDL											
3	Cohort Studies	moderate[o]	not serious	not serious	not serious	none	5646		**MD 2.67 lower (4.81 lower to 0.53 lower)**	⊕◯◯◯ VERY LOW	CRITICAL
HDL											
3	Cohort Studies	moderate[o]	not serious	not serious	not serious	none	5646		**MD 1.97 lower (2.78 lower to 1.15 lower)**	⊕◯◯◯ VERY LOW	CRITICAL
Low HDL											
1	Cohort Studies	moderate[o]	not serious	serious[h]	not serious	none	2443	0.75 [0.39, 1.44]		⊕◯◯◯ VERY LOW	CRITICAL
Triglyceride (mg/dL)											
2	Cohort Studies	moderate[o]	not serious	not serious	serious[q]	none	412		**MD 1.12 lower (1.30 lower to 0.94 lower)**	⊕◯◯◯ VERY LOW	CRITICAL
Metabolic Syndrome											
1	Cohort Studies	moderate[o]	not serious	serious[h]	not serious	none	2443	0.83 [0.45, 1.53]		⊕◯◯◯ VERY LOW	CRITICAL
Overweight and Obesity											
2	Cohort Studies	moderate[o]	not serious	not serious	not serious	none	5224	0.95 [0.73, 1.24]		⊕◯◯◯ VERY LOW	CRITICAL
BMI											
3	Cohort Studies	moderate[o]	not serious	not serious	not serious	none	5631		MD 0.01 higher (0.22 lower to 0.25 higher)	⊕◯◯◯ VERY LOW	CRITICAL
Weight Gain											
4	Cohort Studies	moderate[o]	not serious	not serious	not serious	none	2250		MD 0.21 higher (0.39 lower to 0.80 higher)	⊕◯◯◯ VERY LOW	CRITICAL
Waist Circumference											
3	Cohort Studies	moderate[o]	not serious	not serious	not serious	none	405		MD 0.41 higher (0.29 lower to 1.12 higher)	⊕◯◯◯ VERY LOW	CRITICAL
Maternal Cancer, Bone Health, Mental Health											
Breast Cancer											
4	Cohort Studies	moderate[o]	not serious	not serious	not serious	none	24695	0.83 [0.61, 1.14]		⊕◯◯◯ VERY LOW	CRITICAL
Ovarian Cancer											
1	Cohort Studies	serious[n]	not serious	serious[h]	not serious	none	5480	1.04 [0.46, 2.36]		⊕◯◯◯ VERY LOW	CRITICAL
Cervical Cancer											

1	Cohort Studies	serious[n]	not serious	serious[h]	not serious	none	5480	1.87 [0.93, 3.77]		⊕◯◯◯ VERY LOW	CRITICAL
Cancer Corpus Uteri											
1	Cohort Studies	serious[n]	not serious	serious[h]	not serious	none	5480	1.29 [0.52, 3.20]		⊕◯◯◯ VERY LOW	CRITICAL
All Cancers											
1	Cohort Studies	serious[n]	not serious	serious[h]	not serious	none	5480	0.84 [0.67, 1.06]		⊕◯◯◯ VERY LOW	CRITICAL
Cancer All Except Breast/Genitalia											
1	Cohort Studies	serious[n]	not serious	serious[h]	not serious	none	5480	0.79 [0.60, 1.05]		⊕◯◯◯ VERY LOW	CRITICAL
Osteoporosis											
2	Cohort Studies	serious[n]	not serious	serious[h]	serious[q]	none	1790	3.37 [0.30, 38.07]		⊕◯◯◯ VERY LOW	CRITICAL
Bone Mass Density Total Femur											
1	Cohort Studies	moderate[o]	not serious	serious[h]	not serious	none	1113		**MD 0.06 lower (0.07 lower to 0.06 lower)**	⊕◯◯◯ VERY LOW	CRITICAL
Bone Mass Density Femoral Neck [g/cm2]											
2	Cohort Studies	serious[n]	not serious	serious[h]	not serious	none	1918		**MD 0.06 lower (0.07 lower to 0.05 lower)**	⊕◯◯◯ VERY LOW	CRITICAL
Bone Mass Density Lumbar Vertebrae											
3	Cohort Studies	serious[n]	not serious	serious[h]	serious[q]	none	1948		**MD 0.10 lower (0.11 lower to 0.09 lower)**	⊕◯◯◯ VERY LOW	CRITICAL
Bone Mass Density Mid Radius											
1	Cohort Studies	moderate[o]	not serious	serious[h]	serious[q]	none	30		MD 0.00 higher (0.01 lower to 0.02 higher)	⊕◯◯◯ VERY LOW	CRITICAL
Bone Mass Density Distal Radius											
1	Cohort Studies	moderate[o]	not serious	serious[h]	serious[q]	none	30		MD 0.00 higher (0.01 lower to 0.01 higher)	⊕◯◯◯ VERY LOW	CRITICAL
Maternal cumulative duration											
Diabetes											
11	Cohort Studies	moderate[o]	not serious	not serious	not serious	none	345094	0.96 [0.91, 1.02]		⊕◯◯◯ VERY LOW	CRITICAL
Blood Glucose (mg/dL)											
3	Cohort Studies	moderate[o]	not serious	not serious	not serious	none	14296		MD 0.46 higher (1.70 lower to 2.62 higher)	⊕◯◯◯ VERY LOW	CRITICAL
Hypertension											
9	Cohort Studies	moderate[o]	not serious	not serious	not serious	none	172366	0.99 [0.95, 1.03]		⊕◯◯◯ VERY LOW	CRITICAL
Systolic Blood Pressure											
1	Cohort Studies	moderate[o]	not serious	serious[h]	not serious	none	782		MD 0.53 higher (1.80 lower to 2.86 higher)	⊕◯◯◯ VERY LOW	CRITICAL
Diastolic Blood Pressure											
1	Cohort Studies	moderate[o]	not serious	serious[h]	not serious	none	782		MD 0.64 higher (1.00 lower to 2.28 higher)	⊕◯◯◯ VERY LOW	CRITICAL
All CVD											

4	Cohort Studies	moderate[o]	not serious	not serious	not serious	none	328381	1.01 [0.97, 1.04]		⊕◯◯◯ VERY LOW	CRITICAL
Cardiovascular mortality											
2	Cohort Studies	moderate[o]	not serious	not serious	not serious	none	8613	1.15 [0.95, 1.40]		⊕◯◯◯ VERY LOW	CRITICAL
Hyperlipidaemia											
5	Cohort Studies	moderate[o]	not serious	not serious	not serious	none	38646		**MD 0.91 higher (0.84 higher to 1.00 higher)**	⊕◯◯◯ VERY LOW	CRITICAL
Cholesterol (mg/dL)											
2	Cohort Studies	moderate[o]	not serious	not serious	not serious	none	12449		MD 1.24 lower (4.21 lower to 1.74 higher)	⊕◯◯◯ VERY LOW	CRITICAL
LDL (mg/dL)											
2	Cohort Studies	moderate[o]	not serious	not serious	not serious	none	12278		MD 1.51 lower (4.36 lower to 1.34 higher)	⊕◯◯◯ VERY LOW	CRITICAL
HDL (mg/dL)											
3	Cohort Studies	moderate[o]	not serious	not serious	not serious	none	14296		MD 0.66 higher (0.68 lower to 2.01 higher)	⊕◯◯◯ VERY LOW	CRITICAL
Low HDL											
4	Cohort Studies	moderate[o]	not serious	not serious	not serious	none	7503	0.94 [0.80, 1.12]		⊕◯◯◯ VERY LOW	CRITICAL
Triglyceride (mmol/L)											
3	Cohort Studies	moderate[o]	serious[p]	not serious	not serious	none	14296		MD 0.03 higher (0.06 lower to 0.11 higher)	⊕◯◯◯ VERY LOW	CRITICAL
Metabolic Syndrome											
4	Cohort Studies	moderate[o]	not serious	not serious	not serious	none	7503	0.88 [0.72, 1.08]		⊕◯◯◯ VERY LOW	CRITICAL
Obesity											
7	Cohort Studies	moderate[o]	not serious	not serious	not serious	none	51547	**0.95 [0.90, 1.00]**		⊕◯◯◯ VERY LOW	CRITICAL
BMI											
3	Cohort Studies	moderate[o]	serious[p]	not serious	not serious	none	12463		MD 0.04 lower (0.31 lower to 0.23 higher)	⊕◯◯◯ VERY LOW	CRITICAL
Waist Circumference [cm]											
2	Cohort Studies	moderate[o]	not serious	not serious	not serious	none	13514		MD 0.17 lower (0.53 lower to 0.20 higher)	⊕◯◯◯ VERY LOW	CRITICAL
Maternal cumulative duration: Cancer & Bone health											
Breast Cancer											
6	Cohort Studies	moderate[o]	not serious	not serious	not serious	none	218966	1.07 [0.95, 1.20]		⊕◯◯◯ VERY LOW	CRITICAL
Ovarian Cancer											
2	Cohort Studies	moderate[o]	not serious	not serious	not serious	none	270147	1.12 [0.69, 1.83]		⊕◯◯◯ VERY LOW	CRITICAL
Lung Cancer											
2	Cohort Studies	moderate[o]	not serious	not serious	not serious	none	397789	1.12 [0.69, 1.83]		⊕◯◯◯ VERY LOW	CRITICAL
Osteoporosis											

1	Cohort Studies	serious[n]	not serious	not serious	not serious	none	721	**1.66 [1.21, 2.26]**		⊕◯◯◯ VERY LOW	CRITICAL
Bone Mass Density Total Femur											
1	Cohort Studies	serious[n]	not serious	serious[h]	serious[q]	none	75		MD 0.02 lower (0.07 lower to 0.03 higher)	⊕◯◯◯ VERY LOW	CRITICAL
Bone Mass Density Femoral Neck [g/cm2]											
2	Cohort Studies	serious[n]	not serious	not serious	not serious	none	796		**MD 0.03 lower (0.05 lower to 0.01 lower)**	⊕◯◯◯ VERY LOW	CRITICAL
Bone Mass Density [g/cm2]											
1	Cohort Studies	serious[n]	not serious	serious[h]	serious[q]	none	75		MD 0.01 lower (0.06 lower to 0.04 higher)	⊕◯◯◯ VERY LOW	CRITICAL
Bone Mass Density Lumbar Spine [g/cm2]											
1	Cohort Studies	serious[n]	not serious	serious[h]	not serious	none	721		**MD 0.04 lower (0.06 lower to 0.02 lower)**	⊕◯◯◯ VERY LOW	CRITICAL
Bone Mass Density Tibia											
1	Cohort Studies	serious[n]	not serious	serious[h]	serious[q]	none	75		MD 0.27 lower (0.73 lower to 0.18 higher)	⊕◯◯◯ VERY LOW	CRITICAL

CI: confidence interval; **MD:** mean difference; **OR:** odds ratio; **SMD:** standardised mean difference

Explanations

a. Downgraded by 1 for serious risk of bias. One included study had serious risk of bias due to confounding, classification of exposure, departure from intended exposure and missing data.
b. Downgraded by 2 for very serious risk of indirectness. Evidence from only one study with a small dataset; uncertain about the applicability of the evidence.
c. Downgraded by 2 for very serious risk of bias. Of the 6 included studies 5 had moderate and 1 serious risk of bias due confounding and due to departure from intended exposure, 2 had moderate risk of bias in selection of participants, 3 had moderate risk of bias in classification of exposure, and 4 had moderate risk of bias due to missing data.
d. One study had moderate risk of bias due to confounding, missing data and selection of reported results.
e. One study had moderate risk of bias due to confounding and departure from intended exposure.
f. Downgraded by 3 for serious risk of bias due to confounding and departure from intended exposure, selection of participants, classification of exposure, selection of reported result, and missing data.
g. Downgraded by 1 for moderate risk of bias due to confounding, selection of participants, and departure from intended exposure.
h. Downgraded by 2 for serious risk of indirectness: The available evidence is from one population set and age group from only one country, the findings of which cannot be extrapolated to other settings, locations or age groups.
i. Downgraded by 1 for serious risk of bias due to confounding and departure from intended exposure.
j. Downgraded by 1 for moderate risk of bias due to confounding, departure from intended exposure, selection of participants, missing data and classification of exposure.
k. Downgraded by 1 for imprecision: 95% confidence interval (or alternative estimate of precision) around the pooled or best estimate of effect includes both 1) no effect and 2) appreciable benefit or appreciable harm and/or low cumulative sample.
l. Downgraded by 3 for moderate risk of bias due to confounding, departure from intended exposure
classification of exposure, and missing data.
m. Downgraded by 3 for serious risk of bias due to confounding and departure from intended exposure and classification of exposure.
n. Downgraded by 1 for serious risk of bias due to confounding, departure from intended exposure and missing data.
o. Downgraded by 3 for moderate risk of bias due to confounding, classification of exposure, departure from intended exposure and missing data.
p. There was widely differing estimates of the outcome with significant heterogeneity in results across studies.
q. Total cumulative sample size is low.

Milks for infants 6–11-months of age fed milks other than breast milk

Animal milk versus formula milk

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Animal Milk	Infant Formula	Relative (95% CI)	Absolute (95% CI)		
Anaemia at longest follow up-Randomized Controlled Trials												
2	randomised trials	serious[a]	not serious[b]	serious[c]	not serious[d]	none	12/60 (20.0%)	6/149 (4.0%)	**RR 4.03 (1.68 to 9.65)**	**122 more per 1,000 (27 more to 348 more)**	⊕⊕◯◯ LOW	CRITICAL
Any anaemia at the longest follow up-Cohort studies												
2	observational studies	serious[e]	not serious[f]	serious[c]	not serious	none	20/155 (12.9%)	11/172 (6.4%)	**RR 2.26 (1.15 to 4.43)**	**81 more per 1,000 (10 more to 219 more)**	⊕⊕◯◯ LOW	CRITICAL
Gastrointestinal blood loss at longest follow up - Randomized Controlled Trials												
1	randomised trials	serious[g]	not serious	serious[c]	not serious[h]	none	9/21 (42.9%)	3/22 (13.6%)	RR 3.14 (0.98 to 10.04)	292 more per 1,000 (3 fewer to 1,000 more)	⊕⊕◯◯ LOW	CRITICAL
Gastrointestinal blood loss - Cohort study												
1	observational studies	serious[i]	not serious	serious[j]	serious[k]	none	26/60 (43.3%)	6/21 (28.6%)	RR 1.52 (0.73 to 3.16)	149 more per 1,000 (77 fewer to 617 more)	⊕◯◯◯ VERY LOW	CRITICAL
Weight for age at longest follow up-Randomized Controlled Trials												
3	randomised trials	serious[l]	not serious[m]	serious[n]	not serious[o]	none	194	362	-	SMD 0.02 SD lower (0.26 lower to 0.21 higher)	⊕⊕◯◯ LOW	CRITICAL
Height for age at the longest follow up-Randomized Controlled Trials												
2	randomised trials	serious[p]	not serious[q]	serious[n]	not serious[r]	none	185	344	-	SMD 0.07 SD higher (0.15 lower to 0.3 higher)	⊕⊕◯◯ LOW	CRITICAL
Serum haemoglobin concentration at the longest follow up - Randomized Controlled Trials												
3	randomised trials	serious[s]	not serious[b]	serious[c]	not serious	none	82	168	-	**SMD 0.32 SD lower (0.59 lower to 0.05 lower)**	⊕⊕◯◯ LOW	CRITICAL
Serum haemoglobin level - Cohort studies												

2	observational studies	serious[t]	not serious	serious[j]	not serious[u]	none	148	98	-	SMD 0.37 SD lower (0.78 lower to 0.05 higher)	⊕⊕◯◯ LOW	CRITICAL
Iron deficiency anaemia at the longest follow up-Cohort studies												
2	observational studies	not serious	not serious[f]	serious[c]	not serious	strong association	20/155 (12.9%)	11/172 (6.4%)	**RR 2.26 (1.15 to 4.43)**	**81 more per 1,000 (10 more to 219 more)**	⊕⊕◯◯ LOW	CRITICAL
Constipation-Cohort study												
1	observational studies	not serious	not serious	serious[j]	serious[v]	strong association	7/69 (10.1%)	3/98 (3.1%)	RR 3.31 (0.89 to 12.37)	71 more per 1,000 (3 fewer to 348 more)	⊕◯◯◯ VERY LOW	CRITICAL
Diarrhoea-Cohort study												
1	observational studies	not serious	not serious	serious[j]	not serious	none	21/69 (30.4%)	16/98 (16.3%)	**RR 1.86 (1.05 to 33.10)**	**140 more per 1,000 (8 more to 1,000 more)**	⊕◯◯◯ VERY LOW	CRITICAL
Neurodevelopment outcome (PDI scores) at the longest follow-Randomized Controlled Trial												
1	randomised trials	not serious	not serious	serious[j]	serious[w]	none	160	268	-	SMD 0.18 SD higher (0.02 lower to 0.37 higher)	⊕⊕◯◯ LOW	IMPORTANT
Neurodevelopment outcome (MDI score) at the longest follow up-Randomized Controlled Trial												
1	randomised trials	not serious	not serious	serious[j]	serious[x]	none	160	268	-	SMD 0.16 SD higher (0.03 lower to 0.36 higher)	⊕⊕◯◯ LOW	IMPORTANT

CI: confidence interval; **RR:** risk ratio; **SMD:** standardised mean difference

Explanations

a. One of the two randomized trial studies had "some concerns' for the risk of bias from the Cochrane risk of bias tool (2).
b. No statistical heterogeneity was found in the pooled data. I2 = 0%. There was clinical heterogeneity in the type of formula and animal milk use. We did not downgrade the grade level for clinical heterogeneity as there is no consensus on the type of formula or animal milk that should be used when the breastmilk is not available and that multiple options are available for infant formula and animal milk in the community.
c. All the included studies were from high-income countries. This might limit the applicability of the results to populations from low and middle-income countries. We however think that the direction of effect might remain the same if there were eligible studies from low and middle-income countries and the magnitude of the effect might increase against animal milk.
d. Results were statistically significant and the confidence interval is fairly narrow around the summary estimate.
e. One cohort study had high risk of bias and the second one had some concerns for risk of bias.
f. The I2 statistics was 0 %
g. Study had "some concerns" for risk of bias based on Cochrane risk of bias tool (2).
h. Even though the confidence interval around the summary estimate included 1, the lower limit of the confidence interval was 0.98.
i. The study had 'high risk of bias' from the ROBINS tool.
j. The only include study for this outcome was conducted in high-income country
k. The confidence interval around the summary estimate inlcuded 1 and risk of increased or decreased risk cannot be excluded.
l. All three studies were randomized trials. One of the three randomized trial studies had 'high' and another has "some concerns' for the risk of bias from Cochrane risk of bias tool-2 (ROB 2).
m. The overall unexplained statistical heterogeneity based on 12 statistics was 19 %. The visual inspection of the forest plot showed that three of the included studies had an effect in the same direction and around the mean summary estimate. We did not downgrade the grade level for inconsistency for this outcome.
n. All other than one of the included studies were from high-income countries. This might limit the applicability of the results to populations from low and middle-income countries.

o. The overall magnitude of the effect for the weight for age was small (SMD 0.06). This small statistical effect is not meaningful clinically. Also, even though the confidence interval included 0, the total sample size from the pooled studies was 1216. We think there was optimal information size (OIS) from the sample size of the pooled studies that if there was a true effect, that should have been picked up by this much of sample size. We, therefore, did not downgrade for imprecision.
p. Two studies were randomized trials. One of the two randomized trial studies had "some concerns' for the risk of bias from the Cochrane risk of bias tool-2(ROB 2).
q. Unexplained statistically heterogeneity based on 12 statistics was 17 % only.
r. The overall magnitude of the effect for the weight for age was small (SMD 0.07) and the confidence interval included 0. This is a very small effect clinically. The total sample size in the analysis was 529 which should have been enough to pick a clinically meaningful effect. We, therefore, did not downgrade the level for imprecision.
s. The randomized trial studies had "some concerns' for the risk of bias from the Cochrane risk of bias tool-2.
t. One of the observational studies had 'high' risk of bias and the other had a 'moderate' risk of bias from the ROBINS-1 tool
u. Even though the confidence interval around the summary estimate included a null effect, the upper limit was almost toward the threshold of statistical significance. The data from RCTs showed a similar direction of effect and was statistically significant.
v. The 95 % CI around the summary estimate included 1. The total sample size was 167 which is not large enough to be confident about the summary estimate.
w. The overall magnitude of the effect was small (SMD 0.18) and the confidence interval included 0.
x. The overall magnitude of the effect was small (SMD 0.16) and the confidence interval included 0.

Milks for children 12-23 months of age fed milks other than breast milk

Animal milk vs follow-on formula

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Animal milk (full-fat or lower fat milk)	follow-on formula	Relative (95% CI)	Absolute (95% CI)		
Weight (kg) (All)												
3 [1]	randomised trials	not serious	not serious	not serious	serious [a]	none	250	354	-	MD 0.13 kg higher (-0.11 lower to 0.36 higher)	⊕⊕⊕◯ MODERATE	CRITICAL
Height (cm) (All)												
3 [1]	randomised trials	not serious	not serious	not serious	serious [a]	none	250	354	-	MD 0.20 cm higher (-0.31 lower to 0.72 higher)	⊕⊕⊕◯ MODERATE	CRITICAL
Weight for height z score												
1 [2]	randomised trials	serious	not serious	not serious	serious [a]	none	73	70	-	MD 0.3 higher (0.01 lower to 0.61 higher)	⊕⊕◯◯ LOW	CRITICAL
Head circumference (cm)												
2 [3]	randomized trial	not serious	not serious	not serious	serious [a]	none	157	268	-	MD 0.05 lower (0.36 lower to 0.26 higher)	⊕⊕⊕◯ MODERATE	CRITICAL
Body composition-Body Mass Index												
2 [4]	randomized trial	not serious	not serious	not serious	serious [a]	none	90	86	-	MD 0.28 higher (0.15 lower to 0.7 higher)	⊕⊕◯◯ LOW -	CRITICAL
Body composition-% Body fat												
1 [2]	randomized trial	not serious	not serious	not serious	serious [a]	none	67	67	-	MD 2.4 higher (0.16 lower to 4.96 higher)	⊕⊕◯◯ LOW -	CRITICAL
Nutrient status-Vit D as serum 25-hydroxyvitamin D [25(OH)D], nmol/L												

2 [5]	randomised trials	not serious	not serious	not serious	serious [a]	none	231	224	-	MD 16.27 nmol/L lower (-21.23 lower to 11.31 lower)	⊕⊕◯◯ LOW	CRITICAL
Nutritional status-Vit D deficiency												
2 [5]	randomised trials	not serious	not serious	not serious	serious [a]	none	47/231 (20.3%)	17/224 (7.6%)	RR 2.64 (1.57 to 4.45)	124 more per 1000 (43 more to 262 more)	⊕⊕◯◯ LOW	CRITICAL
Nutritional status-Iron as serum iron (µmol/l)												
1 [2]	randomised trials	serious	not serious	not serious	serious [a]	none	67	67	-	MD 0.7 lower (2.63 lower to 1.23 higher)	⊕⊕◯◯ LOW	CRITICAL
Child development-Bayley psychomotor development index (PDI)												
1 [3]	randomised trials	serious	not serious	not serious	serious [a]	none	155	240	-	MD 1.15 lower (3.07 lower to 0.77 higher)	⊕⊕◯◯ LOW	CRITICAL
Child development-Bayley mental development index (MDI)												
1 [3]	randomised trials	not serious	not serious	not serious	serious [a]	none	155	240	-	MD 1.55 points higher (0.64 lower to 3.73 higher)	⊕⊕⊕◯ MODERATE	CRITICAL
Iron deficiency anaemia (IDA)												
2 [5]	randomised trials	not serious	not serious	not serious	very serious [a]	none	9/223 (4.0%)	1/222 (0.5%)	**RR 6.16** (1.11 to 34.20)	**23 more per 1000** (0 fewer to 150 more)	⊕⊕◯◯ LOW	CRITICAL
Iron deficiency (ID, serum ferritin <12 µg/l)												
2 [5]	randomised trials	not serious	not serious	not serious	very serious [a]	none	45/227 (19.8%)	19/225 (8.4%)	**RR 2.33 (1.40 to 3.86)**	**112 more per 1000 (34 more to 242 more)**	⊕⊕◯◯ LOW	CRITICAL
Haemoglobin (g/L)												
5 [6]	randomised trials	serious	serious	not serious	not serious [a]	none	315	348	-	**MD 2.61 g/L lower (4.86 lower to 0.37 lower)**	⊕⊕◯◯ LOW	CRITICAL
Ferritin (µg/L) (All)												
5 [6]	randomised trials	serious	serious	not serious	not serious [a]	none	362	434	-	MD 9.87 lower (15.02 lower to 4.72 lower)	⊕⊕◯◯ LOW	CRITICAL
Gut health-Stool frequency (per day)												
1 [7]	randomised trials	not serious	not serious	not serious	very serious [a]	none	2/153 (1.3%)	1/153 (0.7%)	RR 2.00 (0.18 to 21.83)	7 more per 1000 (5 fewer to 136 more)	⊕⊕◯◯ LOW	IMPORTANT

I: Confidence interval; **MD:** Mean difference; **RR:** Risk ratio

Explanations

a. Total number of participants is less than 400 (a "rule of thumb") for continuous outcomes and less than 300 for continuous outcomes. Information is likely to be insufficient to precise effect estimate.

1. Duration of the intervention varied between the 3 included studies, from 4, 9 and 12 months. Two trials were 2-arms and on was 3-arms. Evidence was downgraded one level. One trial was assessed as low risk of bias (Morley). Maldonado did not report on the similarity of baseline characteristics.

2. One trial reported on this outcome (Lovell). We downgraded on level for detection bias as final data analysis post data lock was not blinded by treatment group, and no interim analysis was planned for the trial. Trial was downgraded another level for imprecision.

3. One trial reported on this outcome and was downgraded one level due to small sample size (Morley). It is a 3-arm RTC and showed no clear signs of risk of bias.

4.Two trials reported on this outcome (Lovell and Maldonado). Lovell was assessed with unclear risk for performance bias and detection bias. Final data analysis post data lock was not blinded by treatment group, and no interim analysis was planned for the trial. The second trial was assessed with high risk of bias as similarity of baseline characteristics was not reported. Therefore, we downgraded one level the evidence. Both trials had serious imprecision due to small sample size and we downgraded the evidence another level.

5. Two trials reported on this outcome and in one the intervention lasted 5 months (Akkermans) while in the other lasted 12 months (Lovell). In the first one, there was unclear risk for contamination (performance bias) and allocation concealment (selection bias). Final data analysis post data lock was not blinded by treatment group, and no interim analysis was planned for the trial. There were no reported adverse reactions to the study milk and therefore the blinding procedure was maintained until the end of study. One trial (Akkermans) provided low-fat animal milk (1.7g/100 ml) and follow-up period was 5 months; while the other trial (Lovell) provided full-fat animal milk (3.1g/100ml) and follow up period was 12 months.

6. From the five trials reporting on this outcome, two were assessed at high risk of bias. Risk for performance bias in Daly as mothers from both groups on income support were still entitled to claim free cows' milk with milk tokens. However, as not all parents were in receipt of income support, and therefore not entitled to the cows' milk, the cows' milk group received funding to purchase 500 ml cows' milk per day. Risk for selection bias (confounding) as similarity of baseline characteristics was not reported. Age range of children varied between studies: 6-18 mo and followed until 24 mo in Daly; 9 mo and followed until 18 mo in Morley; 12-36 mo in Akkermans and followed for 5 mo; 12 mo and followed until 24 mo in Lovell; and 12-30 mo for 4 mo in Maldonado. There was substantial heterogeneity and therefore, evidence was downgraded another level for this.

7. One trial reported on this outcome (Akkermans). Multi-country, 2-arm study with children 12-24 mo of age visiting hospitals and clinics and the follow-up period was 5 months. Follow-on formula contained 1-2 mg of iron (form not specified) and 1.7 mg vitamin D, compared animal milk. The study was funded by the private industry but analyzed independently and before unblinding the study. Unclear risk for contamination (performance bias) and allocation concealment (selection bias). Results are inconclusive, but the size of the effect included potentially important benefits, test of overall effect (P=0.57). Evidence was downgraded to levels.

Full-fat milk vs lower-fat milk

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Full-fat animal milk	Lower-fat milk	Relative (95% CI)	Absolute (95% CI)		
Nutrient status-serum cholesterol (mmol/l)												
1[1]	randomized trials	serious	not serious	not serious	serious	none	9	8	-	MD 0.17 lower (0.92 lower to 0.58 higher)	⊕⊕◯◯ LOW	CRITICAL
Nutrient status-serum low density lipoprotein (LDL) (mmol/l)												
1[1]	randomized trials	serious	not serious	not serious	serious	none	9	8	-	MD 0.25 lower (0.94 lower to 0.44 higher)	⊕⊕◯◯ LOW	CRITICAL
Nutrient status-serum high density lipoprotein (HDL) (mmol/l)												
1[1]	randomized trials	serious	not serious	not serious	serious	none	9	8	-	MD 0.1 lower (0.3 lower to 0.1 higher)	⊕⊕◯◯ LOW	CRITICAL
Nutrient status-serum triglycerides (mmol/l)												
1[1]	randomized trials	serious	not serious	not serious	serious	none	9	8	-	MD 0.34 higher (0.12 lower to 0.8 higher)	⊕⊕◯◯ LOW	CRITICAL
Nutrient status-serum LDL/HDL												
1[1]	randomized trials	serious	not serious	not serious	serious	none	9	8	-	MD 0.05 lower (0.83 lower to 0.73 higher)	⊕⊕◯◯ LOW	CRITICAL

CI: Confidence interval; **MD:** Mean difference;

Explanations

1. One trial reported on outcomes for this comparison (Svahn). The intervention lasted 6 months. Evidence was downgraded two levels. It assessed as high risk for detection bias and reporting bias, and it had very small sample size so there was not enough information to detect a precise estimate of the effect. There was unclear risk for selection bias and for similarity of baseline characteristics.

Animal milk versus plant-based milk

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Animal milk	Plant-based milk alternatives	Relative (95% CI)	Absolute (95% CI)		
Nutrient status-serum cholesterol (mmol/l)												
1	randomized trials	serious	not serious	not serious	serious	none	9	12	-	MD 0.16 lower (0.76 lower to 0.44 higher)	⊕⊕◯◯ LOW	CRITICAL
Nutrient status-serum low density lipoprotein (LDL) (mmol/l)												
1	randomized trials	serious	not serious	not serious	serious	none	9	12	-	MD 0.03 lower (0.48 lower to 0.54 higher)	⊕⊕◯◯ LOW	CRITICAL
Nutrient status-serum high density lipoprotein (HDL) (mmol/l)												
1	randomized trials	serious	not serious	not serious	serious	none	9	12	-	MD 0.18 lower (0.85 lower to 0.49 higher)	⊕⊕◯◯ LOW	CRITICAL
Nutrient status-serum triglycerides (mmol/l)												
1	randomized trials	serious	not serious	not serious	serious	none	9	12	-	MD 0.08 lower (0.63 lower to 0.47 higher)	⊕⊕◯◯ LOW	CRITICAL
Nutrient status-serum LDL/HDL												
1	randomized trials	serious	not serious	not serious	serious	none	9	12	-	MD 0.33 higher (0.36 lower to 1.02 higher)	⊕⊕◯◯ LOW	CRITICAL

CI: Confidence interval; **MD:** Mean difference;

Explanations

1. One trial reported on outcomes for this comparison (Svahn). The intervention lasted 6 months. Evidence was downgraded two levels. It assessed as high risk for detection bias and reporting bias, and it had very small sample size so there was not enough information to detect a precise estimate of the effect. There was unclear risk for selection bias and for similarity of baseline characteristics.

Age of introduction of complementary foods

Early introduction of CF (<= 4 months of age) compared to 6 months of age

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	With early introduction of CF <=4 months	With >= 6 months	Relative (95% CI)	Absolute (95% CI)		
Length												
4	RCTs	serious[a]	not serious	not serious	serious[b]	none	235	194	-	SMD 0.05 higher (0.16 lower to 0.27 higher)	⨁⨁◯◯ LOW	CRITICAL
Weight												
4	RCTs	serious[a]	not serious	not serious	serious[b]	none	235	194	-	SMD 0.06 lower (0.26 lower to 0.13 higher)	⨁⨁◯◯ LOW	CRITICAL
Head circumference/occipital frontal circumference (OFC)												
3	RCTs	serious[c]	not serious	not serious	serious[d]	none	144	144	-	SMD 0.03 higher (0.2 lower to 0.26 higher)	⨁⨁◯◯ LOW	CRITICAL
BMI												
1	RCTs	not serious	not serious	not serious	very serious[e]	none	50	50	-	MD 0.02 higher (0.41 lower to 0.45 higher)	⨁⨁◯◯ LOW	CRITICAL
BMI for age												
1	RCTs	not serious	not serious	not serious	very serious[e]	none	50	50	-	MD 0.15 lower (0.48 lower to 0.18 higher)	⨁⨁◯◯ LOW	CRITICAL
Overweight												
1	RCTs	not serious	not serious	not serious	very serious[f,g]	none	4/40 (10.0%)	1/37 (2.7%)	RR 3.70 (0.43 to 31.61)	73 more per 1,000 (15 fewer to 827 more)	⨁⨁◯◯ LOW	CRITICAL
Anaemia												
2	RCTs	serious[h]	not serious	not serious	very serious[b,g,i]	none	50/131 (38.2%)	35/85 (41.2%)	RR 0.83 (0.63 to 1.08)	70 fewer per 1,000 (152 fewer to 33 more)	⨁◯◯◯ VERY LOW	CRITICAL

Severe anaemia												
1	RCTs	serious[j]	not serious	not serious	very serious[b,f,g]	none	22/89 (24.7%)	16/50 (32.0%)	RR 0.77 (0.45 to 1.33)	74 fewer per 1,000 (176 fewer to 106 more)	⊕◯◯◯ VERY LOW	CRITICAL
Food acceptance score												
1	RCTs	serious[j]	not serious	not serious	very serious[b,e]	none	80	45	-	MD 0 (0.18 lower to 0.18 higher)	⊕◯◯◯ VERY LOW	CRITICAL

CI: confidence interval; **MD:** mean difference; **OR:** odds ratio; **SMD:** standardised mean difference

Explanations

a. Out of 4 studies, one study is at high risk for sequence generation, another study is at high risk for blinding of participants and personnel and the third study is at high risl for attrition bias
b. Number of participants vary between the groups
c. Out of 3 studies, one study has an unclear risk of bias for sequence generation, one study is at risk for blinding of participants and personnel and one study is at high risk for attrition bias
d. Number of participants <300
e. Number of participants <100
f. 95% CI <0.75 and >1.25
g. Number of events <100
h. Out of 2 studies one study is at high risk for sequence generation and both the studies are at unclear risk of bias for allocation concealment. The other study is also at unclear risk of bias for sequence generation and at high risk for attrition bias
i. 95% CI is <0.75
j. The study was at high risk for sequence generation

Early introduction of CF (< 6 months of age) compared to >=6 months of age

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	With Early introduction of CF (< 6 months of age)	With >=6 months of age	Relative (95% CI)	Absolute (95% CI)		
Stunting												
10	Observational studies	very serious[a]	very serious[b]	not serious	serious[c]	none	3399/6277 (54.2%)	3774/6539 (57.7%)	OR 1.16 (0.77 to 1.75)	36 more per 1,000 (65 fewer to 128 more)	⊕◯◯◯ VERY LOW	CRITICAL
Underweight												
6	Observational studies	very serious[a]	not serious	not serious	serious[c,d]	none	316/2458 (12.9%)	342/3481 (9.8%)	**OR 1.29 (1.08 to 1.53)**	**25 more per 1,000 (7 more to 45 more)**	⊕◯◯◯ VERY LOW	CRITICAL
Wasting												
6	Observational studies	very serious[a]	very serious[b]	not serious	serious[c,e]	none	125/559 (22.4%)	193/1281 (15.1%)	OR 1.55 (0.91 to 2.62)	65 more per 1,000 (12 fewer to 167 more)	⊕◯◯◯ VERY LOW	CRITICAL
Thinness												
1	Observational studies	very serious[a]	not serious	not serious	very serious[f]	none	5/55 (9.1%)	3/84 (3.6%)	OR 2.70 (0.62 to 11.79)	55 more per 1,000 (13 fewer to 268 more)	⊕◯◯◯ VERY LOW	CRITICAL
Height for Age Z-Score												
2	Observational studies	serious[g]	not serious	not serious	serious[h]	none	673	676	-	MD 0.03 higher (0.13 lower to 0.19 higher)	⊕⊕◯◯ LOW	CRITICAL
Weight for Age Z-Scores												
1	Observational studies	serious[g]	not serious	not serious	serious[d]	none	590	123	-	MD 0.08 higher (0.12 lower to 0.27 higher)	⊕⊕◯◯ LOW	CRITICAL
Weight for Height Z-Score												

1	Observational studies	serious[g]	not serious	not serious	serious[d]	none	1224	771	-	MD 0 (0.01 lower to 0)	⊕◯◯◯ VERY LOW	CRITICAL
Height/Length												
6	Observational studies	serious[i]	very serious[b]	not serious	serious[d]	none	1827	732	-	MD 0.12 lower (0.54 lower to 0.3 higher)	⊕◯◯◯ VERY LOW	CRITICAL
Weight												
6	Observational studies	serious[g]	serious[j]	not serious	serious[d]	none	1423	654	-	SMD 0.13 SD higher (0.02 lower to 0.29 higher)	⊕◯◯◯ VERY LOW	CRITICAL
BMI												
3	Observational studies	serious[k]	not serious	not serious	serious[d]	none	2071	1461	-	**SMD 0.13 SD higher (0.05 higher to 0.21 higher)**	⊕⊕◯◯ LOW	CRITICAL
BMI Z score												
1	Observational studies	serious[g]	not serious	not serious	serious[d]	none	1803	644	-	**SMD 0.19 SD higher (0.09 higher to 0.29 higher)**	⊕⊕◯◯ LOW	CRITICAL
Head Circumference												
1	Observational studies	serious[g]	not serious	not serious	very serious[l]	none	21	20	-	MD 0.2 higher (0.34 lower to 0.74 higher)	⊕◯◯◯ VERY LOW	CRITICAL
Overweight												
4	Observational studies	serious[g]	not serious	not serious	serious[d]	none	501/4624 (10.8%)	432/3799 (11.4%)	OR 1.17 (0.89 to 1.54)	17 more per 1,000 (11 fewer to 51 more)	⊕◯◯◯ VERY LOW	CRITICAL
Obesity												
4	Observational studies	serious[g]	not serious	not serious	serious[d]	none	1169/8384 (13.9%)	626/5434 (11.5%)	OR 1.06 (0.95 to 1.19)	6 more per 1,000 (5 fewer to 19 more)	⊕◯◯◯ VERY LOW	CRITICAL
Overweight and Obese												
4	Observational studies	serious[g]	not serious	not serious	serious[c,d]	none	801/7115 (11.3%)	123/1963 (6.3%)	**OR 1.34 (1.09 to 1.65)**	**20 more per 1,000 (5 more to 37 more)**	⊕⊕◯◯ LOW	CRITICAL

Anaemia												
2	Observational studies	serious[m]	very serious[b]	not serious	serious[c,d]	none	551/2637 (20.9%)	2328/16286 (14.3%)	OR 1.72 (0.90 to 3.27)	80 more per 1,000 (12 fewer to 210 more)	⊕◯◯◯ VERY LOW	CRITICAL
Iron deficiency anaemia												
1	Observational studies	serious[a]	not serious	not serious	very serious[d,f]	none	19/597 (3.2%)	24/273 (8.8%)	**OR 0.34 (0.18 to 0.63)**	**56 fewer per 1,000 (71 fewer to 31 fewer)**	⊕◯◯◯ VERY LOW	CRITICAL
Diarrhoea												
2	Observational studies	serious[m]	serious[j]	not serious	serious[d]	none	208/102695 (0.2%)	242/4712 (5.1%)	OR 0.64 (0.21 to 1.97)	18 fewer per 1,000 (40 fewer to 45 more)	⊕◯◯◯ VERY LOW	CRITICAL
Atopic dermatitis												
1	Observational studies	serious[n]	not serious	not serious	serious[d,e]	none	42/391 (10.7%)	107/1034 (10.3%)	OR 1.04 (0.71 to 1.52)	4 more per 1,000 (28 fewer to 46 more)	⊕⊕◯◯ LOW	CRITICAL
Asthma												
3	Observational studies	very serious[o]	not serious	not serious	serious[d,e]	none	135/3062 (4.4%)	66/1876 (3.5%)	OR 0.99 (0.71 to 1.38)	0 fewer per 1,000 (10 fewer to 13 more)	⊕◯◯◯ VERY LOW	CRITICAL
LRTI												
3	Observational studies	serious[p]	not serious	not serious	serious[d]	none	773/105225 (0.7%)	118/4810 (2.5%)	OR 1.11 (0.90 to 1.38)	3 more per 1,000 (2 fewer to 9 more)	⊕◯◯◯ VERY LOW	CRITICAL
Wheeze												
1	Observational studies	very serious[q]	not serious	not serious	serious[d,e]	none	173/1683 (10.3%)	70/637 (11.0%)	OR 0.93 (0.69 to 1.25)	7 fewer per 1,000 (31 fewer to 24 more)	⊕◯◯◯ VERY LOW	CRITICAL
Eczema												
1	Observational studies	very serious[q]	not serious	not serious	serious[d]	none	328/1661 (19.7%)	113/629 (18.0%)	OR 1.12 (0.89 to 1.42)	17 more per 1,000 (17 fewer to 58 more)	⊕◯◯◯ VERY LOW	CRITICAL
Food Allergy												

2	Observational studies	serious[m]	not serious	not serious	serious[d,e]	none	84/1357 (6.2%)	61/1039 (5.9%)	OR 0.90 (0.60 to 1.35)	6 fewer per 1,000 (23 fewer to 19 more)	⊕◯◯◯ VERY LOW	CRITICAL
Gastrointestinal illness												
1	Observational studies	serious[a]	not serious	not serious	very serious[d,f]	none	41/186 (22.0%)	10/47 (21.3%)	OR 1.05 (0.48 to 2.28)	8 more per 1,000 (98 fewer to 169 more)	⊕◯◯◯ VERY LOW	CRITICAL
Respiratory illness												
1	Observational studies	serious[a]	not serious	not serious	very serious[d,f]	none	49/186 (26.3%)	9/47 (19.1%)	OR 1.51 (0.68 to 3.35)	72 more per 1,000 (53 fewer to 251 more)	⊕◯◯◯ VERY LOW	CRITICAL
Rickets												
1	Observational studies	serious[a]	not serious	not serious	very serious[d,f]	none	12/186 (6.5%)	1/47 (2.1%)	OR 3.17 (0.40 to 25.03)	43 more per 1,000 (13 fewer to 331 more)	⊕◯◯◯ VERY LOW	CRITICAL

CI: confidence interval; **MD:** mean difference; **OR:** odds ratio; **SMD:** standardised mean difference

Explanations

a. Most of the studies were not adjusted for confounding/ the study was not adjusted for confounding
b. Downgrade by two levels; high heterogeneity
c. 95% CI <0.75 and/or >1.25
d. Number of participants vary between the groups
e. Downgrade by one level: Number of events <300
f. Downgrade by two levels: Number of events <100
g. Study/studies had an overall moderate risk of bias
h. We downgrade by one level due to broad confidence intervals
i. Out of 7 studies, 6 have moderate risk of bias, 1 study have overall serious risk of bias
j. Heterogeneity exists; we downgrade by one level
k. Three studies have an overall moderate risk of bias and one study have a low risk of bias
l. Sample size very small
m. Studies have an unclear risk of bias for blinding, attrition bias and blinding of outcome assessors
n. The outcome assessors were not blinded to the exposure status of the participants
o. Two studies had an overall serious risk of bias and one study did not blind the assessors to the exposure status of the participants. Thus, we downgrade by two levels
p. One study has an overall serious risk of bias, second study has unclear risk of bias for blinidng, attrition. and time of exposure assessment, while the third study was at unclear risk of bias for attrition. and blinding of outcome assessors. Thus, we downgrade by one level
q. Overall, the study has an serious risk of bias. Thus, we downgrade by 2 levels

Late introduction of CF (> 6 months of age) compared to <=6 months of age

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	With late introduction of CF (> 6 months)	With <=6 months of age	Relative (95% CI)	Absolute (95% CI)		
Stunting												
7	Observational studies	very serious[a]	very serious[b]	not serious	serious[c,d]	none	2424/4023 (60.3%)	4405/7849 (56.1%)	OR 1.19 (0.71 to 2.00)	42 more per 1,000 (85 fewer to 158 more)	⊕◯◯◯ VERY LOW	CRITICAL
Underweight												
4	Observational studies	very serious[e]	very serious[f]	not serious	serious[c,d,g]	none	115/425 (27.1%)	111/682 (16.3%)	OR 1.35 (0.65 to 2.78)	45 more per 1,000 (51 fewer to 188 more)	⊕◯◯◯ VERY LOW	CRITICAL
Wasting												
3	Observational studies	very serious[h]	very serious[b]	not serious	serious[c,d,g]	none	65/299 (21.7%)	71/241 (29.5%)	OR 0.42 (0.07 to 2.56)	145 fewer per 1,000 (266 fewer to 222 more)	⊕◯◯◯ VERY LOW	CRITICAL
Thinness												
1	Observational studies	serious[h]	not serious	not serious	very serious[i]	none	3/84 (3.6%)	5/55 (9.1%)	OR 0.37 (0.08 to 1.62)	55 fewer per 1,000 (83 fewer to 49 more)	⊕◯◯◯ VERY LOW	CRITICAL
Height/Length												
2	Observational studies	serious[j]	not serious	not serious	serious[c]	none	1334	1023	-	**SMD 0.12 lower (0.2 lower to 0.04 lower)**	⊕◯◯◯ VERY LOW	CRITICAL
Weight												
2	Observational studies	very serious[k]	very serious[b]	not serious	serious[c]	none	1892	5696	-	MD 0.11 lower (0.69 lower to 0.48 higher)	⊕◯◯◯ VERY LOW	CRITICAL
BMI												

1	Observational studies	serious[j]	not serious	not serious	serious[c]	none	1260	943	-	**MD 0.14 lower (0.23 lower to 0.05 lower)**	⊕◯◯◯ VERY LOW	CRITICAL
Overweight												
3	Observational studies	serious[l]	not serious	not serious	serious[c]	none	161/1378 (11.7%)	4572/39774 (11.5%)	OR 0.94 (0.69 to 1.29)	6 fewer per 1,000 (33 fewer to 29 more)	⊕⊕◯◯ LOW	CRITICAL
Obesity												
1	Observational studies	serious[j]	not serious	not serious	serious[c]	none	32/1175 (2.7%)	1256/39335 (3.2%)	OR 0.85 (0.59 to 1.21)	5 fewer per 1,000 (13 fewer to 6 more)	⊕⊕◯◯ LOW	CRITICAL
Overweight and Obesity												
2	Observational studies	serious[m]	very serious[b]	not serious	serious[c]	none	69/493 (14.0%)	1211/4746 (25.5%)	OR 0.76 (0.10 to 5.92)	49 fewer per 1,000 (222 fewer to 415 more)	⊕◯◯◯ VERY LOW	CRITICAL
Anaemia												
2	Observational studies	serious[m]	very serious[b]	not serious	serious[c,d]	none	235/468 (50.2%)	129/236 (54.7%)	OR 2.49 (0.02 to 359.68)	204 more per 1,000 (523 fewer to 451 more)	⊕◯◯◯ VERY LOW	CRITICAL
Atopic Dermatitis												
2	Observational studies	serious[h,j]	not serious	not serious	serious[c]	none	106/1887 (5.6%)	972/16942 (5.7%)	OR 0.98 (0.79 to 1.20)	1 fewer per 1,000 (11 fewer to 11 more)	⊕⊕◯◯ LOW	CRITICAL
Asthma												
1	Observational studies	very serious[n]	not serious	not serious	serious[c,g]	none	75/1699 (4.4%)	30/644 (4.7%)	OR 0.95 (0.61 to 1.46)	2 fewer per 1,000 (18 fewer to 20 more)	⊕◯◯◯ VERY LOW	CRITICAL
Wheeze												
1	Observational studies	very serious[n]	not serious	not serious	serious[c,g]	none	173/1683 (10.3%)	70/637 (11.0%)	OR 0.93 (0.69 to 1.25)	7 fewer per 1,000 (31 fewer to 24 more)	⊕◯◯◯ VERY LOW	CRITICAL
Eczema												
1	Observational studies	very serious[n]	not serious	not serious	serious[c]	none	224/1670 (13.4%)	82/631 (13.0%)	OR 1.04 (0.79 to 1.36)	4 more per 1,000 (24 fewer to 39 more)	⊕◯◯◯ VERY LOW	CRITICAL

LRTI												
2	Observational studies	serious[o]	not serious	not serious	serious[c]	none	312/2835 (11.0%)	114/1068 (10.7%)	OR 1.09 (0.86 to 1.37)	8 more per 1,000 (14 fewer to 34 more)	⊕◯◯◯ VERY LOW	CRITICAL
Intestinal Helminth Infection												
1	Observational studies	serious[m]	not serious	not serious	serious[c]	none	346/374 (92.5%)	2/30 (6.7%)	**OR 173.00** (39.17 to 764.05)	**858 more per 1,000** (670 more to 915 more)	⊕◯◯◯ VERY LOW	CRITICAL
Systolic BP												
1	Observational studies	serious[j]	not serious	not serious	serious[c]	none	1260	943	-	**MD 1.2 lower (1.8 lower to 0.6 lower)**	⊕◯◯◯ VERY LOW	CRITICAL
Diastolic BP												
1	Observational studies	serious[j]	not serious	not serious	serious[c]	none	1260	943	-	**MD 0.72 lower (1.22 lower to 0.22 lower)**	⊕◯◯◯ VERY LOW	CRITICAL
Diarrhoea												
1	Observational studies	very serious[e]	serious[p]	not serious	serious[c,g]	none	72/177 (40.7%)	127/420 (30.2%)	**OR 1.58 (1.10 to 2.28)**	**104 more per 1,000 (20 more to 195 more)**	⊕◯◯◯ VERY LOW	CRITICAL
Food hypersensitivity												
1	Observational studies	very serious[e]	not serious	not serious	very serious[i]	none	3/6 (50.0%)	32/50 (64.0%)	OR 0.56 (0.10 to 3.08)	141 fewer per 1,000 (489 fewer to 206 more)	⊕◯◯◯ VERY LOW	CRITICAL

CI: confidence interval; **MD:** mean difference; **OR:** odds ratio; **SMD:** standardised mean difference

Explanations

a. Out of 7 studies, four studies did not control potential confounding
b. Heterogeneity exists (P value is <0.00001). Thus, we downgrade by 2 levels.
c. Number of participants vary between the groups. Thus, we downgrade by one level.
d. 95% CI <0.75 and >1.25
e. Out of 4 studies, three studies did not control potential confounding
f. Heterogeneity exists (P value is <0.005). Thus, we downgrade by 2 levels.
g. Number of events <300. Thus, we downgrade by one level.
h. Studies/study did not control for potential confounding
i. Number of events <100. Thus, we downgrade by 2 levels.
j. Study/Studies had an overall moderate risk of bias. Thus, we downgrade by one level.

k. One study had an overall high risk of bias and the other study was at moderate risk of bias thus we downgrade by two levels
l. Out of three studies, one study did not control for potential confounding, one study did not provide a justification on sample size, and one study had an overall moderate risk of bias
m. Studies/study had an unclear risk of bias for blinding, attrition bias and time of exposure assessment. Thus, we downgrade by one level.
n. Study/Studies had an overall serious risk of bias. Thus, we downgrade by 2 levels
o. Overall, one study has a serious risk of bias but the other study had an unclear risk of bias for blinding, attrition and time of exposure assessment. Thus, we downgrade by one level
p. Heterogeneity exists (P value is <0.01). Thus, we downgrade by one level

Dietary diversity: Animal source foods

Greater frequency of animal source food (ASF) vs. lesser frequency of ASF (observational studies)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	greater frequency of ASF	lesser frequency of ASF	Relative (95% CI)	Absolute (95% CI)		
Stunting in Children Consuming Meat: Assessed by Ahmad et al. 2018; Age at outcomes 6-23 months; Follow-up NA												
1	observational studies (non-comparative cross-sectional study)	serious [a]	not serious [b]	not serious	very serious [c]	none	11/392 (2.8%)	26/392 (6.6%)	RR 1.10 (0.61 to 1.96)	7 more per 1,000 (26 fewer to 64 more)	⊕◯◯◯ VERY LOW	CRITICAL
Stunting in Children Consuming Meat: Assessed by Zhao et al. 2016; Age at outcomes not reported; Follow-up NA												
1	observational studies (cross-sectional study)	very serious [d]	not serious [b]	serious [e]	not serious	none	370/668 (55.4%)	76/139 (54.7%)	RR 1.01 (0.86 to 1.20)	5 more per 1,000 (77 fewer to 109 more)	⊕◯◯◯ VERY LOW	CRITICAL
Wasting in Children Consuming Meat: Assessed by Ahmad et al. 2018; Age at outcomes 6-23 months; Follow-up NA												
1	observational studies (non-comparative cross-sectional study)	serious [a]	not serious [b]	not serious	very serious [f]	none	9/392 (2.3%)	18/392 (4.6%)	RR 1.28 (0.64 to 2.56)	13 more per 1,000 (17 fewer to 72 more)	⊕◯◯◯ VERY LOW	CRITICAL
Wasting in Children Consuming Meat: Assessed by Zhao et al. 2016; Age at outcomes not reported; Follow-up NA												
1	observational studies (cross-sectional study)	very serious [d]	not serious [b]	serious [e]	very serious [f]	none	88/660 (13.3%)	18/136 (13.2%)	RR 1.01 (0.63 to 1.62)	1 more per 1,000 (49 fewer to 82 more)	⊕◯◯◯ VERY LOW	CRITICAL
Underweight in Children Consuming Meat: Assessed by Ahmad et al. 2018; Age at outcomes 6-23 months; Follow-up NA												
1	observational studies (non-comparative cross-sectional study)	serious [a]	not serious [b]	not serious	very serious [g]	none	14/392 (3.6%)	22/392 (5.6%)	RR 1.65 (0.96 to 2.83)	36 more per 1,000 (2 fewer to 103 more)	⊕◯◯◯ VERY LOW	CRITICAL
Underweight in Children Consuming Meat: Assessed by Zhao et al. 2016; Age at outcomes not reported												
1	observational studies (cross-sectional study)	very serious [d]	not serious [b]	serious [e]	very serious [g]	none	269/668 (40.3%)	51/138 (37.0%)	RR 1.09 (0.86 to 1.38)	33 more per 1,000 (52 fewer to 140 more)	⊕◯◯◯ VERY LOW	CRITICAL
Anaemia in Children Consuming Meat: Urkin et al. 2007; Follow up 1-2 months after enrollment, children aged 12-months were the study population												

1	observational studies (non-comparative prospective study)	very serious [h]	not serious [b]	serious [e]	not serious	none			not estimable	not estimable	⊕◯◯◯ VERY LOW	IMPORTANT
Anaemia in Children Consuming Red Meat; Assessed by Silva et al, 2007; children aged 6-12 months were study population; Follow-up NA												
1	observational studies (cross-sectional study)	very serious [i]	not serious [b]	not serious [j]	serious [f]	none	81/153 (52.9%)	37/52 (71.2%)	RR 0.74 (0.59 to 0.94)	**185 fewer per 1,000 (292 fewer to 43 fewer)**	⊕◯◯◯ VERY LOW	IMPORTANT
Anaemia in Children Consuming Organ Meat i.e.. Liver: Assessed by Silva et al, 2007; children aged 6-12 months were study population; Follow-up NA												
1	observational studies (cross-sectional study)	very serious [i]	not serious [b]	not serious [j]	very serious [c]	none	46/83 (55.4%)	72/122 (59.0%)	RR 0.94 (0.74 to 1.20)	35 fewer per 1,000 (153 fewer to 118 more)	⊕◯◯◯ VERY LOW	IMPORTANT
Stunting in children Consuming Eggs: Assessed by Ahmad et al. 2018; Age at outcomes 6-23 months; Follow-up NA												
1	observational studies (non-comparative cross-sectional study)	serious [a]	not serious [b]	not serious	very serious [g]	none	43/392 (11.0%)	39/392 (9.9%)	RR 1.14 (0.79 to 1.65)	14 more per 1,000 (21 fewer to 65 more)	⊕◯◯◯ VERY LOW	CRITICAL
Stunting in Children Consuming Eggs: Assessed by Zhao et al. 2016; Age at outcomes not reported; Follow-up NA												
1	observational studies (cross-sectional study)	very serious [e]	not serious [b]	serious [e]	not serious	none	346/623 (55.5%)	100/182 (54.9%)	RR 1.01 (0.87 to 1.17)	5 more per 1,000 (71 fewer to 93 more)	⊕◯◯◯ VERY LOW	CRITICAL
Wasting in Children Consuming Eggs: Assessed by Ahmad et al. 2018; Age at outcomes 6-23 months; Follow-up NA												
1	observational studies (non-comparative cross-sectional study)	serious [a]	not serious [b]	not serious	very serious [g]	none	34/392 (8.7%)	33/392 (8.4%)	RR 1.07 (0.70 to 1.62)	6 more per 1,000 (25 fewer to 52 more)	⊕◯◯◯ VERY LOW	CRITICAL
Wasting in Children Consuming Eggs: Assessed by Zhao et al. 2016; Age at outcomes not reported; Follow-up NA												
1	observational studies (cross-sectional study)	very serious [e]	not serious [b]	serious [e]	not serious	none	346/625 (55.4%)	100/182 (54.9%)	RR 1.01 (0.87 to 1.17)	5 more per 1,000 (71 fewer to 93 more)	⊕◯◯◯ VERY LOW	CRITICAL
Underweight in Children Consuming Eggs: Assessed by Ahmad et al. 2018; Age at outcomes 6-23 months; Follow-up NA												
1	observational studies (non-comparative cross-sectional study)	serious [a]	not serious [b]	not serious	very serious [g]	none	44/392 (11.2%)	34/392 (8.7%)	RR 1.34 (0.91 to 1.97)	29 more per 1,000 (8 fewer to 84 more)	⊕◯◯◯ VERY LOW	CRITICAL
Underweight in Children Consuming Eggs: Assessed by Zhao et al. 2016; Age at outcomes not reported; Follow-up NA												
1	observational studies (cross-sectional study)	very serious [e]	not serious [b]	serious [e]	serious [k]	none	251/623 (40.3%)	69/181 (38.1%)	RR 1.06 (0.86 to 1.30)	23 more per 1,000 (53 fewer to 114 more)	⊕◯◯◯ VERY LOW	CRITICAL
Stunting in Children Consuming Fish: Assessed by Ahmad et al. 2018; Age at outcomes 6-23 months; Follow-up NA												

1	observational studies (non-comparative cross-sectional study)	serious [a]	not serious [b]	not serious	very serious [g]	none	63/392 (16.1%)	18/392 (4.6%)	RR 1.22 (0.77 to 1.91)	10 more per 1,000 (11 fewer to 42 more)	⊕◯◯◯ VERY LOW	CRITICAL
Wasting in Children Consuming Fish: Assessed by Ahmad et al. 2018; Age at outcomes 6-23 months; Follow-up NA												
1	observational studies (non-comparative cross-sectional study)	serious [a]	not serious [b]	not serious	serious [k]	none	39/392 (9.9%)	26/392 (6.6%)	RR 0.52 (0.34 to 0.80)	32 fewer per 1,000 (44 fewer to 13 fewer)	⊕⊕◯◯ LOW	CRITICAL
Underweight in Children Consuming Fish: Assessed by Ahmad et al. 2018; Age at outcomes 6-23 months; Follow-up NA												
1	observational studies (non-comparative cross-sectional study)	serious [a]	not serious [b]	not serious	serious [k]	none	54/392 (13.8%)	23/392 (5.9%)	RR 0.82 (0.54 to 1.23)	11 fewer per 1,000 (27 fewer to 13 more)	⊕⊕◯◯ LOW	CRITICAL

CI: Confidence interval; **RR:** Risk ratio

Explanations

a. Downgraded once- The use of observational study design increases the risk of biases in the study.
b. It is a single study. Thus, inconsistency cannot be determined.
c. Downgraded twice. The sample size is VERY LOW and the number of events are also < 300.
d. Downgraded twice- Lacks the mentioning of sample size justification and the exposures were also not mentioned clearly.
e. Downgraded once. No proper sample size justification and representativeness
f. Downgraded twice. Number of events < 300.
g. Downgraded twice. CI is very broad and the number of events are < 300
h. Downgraded twice- This study lacks many details on methodology making it very hard to determine whether there is internal validity. In addition, the study is described as an interventional study when it actually appears to be observational in nature. Food exposures are assessed by a questionnaire and blood samples tested for anaemia outcome/ iron stores. There is no assigning of food as an intervention but rather observation of what is eaten by questionnaire.
i. Downgraded twice- Very limited detail in the methodology makes it challenging to tell if there was internal validity of study. No sample size justification provided or mention of loss to follow up or participation rate of eligible participants. Small sample size with no information provided for baseline characteristics of study sample of children.
j. Lack of sample size justification and loss to follow-up of the participants from the baseline.
k. Downgraded once. Number of events <300

Greater amount of animal source food (ASF) vs. lesser amount of ASF (RCTs)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	greater amount of ASF	Lesser amount	Relative (95% CI)	Absolute (95% CI)		
Change in Height in Children Consuming Meat (assessed with: Higher MD indicates improvement.): Assessed by Engelmann et al. 1998; children aged 8-10 months were study population; Follow-up at 8 and 10 months after intervention												
1	randomised trials	serious [a]	not serious [b]	serious [c]	serious [d]	none	21	20	-	MD 0.1 SD lower (1.77 lower to 1.57 higher)	⊕◯◯◯ VERY LOW	CRITICAL
Change in Weight in Children Consuming Meat: (assessed with: Higher MD indicates improvement.): Assessed by Engelmann et al. 1998; children aged 8-10 months were study population; Follow-up at 8 and 10 months after intervention												
1	randomised trials	serious [a]	not serious [b]	serious [c]	serious [e]	none	21	20	-	MD 0.08 higher (0.53 lower to 0.7 higher)	⊕◯◯◯ VERY LOW	CRITICAL
Change in Triceps Skinfolds in Children Consuming Meat: (assessed with: Higher MD indicates improvement.): Assessed by Engelmann et al. 1998; children aged 8-10 months were study population; Follow-up at 8 and 10 months after intervention												
1	randomised trials	serious [a]	not serious [b]	serious [c]	serious [d]	none	21	20	-	MD 0.1 lower (5.39 lower to 5.19 higher)	⊕◯◯◯ VERY LOW	CRITICAL
Haemoglobin in Children Consuming Meat: (assessed with: Higher MD indicates improvement.): Assessed by Engelmann et al. 1998; children aged 8-10 months were study population; Follow-up at 8 and 10 months after intervention												
1	randomised trials	serious [a]	not serious [b]	serious [c]	not serious	none	21	20	-	SMD 0.21 SD higher (0.4 lower to 0.83 higher)	⊕⊕◯◯ LOW	IMPORTANT
Stunting in Children Consuming Lyophilized Beef: Assessed by Kreb et al. 2012; children aged 6-18 months were the study population; Follow-up at 6, 9, 12 and 18 months												
1	randomised trials	serious [f]	not serious [b]	serious [g]	serious [e]	none	184/532 (34.6%)	179/530 (33.8%)	RR 1.02 (0.87 to 1.21)	7 more per 1,000 (44 fewer to 71 more)	⊕◯◯◯ VERY LOW	CRITICAL
Wasting in Children Consuming Lyophilized Beef: Assessed by Kreb et al. 2012; children aged 6-18 months were the study population; Follow-up at 6, 9, 12 and 18 months												
1	randomised trials	serious [f]	not serious [b]	serious [g]	very serious [d,e]	none	38/532 (7.1%)	54/530 (10.2%)	RR 0.70 (0.47 to 1.04)	31 fewer per 1,000 (54 fewer to 4 more)	⊕◯◯◯ VERY LOW	CRITICAL
WAZ in Children Consuming Lyophilized Beef (assessed with: Higher MD indicates improvement.) : Assessed by Kreb et al. 2012; children aged 6-18 months were the study population; Follow-up at 6, 9, 12 and 18 months												
1	randomised trials	serious [f]	not serious [b]	serious [g]	not serious	none	532	530	-	**MD 0.02 lower (0.16 lower to 0.12 higher)**	⊕⊕◯◯ LOW	CRITICAL
HAZ in Children Consuming Lyophilized Beef (assessed with: Higher MD indicates improvement.) : Assessed by Kreb et al. 2012; children aged 6-18 months were the study population; Follow-up at 6, 9, 12 and 18 months												

1	randomised trials	serious [f]	not serious [b]	serious [g]	not serious	none	532	530	-	MD 0.06 lower (0.23 lower to 0.11 higher)	⊕⊕◯◯ LOW	CRITICAL
WHZ in Children Consuming Lyophilized Beef (assessed with: Higher MD indicates improvement.) : Assessed by Kreb et al. 2012; children aged 6-18 months were the study population; Follow-up at 6, 9, 12 and 18 months												
1	randomised trials	serious [f]	not serious [b]	serious [g]	not serious	none	532	530	-	MD 0.01 SD higher (1.04 lower to 0.16 higher)	⊕⊕◯◯ LOW	CRITICAL
HCAZ in Children Consuming Lyophilized Beef (assessed with: Higher MD indicates improvement.) : Assessed by Kreb et al. 2012; children aged 6-18 months were the study population; Follow-up at 6, 9, 12 and 18 months												
1	randomised trials	serious [f]	not serious [b]	serious [g]	not serious	none	532	530	-	MD 0.07 lower (0.2 lower to 0.06 higher)	⊕⊕◯◯ LOW	CRITICAL
WAZ in Children Consuming Pork (assessed with: Higher MD indicates improvement.): Assessed by Tang et al, 2014; children aged 6-18 months were the population; Follow-up at 6, 7, 9, 12, 15 and 18 months of age												
1	randomised trials	serious [h]	not serious [b]	not serious	not serious	none	462	856	-	**MD 0.08 higher (0.01 higher to 0.15 higher)**	⊕⊕⊕◯ MODERATE	CRITICAL
HAZ in Children consuming Pork (assessed with: Higher MD indicates improvement.): Assessed by Tang et al, 2014; children aged 6-18 months were the population; Follow-up at 6, 7, 9, 12, 15 and 18 months of age												
1	randomised trials	serious [h]	not serious [b]	not serious	not serious	none	462	856	-	**MD 0.11 higher (0.03 higher to 0.19 higher)**	⊕⊕⊕◯ MODERATE	CRITICAL
WHZ in Children Consuming Pork (assessed with: Higher MD indicates improvement.): Assessed by Tang et al, 2014; children aged 6-18 months were the population; Follow-up at 6, 7, 9, 12, 15 and 18 months of age												
1	randomised trials	serious [h]	not serious [b]	not serious	not serious	none	462	856	-	MD 0.03 higher (0.06 lower to 0.12 higher)	⊕⊕⊕◯ MODERATE	CRITICAL
HCAZ in Children Consuming Pork ((assessed with: Higher MD indicates improvement.): Assessed by Tang et al, 2014; children aged 6-18 months were the population; Follow-up at 6, 7, 9, 12, 15 and 18 months of age												
1	randomised trials	serious [h]	not serious [b]	not serious	not serious	none	462	856	-	MD 0.01 lower (0.07 lower to 0.05 higher)	⊕⊕⊕◯ MODERATE	CRITICAL
Change in Height in Children Consuming Pork (assessed with: Higher MD indicates improvement.): Assessed by Tang et al, 2014; children aged 6-18 months were the population; Follow-up at 6, 7, 9, 12, 15 and 18 months of age												
1	randomised trials	serious [h]	not serious [b]	not serious	not serious	none	462	856	-	**MD 0.26 higher (0.05 higher to 0.47 higher)**	⊕⊕⊕◯ MODERATE	CRITICAL
Change in Weight in Children Consuming Pork (assessed with: Higher MD indicates improvement.): Assessed by Tang et al, 2014; children aged 6-18 months were the population; Follow-up at 6, 7, 9, 12, 15 and 18 months of age												
1	randomised trials	serious [h]	not serious [b]	not serious	not serious	none	462	856	-	**MD 0.07 SD higher (0 to 0.14 higher)**	⊕⊕⊕◯ MODERATE	CRITICAL
Change in Head Circumference in Children Consuming Pork (Tang) ((assessed with: Higher MD indicates improvement.): Assessed by Tang et al, 2014; children aged 6-18 months were the population; Follow-up at 6, 7, 9, 12, 15 and 18 months of age												

1	randomised trials	serious [h]	not serious [b]	not serious	not serious	none	462	856	-	**MD 2.98 higher (2.9 higher to 3.06 higher)**	⊕⊕⊕◯ MODERATE	CRITICAL
HCAZ in Children Consuming Eggs (assessed with: Higher MD indicates improvement.): Assessed by Stewart et al. 2019; children aged 6-9.9 months were the study population; Follow-up 6 months												
1	randomised trials	serious [i]	not serious [b]	serious [j]	not serious	none	290	305	-	**MD 0.23 higher (0.05 higher to 0.41 higher)**	⊕⊕◯◯ LOW	CRITICAL
Stunting in Children Consuming Eggs: Assessed by Ianotti et al. 2017; children aged 6-9 months were the study population; Follow-up 6 months												
1	randomised trials	serious [i]	not serious [b]	serious [k]	very serious [d,e]	none	21/75 (28.0%)	29/73 (39.7%)	RR 0.70 (0.44 to 1.12)	119 fewer per 1,000 (222 fewer to 48 more)	⊕◯◯◯ VERY LOW	CRITICAL
Anaemia in Children Consuming Eggs: Assessed by Makrides et al, 2002; children aged 6-12 months were the study population; Follow-up at 6,9 and 12 months												
1	randomised trials	not serious	not serious [b]	not serious	very serious [d,e]	none	3/44 (6.8%)	2/23 (8.7%)	RR 0.78 (0.14 to 4.36)	19 fewer per 1,000 (75 fewer to 292 more)	⊕⊕◯◯ LOW	IMPORTANT
Haemoglobin in Children Consuming Eggs (assessed with: Higher MD indicates improvement.): Assessed by Makrides et al, 2002; children aged 6-12 months were the study population; Follow-up at 6,9 and 12 months												
1	randomised trials	not serious	not serious [b]	not serious	serious [e]	none	44	23	-	SMD 0.2 SD higher (0.31 lower to 0.71 higher)	⊕⊕⊕◯ MODERATE	IMPORTANT
Change in Triceps Skinfold in Children Consuming Skimmed Milk (assessed with: Higher MD indicates improvement.): Assessed by Skau et al, 2015; children aged 6-15 months were study population; Follow-up at 9 months of intervention [15 months]												
1	randomised trials	not serious [l]	not serious [b]	not serious	serious [e]	none	106	102	-	MD 0.1 lower (0.38 lower to 0.18 higher)	⊕⊕⊕◯ MODERATE	CRITICAL
Change in Head Circumference in Children Consuming Skimmed Milk (assessed with: Higher MD indicates improvement.): Assessed by Skau et al, 2015; children aged 6-15 months were study population; Follow-up at 9 months of intervention [15 months]												
1	randomised trials	not serious [l]	not serious [b]	not serious	serious [e]	none	106	102	-	MD 0.4 lower (0.76 lower to 0.04 higher)	⊕⊕⊕◯ MODERATE	CRITICAL
Change in Triceps Skinfold in Children Consuming Fish and Concoction (assessed with: Higher MD indicates improvement.): Assessed by Skau et al, 2015; children aged 6-15 months were study population; Follow-up at 9 months of intervention [15 months]												
1	randomised trials	not serious [l]	not serious [b]	not serious	serious [e]	none	85	93	-	MD 0.2 lower (0.48 lower to 0.08 higher)	⊕⊕⊕◯ MODERATE	CRITICAL
Change in Head Circumference in Children Consuming Fish and Concoction (assessed with: Higher MD indicates improvement.): Assessed by Skau et al, 2015; children aged 6-15 months were study population; Follow-up at 9 months of intervention [15 months]												
1	randomised trials	not serious [l]	not serious [b]	not serious	serious [e]	none	85	93	-	MD 0.1 lower (0.52 lower to 0.32 higher)	⊕⊕⊕◯ MODERATE	CRITICAL
Haemoglobin in Children Consuming Fish and Concoction (assessed with: Higher MD indicates improvement.): Assessed by Skau et al, 2015; children aged 6-15 months were study population; Follow-up at 9 months of intervention [15 months]												

1	randomised trials	not serious [l]	not serious [b]	not serious	serious [e]	none	69	73	-	SMD 0.28 SD lower (0.58 lower to 0.02 higher)	⊕⊕⊕◯ MODERATE	IMPORTANT
Stunting in Children Consuming Caterpillars: Assessed by Bauserman et al. 2015; children aged 6-18 months were the study population; Follow-up at 9, 12 and 18 months												
1	randomised trials	serious [m]	not serious [b]	serious [n]	very serious [d,e]	none	54/81 (66.7%)	58/82 (70.7%)	RR 0.94 (0.77 to 1.16)	42 fewer per 1,000 (163 fewer to 113 more)	⊕◯◯◯ VERY LOW	CRITICAL
Wasting in Children Consuming Caterpillar: Assessed by Bauserman et al. 2015; children aged 6-18 months were the study population; Follow-up at 9, 12 and 18 months												
1	randomised trials	serious [m]	not serious [b]	serious [n]	very serious [d,e]	none	6/80 (7.5%)	8/80 (10.0%)	RR 0.75 (0.27 to 2.06)	25 fewer per 1,000 (73 fewer to 106 more)	⊕◯◯◯ VERY LOW	CRITICAL
WAZ in Children Consuming Caterpillar (assessed with: Higher MD indicates improvement.): Assessed by Bauserman et al. 2015; children aged 6-18 months were the study population; Follow-up at 9, 12 and 18 months												
1	randomised trials	serious [m]	not serious [b]	serious [n]	very serious [d,e]	none	80	82	-	MD 0.2 lower (0.57 lower to 0.17 higher)	⊕◯◯◯ VERY LOW	CRITICAL
HAZ in Children Consuming Caterpillar (assessed with: Higher MD indicates improvement.): Assessed by Bauserman et al. 2015; children aged 6-18 months were the study population; Follow-up at 9, 12 and 18 months												
1	randomised trials	serious [m]	not serious [b]	serious [n]	serious [e]	none	81	82	-	MD 0.1 higher (0.35 lower to 0.55 higher)	⊕◯◯◯ VERY LOW	CRITICAL
WHZ in Children Consuming Caterpillar (assessed with: Higher MD indicates improvement.): Assessed by Bauserman et al. 2015; children aged 6-18 months were the study population; Follow-up at 9, 12 and 18 months												
1	randomised trials	serious [m]	not serious [b]	serious [n]	serious [e]	none	79	80	-	MD 0.2 lower (0.62 lower to 0.22 higher)	⊕◯◯◯ VERY LOW	CRITICAL
Anaemia in Children Consuming Caterpillar: Assessed by Bauserman et al. 2015; children aged 6-18 months were the study population; Follow-up at 18 months												
1	randomised trials	serious [m]	not serious [b]	serious [n]	serious [e]	none	20/77 (26.0%)	32/64 (50.0%)	**RR 0.52 (0.33 to 0.81)**	**240 fewer per 1,000 (335 fewer to 95 fewer)**	⊕◯◯◯ VERY LOW	IMPORTANT
Haemoglobin in Children Consuming Caterpillar Assessed by Bauserman et al. 2015; children aged 6-18 months were the study population; Follow-up at 18 months												
1	randomised trials	serious [m]	not serious [b]	serious [n]	serious [e]	none	77	64	**-**	**SMD 0.35 SD higher (0.02 higher to 0.69 higher)**	⊕◯◯◯ VERY LOW	IMPORTANT
WAZ (assessed with: Higher MD indicates improvement.): Assessed by Ianotti et al. 2017 and Stewart et al. 2019; children aged 6-9 months; Follow-up 6 months												
2	randomised trials	serious [o]	not serious	not serious	not serious	none	365	378	**-**	**MD 0.15 higher (0 to 0.3 higher)**	⊕⊕⊕◯ MODERATE	CRITICAL
HAZ (assessed with: Higher MD indicates improvement.): Assessed by Ianotti et al. 2017 and Stewart et al. 2019; children aged 6-9 months; Follow-up 6 months												
2	randomised trials	serious [o]	serious [p]	not serious	not serious	none	365	388	-	MD 0.06 higher (0.1 lower to 0.22 higher)	⊕⊕◯◯ LOW	CRITICAL

WHZ (assessed with: Higher MD indicates improvement.): Assessed by Ianotti et al. 2017 and Stewart et al. 2019; children aged 6-9 months; Follow-up 6 months												
2	randomised trials	serious [o]	very serious [q]	not serious	not serious	none	365	378	-	MD 0.06 lower (0.31 lower to 0.19 higher)	⊕◯◯◯ VERY LOW	CRITICAL

CI: Confidence interval; **MD:** Mean difference; **SMD:** Standardised mean difference; **RR:** Risk ratio

Explanations

a. Downgraded once- Some major concerns over bias that may have arose from using different blood collection methods. Venous blood vs capillary blood have been shown previously to have significantly different levels of haemoglobin and serum ferritin levels. However, the number of infants who received venipuncture vs finger pricks was not significantly different (p=0.9) between the two intervention groups.
b. It is a single study. Thus, inconsistency cannot be determined.
c. Downgraded once. Inability to generalize the findings because of low sample size.
d. Downgraded once. Very broad Confidence interval
e. Downgraded once. Number of events < 300
f. Downgraded once- Allocation sequence not concealed and thus increases the likelihood of bias
g. Downgraded once. The external validity is not guaranteed since the sample size is low and the risk of bias related to poor randomisation reduces the applicability.
h. Downgraded once- Study participants were randomized to an intervention arm. There is no mention of concealment
i. Downgraded once- Lack of blinding in intervention allocation.
j. Downgraded once. The sample size is appropriate. However, lack of blinding is one factor that affects the applicability of the research findings.
k. Downgraded once. The sample size is VERY LOW to make conclusions for the entire population.
l. Participants seem to have been blinded as the food packages were identical. However, one study staff member who was responsible for distribution knew intervention type in order to distribute food appropriately. Does not appear to be any deviations from intented intervention.
m. Downgraded once- No mention of blinding towards participants or those delivering the caterpillar cereal to participants.
n. Downgraded once. Lost an undesirably high percentage of infants to follow-up because many of the participants relocated and therefore outcome data were incomplete.
n. Downgraded once. Lost an undesirably high percentage of infants to follow-up because many of the participants relocated and therefore outcome data were incomplete.
o. Downgraded once. Both the studies had some concerns in regards to the blinding.
p. Downgraded once. The studies had the heterogeneity which can be seen by I2 that is more than 30%% and the p-value which is more than 0.10.
q. Downgraded twice. The studies had the heterogeneity of 63% which can be seen by I2 which is more than 30% and the p-value which is 0.10.

Greater amount of animal source food (ASF) vs. lesser amount of ASF (observational studies)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**greater amount**	**lesser amount ASF**	**Relative (95% CI)**	**Absolute (95% CI)**		
Haemoglobin in Children Consuming Cow Milk: Assessed by Thorsdottir et al. 2003; children aged 0-12 months were the study population; Follow-up at 9-12 months												
1	observational studies (Observational cohort longitudinal study)	very serious [a]	not serious [b]	serious [c]	very serious [d]		17	80	-	0 (0 to 0)	⨁◯◯◯ VERY LOW	IMPORTANT

CI: Confidence interval

Explanations

a. Downgraded twice- Lacked sample size justification, and power description
b. It is a single study. Thus, inconsistency cannot be determined.
c. Confounders not taken into consideration, nor identified.
d. No absolute values can be calculated and the number of events are < 300

More varied animal source foods (ASF) vs. less varied ASF (RCTs)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**more varied ASF**	**less varied ASF**	**Relative (95% CI)**	**Absolute (95% CI)**		
Stunting in Children Consuming varied ASF: Assessed by Krasevec at al. 2017; children aged 6-23 months were the study population; Follow-up NA												
1	observational studies (non-comparative cross-sectional study)	serious [a]	not serious [b]	serious [c]	not serious		6.467/74.548 (8.7%)	26.818/74.548 (36.0%)	**RR 0.17 (0.16 to 0.17)**	**299 fewer per 1,000 (302 fewer to 299 fewer)**	⊕⊕◯◯ LOW	CRITICAL
Stunting in Children Consuming Milk/ Milk Products: Assessed by Marinda et al. 2018; children aged 6-59 months were the study population; Follow-up NA												
1	observational studies (non-comparative cross-sectional study)	serious [d]	not serious [b]	not serious	not serious	none			not estimable		⊕⊕⊕◯ MOERATE	CRITICAL

CI: Confidence interval; **RR:** Risk ratio

Explanations

a. Downgraded once- All the participants were not selected from the same population increasing the risk of bias
b. It is a single study. Thus, inconsistency cannot be determined.
c. Downgraded once. The sample was different and not consistent; thus, it is hard to look for the sample representativeness.
d. Downgraded once- Funding for this study was provided to WorldFish by the German Federal Ministry for Economic Cooperation and Devel. Wondering if there could be some bias or advertising towards consuming fish based on the funder?

Dietary diversity: Fruits and vegetables

More frequent consumption of vegetables compared to less frequent consumption of vegetables

Certainty assessment							Impact	Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations			
Wasting									
1 [a]	observational studies [b]	not serious [c]	serious [d]	not serious [e]	serious [f]	none	Frequency of eating green leafy and orange vegetables was not associated with wasting. 20.5% (n=31), 25.2% (n=31), and 22% (n=26) of those who consumed green leafy and orange vegetables ≥4 times/week, 1-3 times/ week, and never were wasted, respectively (p=0.542 between groups). [g]	⊕◯◯◯ VERY LOW	CRITICAL
Underweight									
1 [a]	observational studies [b]	not serious [c]	serious [d]	not serious [e]	serious [f]	none	Frequency of eating green leafy and orange vegetables was not associated with being underweight. 25.8% (n=39), 26% (n=32), and 27.1% (n=32) of those who consumed green leafy and orange vegetables ≥4 times/week, 1-3 times/ week, and never were underweight, respectively (p=0.969 between groups). [g]	⊕◯◯◯ VERY LOW	CRITICAL
Stunting									
1 [a]	observational studies [b]	not serious [c]	serious [d]	not serious [e]	serious [f]	none	Frequency of eating green leafy and orange vegetables was not associated with stunting. 29.1% (n=44), 30.1% (n=37), and 23.7% (n=28) of those who consumed green leafy and orange vegetables ≥4 times/week, 1-3 times/ week, and never were stunted, respectively (p=0.491 between groups). [g]	⊕◯◯◯ VERY LOW	CRITICAL
Height-for-age Z-score									

1 [h]	observational studies [i]	serious [j]	serious [k]	not serious [l]	serious [m]	none	In breastfed children aged 9-23 months, those who consumed vegetables/ leaves 0-2 days/ week and ≥ 3 days/week had a mean HAZ of -1.01 (SD=0.93) and -0.58 (SD=0.92), respectively. In age-adjusted models, mean HAZ was -1.01 (p=0.052) and -0.59 (p<0.06), respectively, demonstrating a trend that points to lower HAZ with less vegetable consumption. [n]	⊕◯◯◯ VERY LOW	CRITICAL
Linear Growth									
1 [h]	observational studies [i]	serious [j]	serious [k]	not serious [l]	serious [m]	none	Frequent consumption of vegetables had an inverse relationship to linear growth (means: 8.3cm and 7.4cm height increment over the preceding 7 months for rare and frequent consumption, respectively, p=0.041). [o]	⊕◯◯◯ VERY LOW	IMPORTANT
Ferritin <20µg/L									
1 [p]	observational studies [q]	very serious [r]	serious [s]	not serious [t]	serious [u]	none	Those who consumed vegetables once/day, compared to <once/day, were significantly more likely to have low iron stores (ferritin values <20µg/L) (Reg coefficient= -2.7, p=0.02). Although the results suggest a negative effect of feeding vegetables once/day, findings were inconsistent as this relationship was not seen for those eating vegetables more frequently (several times/day) compared to <once/day. Additionally, there was no significant relationship observed between the feeding frequency of vegetables and the likelihood of having VERY LOW iron stores (ferritin <15µg/L). [v]	⊕◯◯◯ VERY LOW	IMPORTANT
Anaemia									
1 [w]	observational studies [x]	very serious [y]	serious [z]	not serious [aa]	serious [bb]	none	Consumption of dark green vegetables was not associated with anaemia prevalence. For those who did not consume dark green vegetables in the past 24 hours, compared to those who did, the odds of having anaemia were 1.21 times greater (95%CI: 0.67-2.21, p=0.502). 55.6% (n=65) and 60.2% (n=53) of those who consumed dark green vegetables in the past 24 hours did not have anaemia, respectively. [cc]	⊕◯◯◯ VERY LOW	CRITICAL
% Stability and Change in Vegetable Consumption									

1 [dd]	observational studies [ee]	very serious [ff]	serious [gg]	not serious [hh]	serious [ii]	none	For boys, overall vegetable consumption at 18 months was positively associated with overall vegetable consumption at 36 months (spearman's rho=0.36) and at 7 years of age (spearman's rho=0.28). For girls, overall vegetable consumption at 18 months was positively associated with overall vegetable consumption at 36 months (spearman's rho=0.37) and at 7 years of age (spearman's rho=0.31). [jj]	⊕◯◯◯ VERY LOW	CRITICAL

Explanations

a. Ahmad A, Madanijah S, Dwiriani CM, & Kolopaking R. Complementary feeding practices and nutritional status of children 6-23 months old: formative study in Aceh, Indonesia. *Nutrition Research and Practice*. 2018; 12(6): 512-520. This cross-sectional study took place in 3 sub-districts in Aceh Besar District, Indonesia with 392 healthy children aged 6-23 months.
b. This was a cross-sectional observational study conducted between May-June 2016. Mother's provided detail on frequency of FV, ASF, and NPS consumption through 3 repeated 24hr recalls. Frequency for vegetable consumption was defined as ≥4 times/week, 1-3 times/ week, and never.
c. NIH tool for observational cohort and cross-sectional studies used. Overall, this study has a low risk of selection bias, information bias, and measurement bias and has been rated as good quality for internal validity of cross-sectional studies.
d. Inconsistencies cannot be determined given the lack of data (n=1 study) contributing to this outcome, thus we have downgraded the certainty of evidence for this outcome by 1 level.
e. This study provides direct evidence on population group, exposure, and outcomes of interest.
f. This study did not provide confidence intervals, limiting the ability to make a judgement about imprecision. It should also be noted that the sample size is relatively small (n=392). We have downgraded the certainty of evidence for this outcome by 1 level.
g. Study did not adjust for confounding variables.
h. Ntab B, Simondon KB, Milet J, Cissé B, Sokhna C, Boulanger D, et al. A Young Child Feeding Index Is Not Associated with Either Height-for-Age or Height Velocity in Rural Senegalese Children. *The Journal of Nutrition*. 2004; 135(3): 457-464. This cross-sectional study took place in rural Senegal with 543 children aged 6-42 months who were included in the Intermittent Preventive Malaria Treatment in children (IPTc) trial.
i. This was a cross-sectional observational study conducted between April-May 2003. Measurements of height and weight were obtained in September 2002 and again 7 months later during survey administration. Height and weight measurements were taken by trained staff using a measuring board and a Seca baby scale, respectively. Caregivers provided detail on frequency of FV and ASF consumption by 24hr recall. Additionally, the number of days each food group had been consumed during the preceding week was also assessed. Frequency was defined as 0-2 days/week and ≥ 3 days/week. A second definition for frequency was characterized as rare and frequent consumption but was never further defined in the paper.
j. NIH tool for observational cohort and cross-sectional studies used. Overall, this study presents some concerns for bias. Firstly, it should be noted that data on food consumption was attained from April to May 2003, while height increment was measured in 2002 and 2003. As food consumed in 2003 does not necessarily mean the child had consumed the same food back in 2002, this may lead to inaccuracies in the association between frequency of food consumption and height increment. Secondly, 7 months may not be long enough to cause a significant change in height. Caution should be taken when interpreting results from this study.
k. Inconsistencies cannot be determined given the lack of data (n=1 study) contributing to this outcome, thus we have downgraded the certainty of evidence for this outcome by 1 level.
l. This study provides direct evidence on population group, exposure, and outcomes of interest.
m. This study did not provide confidence intervals, limiting the ability to make a judgement about imprecision. It should also be noted that the sample size is relatively small (n=543). We have downgraded the certainty of evidence for this outcome by 1 level.
n. Adjusted model by child age.
o. Adjusted model by child age, sex, malaria study intervention group, maternal height, BMI, schooling, and number of children 5 years old.
p. Wandel M, Fagerli RA, Olsen PT, Borch-Iohnsen B, & Ek J. Iron status and weaning practices among Norwegian and immigrant infants. *Nutrition Research*. 1996; 16(2): 251-265. This cross-sectional study included 74 healthy infants born between 1 January 1991 and 31 March 1992, who were brought at 1 year of age to Fjell Health Clinic in Drammen, Norway.
q. This was a cross-sectional observational study conducted between January 1992-March 1992. Frequency of FV and ASF consumption were measured using a FFQ through dietary interviews by a primary health care nurse. The FFQ was constructed based on dietary results from a pilot study of a small sample of Norwegian and immigrant children belonging to the same health center. Frequency of consumption of FV was defined as several times per day, once per day, and less than once per day.
r. NIH tool for observational cohort and cross-sectional studies used. This study reported relatively little methodology and what methods are reported are unclear at times, in addition to a very small sample size. Because of this, it's extremely challenging to evaluate whether there was bias introduced or not and ultimately whether the study has internal validity.
s. Inconsistencies cannot be determined given the lack of data (n=1 study) contributing to this outcome, thus we have downgraded the certainty of evidence for this outcome by 1 level.
t. This study provides direct evidence on population group, exposure, and outcomes of interest.

u. This study did not provide confidence intervals, limiting the ability to make a judgement about imprecision. It should also be noted that the sample size is very small (n=74), however the certainty of the evidence has already been rated as VERY LOW based on the very poor-quality risk of bias rating and thus could not be further downgraded.
v. Study did not adjust for confounding variables.
w. Silva DG, Priore SE, & Franceschini SCC. Risk factors for anaemia in infants assisted by public health services: the importance of feeding practices and iron supplementation. *J Pediatr*. 2007; 83(2): 149-156. This cross-sectional study included 205 children aged 6-12 months living in the municipality of Vicosa, Minas Gerais, Brazil.
x. This was a cross-sectional observational study conducted between July 2002-April 2003. Frequency of FV, ASF, and NPS were measured using a FFQ and 24hr recall. Frequency of FV was defined as consumption of food item in previous 24hrs versus no consumption in previous 24hrs. Anaemia was defined as haemoglobin <11 g/dL.
y. NIH tool for observational cohort and cross-sectional studies used. Overall, this study has a poor to fair rating. Firstly, there is very limited detail provided on the methodology which makes it challenging to evaluate if there is internal validity. Some considerations to note include lack of sample size justification and no details are provided on participation rate of eligible participants. Furthermore, there is a small sample size with no details provided for baseline characteristics of study sample.
z. Inconsistencies cannot be determined given the lack of data (n=1 study) contributing to this outcome, thus we have downgraded the certainty of evidence for this outcome by 1 level.
aa. This study provides direct evidence on population group, exposure, and outcomes of interest.
bb. The confidence interval around the effect is relatively wide (OR=1.21; 95%CI: 0.67-2.21) and the sample size is small (n=205) thus we have downgraded the certainty of evidence for this outcome by 1 level. However, it should be noted that this study already has a certainty rating of VERY LOW based on the very poor-quality risk of bias rating.
cc. Study did not adjust for confounding variables.
dd. Bjelland M, Brantsaeter AL, Haugen M, Meltzer HM, Nystad W, & Andersen LF. Changes and tracking of fruit, vegetables and sugar-sweetened beverages intake from 18 months to 7 years in the Norwegian mother and child cohort study. *BMC Public Health*. 2013; 13:793. This longitudinal cohort study took place in a mixture or urban and rural locations in Norway with 9,025 mother infant pairs.
ee. This was a longitudinal cohort study using data from the nation-wide Norwegian Mother and Child Cohort Study, conducted between 1999-2008, with data collection occurring at three time points over 5.5 years (18 months, 36 months, and 7 years of age). Frequency of FV was measured using a questionnaire at the 3 time points. Frequency was categorized into 3 variables: low, medium, and high. For vegetables, low consumption was defined as ≤ 5 times/week, medium was defined as 5.1-7 times/week, and high was defined as >7 times/week.
ff. NIH tool for observational cohort and cross-sectional studies used. Overall, study has a rating of poor and lacks internal validity. Hard to determine whether information bias has occurred due to missing methods and detail in the report. For example, study population, exposure measures, and outcome measures are never clearly defined. Additionally, the study does not report accounting for any confounding variables, and it appears that measurement bias could have occurred. Results should be interpreted with caution.
gg. Inconsistencies cannot be determined given the lack of data (n=1 study) contributing to this outcome, thus we have downgraded the certainty of evidence for this outcome by 1 level.
hh. This study provides direct evidence on population group, exposure, and outcomes of interest.
ii. This study did not provide confidence intervals, limiting the ability to make a judgement about imprecision. It should be noted that the sample size is large (n=9,025 mother infant pairs), however given the lack of confidence intervals provided we have downgraded the certainty of evidence for this outcome by 1 level.
jj. Study did not adjust for confounding variables.

More frequent consumption of fruit compared to less frequent consumption of fruit

Certainty assessment							Impact	Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations			
Wasting									
1 [a]	observational studies [b]	not serious [c]	serious [d]	not serious [e]	serious [f]	none	Frequency of eating fruits was not associated with wasting. 19.9% (n=28), 26.1% (n=42), and 20% (n=18) of those who consumed fruits ≥3 times/week, 1-2 times/ week, and never were wasted, respectively (p=0.356 between groups). [g]	⊕◯◯◯ VERY LOW	CRITICAL
Underweight									
1 [a]	observational studies [b]	not serious [c]	serious [d]	not serious [e]	serious [f]	none	Frequency of eating fruits was not associated with being underweight. 26.2% (n=37), 26.1% (n=42), and 26.7% (n=24) of those who consumed fruits ≥3 times/week, 1-2 times/ week, and never were underweight, respectively (p=0.995 between groups). [g]	⊕◯◯◯ VERY LOW	CRITICAL
Stunting									
1 [a]	observational studies [b]	not serious [c]	serious [d]	not serious [e]	serious [f]	none	Frequency of eating fruits was not associated with stunting. 30.5% (n=43), 25.5% (n=41), and 27.8% (n=25) of those who consumed fruits ≥3 times/week, 1-2 times/ week, and never were stunted, respectively (p=0.623 between groups). [g]	⊕◯◯◯ VERY LOW	CRITICAL
Height-for-age Z-score									

1 [h]	observational studies [i]	serious [j]	serious [k]	not serious [l]	serious [m]	none	In breastfed children aged 9-23 months, those who consumed fruit 0-2 days/week and ≥ 3 days/week had a mean HAZ of -1.02 (SD=0.93) and -0.75 (SD=0.94), respectively. In age-adjusted models, mean HAZ was -1.04 (p=0.051) and -0.71 (p=0.059), respectively, demonstrating a trend that points to lower HAZ with less fruit consumption. [n]	⊕◯◯◯ VERY LOW	CRITICAL
Linear Growth									
1 [h]	observational studies [i]	serious [j]	serious [k]	not serious [l]	serious [m]	none	Fruit consumption was positively associated with linear growth in fully adjusted models (means: 7.9cm and 8.7cm height increment over the preceding 7 months for rare and frequent consumption, respectively, p=0.027). [o]	⊕◯◯◯ VERY LOW	IMPORTANT
Anaemia									
1 [p]	observational studies [q]	very serious [r]	serious [s]	not serious [t]	serious [u]	none	Fruit consumption was not significantly associated with anaemia prevalence when considering consumption in the past 24 hours. For those who did not consume fruit in the past 24 hours, compared to those who did, the odds of having anaemia were 1.24 times greater (95% CI: 0.25-6.75, p=0.537). 57.4% (n=113) and 62.5% (n=5) of those who consumed fruit in the past 24 hours and not, had anaemia, respectively. When looking at daily vs < than daily frequencies, fruit consumption was significantly associated with anaemia prevalence. In the adjusted model, those who consumed fruit < daily, compared to daily, had 1.88 (95% CI: 1.03-3.42, p=0.003) greater odds of having anaemia. 48.7% (n=55) and 68.5% (n=63) of those who consumed fruit daily and < daily had anaemia, respectively. [v]	⊕◯◯◯ VERY LOW	CRITICAL
% Stability and Change in Fruit Consumption									

1 [w]	observational studies [x]	very serious [y]	serious [z]	not serious [aa]	serious [bb]	none	For boys, overall fruit consumption at 18 months was positively associated with overall fruit consumption at 36 months (spearman's rho=0.36) and at 7 years of age (spearman's rho=0.23). For girls, overall fruit consumption at 18 months was positively associated with overall fruit consumption at 36 months (spearman's rho=0.36) and at 7 years of age (spearman's rho=0.24). [cc]	⊕◯◯◯ VERY LOW	CRITICAL

Explanations

a. Ahmad A, Madanijah S, Dwiriani CM, & Kolopaking R. Complementary feeding practices and nutritional status of children 6-23 months old: formative study in Aceh, Indonesia. *Nutrition Research and Practice*. 2018; 12(6): 512-520. This cross-sectional study took place in 3 sub-districts in Aceh Besar District, Indonesia with 392 healthy children aged 6-23 months.
b. This was a cross-sectional observational study conducted between May-June 2016. Mother's provided detail on frequency of FV, ASF, and NPS consumption through 3 repeated 24hr recalls. Frequency for fruit consumption was defined as ≥3 times/week, 1-2 times/ week, and never.
c. NIH tool for observational cohort and cross-sectional studies used. Overall, this study has a low risk of selection bias, information bias, and measurement bias and has been rated as good quality for internal validity of cross-sectional studies.
d. Inconsistencies cannot be determined given the lack of data (n=1 study) contributing to this outcome, thus we have downgraded the certainty of evidence for this outcome by 1 level.
e. This study provides direct evidence on population group, exposure, and outcomes of interest.
f. This study did not provide confidence intervals, limiting the ability to make a judgement about imprecision. It should also be noted that the sample size is relatively small (n=392). We have downgraded the certainty of evidence for this outcome by 1 level.
g. Study did not adjust for confounding variables.
h. Ntab B, Simondon KB, Milet J, Cissé B, Sokhna C, Boulanger D, et al. A Young Child Feeding Index Is Not Associated with Either Height-for-Age or Height Velocity in Rural Senegalese Children. *The Journal of Nutrition*. 2004; 135(3): 457-464. This cross-sectional study took place in rural Senegal with 543 children aged 6-42 months who were included in the Intermittent Preventive Malaria Treatment in children (IPTc) trial.
i. This was a cross-sectional observational study conducted between April-May 2003. Measurements of height and weight were obtained in September 2002 and again 7 months later during survey administration. Height and weight measurements were taken by trained staff using a measuring board and a Seca baby scale, respectively. Caregivers provided detail on frequency of FV and ASF consumption by 24hr recall. Additionally, the number of days each food group had been consumed during the preceding week was also assessed. Frequency was defined as 0-2 days/week and ≥ 3 days/week. A second definition for frequency was characterized as rare and frequent consumption but was never further defined in the paper.
j. NIH tool for observational cohort and cross-sectional studies used. Overall, this study presents some concerns for bias. Firstly, it should be noted that data on food consumption was attained from April to May 2003, while height increment was measured in 2002 and 2003. As food consumed in 2003 does not necessarily mean the child had consumed the same food back in 2002, this may lead to inaccuracies in the association between frequency of food consumption and height increment. Secondly, 7 months may not be long enough to cause a significant change in height. Caution should be taken when interpreting results from this study.
k. Inconsistencies cannot be determined given the lack of data (n=1 study) contributing to this outcome, thus we have downgraded the certainty of evidence for this outcome by 1 level.
l. This study provides direct evidence on population group, exposure, and outcomes of interest.
m. This study did not provide confidence intervals, limiting the ability to make a judgement about imprecision. It should also be noted that the sample size is relatively small (n=543). We have downgraded the certainty of evidence for this outcome by 1 level.
n. Adjusted model by child age.
o. Adjusted model by child age, sex, malaria study intervention group, maternal height, BMI, schooling, and number of children 5 years old.
p. Silva DG, Priore SE, & Franceschini SCC. Risk factors for anaemia in infants assisted by public health services: the importance of feeding practices and iron supplementation. *J Pediatr*. 2007; 83(2): 149-156. This cross-sectional study included 205 children aged 6-12 months living in the municipality of Vicosa, Minas Gerais, Brazil.
q. This was a cross-sectional observational study conducted between July 2002-April 2003. Frequency of FV, ASF, and NPS were measured using a FFQ and 24hr recall. Frequency of FV was defined as consumption of food item in previous 24hrs versus no consumption in previous 24hrs. Additionally for fruit, frequency was defined as consumed daily versus less than daily. Anaemia was defined as haemoglobin <11 g/dL.
r. NIH tool for observational cohort and cross-sectional studies used. Overall, this study has a poor to fair rating. Firstly, there is very limited detail provided on the methodology which makes it challenging to evaluate if there is internal validity. Some considerations to note include lack of sample size justification and no details are provided on participation rate of eligible participants. Furthermore, there is a small sample size with no details provided for baseline characteristics of study sample.
s. Inconsistencies cannot be determined given the lack of data (n=1 study) contributing to this outcome, thus we have downgraded the certainty of evidence for this outcome by 1 level.

t. This study provides direct evidence on population group, exposure, and outcomes of interest.
u. The confidence interval around the effect is relatively wide (OR=1.88; 95%CI: 1.03-3.42) and the sample size is small (n=205) thus we have downgraded the certainty of evidence for this outcome by 1 level. However, it should be noted that this study already has a certainty rating of VERY LOW based on the very poor-quality risk of bias rating.
v. Adjusted model by family income per capita and consumption of medicated iron supplements.
w. Bjelland M, Brantsaeter AL, Haugen M, Meltzer HM, Nystad W, & Andersen LF. Changes and tracking of fruit, vegetables and sugar-sweetened beverages intake from 18 months to 7 years in the Norwegian mother and child cohort study. *BMC Public Health*. 2013; 13:793. This longitudinal cohort study took place in a mixture or urban and rural locations in Norway with 9,025 mother infant pairs.
x. This was a longitudinal cohort study using data from the nation-wide Norwegian Mother and Child Cohort Study, conducted between 1999-2008, with data collection occurring at three time points over 5.5 years (18 months, 36 months, and 7 years of age). Frequency of FV was measured using a questionnaire at the 3 time points. Frequency was categorized into 3 variables: low, medium, and high. For fruits, low consumption was defined as ≤ 5 times/week, medium was defined as 5.1-13.9 times/week, and high was defined as >14 times/week.
y. NIH tool for observational cohort and cross-sectional studies used. Overall, study has a rating of poor and lacks internal validity. Hard to determine whether information bias has occurred due to missing methods and detail in the report. For example, study population, exposure measures, and outcome measures are never clearly defined. Additionally, the study does not report accounting for any confounding variables, and it appears that measurement bias could have occurred. Results should be interpreted with caution.
z. Inconsistencies cannot be determined given the lack of data (n=1 study) contributing to this outcome, thus we have downgraded the certainty of evidence for this outcome by 1 level.
aa. This study provides direct evidence on population group, exposure, and outcomes of interest.
bb. This study did not provide confidence intervals, limiting the ability to make a judgement about imprecision. It should be noted that the sample size is large (n=9,025 mother infant pairs), however given the lack of confidence intervals provided we have downgraded the certainty of evidence for this outcome by 1 level.
cc. Study did not adjust for confounding variables.

More frequent consumption of fruit and vegetables compared to less frequent consumption of fruit and vegetables

Certainty assessment							Impact	Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations			
Stunting									
1 [a]	observational studies [b]	very serious [c]	serious [d]	not serious [e]	serious [f]	none	Frequency of eating vegetables and fruits was associated with stunting prevalence. For those who consumed vegetables and fruits weekly, monthly, and < once per month or none, compared to daily, the odds of being stunted were 1.739 (p=0.00), 1.698 (p=0.03), and 1.768 (p=0.00) times greater, respectively. [g]	⊕◯◯◯ VERY LOW	CRITICAL
Underweight									
1 [a]	observational studies [b]	very serious [c]	serious [d]	not serious [e]	serious [f]	none	Frequency of eating vegetables and fruits was associated with underweight prevalence. For those who consumed vegetables and fruits weekly, monthly, and < once per month or none, compared to daily, the odds of being underweight were 1.908 (p=0.00), 1.566 (p=0.10), and 1.478 (p=0.01) times greater, respectively. [g]	⊕◯◯◯ VERY LOW	CRITICAL

Explanations

a. Chang S, Chen C, He W, & Wang Y. Analysis on the changes of nutritional status among Chinese infants and young children -- the improvement of complementary feeding. *Journal of Hygiene Research.* 2007; 36(2):207-9. This cross-sectional study took place in a mixture of urban and rural locations in China with 13,107 infants aged 6-24 months old.
b. This was a cross-sectional observational study using data from the China Food and Nutrition Surveillance System (CFNSS). Data collection occurred between 1992-2005. Frequency of FV consumption was defined as daily, weekly, monthly, less than once per month, or none.
c. NIH tool for observational cohort and cross-sectional studies used. This study lacks internal validity. Firstly, the paper does not report a methods section. Thus, no detail is provided on the tools used to measure the exposure or outcome. The sample size is not clearly stated, and thus had to be inferred from tables. In the analysis, the reference arm is not reported within the tables, once again requiring the reader to infer this information from the text. Results should be interpreted with extreme caution.
d. Inconsistencies cannot be determined given the lack of data (n=1 study) contributing to this outcome, thus we have downgraded the certainty of evidence for this outcome by 1 level.
e. This study provides direct evidence on population group, exposure, and outcomes of interest.
f. This study did not provide confidence intervals, limiting the ability to make a judgement about imprecision. It should be noted that the sample size is very large (n=13,107), however given the lack of confidence intervals provided we have downgraded the certainty of evidence for this outcome by 1 level.
g. Study did not adjust for confounding variables.

Dietary diversity: Pulses, nuts and seeds

Certainty assessment							Impact	Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations			
Wasting									
1 [a]	observational studies [b]	not serious [c]	serious [d]	not serious [e]	serious [f]	none	Frequency of eating legumes was not associated with wasting. 25.9% (n=15), 24% (n=37), and 20% (n=36) of those who consumed legumes ≥3 times/week, 1-2 times/ week, and never were wasted, respectively (p=0.542 between groups). [g]	⨁◯◯◯ VERY LOW	CRITICAL
Underweight									
1 [a]	observational studies [b]	not serious [c]	serious [d]	not serious [e]	serious [f]	none	Frequency of eating legumes was not associated with being underweight. 36.2% (n=21), 25.3% (n=39), and 23.9% (n=43) of those who consumed legumes ≥3 times/week, 1-2 times/ week, and never were underweight, respectively (p=0.174 between groups). [g]	⨁◯◯◯ VERY LOW	CRITICAL
Stunting									
1 [a]	observational studies [b]	not serious [c]	serious [d]	not serious [e]	serious [f]	none	Frequency of eating legumes was not associated with stunting. 32.8% (n=19), 27.9% (n=43), and 26.1% (n=47) of those who consumed legumes ≥3 times/week, 1-2 times/ week, and never were stunted, respectively (p=0.618 between groups). [g]	⨁◯◯◯ VERY LOW	CRITICAL
Anaemia									

1 [h]	observational studies [i]	very serious [j]	serious [k]	not serious [l]	serious [m]	none	Consumption of beans was not significantly associated with anaemia prevalence. For those who consumed beans < daily, compared to daily, the odds of having anaemia were 0.8 times less (95% CI: 0.36-1.78, p=0.550). 58.5% (n=100) and 52.9% (n=18) of those who consumed beans daily and <daily had anaemia, respectively. [n]	⊕◯◯◯ VERY LOW	CRITICAL

Explanations

a. Ahmad A, Madanijah S, Dwiriani CM, & Kolopaking R. Complementary feeding practices and nutritional status of children 6-23 months old: formative study in Aceh, Indonesia. *Nutrition Research and Practice*. 2018; 12(6): 512-520. This cross-sectional study took place in 3 sub-districts in Aceh Besar District, Indonesia with 392 healthy children aged 6-23 months.
b. This was a cross-sectional observational study conducted between May-June 2016. Mother's provided detail on frequency of FV, ASF, and NPS consumption through 3 repeated 24hr recalls. Frequency for NPS consumption was defined as ≥3 times/week, 1-2 times/ week, and never.
c. NIH tool for observational cohort and cross-sectional studies used. Overall, this study has a low risk of selection bias, information bias, and measurement bias and has been rated as good quality for internal validity of cross-sectional studies.
d. Inconsistencies cannot be determined given the lack of data (n=1 study) contributing to this outcome, thus we have downgraded the certainty of evidence for this outcome by 1 level.
e. This study provides direct evidence on population group, exposure, and outcomes of interest.
f. This study did not provide confidence intervals, limiting the ability to make a judgement about imprecision. It should also be noted that the sample size is relatively small (n=392). We have downgraded the certainty of evidence for this outcome by 1 level.
g. Study did not adjust for confounding variables.
h. Silva DG, Priore SE, & Franceschini SCC. Risk factors for anaemia in infants assisted by public health services: the importance of feeding practices and iron supplementation. *J Pediatr*. 2007; 83(2): 149-156. This cross-sectional study included 205 children aged 6-12 months living in the municipality of Vicosa, Minas Gerais, Brazil.
i. This was a cross-sectional observational study conducted between July 2002-April 2003. Frequency of FV, ASF, and NPS were measured using a FFQ and 24hr recall. Frequency of NPS consumption was defined as consumed daily versus less than daily. Anaemia was defined as haemoglobin <11 g/dL.
j. NIH tool for observational cohort and cross-sectional studies used. Overall, this study has a poor to fair rating. Firstly, there is very limited detail provided on the methodology which makes it challenging to evaluate if there is internal validity. Some considerations to note include lack of sample size justification and no details are provided on participation rate of eligible participants. Furthermore, there is a small sample size with no details provided for baseline characteristics of study sample.
k. Inconsistencies cannot be determined given the lack of data (n=1 study) contributing to this outcome, thus we have downgraded the certainty of evidence for this outcome by 1 level.
l. This study provides direct evidence on population group, exposure, and outcomes of interest.
m. The confidence interval around the effect is relatively wide (OR=0.80; 95%CI: 0.36-1,78) and the sample size is small (n=205) thus we have downgraded the certainty of evidence for this outcome by 1 level. However, it should be noted that this study already has a certainty rating of VERY LOW based on the very poor-quality risk of bias rating.
n. Study did not adjust for confounding variables.

Unhealthy foods and beverages: Foods high in sugar, salt, and unhealthy fats

Certainty assessment							Impact	Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations			
Mean BMI/BMI z-scores or change in BMI/BMI z-scores in children aged < 2 years									
3	observational studies	very serious[a]	not serious[b]	not serious[c]	not serious[d]	none	Increased BMI (1 study, n = 666): candies, ANOVA F= 3.23, P = 0.03 (Moore 2019); Different effects (0 studies); No significant association (2 studies, n = 1105); 'extra foods' β = -0.10, 95%CI = -0.30, 0.11, p = 0.36 (Garden 2011); sweetened foods BMIz mean difference 0.03 95% CI -0.12, 0.19 (Santorelli 2014)	⨁⨁◯◯ LOW	CRITICAL
Mean BMI/BMI z-scores or change in BMI/BMI z-scores in children aged 2- <5 years at exposure									
6	observational studies	very serious[e]	not serious[b]	not serious[c]	not serious[d]	none	Increased BMI (3 studies, n = 11639); Fast foods aRR: 1.38, 95% CI 1.13, 1.67, P < 0.01 (Emond 2020); High fat foods β = 0.021, 95% CI 0.014, 0.029 P < 0.001 (Millar 2014); Sugar-added to foods 2 < 6 y: boys P = 0.005, girls P = 0.03; 6 < 10 y: boys P = 0.001, girls P > 0.05 (Russo 2018); Different effects (1 study, n = 1175); Ultra-processed food intake at 4 y β=0·028; 95% CI = 0.006, 0.051, intake at 7 y β=0·014; 95% CI = -0.007, 0.036 (Vedovato 2020); No significant association (2 studies, n = 695); Added sugar β = -0.001, SE = 0.010, P = 0.9 (Buyken 2011); Ultra-processed foods β = 0.05, 95% CI = -0.04, 0.15, P = 0.282 (Costa 2019)	⨁⨁◯◯ LOW	CRITICAL
Mean BMI/BMI z-scores or change in BMI/BMI z-scores in children aged 5 - ≤ 10 years at exposure									
4	observational studies	very serious[f]	not serious[b]	not serious[c]	not serious[d]	none	Increased BMI (1 study, n=1414); Snack foods β = 0.71, 95% CI = 0.14, 1.28 (Alviso-Orellana 2018); Different effects (0 studies); No significant association (3 studies, n = 5797); High fat foods β = -0.02, CI = -0.06, 0.03, P = 0.409 (Carlson 2012); Other sugars β = 0.16, SE = 0.10, P > 0.05 (Hur 2015); Fast foods P > 0.05 (parameter estimate from a cross-lagged autoregressive model (Jackson 2017)	⨁⨁◯◯ LOW	CRITICAL
Prevalence of overweight and obesity or prevalence of obesity only in children aged < 2 years (assessed with: %)									

1	observational studies	extremely serious[g]	not serious[b]	not serious[c]	not serious[d]	none	Increased odds of overweight/obesity (0 studies); Different effects (0 studies); No significant association (1 study, n=1871); Fast foods aOR = 1.14, 95% CI = 0.77, 1.67; snack consumption aOR = 0.71, 95% CI = 0.52, 0.98 (Wijga et al 2010)	⊕◯◯◯ VERY LOW	CRITICAL
Prevalence of overweight and obesity or prevalence of obesity only in children aged 2- < 5 years (assessed with: %)									
2	observational studies	very serious[h]	not serious[b]	not serious[c]	not serious[d]	none	Increased overweight/obesity (0 studies); Different effects (0 studies); No significant association (2 studies, n = 4680); Sweet and savoury snacks aOR = 0.76, 95% CI = 0.41, 1.40, P > 0.05 (De Coen 2014); High fat foods boys: aOR = 0.85, 95% CI = 0.6, 1.19; girls: aOR = 0.97, 95% CI = 0.7, 1.35 (Zulfiqar 2019)	⊕⊕◯◯ LOW	CRITICAL
Prevalence of overweight and obesity or prevalence of obesity only in children aged 5 - ≤ 10 years (assessed with: %)									
2	observational studies	very serious[i]	not serious[j]	not serious[c]	not serious[d]	none	Increased overweight/obesity (0 studies); Different effects (1 study, n=2755); savoury snacks never vs everyday aOR = 0.27, 95% CI = 0.10, 0.72, P <0.01; fast food never vs everyday aOR = 0.91, 95% CI = 0.19, 4.31, P >0.05 (Bel-Serrat 2019); No significant association (1 study, n = 1414); Savoury snacks aRR = 1.43 0.78, 2.69 (Alviso-Orellana 2018)	⊕⊕◯◯ LOW	CRITICAL
Percent body fat ≤ 10 years									
4	observational studies	extremely serious[k]	not serious[b]	not serious[c]	not serious[d]	none	Increased percent body fat (1 study, n = 3514); Ultra-processed foods β = 0.05, 95% CI = 0.04, 0.06, P < 0.001 (NOTE fat mass index, not % body fat) (Costa 2020); Different effects (0 studies); No significant association (3 studies, n = 1239); Added sugar β = 0.048, SE = 0.046, P = 0.3 (Buyken 2011); High fat foods β = -0.38, CI = -0.81, 0.05, P = 0.081 (Carlson 2012); Other sugars β = 0.83, SE = 0.72, P > 0.05 (Hur 2015)	⊕◯◯◯ VERY LOW	CRITICAL
Blood lipids									
4	observational studies	very serious [l]	not serious[b]	not serious[c]	not serious[d]	none	Increased risk (2 studies, n = 613); Ultra-processed foods (UPF): change in total cholesterol, β = 0.430 95% CI 0.008, 0.853 P = 0.046; change LDL-C, β = 0.369 95% CI 0.005, 0.733 P = 0.047; change non-HDL-C, β = 0.319, 95% CI 0.059, 0.697 P = 0.098; change triglycerides, β = 0.465, 95% CI 0.955, 0.025 P = 0.06; change HDL-C, β = 0.125, 95% CI 0.026, 0.277 P = 0.105 (Rauber 2015); UPF intake: increased total serum cholesterol β 0·22 mmol/l; 95% CI 0·04, 0·39 and TAG 0·11 mmol/l, 95% CI 0·01, 0·20 (Leffa 2020). No significant association (2 studies, n = 2172); No association between SSB at 13 m and systolic or diastolic blood pressure, pulse wave velocity, blood lipds or insulin at 6 years of age (Leermakers 2015); No association between mean SSB intake and changes in HDL cholesterol or TG changes (Van Rompay 2015)	⊕⊕◯◯ LOW	CRITICAL

Glucose/insulin									
1	observational studies	very serious [m]	not serious [j]	not serious [c]	serious [n]	none	Increased risks (0 studies); No significant association (1 study, n = 315); No significant association between ultraprocessed food consumption and glucose profiles (Costa 2019)	⊕◯◯◯ VERY LOW	CRITICAL
Metabolic syndrome									
1	observational studies	extremely serious [o]	not serious [b]	not serious [c]	serious [n]	none	Increased risks (0 studies); No significant association (1 study, n = 605); No significant association between beverage sugar or non-beverage sugars and metabolic syndrome score (Hur 2015)	⊕◯◯◯ VERY LOW	CRITICAL
Displacement of healthy food items/breastmilk									
3	observational studies	very serious [p]	not serious [b]	not serious [c]	not serious [d]	none	Increased risks (1 study, n=875); energy-providing liquid intake associated with significantly lower energy intake from formula from 2 - 5 mo (Schiess 2010); Different effects (1 study, n = 515); No significant association between SSB intake and fruit and vegetable intake but weak inverse correlation between SSB intake and milk/alternatives at 2 y (r = −0.11, P = 0.015) and 5 y (r = −0.11, P = 0.012) (Byrne 2018); No significant association (1 study, n = 1252); Change in high-caloric drink and change in energy-dense sweet consumption was not significantly correlated with change in fruit or change in vegetable consumption (Bayer 2014).	⊕⊕◯◯ LOW	CRITICAL
Dietary quality and diversity indicators (assessed with: Healthy eating index/healthy dietary assessment score)									
4	observational studies	very serious [q]	not serious [b]	not serious [c]	not serious [d]	none	Increased diversity (1 study, n = 100); increased 100% fruit juice associated with increase healthy eating index (Wan 2020); No significant association (0 studies); Decreased diversity (3 studies, n= 7986); Sugar-added foods associated with lower healthy dietary adherence score (P < 0.001 for trend) (Russo 2018); Daily high energy-dense food intake at 2 y associated with lower healthy eating index score at 4 y, IRR = 0·56, 95% CI 0·41, 0·77 (Vilela 2014); Lower consumption SSB associated with higher odds of healthy eating index (AOR 2.7, 95% CI 1.6 to 4.3, P < 0.001) (Woo 2020)	⊕⊕◯◯ LOW	CRITICAL

CI: confidence interval

Explanations

a. Risk of bias was moderate in 2 studies (Garden 2011, Moore 2019), serious in 1 study (Santorelli 2014). Downrated by 2 levels for risk of bias due to non-randomization in observational studies leading to confounding and selection bias
b. Not downrated for inconsistency but note that interventions and comparators were different across studies
c. Not downrated as study populations, exposures and comparators were relevant to review question, although no studies were from low income country populations
d. Not downrated as no evidence of imprecision (i.e. not wide confidence intervals, small sample size or low number of events)

e. Risk of bias was moderate in 5 studies (Costa 2019, Emond 2020, Millar 2014, Russo 2018, Vedovato 2020) and serious in 1 study (Buyken 2011). Downrated by 2 levels for risk of bias due to non-randomization in observational studies leading to confounding and selection bias.
f. Risk of bias was moderate in 2 studies (Alviso-Orellana 2018, Jackson 2017) and serious in 2 studies (Carlson 2012, Hur 2015). Downrated by 2 levels for risk of bias due to non-randomization in observational studies leading to confounding and selection bias.
g. Risk of bias was serious in all studies (Wijga 2010). Downrated by 2 levels for non-randomization in observational studies leading to confounding and selection bias, and 1 level further due to body of evidence all from studies with serious risk of bias
h. Risk of bias was moderate in 1 study and serious in 1 study. Downrated by 2 levels for inherent risk of bias due to non-randomisation in observational studies.
i. Risk of bias was moderate in 1 study (Alviso-Orellana 2018) and serious in 1 study (Bel-Serrat 2019). Downrated by 2 levels for risk of bias due to non-randomization leading to confounding and selection bias.
j. Not downrated as only 1 study
k. Risk of bias was moderate for 1 study (Leffa 2020) and serious for 3 studies (Leermakers 2015, Rauber 2015, Van Rompey 2015). Downrated by 2 levels for risk of bias due to non-randomization leading to confounding and selection bias and 1 level further due to majority of the body of evidence had serious risk of bias.
l. Risk of bias was moderate in 3 studies, (Leermakers 2015, Leffa 2020, Rauber 2015) and serious in 1 study (Van Rompay 2015). Downrated by 2 levels for risk of bias due to non-randomization in observational studies. leading to confounding and selection bias.
m. Risk of bias was moderate (Costa 2019). Downrated by 2 levels for risk of bias due to non-randomization in observational studies leading to confounding and selection bias.
n. Downrated by 1 level because of uncertainty around the probability of risk occurring within the included sample.
o. Risk of bias was serious for all studies (Hur 2015). Downrated by 2 levels for risk of bias due to non-randomisation in observational studies and 1 further level because body of evidence was from studies with serious risk of bias.
p. Risk of bias was moderate in 2 studies (Byrne 2018, Scheiss 2010) and serious in 1 study (Bayer 2014). Downrated by 2 levels for risk of bias due to non-randomization in observational studies leading to confounding and selection bias
q. Risk of bias was moderate in 3 studies (Russo 2018, Vilela 2014, Woo 2020) and serious in 1 study (Wan 2020). Downrated by 2 levels for risk of bias due to non-randomization in observational studies leading to confounding and selection bias

Unhealthy foods and beverages: Sugar-sweetened beverages

Certainty assessment							Impact	Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations			
Mean BMI/BMI z-score or change in BMI/BMI z-score in children < 2 years at exposure									
3	observational studies	very serious[a]	not serious[b]	not serious[c]	not serious[d]	none	Increased BMI (0 studies); Different effects (2 studies, n = 3138); different effect in boys vs girls (Quah 2019); different effects by age of follow up: from age 18 m to 6 y, β = 0.06, 95% CI = -0.20, 0.31, P = 0.67 and from age 5 y to 6 y β = 0.34, 95% CI = 0.11, 0.58, P = 0.004 (Leermakers 2015); No significant association (1 study, n = 743): Mean BMIz diff -0.10, 95% CI -0·36, 0·16 (Santorelli 2014) [e,f]	⊕⊕◯◯ LOW	CRITICAL
Mean BMI/BMI z-score or change in BMI/BMI z-score in children 2 - < 5 years at exposure									
6	observational studies	very serious[g]	not serious[b]	not serious[c]	not serious[d]	none	Increased BMI (2 studies, n = 4792): β = 0.05, 95% CI = 0.022, 0.079, P = 0.001 (Marshall 2019); β = 0.017, 95% CI = 0.007, 0.027, P < 0.01 (Millar 2014); Different effects (0 studies); No significant association (4 studies, n = 2163): β -0.01, 95% CI -0.05 to 0.04, P = 0.852 (Costa 2019); ANCOVA P = 0.0626 (Hasnain 2014); β = -0.01, SE = 0.02, P = 0.50 (Newby 2004); P > 0.05 (Byrne 2018) [e,f]	⊕⊕◯◯ LOW	CRITICAL
Mean BMI/BMI z-score or change in BMI/BMI z-score in children 5- ≤ 10 years at exposure									
10	observational studies	very serious[h]	not serious[b]	not serious[c]	not serious[d]	none	Increased BMI (2 studies, n=158) β = 0·74, 95% CI = 0·15, 1.33 (Alviso-Orellana 2018); β = 0.10, SE = 0.03, P = 0.003 (Zheng 2015); Different effects (1 study, n = 2371); Positive association for sodas (β = 0.011, SE = 0.005, P < 0.05) but not other SSBs (β = 0.009, SE = 0.007, P > 0.05); No significant association (7 studies, n=6726); β = 0.11, CI = -0.03, 0.25 (Carlson 2012); ANOVA p > 0.05 (Fiorito, 2009); β = -0.02, SE = 0.03, P > 0.05 (Hur 2015); P > 0.05 (parameter estimate from a cross-lagged autoregressive model, Jackson 2017); intake at 6 y and BMI change 6-9 y β = -0.014, 95% CI = -0.063, 0.035, P = 0.55 (Jensen 2013); boys β = -0.037, SE = 0.019, P = 0.707, girls β = 0.086, SE = 0.027, P = 0.450 (Laurson 2008); at 9 y > 1 serve β= 1.42, SE 0.68, P = 0.29 (Zheng 2014) [e,f]	⊕⊕◯◯ LOW	CRITICAL
Mean change in BMI/BMI z-score in children 5- ≤ 10 years at exposure									
1	randomised trials	serious[i]	not serious[j]	serious[k]	not serious[d]	none	1 study (n=1987) BMI change: β = 0.02, 95% CI 0.00, 0.03 with each glass of sugar-containing beverage consumption/day.	⊕⊕◯◯ LOW	CRITICAL

Prevalence of overweight and obesity or prevalence of obesity only in children aged < 2 years (assessed with: %)									
6	observational studies	extremely serious[l]	not serious[b]	not serious[c]	not serious[d]	none	Increased overweight/obesity (3 studies, n = 3372); aOR = 2.99, 95% CI: 1.27, 7.00 (Cantoral 2016); ≥3 times/week aOR = 2.00, 95% CI = 1.02, 3.90 (Pan 2014); aOR = 1.6, CI = 1.04, 1.93, P < 0.01 (Wang 2013); Different effects (2 studies, n = 7567); At 2 y no significant association, at 5 y aOR = 2.3, 95% CI 1.4, 3.7 (Flores 2013); at 18 m no significant association, at 5 y RR = 1.10, 95% CI = 0.67, 1.81, P = 0.204 (Quah 2019); No significant association (1 study, n = 1871); aOR = 0.91, 95% CI = 0.44, 1.88 (Wijga 2010)	⊕◯◯◯ VERY LOW	CRITICAL
Prevalence of overweight and obesity or prevalence of obesity only in children aged 2-< 5 years (assessed with: %)									
5	observational studies	very serious[m]	not serious[b]	not serious[c]	not serious[d]	none	Increased overweight/obesity (1 study, n= 473); aOR = 1.92, 95% CI = 1.19, 3.11, P ≤ 0.01); Different effects (1 study, n = 2986); overweight/obesity no association, obesity only aOR= 1.65, 95% CI = 1.12, 2.44, P = 0.01 (Macintyre 2018); No significant association (3 studies, 17083); Not significant (no estimate, Dubois 2007); aOR = 1.3, 95% CI = 0.8, 2.1 (Welsh 2005); Boys aOR = 1.01, 95% CI = 0.8, 1.29; girls aOR = 1.08, 95% CI = 0.87, 1.35 (Zulfiqar 2019)	⊕⊕◯◯ LOW	CRITICAL
Prevalence of overweight and obesity or prevalence of obesity only in children aged 5- ≤ 10 years (assessed with: %)									
3	observational studies	very serious[g]	not serious[b]	not serious[c]	not serious[d]	none	Increased overweight/obesity (2 studies, n = 1668): aRR= 2·12, 95% CI 1·05, 4·28 (Alviso-Orellana 2018); aOR = 1.04, 95% CI = 1.01, 1.07, P < 0.05 (Lim 2008); Different effects (0 studies); No significant association (1 study, n = 1250) Overweight only aOR = 1.29, 95% CI = 0.84, 1.96, p = 0.246; obese only aOR = 1.57, 95% CI = 0.82, 3.03, p = 0.177 (Traub 2018)	⊕⊕◯◯ LOW	CRITICAL
Prevalence of overweight and obesity or prevalence of obesity only in children aged 5 - ≤10 years (assessed with: %)									
1	randomised trials	serious[i]	not serious[i]	serious[k]	not serious[d]	none	Increased risk (1 study, n = 1987); Each glass/day of sugar-sweetened beverage consumption increased the odds of obesity (aOR 1·22; 95% CI 1·04, 1·44, p= 0.014) but not overweight (P=0.83)	⊕⊕◯◯ LOW	CRITICAL
Mean percent body fat in children aged ≤ 10 years									
7	observational studies	very serious[n]	not serious[b]	not serious[c]	not serious[d]	none	Increased % body fat (3 studies, n = 578); ANOVA P < 0.01 (Fiorito 2009); β = 1.40, CI = 0.09, 2.72, P = 0.036 (Carlson 2012); β = 1.04, SE = 0.32, P = 0.001 (Zheng 2015); Different effects (0 studies); No significant association (4 studies, n = 3436); ANCOVA P = 0.929 (Hasnain 2014); β = 0.02, SE = 0.21, P > 0.05 (Hur 2015); β = -0.15, 95% CI = -0.54, 0.24, P = 0.45 (Johnson 2007); Boys β= 0.05, 95% CI = -0.11, 0.20, P =0.53; girls β = 0.09, 95% CI = -0.06, 0.23, P = 0.25 (Leermakers 2015). [o]	⊕⊕◯◯ LOW	CRITICAL

CI: confidence interval

Explanations

a. Risk of bias was moderate in 1 study (Santorelli, 2014) and serious in 2 studies (Quah 2019, Leermakers 2015). Downrated by 2 levels due to non-randomization leading to confounding and selection bias.
b. Not downrated for inconsistency but note that interventions and comparators were were different across studies
c. Not downrated as study populations, exposures and comparators were relevant to review question, although no studies were from low income country populations
d. Not downrated as no evidence of imprecision (i.e. not wide confidence intervals, small sample size or low number of events)
e. Meta-analysis of 3 studies across different age groups: BMI change effect size 0.01 (-0.00, 0.02) (Jensen 2013, Newby 2004, Laurson 2008)
f. Meta-analysis of 3 studies across different age groups: BMI z-score change effect size 0.10 95% CI -0.11 - 0.31 (Carlson 2012, Marshall 2019, Quah 2019)
g. Risk of bias was moderate for all studies. Downrated by 2 levels due to non-randomization in observational studies leading to confounding and selection bias.
h. Risk of bias was moderate in 5 studies (Alviso-Orellana 2018, Jackson 2017, Jensen 2013, Laurson 2008, Zheng 2015) and serious in 5 studies (Carlson 2012, Fiorito 2009, Hur 2015, Striegel-Moore 2006, Zheng 2015). Downrated by 2 levels due to non-randomization in observational studies leading to confounding and selection bias
i. Some concerns due to missing outcome data and bias in selection of reported result
j. Not downrated as only 1 study
k. Downrated by 1 level as sugar-sweetened beverage consumption was a secondary outcome of the RCT
l. Risk of bias was serious for all 5 studies. Downrated by 2 levels for inherent risk of bias due to non-randomisation and 1 further level due to serious risk of bias in all studies.
m. Risk of bias was moderate in 4 studies (Dubois 2007, Macintyre 2018, Welsh 2005, Zulfiqar 2019) and serious in 1 study (DeCoen 2014)
n. Risk of bias was moderate in 3 studies (Hasnain 2014, Johnson 2007, Leermakers 2015), serious in 4 studies (Carlson 2012, Fiorito 2009, Hur 2015, Zheng 2015). Downrated by 2 levels for risk of bias due to non-randomization in observational studies leading to confounding and selection bias
o. Meta-analysis of 3 studies (Carlson 2012; Hur 2015; Zheng 2015): pooled effect estimate β=1.86 [0.38, 3.34].

Unhealthy foods and beverages: beverages with non-sugar sweeteners

Certainty assessment							Impact	Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations			
Mean BMI/BMI z-score or change in BMI/BMI z-score in children 2- < 5 years at exposure									
2	observational studies	very serious[a]	not serious[b]	not serious[c]	not serious[d]	none	Increased BMI (0 studies); Different effects (0 studies); No significant association (2 studies, n=1443): ANCOVA p = 0.444 (Hasnain 2014); β = 0.01, SE = 0.02, p = 0.83 (Newby 2004)	⊕⊕◯◯ LOW	CRITICAL
Mean BMI/BMI z-score or change in BMI/BMI z-score in children 5 - ≤ 10 years at exposure									
2	observational studies	extremely serious[e]	not serious[b]	not serious[c]	not serious[d]	none	Increased BMI (0 studies); Different effects (0 studies); No significant association (1 study, n=2371); β = 0.01, SE = 0.013, p > 0.05 (Striegel-Moore 2006); Decreased BMI (1 study, n= 158); β = -0.20, SE = 0.07, p = 0.01 (Zheng 2015)	⊕◯◯◯ VERY LOW	CRITICAL
Prevalence of overweight and obesity or prevalence of obesity only in children aged 2-< 5 years (assessed with: %)									
1	observational studies	very serious[a]	not serious[f]	not serious[c]	not serious[d]	none	Different effects (1 study, n=2986); No signficant association with overweight/obesity (aOR = 0.85, 95% CI 0.63, 1.15 p = 0.85) but increased risk of obesity (aOR = 1.57, 95% CI = 1.05, 2.36, p = 0.03) (Macintyre 2018)	⊕⊕◯◯ LOW	CRITICAL
Mean percent body fat in children aged ≤ 10 years (assessed with: %)									
3	observational studies	very serious[g]	not serious[b]	not serious[c]	not serious[d]	none	No significant association (1 study, n=98): ANCOVA p = 0.584 (Hasnain 2014); **Negative** association (1 study, n=362) β = 0.26, 95% CI = -0.004, 0.52, p = 0.05 (Johnson 2007); Negative association (1 study, n=158) β = -1.41, SE = 0.70, p = 0.046, (Zheng 2015)	⊕⊕◯◯ LOW	CRITICAL

CI: confidence interval

Explanations

a. Risk of bias was moderate for all studies. Downrated by 2 levels for inherent bias due to non-randomization in observational studies leading to confounding and selection bias.
b. Not downrated for inconsistency but note that there were differences between interventions and comparators across studies.
c. Not downrated as study populations, exposures and comparators were relevant to review question, although no studies were from low income country populations

d. Not downrated as no evidence of imprecision (i.e. not wide confidence intervals, small sample size or low number of events)
e. Risk of bias was serious for all studies. Downrated by 2 levels for risk of bias due to non-randomization (confounding and selection bias) and 1 further level for serious risk of bias across the body of evidence.
f. Not downrated as only one study
g. Risk of bias was moderate in 2 studies (Hasnain 2014, Johnson 2007) and serious in 1 study (Zheng 2015). Downrated by 2 levels due to risk of bias due to non-randomization in observational studies leading to confounding and selection bias.

Unhealthy foods and beverages: 100% fruit juice

Certainty assessment							Impact	Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations			
Mean BMI/BMI z-score or change in BMI/BMI z-score in children < 2 years at exposure									
1	observational studies	very serious[a]	not serious[b]	not serious[c]	not serious[d]	none	Increased BMI (0 studies); No significant association (1 study, n= 1038) β=0.30, 95% CI = −0.01, 0.61 at 2.1 y follow-up; β=0. 0.27, 95% CI = −0.05, 0.59 at 6.7 y follow-up (Sonneville 2015) [e]	⨁⨁◯◯ LOW[a]	CRITICAL
Mean BMI/BMI z-score or change in BMI/BMI z-score in children 2- < 5 years at exposure									
5	observational studies	very serious[f]	not serious[g]	not serious[c]	not serious[d]	none	Increased BMI (0 studies); Different effects (1 study, n= 6250): Mean BMI z-score change 0.282 (SE 0.028) vs 0.030 (SE 0.037), p = 0.0003 at 2-4 y, 0.034 (SE 0.031) 0.020 (SE 0.021) p = 0.6778 at 4-5 y (Shefferly 2016); No significant association (4 studies, n= 2138): ANCOVA p = 0.062 (Hasnain 2014); β = 0.01 SE = 0.00, p = 0.20 (Newby 2004); β = -0.001, 95% CI = -0.059, 0.057, p = 0.97 (Marshall 2019); β = -0.057, p = 0.099 (SE not stated) (Skinner 1999) [e]	⨁⨁◯◯ LOW	CRITICAL
Mean BMI/BMI z-score or change in BMI/BMI z-score in children 5- ≤ 10 years at exposure									
2	observational studies	extremely serious[h]	not serious[g]	not serious[c]	not serious[d]	none	Increased BMI (0 studies); Different effects (0 studies); No significant association (2 studies, n=412): β = -0.04, CI = -0.21, 0.13, p = 0.631 (Carlson 2012); β = 0.07, SE = 0.05, p = 0.12 (Zheng 2015). [e]	⨁◯◯◯ VERY LOW	CRITICAL
Prevalence of overweight and obesity or prevalence of obesity only in children aged < 2 years (assessed with: %)									
1	observational studies	extremely serious[h]	not serious[b]	not serious[c]	not serious[d]	none	No significant association (1 study, n = 1076); Odds of overweight including obesity, aOR = 1.0, 95% CI = 0.5, 2.0, p = 0.916 (Budree 2017)	⨁◯◯◯ VERY LOW	CRITICAL
Prevalence of overweight and obesity or prevalence of obesity only in children aged 2-< 5 years (assessed with: %)									

2	observational studies	very serious[a]	not serious[g]	not serious[c]	not serious[d]	none	Increased overweight/obesity (0 studies); Different effects (1 study, n=6250); overweight/obesity aOR = 1.30, 95% CI = 1.06-1.59, p = 0.0129 at 2-4 y follow-up; aOR = 0.80, 95% CI = 0.43-1.49, p= 0.473 at 4-5 y follow-up) (Shefferly 2016); No significant association (1 study, n=10904): high vs low intake among normal weight at baseline: aOR = 1.2 95% cI = 0.8-1.7); high vs low intake among at risk of overweight at baseline aOR= 0.8 95% CI = 0.5- 1.1 (Welsh 2005)	⊕⊕◯◯ LOW	CRITICAL
Mean percent body fat in children aged ≤ 10 years (assessed with: %)									
4	observational studies	very serious[i]	not serious[g]	not serious[c]	not serious[d]	none	Increased % body fat (0 studies); Different effects (0 studies); No significant association (4 studies, n= 872); β= -1.06, 95% CI = -2.70, 0.57, p = 0.202 (Carlson 2021); ANCOVA p = 0.119 (Hasnain 2014); β = -0.11, 95% CI = -0.61, 0.38, p = 0.66 (Johnson 2007); β = -0.05, SE = 0.44, p = 0.91 (Zheng 2015)	⊕⊕◯◯ LOW	

CI: confidence interval

Explanations

a. Risk of bias was moderate in all studies. Downrated by 2 levels due to non-randomization in observational studies leading to confounding and selection bias.
b. Not downrated as only 1 study
c. Not downrated as study populations, exposures and comparators were relevant to review question, although no studies were from low income country populations
d. Not downrated as no evidence of imprecision (i.e. not wide confidence intervals, small sample size or low number of events)
e. Meta-analysis of 3 studies across age groups on BMI z-score effect size: 0.01, 95% CI 0.00, 0.01
f. Risk of bias was moderate in four studies (Hasnain 2014, Marshall 2019, Newby 2004, Shefferly 2016) and serious in one study (Skinner 1999). Downrated by 2 levels due to non-randomization in observational studies leading to confounding and selection bias
g. Not downrated for inconsistency but note that interventions and comparators were not the same across studies
h. Risk of bias was serious in all studies. Downrated by 2 levels for inherent risk of bias due to non-randomisation and 1 further level due to body of evidence based on studies at serious risk of bias.
i. Risk of bias was moderate in 2 studies (Hasnain 2014 and Johnson 2007) and serious in 2 studies (Carlson 2012 and Zheng 2015). Downrated by 2 levels due to non-randomization leading to bias due to confounding and selection bias.

Nutrient supplements and fortified food products: Multiple micronutrient powders (MNPs)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Multiple micronutrient powders	No intervention or placebo	Relative (95% CI)	Absolute (95% CI)		
Anaemia												
16	randomised trials	not serious	serious [a]	not serious	not serious	none	2689/5009 (53.7%)	2980/4918 (60.6%)	**RR 0.82 (0.76 to 0.90)**	**109 fewer per 1,000 (145 fewer to 61 fewer)**	⊕⊕⊕◯ MODERATE	CRITICAL
Iron deficiency												
7	randomised trials	not serious	not serious	not serious	not serious	none	134/853 (15.7%)	262/781 (33.5%)	**RR 0.47 (0.39 to 0.56)**	**177 fewer per 1,000 (204 fewer to 147 fewer)**	⊕⊕⊕⊕ HIGH	CRITICAL
Haemoglobin (g/L)												
21	randomised trials	serious [b]	serious [a]	not serious	not serious	none	5307	5202	-	**MD 2.74 g/L higher (1.95 higher to 3.53 higher)**	⊕⊕◯◯ LOW	CRITICAL
Serum ferritin (ug/L)												
7	randomised trials	not serious	serious [a]	not serious	not serious	none	1319	1293	-	**MD 12.93 ug/L higher (7.41 higher to 18.45 higher)**	⊕⊕⊕◯ MODERATE	CRITICAL
Weight-for-age z-score												
10	randomised trials	serious [b]	not serious	not serious	not serious	none	4716	4571	-	MD 0.02 lower (0.03 lower to 0.07 higher)	⊕⊕⊕◯ MODERATE	CRITICAL

CI: confidence interval

Explanations

a. Downgraded one level because heterogeneity was high.
b. Downgraded one level because some studies are at high or unclear risk of selection and attrition bias

Nutrient supplements and fortified food products: Fortified complementary foods

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Fortified complementary food**	**Non-fortified complementary food**	**Relative (95% CI)**	**Absolute (95% CI)**		
Anaemia (follow-up: 3 to 12 months)												
6	randomised trials	serious[a]	not serious	not serious	not serious[b]	none	51/617 (8.3%)	90/588 (15.3%)	**RR 0.57 (0.39 to 0.82)**	**66 fewer per 1 000 (93 fewer to 28 fewer)**	⊕⊕⊕◯ Moderate	CRITICAL
Haemoglobin (follow-up: 3 to 12 months; assessed with: g/L)												
11	randomised trials	serious[c]	not serious[d]	not serious	not serious[e]	none	1110	1065	**-**	**MD 3.43 g/L higher (1.34 higher to 5.52 higher)**	⊕⊕⊕◯ Moderate	CRITICAL
Weight-for-age (follow-up: 6 to 12 months; assessed with: z-scores)												
5	randomised trials	serious[f]	not serious	not serious	not serious[e]	none	607	599	-	**MD 0.01 z-score lower (0.07 lower to 0.06 higher)**	⊕⊕⊕◯ Moderate	CRITICAL
Weight-for-length (follow-up: 6 to 12 months; assessed with: z-scores)												
4	randomised trials	serious[g]	not serious	not serious	not serious[e]	none	560	549	-	**MD 0.05 z-score lower (0.19 lower to 0.1 higher)**	⊕⊕⊕◯ Moderate	CRITICAL
Length-for-age (follow-up: 6 to 12 months; assessed with: z-scores)												
4	randomised trials	serious[g]	serious[h]	not serious	not serious[e]	none	412	399	-	**MD 0.01 z-score lower (0.21 lower to 0.2 higher)**	⊕⊕◯◯ LOW	CRITICAL
Iron status (follow-up: 3 to 12 months; assessed with: ferritin concentrations in ug/L)												
6	randomised trials	serious[i]	serious[j]	not serious	not serious[e]	none	464	439	-	**MD 0.43 ug/L higher (0.14 higher to 0.72 higher)**	⊕⊕◯◯ LOW	CRITICAL

Iron status (follow-up: 12 months; assessed with: body iron in mg/kg)												
1	randomised trials	serious[k]	not serious[l]	not serious	serious[m]	none	97	104	-	**MD 1.47 mg/kg higher (0.63 higher to 2.31 higher)**	⊕⊕◯◯ LOW	CRITICAL
Iron status (follow-up: 3 months; assessed with: free erythrocythe protoporphyrin in µg/L)												
1	randomised trials	very serious[n]	not serious[l]	not serious	serious[m]	none	69	78	-	MD 30 higher (26.06 lower to 86.06 higher)	⊕◯◯◯ VERY LOW	CRITICAL
Serum retinol (follow-up: 3 to 12 months; assessed with: µmol/L)												
5	randomised trials	serious[o]	not serious	not serious	not serious[e]	none	225	250	-	MD 0.03 umol/L higher (0.02 lower to 0.08 higher)	⊕⊕⊕◯ Moderate	CRITICAL
Serum zinc (follow-up: 6 months; assessed with: g/dL)												
2	randomised trials	serious[p]	not serious	not serious	serious[m]	none	170	163	-	MD 0.13 lower (0.82 lower to 0.56 higher)	⊕⊕◯◯ LOW	CRITICAL
Diarrhoea (follow-up: 6 months; assessed with: number of new episoded/ 100 days at risk)												
1	randomised trials	serious[k]	not serious[l]	not serious	very serious[q]	none	47	50	-	MD 0.6 higher (2.16 lower to 3.36 higher)	⊕◯◯◯ VERY LOW	IMPORTANT
Acute respiratory tract diseases (follow-up: 6 months; assessed with: number of new episoded/ 100 days at risk)												
	randomised trials	serious[k]	not serious[l]	not serious	very serious[q]	none	47	50	-	MD 0.3 higher (0.38 lower to 0.98 higher)	⊕◯◯◯ VERY LOW	IMPORTANT
Fever diseases (follow-up: 6 months; assessed with: number of new episoded/ 100 days at risk)												
	randomised trials	serious[k]	not serious[l]	not serious	very serious[q]	none	47	50	-	MD 0.1 higher (1.21 lower to 1.41 higher)	⊕◯◯◯ VERY LOW	IMPORTANT
Mental skill development (follow-up: 10 to 12 months; assessed with: BSID I-III)												
2	randomised trials	serious[r]	not serious	not serious	not serious[e]	none	250	258	-	**MD 0.8 higher (0.12 higher to 1.48 higher)**	⊕⊕⊕◯ Moderate	IMPORTANT
Fine motor score (follow-up: 12 months; assessed with: BSID III)												
1	randomised trials	serious[k]	not serious[l]	not serious	serious[m]	none	55	58	-	MD 0 (0.4 lower to 0.4 higher)	⊕⊕◯◯ LOW	IMPORTANT
Gross motor score (follow-up: 12 months; assessed with: BSID III)												

1	randomised trials	serious[k]	not serious[l]	not serious	serious[m]	none	55	58	-	MD 0.2 lower (0.62 lower to 0.22 higher)	⊕⊕◯◯ LOW	IMPORTANT
Psychomotor development (follow-up: 6 to 10 months; assessed with: BSID I-III)												
2	randomised trials	serious[p]	not serious	not serious	not serious[e]	none	323	338	-	**MD 1.13 higher (0.35 higher to 1.91 higher)**	⊕⊕⊕◯ Moderate	IMPORTANT
Acceptability (follow-up: 3 days; assessed with: 9-point hedonic scale)												
3	randomised + non-randomised trials	very serious[s]	not serious	not serious	serious[m]	none	Acceptability of fortified as compared to unfortified complementary food was measured in three acute studies with a total of 215 children. All described that there were no significant differences between the ratings of children allocated to the two groups				⊕◯◯◯ VERY LOW	IMPORTANT
Iron deficiency (follow-up: 6 to 12 months)												
3	randomised trials	serious[t]	not serious	not serious	not serious[u]	none	48/274 (17.5%)	120/297 (40.4%)	RR 0.39 (0.21 to 0.75)	246 fewer per 1 000 (319 fewer to 101 fewer)	⊕⊕⊕◯ Moderate	CRITICAL
Vitamin A deficiency (follow-up: 3 to 12 months)												
3	randomised trials	very serious[v]	serious[w]	not serious	serious[x]	none	39/148 (26.4%)	30/109 (27.5%)	RR 0.97 (0.24 to 3.90)	8 fewer per 1 000 (209 fewer to 798 more)	⊕◯◯◯ VERY LOW	CRITICAL
Zinc deficiency (follow-up: 6 months)												
1	randomised trials	serious[y]	not serious[l]	not serious	very serious[z]	none	3/30 (10.0%)	1/31 (3.2%)	RR 3.10 (0.34 to 28.17)	68 more per 1 000 (21 fewer to 876 more)	⊕◯◯◯ VERY LOW	CRITICAL

CI: confidence interval; **MD:** mean difference; **RR:** risk ratio

Explanations

a. Downgraded by one level for risk of bias (RoB) since 1 out of 6 studies was rated with a high RoB, and none of the included studies was rated with a low RoB.
b. Not downgraded for imprecision. Although the number of events was low (<400), the outcome was a common event (occurred >1/100), and there were 6 studies with a median sample size of 170 children included. The 95% confidence interval (CI) for the pooled estimate is narrow and is consistent with benefit.
c. Downgraded by one level for RoB since 5 out of 11 studies were rated with a high RoB, and none of the included studies was rated with a low RoB.
d. Not downgraded for inconsistency although I^2 was 55% (driven by the study of Faber et al. 2005), since 95% CI overlaps mainly between studies. In all sub-group analyses heterogeneity was present only in those sub-groups which contained the study Faber et al. 2005, while no heterogeneity was observed in other sub-groups.
e. Not downgraded for imprecision since number of participants was >400.
f. Downgraded by one level for RoB since 1 out of 5 studies were rated with a high RoB, and none of the included studies was rated with a low RoB.
g. Downgraded by one level for RoB since 1 out of 4 studies were rated with a high RoB, and none was rated as low RoB.
h. Downgraded by one level for inconsistency since I^2 was 68%, p-value for heterogeneity was 0.02, point estimates and 95% CI did not overlap between studies. Sub-group analyses did not fully explain heterogeneity.
i. Downgraded by one level for RoB since 3 out of 6 studies were rated with a high RoB, and none was rated as low RoB.
j. Downgraded by one level for inconsistency since I^2 was 86%, p-value for heterogeneity was <0.001, point estimates and 95% CI did not overlap between studies. Sub-group analyses did not fully explain heterogeneity.
k. Downgraded by one level for RoB since the included study was rated with some concerns of RoB.

l. This is a single study so inconsistency cannot be judged.
m. Downgraded by one level for imprecision since total sample size was low (<400).
n. Downgraded by two levels for RoB since the included study was rated with a high RoB.
o. Downgraded by one level for RoB since 3 out of 5 studies were rated with a high RoB, and 1 study was rated as low RoB.
p. Downgraded by one level since 1 out of 2 included studies was rated with a high RoB, and none of the included studies was rated with a low RoB.
q. Downgraded by two levels for imprecision since sample size was VERY LOW (<100).
r. Downgraded by one level for RoB since both included studies were rated with some concerns for RoB.
s. Downgraded by two levels for RoB since 2 out of 3 included studies was rated with high RoB, and none of the included studies was rated with a low RoB.
t. Downgraded by one level for RoB since all included studies were rated with some concerns for RoB.
u. Not downgraded for imprecision. Although the number of included studies is low (n=3), studies had an intermediate sample size with a median of 239 participants, and the outcome was a common event (occurred >1/100).
v. Downgraded by two levels for RoB, as for this outcome 1 out of the 3 included studies was rated with high RoB, and none of the included studies was rated with a low RoB. There were large baseline between-group differences in the number of vitamin A deficient participants in two studies (Nesamvuni 2005: 7 out of 16 in the experimental and 0 out of 20 in the control group; Palmer 2021: 10 out of 51 in the biofortified, 11 out 52 in the fortified and 18 out of 58 in the control group)
w. Downgraded by one level for inconsistency as point estimates did vary widely, 95% CI did not overlap between studies, the direction of effect was not consistent. and the magnitude of heterogeneity was high (I^2 was 74%, p-value for heterogeneity was 0.02). Due to the low number of studies subgroup analyses were not possible.
x. Downgraded by one level for imprecision since total sample size was low (<400).
y. Downgraded by one level for RoB since the included study was rated with some concerns of RoB.
z. Downgraded by two levels for imprecision since results are derived from one study, where total sample size was VERY LOW (n<100).

Nutrient supplements and fortified food products: Fortified milk

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Animal milk (full-fat or	fortified milk (full-fat or	Relative (95% CI)	Absolute (95% CI)		
Weight (kg)												
1[a]	randomised trials	serious	not serious	not serious	serious	none	16	20	-	MD 0.04 kg higher (0.83 lower to 0.91 higher)	⊕⊕◯◯ LOW	CRITICAL
Undernutrition – Stunting and wasting												
1[b]	randomised trials	not serious	not serious	not serious	serious	none	65/187 (34.8%)	68/191 (35.6%)	RR 0.98 (0.74 to 1.28)	7 fewer per 1000 (93 fewer to 100 more)	⊕⊕⊕◯ LOW	CRITICAL
Undernutrition-Stunting												
1[b]	randomised trials	not serious	not serious	not serious	serious	none	65/187 (34.8%)	68/191 (35.6%)	RR 0.98 (0.74 to 1.28)	7 fewer per 1000 (93 fewer to 100 more)	⊕⊕⊕◯ LOW	CRITICAL
Undernutrition-Wasting												
1[b]	randomised trials	not serious	not serious	not serious	serious	none	57/187 (30.5%)	55/191 (28.8%)	RR 1.06 (0.78 to 1.44)	17 more per 1000 (63 fewer to 127 more)	⊕⊕⊕◯ LOW	CRITICAL
Nutrient status-Iron as serum iron (µmol/L)												
2[c]	randomised trials	serious	not serious	not serious	serious	none	57	58	-	MD 0.46 µmol/L lower (4.38 lower to 3.46 higher)	⊕⊕◯◯ LOW	CRITICAL
Nutrient status-Zinc as plasma zinc (µmol/L)												
2[d]	randomised trials	very serious	not serious	not serious	not serious	none	244	249	-	**MD 0.43 higher (0.11 higher to 0.76 higher)**	⊕⊕◯◯ LOW	CRITICAL
Anaemia												
3[e]	randomised trials	very serious	not serious	not serious	not serious	none	147/556 (26.4%)	96/768 (12.5%)	**RR 2.29 (1.12 to 4.69)**	**161 more per 1000 (15 more to 461 more)**	⊕⊕◯◯ LOW	CRITICAL
Iron deficiency anaemia (IDA)												

1[b]	randomised trials	not serious	not serious	not serious	serious	none	128/232 (55.2%)	31/233 (13.3%)	**RR 4.15 (2.93 to 5.87)**	**419 more per 1000 (257 more to 648 more)**	⊕⊕◯◯ LOW	CRITICAL
Iron deficiency (ID, serum ferritin <12 µg/l)												
1[f]	randomised trials	not serious	not serious	not serious	serious	none	10/114 (8.8%)	17/235 (7.2%)	RR 1.21 (0.57 to 2.56)	15 more per 1000 (31 fewer to 113 more)	⊕⊕◯◯ LOW	CRITICAL
Haemoglobin (g/dL)												
6[g]	randomised trials	very serious	not serious	not serious	not serious	none	663	691	-	**MD 5.91 g/L lower (9.84 lower to 1.99 lower)**	⊕⊕⊕◯ MODERATE	CRITICAL
Oral health -caries free (decayed, missing, and filled surfaces -dmfs- index in molars and canines= 0)												
1[h]	randomised trials	serious	not serious	not serious	serious	none	43/76 (56.6%)	85/110 (77.3%)	RR 1.30 (0.37 to 2.23)	**209 fewer per 1000 (317 fewer to 70 fewer)**	⊕⊕◯◯ LOW	IMPORTANT
Morbidity-Respiratory episodes per child per year												
1[i]	randomised trials	serious [e]	not serious	not serious	serious [d]	none	278	276	-	MD 0.03 lower (0.14 lower to 0.20 higher)	⊕⊕◯◯ LOW	IMPORTANT

CI: confidence interval; **MD:** mean difference; **RR:** risk ratio

Explanations

a. One trial reported on this outcome (Svahn) and the intervention lasted 6 months. Evidence was downgraded two levels. It assessed as high risk for detection bias and reporting bias, and it had very small sample size so there was not enough information to detect a precise estimate of the effect.
b. One trial reported on this outcome and the intervention lasted 12 mo (Sazawal). All children who had severe anaemia at baseline (Hb ≤ 70 g/l) were given a therapeutic dose of iron for three months in addition to their milk supplement. In the fortified milk group 9 children left the area, 2 children died, and 16 withdrawn consents. In the milk group 18 children left the area, 2 children died, and 16 withdrawn consents. Participants´ flow chart is not clear and therefore, incomplete outcome data is unclear too. Small sample sizes both in intervention and control groups. Evidence was downgraded two levels accordingly.
c. Two trials reported on this outcome (Stekel and Svahn). Randomization method and similarity of baseline outcome measurements not mentioned in one study. Small sample size both in control and intervention groups so there was not enough information to detect a precise estimate of the effect. Evidence was downgraded two levels.
d. Two trials reported on this outcome (Sazawal and Villalpando). Sazawal was assessed as unclear risk of bias for selection bias and attrition bias. In the fortified milk group 9 children left the area, 2 children died, and 16 withdrawn consents. In the milk group 18 children left the area, 2 children died, and 16 withdrawn consents. Participants' flow chart is not clear. Villalpando was assessed as high risk of bias for selection bias as the randomization procedure did not result in an even distribution of baseline anaemia in the 2 intervention groups (30.0% non-fortified milk, 41.4% fortified milk). Groups did not differ at 6 mo and the changes between baseline and 6 mo did not differ between the groups. However, baseline outcomes showed that fortification milk group was more susceptible to a larger improvement. Evidence was downgraded two levels.
e.Sazawal was assessed as unclear risk of bias for selection bias and attrition bias. In the fortified milk group 9 children left the area, 2 children died, and 16 withdrawn consents. In the milk group 18 children left the area, 2 children died, and 16 withdrawn consents. Participants' flow chart is not clear. Villalpando was assessed as high risk of bias for selection bias as the randomization procedure did not result in an even distribution of baseline anaemia in the 2 intervention groups (30.0% non-fortified milk, 41.4% fortified milk). Groups did not differ at 6 mo and the changes between baseline and 6 mo did not differ between the groups. However, baseline outcomes showed that fortification milk group was more susceptible to a larger improvement. Evidence was downgraded two levels. Stekel 1986 was assessed as high risk of bias as blinding was not reported.
f. One trial reported on this outcome (Villalpando). The trial was assessed as high risk of bias for selection bias as the randomization procedure did not result in an even distribution of baseline anaemia in the 2 intervention groups (30.0% non-fortified milk, 41.4% fortified milk). Groups did not differ at 6 mo and the changes between baseline and 6 mo did not differ between the groups.
g. Five trials reported on this outcome (Stekel1986, Stekel 1988, Svahn1999, Szymlek-Gay and Villalpando 2006). Evidence was downgraded two levels as randomization and blinding were not mentioned in some studies and risk of bias was assessed as high.
h. One cluster RCT reported on this outcome (Stecksén-Blicks). Children were served 150 ml medium-fat milk (1.5%) at lunch. Blinding of outcome assessment was not mentioned and evidence was downgraded one level for this and another level for the sample size in both control and intervention groups.
i. Stekel 1988 reported on this outcome. The trial did not report on blinding of participants and personnel nor of outcome assesment and was downgraded one level for this. The trial had a small sample size in both intervention and control groups and was downgraded another level.

Small-quantity lipid-based nutrient supplements (SQ-LNS)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	SQ-LNS	Control	Relative (95% CI)	Absolute (95% CI)		
Length-for-age z-score												
17	randomised trials	not serious	not serious	not serious	not serious	none	17248	19547	-	**MD 0.14 higher (0.11 higher to 0.16 higher)**	⊕⊕⊕⊕ HIGH	CRITICAL
Weight-for-length z-score												
17	randomised trials	not serious	not serious	not serious	not serious	none	17002	19309	-	**MD 0.08 higher (0.06 higher to 0.10 higher)**	⊕⊕⊕⊕ HIGH	CRITICAL
Stunting (LAZ< -2SD)												
17	randomised trials	not serious	not serious	not serious	not serious	none	17248	19547	**RR 0.88 (0.85 to 0.91)**	**50 fewer per 1000 (59 fewer to 41 fewer)**	⊕⊕⊕⊕ HIGH	CRITICAL
Severe Stunting (LAZ < -3 SD)												
14	randomised trials	not serious	not serious	not serious	not serious	none	17248	19547	**PR 0.83 (0.78 to 0.90)**		⊕⊕⊕⊕ HIGH	CRITICAL
Wasting												
16	randomised trials	not serious	not serious	not serious	not serious	none	17002	19309	**RR 0.86 (0.80 to 0.93)**	**6 fewer per 1000 (10 fewer to 1 fewer)**	⊕⊕⊕⊕ HIGH	CRITICAL
Severe Wasting (WLZ < -3 SD)												
12	randomised trials	not serious	not serious	not serious	not serious	none	15807	18566	**PR 0.69 (0.55, 0.86)**		⊕⊕⊕⊕ HIGH	CRITICAL
Severe Acute Malnutrition (WLZ < -3 SD or MUAC < 115 mm)												
11	randomised trials	not serious	not serious	not serious	not serious	none	30436		**PR 0.76 (0.62, 0.93)**		⊕⊕⊕⊕ HIGH	CRITICAL
Very low MUAC (MUACZ < -3 SD or < 115 mm)												
10	randomised trials	not serious	not serious	not serious	not serious	none	30069		**PR 0.73 (0.56, 0.94)**		⊕⊕⊕⊕ HIGH	CRITICAL

Haemoglobin (g/L)												
14	randomised trials	not serious	not serious	not serious	not serious	none	8541	6857	-	**MD 2.77 higher (2.30 higher to 3.25 higher)**	⊕⊕⊕⊕ HIGH	CRITICAL
Ferritin (ug/L)												
7	randomised trials	not serious	not serious	not serious	not serious	none	1996	1082	-	**MD 1.56 higher (1.48 higher to 1.64 higher)**	⊕⊕⊕⊕ HIGH	CRITICAL
Serum TfR (mg/L)												
6	randomised trials	not serious	not serious	not serious	not serious	none	1542	938	-	**MD 0.83 higher (0.80 higher to 0.85 higher)**	⊕⊕⊕⊕ HIGH	CRITICAL
Zinc Protoporphyrin umol/mol												
4	randomised trials	not serious	not serious	not serious	serious	none	1542		-	**MD 0.80 higher (0.75 higher to 0.85 higher)**	⊕⊕⊕◯ MODERATE	CRITICAL
Anaemia												
14	randomised trials	not serious	not serious	not serious	not serious	none	8541	6857	**RR 0.84 (0.81 to 0.87)**		⊕⊕⊕⊕ HIGH	CRITICAL
Moderate-to-severe anaemia												
13	randomised trials	not serious	not serious	not serious	not serious	none	14375		**RR 0.72 (0.68 to 0.76)**		⊕⊕⊕⊕ HIGH	CRITICAL
Iron deficiency												
7	randomised trials	not serious	not serious	not serious	not serious	none	1996	1082	**RR 0.44 (0.39 to 0.50)**		⊕⊕⊕⊕ HIGH	CRITICAL
Iron deficiency anaemia												
6	randomised trials	not serious	not serious	not serious	not serious	none	1764	908	**RR 0.36 (0.30 to 0.44)**		⊕⊕⊕⊕ HIGH	CRITICAL
Elevated sTfR												
6	randomised trials	not serious	not serious	not serious	not serious	none	1542	938	**RR 0.64 (0.59 to 0.70)**		⊕⊕⊕⊕ HIGH	CRITICAL
Elevated ZPP												
4	randomised trials	not serious	not serious	not serious	serious	none	1542		**RR 0.70 (0.61 to 0.79)**		⊕⊕⊕◯ MODERATE	CRITICAL
Mortality												

13							218/16062 (1.35%)	294/17989 (1.63%_	**RR: 0.73 (0.59 to 0.89)**			CRITICAL
Language z score												
13	randomised trials	not serious	not serious	not serious	not serious	none	10188	14373	**-**	**MD 0.07 higher (0.04 higher to 0.10 higher)**	⊕⊕⊕⊕ HIGH	IMPORTANT
Social-emotional z score												
11	randomised trials	not serious	not serious	not serious	not serious	none	9573	14015	**-**	**MD 0.08 (0.05 higher to 0.11 higher)**	⊕⊕⊕⊕ HIGH	IMPORTANT
Motor z score												
12	randomised trials	not serious	not serious	not serious	not serious	none	9871	14028	**-**	**MD 0.08 higher (0.05 higher to 0.11 higher)**	⊕⊕⊕⊕ HIGH	IMPORTANT
Language lowest decile												
13	randomised trials	not serious	not serious	not serious	not serious	none	10188	14373	**RR 0.84 (0.76 to 0.92)**	**10 fewer per 1000 (20 fewer to 0 fewer)**	⊕⊕⊕⊕ HIGH	IMPORTANT
Social-emotional lowest decile												
11	randomised trials	not serious	not serious	not serious	not serious	none	9573	14015	**RR 0.81 (0.74 to 0.89)**	**20 fewer per 1000 (20 fewer to 10 fewer)**	⊕⊕⊕⊕ HIGH	IMPORTANT
Motor lowest decile												
12	randomised trials	not serious	not serious	not serious	not serious	none	9871	14028	**RR 0.84 (0.76 to 0.92)**	**20 fewer per 1000 (20 fewer to 10 fewer)**	⊕⊕⊕⊕ HIGH	IMPORTANT
Walking without support												
10	randomised trials	not serious	not serious	not serious	not serious	none	7014	6837	**RR 1.09 (1.05 to 1.14)**	**10 fewer per 1000 (20 fewer to 10 more)**	⊕⊕⊕⊕ HIGH	IMPORTANT

CI: confidence interval; **MD:** mean difference; **RR:** risk ratio; **PR**: prevalence ratio

Responsive feeding

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Responsive feeding	No responsive feeding	Relative (95% CI)	Absolute (95% CI)		
FOOD ACCEPTANCE **A. Interventions focused on one component of responsive feeding** ***Intervention A1:* Guidance on step-by-step repeated exposure to vegetables during the introduction of solids, delivered by research staff or health professionals, component C5** **Food acceptance of target vegetables (at ~6 months old). Amount of target vegetables consumed (g), after 24 to 35 days of repeated exposure to vegetables, measured in a laboratory setting**												
1	randomized trial	serious [a]	not serious	not serious	serious [b]	none	17	18	-	**MD 37.6 higher (14.0 higher to 61.2 higher)**	⊕⊕◯◯ LOW	CRITICAL
Food acceptance of novel vegetables (at ~6 to 7 months old). Amount of novel vegetables consumed (g), after ~ 1 month of repeated exposure to vegetables, measured in a laboratory setting												
2	randomized trials	serious [c]	not serious	not serious	not serious	none	61	58	-	**MD 15.6 higher (7.2 higher to 23.9 higher)**	⊕⊕⊕◯ MODERATE	CRITICAL
Food acceptance of novel fruit (at ~6 to 7 months old). Amount of novel fruits consumed (g), after ~ 1 month of repeated exposure to vegetables, measured in a laboratory setting												
1	randomized trial	very serious [d]	serious [e]	not serious	serious [f]	none	44	40	-	MD 0.5 higher (34.2 lower to 35.2 higher)	⊕◯◯◯ VERY LOW	CRITICAL
***Intervention A2:* Advice and regular counseling for promoting the introduction of textured foods, delivered by a research dietitian from 8 to 15 months of age, component C3** **Food texture acceptance of food textures (at ~15 months old). Global texture acceptance score-**from 0 to 8 (highest food texture acceptance), measured in a laboratory setting.												
1	randomized trial	serious [g]	not serious	not serious	very serious [f,h]	none	30	30	-	MD 0.30 higher (0.80 lower to 1.40 higher)	⊕◯◯◯ VERY LOW	CRITICAL
B. Interventions aimed to prevent under-nutrition, including two or more components of RF ***Intervention B1*: Responsive feeding and development stimulation program, delivered by trained women/mothers of the village or family welfare assistants, including 7 components of RF (C1, C3, C4, C5, C7, C9, C10)** **Food acceptance at ~20 to 23 months old. Number of mouthfuls eaten, observation of a midday meal by a research assistant**												
3	randomized trials	not serious	serious [i]	not serious	serious [f]	none	231	227	-	MD 1.98 higher (0.84 lower to 4.8 higher)	⊕⊕◯◯ LOW	CRITICAL
Food acceptance at ~20 to 23 months old. Self-fed mouthfuls (%), observation of a midday meal by a research assistant												
3	randomized trials	not serious	not serious	not serious	not serious	none	231	227	-	**MD 14.42 higher (6.45 higher to 22.39 higher)**	⊕⊕⊕⊕ HIGH	CRITICAL
Food acceptance at ~20 to 23 months old. Child refusals (%), observation of a midday meal by a research assistant												
3	randomized trials	not serious	not serious	not serious	serious [f]	none	231	227	-	**MD 0.69 lower (1.28 lower to 0.09 lower)**	⊕⊕⊕◯ MODERATE	CRITICAL

C. **Interventions for obesity prevention, with two or more components of RF** (HICs: Norway, New Zealand, USA)												
***Intervention C1*: Interventions aimed to prevent obesity (e-health intervention between 6 and 12 months of age), including 5 components of RF (C1, C4, C6, C7, C8)Food acceptance at 12 months old. Enjoyment of Food scale, measured by the Child Eating Behavior Questionnaire (CEBQ), score from 1 to 5 (highest enjoyment of food)**												
1	randomized trial	not serious	not serious	not serious	serious [f]	none	269	264	-	MD 0.10 higher (0.01 lower to 0.21 higher)	⊕⊕⊕◯ MODERATE	CRITICAL
Food acceptance at 24 months old. Enjoyment of Food scale (CEBQ) , score from 1 to 5 (highest enjoyment of food)												
1	randomized trial	not serious	not serious	not serious	serious [f]	none	152	143	-	MD 0.04 lower (0.16 lower to 0.08 higher)	⊕⊕⊕◯ MODERATE	CRITICAL
Food acceptance at 12 months old. Food Fussiness scale (CEBQ), score from 1 to 5 (highest food fussiness)												
1	randomized trial	not serious	not serious	not serious	serious [f]	none	269	264	-	MD 0.00 higher (0.12 lower to 0.12 higher)	⊕⊕⊕◯ MODERATE	CRITICAL
Food acceptance at 24 months old. Food Fussiness scale (CEBQ), score from 1 to 5 (highest food fussiness)												
1	randomized trial	not serious	not serious	not serious	serious [f]	none	152	143	-	**MD 0.04 lower (0.21 lower to 0.13 lower)**	⊕⊕⊕◯ MODERATE	CRITICAL
***Intervention C2*: Interventions aimed to prevent obesity delivered by health professionals (group sessions, home visits, specific advice during well-child visits), including between 6 and 8 components of RF (C1, C2, C4, C5, C7 and others)** **Food acceptance at 12 months old, Enjoyment of Food scale, measured by the Child Eating Behavior Questionnaire (CEBQ), score from 1 to 5 (highest enjoyment of food)**												
1	randomized trial	not serious	not serious	not serious	serious [f]	none	92	81	-	**MD 0.22 higher (0.04 higher to 0.40 higher)**	⊕⊕⊕◯ MODERATE	CRITICAL
Food acceptance from 24 to 30 months old, Enjoyment of Food scale (CEBQ) , score from 1 to 5 (highest enjoyment of food)												
3	randomized trials	serious [j]	not serious	not serious	serious [f]	none	422	435	-	**MD 0.11 higher (0.02 higher to 0.20 higher)**	⊕⊕◯◯ LOW	CRITICAL
Food acceptance at 12 months old, Food Fussiness scale (CEBQ), score from 1 to 5 (highest food fussiness)												
1	randomized trial	not serious	not serious	not serious	not serious [k]	none	92	81	-	**MD 0.31 lower (0.50 lower to 0.12 lower)**	⊕⊕⊕⊕ HIGH	CRITICAL
Food acceptance from 24 to 30 months old, Food Fussiness scale (CEBQ), score from 1 to 5 (highest food fussiness)												
3	randomized trials	serious [f]	not serious	not serious	serious [f]	none	422	435	-	**MD 0.16 lower (0.26 lower to 0.07 lower)**	⊕⊕◯◯ LOW	CRITICAL
Food acceptance at 13.7 months old, Food Fussiness scale (CEBQ), score from 1 to 5 (highest food fussiness)												
1	randomized trial	serious [l]	not serious	not serious	serious [f]	none	250	254	-	MD 0.10 lower (0.24 lower to 0.04 higher)	⊕⊕◯◯ LOW	CRITICAL
Food acceptance at 13.7 months old, Food Fussiness scale (CEBQ), score from 1 to 5 (highest food fussiness)												
1	randomized trial	serious [l]	not serious	not serious	serious [f]	none	213	211	-	MD 0.10 lower (0.24 lower to 0.04 higher)	⊕⊕◯◯ LOW	CRITICAL
HEALTHIER FOOD PREFERENCES ***Intervention C2*: Interventions aimed to prevent obesity delivered by health professionals (group sessions, individual contacts, home visits), including between 6 and 8 components of RF (C1, C2, C4, C5, C7, C9 + C3, C6 in one trial)** **Food preferences (at 24 months old). Perception of liking for vegetables, SMD**												

2	Randomized trials	not serious	not serious	not serious	serious [m]	none	308	320		SMD 0.15 higher (0.01 lower to 0.3 higher)	⊕⊕⊕◯ MODERATE	CRITICAL
Food preferences (at 3.7 years old). Perception of number of vegetables "liked" (% of listed vegetables "liked", out of 22 items)												
1	Randomized trial	serious [n]	not serious	not serious	serious [c]	none	250	254		MD 2.2 higher (1.96 lower to 6.36 higher)	⊕⊕◯◯ LOW	CRITICAL
Food preferences (at 5 years old). Perception of number of vegetables "liked" (% of listed vegetables "liked" , out of 22 items)												
1	Randomized trial	serious [g]	not serious	not serious	serious [c]	none	213	211		MD 1.3 higher (3.13 lower to 5.73 higher)	⊕⊕◯◯ LOW	CRITICAL
Food preferences (at 24 months old). Perception of liking for fruits, SMD												
2	Randomized trials [b]	not serious	serious [o]	not serious	serious [c]	none	308	320		SMD 0.15 higher (0.07 lower to 0.38 higher)	⊕⊕◯◯ LOW	CRITICAL
Food preferences (at 3.7 years old). Perception of fruits "liked" (% of listed fruits "liked" out of 16 items)												
1	Randomized trials	serious [g]	not serious	not serious	not serious	none	250	254		**MD 7.0 higher (3.4 higher to 10.6 higher)**	⊕⊕⊕◯ MODERATE	CRITICAL
Food preferences (at 5 years old). Perception of fruits "liked" (% of listed fruits "liked" out of 16 items)												
1	Randomized trials	serious [g]	not serious	not serious	not serious	none	213	211		**MD 5.2 higher (1.6 higher to 8.8 higher)**	⊕⊕⊕◯ MODERATE	CRITICAL
Food preferences at 24 months old. Perception of liking for meat and fish, mean score on a response scale of 1 (dislikes a lot) to 5 (likes a lot)												
1	Randomized trials	not serious	not serious	not serious	serious [c]	none	86	75		MD 0.10 higher (0.07 lower to 0.27 higher)	⊕⊕⊕◯ MODERATE	CRITICAL
Food preferences at 24 months old. Perception of number of energy-dense sweet and savory foods "liked" (% of listed sweet and savory foods "liked", out of 18 items)												
1	Randomized trials	serious [g]	not serious	not serious	serious [c]	none	222	245		MD 2.5 lower (5.27 lower to 0.27 higher)	⊕⊕◯◯ LOW	CRITICAL
Food preferences at 3.7 years old. Perception of number of energy-dense sweet and savory foods "liked" (% of listed sweet and savory foods "liked", out of 17 items)												
1	Randomized trial	serious [g]	not serious	not serious	serious [c]	none	250	254		MD 1.40 lower (4.45 lower to 1.65 higher)	⊕⊕◯◯ LOW	CRITICAL
Food preferences at 5 years old. Perception of number of energy-dense sweet and savory foods "liked" (% of listed sweet and savory foods "liked", out of 17 items)												
1	Randomized trial	serious [g]	not serious	not serious	serious [c]	none	213	211		MD 0.20 lower (3.25 lower to 2.85 higher)	⊕⊕◯◯ LOW	CRITICAL
HEALTHY FOOD INTAKE **A. Interventions aimed to prevent under-nutrition, including two or more components of RF** ***Intervention B1*: Responsive feeding and development stimulation program, delivered by trained women/mothers of the village or family welfare assistants, including 6-7 components of RF (C1, C3, C4, C5, C7, C9 and in some studies C10)** **Dietary diversity score, between 17 and 21 months old (# of food groups consumed out of 7 critical food groups during the previous day)**												
3	Randomized trials	not serious	not serious	not serious	serious [p]	none	312	313		**MD 0.25 higher (0.04 higher to 0.45 higher)**	⊕⊕⊕◯ MODERATE	CRITICAL
Vegetable intake, at 9 months old, consumption during the previous week (dichotomous outcome, spinach)												
1	Randomized trials	serious [q]	not serious	not serious	not serious	none	18/122 (14.8%)	7/135 (5.2%)	**RR 2.85 (1.23 to 6.58)**	**96 more per 1,000 (12 more to 289 more)**	⊕⊕⊕◯ MODERATE	CRITICAL

Vegetable intake, at 15 months old, consumption during the previous week (dichotomous outcome, spinach)												
1	Randomized trials	serious [e]	not serious	not serious	not serious	none	51/112 (45.5%)	34/129 (26.4%)	RR 1.73 (1.21 to 2.46)	192 more per 1,000 (55 more to 385 more)	⊕⊕⊕◯ MODERATE	CRITICAL
Vegetable intake, between 20 and 23 months old, times/day												
2	Randomized trials	serious [r]	serious [s]	not serious	serious [b]	none	165	149		MD 0.09 higher (0.88 lower to 1.06 higher)	⊕◯◯◯ VERY LOW	CRITICAL
Fruit intake, at 9 months old, consumption during the previous week (dichotomous outcome, banana)												
1	Randomized trials	serious [e]	not serious	not serious	not serious	none	72/122 (59.0%)	52/135 (38.5%)	**RR 1.53 (1.18 to 1.99)**	**204 more per 1,000 (69 more to 381 more)**	⊕⊕⊕◯ MODERATE	CRITICAL
Fruit intake, at 15 months old, consumption during the previous week (dichotomous outcome, banana)												
1	Randomized trials	serious [e]	not serious	not serious	not serious	none	88/112 (78.6%)	80/129 (62.0%)	**RR 1.27 (1.07 to 1.50)**	**167 more per 1,000 (43 more to 310 more)**	⊕⊕⊕◯ MODERATE	CRITICAL
Fruit intake, between 20 and 23 months old, times/day												
2	Randomized trials	serious [g]	not serious	not serious	not serious	none	165	149		**MD 0.23 higher (0.12 higher to 0.35 higher)**	⊕⊕⊕◯ MODERATE	CRITICAL
Egg intake, at 9 months old, consumption during the previous week (dichotomous outcome)												
1	Randomized trials	serious [e]	not serious	not serious	not serious	none	63/122 (51.6%)	24/135 (17.8%)	**RR 2.90 (1.94 to 4.34)**	**338 more per 1,000 (167 more to 594 more)**	⊕⊕⊕◯ MODERATE	CRITICAL
Egg intake, at 15 months old, consumption during the previous week (dichotomous outcome)												
1	Randomized trials	serious [e]	not serious	not serious	not serious	none	87/112 (77.7%)	70/129 (54.3%)	**RR 1.43 (1.19 to 1.73)**	**233 more per 1,000 (103 more to 396 more)**	⊕⊕⊕◯ MODERATE	CRITICAL
Meat (goat) intake, at 9 months old, consumption during the previous week (dichotomous outcome)												
1	Randomized trials	serious [e]	not serious	not serious	serious [b]	none	14/122 (11.5%)	6/135 (4.44%)	**RR 2.58 (1.02 to 6.51)**	**70 more per 1,000 (1 more to 245 more)**	⊕⊕◯◯ LOW	CRITICAL
Meat (goat) intake, at 15 months old, consumption during the previous week (dichotomous outcome)												
1	Randomized trials	serious [e]	not serious	not serious	serious [b]	none	49/112 (43.8%)	42/129 (32.6%)	RR 1.34 (0.97 to 1.86)	111 more per 1,000 (10 less to 280 more)	⊕⊕◯◯ LOW	CRITICAL
Egg intake, at 20 to 23 months old, times/day												
2	Randomized trials	serious [g]	not serious [t]	not serious	serious [b]	none	165	149		MD 0.13 higher (0.00 lower to 0.25 higher)	⊕⊕◯◯ LOW	CRITICAL
Fish intake, at 20 to 23 months old, times/day												
2	Randomized trials	serious [g]	not serious	not serious	serious [b]	none	165	149		MD 0.06 lower (0.30 lower to 0.17 higher)	⊕⊕◯◯ LOW	CRITICAL
A. Interventions for obesity prevention, including two or more components of RF												
***Intervention C1*: Interventions aimed to prevent obesity (e-health intervention between 6 and 12 months of age), including 5 components of RF (C1, C4, C6, C7, C8)**												
Fruit and vegetable consumption at 12 months old, times/day												
1	randomized trial	not serious	not serious	not serious	serious [b]	none	269	264	-	**MD 0.51 higher (0.07 higher to 0.95 higher)**	⊕⊕⊕◯ MODERATE	CRITICAL
Fruit and vegetable consumption at 24 months old, times/day												

1	randomized trial	not serious	not serious	not serious	serious [b]	none	152	143	-	MD 0.21 higher (0.32 lower to 0.74 higher)	⊕⊕⊕◯ MODERATE	CRITICAL
***Intervention C2*: Interventions aimed to prevent obesity delivered by health professionals (group or individual sessions during home visits or at health centers, specific brief advice included in well-child visits), including between 5 and 8 components of RF (C1, C4, C7, and others)**												
Vegetable Intake, at 9 to 12 months old, SMD												
3	Randomized trials	serious [u]	not serious	not serious	serious [b]	none	268	274		SMD 0.04 higher (0.15 lower to 0.23 higher)	⊕⊕◯◯ LOW	CRITICAL
Vegetable Intake, at 20 to 24 months old, SMD												
3	Randomized trials	serious [v]	not serious	not serious	serious [b]	none	491	511		SMD 0.01 lower (0.15 lower to 0.13 higher)	⊕⊕◯◯ LOW	CRITICAL
Vegetable Intake, at 3.5 to 3.7 years old, SMD												
2	Randomized trials	not serious	serious [w]	not serious	serious [b]	none	349	372		SMD 0.07 higher (0.17 lower to 0.31 higher)	⊕⊕◯◯ LOW	CRITICAL
Vegetable Intake, at 5 years old, SMD												
2	Randomized trials	not serious	not serious	not serious	serious [b]	none	353	370		SMD 0.08 higher (0.06 lower to 0.23 higher)	⊕⊕⊕◯ MODERATE	CRITICAL
Fruit Intake, at 9 to 12 months old, SMD												
3	Randomized trials	serious [s]	not serious	not serious	serious [b]	none	268	274		SMD 0.15 higher (0.06 lower to 0.35 higher)	⊕⊕◯◯ LOW	CRITICAL
Fruit Intake, at 20 to 24 months old, SMD												
3	Randomized trials	serious [x]	not serious	not serious	serious [b]	none	491	511		SMD 0.09 higher (0.03 lower to 0.22 higher)	⊕⊕◯◯ LOW	CRITICAL
Fruit Intake, at 3.5 to 3.7 years old, SMD												
2	Randomized trials	not serious	not serious	not serious	serious [b]	none	349	372		**SMD 0.17 higher (0.02 higher to 0.32 higher)**	⊕⊕⊕◯ MODERATE	CRITICAL
Fruit Intake, at 5 years old, SMD												
2	Randomized trials	not serious	not serious	not serious	serious [b]	none	353	370		SMD 0.05 higher (0.09 lower to 0.20 higher)	⊕⊕⊕◯ MODERATE	CRITICAL
Meat, poultry, fish intake, at 12 months old, g/day												
2	Randomized trials	not serious	not serious	not serious	serious [b]	none	422	435	-	SMD 0.00 higher (0.25 lower to 0.25 higher)	⊕⊕⊕◯ MODERATE	CRITICAL
Water Intake, at 9 to 12 months old, ml/day												
3	Randomized trial	not serious	not serious	not serious	serious [b]	none	268	274		MD 0.04 lower (23.3 lower to 0.15.36 higher)	⊕⊕⊕◯ MODERATE	CRITICAL
Water Intake, at 20 to 24 months old, ml/day												
3	Randomized trial	not serious	not serious	not serious	serious [b]	none	491	511		MD 24.2 higher (26.4 lower to 74.8 higher)	⊕⊕⊕◯ MODERATE	CRITICAL
Water Intake, at 3.6 years old, ml/day												

2	Randomized trial	not serious	not serious	not serious	serious [b]	none	349	372		**MD 111.3 higher (17.0 higher to 205.6 higher)**	⊕⊕⊕◯ MODERATE	CRITICAL
Water Intake, at 5 years old, ml/day												
2	Randomized trial	not serious	not serious	not serious	serious [b]	none	353	370		MD 52.6 higher (42.8 lower to 148.0 higher)	⊕⊕⊕◯ MODERATE	CRITICAL
UNHEALTHY FOOD CONSUMPTION ***Intervention B1*: Responsive feeding and development stimulation program, delivered by trained women/mothers of the village, including 7 components of RF (C1, C3, C4, C5, C7, C9, C10)** **Sweet snacks/sugar-dense foods between 20 and 23 months old, times/day**												
2	Randomized trials	serious [x]	serious [y]	not serious	serious [z]	none	165	149		MD 0.11 lower (0.50 lower to 0.28 higher)	⊕◯◯◯ VERY LOW	CRITICAL
***Intervention C1*: Interventions aimed to prevent obesity (e-health intervention between 6 and 12 months of age), including 5 components of RF (C1, C4, C6, C7, C8)Sweet and salty snacks + sweetened beverages, consumption at 12 months old, times/day**												
1	randomized trial	not serious	not serious	not serious	serious [aa]	none	269	264	-	MD 0.02 lower (0.06 lower to 0.02 higher)	⊕⊕⊕⊕ HIGH	CRITICAL
Sweet and salty snacks, consumption at 24 months old, >3.5 times/week vs less frequent												
1	randomized trial	not serious	not serious	not serious	serious [d]	none	101/165 (61.2%)	105/153 (68.6%)	RR 0.89 (0.76 to 1.05)	75 less per 1,000 (165 less to 34 more)	⊕⊕⊕◯ MODERATE	CRITICAL
Sweetened beverages, consumption at 24 months old, >2 times/week vs less frequent												
1	randomized trial	not serious	not serious	not serious	serious [d]	none	88/166 (53.0%)	70/155 (45.2%)	RR 1.17 (0.94 to 1.47)	77 more per 1,000 (27 less to 212 more)	⊕⊕⊕◯ MODERATE	CRITICAL
***Intervention C2*: Interventions aimed to prevent obesity delivered by health professionals (group or individual sessions during home visits or at health centers, specific brief advice included in well-child visits), including between 5 and 8 components of RF (C1, C4, C7 and others)** **Sugar-sweetened beverages (SSBs) intake, at 6 months, ml/day**												
1	Randomized trial	not serious	not serious	not serious	serious [bb]	none	55	42		MD 5.07 lower (10.53 lower to 0.39 higher)	⊕⊕⊕◯ MODERATE	CRITICAL
SSB intake at 9 months old, ml/day												
2	Randomized trials	not serious	serious [cc]	not serious	serious [dd]	none	219	206		**MD 7.45 lower (14.21 lower to 0.68 lower)**	⊕⊕◯◯ LOW	CRITICAL
SSB intake between 12 and 24 months old, SMD												
4	Randomized trials	serious [ee]	serious [ff]	not serious	serious [o]	none	509	529		SMD 0.34 lower (0.78 lower to 0.09 higher)	⊕⊕◯◯ LOW	CRITICAL
SSB intake between 3.6 to 3.7 years old, SMD												
2	Randomized trials	not serious	not serious	not serious	serious [o]	none	338	357		SMD 0 (0.15 lower to 0.15 higher)	⊕⊕⊕◯ MODERATE	CRITICAL
SSB intake at 5 years old, SMD												

2	Randomized trials	not serious	not serious	not serious	serious [o]	none	342	355		SMD 0.08 lower (0.26 lower to 0.11 higher)	⊕⊕⊕◯ MODERATE	CRITICAL
Sweet snacks/sugar-dense food intake between 9 and 16 month old, SMD												
2	Randomized trials	not serious	not serious	not serious	serious [o]	none	318	362		SMD 0.14 lower (0.29 lower to 0.01 higher)	⊕⊕⊕◯ MODERATE	CRITICAL
Sweet snacks/sugar-dense food between at 20 months old, SMD												
1	Randomized trials	not serious	not serious	not serious	serious [o]	none	139	139		**SMD 0.25 lower (0.48 lower to 0.01 lower)**	⊕⊕⊕◯ MODERATE	CRITICAL
Sweet snacks/sugar-dense food intake between 3.6 to 3.7 years old, SMD												
2	Randomized trials	not serious	serious [gg]	not serious	serious [o]	none	238	286		SMD 0.22 lower (0.45 lower to 0.01 higher)	⊕⊕◯◯ LOW	CRITICAL
Sweet snacks/sugar-dense food intake between 5 and 8 years old, SMD												
2	Randomized trials	not serious	not serious	not serious	serious [o]	none	223	263		**SMD 0.22 lower (0.40 lower to 0.04 lower)**	⊕⊕⊕◯ MODERATE	CRITICAL
ENERGY and NUTRIENT INTAKES **A. Interventions aimed to prevent under-nutrition, including two or more components of RF** ***Intervention B1*: Responsive feeding and development stimulation program, delivered by trained women/mothers of the village, including 7 components of RF (C1, C3, C4, C5, C7, C9, C10)** **Energy intake (kcal/day) at 9 months old**												
1	Randomized trial	serious	not serious	not serious	not serious	none	122	135		**MD 122 higher (76.7 higher to 167.3 higher)**	⊕⊕⊕◯ MODERATE	CRITICAL
Energy intake (kcal/day) at 15 months old												
1	Randomized trial	serious [b]	not serious	not serious	not serious	none	112	129		**MD 100 higher (86.7 higher to 113.3 higher)**	⊕⊕⊕◯ MODERATE	CRITICAL
Protein intake (g/day), at 9 months old												
1	Randomized trial	serious [b]	not serious	not serious	not serious	none	122	135		**MD 3.3 higher (2.19 higher to 4.41 higher)**	⊕⊕⊕◯ MODERATE	CRITICAL
Protein intake (g/day), at 15 months old												
1	Randomized trial	serious [b]	not serious	not serious	not serious	none	112	129		**MD 4.6 higher (4.12 higher to 5.08 higher)**	⊕⊕⊕◯ MODERATE	CRITICAL
Iron intake (mg/day), at 9 months old												
1	Randomized trial	serious [b]	not serious	not serious	serious [hh]	none	122	135		**MD 0.20 higher (0.04 higher to 0.36 higher)**	⊕⊕◯◯ LOW	CRITICAL
Iron intake (mg/day), at 15 months old												

1	Randomized trial	serious [b]	not serious	not serious	not serious	none	112	129		**MD 0.30 higher (0.25 higher to 0.35 higher)**	⊕⊕⊕◯ MODERATE	CRITICAL
Zinc intake (mg/day), at 9 months old												
1	Randomized trial	serious b	not serious	not serious	not serious	none	122	135		**MD 0.40 higher (0.29 higher to 0.51 higher)**	⊕⊕⊕◯ MODERATE	CRITICAL
Zinc intake (mg/day), at 15 months old												
1	Randomized trial	serious [b]	not serious	not serious	not serious	none	112	129		**MD 0.29 higher (0.27 higher to 0.31 higher)**	⊕⊕⊕◯ MODERATE	CRITICAL
Calcium intake (mg/day), at 9 months old												
1	Randomized trial	serious b	not serious	not serious	not serious	none	122	135		**MD 50 higher (17.7 higher to 82.3 higher)**	⊕⊕⊕◯ MODERATE	CRITICAL
Calcium intake (mg/day), at 15 months old												
1	Randomized trial	serious [b]	not serious	not serious	serious [h]	none	112	129		**MD 21 higher (1.20 higher to 40.8 higher)**	⊕⊕◯◯ LOW	CRITICAL
C. Interventions for obesity prevention, which includes two or more components of RF ***Intervention C2*: Interventions aimed to prevent obesity delivered by health professionals (group or individual sessions during home visits or at health centers, specific brief advice included in well-child visits), including between 6 and 7 components of RF (C1, C5, C7, C9 and others)** **Energy intake (kcal/day) at <12 months old**												
1	Randomized trial	not serious	not serious	not serious	serious [ii]	none	85	77		MD 28.7 higher (6.7 lower to 64.1 higher)	⊕⊕⊕◯ MODERATE	CRITICAL
Energy intake (kcal/day) at 12 months old												
1	Randomized trial	not serious	not serious	not serious	serious [p]	none	74	70		MD 26.5 higher (33.1 lower to 86.2 higher)	⊕⊕⊕◯ MODERATE	CRITICAL
Energy intake (SMD) at 24 months												
2	Randomized trials	serious [jj]	not serious	not serious	serious [kk]	none	77	76		SMD 0.13 lower (0.46 lower to 0.2 higher)	⊕⊕◯◯ LOW	CRITICAL
Protein intake (g/day) at <12 months												
1	Randomized trial	not serious	not serious	not serious	serious [p]	none	85	77		MD 1.4 higher (0.16 lower to 2.96 higher)	⊕⊕⊕◯ MODERATE	CRITICAL
Protein intake (g/day) at 12 months old												
1	Randomized trial	not serious	not serious	not serious	serious [p]	none	74	70	-	MD 0.9 higher (2.08 lower to 3.88 higher)	⊕⊕⊕◯ MODERATE	CRITICAL
Protein intake (g/day), at 24 months old												
1	Randomized trial	not serious	not serious	not serious	serious [p]	none	57	56	-	MD 0.3 higher (1.86 lower to 2.46 higher)	⊕⊕⊕◯ MODERATE	CRITICAL
Total fat intake (% energy) at <12 months old												
1	Randomized trial	not serious	not serious	not serious	serious [p]	none	85	77		MD 0.50 higher (1.16 lower to 2.16 higher)	⊕⊕⊕◯ MODERATE	CRITICAL

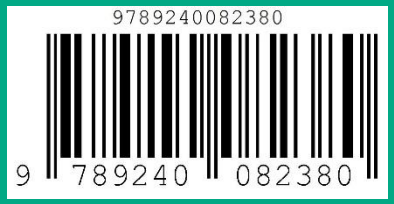
Total fat intake (% energy) at 12 months old												
1	Randomized trial	not serious	not serious	not serious	serious [p]	none	74	70		MD 0.20 lower (1.86 lower to 1.46 higher)	⊕⊕⊕◯ MODERATE	CRITICAL
Total fat intake (% energy) at 24 months												
2	Randomized trials	serious [u]	not serious	not serious	serious [v]	none	77	76		MD 0.21 lower (1.51 lower to 1.09 higher)	⊕⊕◯◯ LOW	CRITICAL
Iron intake (mg/day) at <12 months												
1	Randomized trial	not serious	not serious	not serious	serious [p]	none	85	77		MD 0.20 higher (0.87 lower to 1.27 higher)	⊕⊕⊕◯ MODERATE	CRITICAL
Iron intake (mg/day) at 12 months old												
1	Randomized trial	not serious	not serious	not serious	serious [p]	none	75	68	-	MD 0.00 higher (1.40 lower to 1.40 higher)	⊕⊕⊕◯ MODERATE	CRITICAL
Zinc intake (mg/day) at <12 months												
1	Randomized trial	not serious	not serious	not serious	serious [p]	none	85	77		MD 0.13 higher (0.07 lower to 0.33 higher)	⊕⊕⊕◯ MODERATE	CRITICAL
Zinc intake (mg/day) at 12 months old												
1	Randomized trial	not serious	not serious	not serious	serious [p]	none	75	68	-	MD 0.30 higher (0.10 lower to 0.70 higher)	⊕⊕⊕◯ MODERATE	CRITICAL
Calcium intake (mg/day) at <12 months												
1	Randomized trial	not serious	not serious	not serious	serious [p]	none	85	77		MD 19.0 higher (28.5 lower to 66.4 higher)	⊕⊕⊕◯ MODERATE	CRITICAL
Calcium intake (mg/day) at 12 months old												
1	Randomized trial	not serious	not serious	not serious	serious [p]	none	75	68	-	MD 6 higher (70.5 lower to 82.5 higher)	⊕⊕⊕◯ MODERATE	CRITICAL
Calcium intake (mg/day), at 24 months old												
1	Randomized trial	not serious	not serious	not serious	serious [p]	none	57	56	-	MD 9.00 lower (91.4 lower to 73.4 higher)	⊕⊕⊕◯ MODERATE	CRITICAL

CI: confidence interval; **MD:** mean difference; **RR:** risk ratio; **SMD:** standardised mean difference. **COMPONENTS:** C1. Recognition of hunger and satiety; C2. Infant readiness for introduction of complementary foods, taking into account the child's developmental readiness; C3. Texture/ consistency responsive to child developmental needs; C4. Not pressuring child to eat; praising, encourage self-feeding; C5. Flavor preferences and repeated exposure to certain foods. It includes interventions promoting caregivers to offer a diverse diet with repeated exposure to healthy foods/beverages, and to avoid offering ultra-processed foods and sugar-sweetened beverages; C6. Role modelling of healthy eating; C7. Pleasant and stimulating family eating environment; C8. Appropriate soothing (caregivers do not use food to calm child when s/he is not hungry), sleeping, and play routines (caregivers establish well-structured daily routines for sleeping and playing as part of a responsive parenting/responsive feeding intervention); C9. Positive caregiver verbalization during feeding; C10 Feeding during and after illness

Explanations

a. Risk of bias: Downgraded one level due to some concerns due to possible bias in randomization process, deviations from intended interventions and in the selection of the reported results.
b. The confidence interval is precise but the study does not meet the optimal information size criteria. Downgraded one level.
c. Risk of bias: Fildes: high risk of bias due to possible bias in the randomization process. Hetherington 2015: some concerns due to possible bias in the randomization process, deviations from intended interventions and in the selection of the reported results. Downgraded one level.
d. Risk of bias: Downgraded two levels because of high risk of bias in randomization process in a single study.
e. Inconsistency: Downgraded one level due to high heterogeneity between sub-studies (I2= 61%).
f. Imprecision: Downgraded one level due to wide confidence intervals compatible both with benefit or no effect.
g. Risk of bias: Downgraded one level because of some concerns due to possible bias in randomization process.
h. The study does not meet the optimal information size criteria. Downgraded one level.

i. Inconsistency: Downgraded one level due to moderate heterogeneity (I2=40%).
j. Risk of bias: Daniels 2012: Some concerns due to missing outcome data. Savage 2016: Some concerns due to possible bias in randomization process. Downgraded one level
k. Although it is a single study, we decided not to downgrade its certainty of evidence level because it is of low risk of bias and its sample size exceeded the optimal information size (n=68 each group).
l. Risk of bias: Some concerns due to missing outcome data. Single study. Downgraded one level.
m. Imprecision: Downgraded one level due to wide confidence intervals compatible both with benefit or no effect.
n. Risk of bias: Some concerns due to missing outcome data. Single study. Downgraded one level.
o. Inconsistency: Downgraded one level due to heterogeneity (I2 = 42%). One trial (Fangupo 2016) presented no effect (0.00, 95% CI -0.31 to 0.31), but mean score in intervention and control groups was 4.8 from a maximum of 5. The other study (Daniels 2012) showed an increase in the perception of fruit liking (0.24, 95%CI 0.06 to 0.42), both expressed as SMD. Heterogeneity may be partially explained by heterogeneity in follow-up time.
p. Imprecision: downgraded one level due to wide confidence intervals compatible both with benefit or no.
q. Risk of Bias: Some concerns due to missing outcome data. It is only study, Downgraded one level.
r. Risk of Bias: Some concerns due to possible deviations from intended interventions in both trials. Downgraded one level.
s. Inconsistency: Downgraded one level due to heterogeneity (I 2=91%). Both trials differ slightly with regard to the population. While population in both trials were classified as poor and very poor, family assets and mothers' education in Aboud 2009 were lower than in Aboud 2008. Also, consumption of vegetables during the previous day was more frequent both in control and intervention groups in Aboud 2009 than in Aboud 2008 trial.
t. Inconsistency: Both trials differ slightly with regard to the population. While population in both trials were classified as poor and very poor, family assets and mothers' education in Aboud 2009 were lower than in Aboud 2008. Heterogeneity was moderate (I 2=43%), but we decided not to downgrade because both trials reported estimations in the same direction.
u. Risk of Bias: Downgraded one level due to RoB-2 overall judgement was classified as some concerns or high risk in some of the studies. Risk of bias: French 2012: High risk due to possible bias in randomization process, deviations from intended interventions and missing outcome data. Savage 2016: Some concerns due to possible bias in randomization process. Campbell 2013: Low risk of bias.
v. Risk of Bias: Downgraded one level due to RoB-2 overall judgement was classified as some concerns in some of the studies. Risk of bias: Black 2021: Some concerns due to possible bias in randomization process and deviations from intended interventions. Daniels 2012: Some concerns due to missing outcome data. Campbell 2013: Low risk of bias.
w. Inconsistency: Downgraded one level due to heterogeneity (I^2 = 54%).
x. Risk of Bias: Some concerns due to possible deviations from intended interventions in both trials. Downgraded one level.
y. Inconsistency: Downgraded one level due to heterogeneity (I^2=60%). Both trials differ slightly with regard to the population. While population in both trials were classified as poor and very poor, family assets and mothers' education in Aboud 2009 were lower than in Aboud 2008.
z. Imprecision: downgraded one level due to wide confidence intervals compatible both with benefit or no.
aa. Although it is a single study, we decided not to downgrade its certainty of evidence level because it is of low risk of bias and its sample size exceeded the optimal information size.
bb. Imprecision: Downgraded one level due to wide confidence intervals compatible both with benefit or no. Single study
cc. Inconsistency: Downgraded one level due to heterogeneity. I^2=61%. Heterogeneity may be partially explained by population heterogeneity (see subgroup analysis HIC vs HICs-low resource settings). Estimations from both trials are in the same direction.
dd. Imprecision: Downgraded one level due to wide confidence intervals compatible both with benefit or no.
ee. Risk of Bias: Downgraded one level because of high risk of bias due to possible bias in randomization process, deviations from intended interventions and missing outcome data in one of the trials (French et al).
ff. Inconsistency: Downgraded one level due to substantial heterogeneity (I^2 = 91%) which could be partially explained by population heterogeneity (see subgroup analysis HIC vs HICs-low resource settings).
gg. Inconsistency: Downgraded one level due to moderate heterogeneity (I^2 = 44%).
hh. Imprecision: downgraded one level due to wide confidence intervals compatible both with benefit or no.
ii. Imprecision: Downgraded one level due to wide confidence intervals compatible both with benefit or no and single study.
jj. Risk of bias: Fangupo 2016: Low risk of bias. Harvey-Berino 2003: Some concerns due to possible bias in randomization process and in the selection of the reported results.
kk. Imprecision: Downgraded one level due to wide confidence intervals compatible both with benefit or no.


9789240082380
9 789240 082380